Proceedings

2012 IEEE Control and System Graduate Research Colloquium (ICSGRC 2012)

16 – 17 July 2012
Faculty of Electrical Engineering
UNIVERSITI TEKNOLOGI MARA
Shah Alam, Malaysia

ORGANIZING COMMITTEE

Patron
Dean, Faculty of EE, UiTM

General Chair
Prof. Dr. Mohd Nasir Taib

Secretary
Dr. Rozita Jailani

Finance / Sponsorship Chair
Dr. Mohd Hezri Fazalul Rahiman

Technical Program Chair
Assoc. Prof. Dr. Nooritawati Md. Tahir

Technical Program Co-Chair
Assoc. Prof. Sr Dr. Abd Manan Samad

Local Arrangement Chair
Megat Syahirul Amin Megat Ali

Publication Chair
Assoc. Prof. Dr. Ramli Adnan

Publicity Chair
Ahmad Ihsan Mohd Yassin

Admin Representive
Hamidah Mohd Nasir

Reviewers

The Organizing and Technical Committees of 2012 IEEE Control and System Graduate Research Colloquium (ICSGRC 2012) would like to express gratitude to all reviewers for the volunteering support and contribution in the reviewing process.

Abd Manan Samad
Ali Akbar Shaik
Alishir Moradikordalivand
Faisal Rahiman Pazheri
Fazlina Ahmat Ruslan
Gursharanjeet Singh Kalra
Hirosato Seki
Ihsan Yassin
Megat Syahirul Amin Megat Ali
Michael Jackson Patrick
Mirza Mansoor Baig
Mohammad Farid Saaid
Mohd Azman Abdullah
Mohd Helmy Abd Wahab
Mohd Hezri Fazalul Rahiman
Mohd Nasir Taib
Murtaza Rizvi

Nitish Gupta
Nooritawati Md Tahir
Nordiana Mukahar
Nurlaila Ismail
Prabhjot Kaur
Rajeeb Dey
Ramli Adnan
Rozita Jailani
Sabiha Hanim Saleh
Saiful Aman Hj Sulaiman
Shahril Irwan Sulaiman
Siti Salmah Yasiran
Subhash Chander Dubey
Usman Ahmad
Zairi Ismael Rizman

Table of Contents

Research for Security Strategy of Cloud Service Based On System Survivability

Huanan Liu

Information Engineering Institute
Zhengzhou University
Zhengzhou China
cindyliu118@126.com

Shiqing Wang

Information Engineering Institute
Zhengzhou University
Zhengzhou China
iesqwang@zzu.edu.cn

Abstract—**Along with the ragingly increasing concern about cloud computing due to its appalling potential merits, not only do we envision its blossom prospect, but we also realize that the importance and the major traits of system survivability in cloud computing. Therefore, this paper firstly discusses the basic concept and main characteristics of cloud computing. And in sequence it summarizes the definition and strategies of system survivability in traditional secure engineering. What is more important is that, basing on the analysis system survivability in such nascent progressing trend, this paper proposes certain security strategies and model of cloud service in order to maintain and propel system survivability in cloud computing. (Abstract)**

Keywords-component; cloud computing; system survivability

I. INTRODUCTION

Cloud computing, with cloud service as its core, is no doubt the inevitable trend of the development of computer technology and environment. And as expected, it is soaringly developing.

No wonder more and more concerns gains the cloud computing, the fact is that, it owns huge potentials for application. It obtains tremendous amount of advantages: high speed, elastic, convenience, low power-consuming, etc. However, each coin has two sides, the vulnerabilities of security, privacy and confidentiality of cloud computing also should be taken an eye on. Naturally, the secure cloud service becomes popular.

With the thriving of economy and society, government and business all enforce their requirements about security. They do not just require the ability of resisting attacks. But system survivability should also be strengthened, which means when some system is compromised, it can still supply critical services and make immediate recovery.

Under such background, this paper proposes new security strategies and model for cloud service, basing on maintaining and improving system survivability.

II. SECURITY ISSUE IN CLOUD COMPUTING

A. Cloud Computing and Cloud Service

In today's fast-growing field of computing and information, the concept of service [1] plays a vital role. Service covers a lot of content: kinds of basic data, such as picture, video, graphic; hardware abilities, such as CPU (computing power), hard disk (storage capacity); software; platform, on which the users can be freely compile or run their own applications; and infrastructure, which will provide the users servers and databases. Overall, everything can be service that may be shared, passed, sold, or purchased.

Cloud computing means systems (or say cloud points) may provide services, as well as purchase services through networks.

In cloud, no matter when, no matter what type of service, as long as the user request, these services will be easily combined to complete the work. When in need, these services are grouped together, like water vapor floating together, gathering into clouds; when work is done, each service leaves, the same as the cloud dissipates. For Each point can both provide service and purchase service, "role" in such circumstances is no longer important.

B. New Traits of Cloud Computing:

- *Flexibility:* Flexibility [2] is the cloud can be quickly gathered into the desired size and shape.

- *Self-supply capability:* Users can easily obtain the required combination of services. When the work is finished, the service can be automatically discharged or released.

- *Standard interface:* The standard application programming interface can make free interactions between applications. And naturally, the user is easy to gather and use cloud services.

- *Service's measure and charge:* Measure and charge the service by use, i.e. pay by use.

- *Virtualization:* All the content such as a variety of infrastructures, platforms, applications providing by points that comprising the cloud, can be virtualized as

local providing; i.e. it is comfortable and convenient for system to use colorful services.

C. Security Issue in New Trends:

Under such new environment with new technology, new features of system security and integrity occur.

Firstly, in Cloud Computing, the point can act both as consumer and vendor; in addition, each kind of function of one point depends on various services provided by other points. That means along with higher dependence among points, their security is more strongly constrained by each other. If one gets compromised, all the relative ones may get compromised, too. (see Figure 1)
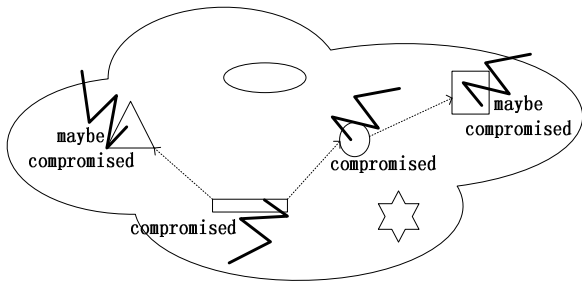
Figure 1. one compromised point may infect others

Secondly, due to cloud traits, each user can be a part of the cloud, and then there will be a raising risk that more points are likely to be attacked. System's security has to be taken full consideration and analysis.

Again and doubtless, under the driven by new demands and nascent traits, system survivability [3] in cloud computing must be given much attention and devoted research.

III. SYSTEM SURVIVABILITY

A. Definition of System Survivability

System survivability is the ability of the system to continue to work and provide critical services when it is in or has been affected by natural or man-made disruption. The "system" here includes not only the commonly understood concept of system, subsystem, it also means equipment, process, procedure, and so on. To ensure the continued provision of key services is the core objective of survivability.

B. Analysis Strategies for System Survivability

Ellison [4] and his colleagues designed a method for analyzing system survivability, known as: Survivable System Analysis. This method can assess system vulnerabilities and design system architecture and features in order to improve system survivability. Their proposed three complementary strategies for system survivability:

- *Attack resistance:* Enhanced relative abilities of the system to avoid attacks. For example, the system can authenticate users by digital certificates to prevent unauthorized users from accessing.

- *Attack identification:* Create for system the ability to detect attacks, system errors or damages. For example,

verify the checksum to determine whether an error occurs and even to locate and correct the mistakes.

- *System recovery:* When the system has been compromised, it is necessary to ensure that the system can continue to provide critical services, and in addition it should recover all the system functions. For example, error-patience mechanism, i.e. the same kinds of services with different implementations, so that when a service has been attacked, others with different implementations can immediately be used to achieve a "substitute "work.

C. Steps for Analyzing System Survivability

According to such analysis, they also propose four steps:

- *Understand the system:* analyze existing or looking-forward-to-building systems; understand the system goals (or known as the mission objectives), system requirements and system architecture.

- *Identify the key services:* basing on the understanding of the system target, identify the specific critical services.

- *Simulate attack:* Basing on a comprehensive understanding of the system goals, needs and architecture, imagine the possible damages and attack-liable services.

- *Analyze survivability:* since having found the critical and compromising services, then analyze and design targeted prevention strategies.

IV. SECURITY STRATEGIES FOR CLOUD SERVICE

A. Definition of System Survivability in Cloud Computing

This paper defines system survivability in cloud computing as: when the system is or has been attacked, firstly, it can continue to supply vital services; secondly, cut off any connection with such compromised region, notifying the relative providers and consumers; thirdly, redirect to new request for substitute services. All these steps pursue the ability that recovering the whole system [5] as quickly as possible.

B. Analysis for Cloud Security

Nascent environment, nascent considerations. We may address the emergent problems by taking the advantages of new features of the fresh environment. In other words, we could use the appropriate countermeasures [6] for different problems.

First, owing to the massive and complex points, vast malicious attackers [7] may damage the system and data by pretending to provide services. The available strategy is: since in the cloud, the notation of role is vague, i.e. one point is either a consumer or a vendor; then we can establish "credit rating" tag for each point (see Figure 2). All clients can evaluate the service providers. Once a certain point fail to supply quality services, even provide malicious services, the consumer can pull the provider into the blacklist, and rating the

provider for its credit, which means the provider is "tainted". Such "tainted" points will be rejected in the future interactions by other points. Thus, in order to survive in the cloud, each point should review and care about their services and behaviors. According to that, there can automatically reduce and eliminate malicious attacks to some extent.
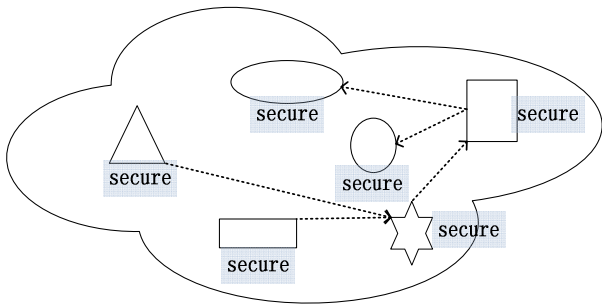
Figure 2. credit rating tag

Secondly, although points are more prone to be compromised due to close relations that one compromised point may quickly contaminate others, such close relations not only can help detect and locate tainted dots, but it can also establish notify mechanism, resulting in quick broadcasting of "bad news" (see Figure 3). When other points get the bad news, they will address certain measures against the assault and damage. In a word, it can fully take advantage of such close relations among points.
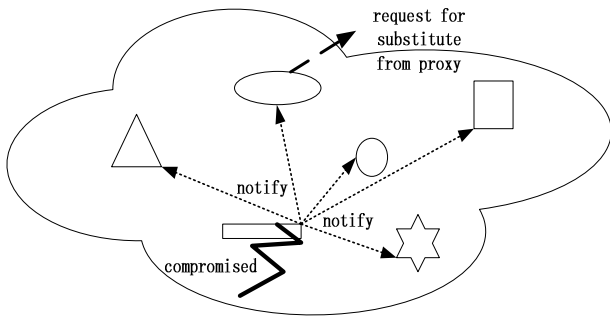
Figure 3. notify the bad news

Finally, agents or proxy can be set up for all the points. Then the scenarios will look like: when one point request for certain service, it can turn to the agents for purchasing such service. There will also be an index list about all the available agents. For each agent, it may present the index of all the vendors corresponding to thus service. Following that, once one vendor fails, the agent may redirect to new vendor point for substitution. Similarly, once one agent fails, the consumer could choose another proxy (see Figure 4). As illustrated, the system guarantees its survivability.
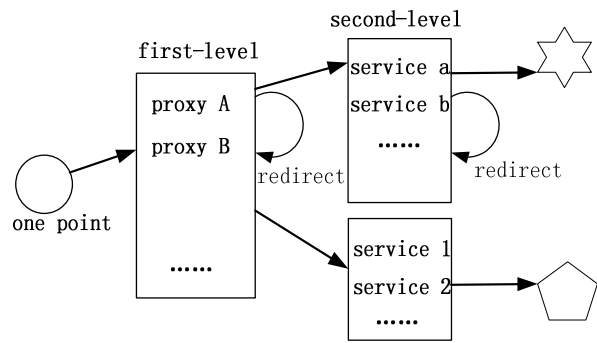
Figure 4. the proxy

C. Strategies for System Survivability in Cloud Computing

To sum up, there are some strategies for maintaining and improving system survivability:

- *Resist attack:* Avoid attacks by enhancing relative abilities of system, such as authenticating legal users by digital certificates and also preventing unauthorized users [8] from accessing.

- *Identify attack:* Establish attack-detect [9] mechanism, detecting assaults, system errors, damages. For example, verifying the checksum can determine whether and where error exists, and maybe even rectify the error.

- *Notify attack:* Establish notify mechanism, when some service in some point has been compromised, such bad news will immediately notify relative points.

- *Evaluate service:* Each point, strictly speaking, each service consumer has the privilege to evaluate the services it consumes. Such point grades the "credit rating" for services. If the service is malicious, the point can pull the service into its blacklist, and starts the notify mechanism. Of course, at the very beginning, the consumer checks the "credit rating" of the service. Low grade means bad quality, which warns the consumer to request for substitute service.

- *Recover system:* Once bad news expands, on one hand, the compromised point should recover [10] itself; on the other hand, the point who was supported by such tainted point must also recover its own.

D. System Work Model in Cloud Computing

- *Self-monitoring:* Each point shoulders the responsibility to monitor and test the living conditions and security threats of the services it provides.

- *Service judge and evaluation:* When beginning to consume, consumers judge the services by their "credit rating". If there are the low grades, redirect for new providers. After using, the consumer needs to evaluate services and grade them in their "credit rating" respectively.

- *Notify attack:* Whilst testing compromised services, firstly insulate such services, and secondly notify other points on time.

- *Recovery:* Each point use first-level proxy directory for selecting appropriate proxy or later redirecting for substitute proxy. And it uses second-level proxy directory for selecting services or later redirecting for substitute services.(see Figure 5)
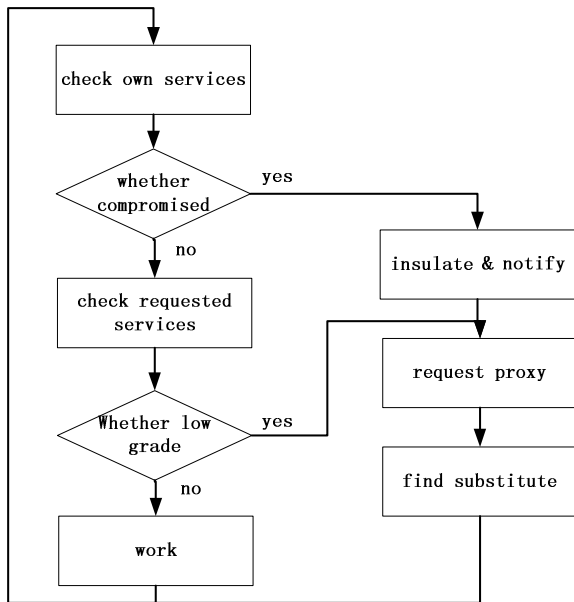
Figure 5. system work model

V. CONCLUSION

With cloud computing gathering increasingly more attention and owing to its own excellent prospects in development, this paper firstly analyze the nascent features of cloud computing by illustrating its basic concepts, traits and categories. And in sequent, the paper discusses the importance and new definition of system survivability through concerning the traditional security engineering. Finally, it proposes the strategies and model for maintaining and improving system survivability in cloud computing.

REFERENCES

[1] www.cloudcomputing.sys-con.com/node/1087426

[2] www.dummies.com/how-to/content/cloud-computing-characteristic.html

[3] www.wikipedia.org/wiki/survivability

[4] chapter 30.4, 'system survivability', <Software Engineering> 8th version, Ian Sommerville ISBN: 7-111-19770-4

[5] Irving Vitra Paputungan, Islamic University of Indonesia ;Azween Abdullah, Universiti Teknologi Petronas. Modelling a Survivable System through Critical Service Recovery Process. Second UKSIM European Symposium on Computer Modeling and Simulation.2008 IEEE.

[6] Partha Pal, Michael Atighetchi, Joseph Loyall,Andrew GronoskyInformation and Knowledge Technologies Unit Raytheon BBN Technologies Cambridge, USA. Advanced Protected Services.2010 13th IEEE International Symposium on Object/ Component/ Service-Oriented Real-Time Distributed Computing Workshops

[7] Shuai Zhang, Shufen ZhangHebei, Xuebin Chen, Xiuzhen Huo Polytechnic University Cloud Computing Research and Development Trend 2010 Second International Conference on Future Networks

[8] Matthew G. Richards Adam M. Ross,Daniel E. Hastings Donna H. Rhodes,Massachusetts Institute of Technology,Cambridge, MA. DESIGN PRINCIPLES FOR SURVIVABLE SYSTEM ARCHITECTURE. 2007 1st Annual IEEE Systems Conference,Waikiki Beach, Honolulu, Hawaii, USA April 9-12, 2007

[9] Huiqiang Wang, Guosheng Zhao, and Jian Wang.School of Computer Science and Technology, Harbin Engineering University, Harbin, China. Survivable Network System: An Immune Approach. 2008 International Conference on Internet Computing in Science and Engineering

[10] Irving Vitra Paputungan, and Azween Abdullah. Specifying a Recovery Model for Survivable System. The 5th Student Conference on Research and Development –SCOReD 2007,11-12 December 2007, Malaysia

The Analysis and Design of Trusted Computing Applied Into Cloud

Huanan Liu

Information Engineering Department of Zhengzhou University

Zhengzhou City,Henan Province,China

cindyliu118@126.com

Shiqing Wang

Information Engineering Department of Zhengzhou University

Zhengzhou City,Henan Province,China

iesqwang@zzu.edu.cn

Abstract—**Although making changes in hardware is not the only means to enhance security, it still holds the undefeatable advantages for security comparing with software. Trusted computing is such a kind of security technology that relies on the basis of making changes in hardware. Moreover, thus technology is already developed relevant standard specifications and it owns advantages of stronger security and more efficiency in information communications. In the cloud computing era, whose environment requires high security for information and services, trusted computing is proposed by this paper with objective analysis, original improvement and novel evolvement by adding two new modules: Service Authentication List Management (SAL), and Configuration Dynamic Update Module, (CUM). We also present and illustrate an abstract model for Trusted Computing Applied into the Cloud.**

Keywords- trusted computing; cloud computing; system survivability; Service Authentication List Management (SAL), and Configuration Dynamic Update Module, (CUM)

I. INTRODUCTION

More than a decade, although the concept of trusted computing just looms, thus technology is always getting progresses. It is such a technology, by adding specific computer hardware and software, to ensure the authenticity of the user, thus improving the safety of the interaction. Trusted Computing is suffering from doubt and criticism. These questions focus on three main areas: user's autonomy being limited, user's privacy being no longer protected and the hardware's limitations on its own. However, "whatever is, is right". Trusted computing, in interaction security, has many advantages: authentication is more secure, more efficient, and it has established standards which have laid a solid foundation for the future evolution of any development. More importantly, in the era of cloud, the security requirements of cloud services, make it useful. This paper will analyze the feasibility of "trusted computing applied into cloud" design the model for it. Section II describes the concept, advantages, and disadvantages of trusted computing; Section III focuses on the characteristics of the era of cloud computing and cloud security; Section IV illustrates how to avoid weaknesses, how to highlight strengths and proposes some appropriate expansions and evolutions of trusted computing for a particular environment, thus making it as an important technical means in the protection of cloud

security; section V designs the model for the concept of trusted computing applied into cloud; section VI concludes the paper.

II. CONTENT RELATED TO TRUSTED COMPUTING

A. Concept and Features of Trusted Computing

In such an intensive information interaction era, one of the biggest problems confronted by computer technology is data security. And thus problem is getting increasingly serious. Many researchers believe that the development of trusted computing, the technology basing on hardware, becomes a necessity, for no need to add software layers to mitigate problems over and over again. In 1999, Hewlett-Packard, IBM, Compaq, Intel and Microsoft set up the Trusted Computing Platform Alliance, short as TCPA [1], focuses on the credibility of computing platform during e-commerce transactions. In 2003, the Trusted Computing Group, TCG was established and it adopted the Trusted Computing Specifications proposed by TCPA. The most prominent trait of TCG is the Trusted Base embedded into the computing platform.

Trusted computing aims at providing customers with a trusted domain, which is guaranteed by special hardware and operating systems provided by various providers of software and hardware, basing on the trusted platform technology. Trusted platform ensures a range of security-related properties [2] by transmitting trusted information chain. Information transmission is a process, viewing a certain system module as a trusted base, used to prove other modules are trusted, thus and thus establishing the trusted chain. In other words, it starts from a trusted root, and gradually expands the trusted border, and ultimately ensures the integrity of the entire computing platform.

Trusted Computing Platform, TCP [3], works through a set of hardware and software. TCP provides two basic services: authentic start-up and encryption which closely cooperate with each other. Authentic start-up means monitoring the process and making auditing records after the machine started. All these jobs done by the special hardware adding into the machine, can be called TPM, Trusted Platform Module, which computes values of components and then stores them into platform configuration register, PCR, through a credible channel. In TCP, TPM is used to report the configuration

parameters securely and trusted platform software stack, TSS, provides interfaces between TPM and other system modules [4].

In trusted computing, the machine seals data and applications with its own configurations. Hence, only the authentic subject could decrypt and acquire the corresponding information. Thus significant characteristic is viewed as a vital basis for developing technology specifications by TCG. "Trust" [5] refers to that, components, operations, or processes, whatever the environment is, maybe resisting any attacks, can perform as expected.

B.   Advantages and Application of Trusted Computing

The most prominent feature [6] of trusted computing——"hardwareness", creating irreplaceable advantages of trusted computing.

- More efficient

Since the instructions are fixed into hardware, without any compiling or interpretation as done by other software applications, measuring and encrypting kinds of information not only faster, but it also ease the pressure of CPU.

- More Secure

After the configuration of some machine has been measured and stored securely, its identity has been strongly formed and confirmed. That is to say, the machine will fail to deceit with such a unique identity, like fingerprints or DNA of human.

- TCG has issued specifications for trusted computing

The deliberate specifications contribute to be as a solid foundation for development, expansion and evolution of trusted computing in future.

According to the illustration of system architecture of trusted computing from Microsoft, the expected hardware-changes, also may be called basic functions of trusted computing, are classified as four facets:

- Memory Curtaining

There is an isolated storage; even the OS has no access to such a domain. Accordingly, malicious attackers even mal-control the OS and even easily get unauthorized access to the memory, still cannot infect thus storage. Therefore, memory curtaining effectively protect sensitive and vital information from leaking. Of course, such hard-change can be simulated by software, but that stands for considerable coding work, changing from OS to various device-drivers and applications.

- Secure I/O

Secure input and output supplies safe channels for data, avoiding keyboard-tapping or screen-theft.

- Sealed Storage

It refers to seal data and applications with machine's configuration, aiming at raising the security and prevent data and applications from leakage. For example, at first, make add-operation between sensitive information and configuration;

then, sign and encrypt the add-result. So that, in order to gain the real content, the recipient (no matter normal or compromised and no matter authentic or counterfeit) not only needs the private key to decrypt the information, but also has to match the configuration. As we know before, the configuration parameters fail to deceive, then, even if interaction information were intercepted, the vital and correct content still cannot be leaked.

- Remote Attestation

The TCP is capable to report errors, has occurred in its machine, to TCG or other related subjects. It owns popular application, such as digital rights management.

There are four major areas for the application of trusted computing [7]:

- Risk Management

Risk management mainly consists of intrusion detection and addresses. One of the work is vulnerability assessment, assess whether the current information assets have been destructed. In order to minimize risk, information manager naturally tries to protect the assets through kinds of methods described as the following: first, check the integrity by computing the encrypted hash value; second, prevent unauthorized party from acquiring information by encrypting with public and private keys; third, authenticate the transmission by digital signature. As presented before, memory curtaining is an isolated subsystem based on hardware. It can scrupulously take responsibility to various sorts of keys and hash values of platform configuration, which are still named as "meta information". Through such a firm wall, information assets are protected fabulously.

- Asset Management

Asset manager intends to prevent the theft and unauthorized using of computing resources. Asset tracking can achieve this goal. Under the control of the owner, TPM is used to create and protect system identity, avoiding being physically unloaded or replaced. Asset database may obtain more information assets by its identity. For such a unique and non-deceivable identity, even if the assets had been stolen, the thief still fails to access it.

- E-commerce

TCG is good at providing e-commerce environment, in which buyers and sellers could build relationships and exchange information. Still counting on the authentic report from TPM, not only the transactions are guaranteed proceeding securely, but also merchants will recognize old customers and recall their preferences.

- Security Monitoring and Emergency Response

IT managers have spent a lot of effort to exploit a variety of methods in response to attacks and threats. Security monitoring and emergency response are the main means and goals. Emergency response requires quick isolating the compromised system. Through trusted computing, TPM storing the non-deceivable configuration and making its report, then it determine whether the system is infected or the system

confront the normal upgrade by scanning the relevant machine's configuration and settings.

C. Shortcomings and Doubts of Trusted Computing

Despite amounts of advantages and prospects of trusted computing, however, in recent years, it suffers doubts and criticisms, mainly due to three areas [8]:

- Restricted User Autonomy

Mainly due to the abuse of remote attest, it would limit the autonomy of users to select their preferred software or to interact with other systems. For example, when a user visit a web site, which requires users to show some software proof, if the user refuses to hand out the proof, the site will deny its requests; if the user presents the proof of other software non-preferred by the site, the site still ignores the requests. Only is showed the proof satisfying the site's taste, user can gain the access. Clearly, to certain extent, users lose its autonomy.

The refusal to non-favor software is irrelevant with any security concern. In fact, the reason may be completely out of "dictatorship", or may be used for advertising, or may be to lessen competition.

- Non-Protected Privacy

Each coin has two sides. Since TCP can succeed in prevent identity from counterfeiting, resulting in the identity get equivalent to fingerprint or DNA, the one with trusted computing technology will be no longer as private as before. For example, the trace of the one who owns TCP cannot make any disguise.

- Hardware Limitation

First of all, there is inevitable drawback of hardware, comparing with software – neither feasible nor dynamic. For example, the normal upgrade is easily likely to be judged as "being compromised"; or, if user has changed its platform temporarily or permanently, its identity fails to be recognized. Secondly, hardware closely relies on its producers, then how to guarantee the merchants produce as strictly as the regulations ask becomes a worry. For example, once the merchants carved a "backdoor" [9] in the chip or add some unknown properties, the risks of insecurity or even "being blackmailed" will definitely raise.

III. RELATED WITH CLOUD

Cloud computing is a kind of model [10] for computing. It allows users access and obtain services non-locally through internet. The computing tasks required by users are deployed dynamically by the cloud into resource pool consists of considerable computers, which enables users gain computing ability, storage space and information at an easy way. For instance, one user may assign its intensive computing tasks to the cloud, for on one hand, it has no need to concern about some bottleneck of computing ability by its own, on the other hand, the user could put much more energy and wisdom into pursuing procedures for innovation and expansion. Another instance is that, users are becoming lighter for storing their data and applications into the cloud with the result of sparing

numbers of space in local databases. Moreover they can access their remote-stored data whenever they wanted. In addition, users are spared to develop or establish their own applications from the zero ground; instead, they absolutely can "purchase" any component whatever they need, in order to mitigating the difficult of developing. All in all, cloud will bring infinite convenience.

However, before people publish their information and services into the cloud, or acquire data and applications from the cloud, the security of the cloud is an inevitable concern and worry. There are many critical technology researches about the security of cloud computing, corresponding to every facet of worries, such as identity authentication, access control, cipher-text retrieval and process, proof of data possession, virtualization security, etc. This paper will mainly discuss about the use of trusted computing against identity authentication and access control.

Identity authentication and access control are tied tightly. Identity authentication aims at recognize the legal identity of any participant. For users, the service provider needs to provide its qualified identity. And for service providers, due to prevent malicious users from attacking, users have to prove their authentic identities.

Access control refers to that, only do the authorized and specific rights-gained users operations on certain objects at some scopes, completely eradicating unauthorized users' illegal accessing and obtaining.

For addressing such two subjects mentioned above, nowadays, the most concerned method is based on encryption. However, pure encryption has flaws: on one hand, encryption costs a lot; on the other hand, the essence of encryption still contains the old problem: store the cipher-text and key together. Under that situation, once the system has been broken into, the key being stolen, then the cipher-text naturally was decrypted and leaked. It likes that a safe and the key for it are stored in one room, once the room being burst into, safe will be no longer safe (see figure 1).

The era of cloud is coming at full speed. While the convenience brought by the cloud rises fabulous eagerness of people, the security of cloud surely acts as the bottleneck hindering its development.

IV. ANALYSIS OF FEASIBILITY OF TRUSTED COMPUTING APPLIED INTO CLOUD

A. Take Full Advantages of Trusted Computing

In the process of identity authentication and access control, the method of pure encryption may, at a high possibility, leak important information by putting cipher-text and key into the same room, as explained before.

Such a defect can be made up with by trusted computing. Because, first, TCP will calculate configuration at each unique instant, (see figure 2); second, the configuration is stored at the isolated domain, it cannot cheat or be stolen. Therefore, by trusted computing, the identity is guaranteed authentic. Then, the following operations will be under control. For example, client A who is authorized by only reading resource 1, when it

is recognized, cannot execute write-operation on resource 1, nor accessing resource 2.

What is more important is that, even if vicious one acquires the cipher-text, due to lacking matches with the origin platform, it still fails to access the vital information.
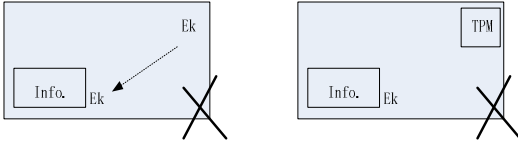
Figure 1.   Cyphertext and key share the same room

B.   Solutions By Turning Around Disadvantages Into Advantages

- Attestation Plus Owner Override

As depicted in section II, for resolving the problem, restricted user's autonomy, we can add owner override to the platform. That means, the computer would let remote parties know if the software on its own has been changed without its knowledge. In addition, owners even have the autonomy to decide whether to unload the TPM chip and to decide whether to proceed trusted authentication under specific environment.

For example, the regular update operation is known by owners, so it will not be misjudged as illegal changes. And remote parties will also update its configuration information for smooth interaction in future.

- Add Software-Token

Facing the users' shouting to protect their privacy, trusted computing can also make changes to meet that demand: to design a trusted software token, which is bound with a module for verifying security. Its goal is to execute various operations upon outsourced sensitive information without any leakage.

- Reinforce Supervision

Establish appropriate laws and regulations, standardize supervision approach, and appoint an independent trusted third party, to reinforce the supervision of trust hardware manufacturers and also to prevent software dictatorship or anti-competitive restrictions. We need to know that some software providers simply by software-approach can arouse incompatibility in the owner's machine. Therefore, for addressing such troublesome problems, both can be caused by hardware and software, the appropriate laws and regulations are emergently required.

C.   Expansion and Evolution

In addition to make changes for facing challenges suffered by trusted computing, it needs to continue evolving under the cloud environment.

As we know, virtualization is one of important features of cloud computing. Several virtual machines are running on each physical machine among cluster servers. Each cloud service may be accomplished by a combination of virtual machines

from various physical machines. Then the service needs not only one authentic platform report.

Another obvious feature of cloud computing is dynamic scalability: for the server, due to hardware or other reasons, a large number of machines, both virtual and physical, may be replaced, re-combined, or revoked. At one time, if machine A, serves service 1, has to be replaced by machine B, machine B's configuration should be recognized and authenticated; for the large client, maybe lots of machines have to be identified; for the small client, due to movement or hardware problems, there is still recognition issue for new machines. (see figure 2) Accordingly, trusted platform replacement technology is needed, for dynamically compute and authenticate platforms.
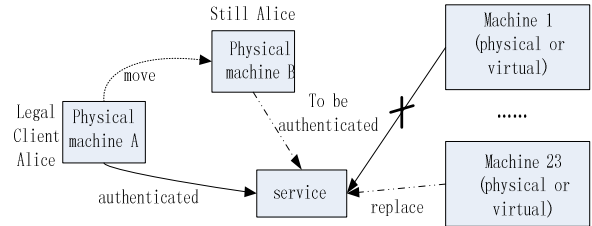
Figure 2.   Dynamic replacement of machines (virtual and physical) in cloud

Basing on the characteristics of the cloud and the security requirements of cloud services, we propose two modules to complement trusted computing: Service Authentication List, SAL , each service owns a list of information of authenticated machines; and Configuration Update Module, CUM, for dynamically replace the authenticated machines.

V.   THE MODEL FOR TRUSTED COMPUTING APPLIED INTO CLOUD

According to the elaboration in section IV about expansion and evolution of trusted computing, the model of trusted computing applied into cloud is designed as follows (see figure 3).
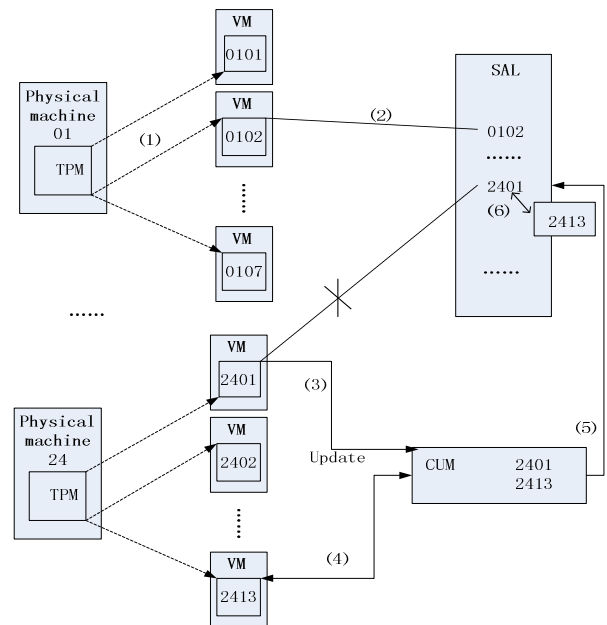
Figure 3.   The abstract model for trusted computing applied into cloud

(1) From the trusted base of a physical machine, it computes all the configurations of its virtual machines, authenticating the identities of all those virtual machines.

(2) For a certain service, it may need multiple virtual machines, then the relative machines register their identities into Service Authentication List, SAL. In future, when the service has to be authenticated, the SAL plays an important role.

(3) For various reasons, a virtual machine can no longer support one service, it needs to submit application to Configuration Update Module, CUM.

(4) After CUM has received the application, it select a new virtual machine replace the old one.

(5) CUM submits the amend information to SAL;.

(6) SAL updates the identity information.

VI.   CONCLUSIONS

This paper describes the pros and cons of trusted computing and depicts the improvements for addressing the challenges and doubts faced by trusted computing. After those, the characteristics of cloud computing and security requirements of cloud services are exhibited. Then, according to the analysis for trusted computing applied into cloud, we propose two modules for implementing trusted computing: Service Authentication List, SAL and Configuration Update Module, CUM. Finally, we give the model for trusted computing applied into cloud. More details upon the application of such model are worth future research.

REFERENCES

[1] Zhidong Shen, Wuhan University; Qiang Tong, Northeastern University; The Security of Cloud Computing System enabled by Trusted Computing Technology; 2010 2ed International Conference on Signal Processing Systems(ICSPS).

[2] Zhiyong Tan, Tiange Si, Duo Liu, Yiqi Dai; Tsinghua University; <Trusted Computing Model Based On Server-end Storage>; Journal of Tsinghua University(Science&Technology); 2009, Volume 49, No.7.

[3] TCG specification architecture overview; Specification Revision 1.4, 2nd August 2007; TCG

[4] <Cloud Computing>; Peng Liu; Electronics Industry Publication; ISBN: 978-7-121-10199-1

[5] Dengguo Feng, Min Zhang, Yan Zhang, Zhen Xu; State Key Laboratory of Information Security; <Study on Cloud Computing Security>; Journal of Software, 2011.

[6] Trusted Computing Group. TPM Main Specification[S]. 29 March 2006.Version 1.2, Revision 94,http://www. trustedcomputinggroup.org.

[7] England P,Lampson B,Manferdelli J,et al. A Trusted Open Platform[J]. IEEE Computer, 2003,36(7):55-62.

[8] TCG Specification Architecture Overview Specification (Revision1.2) [Z]https://www.trustedcomputinggroup. org/downloads/TCG_1_0_Architecture_Overview. pdf,2004.

[9] Chen,L., Landfermann,R., Loehr,H.,Rohe, M.,Sadeghi, A.R.,Stuble,C.: A protocol for property-based attestation.In:STC'06:Proceedings of the First ACM Workshop on Scalable Trusted Computing,ACM (2006)7-16.

[10] Ahmad-Reza Sadeghi,Christian Stüble,Marcel Winandy. Property-Based TPM Virtualization,Proceedings of the 11th international conference on Information Security, September 15-18, 2008,Taipei,Taiwan.

Stability Study of PD and PI Controllers in Multiple Difference Disturbances

M.A. Nur Huda[1,a*], H.A. Kasdirin[2,b] and A.G. Mohd Ruddin[3,c]
Faculty of Electrical Engineering
UTeM
Melaka, Malaysia
[a]nurhudamohdamin@student.utem.edu.my
[b]hyreil@utem.edu.my
[c]dpdruddin@utem.edu.my

Abstract—**This paper discusses the stability study of PD and PI controllers in multiple difference disturbances. The multiple difference disturbances in this paper are added to the inverted pendulum model that based on robotic leg application such as pendubot. By applying the pendubot model via MATLAB/Simulink block diagram, the performances between the model and disturbances are compared for stability in the simulation results. The simulation results showed that the PD controller could reduce and eliminate disturbances more effective than PI controller in the pendubot model. Overall, the simulation results are based on stability analysis for the degree of stability, steady state performance and transient response.**

Keywords-inverted pendulum; pendubot; PD controller; PI controller; disturbance

I. INTRODUCTION

Technology development growth fast onward through over the world including Malaysia. Each part of engineering tools is used to build and design the technology such as robot, all types vehicle, building and others. These technology especially robot application, need a control system approach to organize, monitor and stabilize any movement. Within this control system approach, the robot would operate smoothly. In case, the stability in control system is studied for the robotic leg application.

The robotic leg application gives advantage in the pendubot at which also called the arm-driven that do tasks like manipulating, moving, and painting. The tasks are usually found in industrial automation, architecture and artistic. For the industrial automation as an example, the high quality products are a primary goal to achieve. To achieve the goal, the arm-driven robots are developed more than a workers to maintain the products quality. This quality based on the stability [1] in control system that is being explained by James Clerk Maxwell, Edward John Routh, William Kingdon Clifford, Adam Prize, and Alexandr Michailovich Lyapunov in the latter half of the 19th century. The stability of the robotic leg application in this paper is the main problem for the study purposes with the PD type controller in the multiple difference disturbances.

The PD type controller refers to the familiarity terms of the Proportional (P), Integral (I) and Derivative (D) in the PID controller that are well-known as a conventional controller since the last few years ago. These terms are quite importance in define the gains P, I and D in the root-locus techniques. The root-locus techniques are based on the development of the advanced control systems from the state-space model. The state-space model refers to the inverted pendulum as apply in a robotic leg application such as the pendubot.

In a robotic leg application, the PID controller is not depending by self. Looking through from the previous research, the adapting idea of the PID controller [2], the researcher is preferred to use the PD than the PID controller as the PD controller is based on human naturally acts. Considerably from other authors paper: the study of the development of PI controller for disc speed [3], shows that PI controller is also a quite useful for eliminating and reducing disturbance.

Sometimes, the researcher looking for a wide view of research study, as posted in the article of observer control improves motion [4] by Kristin, found that the observer control enable to eliminate ringing and overshoot, and also solve the problem of the PID loops disable to do. By the way, this paper is focusing on the PD and the PI controllers for comparison stability in the multiple difference disturbances. Based on a few review of the introduction, the minority contents of the paper included the model, the controller design in general view, simulation results, conclusion, acknowledgement and references as following.

II. MODEL

The pendubot of the plant model in this paper is referring to the one of the underactuated model with the only 2encoders. Both encoders measure the speed and position for the DC servomotor and pendulum each. The plant model for a single diagram is showed as following:
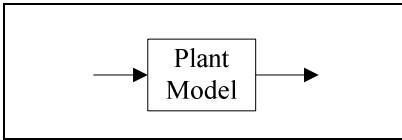
Figure 1. Plant model

Based on the plant model in Fig. 1, the input of the plant model is connected from the output sums of the controller and the disturbance, and the output of the plant model is connected from the output of pendulum. This connection is showed in Fig. 2 as following:
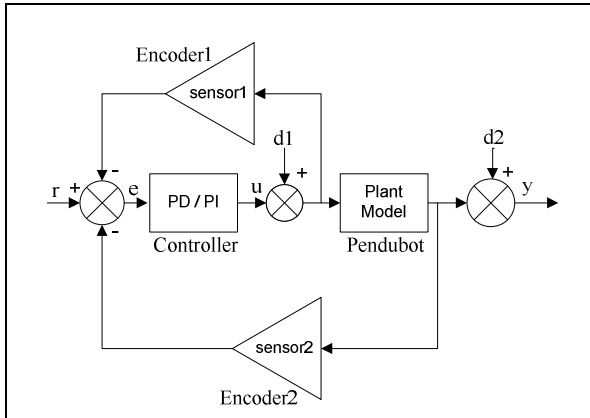
Figure 2. Block diagram for the proposed model of the control system

By the way, the block diagram in Fig. 2 is related to the proposed physical model in Fig. 3 from which is figured as the subsystem in the simulation results. This proposed physical model in Fig. 3 produces the algebra equation of inverted pendulum proposed model by applying the net torque of the DC motor T_m, the torque of disc T_d and the torque of the pendulum T_p at which is expanded into the three separate algebra equations as following:
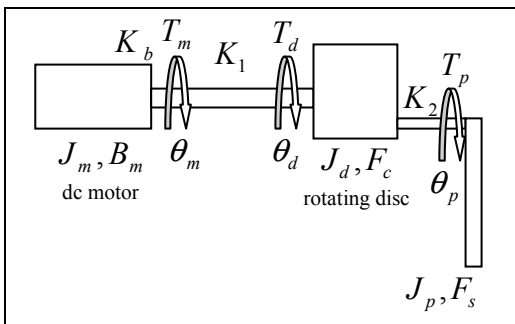
Figure 3. Inverted pendulum proposed physical plant on rotating disc

The torque of the DC motor T_m is derived as:

$$T_m = J_m \ddot{\theta}_m + B_m \dot{\theta}_m + K_1[\theta_m - \theta_d] + d1 + u$$

The torque of the disc T_d is derived as:

$$T_d = J_d \ddot{\theta}_d - K_1[\theta_m - \theta_d] - d1 - u + F_c \frac{\dot{\theta}_d}{|\dot{\theta}_d|}$$

The torque of pendulum T_p is derived as:

$$T_p = J_p \ddot{\theta}_p + K_2[\theta_d - \theta_p] \pm (F_s)|_{\dot{\theta}=0} + d2$$

These derivations are compiled in the state-space equation as following:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -a & a & 0 & -b & 0 & 0 \\ c & -e & 0 & 0 & 0 & 0 \\ 0 & d & f & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{bmatrix}^T + \begin{bmatrix} 0 \\ 0 \\ 0 \\ g \\ h \\ i \end{bmatrix}^T$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{bmatrix}^T$$

Generally, the internal and external disturbances (d_1 or d_2) are placed at the plant input or output, or both points. These disturbances [5, 6] effected to the results output (y) at which considered the closed and the open loop in the control system.

In control system, the disturbances consist like the rejection power sources such as sine, pulse generator and step, environmental source such as wind, machinery source such as vibration, and others. Sometimes, the disturbances that come from the rejection power, environmental and machinery sources are all affected to the one system only.

This paper is focusing to the timer and step disturbances only at which used the PD or the PI controller for the reduction and elimination. By using the timer disturbance, the time period of the system affects with disturbances are identified from the response of the output result. Based on the response of the output result for the timer disturbance, this paper shows that the disturbance is happened for a certain period. Other than the timer disturbance, the step disturbance in this paper presents the final value of the response till infinity, except there is another disturbance coming through the system.

Overall, the sources of disturbances place in the proposed model, with the cascaded [7] loop in control system, give high performance to the results output. The performance of the results output is based on controller (u) design as following section:

III. CONTROLLER DESIGN

This paper proposes the PI and the PD controllers to eliminate or reduce the multiple disturbances. Both controllers are designed by applying the controller transfer function to get the gain values as following:

For the PD type controller,

$$u_1 = K_p + K_d s$$

and for the PI type controller,

$$u_2 = K_p + \frac{K_i}{s}$$

By the way, these controller transfer function is compared the similarity with the compensator of the SISO tool from root locus in MATLAB application. Even though the root locus method quite accurate compared to Ziegler-Nichols Tuning Rule and other methods, but this root locus method is from the figure viewing. This figure viewing gives an advantage to the plant model with multiple disturbances. The plant model is figured as the subsystem for the simulation results as following:

IV. SIMULATION RESULTS

Via the model and the controller design, the rejection power sources or disturbances are added in Fig.4, Fig.5, Fig.6, Fig.7, Fig.8, Fig.9, Fig.10, Fig.11, and Fig.12 to compare the simulation results output. These disturbances are placed at the plant input or output, or both. For the elimination or reduction of disturbances, either the PD or PI controller is placed before the plant in the open loop or closed loop model as following.
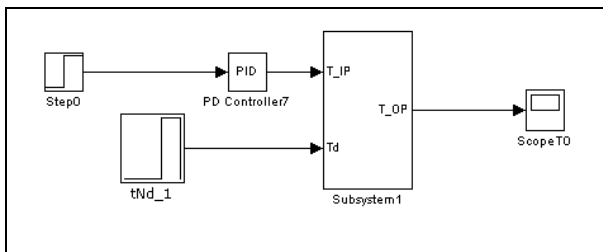
Figure 4. The open loop model with PD controller

By referring to the open loop model with the PD controller in Fig. 4, the disturbance (tnd_1) is placed at the plant input. The plant input is determined the disturbance for the torque of DC motor. Based on control theory, the open loop system does not correct the errors of the model. The errors of the model that are not measured might affect the results output even though the disturbances are eliminated by the PD controller.
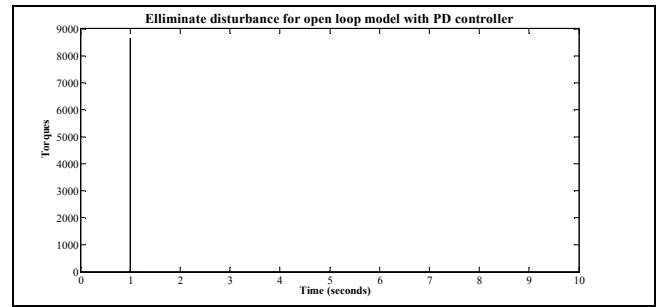
Figure 5. Output disturbance for open loop model with PD controller

The results output in Fig.5 shows that there is no disturbance at all, that gives between 4 to 6 seconds. By the way, there is the highest overshoot of the torque. In control system, the overshoot is not available to use for this model. This is because the highest or maximum torque [8] only available for a wide speed range.
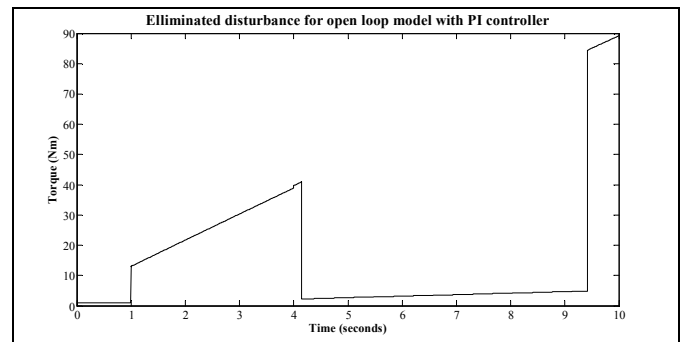
Figure 6. Output disturbance for open loop model with PI controller

In Fig.6 the output response shows a quite difference from Fig.5. The PI controller reduced the maximum overshoot value less than the PD controller for the open loop model with a longer time period after a second to 4 seconds. This response continually shows a gain after 9 to 10 second for at least 90 percent overshoot.
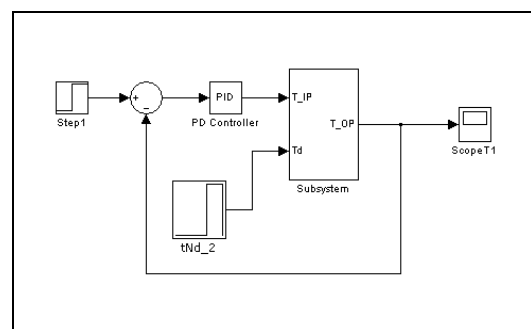
Figure 7. Single disturbance for the closed loop model with PD controller

The single disturbance (tNd_2) of the closed loop model in Fig.7 referred to the timer disturbance. The PD controller for the closed loop model is tuned with a few gains P and D. From the gains tuning, the output response is produced in Fig.8.
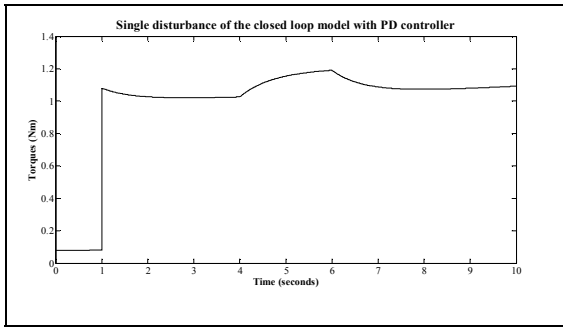
Figure 8. Output of the single disturbance for the closed loop model with PD controller

Based on the single disturbance of the closed loop model in Fig.7, Fig.8 shows that the disturbance presented in between 4 and 6 seconds with the torque range for a zero and one point two. Within the closed loop or also well known as feedback model, the error is corrected. From this output result, the response achieved the steady state of stability more than 6 seconds with more than one Newton meter, torque.
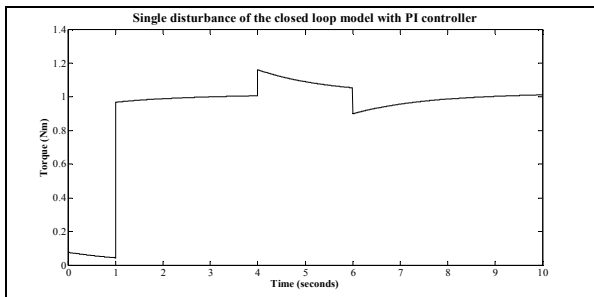
Figure 9. Output of the single disturbance for the closed loop model with PI controller

For comparison output results, the PD controller is replaced by the PI controller in Fig.9. By using the PI controller, the response shows that the increases in the rise time from which is compared to the PD controller based on the tuning of the gains P and I from a high to a low value. From the output result with the PI controller also, the response achieved the steady state earlier than the output result with the PD controller. The output result of a single disturbance with the PI controller only takes a less time from which is compared to the PD controller with the one Newton meter of the torque to stop the simulation.
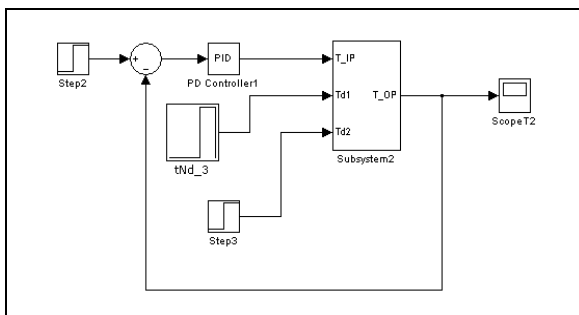
Figure 10. Multiple disturbance of timer and step with PD controller

The improvement from the open loop model to the closed loop model with a single disturbance shows that the PD and the PI controllers affected to the output results. Both models are compared also with multiple disturbances of timer and step as shown in Fig.10.
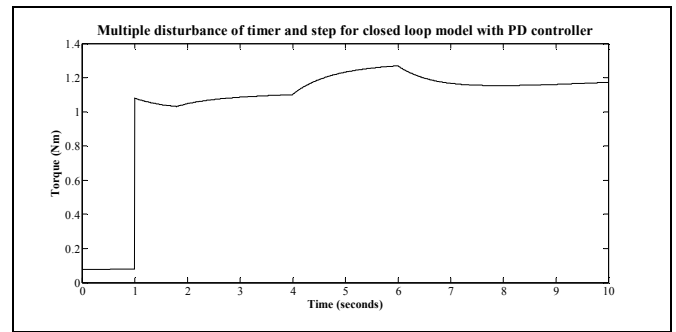
Figure 11. Output for multiple disturbance with PD controller

By using the PD controller to the closed loop model for the multiple disturbances, the output result in Fig.11 shows the response with the disturbances of timer and step. The disturbance of timer began in between 4 and 6 seconds but the disturbance of step began after a second. The multiple disturbances with the PD controller are increased the torque of the response after 6 seconds to achieve the steady state. This steady state shows that a slow response of the time from which is compared to the two previous models.
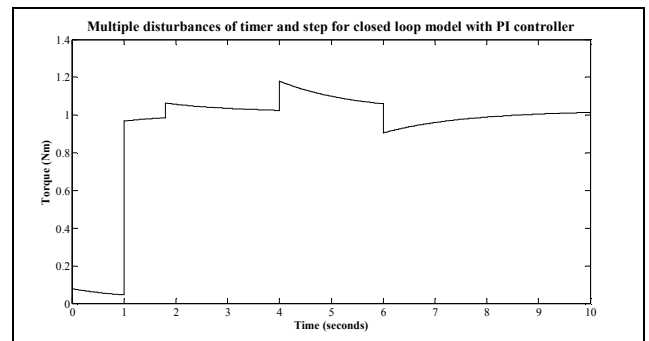
Figure 12. Output for multiple disturbance with PI controller

By the way, the output result for the multiple disturbances with the PI controller in Fig.12 gives a quite similar response of the output result for the single disturbance in Fig.9. For the output result for the multiple disturbances with the PI controller in Fig.12, the response of the two disturbances is figured clearly. The response in Fig.12 with the PI controller is also compared with the response in Fig.11 with the PD controller. In Fig.11, the disturbances are not clearly figured for the response but in Fig.12 the disturbances are clearly figured. These output results for the response in Fig.11 and Fig.12 are caused by the low and the high rise time. Each output results are summarized in Table 1 as following:

TABLE 1. Comparison stability for PD and PI controller

Stability Analysis	Controllers	
	Proportional and Derivative (PD)	*Proportional and Integral (PI)*
degree of system stability	Improves the system stability	As this PI controller increases the compensated by 1, the system show less stable
steady state performance	To satisfy the steady state, the value of Kp must be suited. The constant of the steady state error based on the time, Td=0 and derivative portion provides no input.	Improves the steady state as its infinite gain at zero frequency
transient response	Improves the transient response with the reduction of the rise time. From the transient response, the maximum overshoot is reduced better than PI controller.	The suitable values of Kp and Ti are selected to improve the transient response. From the transient response, the rise time is increasing and the maximum overshoot is reduced but shows clearer signal.

V. CONCLUSIONS

The multiple disturbances for the proposed model show that a better response of the simulation results output with the cascaded loop. These responses reduce disturbance with the gains tuning of the PD or PI controller for the inverted pendulum proposed model, but low performance for the open loop model from which is compared to the closed loop model. By the way, this simulation results output is not denied at all as the single disturbance gives less problems to the response from which is compared to the multiple disturbances for the stability study. Therefore, the experimental results are followed up for further study.

ACKNOWLEDGEMENT

REFERENCES

[1]   N. S.Nise, *Control System Engineering*: John wiley & Sons, 2008.
[2]   A. Kadis, *et al.*, "Modelling, simulation and control of an electric unicycle," 2010.
[3]   M. A. Nur Huda, *et al.*, "Development of PI Controller for Disc Speed," in *IEEE Symposium on Industrial Electronics and Applications (ISIEA 2011)*, 2011, pp. 652-656.
[4]   Kristin. (2008, Observer Control Improves Motion
[5]   B. C.Kuo and F. Golnaraghi, *Automatic Control Systems*, 8 ed.: John Wiley & Sons Inc, 2003.
[6]   K. Ogata, *Modern Control Engineering*, Fourth ed. New Jersey: Prentice Hall, 2002.
[7]   Y. Zhong and Y. Luo, "Comparative Study of Single-loop Control and Cascade Control of Third-order Object," *Procedia Engineering,* vol. 15, pp. 783-787, 2011.
[8]   J. Zhi and H. Fengyou, "Maximum torque control of servo motor for wide speed range," in *Electronic and Mechanical Engineering and Information Technology (EMEIT), 2011 International Conference on*, 2011, pp. 3353-3357.

APPENDIXES

Parameters are defined as following:

J_m = DC motor moment inertia
J_d = disc moment inertia
J_p = pendulum moment inertia
B_m = viscous friction
F_c = Coulomb friction
F_s = static friction
K_1 = torque constant / back e.m.f
K_2 = spring constant
$d1$ = disturbance 1
$d2$ = disturbance 2
u = PD or PI controller

$$x_1 = \theta_m \qquad a = \frac{K_1}{J_m} \qquad f = \frac{(m_p g l_p - K_2 - m_p l_p a_p)}{J_p}$$

$$x_2 = \theta_d \qquad b = \frac{B_m}{J_m} \qquad g = \frac{1}{J_m}$$

$$x_3 = \theta_p \qquad c = \frac{K_1}{J_d} \qquad h = \frac{1}{J_d}$$

$$x_4 = \dot{\theta}_m$$

$$x_5 = \dot{\theta}_d \qquad d = \frac{K_2}{J_d} \qquad i = \frac{1}{J_p}$$

$$x_6 = \dot{\theta}_p \qquad e = \frac{(K_1 + m_d l_d a_d)}{J_d}$$

$$\dot{x}_1 = \dot{\theta}_m = x_4$$

$$\dot{x}_2 = \dot{\theta}_d = x_5$$

$$\dot{x}_3 = \dot{\theta}_p = x_6$$

$$\dot{x}_4 = \ddot{\theta}_m = -\frac{B_m}{J_m}x_4 - \frac{K_1}{J_m}x_1 + \frac{K_1}{J_m}x_2 + \frac{1}{J_m}(T_m \mp \left(F_{s_1}\right)_{\dot{\theta}_p=0} r_1)$$

$$\dot{x}_5 = \ddot{\theta}_d = \frac{K_1}{J_d}x_1 - \frac{1}{J_d}(K_1 + m_d l_d a_d)x_2 + \frac{1}{J_d}F_c r_1$$

$$\dot{x}_6 = \ddot{\theta}_p = \frac{K_2}{J_p}x_2 + \frac{1}{J_p}(m_p g l_p - K_2 - m_p l_p a_p)x_3 \pm \frac{1}{J_p}\left(F_{s_2}\right)_{\dot{\theta}_p=0} r_2$$

Feed-Forward Control of Non-Minimum Phase XY Table System Using Trajectory Fuzzy-RLS Hybrid ZPETC

Ramli Adnan[1], Michael Jackson Patrick[2], Norlela Ishak[3]

Faculty of Electrical Engineering, UiTM Malaysia, 40450 Shah Alam, Selangor, Malaysia
[1]ramli324@salam.uitm.edu.my
[2]mj_29jack@yahoo.com.sg
[3]norlelaishak@salam.uitm.edu.my

Abstract— **A tracking control for non-minimum phase (NMP) XY Table could not use directly the inverse closed-loop transfer function as the feed-forward controller. This is due to unstable zeros of the closed-loop transfer function will become unstable poles at the feed-forward controller. Thus, ZPETC was introduced to encounter this unstable zeros problem. The ZPETC that was introduced by Tomizuka has unity gain for low frequencies, but not at high frequencies. Hence, this paper proposed a scheme called Trajectory Fuzzy-RLS Hybrid ZPETC without factorization of zeros polynomial where the optimum controller parameters are obtained offline using recursive least square (RLS) method and further retune online through fuzzy inference system (FIS). The tracking performance was compared with the design using Trajectory-Adaptive ZPETC. The simulation and real-time controls of the NMP XY Table show that the tracking performances for circular movement are significantly improved.**

Keywords— **Feed-forward Control; Non-minimum Phase System; Motion Control; ZPETC; Adaptive Control, Fuzzy Robust Control.**

I. INTRODUCTION

XY Table has been widely used and of important part of many Computer Numeric Controlled (CNC) processing facilities, e.g. the work feeder of CNC lathe, CNC milling and drill press, the worktable of laser processing, welding, dispenser, bonding, packing, drilling, laser cutting and painting [1]. In general, XY Table is composed of X-axis and Y-axis motion mechanism where each motion axis is driven by individual actuator such as DC servo-motor through high-precision ball-screw.

A two-degree of freedom (2DOF) controller is shown in Figure 1 and the perfect tracking control (PTC) can be achieved by introducing a feed-forward controller that acts as an inverse of system closed-loop transfer function, cancelling all poles and zeros of the closed-loop system.

Figure 1. Two-degrees-of-freedom Controller

The PTC is desired and important element for XY Table control system and has been widely studied in [1-10]. PTC can provide the overall transfer function of unity between the desired output r_k and the actual output y_k. This will be true when the plant is minimum phase. A tracking control for non-minimum phase could not use directly the inverse of closed-loop transfer function as a feed-forward controller. This is due to unstable zeros of the closed-loop transfer function will become unstable poles at the feed-forward controller. Thus, ZPETC was introduced to encounter this unstable zeros problem. The ZPETC that was introduced by Tomizuka [2,3] has unity gain for low frequencies, but not at high frequencies [4]. In resolving for broader frequency tracking control system, there has been many research works in this area [5].

In [5,6] when one of zeros located outside unity circle, low-order ZPETC cannot be used to provide unity gain for the overall transfer function between the desired output and the actual output. In resolving these problems, there has been some research works in this area. Trajectory-Adaptive ZPETC was suggested in [7,8]. Thus, satisfactory performance can still be achieved with relatively low-order ZPETC. By reducing the NMP zeros problems, e.g. gain and phase error, the Adaptive ZPETC shows interesting tracking performance. Unfortunately, due to the feed-forward nature of ZPETC such that the overall performance of the controlled system will be affected if there is plant-model mismatch [7,8,9,10].

Fortunately, to improve the tracking performance by minimizing the plant-model mismatch, a scheme called as Trajectory Fuzzy-RLS Hybrid ZPETC is proposed. To do so, a FIS is designed and added to the Adaptive ZPETC structure. The tracking performances for both controllers are then been compared. Simulation and real-time studies are done on XY Table to verify the effectiveness of the proposed scheme. The

simulation and real-time studies were done in MATLAB environment.

This paper was organized in the following manner: Section II describes plant model; Section III describes reference trajectory; Section IV describes Trajectory Fuzzy-RLS ZPETC; Section V is on results and discussion; and finally, Section VI is the conclusions.

II. PLANT MODEL

The XY Table used in this paper is shown in Figure 2. The NMP plant model was derived from open-loop input-output test. The open-loop transfer function of the plant is approximated using MATLAB system identification toolbox in the form of ARX331 with input-output signals sampled at $80ms$. The X-axis plant model is given by Eq. (1) and for Y-axis plant model is given by Eq. (3).

Figure 2. XY Table system

A. X-Axis Plant Model:

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{0.03698z^{-1} + 0.1217z^{-2} + 0.05178z^{-3}}{1 - 0.6479z^{-1} - 0.5451z^{-2} + 0.1933z^{-3}} \quad (1)$$

From Eq. (1), the zeros polynomial obtained are

$$B_c(z^{-1}) = 1 + 3.2910z^{-1} + 1.4002z^{-2} \quad (2)$$

$$B_c(z) = 1 + 3.2910z^{1} + 1.4002z^{2}$$

When Eq. (2) is factorized, the locations of zero are at $z = -0.5021$ and $z = -2.7889$. This means that the obtained model is a non-minimum phase (NMP) model with one zero situated outside and far away from the unit circle.

B. Y-Axis Plant Model:

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{0.03978z^{-1} + 0.1054z^{-2} + 0.03507}{1 - 0.7495z^{-1} - 0.4557z^{-2} + 0.2056z^{-3}} \quad (3)$$

From Eq. (3), the zeros polynomial obtained are

$$B_c(z^{-1}) = 1 + 2.6496z^{-1} + 0.8816z^{-2} \quad (4)$$

$$B_c(z) = 1 + 2.6496z^{1} + 0.8816z^{2}$$

When Eq. (4) is factorized, the locations of zero are at $z = -0.3902$ and $z = -2.2594$. This means that the obtained model is also a non-minimum phase model with one zero situated outside and far away from the unit circle.

III. REFERENCE TRAJECTORY

A circular movement of $30mm$ in radius is shown in Figure 3. It also shows the reference trajectory of respective individual axis used for plant worktable to move in circular. Both individual axis reference trajectories are sinusoidal wave signal with maximum amplitude of $+30mm$, minimum amplitude of $-30mm$ and frequency of 0.4 radians per second. All simulation studies and real-time experiment will be based on this reference trajectory.
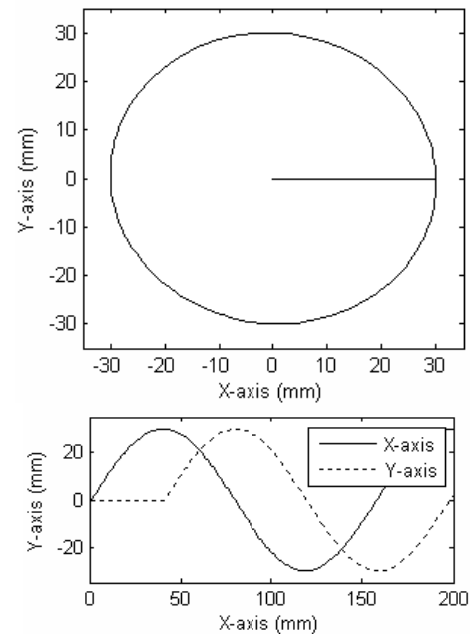
Figure 3. Circular movement and its individual axis trajectory signals

IV. TRAJECTORY FUZZY-RLS HYBRID ZPETC

Nowadays, fuzzy systems are the best choice to transform logic statements (like if-then) into automatic control strategies in order to design controller for uncertainties of controlled systems due to plant-model mismatch caused by parameter variations [11]. However, fuzzy logic systems have some disadvantages, for example, fuzzy rules must be preset by time consuming trial and error methods due to lack of analytical techniques. To do so, a fuzzy inference system (FIS) is designed and added to the previous Adaptive ZPETC as a hybrid system. The ideal hybrid system would be robust against parameter variations and lead to better tracking performances [11].

The structure of the proposed Trajectory Fuzzy-RLS Hybrid ZPETC using the combination of FIS and recursive least square

(RLS) adaptive parameter estimation methods is given in Figure 4 for real-time experiment and Figure 5 for simulation studies. Note that the memory block is the d-step

delay z^{-d} used to represent tracking error e_k and FIS at d-step, which is the same as the gain compensation filter. Thus, it preventing covariance blow-up.
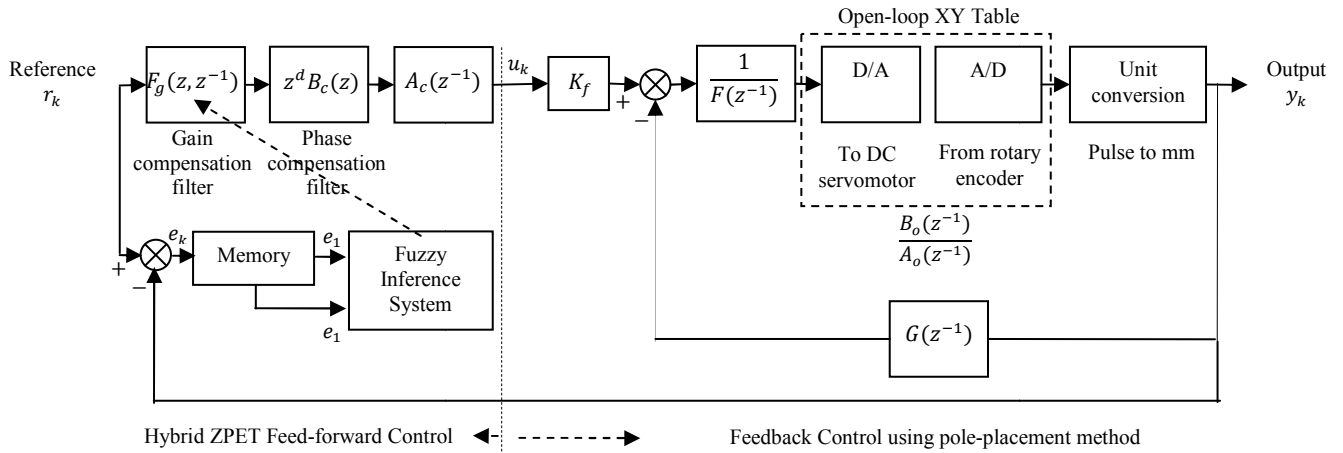
Figure 4. Trajectory Fuzzy-RLS Hybrid ZPETC structure for real-time control
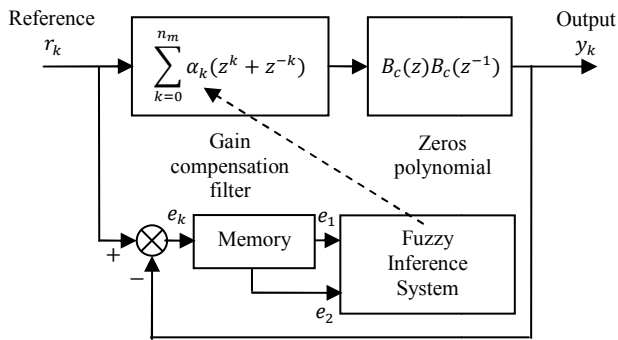
Figure5. Trajectory Fuzzy-RLS Hybrid ZPETC structure for simulation control

There are two inputs to the FIS, the errors e_1 and e_2. The respective membership functions (MF) of each input is shown in Figure 6. For each input and output, five MF were used. The respective MF of each input and output are shown in Figure 6.

Figure 6. The membership functions of α_k

Figure 7. Fuzzy inference system of α_k

There are six outputs from the Fuzzy Inference System (FIS) given in Figure 7 as respectively new optimum set of α_k. Note that the old optimum set of α_k are based on Table I and Table II.

TABLE I.   Optimal set of α_{old} for X-axis

k	α_k
0	0.1182
1	-0.1569
2	0.0746
3	-0.02533
4	0.005487
5	-0.0005671

TABLE II. Optimal set of α_{old} for Y-axis

k	α_k
0	0.5356
1	-0.2606
2	-0.02622
3	0.03185
4	-0.008374
5	0.000865

These old values of optimum set of α_k are then tuned by FIS through respective MF that providing rule maps to the e_1, e_2, and the outputs. New α_{new} are computed using Eq. (5).

$$\alpha_{new} = (\alpha_{max} - \alpha_{min})\alpha_{old} + \alpha_{min} \quad (5)$$

Through MF modifications, Eq. (5) is use to calibrate the old optimum set of α_k through FIS, hence, α_{new} will be computed. Note that each optimum set of α_k has its own calibration. This part is important so that feasible rules based on high frequency efficiency were obtained [11]. The relationship between the input and output of the FIS shown in Figure 7 are listed in Table III. There were five linguistic variables that had been set, so 25 Mamdani rules were applied in the FIS. The aggregation and defuzzification methods used are respectively max-min and centroid method. The whole systems were developed using MATLAB environment.

TABLE III. Fuzzy inference system rules

e_1 \ e_2	NB	NS	ZE	PS	PB
NB	S	S	MS	MS	M
NS	S	MS	MS	M	MB
ZE	MS	MS	M	MB	MB
PS	MS	M	MB	MB	B
PB	M	MB	MB	B	B

V. RESULTS AND DISCUSSION

The simulation studies and real-time control experiments were done in MATLAB environment. The simulation results for Trajectory-Adaptive ZPETC and Trajectory Fuzzy-RLS Hybrid ZPETC using 5th order controller to move in circular motion are shown in Figure 8 and Figure 9. It can be observed that the resulting circular movement for both controllers and reference circular movement could hardly be seen the differences. This is explained by the smooth circular movement in Figure 8 and Figure 9.
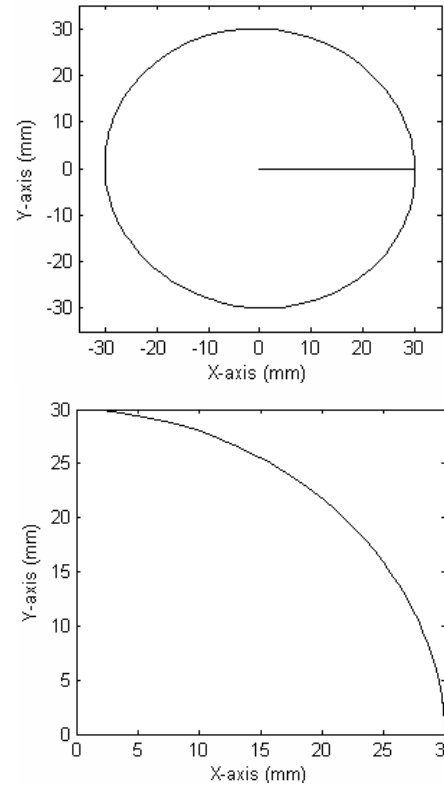
Figure 8. Simulation result of circular movement for 5th order Trajectory-Adaptive ZPETC
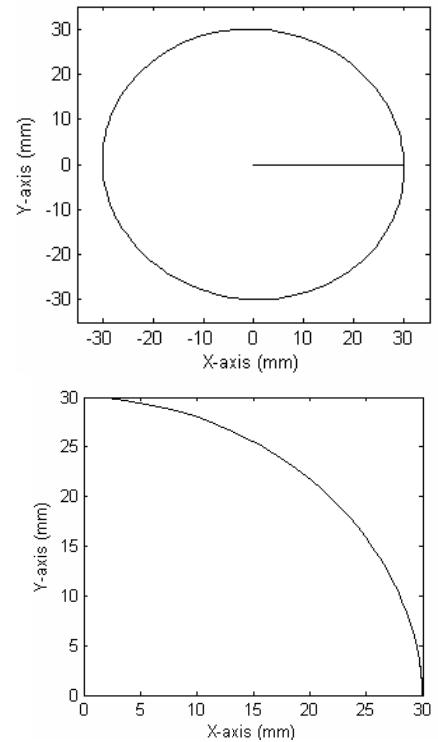
Figure 9. Simulation result of circular movement for 5th order Trajectory Fuzzy-RLS Hybrid ZPETC

The simulation results using both controllers are given in Table IV. We can observe that using the Hybrid ZPETC, the tracking error is much smaller. The proposed Hybrid ZPETC shows better RMSE about 80.03% improvement for X-axis and 75.56% improvement for Y-axis.

TABLE IV. Simulation Results (Tracking Performance)

	NMP XY-Table System	
	X-axis RMSE (mm)	Y-axis RMSE (mm)
Trajectory adaptive ZPETC	8.0414×10^{-5}	9.4267×10^{-5}
Trajectory Fuzzy-RLS Hybrid ZPETC	1.6058×10^{-5}	2.3043×10^{-5}

The real-time experiment results for both controllers using 5th order controller to move in circular motion are shown in Figure 10 and Figure 11. It can be observed that the resulting circular movement for Hybrid ZPETC is better than Adaptive ZPETC. This is explained by the rough circular movement of Figure 10 for Adaptive ZPETC. This is due to the feed-forward nature of ZPETC such that the overall performance of the controlled system will be affected if there is plant-model mismatch. Fortunately, the Hybrid ZPETC produces better circular movement as shown in Figure 11. Thus, it is confirmed that the proposed Hybrid ZPETC is better than the Adaptive ZPETC in terms of tracking performance and also robustness to plant-model mismatch.
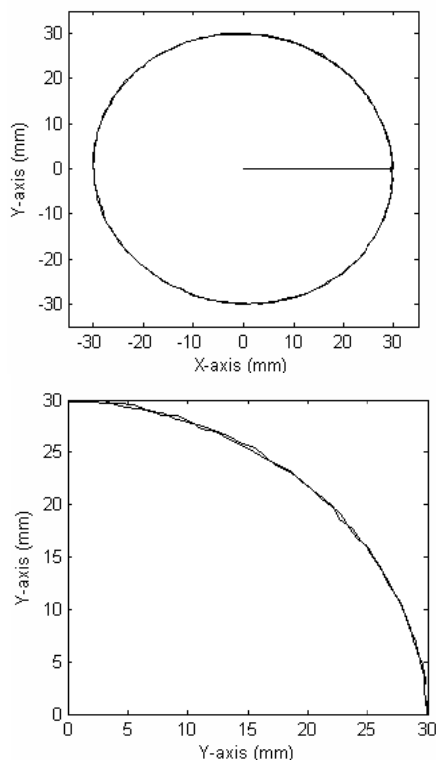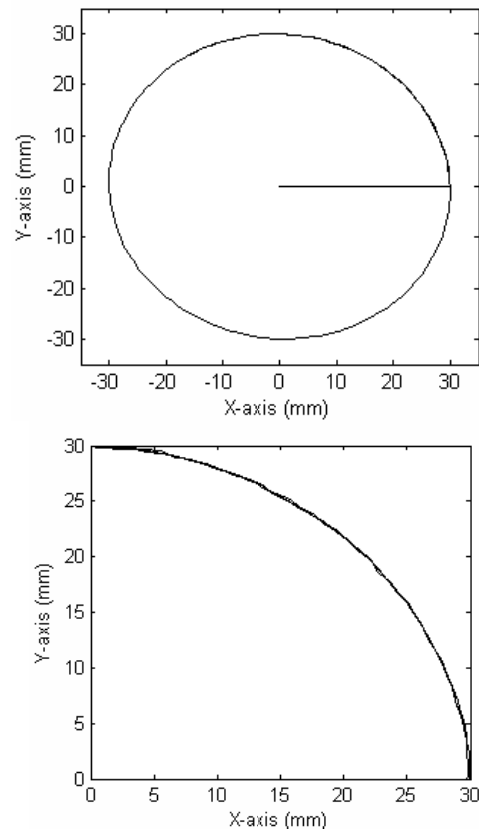
Figure 11. Real-time experiment result of circular movement for 5th order Trajectory Fuzzy-RLS Hybrid ZPETC

The experimental results of using both controllers are given in Table V. We can observe that using the Hybrid ZPETC, the tracking error is much smaller. The proposed Hybrid ZPETC shows better RMSE about 40.28% improvement for X-axis and 52.77% improvement for Y-axis.

TABLE V. Experimental Results (Tracking Performance)

	NMP XY-Table System	
	X-axis RMSE (mm)	Y-axis RMSE (mm)
Trajectory-Adaptive ZPETC	0.1405	0.2206
Trajectory Fuzzy-RLS Hybrid ZPETC	0.0839	0.1042

VI. CONCLUSION

The simulation studies and real-time digital tracking control of NMP XY Table using 5th order Trajectory-Adaptive ZPETC and 5th order Trajectory Fuzzy-RLS Hybrid ZPETC without factorization of zeros polynomials have been presented. In simulation studies, both controllers show good tracking performance by successfully reducing the NMP zeros problems. Hence, better circular movement was achieved. However, in real-time studies, there were large differences

Figure 10. Real-time experiment result of circular movement for 5th order Trajectory-Adaptive ZPETC

between both results. These are due to the nature of ZPETC which is sensitive to plant-model mismatch causes by parameter variations. Fortunately, improvement in tracking performance by minimizing the plant-model mismatch was achieved using the proposed Hybrid ZPETC. Thus, it is confirmed that the proposed hybrid controller is better than Adaptive ZPETC in terms of tracking performance and also robustness to plant-model mismatch.

REFERENCES

[1] Michael, J.P., Norlela, I., Mohd Hezri, F., Mazidah, T. & Ramli, A., "Modeling and Controller Design for Non-Minimum Phase System with Application to XY-Table". IEEE.7th Control & System Graduate Research Colloquium (ICSGRC), pp. 113-118, 2011.
[2] Tomizuka, M. "Zero Phase Error Tracking Algorithm for Digital Control", ASME J. of Dyn, Syst., Meas., and Control, vol. 109, pp. 65-68, 1987.
[3] Tomizuka, M. "On the Design of Digital Tracking Controllers", ASME J. of Dyn, Syst., Meas., and Control, vol. 115, pp. 412-418, 1993.
[4] Haack, B. & Tomizuka, M., "The Effect of Adding Zeros to Feedforward Controller". ASME J. of Dyn, Syst., Meas., and Control, vol. 113, pp. 6-10, 1991.
[5] Norlela, I., Mazidah, T., Hashimah, I., Michael, J.P., Yahaya, M.S. & Ramli, A. "Non-Minimum Phase Model of Vertical Position Electro-Hydraulic Cyclinder for Trajectory ZPETC". Int. Journal on Smart Sensing & Intelligent Systems, vol. 5, pp. 504-520, 2012.
[6] Ramli, A., Abd. Manan, S., Noritawati, M. T., Mohd Hezri, F. & Mohd Marzuki, M., "Trajectory Zero Phase Error Tracking Control Using Comparing Coefficients Method". IEEE.5th International Colloquium on Signal Processing & Its Application (CSPA), pp. 385-390, 2009.
[7] Ramli, A., Abd. Manan. & Mohd Marzuki, M., "Real-time Control of Non-minimum Phase Electro-Hydraulic System using Trajectory Adaptive ZPETC". IEEE.7th International Colloquium on Signal Processing & Its Application (CSPA), pp. 72-76, 2011.
[8] Mustafa, M. M., "Trajectory-Adaptive Digital Tracking Controllers for Non-minimum Phase Systems without Factorisation of Zeros".IEE Proc, Control Theory Appl., vol. 149, no. 2, pp. 157-162, 2002.
[9] Norlela, I., Hashimah, I., Mohd Hezri, F., Mazidah, T., Ramli, A. & Yahaya, M.S., "Tracking Control for Electro-Hydraulic Actuator using ZPETC". IEEE.7th Control & System Graduate Research Colloquium (ICSGRC), pp. 94-97, 2011.
[10] Hashimah, I., Norlela, I., Mohd Hezri, F.,Mazidah, T.&Ramli, A., "Positioning and Tracking Control of an XY Table". IEEE.7th International Conference on Control System, Computing and Engineering (ICCSCE), pp. 343-347. 2011.
[11] Ramli, A., Mazidah,T.,Norlela, I.,Hashimah,I. &Mohd Hezri,FR., "Self-tuning Fuzzy PID Controller for Electro-Hydraulic Cylinder". IEEE.7th International Colloquium on Signal Processing & Its Application (CSPA), pp. 395-398, 2011.
[12] Yeh, S.S. and Hsu, P.L., "An Optimal and Adaptive Design of the Feedforward Motion Controller", IEEE/ASME J. Mechatronics, vol. 4, no. 4, pp. 428–439, 1999.
[13] Torfs, D., Swevers, J. & De Schutter, J., "Quasi-Perfect Tracking Control of Non-Minimal Phase Systems". 30th Conference on Decision and Control, pp. 241-244, 1991.
[14] Torfs, D., Swevers, J. & De Schutter, J., "Extended Bandwidth Zero Phase Error Tracking Control of Non-Minimum Phase System". ASME J. of Dyn, Syst., Meas., and Control, vol. 114, pp. 347-351, 1992.
[15] Xia, J.Z. & Menq, C.H. "Precision Tracking Control of Non-Minimum Phase System with Zero Phase Error", Int J. Control, vol. 61, no. 4, pp. 791-807, 1995
[16] Mohd Nasir, T., Ramli, A., and Mohd Hezri, F., Practical System Identification, Faculty of Electrical Engineering, UiTM, 2007. ISBN: 978-967-5070-03-7.
[17] Landau. I.D, System Identification and Control Design using P.I.M + software, Prentice Hall, Englewood Cliffs, N.J, 1990.

Real-Time Control of Non-Minimum Phase XY Table System using Trajectory-Adaptive ZPETC

Ramli Adnan[1], Michael Jackson Patrick[2], Norlela Ishak[3]

Faculty of Electrical Engineering, UiTM Malaysia, 40450 Shah Alam, Selangor, Malaysia
[1]ramli324@salam.uitm.edu.my
[2]mj_29jack@yahoo.com.sg
[3]norlelaishak@salam.uitm.edu.my

Abstract— **In XY Table system, perfect tracking control (PTC) with zero tracking error is desired. This can be achieved by introducing a feed-forward controller that acts as an inverse of the system closed-loop transfer function and cancelling all poles and zeros of the closed-loop system. Adding feed-forward controller in axis controller can prompt the response of the single axis. Consequently, minimize the tracking error of each individual axis. This result an overall transfer functions of unity between the desired output and the actual output. In non-minimum phase (NMP) system where one of the zeros located outside the unit circle, unfortunately, this feed-forward controller could not be implemented as this would result an unstable tracking control. Thus, this paper proposes a Trajectory-Adaptive ZPETC without factorization of zeros polynomial scheme in which the optimum set of α_k are automatically tuned online to reduce this NMP zeros problem. Simulation and experimental results of the real-time control of XY Table show that the tracking performance for lemniscates movement is significantly improved.**

Keywords— **Feed-forward Control, Non-minimum Phase System, ZPETC, Adaptive Control.**

I. INTRODUCTION

XY Table has been widely used and of important part of many Computer Numeric Controlled (CNC) processing facilities, e.g. the work feeder of CNC lathe, CNC milling and drill press, the worktable of laser processing, welding, dispenser, bonding, packing, drilling, laser cutting and painting [1]. In general, XY Table is composed of X-axis and Y-axis motion mechanism where each motion axis is driven by individual actuator such as DC servo-motor through high-precision ball-screw.

A two-degree of freedom (2DOF) controller is shown in Figure 1 and the perfect tracking control (PTC) can be achieved by introducing a feed-forward controller that acts as an inverse of system closed-loop transfer function, cancelling all poles and zeros of the closed-loop system.

Figure 1. Two-degrees-of-freedom Controller

The PTC is desired and important element for XY Table control system and has been widely studied in [1-13]. PTC can provide the overall transfer function of unity between the desired output r_k and the actual output y_k. This will be true when the plant is minimum phase. But when the plant is non-minimum phase (NMP), the unstable zeros of the plant will become unstable poles to the feed-forward controller transfer function and thus, creating internal instability [8].

In NMP system where one of the zeros located outside the unit circle, the PTC could not be implemented as this would result an unstable tracking control. The phase and gain errors that caused by the NMP or unstable zero are problematic, however, it can be resolved. The zero phase error tracking controller (ZPETC), developed by Tomizuka [9,10] has attracted attention many researchers as an effective and simple remedy to the problem due to NMP zero. By eliminating phase error caused by NMP zeros, ZPETC displays good tracking performance. Unfortunately, ZPETC cancels only the phase error, as the gain error which cannot be eliminated by ZPETC becomes larger for fast tracking control. This gain error causes undesirable effects to the tracking performance. In resolving these problems, there has been many research works in this area [3,4,5,6,7,8,11,12,13].

This paper proposes a Trajectory-Adaptive ZPETC without factorization of zeros polynomial scheme in which the optimum set of α_k are automatically tuned online to reduce this NMP zeros problem. Simulation and real-time studies are done on XY Table to verify the effectiveness of the proposed scheme. The simulation and real-time studies were done in MATLAB environment.

This paper was organized in the following manner: Section II describes plant model; Section III describes reference trajectory; Section IV describes Trajectory-Adaptive ZPETC without factorization of zeros polynomial; Section V describes feedback control; Section VI is on results and discussion; and finally, Section VII is the conclusions.

II. PLANT MODEL

The XY Table used in this paper is shown in Figure 2. The NMP plant model was derived from open-loop input-output test. The open-loop transfer function of the plant is approximated using MATLAB system identification toolbox in the form of ARX331 with input-output signals sampled

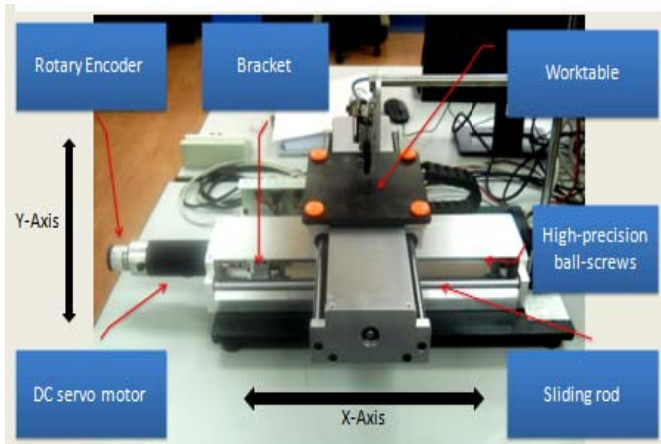at $80ms$. The X-axis plant model is given by Eq. (1) and for Y-axis plant model is given by Eq. (3).

Figure 2. XY Table system

A. X-Axis Plant Model:

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{0.03698z^{-1} + 0.1217z^{-2} + 0.05178z^{-3}}{1 - 0.6479z^{-1} - 0.5451z^{-2} + 0.1933z^{-3}} \quad (1)$$

From Eq. (1), the zeros polynomial obtained are

$$B_c(z^{-1}) = 1 + 3.2910z^{-1} + 1.4002z^{-2} \quad (2)$$

$$B_c(z) = 1 + 3.2910z^1 + 1.4002z^2$$

When Eq. (2) is factorized, the locations of zero are at $z = -0.5021$ and $z = -2.7889$. This means that the obtained model is a non-minimum phase (NMP) model with one zero situated outside and far away from the unit circle.

B. Y-Axis Plant Model:

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{0.03978z^{-1} + 0.1054z^{-2} + 0.03507}{1 - 0.7495z^{-1} - 0.4557z^{-2} + 0.2056z^{-3}} \quad (3)$$

From Eq. (3), the zeros polynomial obtained are

$$B_c(z^{-1}) = 1 + 2.6496z^{-1} + 0.8816z^{-2} \quad (4)$$

$$B_c(z) = 1 + 2.6496z^1 + 0.8816z^2$$

When Eq. (4) is factorized, the locations of zero are at $z = -0.3902$ and $z = -2.2594$. This means that the obtained model is also a non-minimum phase model with one zero situated outside and far away from the unit circle.

III. REFERENCE TRAJECTORY

A lemniscates movement of $30mm$ in radius is shown in Figure 3. It also shows the reference trajectory of respective individual axis used for plant worktable to move in lemniscates. For X-axis, the reference trajectory is a sinusoidal wave signal with maximum amplitude of $+30mm$, minimum

amplitude of $-30mm$ and frequency of 0.2 radians per second. For Y-axis, the reference trajectory is a sinusoidal wave signal with maximum amplitude $+30mm$, minimum amplitude of $-30mm$ and frequency of 0.4 radians per second. All simulation studies and real-time experiment will be based on this reference trajectory.
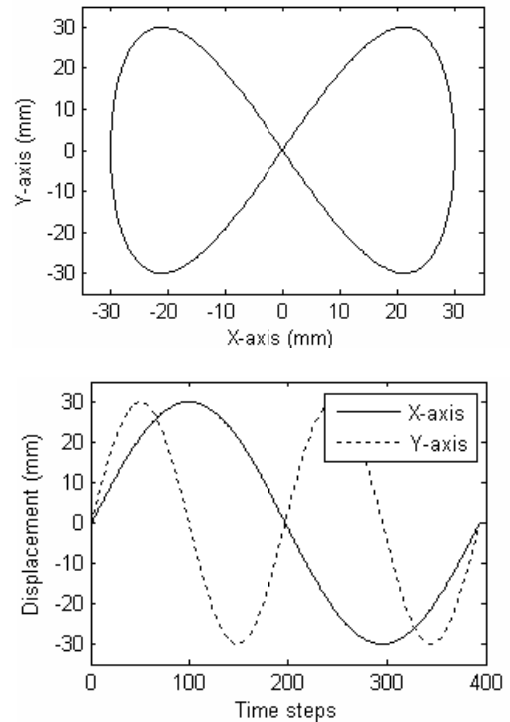
Figure 3. Lemniscates movement and its individual axis trajectory signals

IV. TRAJECTORY-ADAPTIVE ZPETC

Referring to Figure 1, let the closed-loop transfer function (without feed-forward control) be given as:

$$G_c(z^{-1}) = \frac{z^{-d}B_c(z^{-1})}{A_c(z^{-1})} \quad (5)$$

where z^{-d} represents a d-step delay and

$$B_c(z^{-1}) = b_0 + b_1z^{-1} + \cdots + b_{n_b}z^{-n_b}$$

$$A_c(z^{-1}) = 1 + a_1z^{-1} + \cdots + a_{n_a}z^{-n_a}$$

$$b_0 \neq 0$$

The structure of the Trajectory-Adaptive ZPETC without factorization of zeros polynomial is shown in Figure 4 and Figure 5. The controller design mainly focuses on the selection of an appropriate gain compensation filter (F_g), to ensure that the overall gain is unity within the frequency spectrum of reference trajectory. In order to ensure that the F_g does not introduces any phase error, the same structure of symmetrical gain compensation filter $F_g(z, z^{-1})$ as proposed by [6,7] will be used here.
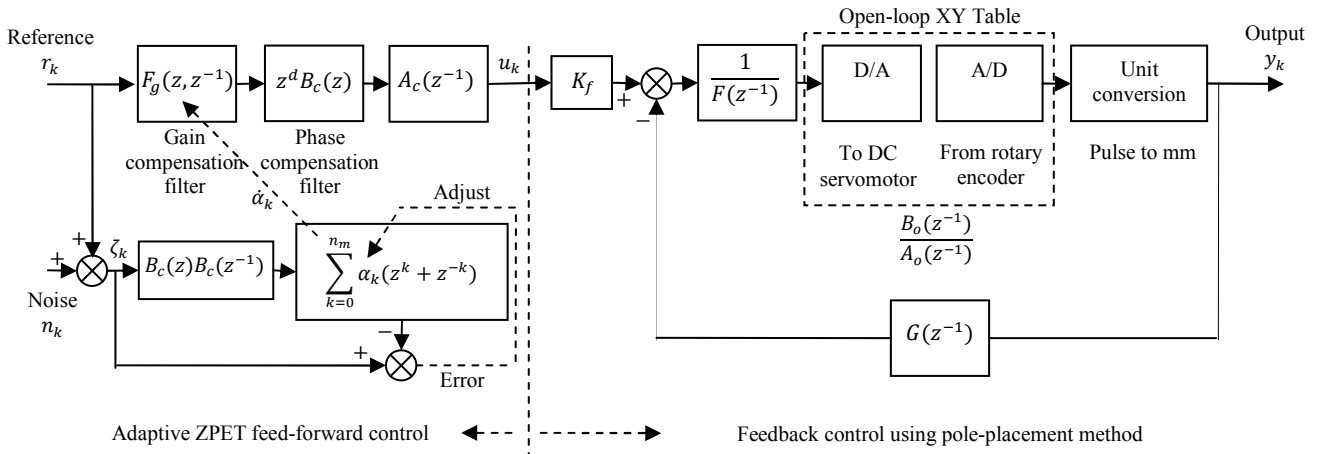
Figure 4. Trajectory-Adaptive ZPETC structure for real-time control

The symmetrical gain compensation filter is represented by the following equation:

$$F_g(z, z^{-1}) = \sum_{k=0}^{n_\alpha} \alpha_k(z^k + z^{-k}) \qquad (6)$$

n_α : Filter order

Due to the effect of poles cancellation to the ZPETC structure of Figure 4, the control structure for simulation purposes can be simplified as given in Figure 5. Thus, the implementation of tracking control by simulation does not require the whole plant model transfer function. What was needed here was only the gain compensation filter and zeros polynomial equations of the plant model.
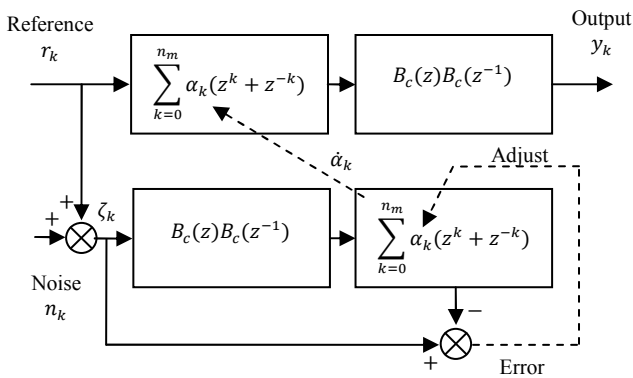
Figure 5. Trajectory-Adaptive ZPETC structure for simulation control

The cost function to represents the error between the desired and actual frequency response of Figure 4 and Figure 5 is given by:

$$J(\alpha_i) = \left\| 1 - B_c(z^{-1})B_c(z) \sum_{k=0}^{n_\alpha} \alpha_k(z^k + z^{-k}) \right\|_{l_2} \qquad (7)$$

$$J(\alpha_k) = \left\| W(z^{-1})\left(1 - B_c(z^{-1})B_c(z) \sum_{k=0}^{n_\alpha} \alpha_k(z^k + z^{-k})\right) \right\|_{l_2} \qquad (8)$$

It is very unlikely that Eq. (7) will vanish for a finite n_α, implying the term inside $\|.\|$ cannot be made zero at all frequencies [6]. To give emphasis to certain frequency band of interest, a frequency weighting function $W(z^{-1})$ can be incorporated in Eq. (8). Since the excitation signal, ζ_k is artificially generated, this freedom can be exploited by choosing ζ_k to have a frequency spectrum similar or close to the frequency spectrum of the reference trajectory. Effectively, this is similar to using a frequency weighting function to give emphasis to the frequency spectrum of interest. The best way to capture the dominant frequency component of the trajectory signal is to use the trajectory signal itself as the excitation signal.

In order to ensure that the excitation signal will always persistently excited, a low-level noise signal which has its frequency spectrum close to the frequency spectrum of the reference trajectory can be superimposed on the reference trajectory. The addition of noise to the reference trajectory will not add disturbance to the system as the "noisy" reference trajectory will not pass through the system [6]. The design objective of the controller is to determine the optimal set of α_k such that the cost function given by Eq. (8) is minimized. Here, the optimal set of α_k are solved using a parameter estimation algorithm, Recursive Least Square (RLS) method. The optimum set of α_k was estimated by running the estimation algorithm online using the scheme shown in Figure 6.
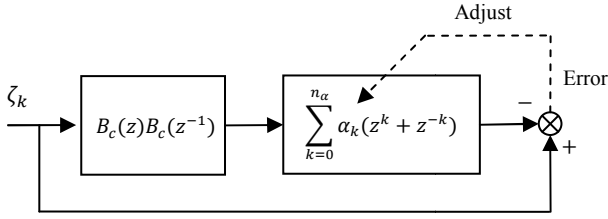
Figure 6. Computing optimal set of α_k using RLS parameter estimation method

V. FEEDBACK CONTROL

The feedback control in Figure 4 was designed using pole-placement method. This method enables all poles of the closed-loop system to be placed at desired location and providing good and stable output performance. The closed-loop transfer function of the feedback system is given by:

$$\frac{y_k(z^{-1})}{u_k(z^{-1})} = \frac{K_f B_o(z^{-1})}{A_o(z^{-1})F(z^{-1}) + B_o(z^{-1})G(z^{-1})} \quad (9)$$

$$K_f = \frac{\sum T}{\sum B_o}$$

$$A_o(z^{-1})F(z^{-1}) + B_o(z^{-1})G(z^{-1}) = 1 + t_1 z^{-1} \quad (10)$$

t_1 : Desired location of pole inside unity circle

The feedback control parameters were obtained by solving the Diophantine equation of Eq. (10) to solve for $F(z^{-1})$ and $G(z^{-1})$. The literature on this material can be obtained in [14].

In this paper, $t_1 = -0.2$ were used since the tracking required fast response. Using developed MATLAB m-file, the following parameters were computed:

A. X-Axis Plant Model:

$$T = 1 - 0.2z^{-1} \quad (11)$$

$$K_f = 3.8012$$

$$F(z^{-1}) = 1 + 0.0203z^{-1} - 0.4938z^{-2}$$

$$G(z^{-1}) = 11.5633 - 9.6061z^{-1} + 1.8433z^{-2}$$

B. Y-Axis Plant Model:

$$T = 1 - 0.2z^{-1} \quad (12)$$

$$K_f = 1.4863$$

$$F(z^{-1}) = 1 + 0.2289z^{-1} + 0.0817z^{-2}$$

$$G(z^{-1}) = 0.8060 + 1.1579z^{-1} - 0.4787z^{-2}$$

VI. RESULTS AND DISCUSSION

The simulation and real-time controls were done in MATLAB environment. The simulation results using 5th and 10th orders Trajectory-Adaptive ZPETC moving in lemniscates are shown in Figure 7 and Figure 8. We can observed that the resulted lemniscates movement for both controllers and reference movement could hardly be seen the differences. This is explained by the smooth lemniscates movement of the figures.
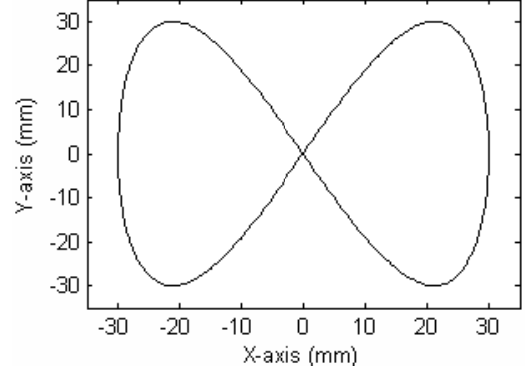
Figure 7. Simulation result of lemniscates movement for 5th order Trajectory-Adaptive ZPETC
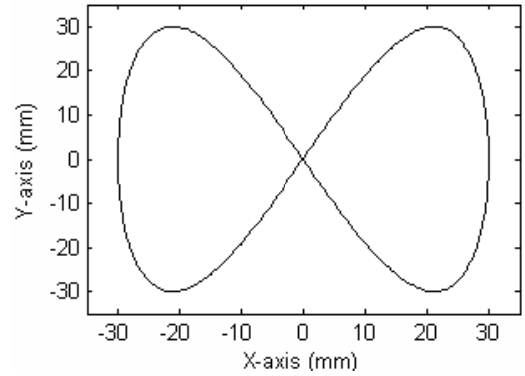
Figure 8. Simulation result of lemniscates movement for 10th order Trajectory-Adaptive ZPETC

Simulation results of individual axis tracking performance are given in Table I. We can observe that by increasing the controller order from 5th to 10th orders, the tracking error is getting much smaller. The 10th order Trajectory-Adaptive ZPETC shows better RMSE with 52.34% improvement for X-axis and 68.08% improvement for Y-axis.

TABLE I: Simulation results of individual axis tracking performance

	NMP XY-table system	
	X-axis RMSE (mm)	Y-axis RMSE (mm)
5th order Trajectory-Adaptive ZPETC	1.1970×10^{-5}	7.2191×10^{-5}
10th order Trajectory-Adaptive ZPETC	5.7051×10^{-6}	2.3043×10^{-5}

The real-time experiment results using 5th and 10th orders Trajectory-Adaptive ZPETC to move in lemniscates are shown in Figure 9 and Figure 10 respectively. We can observe that the resulted lemniscates movement for 10th order controller is better than 5th order controller. This is explained by the rough lemniscates movement of Figure 9 for 5th order Trajectory-Adaptive ZPETC. Fortunately, the 10th order Trajectory-Adaptive ZPETC produces better lemniscates movement as shown in Figure 10.
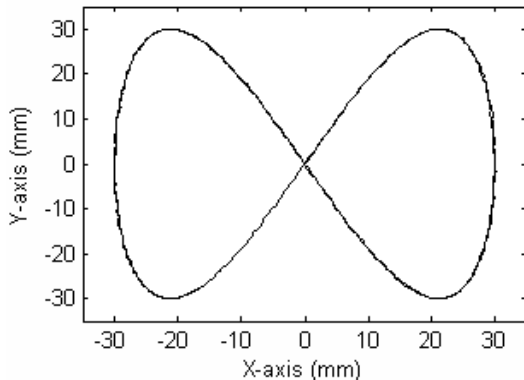
Figure 9. Real-time experiment result of lemniscates movement for 5th Trajectory-Adaptive ZPETC
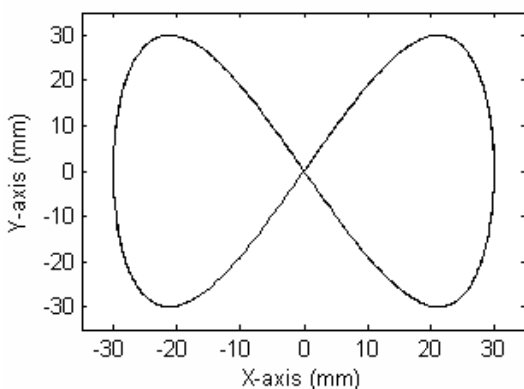
Figure 10. Real-time experiment result of lemniscates movement for 10th Trajectory-Adaptive ZPETC

Experimental results of individual axis tracking performance are given in Table II. We can observe that by

increasing the controller order from 5th to 10th orders, the tracking error is getting smaller. The 10th order Trajectory-Adaptive ZPETC shows better RMSE with 13.51% improvement for X-axis and 10.04% improvement for Y-axis.

TABLE II: Experimental results of individual axis tracking performance

	NMP XY-table system	
	X-axis RMSE (mm)	Y-axis RMSE (mm)
5th order Trajectory-Adaptive ZPETC	0.2014	0.3506
10th order Trajectory-Adaptive ZPETC	0.1742	0.3154

VII. CONCLUSION

The simulation and experimental results of real-time control of NMP XY Table using 5th order and 10th orders Trajectory-Adaptive ZPETC without factorization of zeros polynomials have been presented. In simulation studies, both controllers show good tracking performance by successfully reducing the NMP zeros problems. Hence, better lemniscates movement was achieved. However, in real-time studies, there were large differences between both results. This is due to the nature of ZPETC which is sensitive to plant-model mismatch caused by external disturbance such as parameter variations.

REFERENCES

[1] Michael, J.P., Norlela, I., Mohd Hezri, F., Mazidah, T. & Ramli, A., "Modeling and Controller Design for Non-Minimum Phase System with Application to XY-Table". IEEE.7th Control & System Graduate Research Colloquium (ICSGRC), pp. 113-118. 2011.
[2] Norlela, I., Hashimah, I., Mohd Hezri, F., Mazidah, T., Ramli, A. & Yahaya, M.S., "Tracking Control for Electro-Hydraulic Actuator using ZPETC". IEEE.7th Control & System Graduate Research Colloquium (ICSGRC), pp. 94-97. 2011.
[3] Hashimah, I., Norlela, I., Mohd Hezri, F.,Mazidah, T.&Ramli, A., "Positioning and Tracking Control of an XY Table". IEEE.7th International Conference on Control System, Computing and Engineering (ICCSCE), pp. 343-347. 2011.
[4] Ramli, A., Abd. Manan, S., Noritawati, M. T., Mohd Hezri, F. & Mohd Marzuki, M., "Trajectory Zero Phase Error Tracking Control Using Comparing Coefficients Method". IEEE.5th International Colloquium on Signal Processing & Its Application (CSPA), pp. 385-390, 2009.
[5] Ramli, A., Abd. Manan. & Mohd Marzuki, M., "Real-time Control of Non-minimum Phase Electro-Hydraulic System using Trajectory Adaptive ZPETC". IEEE.7th International Colloquium on Signal Processing & Its Application (CSPA), pp. 72-76, 2011.
[6] Mustafa, M. M., "Trajectory-Adaptive Digital Tracking Controllers for Non-minimum Phase Systems without Factorisation of Zeros".IEE Proc, Control Theory Appl., vol. 149, no. 2, pp. 157-162, 2002.

[7] Yeh, S.S. and Hsu, P.L., "An Optimal and Adaptive Design of the Feedforward Motion Controller", IEEE/ASME J. Mechatronics, vol. 4, no. 4, pp. 428–439, 1999.

[8] Torfs, D., Swevers, J. & De Schutter, J., "Quasi-Perfect Tracking Control of Non-Minimal Phase Systems". 30th Conference on Decision and Control, pp. 241-244, 1991.

[9] Tomizuka, M. "Zero Phase Error Tracking Algorithm for Digital Control", ASME J. of Dyn, Syst., Meas., and Control, vol. 109, pp. 65-68, 1987.

[10] Tomizuka, M. "On the Design of Digital Tracking Controllers", ASME J. of Dyn, Syst., Meas., and Control, vol. 115, pp. 412-418, 1993.

[11] Haack, B. & Tomizuka, M., "The Effect of Adding Zeros to Feedforward Controller". ASME J. of Dyn, Syst., Meas., and Control, vol. 113, pp. 6-10, 1991.

[12] Torfs, D., Swevers, J. & De Schutter, J., "Extended Bandwidth Zero Phase Error Tracking Control of Non-Minimum Phase System". ASME J. of Dyn, Syst., Meas., and Control, vol. 114, pp. 347-351, 1992.

[13] Xia, J.Z. & Menq, C.H. "Precision Tracking Control of Non-Minimum Phase System with Zero Phase Error", Int J. Control, vol. 61, no. 4, pp. 791-807, 1995

[14] Mohd Nasir, T., Ramli, A., and Mohd Hezri, F., Practical System Identification, Faculty of Electrical Engineering, UiTM, 2007. ISBN: 978-967-5070-03-7.

[15] Landau. I.D, System Identification and Control Design using P.I.M + software, Prentice Hall, Englewood Cliffs, N.J, 1990.

Flood Water Level Modelling and Prediction Using Artificial Neural Network : Case Study of Sungai Batu Pahat in Johor

Ramli Adnan [#1], Fazlina Ahmat Ruslan [#2], Abd Manan Samad [*3], Zainazlan Md Zain [#4]

[#]Faculty of Electrical Engineering
[*]Center for Surveying Science and Geomatics, Faculty of Arc., Planning and Surveying
Universiti Teknologi MARA 40450, Shah Alam, Selangor, Malaysia
[1]ramli324@salam.uitm.edu.my
[2]fazlina419@salam.uitm.edu.my

Abstract— **Flood water level prediction has long been the earliest forecasting problems that have attracted the interest of many researchers. Accurate prediction of flood water level is extremely importance as an early warning system to the public to inform them about the possible incoming flood disaster. Using the collected data at the upstream and downstream station of a river, this paper proposes a modelling of flood water level at downstream station using back propagation neural network (BPN). In order to improve the prediction values, an extended Kalman filter was introduced at the output of the BPN. The introduction of extended Kalman filter at the output of BPN shows significant improvement to the prediction and tracking performance of the actual flood water level.**

Keywords ⸺ *Back Propagation Neural Network, Flood Water Level Forecasting, Mean Squared Error, Correlation Coefficient, Error Goal*

I. INTRODUCTION

Flood water level prediction has long been the earliest forecasting problems that have attracted the interest of many researchers [1]. Accurate prediction of flood water level is extremely important as an early warning system to the public to inform them about the possible incoming flood disaster. Flood water levels at downstream areas are strongly affected by upstream conditions [2]. The artificial neural network (ANN) is capable of articulate the underlying relationship between downstream and upstream conditions because of its ability to represent a complex nonlinear system without any priori knowledge of the physical systems itself [3].

McCulloch and Pitts [4] introduced ANN as a representative of the synaptic processes in brain. ANN became popular in the early 1980s with the advent of microprocessors and by late 1990s ANN starts to be used in hydrological modelling. Recently, ANN has been widely applied in various areas to overcome the problem of nonlinear relationship in geographical system. For example, ANN is able to predict all India summer monsoon rainfall with varieties of model inputs [5]. Tsung-Lin Lee [6] successfully forecast the storm surge and surge deviation in coastal areas using back propagation neural network approach. Lee [7] applied back propagation neural network with typhoon characteristics and tide level as the input factors to develop a typhoon surge forecasting model. Lin and Wu [8] proposed rainfall prediction using hybrid ANN model where multilayer networks are employed to map the relationship between input and output.

This paper proposes a modelling of flood water level at downstream station using back propagation neural network (BPN). In order to improve the prediction values, an extended Kalman filter was introduced at the output of the BPN.

This paper was organized in the following manner: Section II describes the methodology; Section III is on results and discussion; and finally, Section IV is on conclusions.

II. METHODOLOGY

A. Study Area and Data

The case study was done to a Sungai Batu Pahat that located at Batu Pahat, Johor. A major contribution of water level at Sungai Batu Pahat basin comes from four upstream rivers as depicted in Figure 1. The upstream river level measuring station names are Station 1929401(Sg Simpang Kiri at Sri Medan), Station 2130422 (Sg Bekok Bt 77 Jln Yong Peng), Station 1831480 (Sg Sembrong at Parit Karjo) and Station 2030402 (Sg Bekok at Yong Peng). The water level measuring station for Sungai Batu Pahat is Station 1829478 (Sg Batu Pahat at Batu Pahat). The water level data is in meters, starting from May 2, 2012 until May 8, 2012. This real-time data is available online from website www.water.gov.my. The water level data is measured using Supervisory Control and Data Acquisition System (SCADA) by Department of Irrigation and Drainage Malaysia.

Figure 1. The location of Sungai Batu Pahat basin and its upstream river (http://telemetryjohor.com/Pages/StationStatus.aspx)

B. Artificial Neural Network

ANN is a mathematical system that emulates the ability of connecting simple neurons and represents it in a biological neural network. In general, ANN contains input layers, hidden layers and output layers. This is illustrated in Figure 2. Input layers receive input signals and then passed to each hidden layer. Hidden layer then represent relationship between input and output layer and finally output layer releases the output signals.
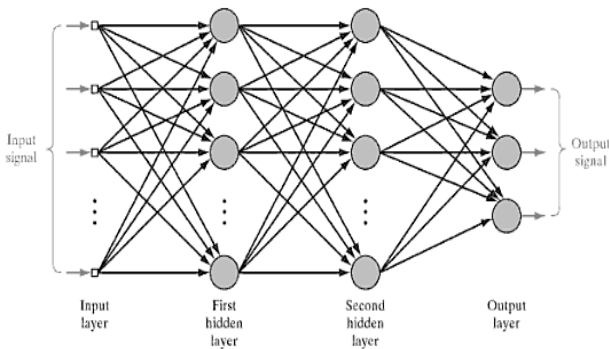
Figure 2. Structure of artificial neural network

The back propagation neural network (BPN) is the most popular supervised model of artificial neural networks which is developed by Rumelhart et al. [9]. BPN repeatedly adjust the parameters (weights and biases) to minimize the error between target output and estimated output according to "generalized delta rule". The parameter modification stops when the performance error goal is met [10].

The crucial factor that leads to the success of ANN model is the selection of appropriate inputs. Too many inputs that are redundant will contribute to the problem of local minima [11]. By understanding the physical system to be model, the good choice of input can be guided. For example, upstream river

levels are expected to be a good downstream river levels predictor in the same river system [12]. Singh et al. [13] reported that the data sets range from 0.1-0.9 or 0.2-0.8 can improve generalization of ANN architecture.

Figure 3 illustrates the basic structure of BPN water level prediction model of this study. The input variable of this model is water level at the upstream river whereas the output variable is the peak water level at the downstream river, Sungai Batu Pahat. The number of neuron in hidden layer is only representative and not the actual number in the model.
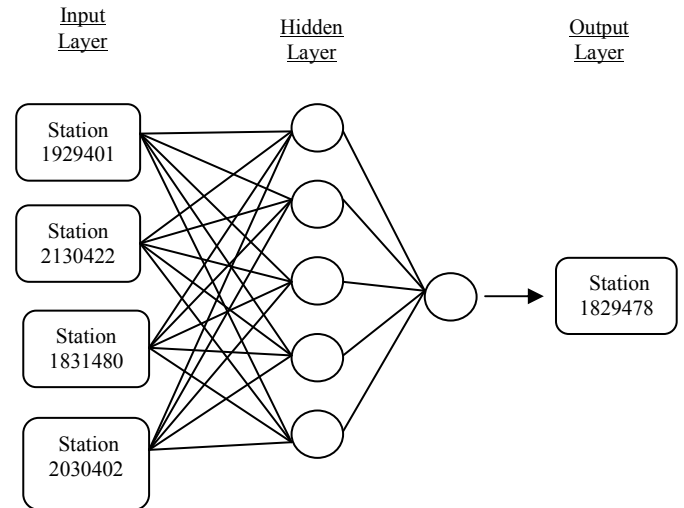
Figure 3. BPN model for flood water level modelling

C. Extended Kalman Filter (EKF)

The EKF algorithm has two main stages: prediction step; and update (filtering) step. In prediction steps, previous state estimate of the system is being used and in the update step, the predicted state is corrected by using feedback correction step. The feedback correction step contains the weight of the measured and estimated output signals. In EKF, by calculating the stochastic properties of the noise, the initial values of the matrices can be arranged correctly. Furthermore, a critical part of EKF is to apply correct initial values for various covariance matrices. Covariance matrices have an important effects on converge time and filter stability. The summary of EKF algorithm is given in Figure 4.

Time Update ("Predict")

(1) Project the state ahead

$$\hat{x}^-_k = f(\hat{x}_{k-1}, u_{k-1}, 0)$$

(2) Project the error covariance ahead

$$P^-_k = A_k P_{k-1} A^T_k + W_k Q_{k-1} W^T_k$$

Measurement Update ("Correct")

(1) Compute the Kalman Gain

$$K_k = P^-_k H^T_k \left(H_k P^-_k H^T_k + V_k R_k V^T_k \right)^{-1}$$

(2) Update estimate with measurement z_k

$$\hat{x}_k = \hat{x}^-_k + K_k \left(z_k - h(\hat{x}^-_k, 0) \right)$$

(3) Update the error covariance

$$P_k = (I - K_k H_k) P^-_k$$
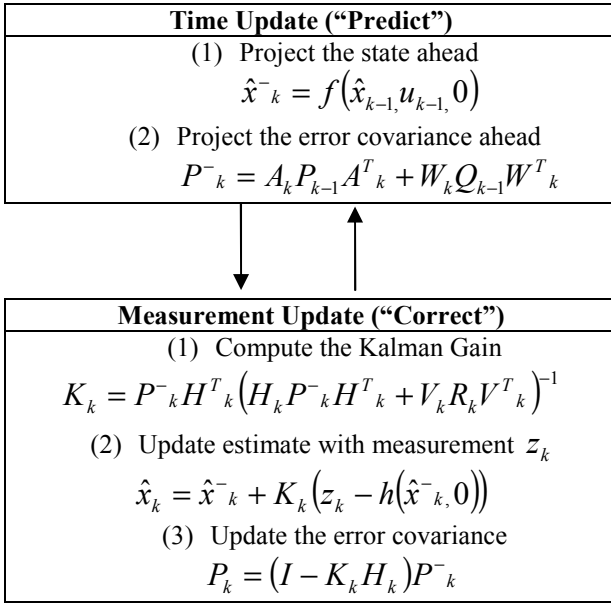
Figure 4. Extended Kalman Filter (EKF) Loop

Figure 4 summarizes the overall operation of EKF algorithm. The prediction stage starts by projecting the state and error covariance ahead using the present value of Jacobian matrix and prior value of error covariance. Then, based on the projected error covariance value, the Kalman gain is calculated and measurement $\hat{x}_k$ is estimated. Finally, the error covariance value is updated. The loop will continue until the desired error covariance value is achieved.

III. RESULTS AND DISCUSSION

The water levels at the upstream and downstream rivers station at Sungai Batu Pahat is applied in the BPN model. Station 1929401, 2130422, 1821480 and 2030402 acts as a feeder to the BPN model because it is located at upstream location whereas Station 1829478 is the output of the model because it is the target station. The nodes of both input and output layers are set as 1.The number of nodes for hidden layer were set to be 15, in following the law of Kolmogorov [14]. Kolmogorov's law state that the number of nodes in the hidden layer is at least 2n+1, where n is the number of nodes in the input layer. Thus, the topology structure of BPN flood modelling and forecasting in this study is (1,15,1).

In order to accelerate the convergence, it is necessary to normalize the original data, but need to remember to denormalize it back to obtain the desired data. After the normalization process, the training parameters used were learning rate η at 0.55 and momentum constant α is fixed at 1. Learning rate and momentum constant is used to accelerate convergence of ANN training process and the value of η and α is selected in the range of 0 and 1.

According to 10, 0000 cycles of training, the error goal is set at 1.00e-05. The training algorithm used is traincgp, conjugate gradient with Polak Ribiere updates. The verification results are given in Figure 5. The BPN verification result shows that the model cannot be considered good because there is slight difference in the simulated and actual water level results. Furthermore, overshoot occurs at the end of the graph. This is maybe because that the training is stop after the minimum step size is reached. However the value of mean squared error (MSE) is 8.89e-05 which is still close to the error goal that is 1.00e-05, so the result is still acceptable.
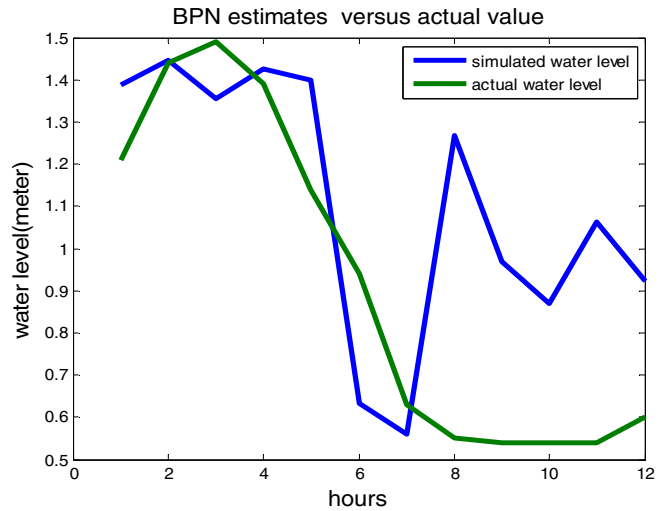
Figure 5. The verification results of the flood water level

In order to improve the prediction results, an Extended Kalman Filter (EKF) is introduced at the output of BPN. EKF is the most commonly used nonlinear Bayesian estimator which is able to filter and track the nonlinear output based on assumptions that are: perturbations from the mean trajectory are small; and the conditional density function of the state is Gaussian [15]. Figure 6 shows the results after the EKF being applied. From the graph, it is clearly shown that the results are improved as the overshoot is reduced and the pattern of the simulated and actual water level between 0— 6 hours is nearly the same as compared to Figure 5.

BPN estimates(with EKF)versus actual value
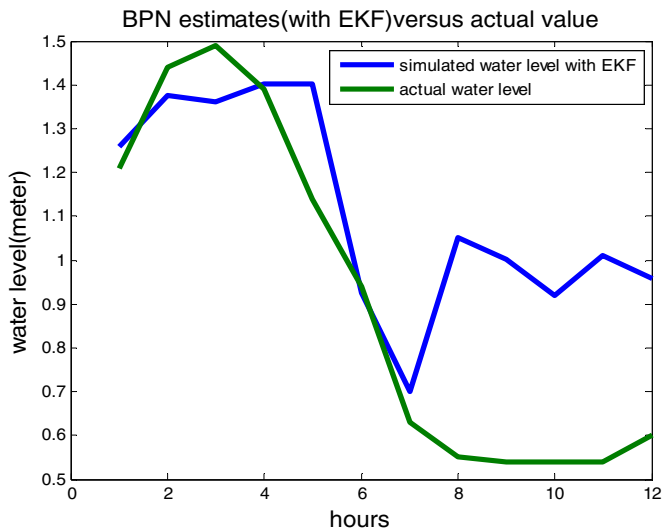
Figure 6.   The verification results of the flood water level with EKF

IV.   CONCLUSIONS

Accurate prediction of flood water level is important as an early warning system to the public to inform them about the possible incoming flood disaster. This paper proposed ANN model to predict the flood water level. A supervised neural network with the back propagation model is adopted in this study. The performances of BPN model have been evaluated but yet the result is not that satisfactory. Hence, Extended Kalman Filter is proposed to improve the results. In conclusion, Extended Kalman Filter is able to provide better result compared with actual result without application of Extended Kalman Filter. The water level is only the parameter in this Back Propagation Neural Network model. So, for future works, more parameters can be added in the model such as rainfall and water flow so that more accurate and realistic prediction results can be achieved.

ACKNOWLEDGEMENT

REFERENCES

[1]   Amir F. Atita et al, "A comparison between neural network forecasting techniques – case study: river flow forecasting," IEEE Transactions on Neural Networks, vol. 10, no.2, pp.402-409, 1999.

[2]   Chang-Shian Chen et al, "Development and application of a decision group Back-Propagation Neural Network for flood forecasting," ELSEVIER Journal of Hydrology, vol. 385, pp. 173-182, 2010.

[3]   Paul Leahy, Ger Kiely & Gearoid Corcoran, "Structural optimisation and input selection of an artificial neural network for river level prediction," ELSEVIER Journal of Hydrology, vol. 355, pp. 192-201, 2008.

[4]   McCulloch, W., Ps, W.H., "A logical calculus of the ideas immanent in nervous activity," Bull. Math. Biophys. 5, pp. 115-133, 1943.

[5]   Venkatesan et al, "Prediction of all India summer monsoon rainfall using error back propagation neural networks," Meteorology and Atmospheric Physics, vol. 62(3-4), pp.225-240, 1997.

[6]   Tsung-Lin Lee, "Predictions of tyhphoon storm surge in Taiwan using artificial neural networks," ELSEVIER Advances in Engineering Software, vol. 40, pp.1200-1206, 2009.

[7]   Lee, T.L, "Neural network prediction of a storm surge," Ocean Engineering, vol.33, pp.483-494, 2006.

[8]   Lin, G.F., Wu. M.C., "A hybrid neural network model for typhoon rainfall forecasting ," Journal of Hydrology, vol. 375(3-4), pp.450-458, 2009.

[9]   Rumelhart et al., "Learning representations by back propagating errors," Nature, 323:533-6, 1986.

[10]   Hsiao, T.R., Lin, C.W., Chiang,H.K., "Partial least-squares algorithm for weights initialization of back propagation network," Neurocomputing, vol. 50, pp. 237-247, 2003.

[11]   Govindaraju, R., "Artificial neural networks in hydrology : preliminary concepts," Journal of Hydrology, vol. 5(2), pp. 115-123, 2000.

[12]   Govindaraju, R., "Artificial neural networks in hydrology : hydrologic applications," Journal of Hydrology, vol. 5(2), pp. 124-137, 2000.

[13]   Singh et al, "Appropriate data normalization range for daily river flow forecasting using an artificial neural networks," Hydroinformatics in Hydrology, Hydrogeology and Water Resources. International Association of Hydrological Sciences Publication, Uk, pp. 51-57, 2009.

[14]   Wabgaonkar,H., Stubberud,A., "The Kolmogorov superposition theorem and functional approximation in neural networks," 24th Asilomar Conference on Signals, Systems and Computers, vol. 2, 1990.

[15]   Jangho Yoon, Yunjun Xu & Prakash Vedula, "A direct quadrature based nonlinear filtering with Extended Kalman Filter update for orbit determination," 2010 American Control Conference Marriot Waterfront, Baltimore, MD, USA, June 30-July 02, 2010.

Implementation of ODFC for angular position control of cart-pendulum dynamic model

Rajeeb Dey, Abhijit Rajkumar, Aditi Bandyopadhyay

Dept. of Electronics & Communication Engineering, Sikkim Manipal University, Sikkim, India
E-mail: rajeeb_de@ieee.org, Tele: +91-9475078606.

Absract— **This paper investigates the implementation of output delayed feedback control (ODFC) technique for controlling the pendulum sway angle of a cart-pendulum control system. Linearized mathematical model of the cart control in state space form is considered for the investigation. The designed ODFC has undergone complete stability analysis with two different approaches for a given controller gain.**

Keywords—Output delayed feedback control (ODFC), Cart-Pendulum dynamics, Stability of time-delay system, Time-delay control (TDC).

I. Introduction

A cart control problem for a cart-pendulum mechanism consists of crane carrying the cart with a movable pendulum hinged at the cart. Such mechanism is a motivation for several modern industrial environments where transportation of heavy payloads from one position to another as fast and as possible and accurately without collision with other equipments is a requirement as per factory safety norms. Thus solutions of vibration control of mechanical systems and structures have been attempted in the past as evident from the literature [1, 2, 3, 6, 8, 12, 14]. In cart-pendulum dynamical system under consideration [17], the basic motion involves crane traversing, and oscillation of the pendulum hinged with it. In case, the crane is traversing at a high velocity, a large sway of the pendulum will take place. The objective of this investigation is to design a controller for controlling sway angle of the pendulum with two different approaches and their comparison.

The controller is designed using output delay feedback mechanism. The inclusion of time-delay in the feedback makes the system dynamics infinite dimensional, thus direct computation of the characteristic roots and consequently deciding about stability is a difficult task.

A detailed review of the research on time-delay stability and stabilization issues using both frequency domain technique and time domain methods can be found in [7, 9, 11,16]. The former technique provides exact stability analysis for time-invariant delay and hence TDC has been used in many control applications [3, 10, 13] to suppress vibration or oscillations of the system. The later technique can treat the natures of delay, time-variant and time-invariant case [11, 14, 15, 16] and references therein. A numerically tractable algorithm exists to solve the problem, but provides conservative analysis compared to the former technique. In this paper, to control this lightly damped system, a signal is generated from the position sensor which

is then combined with the delayed signal from the same sensor and fed back to the system, thus calling it output delayed feedback controller (ODFC). This design involves prior knowledge of the controller gain for which the time-delay is treated as design parameter.

II. Dynamic Model of Cart Control System

The two-dimensional cart-pendulum dynamical system with its pendulum considered in this work is shown in Fig.1 and Fig.2. $F(t)$ represents the force causing translational motion of the cart. The nomenclatures along with the values of the physical parameter are given in Table 3 and can be found in [17].
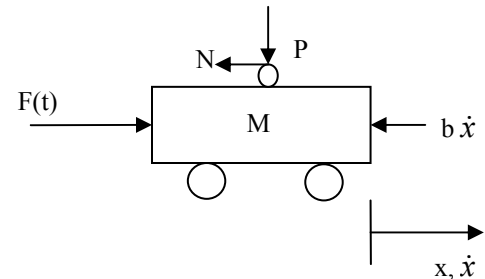
Fig. 1: Free Body Diagram of cart
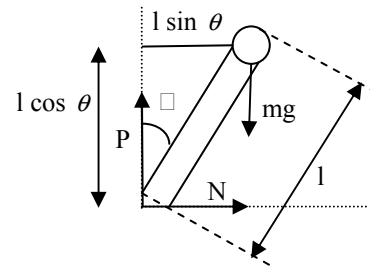
Fig. 2 Free Body Diagram of Pendulum

Following simplification apply in this model, (i) Model does not include hoisting drive for pendulum, thus pendulum length is fixed (ii) cart is assumed to be point mass (iii) pendulum assumed to move in X-Y plane and (iv) force on cart due to pendulum swing is neglected.

The non-linear dynamic model of the cart control [17] using formulation with above simplifications yields

$$(M+m)\ddot{x}+b\dot{x}+ml\ddot{\theta}\cos\theta-ml\dot{\theta}^2\sin\theta=F_x \qquad (1)$$

$$(I+ml^2)\ddot{\theta}+d\dot{\theta}+ml\ddot{x}\cos\theta-mgl\sin\theta=0 \qquad (2)$$

There are two equilibrium points in this system, i.e., $\theta=0$ and $\theta=\pi$ for cart control and inverted pendulum respectively. Assuming small sway angle θ about the equilibrium point $\theta_0=0$, we can write

$$\varepsilon=\theta-\theta_0 \text{ and } \dot{\theta}=\dot{\varepsilon} \qquad (3)$$

Using Taylor's series expansion, first order approximation of any function of θ can be written as:

$$f(\theta)=f(\theta)|_{\theta_0}+\varepsilon\frac{df}{d\theta}|_{\theta_0}+\text{higher order derivatives.} \quad (4)$$

As, all higherorder derivatives are neglected, so $\dot{\varepsilon}^2\approx 0$.
Considering $\theta_0=\pi$,

$$\cos\theta\approx -1 \text{ and } \sin\theta\approx-\theta \qquad (5)$$

Thus, an approximate linear dynamic model for (1) and (2) considering $\theta_0=\pi$ can be represented as

$$(M+m)\ddot{x}+b\dot{x}-ml\ddot{\theta}=F_x \qquad (6)$$

$$(I+ml^2)\ddot{\theta}+d\dot{\theta}-ml\ddot{x}+mgl\theta=0 \qquad (7)$$

Above equations can be written in state space form as

$$\begin{aligned}\dot{X}(t)&=AX(t)+Bu(t)\\ y(t)&=Cx(t)\end{aligned} \qquad (8)$$

where $X(t)\in\Re^{4\times 1}$ is the state vector, $u(t)\in\Re$ is the control input, $y(t)\in\Re$ is the output of the system, further the considered states are $X(t)=[x\ \theta\ \dot{x}\ \dot{\theta}]^T$ and $y(t)=\theta(t)$, the constant system as well as input matrices are given by A and B respectively as,

$$A=\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -\dfrac{m^2l^2g}{den} & -\dfrac{b(I+ml^2)}{den} & -\dfrac{mld}{den} \\ 0 & -\dfrac{(M+m)mlg}{den} & -\dfrac{mlb}{den} & -\dfrac{(M+m)d}{den} \end{bmatrix}$$

$$B=\begin{bmatrix} 0 \\ 0 \\ \dfrac{(I+ml^2)}{den} \\ \dfrac{ml}{den} \end{bmatrix}$$

Where, den= $((I+ml^2)(M+m)- m^2l^2)$

In order to use delayed feedback control method, we chose to feedback the sway angle θ of the pndulum as, $C=\begin{bmatrix}0 & 1 & 0 & 0\end{bmatrix}$.

III. DESIGN OF OUTPUT DELAYED FEEDBACK CONTROLLER

A. *Approach I: [4, 5, 13]*

In this section, we explain the controller design following the technique in [1,4,5,13]. The control law adopted for this controller is shown in Fig. 3 and mathematically given by

$$u(t)=K[y(t)-y(t-\tau)] \qquad (9)$$

Using (9) the closed loop system dynamics of (8) can be written as

$$\dot{X}(t)=A_0 X(t)+A_1 X(t-\tau) \qquad (10)$$

where, $A_0=A+BKC$ and $A_1=-BKC$.

The characteristic equation of (10) is a transcedental equation and one can write it as

$$\left|sI-A_0-A_1e^{-s\tau}\right|=0 \qquad (11)$$

Infinite number of characteristic roots exists for (11) due to presence of the delay term in (10). To carry out the stability analysis of such time-delay system several appraoches have evolved in past. Frequency domain techniques gives exact stability analysis and involves finding roots of (11) whereas time-doamin technique do not involve actually computing roots of (11) and hence provides conservative stability results. We adopt the exact stability analysis technique for the design of ODFC.

The implementation of the technique for cart-pendulum control system under delayed output feedback is described briefly below:

The characteristic equation in (11) is written in the genereic form as

$$\Delta(s,\tau)=\sum_{l=0}^{n} p_l(s)e^{-l\tau s} \qquad (12)$$

Finding the complete root crossing structure for (12) using the Rekasius substitution for $e^{-\tau s}=\dfrac{1-Ts}{1+Ts}$, where T is the psuedo delay parameter term to convert it into resulting polynomial without transcedentality, which takes the form

$$\sum_{l=0}^{2n} b_l s^l=0 \qquad (13)$$

where, $b_l=b_l(T,a_{ij},b_{ij})$, $a_{ij},b_{ij},1\le i,j\le n$ being the elements of A and B matrices.

Next, we apply Rouths-Hurwitz criterion on (13), we determine set of T values i.e, $\{T_c\}$, by equating the element of s^1 row of Routh's array to zero, and for each T_{cl} value when the auxilliary equation (which is formed by the row preceeding s^1 in the Routh's array) is solved it gives either a pair of imaginary roots $(\mp\omega_{cl}i)$ or real and equal roots with opposite sign. Those T_{cl} values that give real roots are overlooked as we are interested only in the imaginary

crossing frequencies. The computed values of T_{cl} and it's corresponding ω_{cl} are placed in Table 1. For each T_{cl} (or corresponding ω_{cl}) there are infinitely many time delays $\{\tau_l\}$, which one can obtain using

$$\tau = \frac{2}{\omega}[\tan^{-1}(\omega T) \mp k\pi], \ k = 0, 1, 2.... \quad (14)$$

For the system in (8) one can get finite number of purely imaginary roots $\{\omega_{cl}\}, l = 1, 2, 3........m$ but infinitely many time delays i.e, $\{\tau_{lk}\}$ with $k = 0, 1, 2......\infty$. The computed delay values are placed in Table 1 for $K = 3$.

To find the stable delay intervals one need to identify the tendency of the roots at the calculated frequency and critical delay values, this is accomplished by computing the root tendency as described in [4], which is briefly outlined here pointwise,

(i) The characteristic roots of system in (8) crosses the imaginary axis at $\{\omega_{cl}\}$ values for infinitely many time-delays $\{\tau_{lk}\}$, that is computed as described above. The stability regions or switches are found by computing the root tendencies (RT) (or root crossing directions) at the point of corresponding delays τ_{lk} by the following equation from

$$\left. RT \right|_{\substack{\omega=\omega_{cl} \\ \tau=\tau_{lk}}} = \text{sgn}\left[\text{Im}\left(\frac{\sum_{j=0}^{n} a'_j e^{-j\tau s}}{\sum_{j=0}^{n} j a_j e^{-j\tau s}} \right)_{\substack{s=\omega_{cl} i \\ \tau=\tau_{lk}}} \right] \quad (15)$$

(ii) The values of RT will be +1 or -1. NU increases by 2 if it comes out to be +1, if it is -1 it decreases by 2. The computation of NU can be done using (20) in [4].

(iii) Now for finding the delay range (or stability switches) for the considered system, we rearrange the computed τ_{lk} values in an ascending order and check for NU by looking at the RT's corresponding to delay values. The complete information about the stability switches for $K = 3$ are placed in Table 2.

Table 1: COMPUTED VALUES OF ω_c AND T_c FOR $K = 3$

ω_{c1} =2.2865, T_{c1} =9.315		ω_{c2} =2.608, T_{c2} =0.0242	
τ_{k0}	0.0483	τ_{k0}	1.332
τ_{k1}	2.4575	τ_{k1}	4.08
τ_{k2}	4.8667	τ_{k2}	6.828
:	:	:	:
$\tau_{k\infty}$	$\tau_{1\infty}$	$\tau_{k\infty}$	$\tau_{2\infty}$

Fig. 3: Block Diagram of ODFC system

B. Approach II: [3]

In this approach, using the technique used in [2,3], the critical values of the time delay τ that result in characteristic roots of crossing the imaginary axes are obtained. This approach suggests that (8) can be written in the form

$$\Delta(s,t) = P(s) + Q(s)e^{-st} \quad (16)$$

$P(s)$ and $Q(s)$ are polynomials in s with real coefficients and deg $(P(s) = n > \deg(Q(s))$ where n is the order of the system. In order to find the critical time delay τ that leads to marginal stability, the characteristic equation is evaluated at $s = j\omega$. Separating the polynomials $P(s)$ and $Q(s)$ into real and imaginary parts and replacing $e^{-j\omega t} = \cos(\omega t) - j\sin(\omega t)$, (11) can be written as

$$\Delta(j\omega, \tau) = (P_R(\omega) + jP_I(\omega)) + (Q_R(\omega) + jQ_I(\omega)) \times Z \quad (17)$$
where,
$$Z = \cos(\omega\tau) - j\sin(\omega\tau)$$

The characteristic equation $\Delta(s,t) = 0$ has roots on the imaginary axis for some values of $\tau \geq 0$ if (17) has positive real roots. A solution of $\Delta(j\omega, \tau) = 0$ exists if the magnitude $|\Delta(j\omega, \tau)| = 0$. Taking the square of the magnitude of $\Delta(j\omega, \tau)$ and setting it to zero lead to the following equation

$$P_R^2 + P_I^2 - (Q_R^2 + Q_I^2) = 0 \quad (18)$$
By setting the real and imaginary parts of (18) to zero, the equation is rearranged as below

$$\begin{bmatrix} Q_R & Q_I \\ Q_I & -Q_R \end{bmatrix} \begin{bmatrix} \cos\beta \\ \sin\beta \end{bmatrix} = \begin{bmatrix} -P_R \\ -P_I \end{bmatrix} \quad (19)$$

where $\beta = \omega\tau$.

Solving for $\sin\beta$ and $\cos\beta$ gives

$$\sin(\beta) = \frac{(-P_R Q_I + P_I Q_R)}{(Q_R^2 + Q_I^2)} \qquad (20)$$

$$\cos(\beta) = \frac{(-P_R Q_R - P_I Q_I)}{(Q_R^2 + Q_I^2)} \qquad (21)$$

In order to determine the critical values of time delay, if a positive root of (18) exists, the corresponding time delay τ can be found by

$$\tau_k = \frac{\beta}{\omega} + \frac{2k\pi}{\omega} \qquad (22)$$

where, $\beta \in [0 \; 2\pi]$. At these time delays, the root loci of the closed-loop system are crossing the imaginary axis of the plane. This crossing can be from stable to unstable or from unstable to stable and can be determined from the root tendency which is given by:

$$S = \operatorname{sgn} F'(\omega) \qquad (23)$$

where

$$F'(\omega) = 2(P_R P_R' + P_I P_I' - Q_R Q_R' - Q_I Q_I') \qquad (24)$$

For the given system,

$$P(s) = s^3 + 0.0572\,s^2 + (6.808 - 0.2638\,k)s$$
$$+ (0.129 - 2.7 \times 10^{-6}\,k) \qquad (25)$$

$$Q(s) = (0.2638\,ks + 2.7 \times 10^{-6}\,k) \qquad (26)$$

Substituting, $s = j\omega$,

$$P_R = -2.7 \times 10^{-6}\,k - 0.0572\omega^2 + 0.1296 \qquad (27)$$

$$P_I = 6.808\omega - \omega^3 - 0.2638k\omega \qquad (28)$$

$$Q_R = 2.7 \times 10^{-6}\,k \qquad (29)$$

$$Q_I = 0.2638k\omega \qquad (30)$$

Where P_R and Q_R are the real counterparts and P_I and Q_I are the imaginary counterparts of P and Q respectively.

The critical time delays for $K = 3$ obtained from the positive real roots are computed from (18) and listed in Table 2. In order to investigate the performance of the closed-loop system under the control action in (9), and in order to find the critical values of the time delay τ, the roots of (11) are obtained for different values of the control gain K. From the expressions of the polynomials Q_R, Q_I, P_I and P_I, (18) will result in three roots of which only the real ones are of interest. The critical values of τ are calculated and tabulated in Table 2 for K=3. Each of the time delays in Table 2 can either add a pair of characteristic roots with positive real parts or it can remove them depending on whether the root loci cross the imaginary axis from left to right or from right to left, which in turn, can be determined mathematically by using (23).By ordering the values of the time delays in an ascending manner, the

critical time delay after which the system remains unstable can be determined. It is seen that the critical time delays obtained by approach I are the same as those obtained by approach II.

IV. SIMULATION RESULTS FOR APPROACH I

The closed loop simulation for the system in (11) is carried using MATLAB for several values of time-delays and for a given gain. For $K = 3$, the system is unstable at $0.0483 < \tau < 1.332$, it is marginally stable at 1.332 and stable in the range $1.332 < \tau < 2.4575$. The simulation result at $\tau = 1 \in (0.048, 1.332)$, $\tau = 1.332$ and $\tau = 2 \in (1.33, 2.4575)$ are given in Fig. 5, 6 and 7 respectively. It is seen that the simulation results complies with the computed critical time-delay values obtained using the analysis.

Remark 1: There is no crossing frequency corresponding to T_{c3} = -52.598 for $K=3$, as it yields real roots and is hence neglected.

Remark 2: We observe from Table 2 the system is stable in the range $0 < \tau < 0.0483$, $1.332 < \tau < 2.4575$, $4.08 < \tau < 4.866$, $6.828 < \tau < 7.275$, $9.57 < \tau < 9.685$. After $\tau = 9.685$, the system becomes unstable and the stability is never regained.

Remark 3: It is also observed that the value of the critical time delay after which the system never regains stability decreases along with the number of stability pockets as the value of gain is increased. This tendency can be easily seen in Fig. 4.

Table 2: STABILITY POCKETS FOR $K = 3$

Critical Time Delay τ_{ck} (sec)	Imaginary Roots $(s = \pm j\omega_{ck})$, ω_{ck}	Root Tendency RT	Number of Unstable Roots (NU)
			NU=0
0.0483	2.608	+1	
			NU=2
1.332	2.286	-1	
			NU=0
2.4575	2.608	+1	
			NU=2
4.08	2.286	-1	
			NU=0
4.866	2.608	+1	
			NU=2
6.828	2.286	-1	
			NU=0
7.275	2.608	+1	
			NU=2
9.576	2.286	-1	
			NU=0
9.685	2.608	+1	
			NU=2
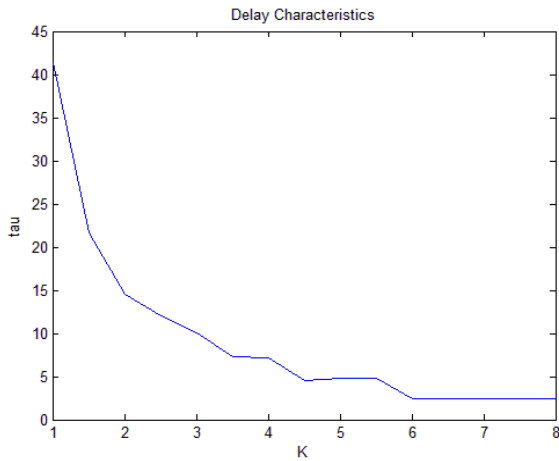
Fig. 4: Delay characteristics
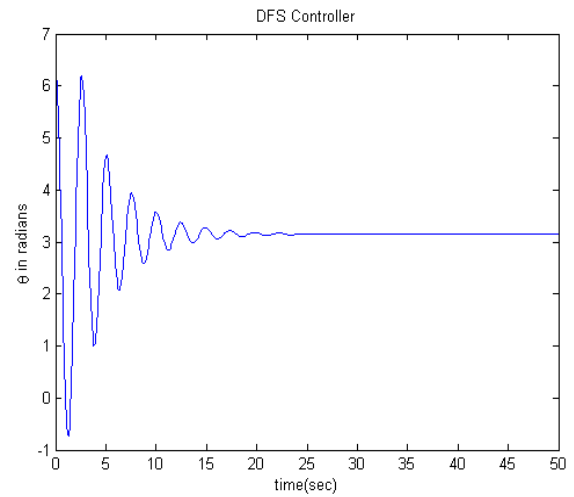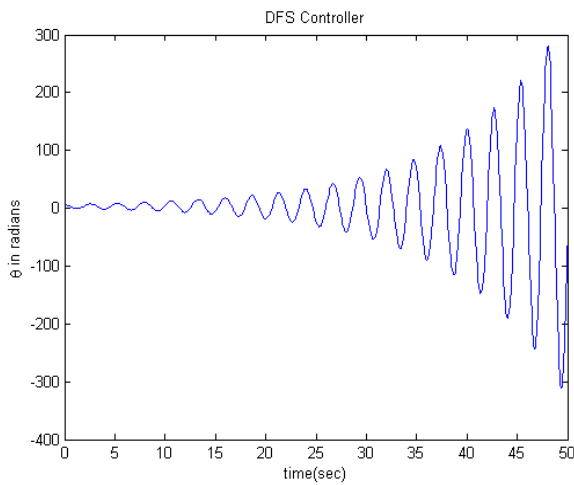
Fig. 5: Unstable system at τ=1 sec, K=3
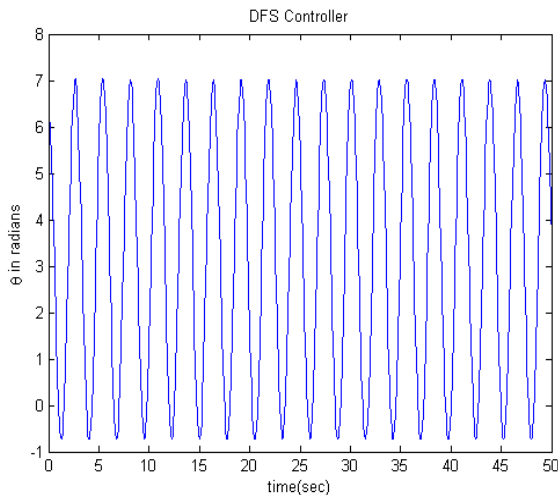
Fig. 6: Critically stable system at τ=1.332 sec, K=3

Fig. 7: Stable system at τ=2 sec, K=3

Table 3: SYSTEM PARAMETERS [17]

Mass of the cart	M= 2.4 Kg
Mass of payload	m = .23 Kg
Acceleration due to gravity	g = 9.81 m/s^2
Cart friction coefficient	b = 0.05 Ns/m
Pendulum damping coefficient	d = 0.005 Nms/rad
Moment of inertia of the pole	I = 0.099 Kgm^2
Pendulum length from the point of suspension to CG	l =0.4 m

V. CONCLUSION

In this work, implementation of the ODFC design for controlling the angular position of the linearized cart-pendulum dynamics using the exact stability analysis in [4, 5] is presented. The stability regions determined using the analysis for a given gain matches exactly with the simulation results as shown in Fig. 3 to Fig. 6. The computed value of the delay ranges is found to match using approach I and II. But the applicability of the approach II is limited by the complications faced for commensurate delays.

The future work is directed towards the implementation of the designed controller to the Digital Pendulum Control System 33-200 [17] shown in Fig. 8.

REFERENCES

[1] Rajeeb Dey, Nishant Sinha, Priyanka Chaubey, "Active sway control of single pendulum gantry crane system using output delayed feedback control technique," *in Proc. 11th Int. Conf. Control, Automation, Robotics and Vision*, Singapore, 2010.

[2] M.A. Ahmad, R.M.T. Raja Ismli, A.N.K. Nasir and M.S. Ramli, "Anti sway control of gantry crane system based on feedback loop approaches",

in *Proc. IEEE/ASME International Conference on Advanced Intelligent Mechatronics,* Singapore, 2009.

[3] Amor Jnifene, "Active vibration control of flexible structures using delayed position feedback," *System & Control Letters,* vol. 56, pp. 215 – 222, 2007.

[4] Nejat Olgac and Rifat Sipahi, "An Exact Method for the Stability Analysis of Time-Delayed Linear Time-Invariant (LTI) Systems," *IEEE Transaction on Automatic Control*, vol. 47(5), pp. 793-797, MAY 2002.

[5] Nejat Olgac and Rifat Sipahi, "An Improved Procedure in Detecting the Stability Robustness of Systems with Uncertain Delay," *IEEE Transaction on Automatic Control*, vol. 51(7), pp. 1164-1165, JULY 2006.

[6] D.Liu,J.Yi,D.Zhao,W.Wang, "Swing free transporting of two dimensional overhead crane using sliding mode Fuzzy control," *in Proc. American Control Conf.*, Boston, 2004, pp. 1764-1768.

[7] J.P Richard, "Time-delay systems: An overview of some recent problems", *Automatica*, vol. 39, pp. 1667-1694, 2003.

[8] J.M. Hyde,W.P.Seering, "Using input command pre-shaping to suppress multiple mode vibration," in Proc. *IEEE Int. Conf. on Robotics and Automation*, Sacramento, CA, 1991, pp. 2604-2609.

[9] J.E Marshall, *Control of Time delay system*. Stevenage, U.K: Peter Peregrinus, 1979.

[10] K. V. Singh, B.N. Datta, M. Tayagi, "Zero assignment in vibration: with or without time-delay," *in Proc. of ASME IDETC/CIE*, Las Vegas, NV, Sep. 2007.

[11] K. Gu, V.L Kharitonov, J.Chen, *Stability of time-delayed system, Boston*. Birkhauser, 2000.

[12] Md. A. Ahmad, "Active sway suppression techniques of gantry crane system", *European Journal of Scientific Research*, vol. 27(4), pp. 322-333, 2009.

[13] M. Ramesh, S. Narayana, "Controlling Chaotic motions in a two dimensional airfoil using time-delay feedback," *J. Sound Vibration*, vol. 239 (5), pp. 1037-1049, 2001.

[14] M. Bodson, "Experimental comparison of two input shaping method for control of resonant system", *in Proc. IFAC world congress*, San Francisco, CA, 1996.

[13] N. Olgac, R. Sipahi, "A new practical stability analysis method for time delayed LTI system, *in Proc. 3rd IFAC workshop on TDS (TDS 2001)*, Santa Fe, 2001.

[14] R.Dey, G.Ray, S.Ghosh, A. Rakshit, "Stability analysis for continuous system with additive time-varying delay: A less conservative result", *Applied Mathematics and Computation*, vol. 215, pp. 3740-3745, 2010.

[15] R. Dey, S.Ghosh, G.Ray, "Delay-dependent stability analysis of linear system with multiple state delays", *in Proc. 2nd IEEE ICIIS*, Srilanka, pp. 255- 260, 2007.

[16] S. Xu, J. Lam, "A survey of Linear matrix inequality techniques in stability analysis of delayed system", *Int. J. of System Science*, vol. 39(12), pp. 1095-1113, 2008.

[17] Feedback Instruments Ltd., *Operating Manual of Digital Control Experimental Set up 33-200*. East Sussex, UK, 2009.
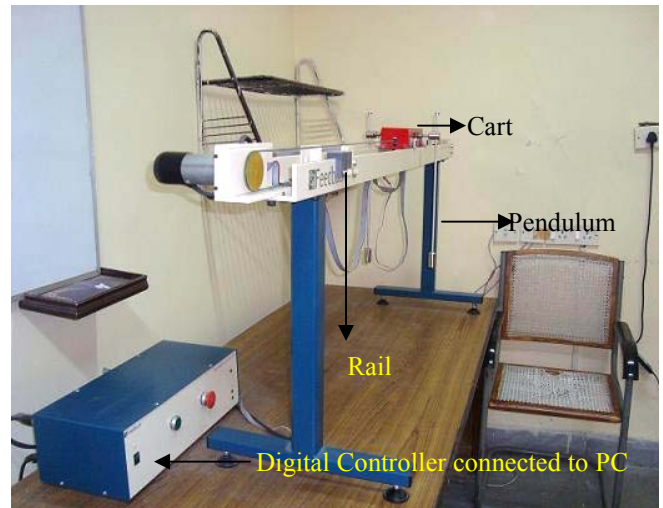
Fig. 8: Cart-Pendulum model of Feedback make Inverted Pendulum Experimental Set up.

Heading Direction Anticipation Using Visual Feedback System

Zairulazha Zainal, Rizauddin Ramli and
Mohd Marzuki Mustafa

Faculty of Engineering and Built Environment

Universiti Kebangsaan Malaysia

43600 UKM Bangi, Selangor

Malaysia

Zairulazha Zainal

Faculty of Mechanical Engineering

Universiti Teknikal Malaysia Melaka

Hang Tuah Jaya, 76100 Durian Tunggal

Melaka, Malaysia

zazain@vlsi.eng.ukm.my

Abstract - **This paper presents a technique of visual feedback system application to interpret heading direction which can be adapted to a moving system such as walking mobile robot. An optical flow is used to extract information obtained from generation of images sequence from the visual feedback system. The generation of optical flow is done by differential based method and the heading direction is represented by angle estimation. The technique was tested in two conditions in which a situation of a static camera with moving object and in a situation where moving camera with static environment. The efficiency of the visual feedback system is evaluated using root mean square error (RMSE) in order to determine the performance of proposed method in angle representation.**

Keyword: Visual feedback system; optical flow; heading direction.

I. INTRODUCTION

In recent days, control engineering becomes an essential component that cannot be avoided in our life. Applications of this area are broad, such as in washing machine, air conditioner, aircrafts autopilots, automated guided vehicles, paper-making machines and countless electro-mechanical servo-mechanisms are among the many examples. In order to analyze and design controller for the plant or system, it is important to simplify complex systems into smaller, interconnected subsystems that can be represented in Figure 1.

The structure of control system shown in Figure 1 is known as a feedback control system. In the system, the feedback mechanism information is provided through measurement of various types of sensor based on their application and the quantity measured.
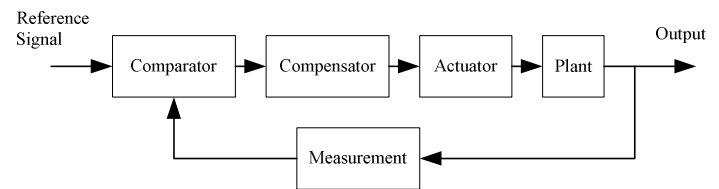
Figure 1: Elements in a control system

However, since the usage of sensors is related to their application, there are few disadvantages such as wear and tear. In addition high cost especially when deal with accuracy issues also occurs. Therefore, many researchers tried to develop new design or concept of sensors that can improve their sensing capability and reduce its cost such as integration of 2D and 3D image sensor and computer vision [1].

When dealing with computer vision, it is usually concerned with the theory behind systems with complex algorithm built-in that extract information from images, where the image data can take many forms ranging from video sequences to multi-dimensional data from a medical scanner. The estimation of motion from an image sensor is an appealing idea because of low-cost, compact and self-contained device [2]. Many studies related to vision system have been done such as in agriculture and motion detection [3-4]. In this paper, we deal with visual feedback system, which is known as visual odometry in the field of robotics. The visual odometry is defined as a process of analyzing associated camera images in order to determine the position and orientation. Visual odometry has significant contribution in the field of robotics especially in its locomotion control. The visual odometry can be classified into two categories which are dense motion algorithm or also known as optical flow and feature tracking method [2]. Between both methods, the approach of optical flow has

captured many interests of researchers in order to justify or estimate the direction of moving system [5]. It is because the optical flow is to extract the information driven from sequence of images in order to estimate current condition or position of moving system.

Study regards on visual odometry has attracted many researchers such as Kazik and Goktogan that suggested the visual odometry method to estimate the location and orientation based on Fourier-Merlin transform and phase-only matched filters [6]. Other works related includes in mobile robots [7-10], unmanned air vehicle (UAV) [11-12], automotive [13-14] and humanoid robots [15-16]. However, there have some assumptions and limitations included with it in order to successfully implement it, such as [6, 7]. Some of them need additional task such as real world map integration [8, 10].

In this paper, we present the potential of visual feedback system to be used as a sensor to estimate an angle, which represents heading direction of the system. The visual feedback system uses Horn-Schunck and Lucas-Kanade method to generate the optical flow for angle estimation from the captured images. The objective of this paper is to ascertain that the generated optical flow can be used to anticipate any current position of the moving system which can be used as a feedback input in control system. Difference from [9], the output from visual feedback system is used as a measurement value that will be used as a feedback information to controller. While in [9], generated optical flow is used to justify curvature weighted depth that will be used to identify collision point, which is useful only for limited task such as collision avoidance.

II. MOTION ANALYSIS AND OPTICAL FLOW

Generally, studies in visual feedback system are focused only on motion processing and analysis. However, in this paper, an optical flow approach has been adapted to realize these objectives. Literally, the optical flow can be defined as a pattern of apparent motion of objects, surfaces or edges in a visual scene caused by relative motion between an observer and the scene. The optical flow gives a description of motion and contributes to image interpretation even if no quantitative parameters are obtained from motion analysis.

The optical flow is derived based on following constraint equation.

$$f\left(x+\Delta x, y+\Delta y, t+\Delta t\right) \approx f\left(x, y, t\right) \qquad (1)$$

where $f\left(x, y, t\right)$ is the intensity of the image at location $\left(x, y\right)$ and t is the time of event, $\Delta x, \Delta y$ is the alteration in location, as well as Δt is the time variance.

By applying Taylor series expansion to (1) and with the assumption that the movement is so small, (1) can be expressed as,

$$\frac{\partial f}{\partial x}\Delta x+\frac{\partial f}{\partial y}\Delta y+\frac{\partial f}{\partial t}\Delta t$$
$$=\frac{\partial f}{\partial x}\frac{\Delta x}{\Delta t}+\frac{\partial f}{\partial y}\frac{\Delta y}{\Delta t}+\frac{\partial f}{\partial t}$$
$$=I_x u+I_y v+I_t = 0 \qquad (2)$$

where,

$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial t}$ is defined as spatial and temporal gradient I_x, I_y, I_t and $\frac{\Delta x}{\Delta t}, \frac{\Delta y}{\Delta t}$ is defined as optical flow vector u, v in horizontal and vertical direction, respectively.

From (2), by applying the flow vector to the spatial gradient of the image, the flow vector will be exactly cancelled by the temporal gradient, in which are logical since we assumed that there will be no changes in the brightness of the image.

Many methods have been suggested for the computation of optical flow but the most dominant are differential-based methods such as Horn-Schunck method and Lucas-Kanade method [17]. The Horn-Schunck method basically explores a motion field that satisfies the optical flow constraint with the minimum pixel-to-pixel variation among the flow vectors. On the other hand, the Lucas-Kanade method uses local method which calculates the flow vector using the constraints of a neighborhood around the pixel.

III. METHODOLOGY

Figure 2 shows the flows of the methodology used in this work. The test was done in two conditions of image sequences. First condition is where a static camera with a moving object as shown in Figure 3 (note that the object moves from right to left). On the other hand, the second condition is a static environment with moving camera as shown in Figure 4 where camera moves from left to right. Time difference between each image frame is 0.1 second. The objective of the first condition is to confirm that our hypothesis of using this condition can be used even in the condition which camera moves with static environment.

The optical flow is generated by using Horn-Schunck and Lucas-Kanade method. In addition, the performance evaluation we used a criterion from the percentage of root mean square error (RMSE). The RMSE is selected because it is frequently used to measure differences between value predicted by the model or estimator with the

value actually observed, as well as it is a good measure of accuracy.

Figure 3: Moving object with static camera (object moves from right to left)

Figure 2: Methodology workflow
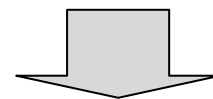
Figure 4: Moving camera with static environment (camera moves from left to right)

IV. RESULTS AND DISCUSSION

In proposed method, using Horn-Schunck and Lucas Kanade method, optical flow field is generated where a sample of optical flow field generated from sequence of images of moving camera with static environment is shown in Figure 5. Each flow vector is then decoupled into horizontal and vertical component and vector addition and subtraction operation is implemented for each component. From total vectors for each component, information regards on angle of the resultant vector is extracted.

Experimental results are shown in Figure 6, 7, 8 and 9 respectively. In Figure 6 and Figure 7, the angular value estimation is done in the first condition, i.e. static camera with moving object using both Horn-Schunck and Lucas-Kanade method, respectively. Figure 8 and Figure 9 shows the angle estimation in condition of dynamic camera with static environment using Horn-Schunck and Lucas-Kanade method, respectively. Notice that in graph, RV means resultant vector, MA means moving average, EMA means exponential moving average and CMA means cumulative moving average.

By using the RMSE criterion, performances for both conditions are evaluated and concluded as in Table I.
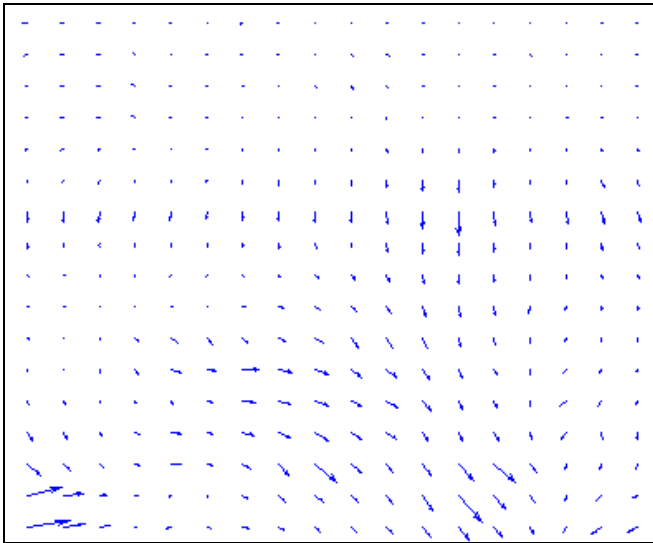
Figure 5: Sample of generated optical flow from image sequences
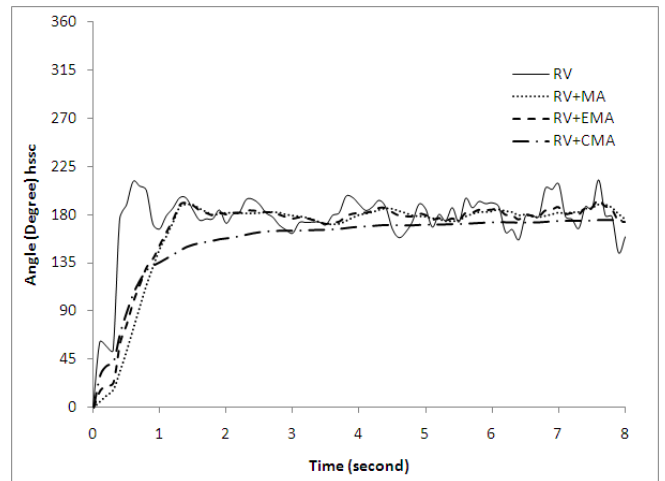
Figure 6: Angle estimation from optical flow via Horn-Schunck method (static camera)
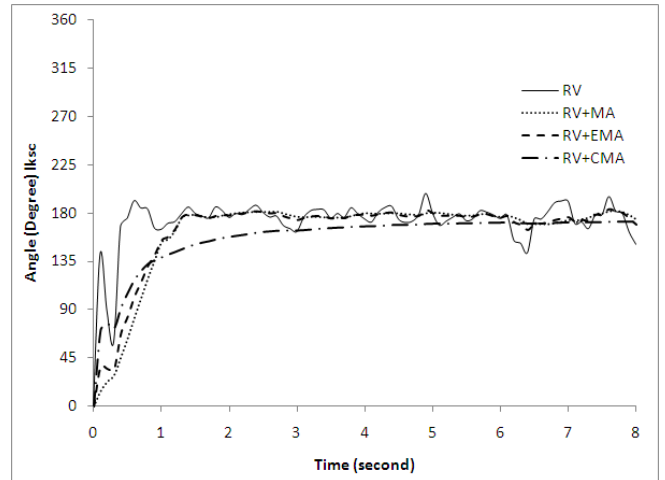
Figure 7: Angle estimation from optical flow via Lucas-Kanade method (static camera)
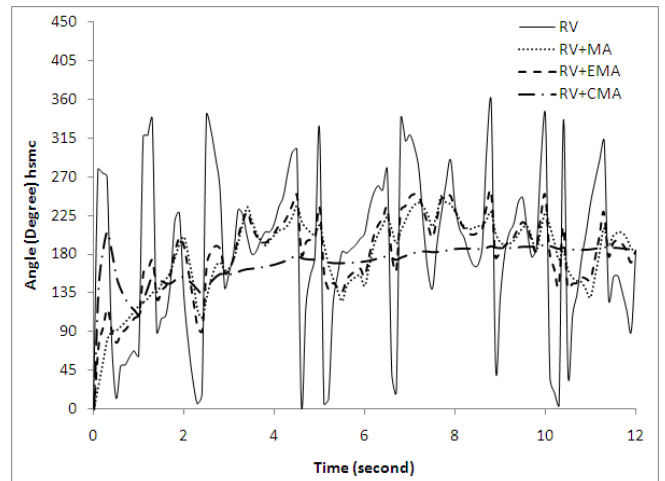
Figure 8: Angle estimation from optical flow via Horn-Schunck method (moving camera)
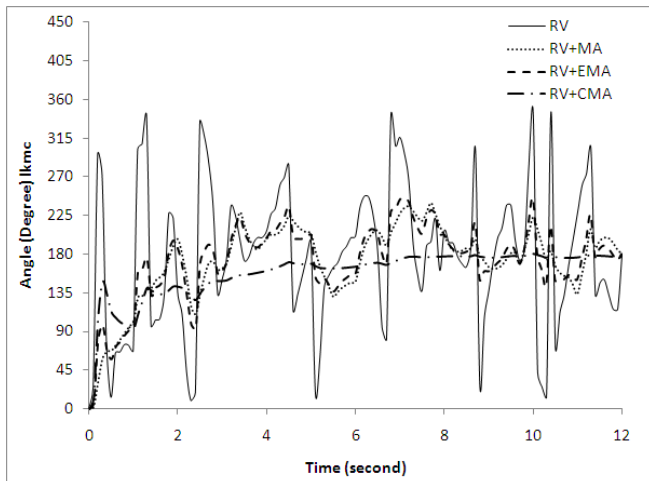
Figure 9: Angle estimation from optical flow via Lucas-Kanade method (moving camera)

From Figure 6 to Figure 9, for RV, it shows fluctuation, and this fluctuation is reduced through implementation of filtering techniques, which the reduction can be observed through results for RV+ MA, RV + EMA and RV + CMA. For evaluation purpose, from graphs, each single estimation is treated as a single measurement. In this case, for situation where static camera is involved, 80 single measurements has been done, while for situation where moving camera is involved, 120 single measurement has been done. Using these data, RMSE computation has been implemented, and results are shown as in Table I.

TABLE I.   PERCENTAGE OF RMSE FOR ANGLE ESTIMATION

Condition	Static Camera		Moving Camera	
Method	Horn-Schunck	Lucas-Kanade	Horn-Schunck	Lucas-Kanade
Resultant Vector	15.1	11.1	51.9	46.3
Resultant Vector + Moving Average	24.1	22.4	23.7	24.2
Resultant Vector + Exponential Moving Average	21.4	19.7	23.7	23.7
Resultant Vector + Cumulative Moving Average	20.9	17.3	11.3	17.7

From Figure 6 to Figure 9, the heading direction by angle estimation can be anticipated with the same method. In Table I, it can be ascertained that from the RMSE calculation, it shows that the static camera gave better results by using resultant vector (both Horn-Schunck and Lucas Kanade method) directly as a method to represent heading direction of the system. Even though from Figure 6 and Figure 7, resultant vectors shows fluctuation, if considering as a single measurement at each 0.1 second, error generated between estimated value and actual value for overall data is relatively small, as shows in Table I. This is due to the fact that RMSE computation is more sensitive to occasional large error, thus give large percentage of error in case of RV+MA, RV+EMA and RV+ CMA.

It is assumed that because of the environment is controlled and static camera is used, less noise is observed from the image sequences; hence the result generated using this technique show a good accuracy without filtering. In the other hand, for moving camera, noise is observed while camera is moving. Table I shows that for situation where moving camera is involved, the result generated using resultant vector gave less accuracy than the result generated using static camera and moving object. Therefore, filtering is needed for this case. From all the three types of filter that has been used, the combination of resultant vector and cumulative moving average technique (for both Horn-Schunck and Lucas-Kanade method) shows the best performance.

V.   CONCLUSION

In this paper, the performances of the visual feedback system have been evaluated. In conventional way, output from plant or system is measured using sensor, which measurement input usually takes form of quantitative value. However, in visual feedback system, which is used as measurement instrument, measurement input takes form of qualitative value, and through processing that is completed in the visual feedback system, the output of visual feedback system is represented in form of quantitative value, in this case, angle estimation, where this value can be used as feedback information to control a system. Two conditions of visual feedback systems have been considered that is a condition of static camera with moving object and a condition of static environment with moving camera. The Lucas-Kanade method gave better performance in representing the heading direction of the moving object. Furthermore, for moving camera with static environment, combination of resultant vector and cumulative moving average based on Horn-Schunck method shows best performance. Finally, we ascertained that from the extraction of information captured by the visual system via Horn-Schunck and Lucas-Kanade's optical flow technique, the heading direction of a moving system such as mobile robot can be anticipated.

VI.   ACKNOWLEDGEMENT

REFERENCES

[1]   M. C. Lin, C. N. Chan and O.T.-C. Chen, "2D and 3D-integrated image sensor," 53rd IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), 2010, pp. 292–295.

[2]   K. Konolige, M. Agrawal and J. Sola, "Large scale visual odometry for rough terrain," International

Symposium on Research in Robotics 2007 (ISRR 2007), 2007, pp. 1-12.

[3] K. H. Ghazali, M. M. Mustafa and A. Hussain, "Machine vision system for automatic weeding strategy using image processing technique," American-Eurasian Journal of Agricultural & Environmental Sci., 3 (3), 2008, pp. 451-458.

[4] J. Kanawathi, S. S. Mokri, N. Ibrahim, A. Hussain and M. M. Mustafa, "Motion detection using Horn Schunck algorithm and implementation," International Conference on Electrical Engineering and Informatics 2009 (ICEEI '09), 2009, pp. 83-87.

[5] N. G. Hatsopoulos and W. H. Warren, Jr, "Visual navigation with a neural network," Neural Networks vol. 4, 1991, pp. 303-317.

[6] T. Kazik and A. H. Goktogan, "Visual odometry based on the Fourier-Mellin transform for a rover using a monocular ground-facing camera," Proceedings of the 2011 IEEE International Conference on Mechatronics, 2011, pp. 469-474.

[7] E. Bayramoglu, N. A. Andersen, N. K. Poulsen, J. C. Andersen and O. Ravn, "Mobile robot navigation in a corridor using visual odometry," International Conference on Advanced Robotics 2009 (ICAR 2009), 2009, pp. 1–6.

[8] A. Levin and R. Szeliski, "Visual odometry and map correlation," Computer Vision and Pattern Recognition, 2004, *v*ol. 1 pp. 611-618.

[9] B.J.A Krose, A. Dev and F.C.A. Groen, "Heading direction of a mobile robot from the optical flow," Image and Vision Computing 18, 2000, pp. 415–424.

[10] S. Indu, M. Gupta and A. Bhattacharyya, "Vehicle tracking and speed estimation using optical flow method," International Journal of Engineering Science and Technology (IJEST), 2011, Vol. 3, No. 1, pp. 429–434.

[11] M. V. Srinivasan, S. Thurrowgood and D. Soccol, "An optical system for guidance of terrain following in UAVs," Proceedings of the IEEE International Conference on Video and Signal Based Surveillance AVSS '06, 2006, pp. 51–51.

[12] D. S. Thurrowgood and M. Srinivasan, "A vision system for optic-flow-based guidance of UAVs," Proceedings of the 2007 Australasian Conference on Robotics & Automation, 2007.

[13] N. Nourani-Vatani, J. Roberts and M. V. Srinivasan, "Practical visual odometry for car-like vehicles," Proceedings of the 2009 IEEE International Conference on Robotics and Automation (ICRA '09), 2009, pp. 3551–3557.

[14] B. Kitt, A. Geiger and H. Lategahn, "Visual odometry based on stereo image sequences with RANSAC-based outlier rejection scheme," IEEE Intelligent Vehicles Symposium (IV), 2010, pp. 486–492.

[15] A. Pretto, E. Menegatti, M. Bennewitz, W. Burgard and E. Pagello, "A visual odometry framework robust to motion blur," IEEE International Conference on Robotics and Automation 2009 (ICRA '09), 2009, pp. 2250-2257.

[16] S. P. N. Singh, P. J. Csonka and K. J. Waldron, "Optical flow aided motion estimation for legged locomotion," IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006, pp. 1738-1743

[17] A. Murat Tekalp, Digital Video Processing, Prentice Hall Inc, 1995, pp. 83-84.

Customized Instruction Set Simulation for Soft-core RISC Processor

Ahmad Jamal Salim, Sani Irwan Md Salim, Nur Raihana Samsudin, Yewguan Soo
Faculty of Electronics and Computer Engineering
Universiti Teknikal Malaysia Melaka
Melaka, Malaysia
e-mail: shaj@utem.edu.my, sani@utem.edu.my

Abstract—**This paper presents the instruction set simulation process for soft-core Reduced Instruction Set Computer (RISC) processor. The aim of this paper is to provide reliable simulation platform in creating customizable instruction set for Application Specific Instruction Set Processor (ASIP). The targeted RISC processor is based on basic 8-bit PIC16C5X-compatible processor where the architecture is reconfigurable through Hardware Description Language (HDL). Instruction set architecture (ISA) has been modified in term of instruction width and machine instruction. Memory address remapping algorithm is introduced to remap the memory address to correct physical memory address due to memory banking scheme being applied. A total number of 34 instruction sets are simulated and verified its operations. Selected instruction set has been reconfigured from its original operation to demonstrate the ability to modify current instruction set to suit certain specialized application. The simulation is done using a Java-based CPU architecture simulation program and data movements at file register array and memory registers are monitored to verify the correct working operation of each instruction set. The instruction set simulation process is proved capable to be the starting point in creating a reconfigurable RISC processor with customized instruction set, inline with ASIP methodology.**

Keywords- RISC; instruction set; ASIP

I. Introduction

The importance of reconfigurable processor in embedded system design has significantly risen from the past two decades. Due to the advancement of field programmable gate array (FPGA), the term soft-core processor was introduced to reflect the capability of the processor to change its architecture instantaneously [1]. A soft-core processor also provides designers the ability to tweak, manage and structure the architecture to achieve specific performance and resource usage. Xilinx's MicroBlaze and Altera's NIOS II processor are two popular soft-core processors due to its manufacturer support and respective platform optimization. However, there are variants of soft-core processors being developed by researchers worldwide in the form of licensed IP or open-source [2, 3]. The emergence of ARM-Cortex processor, which is embedded in FPGA chip, also brings new alternative in embedded system development.

Soft-core processor in FPGA normally utilized as the main controller of the overall system [4]. System peripheral and co-processor are implemented in FPGA modules and synchronize by the processor. Therefore, the modules could be executed in parallel while the processor program handles the control algorithm sequentially. Programming a soft-core processor is no different from any discrete microprocessor available where the resulting programming code is initiated in the processor's ROM prior to the FPGA implementation. Its reconfigurable features certainly suit the customizable instruction set of ASIP as it introduces more flexibility in the design [5].

Among all soft-core processors, RISC design is widely adopted for its single clock cycle instructions and less resource requirement compared to CISC approach. Although there is a system that successfully incorporates CISC processor [6], RISC processor offers more flexibility in instruction set extension to suit specific application. Other than that, RISC's simple instruction set architecture is expandable and reconfigurable to accommodate new instructions that can perform more complex function [7, 8].

This paper describes the custom simulation of a RISC soft-core processor's instruction set that is based on Microchip PIC16C5X architecture. All the instructions are simulated with modified instruction set architecture to accommodate instruction set extension process at later stage. The simulation also considered memory address re-mapping where memory bank organization is implemented to the architecture. Addressing mode for the instruction, including direct and indirect addressing is also configured with reference to the respective file and memory register. Simulation process is done on a highly customizable Java application computer architecture simulator [9]. It provides features to insert customized instruction in an assembly language and have the ability to perform simulation at microcode level to variety of CPU architectures.

II. Simulation Method

A. Instruction Set Architecture (ISA)

Essentially, ISA specifies all the executed instruction including its fetched operand, arithmetic or logic operation and the results' destination [10]. For PIC16C5X, the instruction set width is set to 12-bit wide encompassing opcode and literal bit or file register or memory address [11]. Instructions' opcode are varied between 3-6 bits depending on its mode of operation. The next corresponding bits in the instruction code are file register, literal value or bit address. Instead, in the

conducted simulation process, the ISA is modified from 12-bit to 16-bit width in view to accommodate any new instruction set extension in the future. Fig. 1 shows the general instruction set format that is used in the simulation.

The opcode consists of 6 bits that can provide maximum 2^6 of instruction set. This is followed by 9 bits of operand that can represent memory/register address or literal values. The final 15th bit is the destination bit where '1' indicates the result is stored in a designated file register while '0' indicates the result is stored in working register (accumulator). The bit segregation is varied according to its instruction operation that includes byte-oriented, bit-oriented and literal and control.

B. Simple Computer Architecture

In the simulator's CPU architecture, a simple hypothetical computer system called SHAJ01PC is utilized. The system is based on Harvard architecture that resembles a 16-bit RISC processor with 128 words of RAM as shown in Fig. 2.

C. Memory Organization

Generally, the memory organization adopts Harvard architecture where data memory and program memory are separated in two modules. A banking scheme is used that consists of 4 banks in total and data memory is accessible through File Selection Register (FSR). Special Function Registers (SFR) is fixed from address 00 to 07 regardless of memory bank selection. Other addresses are configured as general purpose registers. Therefore, an algorithm is developed to remap the physical memory address to adopt the memory bank requirement of this scheme.
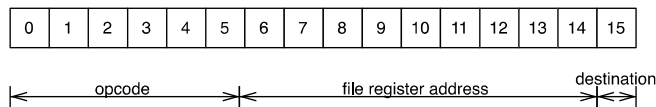
Figure 1.   General instruction format.
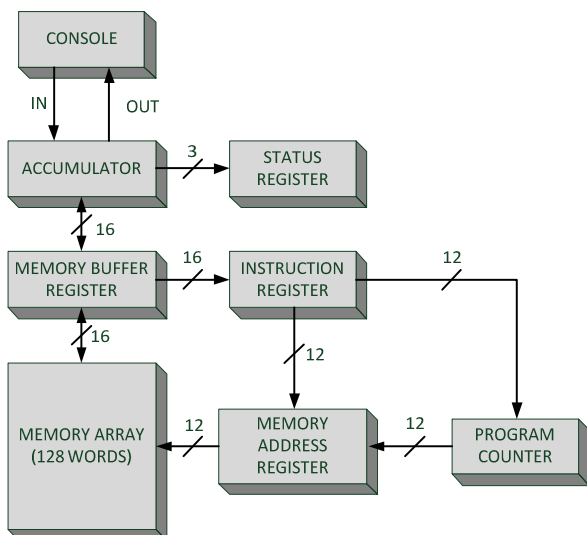
Figure 2.   Simple computer architecture.

III. INSTRUCTION SET LIST

Table 1 shows the total instruction set and its description that are successfully simulated.

TABLE 1.   LIST OF ALL INSTRUCTIONS

Mnemonics	Descriptions	STATUS affected
clrwdt k	clear watchdog timer	~TO, ~PD
movwf f	move w to f, w=>f	none
sleep	go to standby mode	~TO, ~PD
tris f	load tris register	none
nop	no operation	none
option k	load option register	none
clrw	clear w	Z
clrf f	clear f	Z
subwf f,d	subtract w from f, d=f-w	C,DC,Z
decf f,d	decrement f	Z
iorwf f,d	inclusive OR w with f	Z
andwf f,d	AND w with f	Z
xorwf f,d	exclusive OR w with f	Z
addwf f,d	add w and f	C,DC,Z
movf f,1 rename to swapfw	swap f with w, f⇔w	none
movf f,0	move f to w, f=>w	Z
comf f,d	complement f	Z
incf f,d	increment f	Z
decfsz f,d	decrement f, skip if 0	none
rrf f,d	rotate right through carry	C
rlf f,d	rotate left through carry	C
swapf f,1	swap nibble f and store in f	none
swapf f,0 rename to mulw	multiply f & w	none
incfsz f,d	increment f, skip if 0	none
bcf f,b	bit clear f	none
bsf f,b	bit set f	none
btfsc f,b	bit test f, skip is clear	none
bifss f,b	bit test f, skip if set	none
retlw k	return, place literal in w	none
call k	call subroutine	none
goto k	unconditional branch	none
movlw	move literal to w	none
iorlw k	inclusive OR literal with w	Z
andlw k	AND literal with w	Z
xorlw k	exclusive OR literal with w	Z

IV. RESULTS AND DISCUSSIONS

To test all the instruction, an assembly language test program is created inside the simulator program. The program includes short routines to implement the instructions with difference operands and addressing mode. The test program also evaluates the remapping of memory address with regards to physical memory addressing in banking scheme.

In the instruction set simulation, two instruction sets' operation have been modified to demonstrate the capability of the simulator in verifying operation of new instruction. Those instruction sets are "movf f,1" and "swapf f,1". In "movf f,1" (rename as "swapfw" in simulator), instead of moving data from file register to its own, the microinstruction is modified to perform swap operation between file register and working register. In addition, instruction "swapf f,1" (rename as "mulw" in simulator) is modified to perform multiplication of file register and working register. The 16-bit multiplication result is stored back in f and w register respectively.

A. Instruction Set Modification – "swapfw"

Fig. 3 shows the machine instruction modification for "movf f,1" that is renamed as "swapfw". The instruction opcode remains the same yet the instruction's operation has been modified to perform a byte-swap operation between working register and FSR in indirect addressing mode. The value of FSR is stored first in register called "dum" while data from working register overwrite the FSR. Data from "dum" then is transferred to working register to complete the swap.

An assembly language test program is developed as shown in Fig. 4. It demonstrates a swap between signed number -6 (stored in FSR) and signed number -2 (stored in working register). Data are observed between registers window (Fig. 5) and register array window (Fig. 6). The first 5 lines in the assembly test program initiated the value for FSR and working register. Then the "swapfw" instruction is executed in a loop and data movement is monitored at their respective registers.

Figure 3.   Machine instruction modification - swapfw
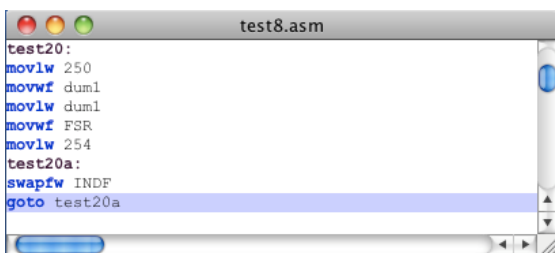
Figure 4.   Assembly test program.
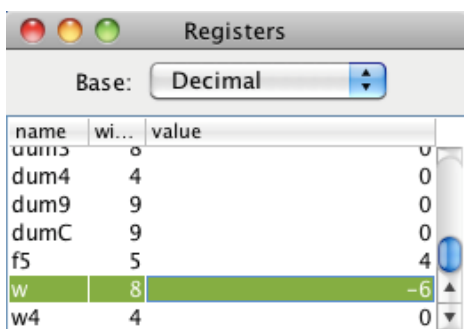
Figure 5.   Registers window.

Figure 6.   Register array window.

B. Instruction Set Modification – "mulw"

The original set of instruction in PIC architecture does not contain multiplication although such operation can be achieved with combination of basic instruction such as bit rotation and bit arithmetic. However, multiplication can be implemented by tweaking the machine instruction to determine the function, source and destinations.

In this case, new instruction "mulw" is created from original "swapf f,0" instruction. In order to perform multiplication, the machine instruction is change as shown in Fig. 7. The multiplication involved two 8-bits data stored in file register and working register. The result, which is a 16-bit data, is temporarily stored in register "dum16". As the architecture permits 8-bit data bus only, the result is distributed to two registers. The first 8 bits are stored in working register while the second 8 bits are stored in file register.

Fig. 8 shows the assembly test program to implement "mulw" instruction. In this code, the multiplicand is stored in register FSR while the multiplier is in working register. The multiplication result is stored in both INDF register (memory address 00) and working register. The data pointed by FSR then is incremented by 1 and the updated data is transferred to working register for the next cycle.
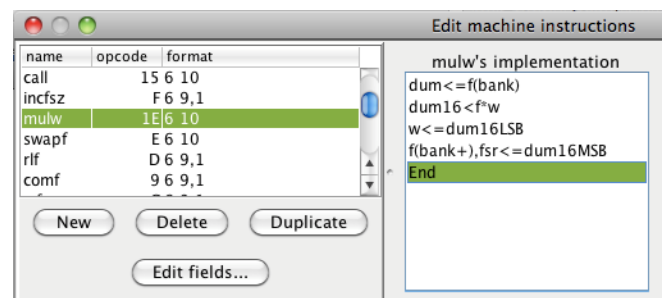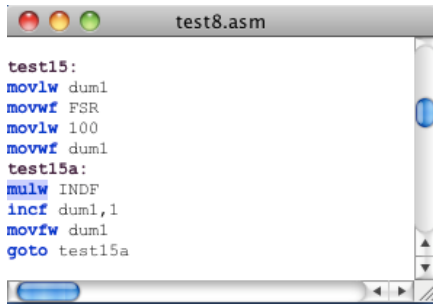
Figure 7.   Machine instruction modification - mulw.

For the first past, the both FSR and working register are loaded with 64_{10}. The result, which is 2710_{16}, is initially stored in register "dum16". Then the result is divided into two 8-bit data and stored in INDF register and working register respectively. Fig. 9 and Fig. 10 show all the affected registers and the updated values during the "mulw" instruction execution. For the next iteration, the data pointed by FSR is incremented to 65_{16} and copied to the working register. The multiplication instruction is executed again with 65_{16} x 65_{16} operation.
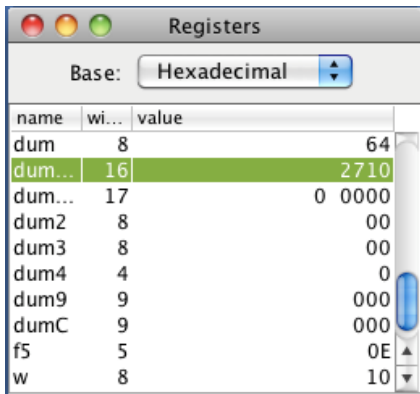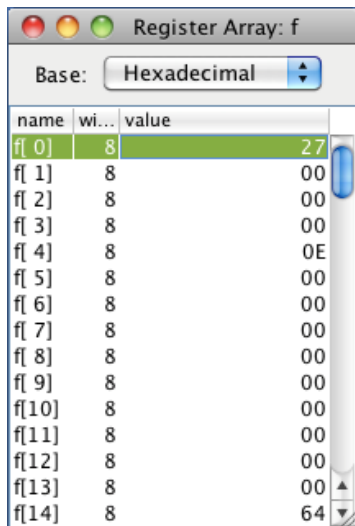
Figure 8.   Assembly test program - mulw.

Figure 9.   Registers window - mulw.

Figure 10. Registers array window - mulw.

The same method is applied to all 34 instructions and all instruction operations are observed and verified through registers and memory transfer activities. Each instruction is assigned by its own assembly test routine that include direct and indirect data addressing mode. Indirect data addressing mode is supported through file select register and INDF register.

In addition, all the simulated instruction sets are based on PIC instruction set RISC architecture. The simulated instruction opcode are consistent to its PIC counterpart while at the same time maintaining the correct operation and affected STATUS register. This ensures direct compatibility of instruction set architecture when implementing the new or modified instruction on soft-core RISC processor in FPGA. Furthermore, comparison can be made between the simulated instruction and real execution of PIC instruction set for performance benchmarking.

V.   CONCLUSIONS

Soft-core processors offer a viable platform in customizing the architecture of a microprocessor to suit specific requirement of a project. The ability to modify the instruction set architecture of a processor also being regarded as one of the method in ASIP. This paper has described the method in simulating instruction set for a RISC processor on top of Java-based CPU architecture simulator. All instruction set has been verified in working condition with machine instruction modification capability and compatible with memory bank scheme. To demonstrate the ability to create/modify new instruction, selected instruction set has been configured to perform custom operations and the simulation results are evaluated at register level. For future work, the simulated instruction set, as well as the modified instruction set will be programmed in reconfigurable platform such as FPGA, to be executed as a customized RISC processor prototype.

REFERENCES

[1]   J. Ball, "Designing Soft-Core Processors for FPGAs Processor Design," in *Processor Design: System-on-Chip Computing for ASICs and FPGAs*, J. Nurmi, 1st Ed: Springer Netherlands, 2007, pp. 229-256.

[2]   L. Barthe, L. V. Cargnini, P. Benoit, and L. Torres, "The SecretBlaze: A Configurable and Cost-Effective Open-Source Soft-Core Processor," in *IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW)*, 2011, pp. 310-313.

[3]   G. Kuzmanov, G. Gaydadjiev, and S. Vassiliadis, "The MOLEN processor prototype," in 12th Annual IEEE Symposium on Field-Programmable Custom Computing Machines, 2004, pp. 296-299.

[4]   J. McAllister, "FPGA-based DSP," in *Handbook of Signal Processing Systems*, S. S. Bhattacharyya, E. F. Deprettere, R. Leupers, and J. Takala, 1st Ed: Springer US, 2010, pp. 363-392.

[5]   C. Galuzzi and K. Bertels, "The Instruction-Set Extension Problem: A Survey," *ACM Trans. Reconfigurable Technol. Syst.,* vol. 4, 2011, pp. 1-28.

[6] A. J. Salim, M. Othman, and M. A. M. Ali, "The Development of a Programmable DSC Chip: UKM8051DSC," *European Journal of Scientific Research,* vol. 19, 2008, pp. 350-361.

[7] W. Wenxiang, L. Ling, Z. Guangfei, L. Dong, and Q. Ji, "An Application Specific Instruction Set Processor optimized for FFT," in *2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2011, pp. 1-4.

[8] X. Guan, Y. Fei, and H. Lin, "Hierarchical Design of an Application-Specific Instruction Set Processor for High-Throughput and Scalable FFT Processing," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 20, 2011, pp. 551-563.

[9] D. Skrien, "CPU Sim 3.1: A tool for simulating computer architectures for computer organization classes," *Journal on Educational Resources in Computing (JERIC),* vol. 1, 2001, pp. 46-59.

[10] D. Liu, *Embedded DSP Processor Design, : Application Specific Instruction Set Processors*: Morgan Kaufmann, 2008.

[11] Microchip Technology Inc., "PIC16C5X Datasheet", 2002

GMT feature extraction for representation of BIM sign language

Noor Saliza bte. Mohd Salleh

Centre for Graduate Studies
National Defence University of Malaysia
Kuala Lumpur, Malaysia
noorsaliza@uniten.edu.my

Suzaimah Ramli

Faculty of Defence Science and Technology
National Defence University of Malaysia
Kuala Lumpur, Malaysia
suzaimah@upnm.edu.my

Abstract— **This paper focuses on the features extraction from the hand region segmentation involved in the earlier pre-processing module for Bahasa Isyarat Malaysia (BIM) sign language recognition. This Human Computer Interaction (HCI) research intention is to provide a system for aiding those with impaired hearing community to communicate with the public community in Malaysia specifically. Vision based technique has been proposed for data collection technique. Gestural Motion Trajectories (GMT) are used to keep track the movement of the hand. The objectives of this paper is to determine whether the feature extraction methods used can produced good features to get the discriminate value to distinguish each sign language.**

Keywords- BIM, gesture, hand-detection, feature extraction, motion trajectories

I. INTRODUCTION

Gesture recognition is one of the most important fields in Human Computer Interaction (HCI). With gesture recognition especially in the area of Automatic Sign Language Interpretation, deaf people can communicate easily with a non-deaf person without a human interpreter [1]. "Bahasa Isyarat Malaysia" (BIM) or Malaysian Sign Language has become an identity for deaf people in Malaysia as a common tool for them to communicate [3]. As per current, less research has been done on BIM domain especially in developing BIM recognition system. For example, a research on BIM has been done by [4] and [5].

Organization of this paper is as follows; Section 2 will review some related works that has been done in specific area, sign language recognition, hand detection and motion analysis. Methodology of the research is reviewed in Section 3. Then, our current stage of development, our hand detection technique and also feature detection that have been extracted from hand segmentation. Current phase of our system development and future direction also will be discussed.

II. PREVIOUS WORK

While there are many interesting domains for gesture recognition, one of the most structured sets of gestures is those belonging to any of the several sign languages [6]. "Bahasa Isyarat Malaysia" (BIM) or Malaysian Sign Language has become an identity for deaf people in Malaysia as a common tool for them to communicate [3]. BIM is one of the communication systems, oriented on visual and hand gesture [2]. Sign language recognition has become a popular field in research area especially outside Malaysia such as American Sign Language [7], Korean Sign Language [8] and Arabic Sign Language [9].

Several sign language recognition systems have been developed using various features computed from static images or image sequences [10]. And several of the systems are being developed based on the extraction of some of the properties that are associated with the images, such as image geometry parameter using image moments. The image geometry parameter was employed by [7] in their research work.

Vision based approach utilizes cameras to capture images of the signer performing sign language movement. The input gesture video sequence is processed to extract features that are appropriate for classification of sign language recognition [11]. He also added that the basic components of a sign gesture are (i) the handshape, (ii) hand orientation, (iii) its location and (iv) its movement. Handshape refers to the finger configuration, orientation refers to the direction in which the palm and fingers are pointing, and location refers to hand position relative to the body. Hand movement traces out a trajectory in space.

Feature extraction is a process of studying and deriving useful information from the filtered input patterns. The derived information may be general features, which are evaluated to ease further processing [12], [18]. He also added that the selection of features is very important because there might be only one or two values, which are significant to recognize a particular segmented character or sign.

However, vision based approach has its disadvantage. [7] explained that, video based approach constraints conflict with recognition in a natural context, either by requiring simple, unchanging backgrounds like clothing, not allowing occlusion, requiring carefully labeled gloves if it has been used or even being difficult to run in real time.

In recent research and development, [13, 17] implemented a motion trajectories as what they defined as statistical modeling. The process involved the mass recordings from visual surveillance that are processed to extract objects trajectories. By having a trajectory classes defined earlier, the trajectories classes in video sequence are marked and models are adapted according to them. The outcome of the method is discussed and the researchers concluded that a successful classification of trajectories and for search of abnormal trajectories dependable on the well defined scene with well defined scenarios. Scenarios that are not well defined will affect the anomalous of trajectories as well as recognition.

Hence, in sign language recognition system development, the feature extraction can be concluding as one of the most crucial part that needs to be focused on before it's being used for classification process. Therefore, by developing this system, it will be useful for others researcher to adopt the appropriate feature extraction methods for classification process in developing such sign language recognition system. Indirectly, these can be a contribution to the society especially for deaf people in Malaysia by developing the BIM sign language recognition system to ease their daily communication life.

III. METHODOLOGY

In this research, we focused on two sub-topic, food and festival based on the BIM Books (1st Edition).

Each image frames has a spatial resolution of 768x576 pixels and a grayscale resolution of mono 8 bits. The samples were taken from same distances by digital camera, with same orientations and fix background.

A sequence of raw image frames from a video is processed using a sequence of image processing methods into a raw feature vector representing the information of the raw images. The raw feature vectors is then analyzed with the gestural motion trajectories and scaled into an appropriate feature vector values. Such a vector values is appropriate to represent a sequence of images as the gestural motion analysis is found to be efficient in recognizing value of image motions.

A. Pre-Processing

Generally hand detection comprises of subtracting the signers image from the image background. Finding hand region using colour images is not difficult because the nature of skin colour has its own unique value and can easily being processed. However, dealing with grayscale input image forced us to automate input threshold value to measure the information thus extracted the hand region from the whole image.

The goal of image extraction is to extract the raw image from the original video into filtered images for future feature extraction process. Each image frames extracted as a raw data from the video extraction process is processed by image extraction producing filtered images for the feature extraction.

Hand detection pre-processing flow shows in figure 1.
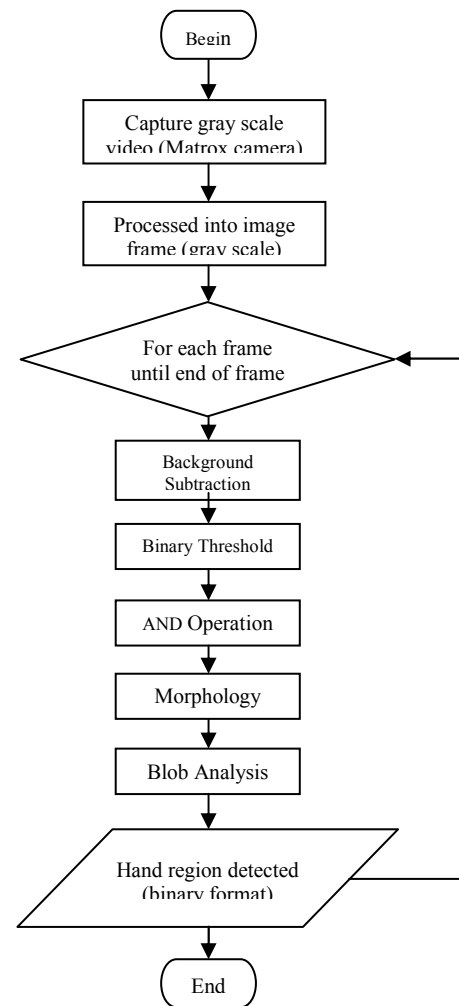
Figure 1. Hand Detection pre-processing

System starts by capturing an image without any signers with a still camera setup towards a certain view. Image captured as in gray scale image input. System stored the image of the background. The next process will captured the background image with the existence of a signer. To proceed with hand detection, image captured will be subtracted by the stored image background to remove the background. By having the only image of the signer's, we now proceed with binary thresholding. Automated function of input binary value is available as system not dealing with color image acquisition which is easier to deal with skin color for hand detection. Bitwise AND operation sets the resulting bit to 1 if the corresponding bit in both image of result from subtraction and binary threshold is 1. Bitwise manipulation enhanced the wanted image of the hand region. Further processing with morphological filters may be used to clean up the segmented hand region. However, due to the nature of same hand and face color, we might get the results of 3 regions of face and both hands. Blob analysis will identify connected regions of pixels within an image, and then calculated selected features of those regions. For this we use blob analysis and find the maximum two values of regions which is those two hand's blob.

B. Feature Extraction

Importance of the hand recognition is the ability to recognize and track the hand region. Process taken place after the hand region segmentation is the feature detection analysis.

The goal of feature extraction is to extract the vector of the processed image and translate it into some meaningful value thus to be used as inputs to the classification process. Each frame is processed in sequence of mentioned flow process. Figure 2 shows the flow.
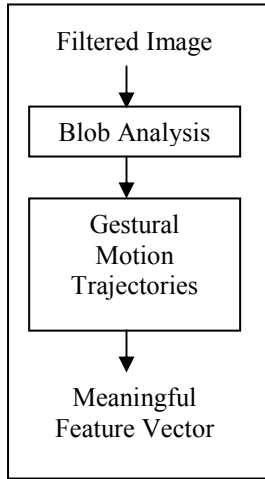
Figure 2. Feature Extraction Flow Process

1) Blob Analysis: Detection of the position, size and angle of hand image is subjected to blob analysis. Research efforts have investigated this issue, for example, [14, 15 and 16]. Each frame image is analyzed using blob. With this implementation, the grayscale Blob analysis count, label and measure image features.

2) Gestural Motion Trajectories (GMT): In our current research, we have evaluated the image moments coordinates (1 to 5) to find the center of region (6) and central moments (7 to 10). The calculation result to the major and minor axes (11). The central moments are normalized with respect to the zeroth moment to yield the normalized central moment (12). The normalized and the central normalized moments were normalized with respect to scale (area) and translation (location). Normalization with respect to orientation is provided by moment invariants (13 to 14).

$$M_{00} = \sum_x \sum_y I(x, y) \qquad (1)$$

Measurement to Find Image Moment Coordinates 00

$$M_{10} = \sum_x \sum_y x I(x, y) \qquad (2)$$

Measurement to Find Image Moment Coordinates 10

$$M_{01} = \sum_x \sum_y y I(x, y) \qquad (3)$$

Measurement to Find Image Moment Coordinates 01

$$M_{20} = \sum_x \sum_y x^2 I(x, y) \qquad (4)$$

Measurement to Find Image Moment Coordinates 20

$$M_{02} = \sum_x \sum_y y^2 I(x, y) \qquad (5)$$

Measurement to Find Image Moment Coordinates 02

$$x_c = \frac{M_{10}}{M_{00}} \qquad y_c = \frac{M_{01}}{M_{00}} \qquad (6)$$

Measurement to Find Center of Region

$$\mu_{10} = \mu_{01} = 0 \qquad (7)$$

Measurement to Find Central Moments 10

$$\mu_{11} = M_{11} - \frac{M_{10}M_{01}}{M_{00}} \qquad (8)$$

Measurement to Find Central Moments 11

$$\mu_{20} = M_{20} - \frac{M_{10}^2}{M_{00}} \qquad (9)$$

Measurement to Find Central Moments 20

$$\mu_{02} = M_{02} - \frac{M_{01}^2}{M_{00}} \qquad (10)$$

Measurement to Find Central Moments 02

$$\tan\theta = \frac{1}{2}\left(\frac{\mu_{02} - \mu_{20}}{\mu_{11}}\right) \pm \frac{1}{2\mu_{11}}\sqrt{\mu_{02}^2 - 2\mu_{02}\mu_{20} + \mu_{20}^2 + 4\mu_{11}^2}$$

(11)

Measurement to Find Major and Minor Axes

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}, \gamma = \frac{(p+q)}{2+1}$$

(12)

Measurement to Find Normalized Central Moment

$$\varnothing_1 = \eta_{20} + \eta_{02}$$

(13)

Measurement to find Moment Invariants 1

$$\varnothing_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

(14)

Measurement to find Moment Invariants 2

From the image processing module, the hand image of interest is established based on region segmentation and moment invariants which are calculated for each of the image. The value is later being normalized into the range of -1 to 1 to suits the input for the classification. Our system uses these normalized moment invariants as final input for the classification phase. The sample of sequence frames image used in this paper is shown in figure 3. Table 1 shows the final feature vector result and figure 3 shows the plotted gestural motion trajectories using the result obtained in figure 4.
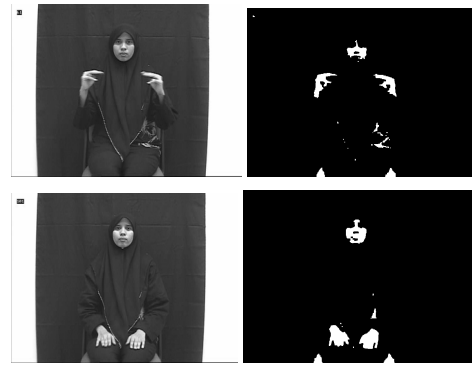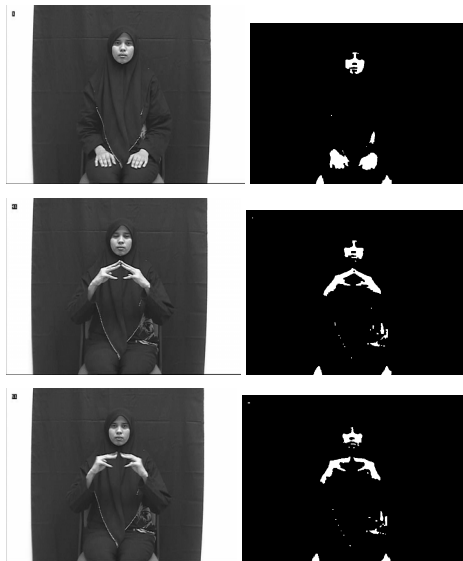
Figure 3. Left Column: Raw Sequence Frames (Frame 1, 40, 50, 60,100) of Data 638–Signer 1
Right Column: Processed Image Frames (1, 40, 50, 60,100) of Data 638–Signer 1

TABLE I. FINAL FEATURE VECTORS (FRAME 1, 40, 50, 60,100) OF DATA 643 (NASI LEMAK) – SIGNER 1

Frame Number	Normalized Feature Vectors			
	Left Hand Palm, X Axis	Left Hand Palm, Y Axis	Right Hand Palm, X Axis	Right Hand Palm, Y Axis
1	0.00451779708954714	1	1	1
40	1	-0.913720394602104	-0.74322010513 9274	-0.2307 858655 9949
50	-0.0935692287996505	-0.902962568123386	-1	-1
60	-0.140079551369688	-0.890760568970285	-0.71268 4460157 438	-0.2190 786532 16709
100	-0.00599436826390998	-0.879724456343415	-0.74317 0818148 656	-0.2117 440536 58403
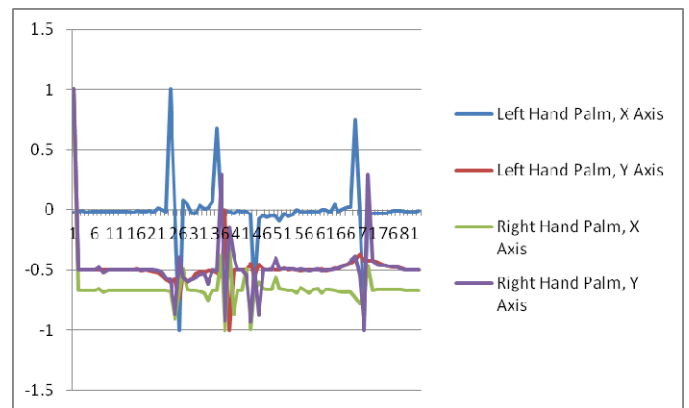
Figure 4. Graph of Velocity of the Feature Vectors relatives to the Processed Image Frames

C. Feature Selection

The objective of feature selection is to select the features that can show discriminate values to differentiate between each sign language performed.

A few features have been analyzed for each feature vector as shown in table 2 and table 3. From the data, the minimum and maximum point (values) and their occurrence (frame no.) shows discriminate values that shows different pattern for each sign language. The graph is plotted as shown in figure 4 and figure 5 to see the pattern for each sign language.

TABLE II.   ANALYZED OF FEATURES VECTORS OF DATA 643 (NASI LEMAK) – SIGNER 1

	643-Nasi Lemak			
	Left Hand Palm, X Axis	Left Hand Palm, Y Axis	Right Hand Palm, X Axis	Right Hand Palm, Y Axis
Min Point	-1	-1	-1	-1
Min at frame	25	37	36	69
Max Point	1	1	1	1
Max at frame	23	1	1	1
Median	-0.0149485	-0.4980391	-0.667432	-0.4975684
Mode	#N/A	#N/A	#N/A	#N/A
Slope Y to X	0.029481133		0.972538324	
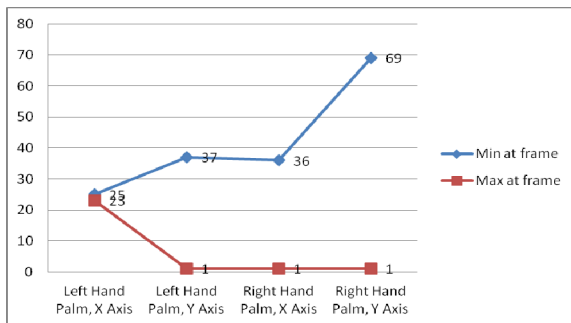
Figure 5.   Graph of Minimum and Maximum point (value) for each feature vectors of Data 643 (Nasi Lemak) – Signer 1

TABLE III.   ANALYZED OF FEATURES VECTORS OF DATA 568 (HARI RAYA AIDILFITRI) – SIGNER 1

	568-Hari Raya Aidilfitri			
	Left Hand Palm, X Axis	Left Hand Palm, Y Axis	Right Hand Palm, X Axis	Right Hand Palm, Y Axis
Min Point	-1	-1	-1	-1
Min at frame	56	94	54	54
Max Point	1	1	1	1
Max at frame	53	101	57	55
Median	0.0114814	0.0008768	-0.0062432	0.1718547
Mode	#N/A	#N/A	#N/A	#N/A
Slope Y to X	-0.026283845		0.48582739	
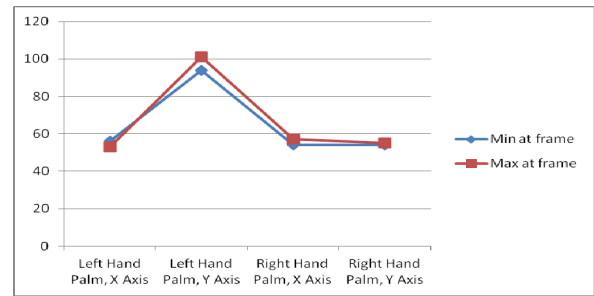
Figure 6.   Graph of Minimum and Maximum point (value) for each feature vectors of Data 568 (Hari Raya Aidilfitri) – Signer 1

IV.   RESULT AND DISCUSSION

Initial finding suggest that the GMT based feature vectors are able to show discriminate values and therefore can be used to differentiate between each sign language towards BIM recognition system. Further, the feature selection process was put in place to confirm that there are discriminate values for each sign language. For future improvement, we will continue with more data set to prove of its' robustness and stability.

REFERENCES

[1] Yasushi Hamada, Nobutaka Shimada, and Yoshiaki Shirai, "Hand shape estimation under complex backgrounds for sign language recognition," In Proceedings of the Sixth IEEE international conference on Automatic face and gesture recognition (FGR' 04). IEEE Computer Society, Washington, DC, USA, pg 589-594, 2004.

[2] Sazali, S. M., "Bahasa Isyarat Malaysia (first edition)," Persatuan Orang Pekak Malaysia (Malaysian Federation of Deaf), 2000.

[3] Sazali, S. M., "Antara BIM Dan Maruah Orang Pekak Malaysia", retrieved 25 July 2011, from http://www.epekak.net.my/bicara_sazali/2005/02_JBI.htm, 2005.

[4] Rini, A., Melanie, P.O. and Kuang, Y.C., "Real-Time Malaysian Sign Language Translation using Colour Segmentation and Neural Network," in Proc. Instrumentation and Measurement Technology Conference, 2007.

[5] Tan Tian Swee, Sh-Hussain Salleh, A.K. Ariff, Chee-Ming Ting, Siew Kean Seng and Leong Seng Huat, "Malay Sign Language Gesture Recognition System," International Conference on Intelligent and Advanced Systems 2007, pg 982-985, 2007.

[6] Starner T. and Pentland A., "Real-time American Sign Language recognition from video using hidden Markov models," Perceptual Computing Section Technical Report No. 375, MIT Media Lab, Cambridge, MA, 1996.

[7] Starner T. and Pentland A., "Visual recognition of American sign language using Hidden Markov models," In International Workshop on Automatic Face and Gesture Recognition, pg 189-194, 1995.

[8] Jong-Sung Kim, Won Jang, Zeungnam Bien, "A dynamic gesture recognition system for the Korean sign language (KSL)," In proceedings of Systems, Man and Cybernetics, Part B, IEEE Transactions on Volume 26, Issue 2, pg 354 – 359, April 1996.

[9] Mohandes, M.; A-Buraiky, S.; Halawani, T.; Al-Baiyat, S. , "Automation of the Arabic sign language recognition," In Proceeding of Information and Communication Technologies: From Theory to Applications, pg 479 – 480, 19-23 April 2004.

[10] T. S. Huang and A. Pentland, "Hand gesture modeling,analysis, and synthesis," Proceedings of Workshop on Automatic Face- and Gesture-Recognition, pp.73-79, Zurich, Switzerland, June 1995.

[11] Sylvie, C.W.O., Surendra, R.and Venkatesh, Y.V., "Understanding Gestures with Systematic Variations in Movement Dynamics," in Proc. Pattern Recognition, 2006.

[12] Amit Choudary, Rahul Rishi, Vijaypal Singh Dhaka,Savita Ahlawat, "Influnce of Introducing an Additional Hidden Layer on the Character Recognition Capability of a BP Neural Network having One Hidden Layer," in International Journal of Engineering and Technology Vol. 2(1), 2010, pg 24-28.

[13] Mlich, J. and Chmelar, P., "Trajectory Classification based on Hidden Markov Models," in Proc. International Conference on Computer Graphics and Vision, 2008.

[14] Lu, S., Tsechpenakis, G., Metaxas, D.N., Jensen, M.L. and Kruse, J., "Blob Analysis of the Head and Hands: A Method for Deception Detection," in Proc. Hawaii International Conference on System Sciences, 2005.

[15] Jung-Bae Kim, Kwang-Hyun Park, and Z. Zenn Bien, "Vision-Based Korean Manual Alphabet Recognition for the Hearing Impaired," Proceedings of the 8th International Conference on Rehabilitation Robotics, pp. 163 – 166, 2003.

[16] Wei Du and Hua Li, "Vision Based Gesture Recognition System with Single Camera," Proceedings of the 5th International Conference on Signal Processing, pp. 1351 – 1357 vol. 2, 2000.

[17] Yang, M.H., Ahuja, N. and Tabb, M., "Extraction of 2D Motion Trajectories and Its Application to Hand Gesture Recognition," in Proc. Trans. on Pattern Analysis and Machine Intelligence, 2002.

[18] Suzaimah Ramli, Mohd Marzuki Mustafa & Aini Hussain, "Feature Vector Selection Based on Freeman Chain COde for Shape Classification," Proc. of the M2USIC 2006, MMU International Symposium of Information and Communication Technologies, pp. 280-283, Kuala Lumpur, 16-17 November 2006.

New Hybrid Model Reference Adaptive Supervisory Fuzzy Logic Controller for Shell-and-Tube Heat Exchanger Temperature System

M.A. Ahmad, A.A. Ishak* and N.K. Ismail
*Fakulti Kejuruteraan Kimia
Universiti Teknologi Mara (UiTM) Shah Alam
40000 Shah Alam , Selangor Darul Ehsan
mohd.aizad.ahmad@gmail.com

Abstract—The control of the outlet temperature of a co-current shell-and-tube heat exchanger with the new hybrid model reference adaptive supervisory fuzzy controller (MRASFC) is presented in this paper. The outlet temperature of the cold fluid is controlled by manipulate the flow of the hot fluid while both temperature fluids keep constant. The shell-and-tube heat exchanger system is modeled mathematically and simulated using MATLAB/Simulink software. The Fuzzy Inference Structure (FIS) used is Sugeno-type. The normalized fuzzy controller using Gaussian membership function (MF) with 3x3 rule bases while hybrid MRAFC also using Gaussian MF with 3x3 rule bases. The performances on set point test of the fuzzy logic controllers with and without hybrid design are compared to PID controller based on maximum overshoot, settling time, number of oscillations and IAE. The results showed that MRASFC produced reduced overshoot, shorter settling time with less oscillation and minimum IAE compare to fuzzy and PID controller.

Keywords – shell-and- tube heat exchanger; set point test temperature control; fuzzy logic controller; model reference adaptive supervisory fuzzy controller;

I. INTRODUCTION

Shell-and-tube heat exchanger is widely used in industrial processes, either to heat or cool the process fluids [1]. It is because it can sustain wide range of temperature and pressure. The main purpose of a heat exchanger system is to transfer heat from hot fluid to cooler fluid.

Heat exchangers, which exhibit thermal process, are examples of nonlinear system, and difficult to achieve stability. In spite of nonlinearity problems, heat exchangers also exhibit significance time delay behavior. The time delays make the temperature loop hard to tune, and need to be tune manually if using conventional PID controller. Poor control using conventional PID controller lead to research of advanced controllers by many authors.

Fuzzy (FL) controller is one of the advanced controllers used for temperature control in different types of heat exchanger system. Many researchers have shown the effectiveness of hybrid FL controllers in their studies. Hybrid FL controllers are using another controller algorithm such as conventional PID, neural network, genetic algorithm and etc, combine with fuzzy logic controllers.

Hongquan et al [2] have introduced a new control scheme based on a Fuzzy-PI algorithm was to control the temperature control and check valve (TCCV) in a condensing heat exchanger (CHX) in manned module. There was a control threshold switching between PI and Fuzzy control. In this system, a continuous control approach to control the habitat temperature was needed by altering the amount of atmospheric flow through Temperature Control and Check Valve (TCCV) in the Condensing Heat Exchanger (CHX). The Fuzzy control of the TCCV is implemented as follows: Simple triangular membership functions were chosen in this design. The variable e and Δe has seven linguistic values. The output variable Δu has five linguistic values. There were 35 pieces of control rules. The proposed Fuzzy-PI control, were compared with those of PID and Fuzzy control. It can concluded that, high precision temperature control can be realized only with a Fuzzy controller where there were, minute oscillations for adopting incremental control. Fuzzy-PI controller system have minor overshoot, better resistance to varying parameters and better control precision.

In the same year, Maidi et al [3] have studied in at improving the control of a heat exchanger, described by a partial differential equation, by optimizing a linear proportional–integral, PI- fuzzy controller. The simulated system consists of a fluid flows through the tube of a shell-and-tube heat exchanger. The steam temperature on the shell side T_{st} (t) considered as a manipulated variable and the inlet temperature of the tube stream T (0, t) was considered as a disturbance. For the PI-FLC, the input variables, e and Δe, and output variable, T_{st}, were partitioned into five fuzzy sets using triangular membership function and have 25 rules. Simulations were performed on set point tracking and disturbance rejection. The obtained set point tracking responses was satisfactory for the linear PI-fuzzy logic controller compare to conventional PI. The fuzzy controller also rejects the boundary condition disturbance faster than the conventional PI controller.

Padhee and Singh in [4] have proposed an intelligent fuzzy PID controller to control the outlet temperature of a shell-and-tube heat exchanger. A typical interacting chemical process for heating consists of a chemical reactor and a shell-and-tube heat exchanger system has been studied. 49 fuzzy rule bases have been developed using triangular membership function.

Fuzzy PID controller has been compared with feedback PID and feedback-feed forward controller. The performance of controllers has been evaluated by two parameters which are maximum overshoot and settling time. The hybrid fuzzy PID controller in the feedback loop produced the less overshoot and fastest settling.

A new fuzzy-based predictive control (FPC) method within overall fuzzy-based predictive model (OFPM) based on the multiple models strategy was represented to control a specified tubular heat exchanger system, by Mazinan and Sadati in [5].The proposed control strategy identifies the best OFPM, at any instant of time, using a suitable performance criterion, which was realized by an intelligent decision mechanism (IDM). The purpose was to control the outlet water temperature that flows in the inner tube of this tubular heat exchanger system by means of the fluid flow pressure, manipulated using commanded valve pressure (P_v) on the shell tube of the system. The fuzzy predictive control based on multiple models strategy (FPCBMMS) has been compared to the single model linear generalized predictive control (SMLGPC) scheme. It was observed that the FPCBMMS performance is as good as the SMLGPC performance, when both the system coefficients and the desired set point were changed abruptly.

The work of Jain et al in [6] was discussed the comparative study of self tuning fuzzy controller with a standard Takagi-Sugeno (TS) fuzzy controller with application to a shell-and-tube heat exchanger. Online tuning of the membership functions and control rules of fuzzy controller was carried out using simulated annealing (SA) to obtain improved performance by minimizing the error function. The hot water outlet temperature was considered as the process or controlled variable and the cold water flow rate as the manipulated variable. The flow rate of hot water was treated as disturbance variable. The TS fuzzy controller was designed with two input variables; temperature, and error in temperature and one output variable, i.e. current, which was using the triangular function equation with 9x7 rule bases. The designed SA-FLC compared to TS-fuzzy controller for set point and load changes. The SA-FLC tracks the set points effectively and the settling time was smaller. The SA-FLC also produced minimum value of Integral Square Error (ISE) and Integral Absolute Error (IAE) error when sudden disturbance was introduced.

Another rule-based fuzzy PI control system has been designed, by Zhang et al in [7] which can self-adjust the values of the parameters, in a boiler temperature system. Fuzzy logic controller has two inputs and one output: the position error, change of the error and the temperature of the boiler. 7x7 rule base were used with triangular membership function in this controller. It can be seen that the fuzzy PI controller are better than the traditional PID controller. The controller was flexible because all parameters can be tuned by the fuzzy rules. The fuzzy PI controller has reached desired performance.

The preceding studies are summarized in table I.

TABLE I. TYPES OF HYBRID FUZZY CONTROLLER

Range of Control	Controller	Reference
20°C	FL-PI	Q.Hongquan et al (2008)
25°C to 50°C	FL-PI	A. Maidi et al (2008)
32°C to 52°C	FL-PID	S. Padhee and Y. Singh (2010)
40°C to 70°C	FPC	Mazinan and Sadati (2010)
40°C to 45°C	SA-FL	R. Jain et al (2011)
100°C to 180°C	FL-PI	W. Zhang et al (2011)

The present paper deals with the new design of a hybrid fuzzy controller, which combines the normalized fuzzy controller with model reference adaptive supervisory fuzzy controller (MRASFC). This controller applied to a co-current shell-and-tube heat exchanger system.

II. MODELING AND CONTROL

A. Shell and tube heat exchanger system

The system studied in this work corresponds to a water-water co-current shell and tube heat exchanger system as described on Figure 1. Water is the fluid that flows on both the sides of the shell and tube heat exchanger on the temperature system. A conventional electric water heater is used to heat the water. The flow rate of the hot water can be regulated by varying the speed of the hot water pump. The cold-water flow can be varied by changing the percentage opening for the flow control valve. The hot water flows in the tube side while the cold water flows in the shell side. The flow pattern is co-current, that means the hot water and cold water flows in same direction.
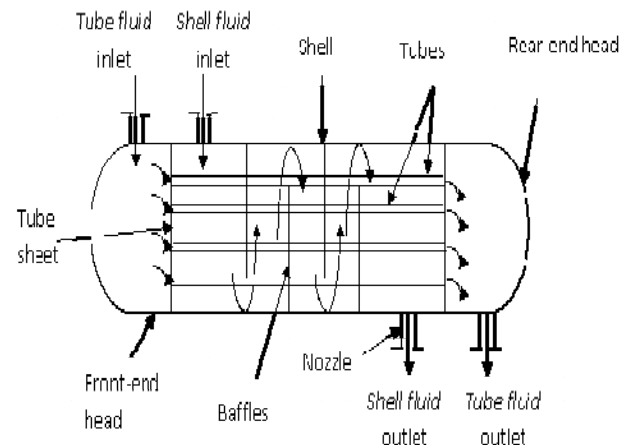
Figure 1. A shell-and-tube heat exchanger system

The control objective in this work is to control the outlet temperature of the cold water by manipulating the inlet flow rate of hot water. The inlet flow rate of cold water and inlet hot water temperature are disturbances in this process system. In this case study, the desired set point temperature is 38°C. The nominal inlet flow rate of cold water is 0.06 kilograms/seconds (kg/s), which is 50% of control valve opening position. The inlet temperature of cold water is set at 10°C, while the inlet temperature of hot water is set at 80°C. The hot water is used to heat cold water from 10°C to set point temperature at 38°C. The maximum flow rate of hot water is 0.076 kg/s. The physical parameters of heat exchanger are reported in table II.

TABLE II.　　PHYSICAL PARAMETERS OF HEAT EXCHANGER

Inputs	Value
Density of Fluid (water), ρ	1000 kg/m^3
Density of Metal (steel), ρ_m	8238 kg/m^3
Specific Heat Capacity, $C_{p, metal}$	468 J/kg^{o}C
Specific Heat Capacity, $C_{p,water}$	4230 J/kg^{o}C
Shell Heat Transfer Area, A_s	0.281 m^2
Tube Heat Transfer Area, A_t	0.253 m^2
Shell side volume, V_s	2.62 X 10^{-4} m^3
Volume of Metal, V_m	4.27 X 10^{-5} m^3
Tube side volume, V_t	1.43 X 10^{-4} m^3

B. Mathematical modeling

Shell and tube heat exchanger can be divided into three sections, which are shell, metal and tube. These sections are further divided into control volumes.

The following assumptions are made while designing the mathematical model:

- The control volumes are small and assumed to have a constant temperature.

- The heat exchanger is insulated and there is no heat loss from the heat exchanger to the surrounding.

A set of differential equations is used to calculate the outlet temperature on the shell and tube side for each control volume on the shell, metal and tube side.

1) Shell control volume (CV) energy balance: The energy storage rate in the shell control volume represented by S_0. As the fluid flows in and out of the control volume, the energy gained by the control volume due to the change in temperature of the fluid flow is give by $\dot{m}_s C_s(T_{si-1} - T_{si})$. This control volume also gain energy from process of convection due to temperature difference between the shell fluid and the metal which is $\frac{h_s A_s}{N}(T_{mi} - T_{si})$. Therefore, the equation for energy balance on the shell control volume is given by

$$\frac{(\rho CV)_s}{N} * \frac{dT_{si}}{dt} = \dot{m}_s C_s(T_{si-1} - T_{si}) + \frac{h_s A_s}{N}(T_{mi} - T_{si}). \quad (1)$$

2) Metal control volume (CV) energy balance: The energy storage rate in the metal control volume is given by $\frac{(\rho CV)_m}{N} * \frac{dT_{mi}}{dt}$. Heat energy is gained from the shell and the tube side is given by $\frac{h_s A_s}{N}(T_{si} - T_{mi})$ and $\frac{h_t A_t}{N}(T_{ti} - T_{mi})$, while heat energy is also gained by conduction within metal, given by $\frac{k A_m}{L}(T_{mj} - T_{mi})$. The energy gain by conduction is ignored in this calculation because temperature difference between the two metal nodes is very small and negligible. Therefore, the equation for energy balance on the metal control volume is given by

$$\frac{(\rho CV)_m}{N} * \frac{dT_{mi}}{dt} = \frac{h_s A_s}{N}(T_{si} - T_{mi}) + \frac{h_t A_t}{N}(T_{ti} - T_{mi}). \quad (2)$$

3) Tube control volume (CV) energy balance: The energy storage rate in the tube control volume represented by T_0. As the fluid flows in and out of the control volume, the energy gained by the control volume due to the change in temperature of the fluid flow is given by $\dot{m}_t C_t(T_{ti-1} - T_{ti})$. This control volume also gains energy from process of convection due to temperature difference between the tube fluid and the metal which is $\frac{h_t A_t}{N}(T_{mi} - T_{ti})$.

Therefore, the equation for energy balance on the tube control volume is given by

$$\frac{(\rho CV)_t}{N} * \frac{dT_{ti}}{dt} = \dot{m}_t C_t(T_{ti-1} - T_{ti}) + \frac{h_t A_t}{N}(T_{mi} - T_{ti}). \quad (3)$$

C. PID controller

In this work, digital PID controller in velocity form is used which calculate the change in controller output, Δp_k.

$$\Delta p_k = p_k - p_{k-1}$$

$$= K_c \left[(e_k - e_{k-1}) + \frac{\Delta t}{\tau_I} e_k + \frac{\tau_D}{\Delta t}(e_k - 2e_{k-1} + e_{k-2}) \right]. \quad (4)$$

PID controller tuning is using trial and error method. The controller parameter values are controller gain, $K_c = 0.00003$, $1/\tau_I = 12$ and $\tau_D = 0.1$.

III. IMPLEMENTATION OF MODEL REFERENCE ADAPTIVE SUPERVISORY FUZZY CONTROLLER

A. Design of normalized fuzzy controller

A normalized fuzzy controller with adjustable input and output scaling factor is proposed. The design of the fuzzy logic controller begins with the selection on input and output. The fuzzy logic controller consists of two inputs and one output. The inputs are error (e) which is the difference between temperature set point (SP) and the actual temperature (PV), and change in error (Δe) which is the difference between the current error e (t) and the error of the previous iteration e (t-1). The output of the fuzzy logic controller (Δu) is proportional to input mass flow rate of hot water, denoted as $\dot{m}_s$. The following equation describe e and Δe

$$e \ (t) = SP - PV. \quad (5)$$

$$\Delta e = e \ (t) - e \ (t-1). \quad (6)$$

The normalized fuzzy controller designed has the following parameters:

- The membership function (MF) used is Gaussian membership function as shown in Figure 2.

- The universe of discourse of both inputs signal is

  [-1 1] while output signal is [-2 2]

- Three MF used for each input, detailed in Table III.

- There are nine rules base, summarized in Table IV.

- Fuzzy inference system is Sugeno.

- Fuzzy inference method are product for AND, max for OR, and weighted average for defuzzification.

- Four scaling gains used which are GE, GCE, GIE and GU. GE =1;GCE = 0.1; GIE = 12 and GU = 0.0003

- Figure 3 show the structure of block diagram for fuzzy controller design.
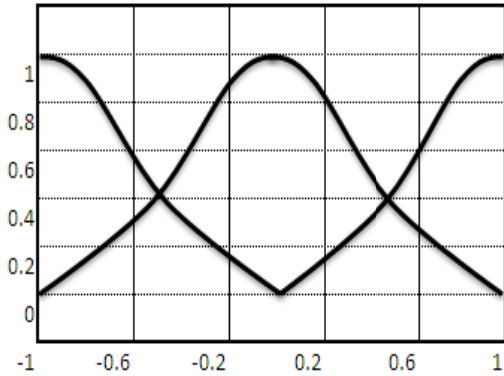
Figure 2.   Gaussian membership function

TABLE III.   LINGUISTIC VARIABLES FOR FUZZY CONTROLLER

Error, e (t)		Change in Error, Δe		Controller Output, Δu	
N	Negative	N	Negative	N	Negative
Z	Zero	Z	Zero	Z	Zero
P	Positive	P	Positive	P	Positive

TABLE IV.   IF-THEN RULE BASE FOR FUZZY CONTROLLER

Controller Output, Δu		Change in Error, Δe		
		N	Z	P
Error, e (t)	N	P	P	Z
	Z	P	Z	N
	P	Z	N	N
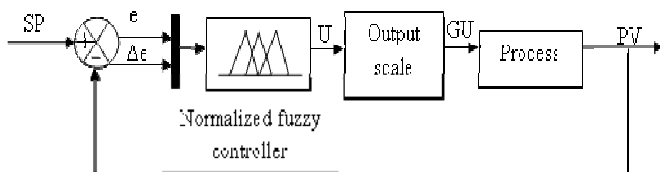
Figure 3.   Block digram of fuzzy controller

B. Design of hybrid model reference adaptive supervisory fuzzy controller

The hybrid model reference adaptive supervisory fuzzy controller has been designed with two inputs variable, error, e_r (t) and change of error, Δe_r and one output, Δu_r. e_r is the difference between model reference desired set point, SP_r and actual temperature. Δe_r is the difference between current error e_r (t) and the error of the previous iteration e_r (t-1). Δu_r is the output from model reference adaptive supervisory fuzzy controller. Figure 4 shows the block diagram of hybrid MRASF model.

Detailed design of this controller has the same structure as normalized fuzzy controller, except the scaling gain parameters. Three scaling gains used, which were GE2, GCE2 and GU2. GE2 = 1; GCE2 = 0.001 and GU2 = 0.0001.
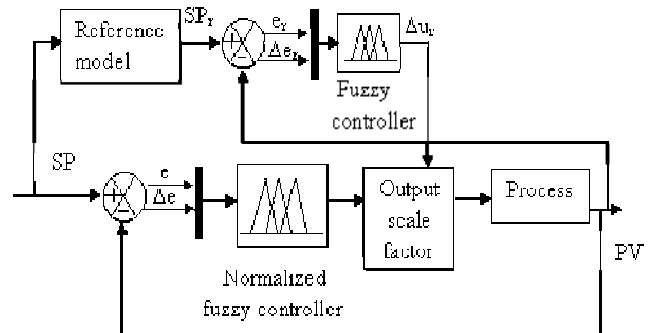
Figure 4.   Block diagram of MRASFC

A reference model has been developed to force the plant follow the required response (overshoot, rise time, etc.). The desired specifications of the system designed are maximum overshoot which is 2% and settling time, which is 70 seconds. The desired response which achieves the desired specification is described by equation:

$$SP_r(t)=A*[1- (1.597\, e^{-0.12t} \sin (0.0964t+38.76*\pi/180))]. \quad (6)$$

where A is step required, ξ is damping ratio, ω_n is natural frequency. This equation based on general equation:

$$1 - \frac{1}{\sqrt{1-\xi^2}} e^{-\xi \omega_n t} \; \sin(w_n t \sqrt{1-\xi^2} + \theta) \quad (7)$$

where $\theta = \cos^{-1} \xi$,

IV.   RESULTS AND DISCUSSIONS

The performances of controllers were tested for set point change from 43°C to 38°C, 38°C to 33°C and from 43°C to 33°C. Evaluation criteria are using maximum undershoots, number of oscillations, settling time and integral absolute error (IAE).
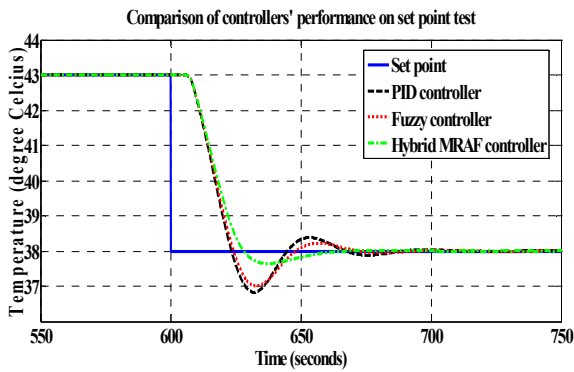
Figure 5.   Process response for set point test (43°C to 38°C)

Figure 5 shows the performances of each controller for set point change from 43°C to 38°C. The outlet temperature using PID controller resulting one oscillation, having 3.08% of maximum undershoots before settled in 67 seconds. The response using normalized fuzzy controller also have slightly similar trend to the PID controller, settled in 72 seconds. The proposed hybrid fuzzy controller performs better than PID and normalized fuzzy controller. The MRASFC produced half oscillation and only 0.95% undershoots and faster settling time.

TABLE V.   COMPARISON OF DIFFERENT PARAMETERS IN CONTROLLERS FOR SET POINT CHANGE FROM 43°C TO 38°C

Type of Controller	Maximum Undershoots	Settling Time,s	No. of Oscillations	IAE
PID	3.08%	67	1	103
Fuzzy	2.63%	72	1	100.2
MRASF	0.95%	63	0.5	93.4

Table V above summarized the results of each controller with all those criterion performances. The most improved performance criteria is the maximum undershoots, where there are 70% reduced using proposed hybrid controller.
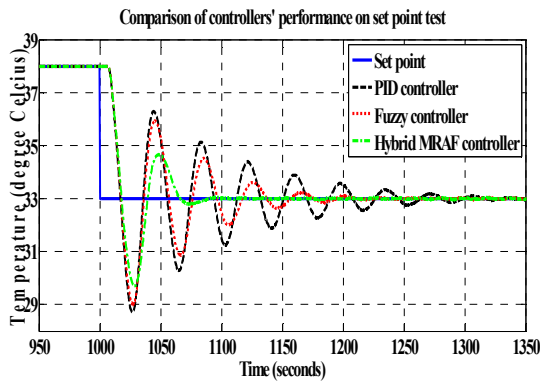
Figure 6.   Process response for set point test (38°C to 33°C)

Figure 6 shows the performances of each controller for set point change from 43°C to 38°C. The outlet temperature using PID controller resulting six oscillations, having 13% of maximum undershoots before settled in 250 seconds. The response using normalized fuzzy controller have better performance compare PID controller. It oscillates four times, a bit reduced undershoots and settled in 176 seconds. The

proposed hybrid fuzzy controller performs better than PID and normalized fuzzy controller. The MRASFC produced only one oscillation, have 10% maximum undershoots and settling time of only 90 seconds.

TABLE VI.   COMPARISON OF DIFFERENT PARAMETERS IN CONTROLLERS FOR SET POINT CHANGE FROM 38°C TO 33°C

Type of Controller	Maximum Undershoots	Settling Time,s	No. of Oscillations	IAE
PID	13%	250	6	313.1
Fuzzy	12.12%	176	4	233.4
MRASF	10%	90	1	147.3

Table VI summarized the results of each controller with all those criterion performances. The proposed hybrid MRASFC achieved improvement about 57% for settling time compare to PID controller. It also made 53% reduced of IAE value, which is good for energy cost and efficiency.
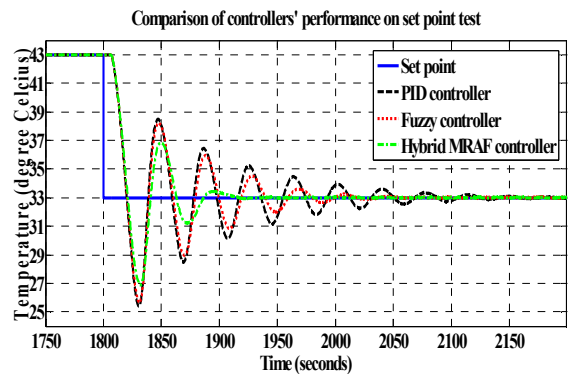
Figure 7.   Process response for set point test (43°C to 33°C)

Figure 7 shows the performances of each controller for set point change from 43°C to 33°C. The outlet temperature using PID controller resulting seven oscillations, having 23% of maximum undershoots before settled in 280 seconds. The response using normalized fuzzy controller have better performance compare PID controller. It oscillates four times, a bit reduced undershoots and settled in 200 seconds. The proposed hybrid fuzzy controller performs better than PID and normalized fuzzy controller. The MRASFC produced only two oscillations, have 18.5% maximum undershoots and settling time of only 116 seconds.

TABLE VII.   COMPARISON OF DIFFERENT PARAMETERS IN CONTROLLERS FOR SET POIN CHANGE FROM 43°C TO 33°C

Type of Controller	Maximum Undershoots	Settling Time,s	No. of Oscillations	IAE
PID	23%	280	11	568
Fuzzy	22.3%	200	4	473.6
MRASF	18.5%	116	2	324.4

Table VII summarized the results of each controller with all those criterion performances. The proposed hybrid MRASFC achieved improvement about 54% for settling time compare to PID controller. It also made 43% reduced of IAE value, which is good for energy cost and efficiency

V. CONCLUSIONS

The hybrid model reference adaptive supervisory fuzzy controller (MRASFC) is applied to a co-current shell-and-tube heat exchanger system for different set point values. The performance of proposed hybrid fuzzy controller is studied and compared to normalized fuzzy controller and PID

Simulation results show that this new hybrid MRAFC has better performance in term of percentage undershoots, settling time, number of oscillations and IAE.

The results also indicate that the process has better controller performance in higher set point temperature value, which produce reduced undershoots and less oscillations compare to lower set point value. i.e. 48^oC to 38^oC compare to 38^oC to 33^oC.

The energy consumed also reduced while temperature change from 43^oC to 33^oC in two step rather than directly change in one step. It was shown by IAE values at these two different conditions.

However, proposed hybrid MRASFC can't reduce the maximum undershoots significantly and another new design approach needed to suppress those undershoots especially when subjected to lower set point value i.e. set point change to 33^oC.

Further studies recommended on using this new design approach for wider range of temperature control. It is also advised to check its controllability on subjected load disturbances and compare with PID and fuzzy controller. Another future work is to implement this new design for real plant of shell-and-tube heat exchanger temperature system.

REFERENCES

[1] R.K. Shah, "Industrial heat exchangers—functions and types, in Industrial Heat Exchangers," J-M. Buchlin, ed., von Ka´rma´n Institute for Fluid Dynamics,Belgium, Lecture Series No. 1991-04, 1991.

[2] Q. Hongquan and P. Li- ping, "Fuzzy-PI algorithm for manned module thermal control," Proceedings of the 7th World Congress on Intelligent Control and Automation, Chongqing, China, 978-1-4244-2114-5, 2008.

[3] A .Maidi, M.Diaf, and J.Corriou, "Optimal linear PI fuzzy controller design of a heat exchanger," Journal of Chemical Engineering and Processing, vol. 47 , pp. 938–945, 2008.

[4] S.Padhee and Y.Singh, "A comparative analysis of various control strategies implemented on heat exchanger system: A case study," Proceedings of the World Congress on Engineering, London, U.K, 978-988-18210-7-2, vol. II, 2010.

[5] D.Zhenjun, F.L.S. Zhanbiao and Z.Yameng, "Design and simulation of fuzzy control system for water temperature of heat exchanger," Journal of IEEE, 978-1-4244-5586-7, vol. 4, 2010.

[6] R.Jain, N.Sivakumaran and T.K. Radhakrishnan, "Design of self tuning fuzzy controllers for nonlinear systems," Expert Systems with Applications, vol.38, pp. 4466–4476, 2011.

[7] W.Zhang, X.Li, L.Li, J.Lv, Y.Chen and X.Mao,"Design and application of fuzzy controller," Key Engineering Materials, vol. 464, pp. 107-110, 2011.

Designing Cavity-Enclosed Microstrip Antenna for Gain Enhancement

N.H.Khairul Anuar[1], M.T.Ali[2], S.H.Anuar[3], S. Subahri[4] and N. Ya'acob[5]

Microwave Technology Center (MTC)

Faculty of Electrical Engineering

University Technology Mara (UiTM)

40450, Shah Alam, Selangor, MALAYSIA

[1]hafizah.anuar@gmail.com, [2]mizi732002@yahoo.com, [3]hajaranuar@gmail.com, [4]eilasuhaila@yahoo.com and [5]norsuzilayaacob@yahoo.com

Abstract — **The antenna has been designed to observe and analyze the effect of the addition of perfect electric conductive (PEC) cavity-enclosed to a rectangular microstrip antenna. The calculated parameters have been optimized using Computer Simulation Technology (CST) software. The design prototype has been fabricated and the output was measured using Vector Network Analyzer (VNA). The simulated and measured results are then compared. It is observed that the gain is improved when PEC cavity is added to the microstrip antenna.**

Keywords: Cavity-enclosed; microstrip antenna; gain; return loss; CST.

I. INTRODUCTION

The designing of various applications by using microstrip antenna has becomes high in demand due to its advantages of low weight, small volume, and ease of fabrication which suit personal and mobile communications applications today [1]. Regardless of the many benefits, microstrip antennas still suffered some drawbacks such as narrow bandwidth, lower gain, and low power handling capability [2]. In fact, David Pozar [3] states that the impedance bandwidth of traditional microstrip antenna configuration is only a few percent compared to other types of antenna for instance dipole, slot and waveguide horn where the bandwidth produced is from 15% up to 50%. There are various gain enhancements techniques that can be applied to an antenna such as by adding a high permittivity superstrate-substrate to an existing microstrip antenna. It is applied to the upper layer of the microstrip antenna at an optimized height as conducted in [4, 5].

There are only few researches that have studied the effect of cavity to microstrip antenna. Nemai Chandra [6] states that cavity theory could exhibit high order resonance where it could eliminate the surface waves and mutual coupling in thick substrates which is suitable to be applied in an array environment. S.Noghanian and L.Shafai [7] have also stated that there is an increased of 3 dB gain when PEC cavity is added to their designed antenna. This addition suppressed the back lobe radiation of aperture feeding and resulted to the gain enhancements. In addition, Weihua Tan et al. [8] have applied

cavity-backed slot and a high permittivity superstrate-substrate technique (same technique as applied in [4, 5]) to their designed microstrip antenna and it is proven that the addition of cavity have resulted to an increase in gain up to 7 dB to their designed antenna.

Based on the observation these past papers, mostly has applied a circular PEC cavity at the bottom of the ground plane. Thus in this paper, a different design of cavity addition is introduced where a rectangular PEC cavity frame is applied on top of the substrate as shown in figure 4 instead of a circular cavity at the bottom of the ground plane. This project is carried out to observe the effect of gain improvement to the microstrip antenna when PEC cavity is added.

II. ANTENNA GEOMETRY AND THEORETICAL FORMULATION

A. Rectangular patch

The designed antenna is based on the standard rectangular patches with operating frequency at 2.5 GHz. The substrate used is Flame Retardant 4 (FR-4) with a relative dielectric constant, ε_r of 4.7, height, h of 1.6mm and tangent loss, $tan\delta$ of 0.022.
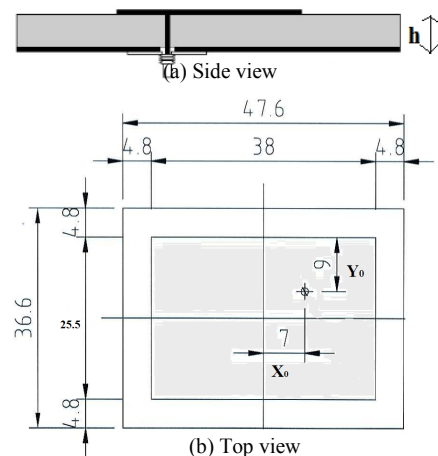
Figure 1. The designed rectangular microstrip antenna.

Figure 1(a) shows the side view of the designed rectangular microstrip antenna with coaxial probe feeding technique. Figure 1(b) shows the dimensions of the antenna from the top view. The value of width and length of the rectangular shape of the radiating patch is calculated using (1) until (7). The effective relative permittivity or effective dielectric constant, εreff is as shown in (1) while the width, W for f_r = 2.5 GHz is as shown in (2) where c is the light velocity. The length, L can be obtained by using (3)-(5).

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{W}\right]^{-\frac{1}{2}} \tag{1}$$

$$W = \frac{v_0}{2f_r}\sqrt{\frac{2}{\varepsilon_r + 1}} \tag{2}$$

$$L = L_{eff} - 2\Delta L \tag{3}$$

$$L_{eff} = \frac{c}{2f_r\sqrt{\varepsilon_{eff}}} \tag{4}$$

$$\frac{\Delta L}{h} = 0.412\left[\frac{(\varepsilon_{eff} + 0.3)(\frac{W}{h} + 0.264)}{(\varepsilon_{eff} - 0.258)(\frac{W}{h} + 0.8)}\right] \tag{5}$$

The ground size and substrate size as well were used following (6) and (7);

$$L_g = 6h + L \tag{6}$$

$$W_g = 6h + W \tag{7}$$

The feeding used is coaxial probe which is located at the point of $y_0 = \frac{W}{2}$ (in linear polarization) and $X_0 = \frac{L}{2}$ from the patch center as the point of origin. The parameters for patch are presented in table 1. The calculated values are obtained from equations and the simulated values are obtained via optimization from CST software. There are slight difference in each value due to optimization that is essential during simulation in order to obtain good return loss and resonance frequency.

TABLE I.   CALCULATED AND SIMULATED PARAMETER VALUE OF DESIGNED ANTENNA

Antenna	Calculated value (mm)	Optimized simulated value (mm)
Patch Width, W	35.54	38
Patch Length, L	23.71	25.5
Feed Point (Y_0, X_0)	(8.89,5.93)	(9.0,7.0)

B. Cavity-Enclosed Effect

A perfect electric conductor (PEC) cavity enclosed was stated in [10] and [11] which resulted to the shifting in the resonance frequency, suppressed of the surrounding surface wave which is best for array applications. This also reduces the antenna backward radiation. The suppression of backward radiation gives advantage to the increasing main lobe radiation as well as gain of the microstrip antenna. Based on graph in figure 2, the best height, h_c of the PEC cavity was at h_c=3.5mm which gives a maximum peak of gain value of up to 6.246 dB.
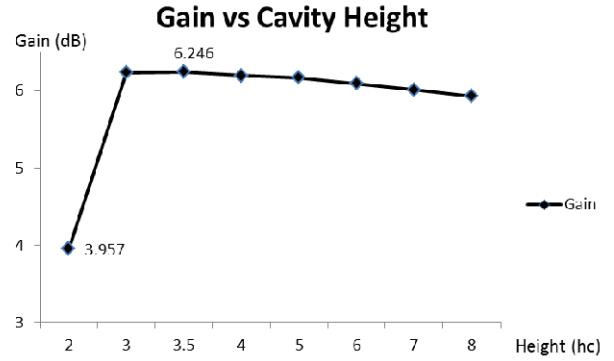
Figure 2.   Microstrip antenna gain versus cavity-enclosed height.
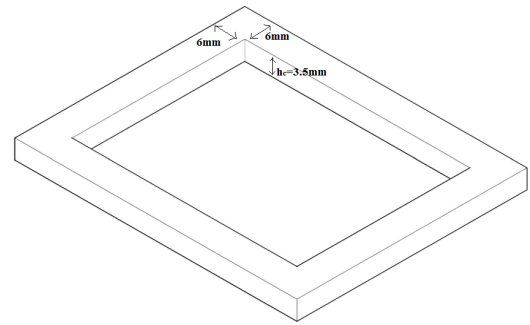
Figure 3.   Dimension of the PEC cavity.

Figure 3 illustrated the dimension of the designed PEC cavity which is enclosed around the rectangular microstrip antenna as illustrated in figure 1 previously. The material used for PEC cavity is the conductive stainless steels. Figure 4 shows the fabricated rectangular microstrip antenna with the PEC cavity applied on top of it.

Figure 4.   The fabricated microstrip antenna with PEC cavity applied.

III. RESULT AND DISCUSSION

The results were compared between the simulated result (obtained from simulation in CST software) and measured result (obtained from the fabricated antenna by using VNA in lab). Figure 5 shows the simulated and measured return loss of the antenna without PEC cavity. In simulation, the value of S_{11} is at -23 dB operating at 2.499 GHz. By measurement, the resonance is at 2.525 GHz with the value of S_{11} at -12dB.
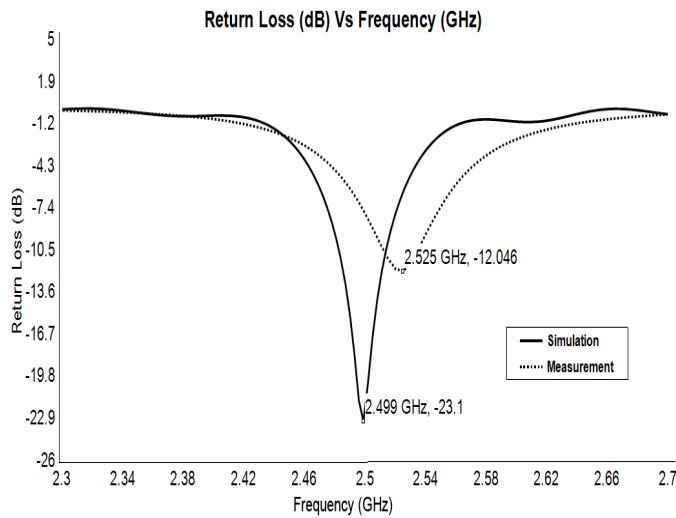
Figure 5.   Simulated and measured return loss value for antenna without cavity applied.
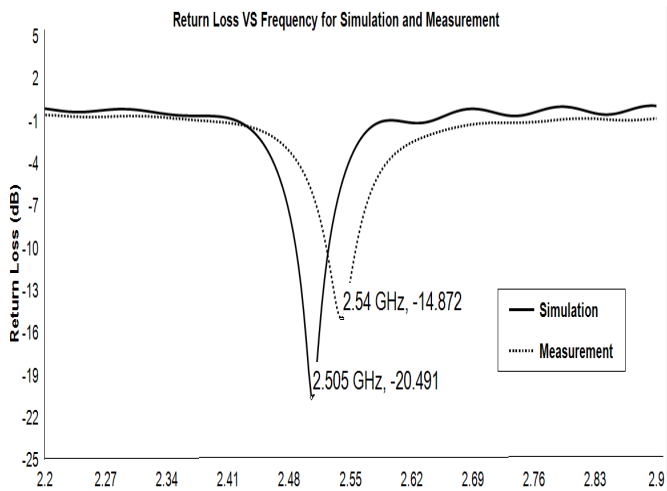
Figure 6.   Simulated and measured return loss value for antenna with cavity applied.

Figure 6 shows the simulated and measured value of return loss with PEC cavity applied to the microstrip antenna. In simulation, S_{11} is at -20 dB operating at 2.505 GHz while there is slight difference in measured value of S_{11} of -14dB resonating at 2.54GHz frequency. The values of simulated and measured result have slightly shifted which could be caused by equipment, human and substrate material (FR-4) loss during the fabrication process. The best material for substrate to be used is the one with the lowest relative dielectric constant, ε_r.
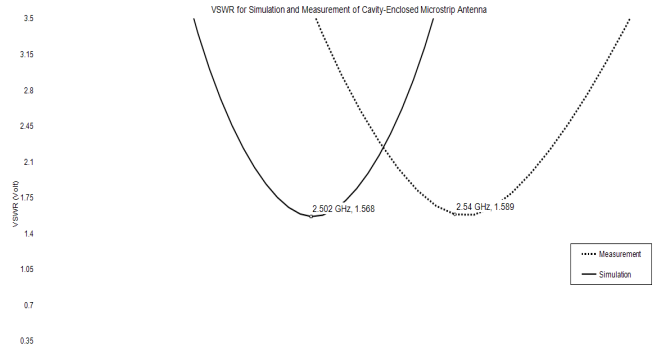
Figure 7.   Simulated and measured VSWR with cavity applied.

Figure 7 shows the simulated and measured VSWR of the microstrip antenna with PEC cavity applied. The result obtained for VSWR is 1.5 for both simulation and measurement.

Figure 8(a) shows the radiation pattern in 3D views for rectangular patch without applied PEC cavity while figure 8(b) shows the radiation pattern in 3D views with installed PEC cavity.
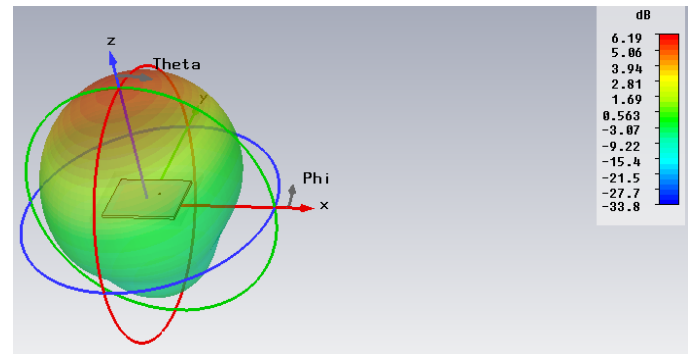
(a) Without PEC cavity
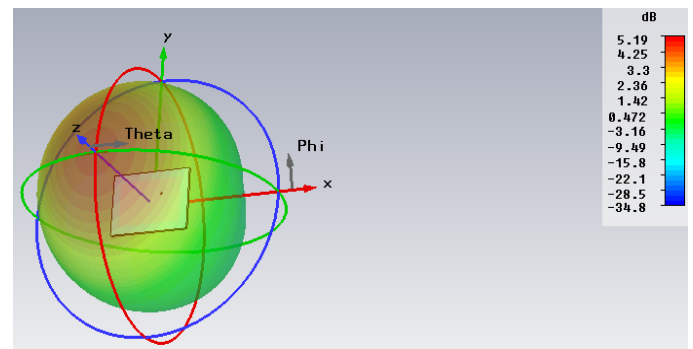
(b) With PEC cavity

Figure 8.   Simulated radiation pattern.

Figure 9 shows the comparison of simulated polar pattern and gain value for both with and without the application of PEC cavity respectively. From the pattern, without cavity (dashed line) resulted to the gain of 5.2 dB while after adding PEC cavity with height of h_c = 3.5mm, (solid line) the gain has increased up to 1 dB. The back lobe pattern also reduced (solid

line) due to the PEC cavity installments which effects to the gain increment. The gain obtained after the addition of cavity is 6.2dB.
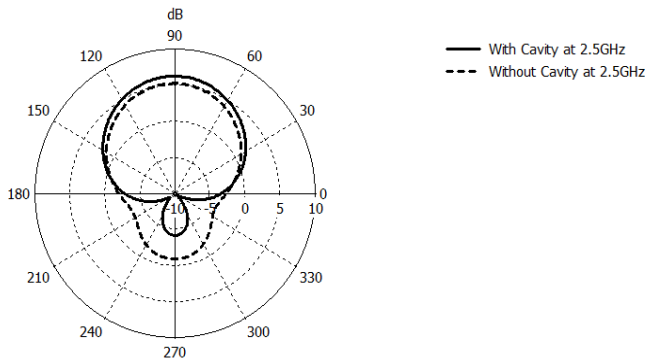
Figure 9. Comparison of simulated polar pattern and gain value of microstrip antenna.

IV. CONCLUSION

A new technique of microstrip antenna gain enhancement by using PEC cavity enclosed was introduced in this paper. The best height of cavity structure which gives maximum gain was determined at h_c = 3.5mm with a gain increment of up to 1dB. In simulation, the operating frequency is set to 2.5 GHz while after fabrication; the measured operating frequency has slightly shifted to a frequency of 2.54 GHz. There are some limitations during the fabrications which may be caused by human error, equipment error and also material loss. In order to optimize the material loss, the best choice of substrate material is by using the low dielectric constant substrate such as Duroid with ε_r of 2.2 and air with ε_r of 1.0. The design can be applied for various applications in wireless communication systems due to the compact size, light weight and also high gain.

ACKNOWLEDGEMENT

REFERENCES

[1] Tomasz Rutkowski, Custodio Peixeiro, " *Dual-band omnidirectional microstrip patch array antenna for a mobile communication system*". Asia Pacific Microwave Conference,1997, in press.

[2] Girish Kumar, K.P. Ray, "*Broadband Microstrip Antenna*", ARTECH HOUSE INC. (2003) ISBN 1-58053-244-6. Chapter 1(pp1-8), Chapter 4(pp131-167). (references)

[3] David M.Pozar. "*A review of bandwidth enhancements techniques for microstrip antennas*" google online books [available at]: http://books.google.com.my/books : 01April2012 .(references)

[4] Chisang You, Manos M. Tentzeris, Woonbong Hwang, "*Multilayer effects on Microstrip Antennas for their integration with mechanical structures*". IEEE transactions on antennas and propagation, vol. 55, no 4, April 2004, in press.

[5] David Jackson and N.Alexopoulos, "*Gain enhancement methods for printed circuit antenna*" IEEE Trans. Antenna and Propagation, vol.AP-33, No9, pp976-987, 1985, in press.

[6] Nemai Chandra. "*Investigation into a cavity-backed circular-patch antenna*"IEEE transactions on antennas and propagations, vol 50, no 12 December 2002, in press.

[7] S.Noghanian and L.Shafai, "*Control of microstrip antenna radiation characteristics by ground plane size and shape*" proc. Inst. Elect. Eng. Microwaves Antennas Propagation, vol145, no3. Pp207-212, June 1998, in press.

[8] Weihua Tan, Zhongxiang Shen and Mayuk Fujise, "*A Gain Enhanced Microstrip-Fed Cavity-Backed Slot Antenna*", IEEE APMC 2005 Proceedings, in press.

[9] Constantine Balanis. "*Antenna Theory Analysis and Design*", Third Edition. John Willey and Sons Publication. Chapter 14. Pp 816-843 (references)

[10] Debatosh Guha and Jawad Y.Siddiqui. "*Effect of a cavity enclosure on the resonant frequency of inverted microstrip circular patch antenna*". IEEE Transactions on Antenna and Propagation, vol 52, no 8, August 2004, in press.

[11] Manotosh Biswas and Jawas Siddiqui. " *Effect of a cyclindrical cavity on the resonance of a circular microstrip patch with variable air-gap*". IEEE Aqnhtennas and Wireless Propagation Letters, Vol5, 2006, in press.

Computer Security Self-Efficacy Effect

an Extention of Technology-to-Performance Chain Model

Mahmoud Al-Shawabkeh
International Islamic University Malaysia (IIUM)
Kuala Lumpur, Malaysia

Madihah Mohd Saudi
Najwa Hayaati Mohd Alwi
Universiti Sains Islam Malaysia (USIM)
Nilai, Malaysia

Abstract— given a paucity of research and apparent lack of coherence in information system research, it seems that there is no consensus in the information system field as to how security fits into the information system acceptance, usage, success, utilization, and/or performance impact (effectiveness, efficiency, and satisfaction). This paper is part of an ongoing research project designed to extend the Technology-to-Performance Chain (TPC) model by including the Computer Security Self-Efficacy (CSSE) construct, a strategy of model extension suggested by several researchers. This project aims to examine the research conducted in the last decade in information system journals regarding security issues then based on social cognitive theory, to propose a construct to measure individuals' computer security self-efficacy. Based on the Technology-to-Performance Chain (TPC) model, this study design expected to models and tests relationship among computer security self-efficacy and secures online banking system performance impact. The study will try to answer the question "to what extent has the computer security self-efficacy affected user's perception of secure online banking system effectiveness". After this research finished, the researcher assume that this study findings will provides an initial step towards understanding of the applicability of social cognitive theory in information system security domain and helps information security professionals design information systems considering the effect of computer security self-efficacy on secure information system.

Keywords- computer security self-efficacy; self-efficacy; task-technology fit; information system; social cognitive theory; online banking; utilization; effectiveness; performance impact.

I. INTRODUCTION

Over the years, the banking industry has been using a variety of information systems technology to provide their services to customers [1], [2], [3], [4]. Today, these technologies form a consortium of banking information systems and enable services such as online banking, mobile banking and online credit card management. Some researchers have observed that while banking information systems are readily available, they are not fully utilized by some customers as anticipated [5], [1], [6], [4]. A survey worth mentioning is the one by Thorton Consulting, which concluded that the perceived lack of security is one of the main obstacles of growth in the number of internet banking users [8]. Furthermore, according to Howcroft [9] the perceived security is one of the principal factors that inhibit online banking adoption. Hence, there is a heightened need in information system research to understand the factors that impact an individual's decision to use such technologies [9], [10], [11], [12].

Generally, elements of computer security research focus on the technology itself, such as the malware and spyware [13],[14],[15],[16],[17], semantic web [18], internet voting [20], mobile and internet banking [20], cloud computing security [21]. However, more current studies seek to understand the effects and perceptions of using the computer security technology. There is not much literature on security factor effects in information system research [22]. The main reason is the difficulty experienced in conducting this type of research [23] due to the sensitive nature of collecting data on secure tasks done by the users using this type of secure information systems [24].

Attention has been more focusing on technology-specific antecedents that may provide significantly stronger guidance for the successful design and implementation of specific types of information system. In recent years, researchers have noted that developing theory that is more focused and context specific (technology specific) is considered an important frontier for advances in information system research [25], [26], [12]. As well as others identified, there are several important directions for future research with suggestions that one of the most important directions for future research is to extent the information system stream of research into other established streams of work [27], [26], [28].

A model that integrates knowledge from computer security research and information system effectiveness research is lacking, a gap which this research seeks to fill by integrating the information system stream with another dominant research stream [12]. The purpose of this study is to address this gap specifically to find "to what extent has the computer security self-efficacy affected user's perception of secure online banking system effectiveness ". The research findings attempt to answer the three main questions, which are:

1. What are the factors that measure a Computer Security Self-Efficacy construct?
2. What are the main constructs affecting the secured Information System utilization and effectiveness?
3. How computer Security Self-Efficacy affects the secure Information System utilization and effectiveness?

Several objectives of this research have been identified. They are as follows:

- To study the importance and impact of Computer Security Self-Efficacy.
- To find the most important constructs that affects the Information System utilization and effectiveness.
- To measure the relationship between the Computer Security Self-Efficacy and Information System utilization, as well as, to measure the relationship between Computer Security Self-Efficacy and Information System effectiveness.

The paper is organized in the following manner: introduction, literature review, research framework, research methodology and conclusion. This section, the introduction has introduced the importance for this research, the research questions, and the research objectives. The following literature review section reviews the information system effectiveness, utilization, technology-to-performance chain model, and social cognitive theory, together with the concepts of self-efficacy and computer self-efficacy in the information system research. Next is a description of the proposed research framework with its hypotheses and constructs definition. The research methodology is then presented, and finally, the conclusion.

II. LITERATURE REVIEW

A. Information System Effectiveness

The information system research outcomes that researchers have looked for can be categorized as either pertaining to performance or fit. Studies done to see if the information system technology use leads to improved outcomes, or motivation, or changes users' behavior are considered performance oriented. Studies which gather and analyze users' perception of how well a technology will help an individual to complete a specific task or set of tasks is defined as fit oriented.

Task-Technology Fit is a key, but Fit is often overlooked as a construct in understanding the impact of technology on individual performance [41]. The Fit is a way to measure the performance impact (effectiveness) of information systems [29]. Information system effectiveness can be difficult to measure; thus, the user evaluations are commonly used as the measurement. User evaluations based on the fit between task and technology have been an effective measures of information systems effectiveness. While researchers have carried out several studies on fit effect on effectiveness of information system, there is still room for further research in assessing fit and in how to best measure information systems effectiveness [41].

B. TInformation System Utilization

Utilization of a system has been used as a dependent variable. It has been modeled as an outcome construct that can be influenced by the process of design and implementation and by characteristics of the information system, the task, the individuals and their interactions. To measure the information

system utilization, some researchers adopt user satisfaction model [30], [31], [32], while others measure the effectiveness of system on a basis of the DeLone and McLean information system success model [33], [34], [35], [36], [37], [38]. Additionally, some others explore the personality trait of self-monitoring in the information system usage by employing the extended theory of planned behavior (TPB) model of Pavlou and Fygenson [39], [40].

Utilization research is one of two major research streams linking technology to performance impact. The utilization research stream is based on theories of user attitudes, beliefs, and behaviors. The implication of the utilization research stream is that increased utilization will lead to better performance impact (see Fig. 1). The other research stream is the Task-Technology Fit model, as shown in Fig. 2. The Task-Technology Fit model indicates that performance will be increased when a technology provides features and support that Fit the requirements of the task [41], [42].

C. Technology-to-Performance Chain Model

Goodhue and Thompson [42] have proposed a model that combines both the utilization model and the Task-Technology Fit model into one model, called the Technology-to-Performance Chain model (TCP) Fig. 3. The Technology-to-Performance Chain model recognizes that technology must be utilized and Fit the task it supports in order to have a performance impact. The Technology-to-Performance Chain model gives a more accurate picture of the way in which technologies, user tasks, and utilization relate to affect a performance impact. According to Goodhue and Thompson, in order for an information system to have a positive performance impact, first, the technology must be utilized, and second, there must be a positive Task-Technology Fit. If either the Task-Technology Fit or utilization is lacking, the technology will not improve the performance impact. This model has been formally recognized in several studies [43], [44], [39], [45] [46]. one of the research finding on Technology-to-Performance Chain model was that for information system users the Task-Technology Fit is more important than the user interface [47]. Other research combined the technology acceptance model (TAM) with the Task-Technology Fit model to study the information system adoption [48], [49]. The technology acceptance model is widely considered in information system research, but will not be considered as the base model for this research because of the lack of task focus, a weakness of the technology acceptance model for understanding utilization and effectiveness is the lack of task focus.
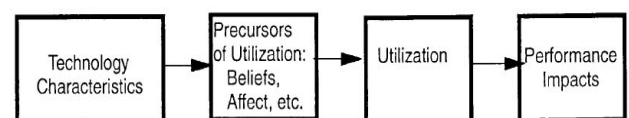
Figure 1. Utilization Model Research Stream

D. Social Cognitive Theory

The social cognitive theory is concerned with how perceptions of self-efficacy affect and individual's actions and motivations. According to Bandura [51], Self-efficacy is an individual belief in his abilities to mobilize the actions, motivations, and cognitive resources needed to exercise control over given events [50], [51], [52]. The theory is clearly well suited to studying the individual's behavior in the domain of information system, because self-regulated behavior in terms of information systems seems critically important for ensuring the information system's effectiveness. Self-efficacy has been argued as the one of the most important factors which motivates and regulates individual behavior [53]. Bandura defined self-efficacy as ''People's judgments of their capabilities to organize and execute courses of action required to attain designated types of performance [51]. Self-efficacy is concerned not with the skills one has but with judgments of what one can do with whatever skills one possesses. In other words, self-efficacy describes an individual's belief in his ability to perform a specific behavior.

Computer Self-Efficacy (CSE) is derived from Self-Efficacy [49], [54]. As noted by Compeau and Higgins [54], computer self-efficacy refers to self-assessment of an individual's ability to apply computer skills to complete the specified tasks. Computer self-efficacy has been related to various individual's computing behavior, such as adoption of an information system [54], [55], [12].
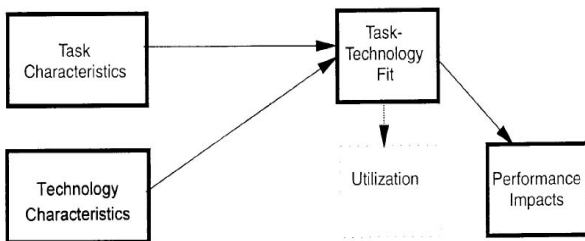
Figure 2.   Task-Technology Fit Model Research Stream
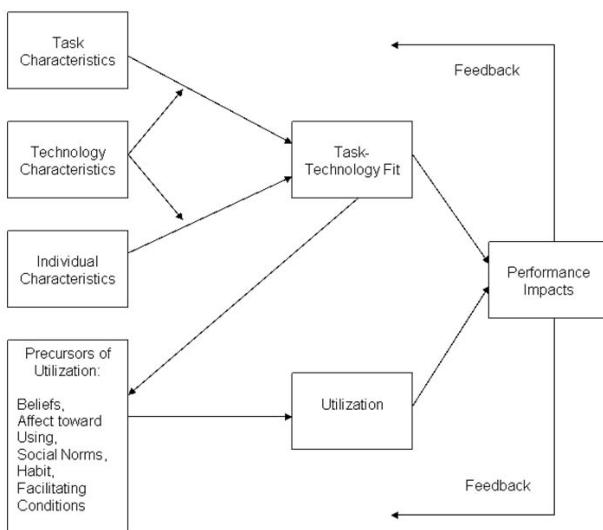
Figure 3.   Technology-to-Performance Chain Model

The three key attributes of information security are confidentiality, integrity, and availability [57]. Based on these attributes and in line with Compeau and Higgins [54], and Marakas [56], this study defines computer security self-efficacy as the "perceived efficacy of individual ability to use secure online banking technology to perform specific secure tasks".

III.   PROPOSED RESEARCH FRAMEWORK

To measure an information system's perceived utilization and effectiveness, this research will adapt the technology-to-performance chain model proposed and validated by Goodhue and Thompson [42], [58]. The Task-Technology Fit defined and measured the FIT between the task and the technology as an alternative user evaluation measure to the user information performance impact [31], [30]. That is, Goodhue and Thompson developed a perceived FIT measure as a substitute for measuring FIT as a moderator (interaction), which is the method suggested by Venkatraman for FIT as moderation models [59]. Goodhue and Thompson extended this approach along the theory dimension by adding utilization as a dependent variable. According to Goodhue and Thompson, there is a limitation of the utilization focus model they proposed [42], [58]. Utilization is not always voluntary for many system users and to the extent that utilization is not voluntary, performance impact will depend increasingly upon Task-Technology Fit rather than utilization. This limitation has been considered in this research by studying the secure online banking system with voluntary utilization only.

In an attempt to address this project's research questions, the intention is to ensure that Goodhue and Thompson technology-to-performance chain base model applies, with the collected data set first test without computer security self-efficacy Fit. The base model includes four direct effects on effectiveness, the task characteristics, technology characteristic, task-technology fit, and utilization. The relationship between technology to computer self-efficacy fit and utilization was tested by previous research [46] in which the finding was that computer self-efficacy has a direct effect on utilization, but no significant FIT interaction effects. This research however, intends to test technology-to-performance chain model on online banking context that have tasks with security characteristics. The relationship between computer security self-efficacy fit and utilization need to be tested. The analysis then tests the additions of the computer security self-efficacy (CSSE) construct as both a direct effect and a moderator fit with computer self-efficacy, that is, the proposed model intends to test both the effect of Task-Technology Fit (TTF), and the effect of Computer Security Self-Efficacy FIT (CSSEF) on Utilization and effectiveness. The full proposed framework is shown in Fig 4.

While a variety of definitions of the task and technology constructs have been suggested, this paper will use the definition suggested by Goodhue and Thompson [42] who defined task as "Actions carried out by individuals in turning inputs into outputs". When using the secure online banking system, the three main tasks which will be considered in this research are to view account details, to transfer funds, and to pay bills.
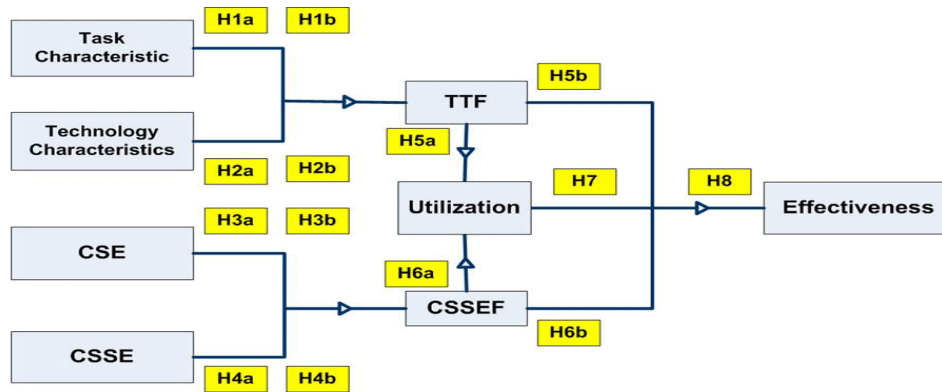
Figure 4. Proposed Research Framework

Goodhue and Thompson defined technology as "tools used by individuals in carrying out their tasks". Secure online banking is the technology used by these research respondents. Table 1 shows research hypothesis. There are three gaps this research intends to study. Propose measurement for computer security self-efficacy (gap1); Effect of computer security self-efficacy fit on utilization (gap 2) addressed at hypothesis number H6a; and Hypothesis number H8 summarizes the main research contribution this research intends to study (gap 3) which is "In order for an information system with secure concerned tasks, to have a positive effectiveness outcome, first, the technology must be utilized; second, there must be a positive Task-Technology Fit; and third there must be a positive Computer Security Self-Efficacy Fit (CSSEF)". The Proposed framework constructs with their definitions are shown in Table 2.

TABLE I. RESEARCH HYPOTHESIS

H1a	There is a direct impact of Task Characteristics on Utilization.
H1b	There is a direct impact of Task Characteristics on Effectiveness.
H2a	There is a direct impact of Technology Characteristics on Utilization.
H2b	There is a direct impact of Technology Characteristics on Effectiveness.
H3a	There is a direct impact of Computer Self-Efficacy (CSE) on Utilization.
H3b	There is a direct impact of Computer Self-Efficacy (CSE) on Effectiveness.
H4a	There is a direct impact of Computer Security Self-Efficacy (CSSE) on Utilization.
H4b	There is a direct impact of Computer Security Self-Efficacy (CSSE) on Effectiveness.
H5a	Increased Task-Technology Fit (TTF) has a positive impact on Utilization.
H5b	Increased Task-Technology Fit (TTF) has a positive impact on Effectiveness.
H6a	Increased Computer Security Self-Efficacy Fit (CSSEF) has a positive impact on Utilization.
H6b	Increased Computer Security Self-Efficacy Fit (CSSEF) has a positive impact on Effectiveness.
H7	There is a direct impact of Utilization on Effectiveness.
H8	In order for an information system with secure concerned tasks, to have a positive effectiveness outcome, first, the technology must be utilized; second, there must be a positive Task-Technology Fit; and third there must be a positive Computer Security Self-Efficacy Fit (CSSEF).

TABLE II. CONSTRUCT DEFINITIONS

Construct	Definition
Task Characteristics	Actions carried out by individuals in turning inputs into outputs.
Technology Characteristics	Tools used by individuals in carrying out their tasks.
Computer Self Efficacy	Perceived efficacy of individual ability to use computing technology to perform specific tasks.
Computer Security Self Efficacy	Perceived efficacy of individual ability to use online technology to perform specific secure tasks.
Task-Technology Fit (TTF)	Degree to which a technology assists an information system user in performing task. More specifically it's the correspondence between Task characteristics, User abilities, and technology characteristics.
Computer Security Self-Efficacy Fit (CSSEF)	The fit between individual judgment on his ability to use computer and his ability to securely use it for secure tasks.
Utilization	An outcome construct that can be influenced by the process of design and implementation and by characteristics of the information system, the task, the individual user and their interaction.
Effectiveness	User's perception of how well a technology will help an individual complete a specific task or set of tasks.

IV. RESEARCH METHODOLOGY

The methodology of this study mainly involves a multi-method approach. Interviews with computer security experts and academics are needed to support literature and develop questions related to the computer security self-efficacy construct. A pilot study needs to be carried out to assess the feasibility of the survey. After validation of the survey questionnaire, the quantitative survey method will be used for data collection. For the proposed model constructs, this research started from a published technology-to-performance chain model questionnaire [42], and used by [60], [46], [61], [62]. This research chose those sub-constructs found in these previous studies that are relevant to this research, and then generalized the associated questionnaire items for our context.

Fit is an interaction (Moderator) and thus will not be collected via questionnaire.

A multi-method approach was chosen since it facilitates explanation and prediction as well as assisting in developing a more holistic view of the aspects under investigation [63]. The proposed framework constructs with their sources are shown in Table 3. We expect to have a sample size of not less than 384. The Structural Equation Modeling using Analysis of Moment Structures (AMOS) will be used as a data analysis tool in this research.

TABLE III. CONSTRUCT SOURCES

Construct	Source
Task Characteristics	[43], [65], [59], [42], [66], [67], [68] Main Source: [42]
Technology Characteristics	[43], [65], [59], [42], [66], [67], [68] Main Source: [42]
Computer Self Efficacy	[69], [29], [13], [28], [27], [54], [51, 52], [47] Main Source: [70]
Computer Security Self Efficacy	Literature+ Interview [57], [55], [51], [52], [53] ,[71], [72], [73], [74] Main Source: [75]
Task-Technology Fit (TTF)	[60], [43], [65], [59], [42], [66], [67], [68] Main Source: [43]
Utilization	[47], [43], [65], [59], [42], [66], [67], [68] Main Source: [43]
Effectiveness	Main Source: [42]

V. CONCLUSION

This paper is part of a larger research project to study the computer security self-efficacy effect on secure information system effectiveness. The study intends to develop and validate a measurement instrument for measuring effect of computer security self-efficacy on secure information system effectiveness. After validation of its constructs, this instrument is expected to show suitability for use in the study and may be used in similar secure information system contexts. In its use, the instrument will advance studies in computer security and information system effectiveness research. Further empirical research will be required for the purpose of validation. Such a study will strengthen or refute claims of other related studies and will offer theoretical contributions to the field of information system Security.

ACKNOWLEDGMENT

REFERENCES

[1] S. Chan and M. Lu, "Understanding internet banking adoption and use behavior: A Hong Kong perspective," *Journal of Global Information Management,* vol. 12, p. 21, 2004.

[2] V. S. Lai and H. Li, "Technology Acceptance Model for Internet Banking: An invariance Analysis," *Information & Management,* vol. 42, p. 13, 2005.

[3] A. Sachan and A. Ali, "Competing in the age of information technology in a developing economy: Experiences of an Indian Bank," *Journal of Cases on Information Technology,* vol. 8, p. 19, 2006.

[4] W. Wresch and S. Fraser, "Managerial strategies used to overcome technological hurdles: A review of e-commerce efforts used by innovative Caribbean managers," *Journal of Global Information Management,* vol. 14, p. 16, 2006.

[5] J. A. Cazier, B. M. Benjamin, and R. D. Louis, "E-business differentiation through value-based trust," *Information and Management,* vol. 43, p. 9, 2006.

[6] Y. S. Wang, Y. M. Wang, H. H. Lin, and T. I. Tang, "Determinants of user acceptance of internet banking: An empirical study," *International Journal of Service Industry Management,* vol. 14, p. 18, 2003.

[7] Thorton Consulting, "Thorton consulting online banking: a success," *Australian Banking and Finance,* vol. 5 1996.

[8] B. Howcroft, R. Hamilton, and P. Heder, "Consumer attitude and the usage and adoption of home-based banking in the United Kingdom," *International Journal of Bank Marketing,* vol. 20 p. 10, 2002.

[9] H. Amin, "Internet banking adoption among young intellectuals," *Journal of Internet Banking and Commerce,* vol. 12, p. 13, 2007.

[10] D. R. Compeau, C. A. Higgins, and S. Huff, "Social cognitive theory and individual reaction to computing technology: A longitudinal study. MIS Quarterly," *MIS Quarterly,* vol. 23, p. 13, 1999.

[11] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: A comparison of two theoretical models," *Management Science,* vol. 35, p. 19, 1989.

[12] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Quarterly,* vol. 27, pp. 425-478, Sep 2003.

[13] M. M. Saudi, "A NEW MODEL FOR WORM DETECTION AND RESPONSE," PhD, Department of Computing, School of Computing, Informatics and Media, University of Bradford, 2011.

[14] M. M. Saudi, A. J. Cullen, and M. Woodward, "Efficient STAKCERT KDD Processes in Worm Detection," *World Academy of Science, Engineering and Technology Journal,* p. 3, 2011.

[15] Lee and Kozar, "Investigating Factors Affecting the Adopting of Anti-Spyware Systems," *Communications of ACM,* vol. 48, p. 5, 2005.

[16] Zang, "What do Consumers Really Know about Spyware?," *Communications of ACM,* vol. 48, p. 4, 2005.

[17] R. Power, "CSI computer crime & security survey," Computer Security Institute2008.

[18] Lee, Shambhu, Raghav, and Sharman, "Secure Knowledhe Management and The Semantic Web," *Communications of the ACM,* vol. 48, p. 6, 2005.

[19] Jefferson, Rubin, Simons, and Wagner, "Analyzing Internet Voting Security," *Communications of the ACM,* vol. 47, p. 5, 2004.

[20] Gimun Kim, BongSik Shin, and H. G. Lee, "Understanding dynamics between initial trust and usage intentions of mobile banking," *Information Systems Journal,* vol. 19, p. 28, 2009.

[21] D. Zissis and D. Lekkas, "Addressing cloud computing security issues," *Future Generation Computer Systems,* vol. 28, p. 9, 2012.

[22] Hong, Kwo-Shing, Chi, Yen-Ping, Chao, Louis, and Tang, "An Integrated System Theory of Information Security Management," *Information Management and Computer Security,* vol. 11, 2003.

[23] Kotulic, Andrew, and Clark, "Why there aren't more information security research studies," *Information and Management,* vol. 41, p. 13, 2004.

[24] C. Sherrie, C. P. Prashant, and S. Richard, "A Research Framework for Information Systems Security," *Journal of Information Privacy & Security,* vol. 2, p. 27, 2006.

[25] W. Orlikowski and C. Iacono, "Research Commentary: Desperately Seeking the "IT" in IT Research—A Call to Theorizing the IT Artifact," *Information System Research,* vol. 12, pp. 121-134, 2001.

[26] V. Venkatesh and H. Bala, "Technology Acceptance Model 3 and a Research Agenda on Interventions," *Decision Sciences,* vol. 39, 2008.

[27] S. A. Brown, A. R. Dennis, and V. Venkatesh, "Predicting Collaboration Technology Use: Integrating Technology Adoption and Collaboration Research," *Journal of Management Information Systems,* vol. 27, pp. 9-53, Fal 2010.

[28] A. R. Dennis, V. Venkatesh, and V. Ramesh, "Adoption of Collaboration Technologies: Integrating Technology Acceptance and

Collaboration Technology Research," *Working Papers on Information Systems,* vol. 3, 2003.

[29] D. L. Goodhue, "Understanding User Evaluations of Information Systems," *Management Science,* vol. 41, p. 17, 1995.

[30] M. L. Irick, "Task-technology fit and information systems effectiveness," *Journal of Knowledge Management Practice,* vol. 9, 2008.

[31] B. Ives, M. Olson, and J. Baroudi, "The measurement of user information satisfaction," *Communications of the ACM,* vol. 26, p. 8, 1983.

[32] J. E. Bailey and S. W. Pearson, "Development of a Tool for Measuring and Analyzing Computer User Satisfaction," *Management Science,* vol. 29, 1983.

[33] C. Coral, "An Applicable Data Quality Model for Web Portal Data Consumers," *World Wide Web archive,* vol. 11, pp. 465-484, 2008.

[34] A. Parasuraman, V. A. Zeithaml, and A. Malhotra, "E-S-QUAL - A multiple-item scale for assessing electronic service quality," *Journal of Service Research,* vol. 7, pp. 213-233, Feb 2005.

[35] W. H. DeLone and E. R. McLean, "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update," *Journal of Management Information Systems,* vol. 19, p. 9, 2003.

[36] W. H. DeLone and E. R. McLean, "Information systems success: the quest for the dependent variable," *Information Systems Research,* vol. 3, pp. 60-95, 1992.

[37] U. Nils, S. Stefan, and Gerold, "The State of Research on Information Systems Success," *A Review of Existing Multidimensional Approaches, in: WIRTSCHAFTSINFORMATIK,* vol. 51, p. 12, 2009.

[38] N. Urbach, S. S, and G. Riempp, "Development and Validation of Model for Assessing the Success of Employee Portals," 2009.

[39] N. Mohamad, "Measuring campus portal effectiveness and the contributing factors," *Campus-Wide Information Systems,* vol. 24, pp. 342-354, 2007.

[40] P. A. Pavlou and M. Fygenson, "Understanding and predicting electronic commerce adoption: An extension of the theory of planned behavior," *Management Information Systems Quarterly,* vol. 30, p. 38, 2006.

[41] S. L. Cathy, W. Sheng, and C. Tsung-Mei, "The Impact of Self-Monitoring on theory of Planned Behavior: Study of Web Portal Usage," *Southern Association for Information Systems Conference,* 2007.

[42] D. L. Goodhue and R. L. Thompson, "Task-Technology Fit and Individual-Performance," *MIS Quarterly,* vol. 19, pp. 213-236, Jun 1995.

[43] B. T. Pentland, "Use and Productivity in Personal Computers: An Empirical Test," in *Proceedings of the Tenth International Conference on Information Systems,* Boston, 1989, pp. 211–222.

[44] K. D. Jones, "Classroom response systems: using task technology fit to explore impact potential," MBA, University of Texas at San Antonio, MSIS, 2010.

[45] N. W. Reay, L. Pengfei, and L. Bao, "Testing a new voting machine question methodology," *The American Journal of Physics,* vol. 76, pp. 171-178, 2008.

[46] D. M. Strong, M. T. Dishaw, and D. B. Bandy, "Extending Task Technology Fit with Computer Self-Efficacy," *Database for Advances in Information Systems,* vol. 37, pp. 96-107, 2006.

[47] M. Keil, P. M. Beranek, and B. R. Konsynski, "Usefulness and Ease of Use: Field Study Evidence Regarding Task Considerations," *Decision Support Systems,* vol. 13, p. 16, 1995.

[48] M. T. Dishaw and D. M. Strong, "Extending the technology acceptance model with task-technology fit constructs," *Information & Management,* vol. 36, p. 9, 1999.

[49] F. Davis and P. Warshaw, "User acceptance of computer technology: A comparison of two theoretical models," *Management Science,* vol. 35, pp. 982-1103, 1989.

[50] A. Bandura, *Social foundations of thoughts and action: a social cognitive theory.* Englewood Cliffs: Prentice Hall, 1986.

[51] A. Bandura, *Self-efficacy: The exercise of control.* New York: Freeman, 1997.

[52] E. Ozer and A. Bandura, "Mechanisms governing empowerment effects: a self-efficacy analysis," *Journal of Personality and Social Psychology,* vol. 58, p. 14, 1990.

[53] A. Bandura and F. Jourden, "Self-regulatory mechanisms governing the impact of social comparison on complex decision making," *Journal of Personality and Social Psychology,* vol. 60, p. 10, 1991.

[54] D. Compeau and C. Higgins, "Computer self-efficacy: development of a measure and initial test," *MIS Quarterly,* vol. 19, p. 12, 1995.

[55] P. Ellen, W. Bearden, and S. Sharma, "Resistance to technological innovations: an examination of the role of self-efficacy and performance satisfaction," *Journal of the Academy of Marketing Science,* vol. 19, p. 10, 1991.

[56] M. Smith, "Computer security-threats, vulnerabilities, and countermeasures," *Information Age,* vol. 11, p. 5, 1989.

[57] G. Marakas, M. Yi, and R. Johnson, "The multilevel and multifaceted characteristics of computer self-efficacy," *Information Systems Research,* vol. 9, p. 26, 1998.

[58] D. L. Goodhue, "Development and measurement validity of a task-technology fit instrument for user evaluations of information systems," *Decision Sciences,* vol. 29, p. 34, 1998.

[59] N. Venkatraman, "The Concept of Fit in Strategy Research: Toward Verbal and Statistical Correspondence," *Academy of Management Review,* vol. 14, pp. 423-444, 1989.

[60] M. T. Dishaw and D. M. Strong, "Assessing software maintenance tool utilization using task-technology fit and fitness-for-use models," *Journal of Software Maintenance-Research and Practice,* vol. 10, pp. 151-179, May-Jun 1998.

[61] M. T. Dishaw and D. M. Strong, "The Effect of Task and Tool Experience on Maintenance CASE Tool Usage," *Information Resources Management Journal,* vol. 16, pp. 1-16, 2003.

[62] M. T. Dishaw and D. M. Strong, "Supporting Software Maintenance with Software Engineering Tools: A Computed Task-Technology Fit Analysis," *The Journal of Systems and Software,* vol. 44, pp. 107-120, 1998.

[63] L. Bradley and K. Stewart, "The diffusion of online banking," *Journal of Marketing Management Information Systems Quarterly,* vol. 19, p. 22, 2003.

[64] D. L. Goodhue, "Supporting users of corporate data : the effect of I/S policy choices," Ph D, Massachusetts Institute of Technology, Sloan School of Management, 1988.

[65] D. L. Goodhue, "User evaluations of MIS success: What are we really measuring?," in *Proceedings of the Hawaii Twenty-Fifth International Conference on Systems Sciences,* 1992, pp. 303-314.

[66] D. Goodhue, R. Littlefield, and D. Straub, "The Measurement of the Impacts of the IIC on the End-Users: The Survey," *Journal of the American Society for Information Science,* vol. 48, p. 12, 1997.

[67] D. Goodhue, "The model underlying the measurement of the impacts of the IIC on the end-users," *Journal of the American Society for Information Science,* vol. 48, pp. 449-453, May 1997.

[68] V. Venkatesh, C. Speier, and M. G. Morris, "User acceptance enablers in individual decision making about technology: Toward an integrated model," *Decision Sciences,* vol. 33, pp. 297-316, Spr 2002.

[69] D. Compeau, J. Gravill, N. Haggerty, and H. Kelley, "Computer selfefficacy: a review. In: Zhang P, Galletta D, editors. ," in *computer interaction and management information systems* 2006

[70] N. E. Miller and J. Dollard, *Social Learning and Imitation.* New Haven, CT: Yale University Press, 1941.

[71] A. Bandura and R. H. Walters, *Social Learning and Personality Development.* New York: Holt, Rinehart & Winston, 1963.

[72] K. Glanz, B. K. Rimer, and F. M. Lewis, *Health Behavior and Health Education. Theory, Research and Practice.* San Fransisco: Wiley & Sons, 2002.

[73] A. Bandura, "Social cognitive theory: An agentive perspective," *Annual Review of Psychology,* vol. 52, p. 26, 2001.

[74] R. Hyeun-Suk, K. Cheongtag, and U. R. Young, "Self-efficacy in information security: Its influence on end users' information security practice behavior," *Computers & Security,* vol. 28, p. 10, 2009.

Design and Analysis of a Log Periodic Antenna by using Proximity Fed Network

S. H. Anuar[1], M.T.Ali[2], N. H. Khairul Anuar[3], A. L. Yusof[4] and M. N. Md Tan[5]

Microwave Technology Centre (MTC)
Faculty of Electrical Engineering
Universiti Teknologi MARA (UiTM)
40450, Shah Alam, Selangor, Malaysia
[1]hajaranuar@gmail.com, [2]mizi732002@yahoo.com, [3]hafizah.anuar@gmail.com, [4]azita968@salam.uitm.edu.my and [5]mnor1408@yahoo.com.

Abstract — **Log periodic antenna (LPA) has been designed using rectangular microstrip patches with inset feed using five elements. The antenna elements and feeding networks is designed using Computer Simulation Technology (CST) Microwave Environment software. The manipulative variable is the spacing between the patches. The simulated and measured results are compared and it is observed that the value of return loss, VSWR and gain of the LPA improved when the value of spacing between patches, *d* is decreased.**

Keywords: Log periodic antenna; microstrip antenna; scaling factor; proximity feed; FR-4 substrate; spacing between patches; CST.

I. INTRODUCTION

In past decades, wireless communications has been developing rapidly and one of the technologies that have been developed to cater this need is microstrip antenna. Wireless communication microstrip antennas operate in the microwave frequency range and are widely used on base stations and as well as in mobile handsets. Microstrip patches are often used as a single element antenna. It has gain popularity because of its main advantages of its low volume and light in weight. But, the most serious limitation of microstrip antenna is the narrow bandwidth. The bandwidth of a single element antenna is usually in the range of 1-3%. This narrow bandwidth is related to the high wave trapping action of the microstrip transmission line.

One way to improve the bandwidth of the single element microstrip antenna is by increasing the thickness of the substrate or by lowering the dielectric constant. However, this can cause maximized power loss due to the increase surface wave loss of the side lobe. It can also caused radiation and reactance associated with the feed junction [1].

Alternatively, a lossy substrate can be used but the resulting low efficiency is usually unacceptable [2]. It can be concluded that microstrip antennas that operates as a single element usually have a relatively large half power beamwidth, low gain and low radiation efficiency. Thus, in order to avoid these problems, a number of different methods have been introduced to improve the bandwidth of the microstrip antenna. Microstrip antennas are used in array configurations to improve the gain and the range of the radiating structure [3].

So one of the methods that can be used in order to improve the bandwidth of the microstrip antenna is by applying the log periodic technique and it is called the log periodic antenna (LPA). Its impedance and radiation characteristics repeat periodically as the logarithm of frequency thus the name log periodic antenna [4, 5]. LPA is in the array form with the purpose to increase the bandwidth as developed by P. S. Hall [3]. According to work done by V. B. Romodin et al [6], it is also has been concluded that the log periodic microstrip array antenna is a good choice for communications application due to the various frequencies obtained by only using a single antenna.

Based on works done by past researchers, it has been proven that the log periodic technique has contributed to the wider bandwidth in antennas. Thus, in this paper, the LPA is studied by doing an analysis on the effect of spacing between the patches to the return loss, gain, voltage standing wave ratio (VSWR) and the radiation pattern of the antenna. This is essential as the size of the designed antenna directly affects the cost of its production in the fabrication stage. Apart from that, the simulated and measured data of the fabricated antenna is then compared. According to work done by P. S. Hall [3], the input return loss can be improved by altering the patch spacing, *d* in order to scan the beam away from the broadside direction.

The log periodic antenna structure is similar to a proximity coupled antenna; however the elements are designed such that they are a log size and spacing apart. The design is based on the linear array of coplanar patch antennas with the size and spacing of the patches increasing in a log periodic manner. The principle of LPA requires that all the dimensions of the array elements to be scaled in a log periodic manner, to achieve a corresponding scaling in frequencies [1].

II. SCOPE OF WORK

The antenna scope of work is to design, to analyze and to measure the log periodic antenna for frequency bands of 2 GHz – 10 GHz. The comparison between the simulated and measured data (measured data is obtained from the fabricated antenna) is done at the resonant frequency of 2.3 GHz for easier and clearer data comparison. SubMiniature version A (SMA) connector is used as the inset feed port of the antenna.

Figure 1 shows the design flow chart of the LPA. After all the specifications are known and the parameters are calculated, the design of antenna and simulation is done using the Computer Simulation Technology (CST) software to obtain the simulated result. Optimization is then done to obtain the maximized value of return loss, bandwidth and return loss. The simulation is repeated again in case the optimization value is not within the expected result. Fabrication is then done when the simulation result is satisfactorily and is producing the optimized result. Lastly, the measured result obtained from the fabricated antenna is then compared with the simulated result.
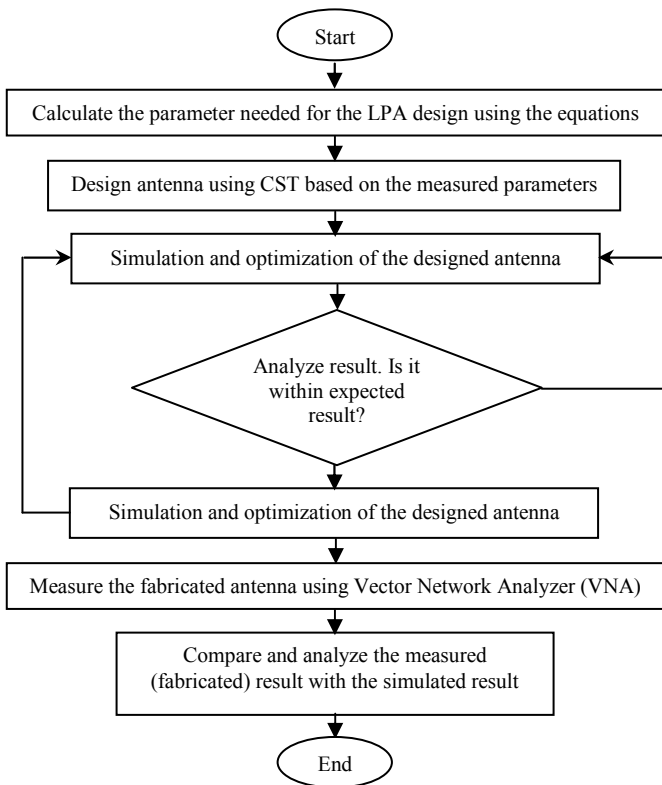
Figure 1.   Design flow chart of the log periodic antenna.

III.   METHODOLOGY
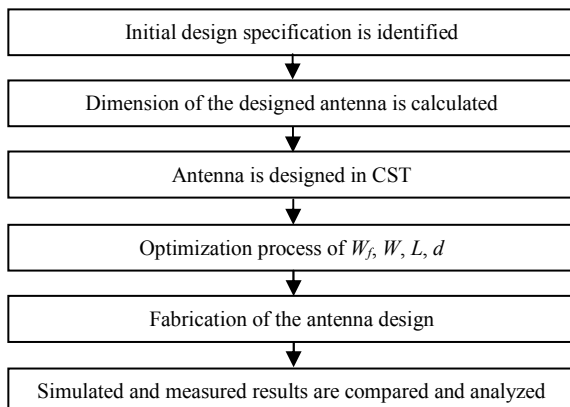
Figure 2.   Methodology of the log periodic antenna.

Figure 2 shows the methodology of the designed antenna. First, the initial design specification is identified. An appropriate frequency range of 2 GHz – 10 GHz is selected so that the designed antenna is able to operate in this chosen frequency range. The resonant frequency chosen is at 2.3 GHz for easier comparison purpose. The dielectric constant of the substrates, ε_r of 4.7 is used with the material type Flame Retardant 4 (FR-4) and height, h_1 & h_2 of 1.6 mm each. Then, the dimension of the antenna design which are the length, width and spacing of the patches are calculated using the appropriate equations according to the scaling factor, τ of 1.05.

The antenna is then designed in CST. Next comes the optimization process which is the most crucial step in the design. First, the location of the inset feed, W_f is optimized in the simulation to achieve the best impedance matching of 50 ohms. Then, the width of the substrates, W is optimized followed by the optimization of the length, L. After obtaining the initial antenna performance, the spacing between patches, d is then varied as the manipulated variable in this design. The effect of it to the overall antenna performance is observed and analyzed. The best simulated antenna design is then fabricated to produce the measured result. The simulated and measured results is then compared for further analysis.

IV.   ANTENNA LAYOUT AND DESIGN

TABLE I.   LPA DESIGN SPECIFICATION

Parameter	Value
Relative dielectric constant, ε_r	4.7
Tangent loss, $tan\delta$	0.022
Thickness of substrate 1 & 2, h_1 & h_2	1.6 mm
Thickness of copper layer (ground)	0.035 mm
Centre frequency, f_o (for comparison purpose between the simulated and measured result)	2.3 GHz
Frequency range	2 GHz – 10 GHz
Scale factor, τ	1.05

Table 1 shows the LPA design specification. The design consists of five rectangular microstrip patches which have different lengths and different relative spacing between them.

According to Valdez A. de A. Filho et al. [7], it is concluded that among circular, rectangular and square geometries; rectangular patch antenna has presented the best antenna performance. It is also concluded that the lower cut off frequency of the antenna arrays depended on the choice of patch geometry. Thus, the rectangular patches are implemented in the LPA. The location of the inset feed is optimized in the simulation in order to achieve the best impedance matching of 50 ohms [8].
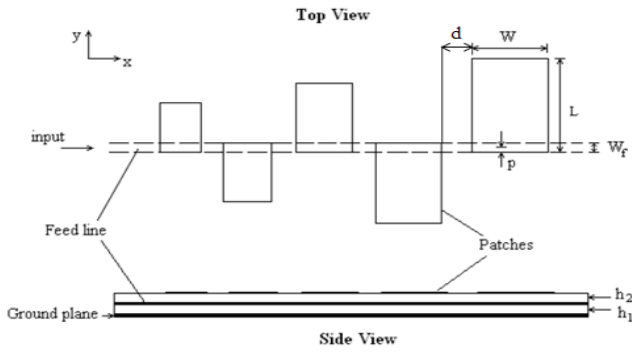
Figure 3.   Basic configuration of the LPA.

Figure 3 shows the basic configuration of the LPA. Figure 4(a) and 4(b) shows the schematic diagram of the antenna from the top and the side view designed in the simulation in CST. Figure 5 shows the fabricated antenna used in the measurement stage that is tested and measured in lab using the Vector Network Analyzer (VNA).
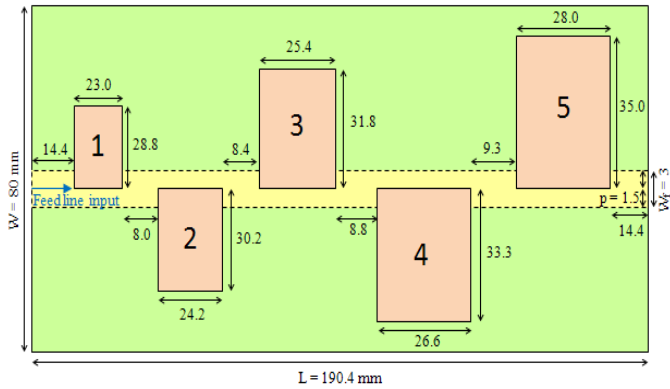
(a) Top view
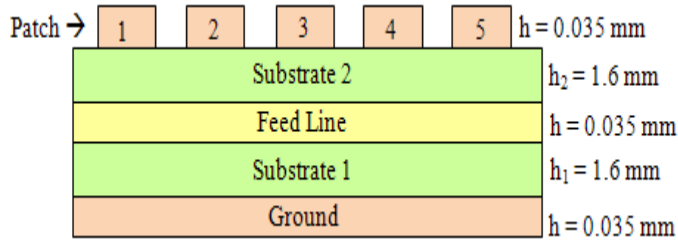
(b) Side view

Figure 4.   The schematic diagram of the LPA.
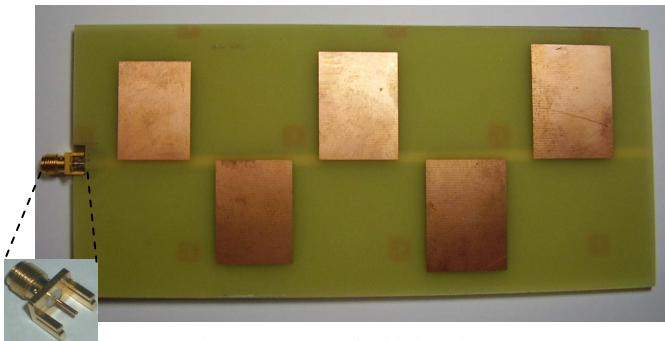
Figure 5.   The fabricated LPA.

The antenna consists of two isotropic dielectric substrate layers. There is a conventional microstrip line that is separated from the ground plane by dielectric substrate 1 in the bottom layer and in the upper layer, the patch resonators are located and displaced in alternate directions from the feed line as shown in figure 4(a). FR-4 is chosen to be used for both substrates due to its low in cost.

Five patches are used in the antenna. The length (L_m), width (W_m), and spacing between patches (d_m), are given by (1) where τ is the scale factor. If all dimensions of the array are multiplied by τ, it scales into itself with element m becoming element $m+1$, element $m+1$ becoming element $m+2$ etc. The height of both substrate layers and feed line width is kept constant. The basic guidelines for the design are as follow.

$$\tau = \frac{L_{m+1}}{L_m} = \frac{W_{m+1}}{W_m} = \frac{d_{m+1}}{d_m} \qquad (1)$$

- The substrate layers are selected: which are the FR-4.

- The upper and lower patch length is determined and then using (1), all the corresponding patch widths and elements are obtained.

- The width of the initial patch is chosen as $W = 0.8 \times L$ in order to prevent higher order modes.

- The width of the substrates, W is varied until the required resonant frequency of 2.3 GHz is obtained.

- The length of the substrates, L is varied until the optimized value of return loss is obtained.

- For the initial design, the value of $d = 1.05 \times L$ is used as an initial estimate then is varied in the simulation to determine the optimum spacing. The input return loss and gain is optimized by changing the patch spacing, d.

Series feed network is chosen to be used in the log periodic antenna over the corporate network. According to P. S. Hall [9], when a corporate feed is used, the bandwidth achieved by the log periodic patch array can have an excitation difficulties and the feed line can introduce significant losses. This can be overcome by using the series feed network as implemented in this LPA. The value of the feed line width of the microstrip line used in the design is $W_f = 2.9806$ using (2).

$$W_f = \left[\frac{expH}{8} - \frac{1}{4expH}\right]^{-1} \qquad (2)$$

The displacement, p is obtained by taking the halves of W_f as shown in (3). This displacement produces a maximum coupling that occurs when the patch edge was above the feed line edge and as p decreases, the coupling reduces smoothly [3].

$$p = \frac{W_f}{2} = \frac{2.9806 \, mm}{2} = 1.4903 \, mm = \sim 1.5 \, mm \qquad (3)$$

V. RESULT AND DISCUSSION

The five elements of the patch is scaled log periodically with scaling factor 1.05. The inset feed distance of the microstrip antenna is matched to 50 ohm input impedance.

TABLE II.   SIMULATED RESULT FOR FIVE PATCH ELEMENTS WITH 10 MM, 8 MM AND 6 MM PATCHES SPACING

Patch Element	Spacing between patches, d (10 mm)	Spacing between patches, d (8 mm)	Spacing between patches, d (6 mm)
Patch 1 (mm)	10.0	8.0	6.0
Patch 2 (mm)	10.5	8.4	6.3
Patch 3 (mm)	11.0	8.8	6.6
Patch 4 (mm)	11.6	9.3	6.9
Patch 5 (mm)	-	-	-
Frequency (GHz)	2.3	2.3	2.3
S11 (dB)	-10.0722	-24.9109	-11.6649
VSWR (Volt)	1.9749	1.1204	1.7067
Gain (dB)	7.969	10.570	10.170

From table II, it can be seen that when the spacing between patches, d is decreased, the increment does not occur linearly the same way to S11, VSWR and gain. They only improved until d is decreased to 8mm and then become worse when d is decreased further to 6mm. The optimum antenna performance occurs to d of 8 mm. This means that the performance of any LPA does not improve monotonically as the spacing is decreased or increased, but an optimum value must be determined through analysis during simulation. Thus, the design of spacing between patches, d of 8 mm with scale factor 1.05 is chosen to be used in the fabrication of the LPA due to its optimum antenna performance.

TABLE III.   CALCULATED DIMENSION OF THE FIVE PATCH ELEMENTS WITH 8 MM PATCHES SPACING

Patch Element	Length, L (mm)	Width, W (mm) (0.8 x L)	Spacing between patches, d (mm)
Patch 1	28.8	23.0	8
Patch 2	30.2	24.2	8.4
Patch 3	31.8	25.4	8.8
Patch 4	33.3	26.6	9.3
Patch 5	35.0	28.0	-

Table III shows the calculated dimension of the five patch elements of the log periodic antenna with the frequency range of 2 GHz – 10 GHz but for comparison purpose between the simulated and measured data, a single resonant frequency of 2.3 GHz is chosen.

A. Return Loss

Return loss is the portion of a signal that cannot be absorbed by the end of the line transmission or cannot cross an impedance change at some point in the transmission system. This component of the signal is reflected from the impedance discontinuity and travels back up the line from that point, since it cannot be absorbed by the termination or traverses the impedance irregularity and signal may be lost forever. S-Parameters or S11 or return loss shows how much power is reflected form the transmitter to the antenna. The value should be lower than -10 dB for the best result.

Figure 6 shows the available frequencies within the frequency band of 2GHz – 10 GHz which shows the characteristic of a log periodic antenna. Table IV shows the comparison of value between the simulated (from CST designed antenna) and measured (from fabricated antenna) frequencies to its corresponding S-parameter. It can be seen that there are slight difference between the simulated and measured data. This may be due to the usage of FR-4 substrate that has a low efficiency and thus produces losses. These losses might also be caused by the soldering activities where it could affect the copper wire's efficiency. It is known that a direct hot contact on any copper material would decrease the conductive features inside it and affects the measured result from the simulated result.
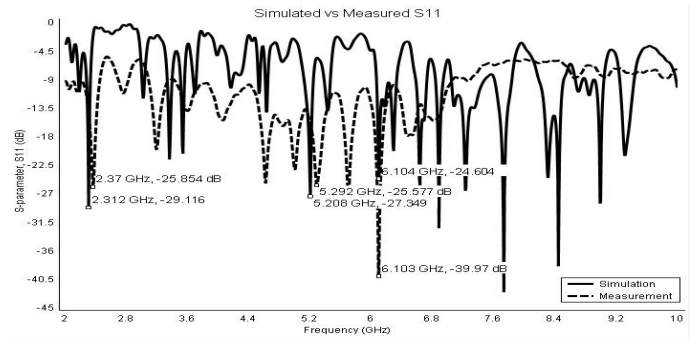
Figure 6.   S-parameter for frequency range of 2 GHz – 10 GHz.

TABLE IV.   COMPARISON BETWEEN THE SIMULATED AND MEASURED RESULT

Simulation		Measurement	
Frequency (GHz)	S11 (dB)	Frequency(GHz)	S11 (dB)
2.31	-29.116	2.37	-25.854
4.17	-10.282	4.07	-15.355
4.64	-13.990	4.63	-25.174
5.21	-27.349	5.29	-25.577
6.10	-24.604	6.10	-39.970
6.64	-25.608	6.54	-17.681

In order to do a comparison between the simulated and measured data, a resonant frequency at 2.3 GHz is chosen. Figure 7 shows the comparison between the simulated and measured S-parameter. The simulated S-parameter is -24.911 dB at 2.3 GHz while the measured S-parameter is -25.854 dB at a slightly shifted frequency of 2.37 GHz. This may be caused by the errors and losses that may happened during the fabrication process caused by soldering and etching where the dimension of fabricated design is not as precise as the dimension of the designed antenna in the simulation in CST.
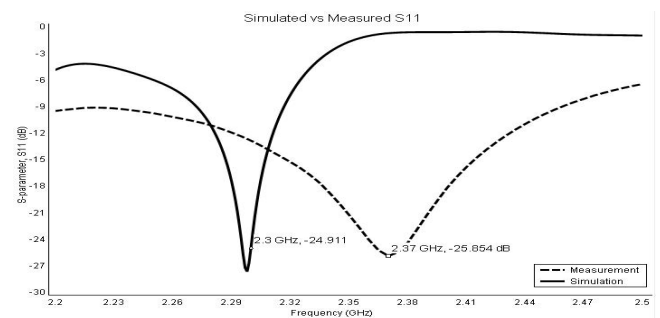
Figure 7.   Comparison of simulated and measured S-parameter at 2.3 GHz.

B. Gain

Gain of an antenna is closely related to directivity and it is a measure that takes into account the efficiency of the antenna as well as its directional capabilities. It can be defined as the ratio of the intensity (measure of the concentration of the radiated power) in a given direction, to the radiation intensity that would be obtained if the power accepted by the antenna were radiated isotropically [4]. High gain antenna provides longer range and better signal quality. The higher the gain obtained, the better the performance of the antenna. By referring to table II, it can be seen that the spacing size between patches at 8 mm produces the best value of gain which is 10.57dB.

C. Voltage Standing Wave Ratio (VSWR)

VSWR is a function of the reflection coefficient, Γ which describes the power reflected from the antenna and is determined from the voltage measured along a transmission line leading to the antenna. The smaller the VSWR is, the better the antenna is matched to the transmission line and the more power is delivered to the antenna. A perfectly matched antenna would have a VSWR of 1:1 which indicates that no power is reflected from the antenna [10].

The obtained VSWR is 1.9749 volt at 10 mm spacing and 1.1204 volt at 8 mm spacing at 2.3 GHz frequency as shown in figure 8. There is an improvement in VSWR as the spacing between patches, d is decreased.

(a) 10 mm   (b) 8 mm
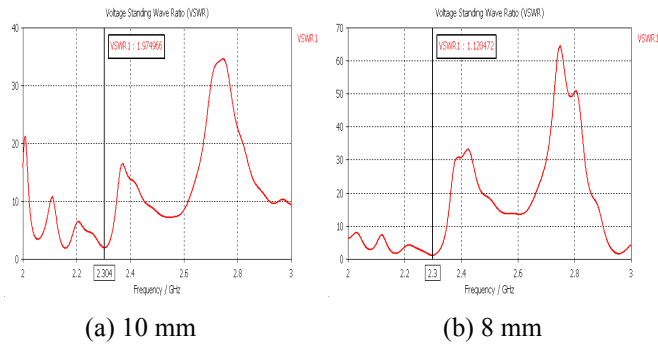
Figure 8.   Simulated VSWR at 2.3 GHz using spacing of 10 mm and 8 mm.

Figure 9 shows the comparison between the simulated and measured VSWR. The simulated VSWR is 1.12 volt at 2.3 GHz while the measured VSWR is 1.107 volt at a slightly shifted frequency of 2.37 GHz.
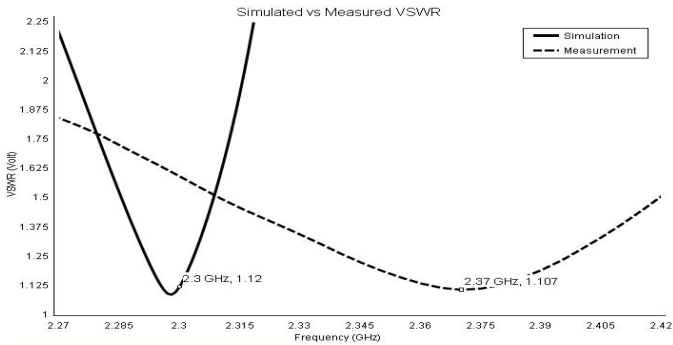
Figure 9.   Comparison of the simulated and measured VSWR.

D. Radiation Pattern

Radiation pattern is a defined as a mathematical function or a graphical representation of the radiation properties of the antenna as a function of space coordinated and is determined in the farfield region [4]. Figure 10 and 11 shows the radiation pattern for different spacing between patches, d of 10 mm and 8 mm. For d of 10 mm, the gain obtained is 7.969 dB and for d of 8 mm, the gain obtained is 10.57 dB. An increment of 2.601 dB is achieved when the spacing is decreased. All the simulated parameters are compared in table V.
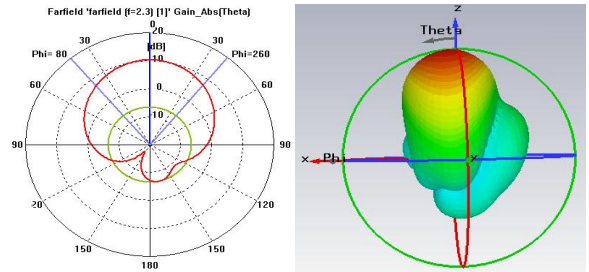
Figure 10. Radiation pattern at 2.3 GHz using spacing of 10 mm.
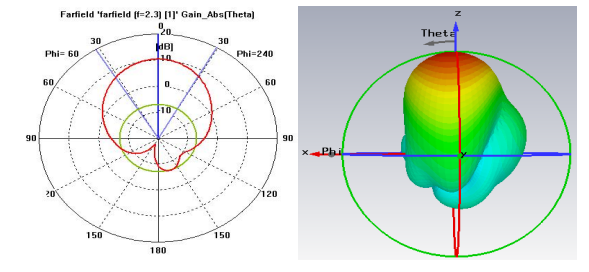
Figure 11. Radiation pattern at 2.3 GHz using spacing of 8 mm.

TABLE V.   COMPARISON BETWEEN PATCHES SPACING OF 10 MM AND 8MM

Parameter	Spacing between patches, d (10 mm)	Spacing between patches, d (8 mm)
S11 (dB)	-10.0722	-24.9109
VSWR (Volt)	1.9749	1.1204
Gain (dB)	7.969	10.570

VI.   CONCLUSION

It can be concluded that the value of return loss, VSWR and gain of the log periodic antenna improved when the value of spacing between patches, d is decreased. Based on the analysis, the optimum antenna performance occurred to spacing between patches, d of 8 mm. Note that table II shows that a further decrease of the spacing to 6 mm does not improve the antenna performance. This means that the performance of any log periodic antenna does not improve monotonically as the spacing is decreased or increased, but an optimum value must be determined through analysis during simulation. To obtain the required resonant frequency of 2.3 GHz, the width of the substrates, W is optimized. To obtain the optimized value of the return loss, the length of the substrates, L is varied. This 2.3 GHz frequency log periodic antenna can be applied in Wireless Fidelity (WiFi) and Worldwide Interoperability for Microwave Access (WiMax). An example of the application is in an indoor panel antenna [11].

REFERENCES

[1] Mohamad Kamal A. Rahim, Peter Gardner, *"Microstrip bandwidth enchancement using log periodic technique with inset feed,"* 2004, in press.

[2] Megha Dadel, Shweta Srivastava, K. P. Tiwary, *"Analysis of log periodic frequency independent arrays of rectangular patch antenna,"* IEEE, 2011, in press.

[3] P.S. Hall, *"Multioctave bandwidth log periodic microstrip antenna array,"* IEEE Proceedings, Vol. 133, No. 2, 1986, in press.

[4] C.A. Balanis, *Antenna Theory*, Third Edition, John wiley & sons, inc., New York, 2005. (references)

[5] Pozar and Schaubert, "*Microstrip antennas,*" Proceedings of the IEEE, vol. 80, 1992, in press.

[6] V. B. Romodin, V. I. Oznobikhin, V.V. Kopylov, *"Log periodic microstrip array,"* IEEE, 2000, in press.

[7] Valdez A. de A. Filho, Paulo H. da F. Silva, Adaildo G. D'Assuncao, *"A comparative study of three ultra-wideband log periodic microstrip antenna arrays,"* IEEE Proceedings, 2010, in press.

[8] S.H Zainud-Deen, K.R Mahmoud, M.El-Adawy, Sabry M.M Ibrahem, "*Design of yagi-uda antenna and electromagnetically coupled curl antenna using particle swarm optimization algorithm,*" Twenty Second National Radio Science Conference (NRSC 2005) Coiro-Egypt, 2005, in press.

[9] P. S. Hall, *"New wideband microstrip antenna using log-periodic technique,"* IEEE, Vol. 16, No. 4, 1980, in press.

[10] http://www.antenna-theory.com/definitions/vswr.php. (references)

[11] http://www.rfwel.com/4g-data/wimax-antennas.php (references)

Analyzing large Dynamic Set-point change tracking of MRAC by exploiting Fuzzy Logic based Automatic Gain Tuning

R. Karthikeyan, Rahul kumar yadav, Shikha Tripathi, Hemanth kumar.G

Electronics and Communication Engineering Department,
Amrita Vishwa Vidyapeetham, Amrita School of Engineering,
Bangalore, Karnataka, India - 560035
r_karthikeyan@blr.amrita.edu, rahul-yadav1990@in.com, {t_shikha, g_hemanth}@blr.amrita.edu

Abstract—**A Model reference adaptive control (MRAC) belongs to the class of adaptive servo system in which the desired performance is expressed with the help of a reference model. MRAC aims to create a closed loop controller with parameters that can change the response of the system to mimic a desired response. However, analysis unravels that there is a tolerance band for the set point change which defines the effectiveness of a particular adaptive gain (γ). Any changes in the set point which is beyond this band calls for a γ-readjustment. We propose a method which aims to overcome this pitfall in conventional MRAC by fusing fuzzy logic to dynamically vary γ. In essence, a fixed γ which fails to stabilize the system response in the advent of a large change in the average value of the set point shall be empowered with fuzzy logic to do the needful. Also, this concept never requires human interference for gain adjustment. A second order Linear Time Invariant system has been considered for all illustration. The results show considerable improvement in performance over the existing conventional MRAC system.**

Keywords-Model Reference Adaptive Control (MRAC); Fuzzy logic; Adaptive control; set point tracking.

I. INTRODUCTION

Non linear control technology aims to implement high performance control systems when plant dynamics are poorly known or not known at all [1]. It spans to include four classes of controllers namely, Robust controllers, Adaptive controllers, Fuzzy logic controllers and Neural controllers. MRAC belongs to the adaptive control genre, which consists of adjustable parameters and a mechanism to adjust such parameters to maintain consistent system performance in presence of uncertainty or unknown variation in plant parameters. More elaborately, for a first order system given by (1),

$$\dot{x} = Ax + Bu. \tag{1}$$

where x stands for the measured output and u is the control variable, an MRAC with a finely tuned gain γ, is capable of driving the closed loop system described by a model given by (2), where x_m stands for model's response and r, the reference signal.

$$\dot{x}_m = A_m x_m + B_m r. \tag{2}$$

In this paper, we use fuzzy logic as an inference engine and a carefully designed rule-base, which are typically IF-THEN rules [2].

In our study, we found that a particular gain value γ can yield effective set point tracking only if the average change in the set point or reference signal is within a band that a given γ can handle. Violation of the above value requires another γ to allow the needed tracking. This task of redesign can get sufficiently tedious and time consuming. To overcome this problem, exploitation of fuzzy inference can provide potential solution. To the best of our knowledge, no work has reported such a technique. Besides, trying to improve an MRAC performance in its native state has been relatively ignored in the decade gone by, which saw more of hybridizing intelligent control techniques with MRAC. Moreover, an explicit study of fuzzy logic handling any input size has not been elaborated so far. It shall be worthwhile to note that, while MRAC drives the plant's response in a manner that makes it mimic the output of the model, the incorporated fuzzy logic control claims the responsibility of providing the needed gain, allowing the plant to follow the specified model. Also, no necessity of tuning the MRAC in the advent of large step input is the merit of this fuzzy-MRAC hybrid genre. Standard metrics of performance evaluation have been provided for proving the efficacy of our study and solution.

The organization of the paper is as follows. Section 2 presents a description in brief about some of the related work proposed thus far. Section 3 brings out a short discussion on constructing MRAC layout for a second order system, using MIT rule. Section 4 introduces the details of the study of the proposed problem and the solution. The simulation result of our proposed technique is presented in Section 5. We conclude the paper by posting our conclusion and few remarks along with the directions for future work in Section 6.

II. BRIEF REVIEW OF RELATED WORKS

One of the earliest related works by Cheung [3] addresses the issue of tuning the adaptation gains of MRAC systems using fuzzy logic technique. The technique was implemented and applied to control a vehicle suspension system under severe operating conditions and the results exhibited both good ride and handling performance as required in vehicles. But this work did not show any simulation results on system's response for dynamically varying large changes in the set point of the system, under the MIT-based MRAC control. It shows load changes separately which does not dynamically vary during a single course of the system. It focuses on presenting a methodology of tuning the adaptation gains by fuzzy logic technique.

Another work proposed by Yin & Lee [4], aims to incorporate fuzziness in MRAC, not for gain scheduling, but to represent the unknown parameters using the fuzzy bases function expansion (FBFE) and then identify the coefficients of the FBFE. With its aid, the unknown plant parameters are estimated with good precision. The fuzzy-MRAC created, in essence, has been shown to perform better than conventional MRAC. However, their investigation does not explicitly expose or promote any idea on exploiting such a novel method for ensuring good response irrespective of the size of the input.

One of the recent work presented by Shao et al. [5], proposes a Neurofuzzy control based Model reference adaptive inverse control for induction machines, which are highly nonlinear and has time-varying parameters. It aims to resolve the shortage of MRAC by using fuzzy logic and neural network, based on rotor field orientation motion model of the induction motor. However, this exceptional proposition lacks the demonstration of a detailed analysis of the drawback that we shall discuss in the succeeding sections.

The work by Liu et al. [6] is a significant proposition in that it aims to fuse Neurofuzzy networks with MRAC to implement a successful nonlinear MRAC. To validate the effectiveness, varying set point changes has been considered, but the case of system response in the advent of large change in average value of the step size has not been demonstrated.

Evidently, the investigation that prevailed in the last decade focused on hybridizing intelligent control systems with conventional MRAC and exploiting it on SISO or MIMO systems, thereby showing the effectiveness of the created novel idea, only for a specific application. A void created by the absence of a generic explicit study of automatic gain scheduling to handle any input size has been relatively ignored. Our proposition aims to fill this gap between current trends in MRAC investigation and former findings pertaining to MIT-rule based MRAC. The succeeding sections justify the claim.

III. MRAC USING MIT-RULE FOR A SECOND ORDER SYSTEM

The classical approach to MRAC design of a second-order linear time-invariant system is presented below. The architecture is obtained based on the MIT rule. MRAC for first-order LTI system such as (1) can be found in most of the textbooks on adaptive control (e.g. [1]).

MRAC begins by defining the tracking error e, which is the difference between the plant output and the reference model output as shown in (3).

$$e = y_{plant} - y_{model}. \tag{3}$$

The error enables one to define a cost function, usually denoted by J, which depends on parameter θ adapted within the controller. The cost function dictates how the parameters are to be updated. Equation (4) shows a typical cost function.

$$J(\theta) = \frac{1}{2} e^2(\theta). \tag{4}$$

Equation (5) gives the update law.

$$\frac{d\theta}{dt} = -\gamma \frac{\delta J}{\delta \theta} = -\gamma e \frac{\delta e}{\delta \theta}. \tag{5}$$

The above relation of dθ/dt and J(θ) is known as the MIT rule and this rule provides the adaptive nature to the controller. The structure of MRAC as a function of θ1 and θ2 is displayed in Fig.1. The control law is defined in (6).

$$u = \theta_1 r - \theta_2 y_{plant}. \tag{6}$$

We now use (3) to derive the useful expression of the given error in (7).

$$e = y_{plant} - y_{model} = G_{plant} u - G_{model} r. \tag{7}$$

G_{plant} and G_{model} represent the plant and the model transfer functions, respectively. The change of error with respect to the control parameters, known as sensitivity derivative, is obtained in (8) and (9), by using the model parameters, instead of that of the plant.

$$\frac{\partial e}{\partial \theta_1} = \frac{a_{1m}s + a_{0m}}{s^2 + a_{1m}s + a_{0m}} r, \tag{8}$$

$$\frac{\partial e}{\partial \theta_2} = \frac{a_{1m}s + a_{0m}}{s^2 + a_{1m}s + a_{0m}} y_{plant}. \tag{9}$$

Now the MIT rule can be applied to obtain the update rule for each θ. This is illustrated in (10) and (11).

$$\frac{d\theta_1}{dt} = -\gamma \frac{\partial e}{\partial \theta_1} e = -\gamma \left(\frac{a_{1m}s + a_{0m}}{s^2 + a_{1m}s + a_{0m}} r \right) e, \qquad (10)$$

$$\frac{d\theta_2}{dt} = -\gamma \frac{\partial e}{\partial \theta_2} e = -\gamma \left(\frac{a_{1m}s + a_{0m}}{s^2 + a_{1m}s + a_{0m}} y_{plant} \right) e. \qquad (11)$$

The expressions in (10) and (11) directly translate to the MRAC architecture shown in Fig.2. The reference model should be designed next.

For the purpose of illustration, we chose the model to have a $T_s = 3$ seconds and a $\zeta = 0.707$, which is an industry accepted standard. The acceptable system response for gains is presented in Fig.3. It is worthwhile to note that increasing γ makes the system respond much faster, but the system threatens to become unstable. While smaller γ leads to longer adaptation time. The gain has to be manually tuned, based on the application.
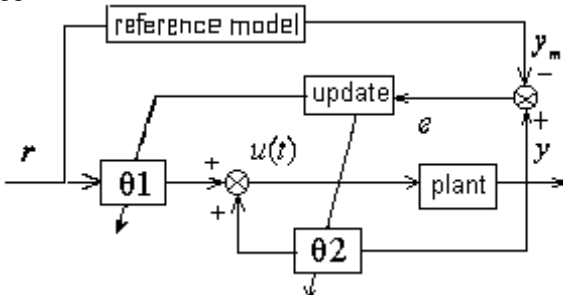
Fig. 1 Generalized Structure of MRAC.
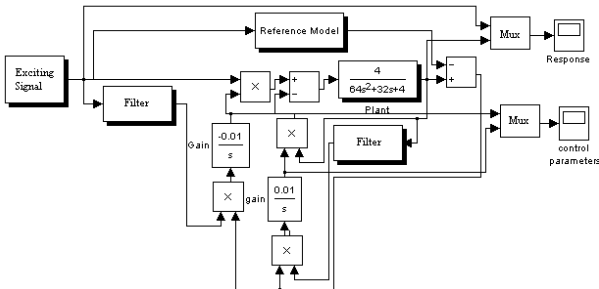
Fig. 2 MRAC layout for second order system.
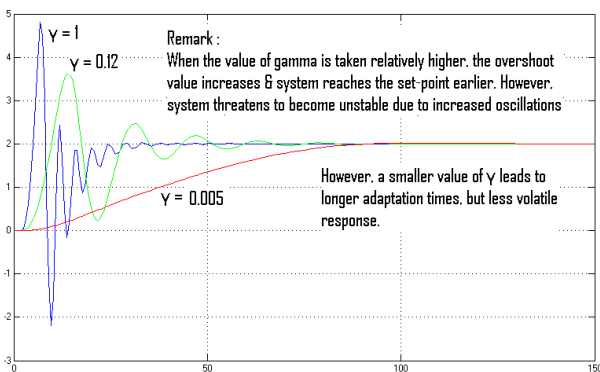
Fig. 3 System response for various γ

IV. PROPOSED TECHNIQUE AND ANANLYSIS

A detailed analysis of the control strategy exploiting MRAC reveals that MIT rule by itself does not guarantee convergence or stability. An MRAC designed using MIT rule is very sensitive to the amplitudes of the signals. And thus, as a general rule the value of γ is kept small. Tuning of the same is crucial to the adaptation rate and stability of the controller. However, as the signals in a system increase, the likelihood of the system entering nonlinear regions of operation increases. For very large reference signals, this nonlinear operation will occur for almost every physical system. This is where the drawback in the conventional MIT-based MRAC will show up, which we aim to eliminate by including fuzzy logic control. Since conventional MRAC completely fails to handle a plant of nonlinear behavior, we consider exploiting fuzzy logic to vary the gain so as to help prevent instability in the control strategy of MRAC. To validate its efficacy we study the system response when subjected to large variation in the set-point within the ambit of a single course of the system's study. The succeeding subsection exhibits the fore said.

A. Potential problem in MIT based MRAC

Consider the response of the second order system, presented in Fig.2, when subjected to multiple fluctuations in the set-point with $\gamma = 0.01$. This is shown in Fig.4. Evidently, large oscillations seen with γ being the same at t = 400 units, while reasonably acceptable response for t ≈ 100 units unfolds the fact that, around the set-point about which the γ was designed, there exist a tolerance band which defines the effectiveness of the chosen γ. The near marginal stability is demonstrated in Fig.5 for the same γ.
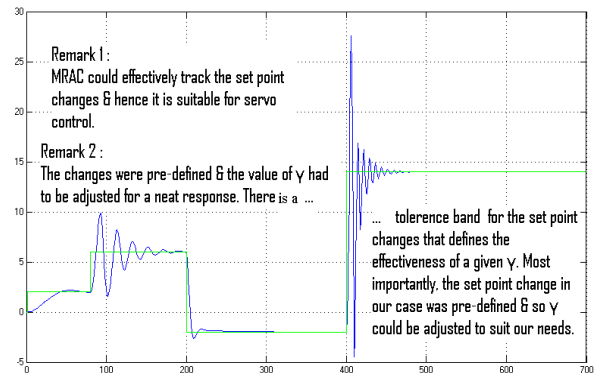
Fig. 4 Response of the second order system for $\gamma = 0.01$.
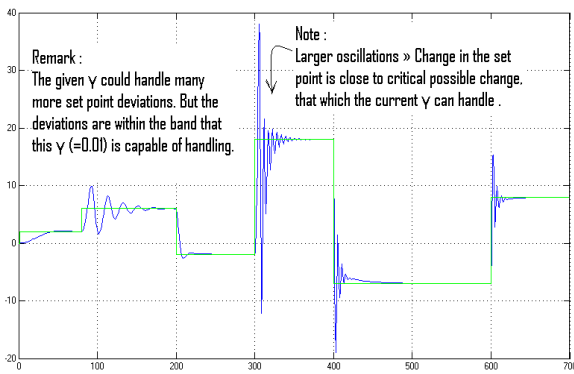
Fig. 5 Near marginal stability during the course of set-point tracking for slightly larger step, with γ = 0.01.

Clearly, a marginal increase in the set-point fluctuation induces large oscillations in the transient period of MRAC control (around t = 300 units). The instability aspect is exhibited in Fig.6, showing that further increase in the input size drives the system to fluctuate about the declared reference point. One can now see that on adjusting γ to 0.005, the tolerance band seems to have increased, as the revised γ is worthy of handling an even greater size of reference amplitude. Fig.7 illustrates this fact.

This behavior of MIT-MRAC with respect to the gain is a matter of concern. It is known that a lower γ leads to longer adaptation time which in turn leads to extended settling time of the system's response.
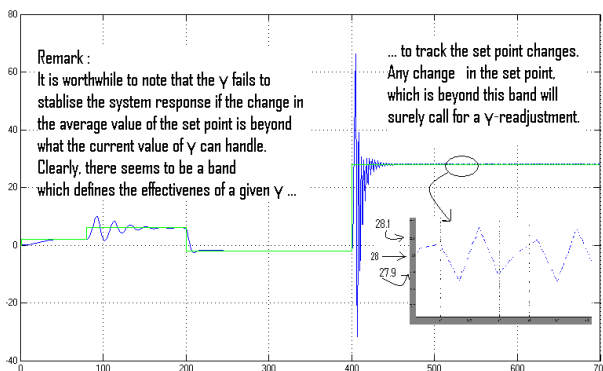
Fig. 6 Failure of γ = 0.01 to handle the change in the set-point beyond its tolerance band.
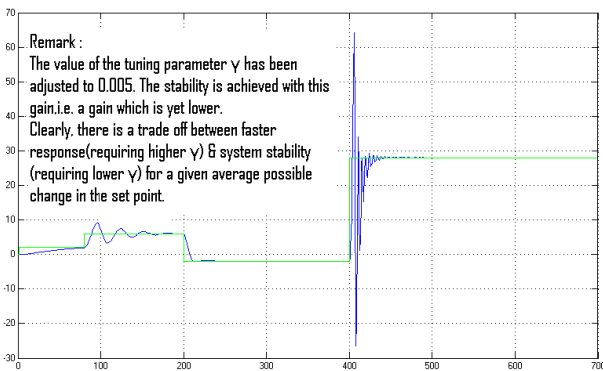
Fig. 7 Response with revised γ = 0.005.

Evidently, one cannot simultaneously handle the large variation in set-point value and its high frequency of occurrence, with a conventional MRAC. Our proposed technique addresses this drawback of the conventional MRAC & provides a feasible solution. This is presented in the succeeding subsection.

B. Proposed solution

We aim to resolve the above mentioned problem by incorporating fuzziness in MRAC. The layout of this fuzzy-MRAC is realized in next section. We shall now turn our focus to the exploited fuzzy control for the purpose of automatic gain scheduling. The error in response between the model and the plant, given in (1), and the time rate of change of this error ė, are the parameters fed to the fuzzy logic block for the purpose of decision making on gain selection. Since these parameters are continuously varying, the γ value also varies throughout the course of control, till it arrives at a gain that suppresses the error to zero. Being able to make γ as γ(e,t) is the crux of the concept of handling large input size. The γ-updation derives more from art than from scientific basis, and is mathematically stated by (12).

$$\gamma_{revised} = \gamma_{previous} \pm \Delta\gamma \qquad (12)$$

The factor $\pm\Delta\gamma$, which we shall call as the correction factor for γ-revision, takes care of updating old γ to obtain a new gain as demanded by the control.

The fuzzy control has been realized using the following structure of rules.

IF (e is x) and (ė is k) THEN $\gamma_{revised}$ is m,

where x, k, and m represents suitable linguistic variables. The designed rules are intuitive and are based on the following idea:

[IF e is present and ė is lowered, THEN decrease γ]
[IF e is present and ė is higher, THEN increase γ]
[IF e is zero, THEN no modification required in γ]

With a particular rule, Defuzzification provides the necessary numerical value to $\gamma_{revised}$. "Centriod" is the method chosen for defuzzification in our simulation. The implementation of the proposed solution is presented in the next section along with the analysis of the associated statistics.

V. EXPERIMENTATION AND STATISTICS

The experimentation of the proposed solution yielded promising results. Fig.8 shows the block diagram of the implementation. The MRAC block has been enclosed within the subsystem "MRAC", due to space constraint. The fuzzy controller takes the error and the change in error, to yield a suitable gain in the specified range.
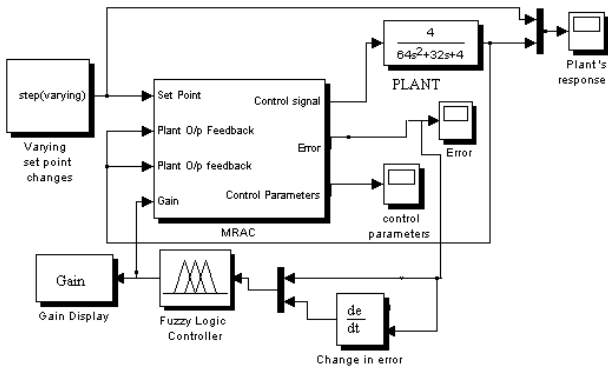
Fig. 8 Fuzzy-MRAC structure.

This helps the MRAC provide the necessary control for large reference changes, that a mono-gain MRAC can't handle. The architecture of the membership function of gain is presented in Fig.9. The linguistic rule base designed for the fuzzy inference engine is presented in Table I. The acronyms used are VL-Very Low, L-Low, ME-MEdium, MO-MOre, H-High, and VH-Very High.

TABLE I.   LINGUISTIC RULE BASE OF THE INFERENCE ENGINE

$\dot{e}$ \ e	VL	L	ME	H	VH
VL	VL	VL	ME	ME	H
L	L	L	ME	H	H
ME	ME	ME	MO	H	H
H	H	H	H	VH	VH
VH	M	H	VH	VH	VH

The design of the membership function chosen for change in error and the gain are presented Fig.10 and Fig.11, respectively. Rigorous efforts toward the choice of the range of action of the fore mentioned parameters can help suppress occasional oscillations & should certainly yield more appealing results, one of which is shown in Fig.13.
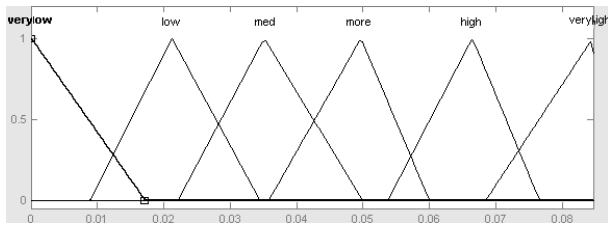
Fig. 9 Membership function of gain used for the control action
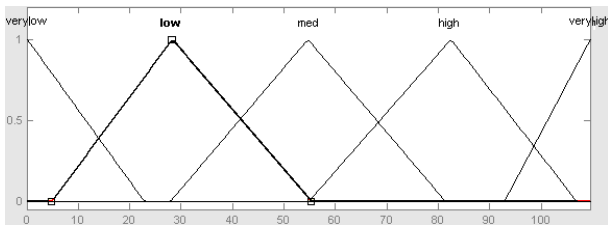
Fig. 10 Membership function of error used for the control action
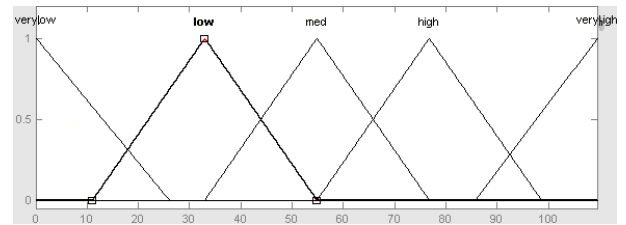
Fig. 11 Membership function of change in error, used for the control action.

Noteworthy remarks highlighting the areas that our proposition improves is included in Fig.13. The response of MIT-MRAC for γ=0.01 is shown in Fig.12. Table II presents the performance-indices as the standard metrics to validate the efficacy of our experimentation.
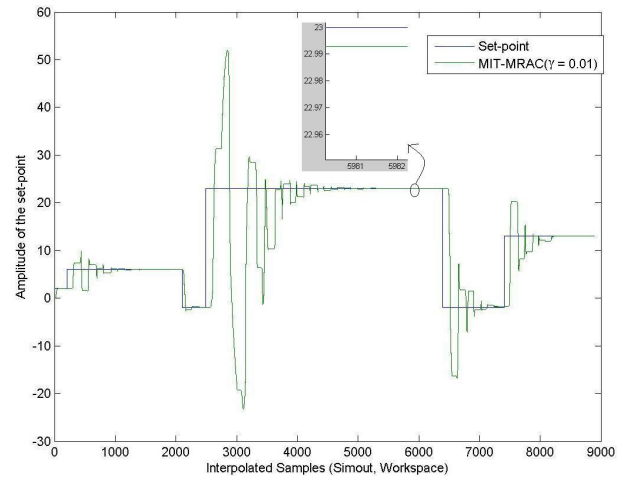
Fig. 12 System response for MIT-MRAC with γ = 0.01.

It can be inferred from this table that though a marginal & insignificant degradation in ITAE has suppressed ISE by a significant order. Clearly, apart from the adaptive nature that MRAC provides, the fuzzy logic makes the entire set-up double adaptive in that the MRAC can now handle a signal of any amplitude. The constructed severely varying conditions show the efficacy of the concept.
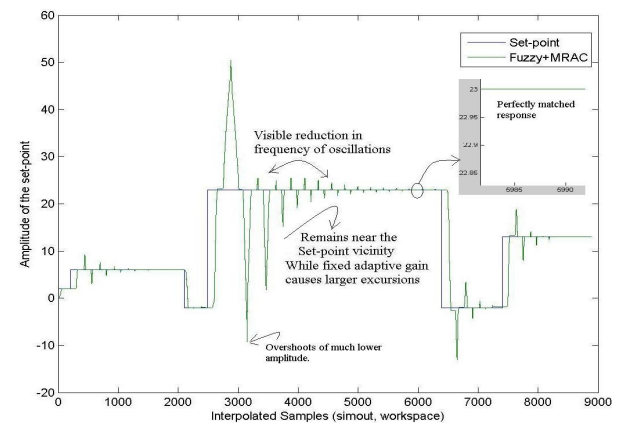
Fig. 13 System response for the considered design of the membership functions and the rule-base.

TABLE II.   VALIDATION USING PERFORMANCE INDICES

Performance- Index	MIT-MRAC ($\gamma = 0.01$)	Fuzzy-MRAC (automatic-tuning)
ISE	6762	4984
ITAE	2.76E+005	2.77E+005

The displayed results justify that fuzzy logic empowers MIT-MRAC.

VI.   CONCLUSION AND FUTURE WORK

The system proposed in this paper aims to fuse fuzzy logic with conventional MIT-rule based MRAC to handle large dynamic set-point changes. As far as the author's knowledge goes, this has not been explored so far. The result obtained by using this concept has demonstrated significant potential in handling great fluctuations in the input size. The concept also relieves the operator from manually tuning the gain, either during the course of control action, or in the advent of a large step size. Elimination of gain selection goes a long way in saving operator's time and operation cost. However, only a deliberately designed rule-base of the exploited inference engine can obtain the desired response. It becomes explicit from the obtained results that frequency of the reference fluctuation becomes a bottleneck for this genre of hybrid. The problem lies in the fact that, during the course of control with larger gain, the overshoots though of lower magnitude become apparent. Due to the nature of inferred control, damping occurs at a slower but acceptable rate. This shows up as slightly long settling time. It shall be interesting to determine if the current advances in the area of MRAC should intersect with the presented fuzzy-MRAC to suppresses the adaptation time and also improve the overall response of conventional MRAC. Our future work shall explore the fore mentioned scalability, along with the investigation of the avenues that should further improve the parameters that govern the accuracy of the MRAC's response.

REFERENCES

[1] J. J. E. Slotine and W. Li, *Applied Nonlinear Control,* Englewood cliffs, NJ:Prentice-Hall, 1991.

[2] T. J. Ross, *Fuzzy Logic with Engineering Applications*, Wiley and Sons, 2008.

[3] J. Y. M Cheung, "A fuzzy logic model reference adaptive controller," ," IEE colloquim on Adaptive controllers in Practice, pp. 1-6, 1996.

[4] T. K. Yin and C. G. S. Lee, "Fuzzy model-reference adaptive control," *IEEE Trans. Syst., Man, Cybern.*, Vol. 25, pp. 1606-1615, Dec. 1995.

[5] Z. Shao, Y. Zhan and Y. Guo, "Fuzzy neural network-based model-reference adaptive inverse control for induction machines," *IEEE Int.confrence on Applied superconductivity and electromagnetic devices*, Chengudu, China, Sept.25-27, 2009.

[6] X. J. Liu, F. L. Rosano, and C. W. Chan, "Model-reference adaptive control based on Neurofuzzy networks," *IEEE Trans. Syst., Man, Cybern.*, Vol. 34, pp. 1094-6977, June, 2004.

[7] S. Sastry and M. Bodson, *Adaptive Control: stability, Convergence and Robustness.* Upper Saddle River, NJ: Prentice-Hall, 1989.

[8] Y, D. Landau, *Adaptive Control: The Model Reference Approach.* New York: Marcel Dekker, 1979.

[9] S. Tong, J. Tang, and T. Wang, "Fuzzy adaptive control of multivariable nonlinear systems*," IEEE Trans. Fuzzy Sets, Syst.,* Vol. 111, pp. 153-167, 2000.

[10] X.-J. Liu and X. Zhou, "Structure analysis of fuzzy controller with Gaussian membership function," *Proc. Of 14th IFAC World Congr.,* Vol. K, pp. 201-206, Beijing, China, July 5-9, 1999.

A Rainfall Prediction Model using Artificial Neural Network

Kumar Abhishek[1*], Abhay Kumar[2#], Rajeev Ranjan[3#], Sarthak Kumar[4$]

[*]Dept. of Computer Science and Engineering, NIT Patna-800005

[#]Dept. of Information Technology, NIT Patna-800005

[$]Deolitte India, Pune.

kumar.abhishek@nitp.ac.in,abhay.kumar@nitp.ac.in,rrnitp@gmail.com,sar1989thak@gmail.com

Abstract-- **The multilayered artificial neural network with learning by back-propagation algorithm configuration is the most common in use, due to of its ease in training. It is estimated that over 80% of all the neural network projects in development use back-propagation. In back-propagation algorithm, there are two phases in its learning cycle, one to propagate the input patterns through the network and other to adapt the output by changing the weights in the network..**
The back-propagation-feed forward neural network can be used in many applications such as character recognition, weather and financial prediction, face detection etc. The paper implements one of these applications by building training and testing data sets and finding the number of hidden neurons in these layers for the best performance.
In the present research, possibility of predicting average rainfall over Udupi district of Karnataka has been analyzed through artificial neural network models. In formulating artificial neural network based predictive models three layered network has been constructed. The models under study are different in the number of hidden neurons.
Keywords: Monsoon rainfall, Udupi, prediction, artificial neural network, back propagation algorithm, multilayer artificial neural network.

I. INTRODUCTION

Weather forecasting is one of the most imperative and demanding operational responsibilities carried out by meteorological services all over the world. It is a complicated procedure that includes numerous specialized fields of know-how. The task is complicated because in the field of meteorology all decisions are to be taken in the visage of uncertainty. Different scientists over the globe have developed stochastic weather models which are based on random number of generators whose output resembles the weather data to which they have been fit.
The reason is that ANN (Artificial Neural Network) model is based on 'prediction' by smartly 'analyzing' the trend from an already existing voluminous historical set of data. Apart from ANN , the other models are either mathematical or statistical. These models have been found to be very accurate in calculation, but not in prediction as they cannot adapt to the irregularly varying patterns of data which can neither be written in form of a function, or deduced from a formula. These real-life situations have been found to be better interpreted by 'artificial neurons' which can learn from experience, i.e by back-propagation of errors in next guess and so on. This may lead to a compromise in accuracy , but give us a better advantage in 'understanding the problem', duplicating it or deriving conclusions from it.

Amongst all weather happenings, rainfall plays the most imperative part in human life. Human civilization to a great extent depends upon its frequency and amount to various scales. Several stochastic models have been attempted to forecast the occurrence of rainfall, to investigate its seasonal variability, to forecast yearly/monthly rainfall over some geographical area.

The paper endeavors to develop an ANN model to forecast average monthly rainfall in the Udupi district of Karnataka. Indian economy is standing on the summer monsoon. So prediction of rainfall is a challenging topic to Indian atmospheric scientists. Back propagation ANN to forecast the average summer monsoon rainfall over Udupi district and aroma of newness further lies in the fact that here various MANN models are attempted to find out the best fit.

II. LITERATURE SURVEY

Hu(1964) initiated the implementation of ANN, an important soft computing methodology in weather forecasting. Since the last few decades, ANN a voluminous development in the application field of ANN has opened up new avenues to the forecasting task involving environment related phenomenon (Gardener and Dorling, 1998; Hsiesh and Tang, 1998). Michaelides et al (1995) compared the performance of ANN with multiple linear regressions in estimating missing rainfall data over Cyprus. Kalogirou et al (1997) implemented ANN to reconstruct the rainfall over the time series over Cyprus.Lee et al(1998) applied ANN in rainfall

prediction by splitting the available data into homogenous subpopulations. Wong et al (1999) constructed fuzzy rules bases with the aid of SOM and back-propagation neural networks and then with the help of the rule base developed predictive model for rainfall over Switzerland using spatial interpolation. Toth et al. (2000) compared short-time rainfall prediction models for real-time flood forecasting. Different structures of auto-regressive moving average (ARMA) models, ANN and nearest-neighbors approaches were applied for forecasting storm rainfall occurring in the Sieve River basin, Italy, in the period 1992-1996 with lead times varying from 1 to 6 h. The ANN adaptive calibration application proved to be stable for lead times longer than 3 hours, but inadequate for reproducing low rainfall events. Koizumi (1999) employed an ANN model using radar, satellite and weather-station data together with numerical products generated by the Japan Meteorological Agency (JMA) Asian Spectral Model and the model was trained using 1-year data. It was found that the ANN skills were better than the persistence forecast (after 3 h), the linear regression forecasts, and the numerical model precipitation prediction. As the ANN model was trained with only 1 year data, the results were limited. The author believed that the performance of the neural network would be improved when more training data became available. It is still unclear to what extent each predictor contributed to the forecast and to what extent recent observations might improve the forecast.

Abraham et al. (2001) used an ANN with scaled conjugate gradient algorithm (ANN-SCGA) and evolving fuzzy neural network (EfuNN) for predicting the rainfall time series. In the study, monthly rainfall was used as input data for training model. The authors analyzed 87 years of rainfall data in Kerala, a state in the southern part of the Indian Peninsula. The empirical results showed that neuro-fuzzy systems were efficient in terms of having better performance time and lower error rates 5 compared to the pure neural network approach. Nevertheless, rainfall is one of the 20 most complex and difficult elements of the hydrology cycle to understand and to model due to the tremendous range of variation over a wide range of scales both in space and time (French et al., 1992).

III. ARTIFICIAL NEURAL NETWORK

A neural network is a computational structure inspired by the study of biological neural processing. There are many different types of neural networks, from relatively simple to very complex, just as there are many theories on how biological neural processing takes place.

A) Feed Forward Network
A layered feed forward neural network has layers, or subgroups of processing elements. A layer of processing elements makes independent computations on data that it receives and passes the result to another layer. The next layer may in turn make its independent computations and pass on the result to yet another layer. Finally, a subgroup of one or more processing elements determines the output of the network. Each processing element makes its computation based upon a weighted sum of its inputs. The first layer in the input layer and the last layer is the output layer. The layers that are in between these two layers are the hidden layers. The processing elements are seen units that are similar to neurons working in the brain, and hence, they are referred to as cells, neuromines, or artificial neurons. A threshold function is sometimes used to qualify the output of a neuron in the output layer. Even though our subject matter deals with artificial neurons, we will simplify them as neurons. Synapses between neurons are referred to as connections, which are represented by edges of a directed graph in which the nodes are the artificial neurons.
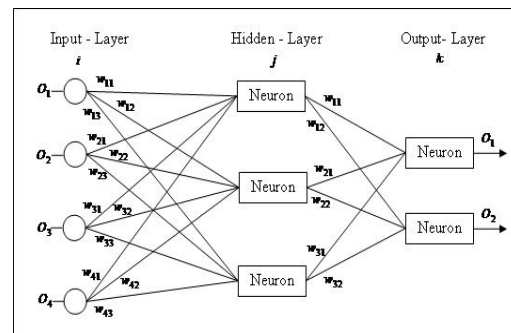
Fig.1—Layout of feed forward neural network

IV. EXPERIMENTAL SETUP

The model building process consists of four sequential steps:
1. Selection of the input and the output data for the supervised BP learning.
2. Normalization of the input and the output data.
3. Training of the normalized data using BP learning.
4. Testing the goodness of fit of the model.
5. Comparing the predicted output with the desired output.

A) Selection Of The Input And The Output Data

In Udupi, Karnataka, the months of April to November are identified as the rainfall season with May, June, July, August, and October as the main monsoon seasons. Thus the present study explores the data of these 8 months from 1960 to 2010. There will be 400 entries in the output and the input files. The Input parameters are the average Humidity and the average Wind Speed for the 8 months of 50 years from 1960-2010 making it a 2*400 matrix. In the output parameter is average

rainfall in the 8 months of every year from 1960 - 2010. The input file consists of 2 rows and 400 columns while the output file consists of 1*400 matrix. The data stated above was retrieved from www.Indiastat.com and the IMD website.

The unknown values were randomized keeping in mind the average value of the data.

B) Normalization of Data:

The input and the output data obtained have to be normalized because they are of different units and otherwise there will be no correlation between the input and the output values.

First the mean of all the data separately were taken for humidity, wind speed and rainfall.

Let the mean be M.

M= sum of all entries/number of entries

Then the standard deviation, SD, for each of these parameters individually were calculated.

Now after having the values of mean and SD for every parameter, the values for each parameter were normalized

Normalized value = (x-M)/SD

C) Training of input Data-

After obtaining the normalized data, the next step is to train the input data using matlab Back-propagation Algorith(BPA). The proposed ANN model is basically a three layered ANN back propagation learning. The algorithm takes only 70 percent of the input data for training. So out of 400 samples only 280 are taken for training and these are selected randomly from the set of data. For every attempt of training the data, the algorithm selects the training sample randomly from the whole set and not a fixed set of data and so every time you train the data, we get different values of mean square error (MSE) depending upon which 70 percent of the input data is chosen for training. Rest 60 samples are kept for validation and the remaining 60 samples are kept for testing.

D) Testing and Validation

Testing is done after the training of the data is complete and the error is below the tolerance levels. The BPA keeps 30 % of the input data for testing and validation. So out of 400 samples, 60 are used in testing and another 60 are used for validation.

E) Comparison Of Actual Data And Predicted Data

After the testing is done, the results are saved in the workspace and a graph is plotted between the actual output and the predicted output so that a comparison can be made. The graph is an efficient way of comparing the two types of data available with us. It can also be used to calculate the accuracy of the model.

In this paper when a graph was plotted between the actual and the predicted values , it showed high degree of similarity between them hence proving that our ANN model is quite accurate in prediction. The following is an example of the the graph that is plotted after the testing and validation part is over
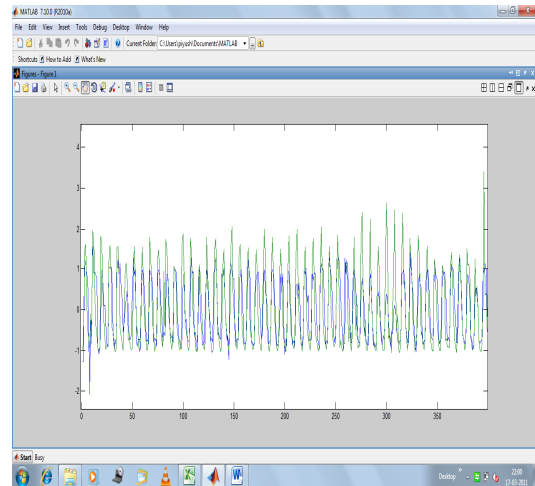
Fig. 2- Snapshot of a comparison graph between the actual data and the predicted data

The experiments were performed on the following Neural Network :
a) Feed Forward with Back –Propagation
b) Layer Reccurent
c) Cascaded feed Forward back Propagation

V. RESULT AND DISCUSSION

There are two tools for implementing the algorithms in matlab. They are-

- **Nntool – open network/data manager.** The single layer and the multi layer algorithms are implemented in the *nntool-open network/ data manager.*
- **Nftool – Neural network fitting tool.** Only back propagation algorithm is implemented in this matlab tool.

Back Propagation Algorithm (BPA) was implemented in the nftools and a minimum MSE was obtained and a graph was plotted between the predicted values and the target values.

The following are the values recorded using the nftools→

MSE = 3.6456

The regression can be plotted as follows→
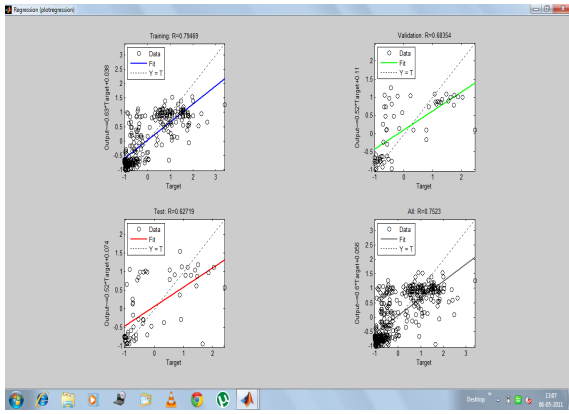
Fig. 3 Regression plot in nftools of Back-Propagation Algorithm (Single Layer)

The performance can be plotted as→
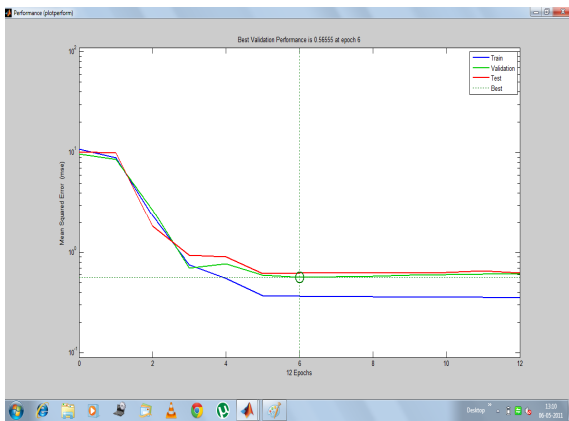
Fig .4 Performance of Back-Propagation Algorithm (Single Layer)

The predicted values by BPA in nftool can be compared with the actual values in the following graph→
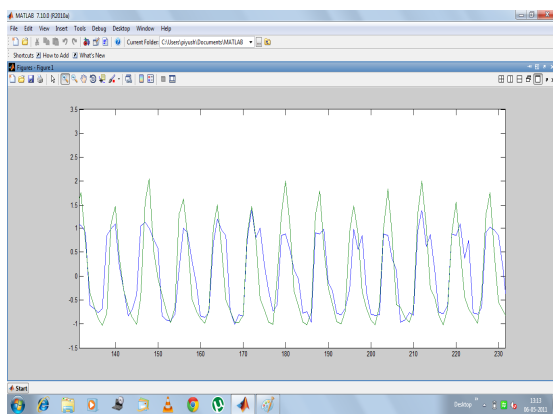
Fig 5. Back Propagation Algorithm results with target results (single-layer)

The results were fairly good and high degree of accuracy was obtained by observing the graph above. The MSE was also within the tolerance levels and the BPA algorithm turned out to be a great success.

The implementation of multi-layer architecture was done using NNTOOL in MTALAB

Three algorithm were tested in multi-layer architecture:
1. Back Propagation Algorithm (BPA)
2. Layer Recurrent Network(LRN)
3. Cascaded Back-Propagation (CBP)

The BPA implemented in the nntools showed great consistency and accuracy with the target data. The following table gives the results of different cases of BPA as observed

For multilayer architecture three hidden were used along with one input and one ouput layert and 10-20 neurons per layer were deployed to be used in experimental for Back-Propagation, Layer recurrent and Cascaded Back propagation

First Nntool was used for testing Back-Propagation Algorithm with the sample data and following obersation were made:

TABLE-I CASES FOR BPA

S.No.	Training function	Adaptive Learning Function	No. Of Neurons	Mean Square Error(MSE)
Case1	TRAINLM	Learngdm	10	0.47
Case2	TRAINLM	Learngd	10	0.52
Case3	TRAINLM	Learngdm	20	0.44
Case 4	TRAINLM	Learngd	20	0.48
Case 5	TRAINRP	Learngdm	10	0.46
Case 6	TRAINRP	Learngd	10	0.57
Case 7	**TRAINRP**	Learngdm	20	**0.42**
Case 8	TRAINRP	Learngd	20	0.46

Case 7 in the table above has the minimum MSE and so is the best case of BPA. Its performance can be plotted as follows→
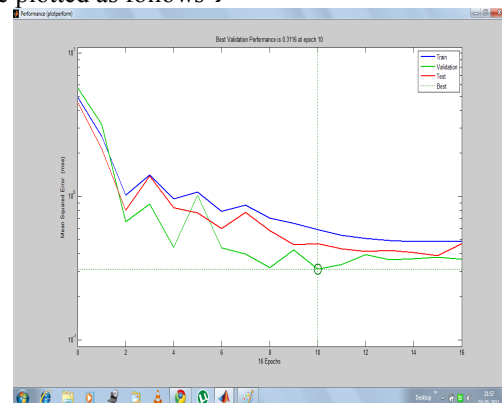
Fig 6 Back Propagation Algorithm case 7 performance

The graph is plotted between MSE and epochs. Train, validation and the test parameters are plotted against the best case. The figure clearly shows that the best validation check occurred at epoch 10.

The final result can be ploted in the following figure→
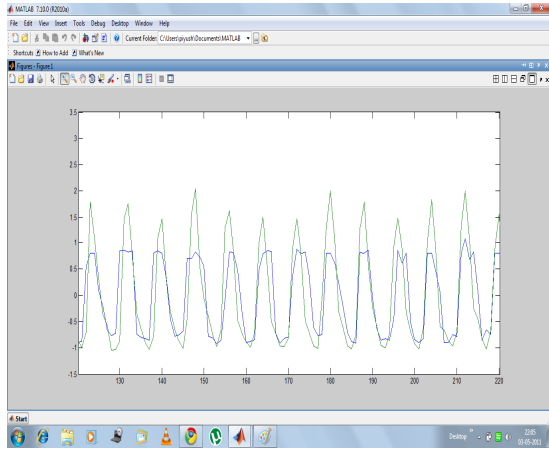
Fig 7 Back Propagation Algorithm case 7 final results with target results.

The figure above clearly shows high level of accuracy and precision. Also the predicted values follow the same trend as that of the target values and the deviation or fluctuations in the graph is the least.

Hence BPA is the best algorithm for training data in the NNTOOLs.

The results of various cases of LRN are shown below. All the cases are similar to BPA in training function used and the number of neurons tested with except the MSE which has been a bit different than that of BPA.

TABLE II LRN CASES

S.No.	Training function	Adaptive Learning Function	No. Of Neurons	Mean Square Error(MSE)
Case1	**TRAINLM**	Learngdm	**10**	**0.44**
Case2	TRAINLM	Learngd	10	0.45
Case3	TRAINLM	Learngdm	20	2.04
Case4	TRAINLM	Learngd	20	0.49
Case5	TRAINRP	Learngdm	10	1.48
Case6	TRAINRP	Learngd	10	1.21
Case7	TRAINRP	Learngdm	20	0.47
Case8	TRAINRP	Learngd	20	1.30

Clearly the best case in the above figure is case 1 with the least MSE. Its performance can be plotted as follows➔
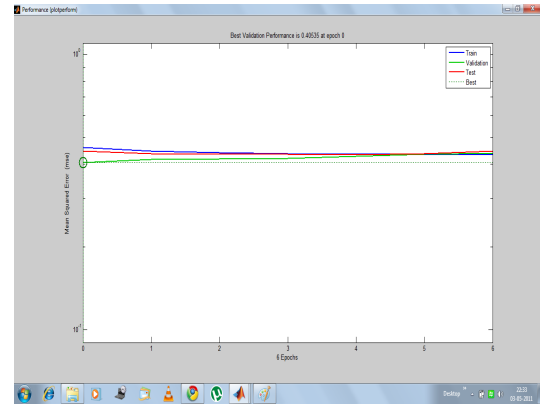
Fig 8 Layer Recurrent case 1 performance

This figure shows the best case of LRN. The best validation check occurs at epoch 0.
The final results of case 1can be plotted as follows➔
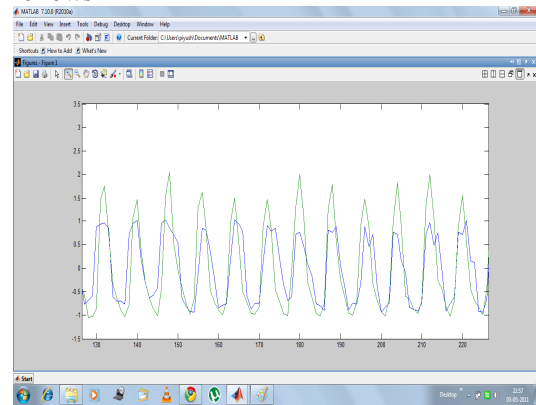
Fig 9 LRN case 1 final result with target result

This algorithm also shows high level of accuracy but in some cases the MSE was quite high as compared to BPA.
The different cases of the CBP algorithm have been listed in the given table.

TABLE III CBP CASES

S.No.	Training function	Adaptive Learning Function	No. Of Neurons	Mean Square Error(MSE)
Case 1	TRAINLM	Learngdm	10	0.50
Case 2	TRAINLM	Learngd	10	0.52
Case 3	**TRAINLM**	**Learngdm**	**20**	**0.47**
Case 4	TRAINLM	Learngd	20	0.61
Case 5	TRAINRP	Learngdm	10	0.54
Case 6	TRAINRP	Learngd	10	1.47
Case 7	TRAINRP	Learngdm	20	0.50
Case 8	TRAINRP	Learngd	20	1.52

The best case of CBP algorithm is case 3 with the least MSE. Its performance is plotted as follows➔
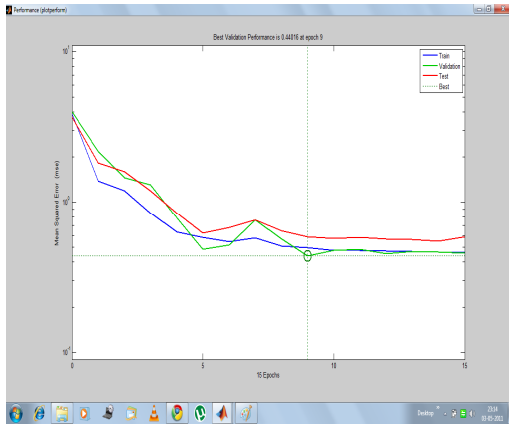
Fig 10 Cascaded Backprogation Algorithm case 3 performance

The best validation check is at epoch 9. The final results can be plotted as follows→
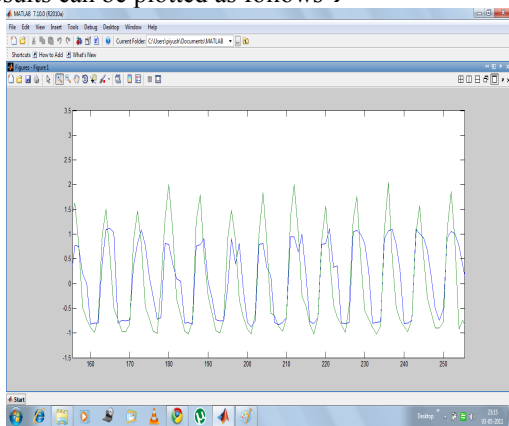
Fig 11 Cascaded Backprogation Algorithm case 3 final results with target results

The CBP algorithm showed high MSE in almost every case as compared to BPA or LRN.
Also, its results were deviating from the fact that as number of neurons increases, the MSE decreases.

VI. CONCLUSION

From the experimental study the following observation were made :

- As the number of neurons increases in an ANN, the MSE decreases.

- BPA is the best algorithm out of the three tested.

- LEARNGDM is the best learning function to train your data with.

- LEARNGD is a bit time consuming.

- TRAINLM is the best training function.

- Multi-layer Alogorithm   is better than Single-layer algorithm terms of performance.

- NNTOOLS should be used to implement the prediction algorithms as a it gives an option of implementing algorithms other than BPA.

- Larger the amount of input data, lower is the MSE after training.

- The input /output data should be normalized if they are of very high order.

REFERENCES

[1] M.J.C., Hu, Application of ADALINE system to weather forecasting, Technical Report, Stanford Electron, 1964
[2] Michaelides, S. C., Neocleous, C. C. & Schizas, C. N. "Artificial neural networks and multiple linear regression in estimating missing rainfall data." In: Proceedings of the DSP95 International Conference on Digital Signal Processing, Limassol, Cyprus. pp 668–6731995.
[3] Kalogirou, S. A., Neocleous, C., Constantinos, C. N., Michaelides, S. C. & Schizas, C. N.,"A time series construction of precipitation records using artificial neural networks. In: Proceedings of EUFIT '97 Conference, 8–11 September, Aachen, Germany. pp 2409–2413 1997.
[4] Lee, S., Cho, S.& Wong, P.M.,"Rainfall prediction using artificial neural network.",J. Geog. Inf. Decision Anal. 2, 233–242 1998.
[5] Wong, K. W., Wong, P. M., Gedeon, T. D. & Fung, C. C. , "Rainfall Prediction Using Neural Fuzzy Technique." 1999
[6] E.Toth, A.Brath, A.Montanari," Comparison of short-term rainfall prediction models for real-time flood forecasting", Journal of Hydrology 239 (2000) 132–147
[7] Koizumi, K.: "An objective method to modify numerical model forecasts with newly given weather data using an artificial neural network", Weather Forecast., 14, 109–118, 1999.
[8] N. Q. Hung, M. S. Babel, S. Weesakul, and N. K. Tripathi "An Artificial Neural network Model for rainfall Forercastingin Bangkok,Thailand", Hydrol. Earth Syst. Sci., 13, 1413–1425, 2009
[9] Kyaw Kyaw Htike and Othman O. Khalifa, "Research paper on ANN model using focused time delay learning", International Conference on Computer and Communication Engineering (ICCCE 2010), 11-13 May 2010, Kuala Lumpur.
[10] Ajith Abraham, Dan Steinberg and Ninan Sajeeth Philip,"Rainfall Forecasting Using Soft Computing Models and Multivariate Adaptive Regression Splines", 2001.
[11] Rainfall data --http://www.Indiastat.com/karnataka/rainfall
[12] Ben Krose and Patrick van der Smagt , "An introduction to neural networks", Eighth edition, November 1996.
[13] Dr S. Santosh Baboo and I. Khadar Shareef, "An efficient Weather Forecasting Model using Artificial Neural Network", International Journal of Environmental Science and Development, Vol. 1, No. 4, October 2010.
[14] Enireddy Vamsidhar et. al.,"Prediction of rainfall Using Back-propagation Neural Network Model", International Journal on Computer Science and Engineering Vol. 02, No. 04, 2010, 1119-1121
[15] Paras, Sanjay Mathur, Avinash Kumar, and Mahesh Chandra, "A feature based on weather prediction using ANN"World Academy of Science, Engineering and Technology 34 2007

Sizing Ratio Of Inverter and PV Array For a-Si FS GCPV System in Malaysia's Perspectives

M. Z. Hussin, A. M. Omar, Z. M. Zain
Faculty of Electrical Engineering
Universiti Teknologi MARA
40450 Shah Alam, Selangor Darul Ehsan
mohamadzhafran@gmail.com

S. Shaari
Faculty of Applied Sciences
Universiti Teknologi MARA
40450 Shah Alam, Selangor Darul Ehsan
Malaysia

Abstract—**System design's sizing and natural environmental analysis are two common issues and considered to be among the most crucial part of the study with the objective of enhancing the Grid-connected Photovoltaic (GCPV) performance in Malaysia's perspective. This a-Si GCPV system located near to the Green Energy Research Centre (GERC) Building at Universiti Teknologi MARA (UiTM), Shah Alam, Malaysia. This a-Si GCPV system installation was referred to Malaysian Standard MS1837:2010 and designed based on initial data values of PV module characteristics. To date, various of GCPV system applications using thin-film (TF) PV module technologies in Malaysia were designed at stabilized data value. So, this system is conducted to study the derating factor for matching inverter and PV array sizing and behavior of a-Si FS GCPV system under Malaysia's real condition. The P-V analysis fully utilizes the in-field data as a guidelines for sizing inverter and PV array, with the aim of guiding the system designer during initial design stage for Malaysia's climate.**

Keywords- amorphous-Silicon (a-Si); Free-standing (FS); Grid-connected inverter; Initial system design; Photovoltaic (PV); Thin-film (TF)

I. INTRODUCTION

Debates among the PV experts in terms of economical viability for optimization of the system design are still ongoing to date. Some of the researcher introduces a rule-of-thumb method to quantify the optimum sizing and other researcher's come out with their concrete explanation to estimate the desirable sizing factor in their system respectively. However, this problem must be decided whether to achieve the high-energy yield for economic gain in terms of system cost installation and energy payback period. Though, in view of these days, the consideration on the system design become more crucial in order to make sure the system is reliable and optimized. Therefore, the system designer should take precedence the reasonable decision on the selection on the BOS components during initial design stage.

The current practice of under-sizing the inverter's nominal power with respect to the PV array power may not be the best choice for thin-film PV module technologies under hot and tropical than a cooler climate because of the different irradiation level and operating temperature based on the geographical site[1,2]. Previously research for cooler climate was suggested that under-sizing up to 40% of the inverter nominal power with respect to PV array power is the best choice to maximize the energy injected to the utility grid [2,3,4,5]. However, the "rule of thumb" is widely used that

approximately 10% to 40% under-sizing of inverter power is not suitable especially in the different geographical location and local climate condition[1,4]. This rationale contradicts to the "rule-of-thumb" for cooler climate, the inverter nominal power can be considered lower than the PV array power at Standard Test Condition (STC), which might lead to energy losses, especially operating at warm and tropical climate such as in Brazil, where there is a high incidence of clear skies and the energy distribution of sunlight was shifted to higher irradiation levels [1,6,7] and the temperature influenced by a smaller of temperature coefficients of power such as thin-film modules made thin-film module reacts with a better performance than crystalline technology in hot and tropical climate. Typically, thin-film modules present a negative temperature coefficient of power value of $\gamma_{Pmp} \approx$ -0.20%/°C (a-Si) until $\gamma_{Pmp} \approx$ -0.38%/°C (CIGS) [8,9]. The size of the inverter chosen either too small or too large with respect to the PV array rating will give the direct impact to the considerable energy losses will occur at high irradiance levels due to inverter clipping or the limited power output, which required to avoid damage to the inverter, whereas, if the size of the inverter is too small it will lead to a lower efficiency during the low irradiance levels which increased the cost of the inverter prices and reduce the financial viability of system. Besides that, the effects of the initial module degradation were taken in consideration in terms of the sizing ratio especially in the thin-film PV module technologies[5,10,11].

The paper represents a P-V analysis technique which can help in determining the preliminary of power sizing in terms of GCPV system design and the results for one-year period at outdoor field of a 1.179 kWp FS GCPV system using a-Si PV modules technology located at UiTM Shah Alam, Malaysia. Data were collected on June 2011 until May 2012 and analysed to estimate the sizing derating factor for inverter and PV array power of a-Si FS GCPV system in Malaysia's field.

II. SYSTEM DESCRIPTION

A. Charcateristics of 1.179 kWp a-Si FS GCPV system

The a-Si GCPV system is located at the latitude of 3° 04' 08.79''N and longitude of 101° 29' 49.66''E and the installation meet all requirements of the Malaysian Standard (MS 1837:2010) for installation of Grid-connected Photovoltaic (PV) system [12]. This a-Si FS GCPV system installation was designed based on initial data values of PV module characteristic used a-Si PV modules with nominal power capacity of about 1.179 kWp as shown in TABLE II.

TABLE I. A-Si FS GCPV SYSTEM DESCRIPTIONS.

System	Description
Location	Green Energy Research Centre, Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM)
City / Country	Shah Alam, Malaysia
Tilt angle / orientation	15° / South
System / Mounting type	Grid-Connected/ Free-standing
Array Configuration	5 parallel x 3 series
Nominal power	1.179 kWp (15 Units)

TABLE II. KANEKA K60 PV MODULE SPECIFICATIONS [13].

PV module	Specification	
PV technology	Thin-film (TF)	
PV module type	Amorphous Silicon (a-Si)	
Model	Kaneka K60	
	Initial value	Stabilized value
Pmp_stc (W)	78.6	60
Voc_stc (V)	95.6	92
Isc_stc (A)	1.22	1.19
Vmp_stc (V)	74	67
Imp_stc (A)	1.06	0.9
Module efficiency (%)	n.a	6.31

TABLE III. GRID-CONNECTED INVERTER SPECIFICATIONS [14].

Inverter	Subject	Specification
Input	Max. DC power	1320 W
	Max. DC voltage	400 V
	MPP– voltage range	100 V- 320 V
	MPP– voltage range	100 V- 320 V
	Max. input current	12.6 A
Output	Max. AC power	1200 W
	Nominal AC output	1200 W
	Max. output current	6.1 A
Efficiency	Max. efficiency/ Euro-eta	92.1%/90.9%
General Data	Operating temperature range	–25 °C … +60 °C
	Topology	LF transformer

III. METHODOLOGY

The monitoring system complies to meet the requirements of IEC 61724 international standard [15] and within the International Energy Agency Photovoltaic Power System (IEA-PVPS) Program Task 2 framework [16]. The electrical system installation procedure of this GCPV system was conducted by a certified electrician and referred to the Malaysian Standard: MS 1837:2010. An irradiance sensor and two platinum sensors (PT100) for measuring temperature are used to obtain the meteorological parameters for the purpose of performance monitoring. For general data acquisition, SMA system monitoring measurement are presented in TABLE IV. The parameters being monitored as shown below;

TABLE IV. PARAMETERS MEASUREMENT OF a-Si FS GCPV SYSTEM.

Electrical	Meteorological
DC current, DC voltage	Total irradiance, G (W/m^2)
AC current, AC voltage	Ambient temperature, T_{am} (°C)
AC power, P_{ac} (W)	Module temperature, T_m (°C)
Array output energy, E_A (kWh)	

By multiplying the peak rating of the PV array power at STC, kWp , peak sun hour (PSH), h for specific tilt angle and de-rating factors that affected the PV performances, the predicted energy yield, E_{sys}, expressed in kWh of the system can be obtained by following equation;

$$E_{sys} = P_{array\_stc} \times f_{de\_rating} \times PSH_{pa} \qquad (1)$$

Besides that, f_{de_rating} calculated using equation (2);

$$f_{de\_rating} = f_{temp} \times f_{mm} \times \eta_{inv} \times \eta_{pv\_inv} \times f_{dirt} \times f_{aging} \qquad (2)$$

Where f_{de_rating} is the expected total amount of reduction factor in various losses involved and affected the power output generated by the PV array are included in f_{de_rating} which can be expressed in equation (2). This amount is very important to estimate the energy supplied from PV system to energy required after a specific year. In addition, a good system designer will be taking the aging effect factor in the energy generation over the lifetime of the system. Among those parameters, the de-rating factor value of the inverter-to-PV array power, k depends on the type of PV module technologies used. It is an important parameter that is needed by matching the inverter and PV array for sizing purposes. The inverter nominal power, P_{nom_inv} can be obtained by calculating the peak power rating of PV array at STC, by multiplying with the value of k. The ratio between these two powers is called as "Inverter-to-PV array" and is defined as;

$$P_{nom\_inv} = k \times P_{array\_stc} \qquad (3)$$

Where;

P_{array_stc} Peak power rating of the PV array at STC, in kWp
P_{nom_inv} Inverter nominal power, in kW
k Inverter-to-PV array power of de-rating factor, dimensionless

In Malaysia's climate, the derating factor of 0.75–0.80 for crystalline modules. However, for thin-film modules, the de-rating factor for inverter-to-PV array ratio range is recommended between 1.0 to 1.3 based stabilized system design [17]. It is very important to determine the appropriate de-rating value for inverter-to-PV array power sizing to avoid any power losses in the PV system. An optimal inverter-to-PV array power can help giving the best estimation for matching inverter and PV array design in terms of averting energy wastage and maintain minimal cost of inverter from over-sizing or under-sizing when comparing with the PV array power.

Generally, about 51367 continuous data were taken for this system with average samples for every 5- minutes via SMA system monitoring. All data set were collected and analyzed for one-year period range from June 2011 until May 2012.

IV. RESULTS AND DISCUSSION

A. Solar Irradiance And Temperature Analysis

The power generates depend mainly on the irradiance and temperature at PV site installation. So, the geographical distribution at the specific site is very important parameters to identify and gather related information such as solar irradiance and temperature before any PV installation implemented. Figure 1 below shows the distribution of irradiance and average operating module temperature throughout the period monitored.
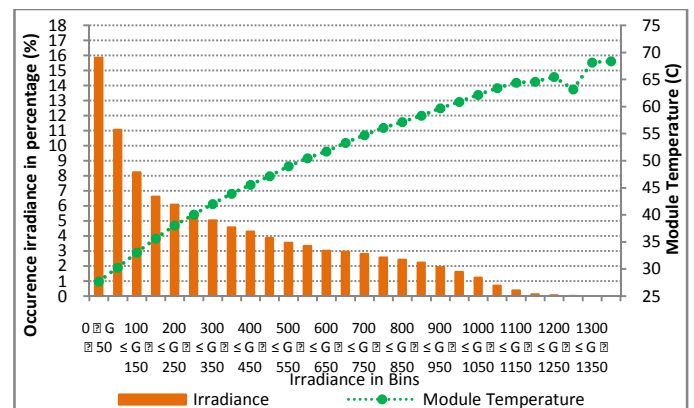
Figure 1. Distribution occurrence of the irradiance in bins representing in percentage (June 2011 – May 2012).

Figure 1 represents the distribution of the irradiance received and average operating module temperature using 5-minute averages. Each bar corresponds to the percentage of the irradiance content in interval width of 50 W/m^2. Looking at the high peak of the irradiance level shows that only some 4% of the daytime have instantaneous irradiance levels more than G ≥ 950 W/m^2. Within one-year period's data were recorded, 96% of the irradiance percentage occurred lies in the range 900-950 W/m^2 with an average for module temperature operated at 59.7°C.
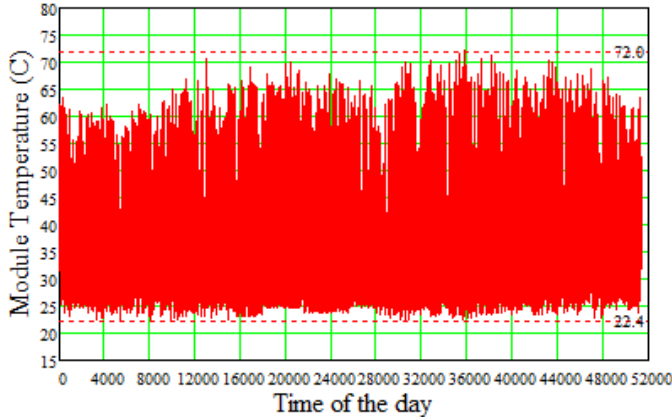
Figure 2. The distribution of operating module temperature over monitored period (June 2011-May 2012).

Referring to Figure 2 above, the results indicate that the maximum operating module temperature was measured at 72.0°C and the minimum about 22.4°C. In current practice in Malaysia's climate, it is recommended that the highest expected cell temperature, T_{cell_max} be set at 75°C and the lowest expected cell temperature, T_{cell_min} at 20°C [17], which is commonly used in the morning to ensure the inverter operates well and safely at its maximum inverter efficiency during peak irradiance occurred in Malaysia's field condition.

However, the lowest of operating module temperature recorded shown that the operating module temperature had not reach less than 20°C throughout the one-year period. So, it is recommended as the lowest expected cell temperature for T_{cell_min} set to be 20°C. Other than that, T_{cell_max} = 75°C is recommended to use for safety reason because the measured data for operating module temperature over one-year period is reached as high as 72.0°C at this location.

B. Derating factor of AC Inverter-to-PV array ratio (System Performance)

Figure 3 explains the relationship between power output and operating array voltage over the monitored period. This system is purposely designed to assess the behaviour of a-Si PV modules and estimate the actual value of the derating factor in terms of the matching the inverter with respect to the PV array power, k. So, the inverter was designed with oversized around 102% of the PV array power to determine on how much power generates by a-Si modules under Malaysia's climate. In addition, this system was designed based on initial data values to know the reliability and durability to handle the behaviour of initial higher in power level of a-Si thin-film that the typically process of stabilization period claimed more than one-year in hot and humid climate [6,7]. Most of the a-Si PV modules in GCPV system in Malaysia were designed based on stabilized data values.

Based on calculation design for a-Si FS GCPV of a-Si thin-film modules after considering power losses about k_design 1.02 is higher than the actual value is around k_actual = 0.74 in June, 2011. In other words, the actual value of inverter-to-PV array ratio, k is ~74% during June, 2011. It meant that the inverter is oversized than the peak power rating

of PV array, hence the inverter operates well without clamping the power produced by PV array with daily inverter efficiency ranged from 90.5% to 91.2%. The "clamp" phenomenon of the power will occur if the power generated exceeds than k_design = 1.02 of the inverter power sizing.

If the operating array voltage is lower than the minimum window voltage of the inverter, $V_{min\_win\_inv}$, it will automatically turn off itself. In addition, the operating array open-circuit voltage, $V_{oc\_array\_stc}$ of the PV array should be never greater than the maximum window voltage of the inverter, $V_{max\_win\_inv}$ to avoid the maximum open circuit voltage harm the inverter components. The input MPP-voltage limit of the SMA SB 1200 inverter's ranges $100V_{dc} - 320V_{dc}$, whereas the maximum input DC voltage of the inverter, $V_{max\_inv\_input}$ about $400V_{dc}$ [14]. The sizing informations are tabulated in TABLE V below.

TABLE V. A-SI FS GCPV SYSTEM SPECIFICATIONS.

Inverter/PV array	Specification
$P_{nom\ inv}/P_{array\ stc}$ (k_design)	1.02
$V_{max\ inv\ input}/V_{mp\ array\ stc}$	1.80
$V_{max\ win\ inv}/V_{mp\ array\ stc}$	1.44
$V_{min\ win\ inv}/V_{mp\ array\ stc}$	0.45
$V_{oc\ stc\ array}/V_{mp\ array\ stc}$	1.29
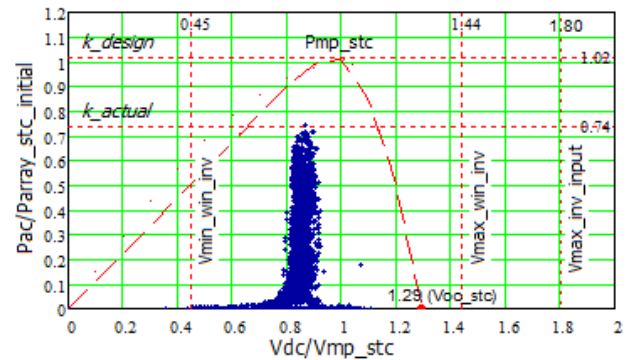
Figure 3. Monthly AC power-voltage characteristic of a-Si thin-film modules in June 2011.
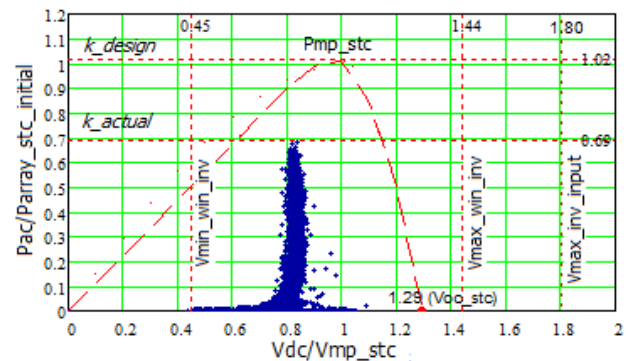
Figure 4. Monthly AC power-voltage characteristic of a-Si thin-film modules in July 2011.
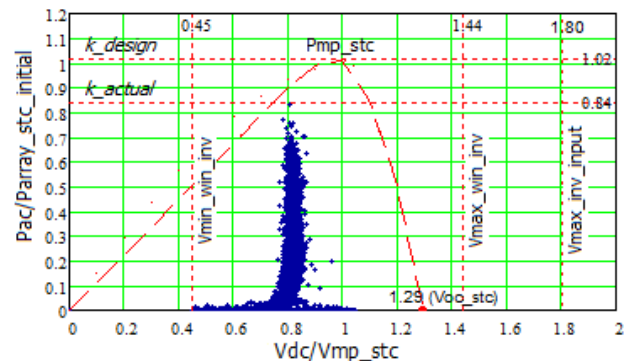
Figure 5. Monthly AC power-voltage characteristic of a-Si thin-film modules in August 2011.
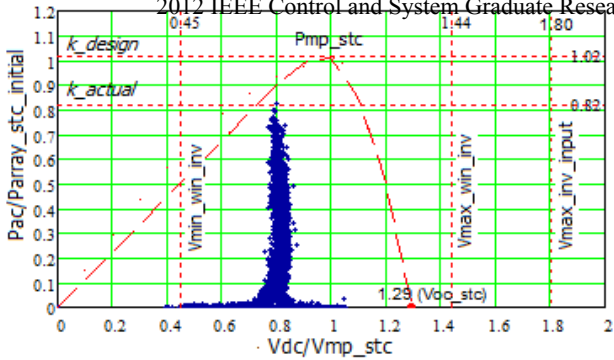
Figure 6.   Monthly AC power-voltage characteristic of a-Si thin-film modules in September 2011.
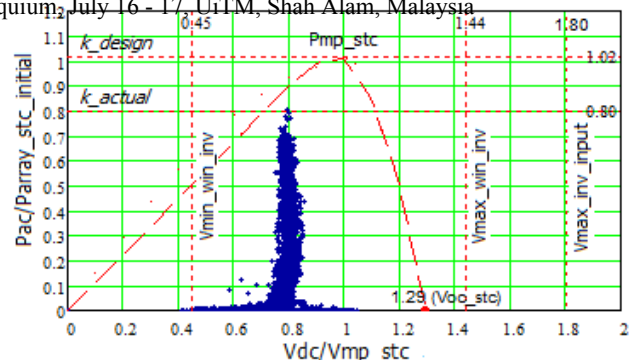
Figure 10. Monthly AC power-voltage characteristic of a-Si thin-film modules in January 2012.
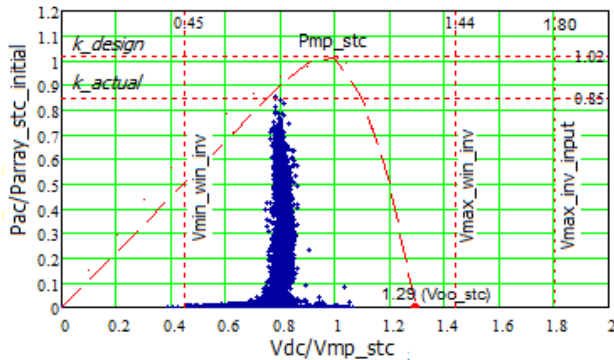
Figure 7.   Monthly AC power-voltage characteristic of a-Si thin-film modules in October 2011.
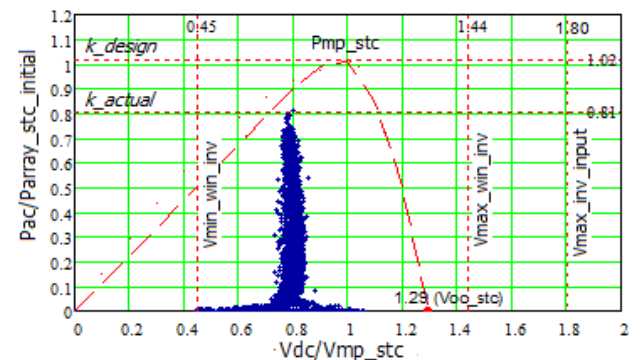
Figure 11. Monthly AC power-voltage characteristic of a-Si thin-film modules in February 2012.
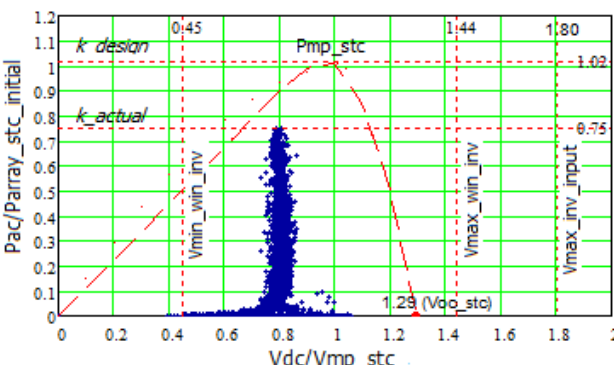
Figure 8.   Monthly AC power-voltage characteristic of a-Si thin-film modules in November 2011.
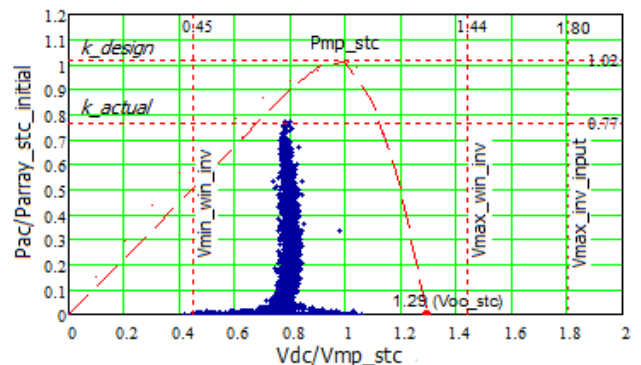
Figure 12. Monthly AC power-voltage characteristic of a-Si thin-film modules in March 2012.
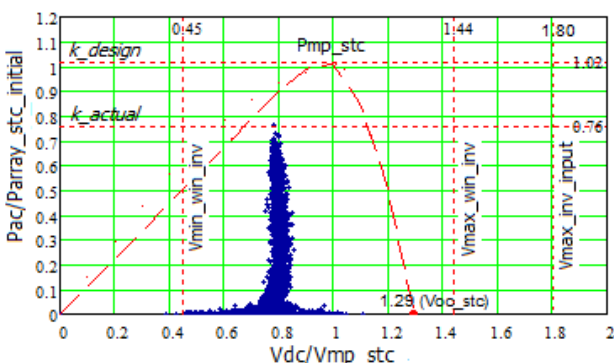
Figure 9.   Monthly AC power-voltage characteristic of a-Si thin-film modules in December 2011.
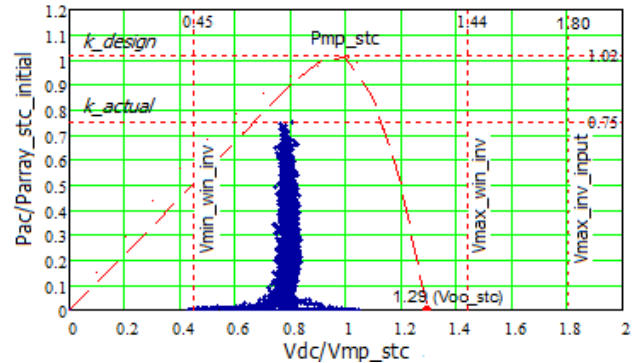
Figure 13. Monthly AC power-voltage characteristic of a-Si thin-film modules in April 2012.
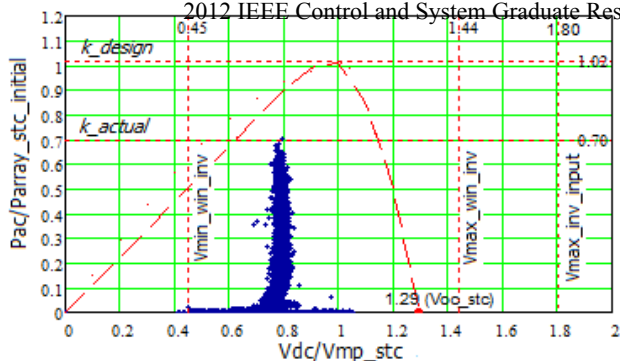
Figure 14. Monthly AC power-voltage characteristic of a-Si thin-film modules in May 2012.

From the scatter graph in Figure 3 - 14, it can be seen that the scatter data point for operating DC voltage was operating safe within the allowable window voltage range of the input inverter MPP-voltage limit. It is very important to remain the operating DC voltage fall within the MPP-voltage range to ensure the inverter work at it best efficiency under all expected weather conditions.

C. Derating factor of DC Inverter-to-PV array ratio (Module Performance)

As PV module and PV array output are degraded on the performance over the lifetime of the PV system, the designer must be careful to keep the array voltage matched within the rated range of the chosen inverter. Apart from that, the highest power was generated by a-Si PV modules during the month of September 2011 and October 2011 with 0.91 and 0.94 respectively, i.e. the actual value derating factor of k_{dc} about 91% and 94% as shown in Figure 15 and Figure 16.
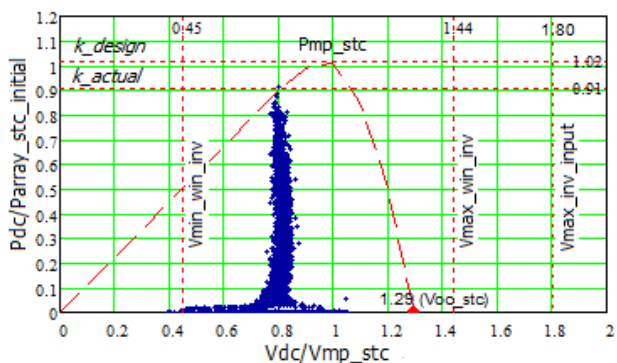
Figure 15. DC power-voltage characteristic of a-Si thin-film modules in September 2011
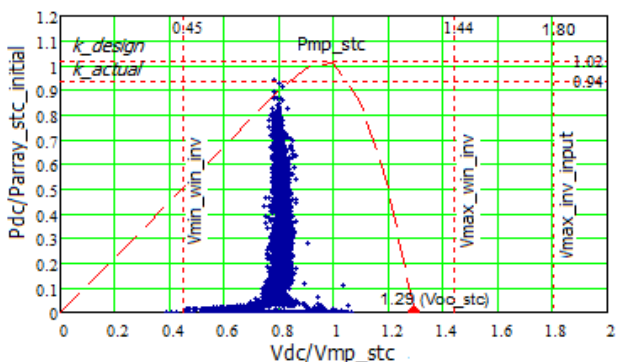
Figure 16. DC power-voltage characteristic of a-Si thin-film modules in October 2011.

D. Summary of AC/DC Inverter-to-PV array ratios

The derating factor actual values in terms of AC/DC inverter-to-PV array ratios are summarized in TABLE VI.

TABLE VI. DERATING FACTOR FOR INVERTER-TO-PV ARRAY RATIOS BASED INITIAL DATA.

MONTH	MONTHLY IRRADIATION (KWH/M²)	AC INVERTER-TO-PV ARRAY RATIO (%)	DC INVERTER-TO-PV ARRAY RATIO (%)
JUNE	109.85	0.74	0.82
JULY	123.87	0.69	0.75
AUGUST	126.76	0.84	0.92
SEPTEMBER	127.05	**0.82**	**0.91**
OCTOBER	127.33	**0.85**	**0.94**
NOVEMBER	113.80	0.75	0.83
DECEMBER	103.44	0.76	0.84
JANUARY	133.85	0.80	0.89
FEBRUARY	124.63	0.81	0.90
MARCH	126.83	0.77	0.85
APRIL	134.18	0.75	0.83
MAY	121.79	0.66	0.72

To date, many inverters have equipped with higher efficiency. By multiplying the derating factor of module performance, k_{dc} with maximum efficiency of the inverter selected. The system designer can obtain the derating factor for overall performance within the system, k_{ac} including power losses. In calculating the overall system for derating factor, k_{ac} with different efficiency of the inverter selection, k_{ac} can be calculated using equation (4);

$$k_{ac} = k_{dc} \times \eta_{inv} \qquad (4)$$

Where;

η_{inv} Maximum efficiency of the selected inverter, in decimal.

E. Data Availability

The a-Si GCPV system was operated for one-year period with data availability of the system achieved of 99%. By the end of May 2012, this a-Si PV system was operated for about 362 days out of 366 days since commissioning as tabulated in TABLE VII below. The a-Si FS GCPV system was commissioned on 3rd June 2011 and the overall system performance is still monitored until today.

TABLE VII. DATA AVAILABILITY OF THE SYSTEM OVER MONITORED PERIOD.

Year	Month	Data availability
2011	June	28 operating days
	July	31 operating days
	August	31 operating days
	September	30 operating days
	October	31 operating days
	November	30 operating days
	December	31 operating days
2012	January	31 operating days
	February	29 operating days
	March	30 operating days
	April	30 operating days
	May	30 operating days

V. CONCLUSION

The optimal sizing of an inverter with respect to the peak power of the PV array rating at Standard Test Conditions (STC) depends on the geographical and climate. It shown that the sizing values are differently between months based on solar irradiation, and temperature received. In this system, the optimum inverter-to-PV array ratio, k_{ac} were ranged in between 0.82 and 0.85 after considering power losses. If the system designer proposes with different of inverter efficiencies, the recommended values of the inverter-to-PV

array ratio, k_{dc} for thin-film modules before excluded power losses were ranged from 0.91 to 0.94. To get the values of k_{ac}, the system designer can use the equation (4). Other than that, for the lowest expected cell temperature T_{cell_min} is recommended to be set at 20°C and T_{cell_max} = 75°C because the operating module temperature for free standing type is lower than building-integrated normally set to be at 80°C. On top of attaining experience on the sizing, design and installation of GCPV system, this research study aims at performing an analysis on the behaviour of one of the thin-film PV module technologies available in order to establish the optimized GCPV system under Malaysia's field.

ACKNOWLEDGMENTS

REFERENCES

[1] B. Burger, and R. Rühter, "Inverter sizing of grid-connected photovoltaic systems in the light of local solar resource distribution characteristics and temperature". Solar Energy 80 (2006), page 32-45. 2006.

[2] S. Chen, P. Li, D. Brady, and B. Lehman, "The impact of irradiance time behaviors on inverter sizing and design", IEEE Compel Workshop, Boulder, CO, June 2010.

[3] J. D. Mondol, Y. G. Yohanis, and B. Norton , "Optimal sizing of array and inverter for grid-connected photovoltaic systems",Solar Energy Vol. 80, Issue. 12, pp. 1517-1539, 2006.

[4] I. Balouktsis, J. Zhu, R. Bründlinger, T.R. Betts, and R. Gottschalg. "Optimised inverter sizing in the UK". Presented at 4th Photovoltaic Science Application and Technology (PVSAT-4) Conference, University of Bath, 2-4 April 2008.

[5] J. Zhu, R. Brundlinge, T. R. Betts, and R. Gottschalg, "Effect Of Module Degradation On Inverter Sizing", 33rd IEEE Photovoltaic Specialists Conference, San Diego, CA, pp. 1 – 6, 11-16 May 2008.

[6] R. Rüther, H.G Beyer, A.A Montenegro, M.M Dacoregio, I.T Salamoni, P. Knob. "Performance results of the first grid-connected, thin-film PV installation in Brazil: temperature behaviour and performance ratios over six years of continuous operation". 19th European Photovoltaic Solar Energy Conference, Paris, France, pp. 3091–3094, 2004

[7] R. Ruther, Arthur, J.G. da Silva, A. A. Montenegro, I. T. Salamoni, M. Kratzenberg and R. G. Araujo, "Assessment of thin-film technologies most suited for BIPV applications in Brazil: the Petrobras 44 kWp project", 3rd World Conference on Photovoltaic Energy Conversion, Osaka, Japan, May 11-18 , 2003.

[8] Amplesun Solar Technology. Product specification for ASF 100. [Online].Available:http://www.amplesun.com/data/article/ASF100%20Specification.pdf

[9] Q-Cell SE, Product specification for Q.SMART UF 75-95 .[Online]. Available: http://www.q-cells.com/uploads/tx_abdownloads/files/Q-Cells_QSMART_UF_G1-3_Data_Sheet_EN_2011-09_Rev03_WEB_03.pdf

[10] Sharp, Installation Manual   Photovoltaic Modules NA-Series, KTA09001 Ver. 3.11.

[11] Solarex, Millennia MST-56, -50 & -43 Photovoltaic Modules. 6087-2 (issue 4/97)

[12] Malaysian Standard MS1837:2010. Installation of Grid-Connected Photovoltaic (PV) Systems, 2010.

[13] IBC Solar AG. Product information sheet for Kaneka K60. [Online]. Available: http://www.ibc-solar.com/ (issue 08/2008)

[14] SMA Solar Technology AG. Technical Data on SMA SB Series of String Inverter. [Online]. Available: http://www.sma.de/en (issue 12/09).

[15] International standard IEC 61724, Photovoltaic system performance monitoring-guidelines for measurement, data exchange and analysis, 1998.

[16] IEA-PVPS Task 2, Analysis of Photovoltaic Systems. Report IEA-PVPS T2-01, 2000. [Online]. Available: http://www.iea-pvps-task2.org/

[17] S. Shaari, A.M Omar, A.H. Haris and S.I. Sulaiman, Solar Photovoltaic power: designing Grid-Connected systems, 1st Edition, Ministry of Energy, Green Technology and Water, Malaysia /MBIPV, Malaysia, 2010

Femtocell Interference Mitigation

Al-Hareth Zyoud, Mohamed Hadi Habaebi, Jalel Chebil and Md. Rafiqul Islam

Electrical and Computer Engineering Department, Faculty of Engineering,
International Islamic University Malaysia (IIUM),
53100 Gombak, Kuala Lumpur, Malaysia
Email: al7_z@yahoo.com

Abstract— **The demand for the Femtocells deployment is in the rise since it enhances the wireless link and increase the capacity for the wireless network. Femtocell is a home base station designed for use in residential and business areas such as houses, apartments, offices, schools, and markets. It connects to the service provider's network via broadband and normally supports 2 to 6 mobile phones in a small area. Using femtocell will enhance the indoor coverage, save power since it reduces the transmitted power, and in the same time reduce the traffic on microcell. However, the most important challenge in femtocell deployment is interference mitigation and management. In this paper, a state of the art literature review on mitigating interference is presented. First, interference in femtocell environment is classified from different point of views. Then, open issues for femtocell interference research are discussed.**

Keywords- Femtocell, Macrocell, Interference Manaement, Frequency reuse, Cognitive Radio, Resource Management, OFDMA.

I. INTRODUCTION

The rapid development of smart phones and tablet PCs has caused the shifting of Cellular networks from being essentially voice networks to become mostly data networks. This leads to unpredictable growth in global mobile and internet data traffic rate. It is expected that monthly global mobile and tablet data traffic will exceed 11 exabytes in 2016 [1]. Moreover, the attenuation of cellular signal inside buildings reduces the signal strength and makes the signal not available in some regions, since the macrocell signal does not penetrate walls. Thus, to overcome the increased in the traffic and to solve the indoor coverage problem, small cells called femtocells are introduced.

Femtocell is a small range (10-30 m), low-power access point (10-100 mW) that can be installed by end users to connect them to the network via the internet as shown in Figure 1 [2, 3]. It has been introduced as a solution to the indoor coverage problems, since it increases the network's coverage and capacity, and reduces both cost and power. Reducing the transmission distance improves the link capacity, and save the mobile battery. Moreover deployment of femtocells in an existing macrocell reduces the indoor traffic, thus the network capacity and reliability will improve. However, there are many challenges that face the success of deployment of femtocell. One of the most important challenges is the interference. There are two types of interference in femtocell networks; the first one is interference between femtocell and macrocell which called cross-layer interference, and the other one is the interference between femtocell and others femtocells that deployed in the coverage of the main macrocell, this type of interference called co-layer interference.

Since January 2012, eight of the top 10 mobile operator groups provide femtocell services [4]. By 2014 there will be 114 million worldwide users on 50 million home base stations [5]. Femtocell is designed to work with a wide range of interface standards [6] such as Global System for Mobile Communications (GSM), Wideband Code Division Multiple Access (WCDMA), Universal Mobile Telecommunications System (UMTS), Long Term Evolution (LTE) which uses Orthogonal Frequency Division Multiple Access (OFDMA), and Worldwide Interoperability for Microwave Access (WiMAX).
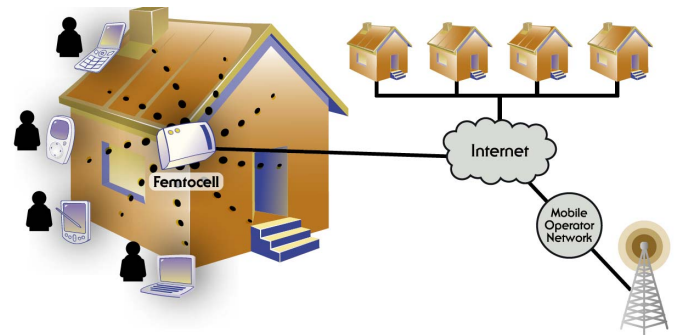
Figure 1: Femtocell network architecture [3].

Generally, femtocell operates in three different access modes [7]. Closed Subscriber Group (CSG), where users defined by the femtocell owner have the right to connect to the femtocell. This mode is normally applied in residential scenarios. Open Subscriber Group (OSG), where any user can connect to the femtocell. This mode is applied in public places such as airports and shopping malls. Hybrid Subscriber Group (HSG), where nonsubscriber users can only connect to the femtocell in emergency cases.

Normally two deployment configurations are considered in the use of femtocell in an existing macrocell: the orthogonal and the co-channel approaches. In orthogonal approach, the available spectrum is divided between the femtocell and the macrocell. Although this approach completely eliminates the cross-layer interference, it is ineffective in terms of spectrum reuse. Therefore, the co-channel approach where the femtocell and the macrocell share the available spectrum is preferable,

even though this approach suffers from cross-layer interference.

This paper presents the state of the art literature review on mitigating interference techniques for OFDMA-based schemes. The rest of this paper is organized as follows. In Section II, the interference mitigation techniques are classified from different point of views. Then, open issues for femtocell interference researches are presented in Section III. Finally, the paper is concluded in Section IV.

II. INTERFERENCE MITIGATION TECHNIQUES

The available interference mitigation techniques in the literature could be divided to centralized and distributed schemes. In centralized schemes, the macrocell controls the femtocell and assigns its frequency band. While in the distributed schemes the femtocell has the ability to choose its frequency band. Next subsections show the centralized and distributed schemes.

A. Centralized Schemes

1) Clustering or Grouping

An Interference mitigation scheme based on femtocells grouping in low duty operation (LDO) is proposed in [8]. After installing the femtocell and starting the operational mode, femtocell becomes active and starts sending control massages and signals. If any user equipment UE locates near the femtocell, it will receive the signal and know about the existence of the femtocell and try access it. To reduce the interference, the femtocell may enter a LDO mode [9]. Femtocell in LDO mode has two intervals, available interval (AI) and unavailable interval (UAI). In AI, the femtocell will be active with the surrounding, while it will be inactive in the UAI. The length of the AI and UAI can be determined based on the low duty cycle (LDC) pattern.

The authors in [8] try to reduce the unnecessary AIs in LDO mode by organizing all femtocells deployed in the coverage of the macrocell into groups. Each group has header and members. When anew femtocell is installed, it will scan the neighborhood area looking for the existing femtocells. Then it will send a report to the macrocell. The macrocell has the topology of the network, so it will check whether there is header or not in that area. If the signal of the header is more than the threshold value, the new femtocell will join the existing group. Otherwise, it will be the header for a new group. In this case the header only has AI and UAI, while the rest of the group members will be in the UAI mode. Once a UE enter the coverage of the femtocell group, it can know about the group by detecting the signal of the header, even though the rest of the group still in the UAI. If the header has been detected, the UE reports to macrocell, which based on the header information, determine which group has been accessed. Furthermore, it can determine the nearest femtocell and send a message to the femtocell to activate the AI mode in the (LDC) pattern. Therefore, the UE can detect the femtocell in the next scan and access to it. It was shown that the proposed model

reduced the interference time and save the energy of the femtocell.

A combination of frequency bandwidth dynamic division and clustering scheme for cross-layer and co-layer interference is proposed in [10]. In this scheme, CSG is used and the femtocell position is known to the macrocell. Each macrocell has a femtocell system controller, which provides the macrocell with a configuration massage includes the position and the required information of femtocells and femtocells users. The total bandwidth is divided into two bands. The first band is shared with femtocells and the macrocell. While the other is the macrocell dedicated band. The ratio of these two bands depends on the number of clusters in the macrocell, which calculated using clustering algorithm. If the distance between two femtocells is less than the threshold distance, so these two femtocells appointed to deferent clusters. Thus the co-layer interference can be avoided. On the other hand, the cross-layer interference between femtocell and macrocell users can be curbed by assigning the macrocell dedicated band for the macrocell users nearby the femtocell. Simulation results show that user capacity has improved, and reusable probability of femtocell and macrocell reuse band is as high as 97.6%. Therefore, cross-layer interference is effectively inhibited.

In [11], a centralized radio resource management was proposed. In this model a central femtocell chosen by the macrocell, where this cell has connectivity with other femtocells and can share information with them. So the central femto makes the decision of resource allocation. After the central femto is chosen, power is assumed to be allocated in binary power allocated form. The proposed scheme improves system capacity by 5-15% compared to available dynamic resource management schemes.

The backwards of the previous models is assuming a specific location and uniformly distributed for the femtocell and this is not the scenario in the real situation.

2) Dynamic Sub-channel Allocation

The work in [12] presents a downlink sub-channels allocation scheme. The proposed scheme is based on two different algorithms; the first algorithm conducted at the sector level. The first step is when the sector received the channel status from the UEs including the interference. Then the sector prepares a utility matrix and solves this matrix using iterative Hungarian algorithm. Finally, the sector prepares a tentative physical resource block (PRB), and forwards the PRB to the central controller where the second algorithm conducted. The central controller resolves the conflict in the PRB, and sends back a restriction list to the sectors. So the PRBs will be scheduled and transmission only on the qualified PRB. This scheme was developed later in [13], where the preparation way of the utility matrix has been modified, and using a lower transmit power restricted PRBs has been considered. The proposed scheme improved the throughput of the cell edge UEs.

3) Self Organization

A self organization scheme named Q-learning is proposed in [14]. Femtocells integrate themselves in the macrocell network, discover their surrounding, and adjust their parameters accordingly. In this scheme the macrocell users update the macrocell with measurement results about neighboring femtocells, so the macrocell report a self organization massage to the interfering femtocells forcing them to change their carrier or transmit power.

B. Distributed Schemes

Distributed schemes could be divided to two parts. Channel state aware schemes where there is cooperation between the macrocell and the femtocell, and channel state non aware schemes where the macrocell and the femtocell are non cooperative.

1) Aware Schemes

a) Cognitive Radio

In cognitive radio schemes, there is cooperation between the macro and femto cells in the network. Thus each femtocell shares its specification like the number of users, power and frequency with other macrocell and femtocells. The work in [15] presents a resource location scheme based on cognitive radio. In this scheme, the femtocell is required to perform channel sensing and detect the occupied channels by estimating the uplink energy before communicate with femto users and assign the free sub-channels for them. It was assumed that macrocell is divided into two parts, the central where the full band is used and boundary where the frequency is divided into three parts.

In [16], a closed-access cognitive resource allocation scheme for interference management on the downlink was proposed. This scheme is for both co-layer and cross-layer interference scenarios. In this scheme the UEs for both Macro and femtocells are classified as 'safe' for their serving cell and 'victim' for the rest. This classification is based on the signal strength that the UE received from the neighboring macro and femtocells. Therefore, if the interference is strong, then the UE tagged as 'victim' for these cells.

b) Spectrum Arrangement Aware

A Spectrum arrangement aware scheme for uplink cross-layer interference was proposed in [17]. In this scheme, macrocell shares some frequency with femtocells within its coverage. So the whole spectrum is divided into two parts, the macrocell dedicated spectrum part and the femtocell sharing spectrum part. In addition, the macrocell has the ability to determine the macrocell UEs speed. Therefore, UEs moving with high speed will be excluded from the femto-interference pool, which includes the macrocell UEs that close to the femtocell and their performance may be affected by the femtocell. Then, all the UEs which are included in the femto-interference pool will be arranged within the macrocell dedicated spectrum part, and the rest that are not included

could be freely using the whole spectrum. Thus, the cross-layer interference could be avoided.

c) Frequency Scheduling

In [18], a frequency scheduling scheme for cross-layer interference during uplink and downlink is proposed. The idea behind this scheme is that, the femtocell use frequency resource for the far macrocell UEs, and avoids using the frequency resources of the nearby macrocell UEs. The femtocell receives the macrocell UEs scheduling information from the macrocell and it also performs a sensing process for the nearby microcell UEs to determine the occupied part of the spectrum. Then by comparing the received part with the occupied one, it can determine the free frequency resource. The proposed model may help to mitigate the interference issue in the CSG mode.

2) Non-Aware Schemes

a) Frequency Reuse

The idea behind frequency reuse is to divide the network to different sectors, then divide the total frequency spectrum to different sub-bands that assigned to the network sectors. In literature, many techniques of frequency reuse were proposed such as fractional frequency reuse (FFR) [19-25], soft frequency reuse (SFR) [26] and dynamic frequency reuse [27] schemes. In [19-23], a co-channel interference reduction scheme was proposed based on frequency reuse technique, where the subcarriers are divided into two groups, (super group and regular group) as shown in Figure 2. Most of the subcarriers will be assigned for the central region of the cell where a reuse factor of one is applied, while the edge region adopts the factor of three. Using this scheme the cross-layer interference can be avoided, and the co-layer interference can be greatly reduced.
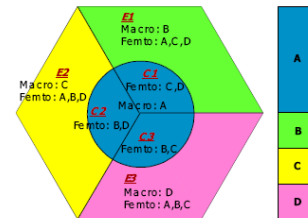
Figure 2: Interference management scheme using fractional frequency reuse [20].

Another idea of frequency reuse scheme is proposed in [24], a fractional frequency reuse scheme for macrocells and femtocells was proposed. In this scheme the cell is divided into two regions inner and outer as shown in Figure 3. A frequency reuse factor of 1 used to the inner region and of 3 for the outer one. So the whole sub-channels can be allocated to the inner region while the outer region uses just 1/3 of all sub-channels in a cell. Therefore the remains 2/3 unused sub-cahnnels can be allocated to the inner femtocells.
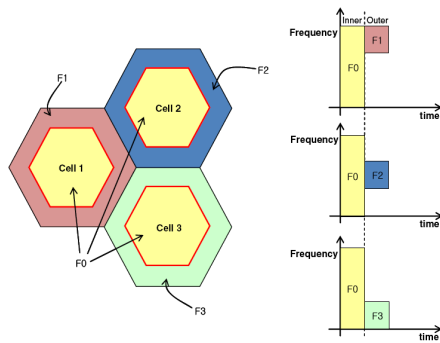
Figure 3: Proposed frequency band/time slots allocation with fractional frequency reuse for macrocells [24].

In [25], a mechanism to reduce cross-layer co-channel interference between macro and femto cells was proposed. This mechanism is based on frequency reuse and pilot sensing method. As shown in Figure 4 a frequency reuse factor of 3 were employed in the network. A pilot sensing is performed as soon as the femtocell turned on, so the femtocells will select the frequency band among the frequency band of the network.
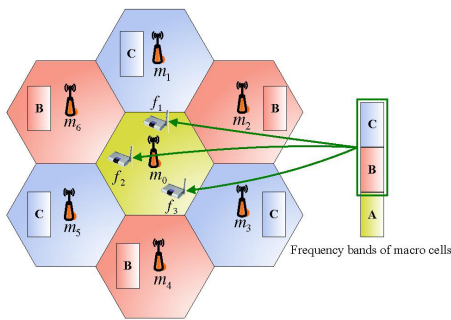
Figure 4: Proposed sub-channel allocation [25].

All of the previous schemes considered as integer frequency reuse since the reuse factor is integer. Therefore new studies go with a non-integer frequency reuse factor [26]. In this scheme the cell-edge region is divided to three virtual tiers each using a different part of spectrum. Hence the cell area is divided up into four distinct sub-cell regions.
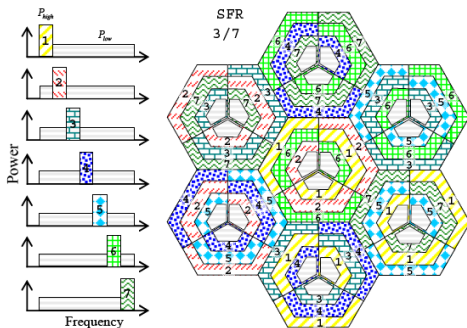
Figure 5: soft frequency reuse 3/7 scheme [26].

The previous models deal with static FFR schemes. So in order to dynamically reuse the frequency resource among the macrocell and femtocells a Dynamic frequency reuse scheme is presented in [27]. This scheme only considered the downlink scenario of the cross-layer interference from the femtocell to the macrocell UEs. The macrocell area is divided into inner and outer regions. Then the outer region is more divided into multiple sectors. In this model the scheme depends not only in the location of the femtocell but also on the density of femtocell in the macrocell region. The scheme adopts hoping FFR for low density and pattern FFR for high density.

b) Distributed Spectrum Allocation Schemes

In [28], two spectrum allocation schemes were investigated. The first one is the shared-spectrum allocation scheme, where the femtocell shares the same spectrum with macrocell system. Thus, the cross-layer interference should be considered. Contrary, in the second one - exclusive-spectrum allocation scheme- co-layer interference should be considered, since the femtocell is completely allocated with different allocation about macrocell system. Two sub-channels selection methods were developed. The channel gain oriented and interference avoidance oriented, where the data transmitted in the sub-channel with higher gain and lower interference respectively. So femtocell users measure the channel status and report to the femtocell the selected sub-channel based on the previous selection methods.

c) Directional Antenna

In [29], the effects of access methods and directional antenna on femtocell were studied. Simulation results show that the open access method can improve the femtocell capacity over the traditional closed access method. Moreover a comparison between switched-beam directional antenna and Omni-directional antenna is conducted. The results show that using a switched-beam directional antenna reduced the interference and achieved better results than using the Omni-directional antenna.

The work in [30] shows that using sectorized antenna and time-hopped CDMA allows about a 7x higher femtocell density than spectrum splitting techniques. The E-plane horns based reconfigurable antenna was used in [31] to reduce the interference and enhance the link reliability and capacity of OFDMA based femtocell systems. The results show that E-plane horns based reconfigurable antenna has 58% higher femtocell capacity than the Omni-directional antenna, and 13.5% higher capacity than the IEEE 802.16m sector antenna.

d) Beam Subset Selection

In [32], the multiuser multiple input multiple output (MIMO) scheme at the macrocell was adopted and an orthogonal random beamforming-based for cross-layer interference scheme was proposed. The most important thing in the orthogonal random beamforming-based is the strategy of beam subset selection in which a subset of the active beams, selected based on the number of users, femtocell density and

SNR, is used to perform the transmission. Adoption the beam subset selection strategy in the macrocell will maximize the throughput by optimizing the trade-off between the multiplexing gain and the multiuser interference. Results show that the proposed strategy protect both macrocell and femtocell users from cross-layer interference.

e) Power Control

Two schemes of interference mitigation based on the adjusting of the power transmitted by the femtocell users were proposed in [33, 34]. These two ways curb the cross-layer interference at the macrocell. The first strategy is the open loop control, where the transmitted power of the femtocell user adjusted to control the interference below a fixed threshold value. While in the second technique, the closed loop control, in addition to the maximum power control, the users also adjusted the noise and interference level at the macrocell. It was shown that the proposed schemes improved the uplink throughput of the macrocell. Moreover the closed loop scheme provides better results. A cross-layer interference mitigation model for the downlink scenario was proposed in [35]. The authors considered the Quality of Service (QoS) at both the Macrocell and femtocell users in terms of their received signal-to-interference-plus-noise ratio (SINR)

In [36], three downlink power control algorithms were developed and investigated. The first algorithm is the fixed power level. In this algorithm, all the femtocells transmit a certain power level. This level determined by the femto itself or by the network coordinators based on the femtocell density. The second algorithm is Femto-Quality of Service (Femto-QoS) power control, where the femtocell transmits the minimum power to meet a certain QoS determined for their users. The last algorithm is Macro-QoS Power Control. In this algorithm the transmitted power of the femtocell is controlled and modified to obtain minimum interference to the neighboring macrocell users. The results show that, by minimizing the transmitted power the data rate of the femtocell will be degraded. Moreover, a high density of femtocells can still affect the performance of the macro-network.

III. OPEN RESEARCH TRENDS

In the previous section, the available interference mitigation techniques were presented. In general centralized techniques shows better results in mitigating interference since there is cooperation between the femtocell and the macrocell. However, these schemes are more complex to implement than the distributed ones. Moreover, the distributed schemes are more practical and realistic. Some techniques proposed for the downlink scenarios only and others are specified for the uplink. Most of the techniques proposed for the cross-layer interference and neglect the co-layer interference. Thus in a future work a comprehensive technique is required to mitigate the interference with high efficiency and low complexity.

A hybrid technique that combines the strengths of both centralized and distributed techniques is an interesting approach for future works. Switched-beam antenna scheme

[29] still needs further investigation to improve the overall femtocell capacity. In addition, MIMO technology looks very attractive to be adopted in femtocell networks [37] with cost being the limiting factor.

It was observed from the literature that most of the proposed schemes assumed specific locations and uniform distributions for the femtocells in the network. On the contrary, as a future work we will consider more practical scenarios where the femtocell is randomly distributed in a dense 3-D environment to accommodate interference from cells spaced horizontally on a terrestrial access or stacked vertically as in the case for office or residential towers. The vertical interference is never considered in the open literature yet.

Also, a hybrid scheme that utilized advances in cognitive radio technology and classical centralized topology (e.g., between Macrocell and its associated femtocells) in conjunction with beam forming will be studied. The femtocell would have the ability to sense and assign the frequency bands for the UEs and also cooperates with the macrocell to share its specification like the number of users, power and frequency. Thus, both cross-layer and co-layer interference can be mitigated. Furthermore, the antenna radiation beam-forming effect on the interference mitigation should be studied in detail as earlier results in the literature looks promising.

Finally, most of the models available in the literature simplify the problem of interference by considering a sinlge-tier of one macrocell and its associated femtocells. This is not accurate enough as several macrocells do exist in any network and each one is associated with its femtocells and users. Hence, our model will include more than one macrocell in the simulation model. On a final note, a small but very important issue is never considered in the literature although it was highlighted before in [38]. The contribution of antenna radiation beam irregularities (e.g., ripple ad small side/back lopes in the radiation pattern) to interference should also be studied.

IV. CONCLUSION

Femtocells are a promising technology to increase the capacity, efficiency and the coverage of the cellular networks. However, several challenging issues must be addressed and resolved before femtocell technology becomes a reality. One of the most important challenges is assessing and mitigating the interference. This paper reviewed different techniques to mitigate the interference in OFDMA femtocell networks.

REFERENCES

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011-2016," Feb. 14, 2012. Available online. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.pdf

[2] airvana, "Femtocells: Transforming The Indoor Experience," October, 2007. http://www.airvana.com/files/Airvana_Femtocell_White_Paper_Oct_2007.pdf.

[3] Small cell forum, "Femtocells and Health," Small cell forum white paper, 2008.

[4] Small cell forum, "Small cells to make up almost 90% of all base stations by 2016," Small cell forum press releases, 28 february 2012.

[5] Informa Telecoms & Media, "Femtocell Market Status," Small cell forum whitepaper, February, 2011.

[6] Airvana, "Femtocell Standardization," May, 2010. Available online. http://www.airvana.com/default/assets/File/WHITEPAPER-Femtocell_ Standardization.pdf

[7] A. Golaup, M. Mustapha, and L. B. Patanapongipibul, "Femtocell Access Control Strategy in UMTS and LTE," IEEE Communications Magazine, vol. 47, no. 9, pp. 117–123, Sept. 2009.

[8] H. Widiarti, S. Pyun, and D. Cho, "Interference Mitigation Based on Femtocells Grouping in Low Duty Operation," Proc. IEEE 72nd Vehicular Technology Conference Fall (VTC'10-Fall), pp. 1–5, 6-9 Sept. 2010.

[9] IEEE P802.16m/D4, "DRAFT Amendment to IEEE Standard for Local and metropolitan area networks, Part 16: Air Interface for Broadband Wireless Access Systems - Advanced Air Interface", Feb. 2010.

[10] H. Li, X. Xu, D. Hu, X. Qu, X. Tao, and P. Zhang, "Graph Method Based Clustering Strategy for Femtocell Interference Management and Spectrum Efficiency Improvement," Proc. IEEE 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM), pp. 1–5, 23-25 Sept. 2010.

[11] J. Kim and D.-H. Cho, "A Joint Power and Subchannel Allocation Scheme Maximizing System Capacity in Indoor Dense Mobile Communication Systems," IEEE Trans. Veh. Technol., vol. 59, no. 9, 2010, pp. 4340-4353.

[12] M. Rahman and H. Yanikomeroglu, "Interference avoidance through dynamic downlink OFDMA subchannel allocation using intercell coordination," in Proc. IEEE Veh. Technol. Conf. (VTC2008-Spring), May 2008, pp. 1630-1635.

[13] M. Rahman and H. Yanikomeroglu, "Enhancing cell-edge performance: a downlink dynamic interference avoidance scheme with inter-cell coordination," IEEE Trans. Wireless Commun., pp. 1414–1425, April 2010.

[14] M. Bennis and D. Niyato, "A Q-learning based approach to interference avoidance in self-organized femtocell networks," 1st IEEE International Workshop on Femtocell Networks (FEMnet) in conjunction with IEEE GLOBECOM 2010, Miami, FL, USA.

[15] D.-C. Oh, H.-C. Lee, and Y.-H. Lee, "Cognitive Radio Based Femtocell Resource Allocation," in Proc. ICTC 2010,pp. 274 - 279, Nov. 2010.

[16] S. Kaimaletu, R. Krishnan, S. Kalyani, N. Akhtar, and B. Ramamurthi, "Cognitive interference management in heterogeneous femto-macro cell networks," in Proc. IEEE ICC'11, Kyoto, Japan, June 2011, pp. 1–6.

[17] W. Yi, Z. Dongmei, J. Hai, and W. Ye, "A Novel Spectrum Arrangement Scheme for Femtocell Deployment in LTE Macrocells," Proc. IEEE 20th Symposium on Personal, Indoor and Mobile Radio Communications, pp. 6–11, 13-16 Sept. 2009.

[18] M. E. Sahin, I. Guvenc, Moo-Ryong Jeong, H. Arslan, "Handling CCI and ICI in OFDMA Femtocell Networks through Frequency Scheduling," IEEE Transactions on Consumer Electronics, vol. 55, no. 4, pp. 1936–1944, Nov. 2009.

[19] H. Lei, L Zhang, X. Zhang, and D. Yang, "A Novel Multi-cell OFDMA System Structure Using Fractional Frequency Reuse," IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Sept. 2007.

[20] T. Lee, H. Kim, J. Park, J. Shin, "An efficient resource allocation in ofdma femtocells networks," In IEEE 72nd Vehicular Technology Conference (2010).

[21] P. Lee, T. Lee, J. Jeong, and J. Shin, " Interference management in LTE femtocell systems using fractional frequency reuse," In ICACT 2010 Proceeding of the 12th International Conference on Advanced Communication Technology (2010).

[22] T. Lee, J. Yoon, S. Lee, and J. Lee, "Interference management in OFDMA Femtocell systems using Fractional Frequency Reuse," In International Conference on Communications, Circuits and Systems, Chengdu, pp. 176–180 (2010).

[23] T. Lee, J. Yoon, S. Lee, and J. Shin, "Resource allocation analysis in OFDMA femtocells Using fractional frequency reuse," in Proc. IEEE PIMRC, Istanbul, Turkey, Sept. 2010.

[24] T. Kim and T. Lee, "Throughput Enhancement of Macro and Femto Networks by Frequency Reuse and Pilot Sensing," Proc. IEEE International Performance, Computing and Communications Conference (IPCCC), pp. 390–394, Dec. 2008.

[25] C.-Y. Oh, M.-Y. Chung, H.-S. Choo, and T.-J. Lee, "A novel frequency planning for femtocells in OFDMA-based cellular networks using fractional frequency reuse," in Proc. of ICCSA, 2010.

[26] C. Kosta, A. Imran, A.U. Quddus, and R. Tafazolli, "Flexible Soft Frequency Reuse Schemes for Heterogeneous Networks (Macrocell and Femtocell) ," in Proc. VTC Spring, 2011, pp.1-5.

[27] F. Tariq, L. S. Dooley, A. S. Poulton and Y. Ji, "Dynamic Fractional Frequency Reuse Based Hybrid Resource Management in Femtocell Networks," 7th International Conference on International Wireless Communication and Mobile Computing (IWCMC), Istanbul, Turkey, July, 2011.

[28] C. Lee, J.-H. Huang, and L. Wang, "Distributed channel selection principles for femtocells with two-tier interference," in Proc. IEEE VTC, May 2010, pp. 1–5.

[29] A. Tsai, J. Huang, and L. Wang, "Capacity Comparison for CSG and OSG OFDMA Femtocells," IEEE Globecom FEMnet Workshop, 2010.

[30] V. Chandrasekhar and J. G. Andrews, "Uplink Capacity and Interference Avoidance for Two-Tier Femtocell Networks," IEEE Trans. Wireless Commun., Feb. 2008.

[31] A. Tsai, J.-H. Huang, L. Wang, and R.-B. Hwang, "High capacity femtocells with directional antennas," in Proc. IEEE WCNC, Apr. 2010, pp. 1–6.

[32] S. Park, W. Seo, Y. Kim, S. Lim, and D. Hong, "Beam Subset Selection Strategy for Interference Reduction in Two-tier Femtocell Networks," IEEE Transactions on Wireless Communications, vol. 9, no. 11, pp. 3440–3449, Nov. 2010.

[33] H.-S. Jo, Jong-Gwan Yook, C. Mun and J. Moon, "A Self-Organized Uplink Power Control for Cross-Tier Interference Management in Femtocell Networks", Military Communications Conference, pp 1-6, Nov. 2008.

[34] H.-S. Jo, C. Mun, J. Moon, and J.-G. Yook, "Interference mitigation using uplink power control for two-tier femtocell networks," IEEE Trans. Wireless Commun., vol. 8, no. 10, pp. 4906–4910, Oct. 2009.

[35] X. Li, L. Qian, D. Kataria, "Downlink Power Control in Co-Channel Macrocell Femtocell Overlay," IEEE ICCS, pp 383-388, March 2009.

[36] S. P. Yeh, S. Talwar, N. Himayat, and K. Johnsson, "Power control based interference mitigation in multitier networks," IEEE Globecom Workshops (GC), 2010.

[37] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: a survey," IEEE Commun. Mag., vol. 46, no. 9, pp. 59–67, Sep. 2008.

[38] L.-C. Wang, A. Chen and S.-Y. Huang, "A Cross=Layer Investigation for the Throughput Performance of CSMA/CA-Based WLANs with Directional Antennas and Capture Effect," IEEE Transactions on Vehicular Technology, Vol. 56, No. 5, pp.: 2756-2766, September 2007.

Statistical Analysis on Human Body Radiation Points

Mohamad Hushnie Haron[1], Mohd Nasir Taib[1], Megat Syahirul Amin Megat Ali[1], Megawati Mohd Yunus[1]

[1]Faculty of Electrical Engineering
Universiti Teknologi MARA
Selangor, Malaysia
hushnie_haron@yahoo.com.my, dr.nasir@ieee.org, megatsyahirul@salam.uitm.edu.my

Siti Zura A. Jalil[1,2]
[2]Razak School of Engineering
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
zura@ic.utm.my

Abstract – **Recently, there has been an increasing concern about scientific investigations of the endogenous electromagnetic fields generated by and contained within the human body. Chakra points are defined as focal spots for reception, absorption and transmission of radiation wave of the human body. Meanwhile brain regions are defined as areas of the brain that represent the cerebrum, cerebellum and brainstem. This paper discusses on frequency characteristic of human body radiation through chakra points and brain regions. A total of 9 sites were investigated for the measurements. The readings from each site were recorded via a frequency counter. A total of 34 healthy subjects comprising of 17 males and 17 females were used in the study. Data analyses were conducted using exploratory and descriptive statistics. Results show that out of the nine sites, three were found to be significant.**

Keywords: Electromagnetic wave, human radiation wave, chakra points, brain regions

I. INTRODUCTION

Every human being in this world emits electromagnetic radiation around their bodies. This radiation vibrates at specific frequencies and generally takes a form of an oval pattern. This phenomenon is known as aura or chakra [1],[2]. An electromagnetic radiation is defined as self-propagating wave in space and it consists of electric and magnetic components. Due to the oscillation and acceleration of an electric charge, radiation waves are produced. Like any other waves, it can be characterized by three properties such as amplitude, wavelength and frequency. These properties are interrelated. By knowing the frequency, the direction, the energy and the intensity of radiation could be defined [1],[3].

Aura is described as luminous body that surrounds and penetrates physical body and emit frequency [2]. The word chakra originates from Sanskrit which means wheel of light or ring of light. These points represent relationship between the mind and body with one's psychological and physiological state. There are seven chakra points located along central axis of human body, namely the root, sacral, solar plexus, heart, throat, forehead, and crown chakra [2],[3]. The location of chakra points are illustrated in Fig. 1.
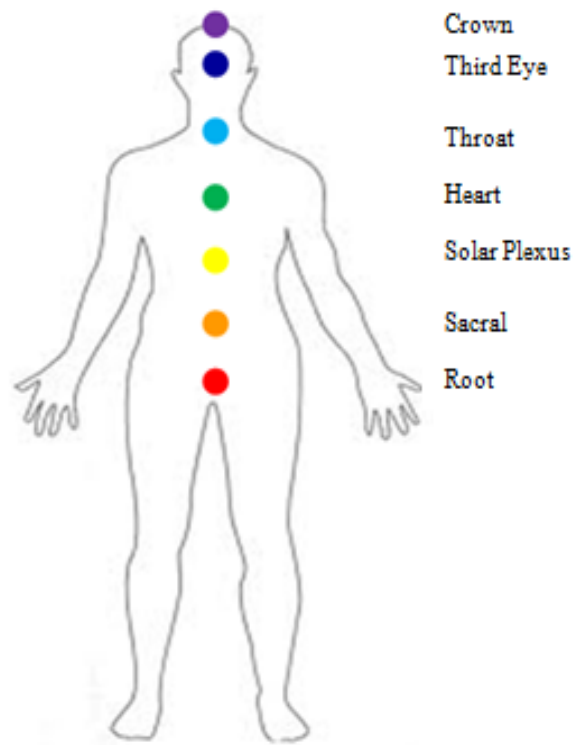
Figure1: Chakra points on human body

The brain can be divided into the cerebrum, cerebellum, and the brainstem, only the first two regions are considered for the measurement in this study. Overlapping points have been identified between frontal lobe and forehead chakra point, and parietal lobe and crown chakra point. Hence, the occipital and cerebellum region were added in order to provide an additional measurement points for obtaining the endogenous electromagnetic fields [4]. The respective brain regions are shown in Fig. 2.
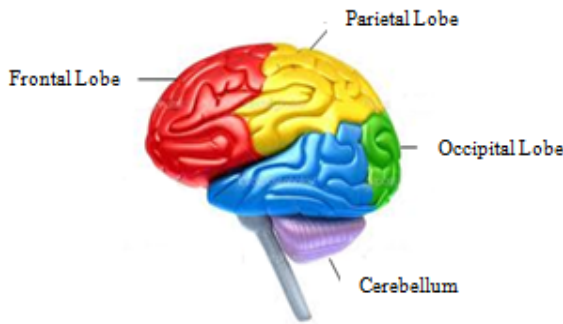
Figure 2: Regions of human brain

This paper discusses on statistical characteristics and relationships of human body radiation points. The characteristics are analyzed using box plots, normality test, standard deviation, skewness and kurtosis. In addition, the relationships are obtained using correlation test. The discussion on the result can be splitted into 2, using all human body radiation points and human body radiation points according to gender.

II. METHOD

A. Measurement

In this study, the body radiations have been measured using frequency counter. The range of measurement for the equipment is between 1 MHz to 3 GHz, with a resolution of 0.1 Hz. A total of 9 measurement spots have been selected for the purpose of the study. The frequency measurements were conducted in an anechoic chamber with ferrite floor, where the temperature was held constant at 23ºC. A total of 34 subjects comprising of 17 males and 17 females have participated for data collection. The age range of the participants is between 19 and 26 years. 10 readings per chakra point or brain region were recorded for every participant. In all readings, the antenna has been placed parallel with the ground and the frequencies were captured at a distance of 1 to 5 cm from the body. The description of the points is represented in Table I.

TABLE I: HUMAN BODY RADIATION POINTS AND THEIR DESCRIPTION

Points	Description	Points	Description
CA	Root	CF	Third Eye / Frontal Lobe
CB	Sacral	CG	Crown / Parietal Lobe
CC	Solar Plexus	CH	Occipital Lobe
CD	Heart	CJ	Cerebellum
CE	Throat		

B. Data Analysis

Data analysis consists of three main works which include pre-processing of raw data, exploratory statistical analysis and descriptive statistics. Fig. 3 shows the flow of statistical analysis conducted in this study.
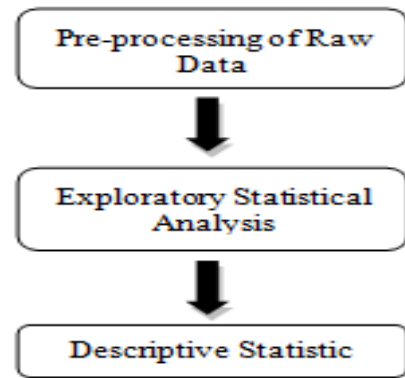
Figure 3: Methodology of statistical data analysis

In pre-processing of raw data, the mean of 10 readings per point per sample were initially calculated using (1) [5].

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

The next step is to rescale the averaged data to a value between 0 and 1. Here, min-max normalization has been used and in theory, the min-max normalization performs linear transformation on the original data. This method has been chosen because it has higher accuracy compared to other normalization methods [6]. The transformation process can be expressed by (2).

$$v' = \frac{(v - \text{minimum value of } v)}{(\text{maximum value of } v - \text{minimum value of } v)} \qquad (2)$$

The second work involves exploratory statistical analysis using box plot. The method was chosen due to its ability to provide relevant information on the maximum and minimum value, upper and lower quartile, as well as median. The advantages of box plot are the simplicity of its design, critical information of dataset is quickly expressed and it also shows the data distribution [7].

The third work employs descriptive statistics. Normality and correlation tests form the two major contributions in this stage. In normality test, the Shapiro-Wilk method has been selected because the technique is capable of detecting non-normality for a wide variety of statistical distributions, including those with Gaussian kurtosis values. It has been recommended as a powerful omnibus test of normality, regardless of distribution and sample size [8],[9]. If the p-value is less than the chosen level of significance, the null hypothesis can be rejected at that level and this can be taken as evidence supporting non-Gaussianity [10]. The Shapiro-Wilks formula can be represented by (3) [8],[9].

$$w = \frac{\left(\sum_{i=1}^{N} a_i y_i\right)^2}{\sum_{i=1}^{N}(y_i - m_1)^2} \qquad (3)$$

As an addition to normality test, standard deviation, skewness and kurtosis values of the normalized data have also been calculated. In general, they represent the shape of a distribution. In Gaussian distribution, normality can be defined with skewness = 0 and kurtosis = 3 or excess kurtosis = 0. Here, traditional sample skewness and kurtosis have been used. The parameters can be expressed by (4), (5) and (6), respectively [5],[11],[12].

$$s = \sqrt{\frac{1}{(N-1)}\sum_{i=1}^{N}(x_i - \bar{x})^2} \qquad (4)$$

$$G_1 = \frac{(\sqrt{n(n-1)})g1}{(n-2)} \qquad (5)$$

$$G_2 = \frac{(n-1)[(n+1)g_2 + 6]}{(n-2)(n-3)} \qquad (6)$$

The final step in descriptive statistics is to perform correlation test. It is used to investigate the relationship between points and with gender. In this case, Pearson's correlation test has been used in correlating between points and point biserial correlation test in correlating between points and gender. The correlation coefficient, r, describes a linear correlative degree between two variables. The value ranges from -1 to 1. If r is positive, both variables will experience a reciprocal effect. If r is negative, the variables would have an opposing effect on each other. The correlation coefficient can be represented by (7) [13]. The interpretation of Pearson's correlation coefficients are as described in Table II [14].

$$r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{(x - \bar{x})^2}\sqrt{(y - \bar{y})^2}} \qquad (7)$$

TABLE II: INTERPRETATION OF PEARSON'S CORRELATION COEFFICIENTS

Correlation Coefficient, r	Interpretation
± (0.80 - 1)	Very Strong Correlation
± (0.60 - 0.79)	Strong Correlation
± (0.40 - 0.59)	Moderate Correlation
± (0.20 - 0.39)	Weak Correlation
± (0 - 0.19)	Very Weak Correlation

III. RESULT AND DISCUSSION

Fig. 4 shows the data for human body radiation points. The box plot shows that points, CF, CG, CH and CJ, which are located above the throat has higher median than points located below throat. In terms of distribution, all points are positively skewed except for CB and CE that are negatively skewed.

There are two outliers at point CB and one outlier each at points CC, CF and CH.
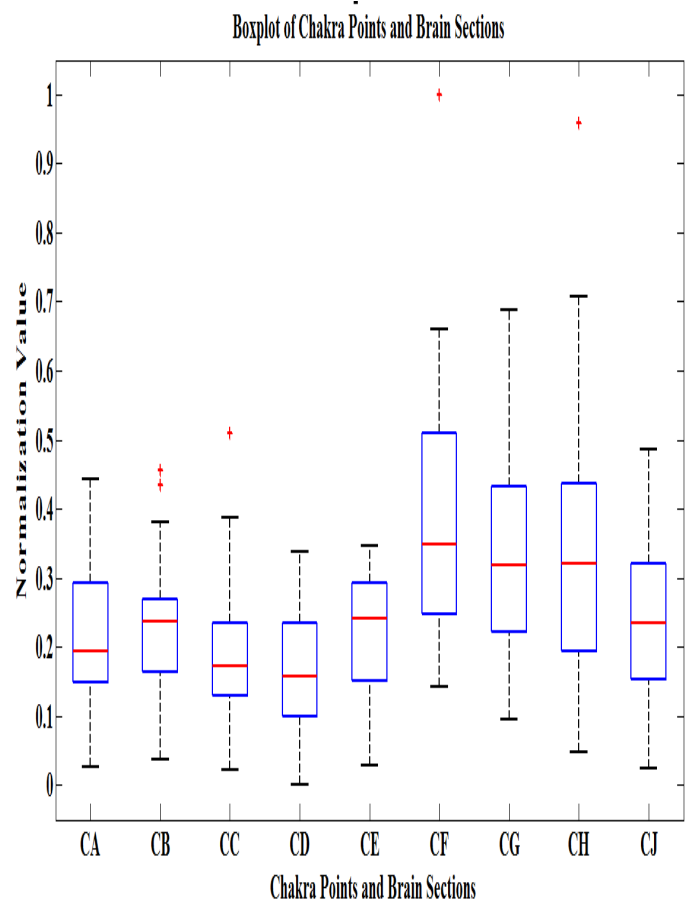
Figure 4: Box plot of human body radiation points

Fig. 5 shows the data of human body radiation points according to gender (male and female). Males have higher maximum frequency than females for points CB, CC, CH and CJ. Meanwhile, females have lower minimum frequency than males for all points except for CG. On the other hand, males have lower median than females for points CF, CH and CJ. In terms of distribution, the skewness between both genders opposes each other for every point analyzed. There is one outlier for females located at CC, while two outliers exist at CF and CH for males.

Boxplot of Chakra Points and Brain Sections Between Gender
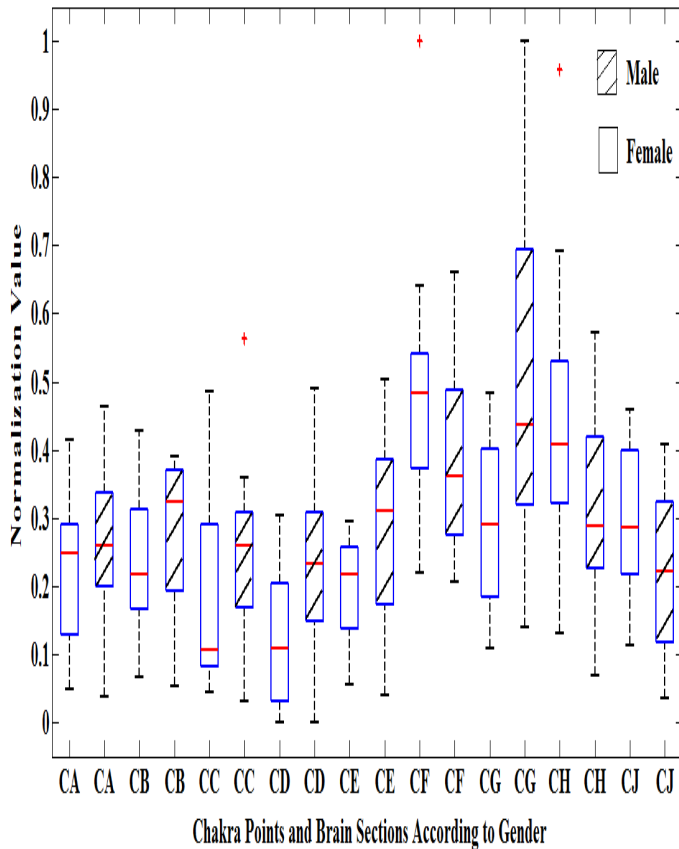
Figure 5: Box plot body radiation points according to gender

Table III shows the result of Shapiro-Wilk normality test for all human body radiation points. In this test, 95% significance level, p, has been chosen. The significance values indicate that the dependent variables, CA, CB, CD, CE, CH and CJ, are above 0.05 and therefore are normally distributed. The other three dependent variables, CC, CF and CG, are not normally distributed.

TABLE III: SHAPIRO-WILK NORMALITY TEST OF HUMAN BODY RADIATION POINTS

Points	Shapiro-Wilk	
	Statistic	*Sig.*
Base (CA)	0.9651	0.3407
Sacral (CB)	0.9797	0.7644
Solar Plexus (CC)	0.9391	0.0578
Heart (CD)	0.9736	0.5688
Throat (CE)	0.9509	0.1307
Third Eye/Frontal Lobe (CF)	0.9104	0.0087
Crown/Parietal Lobe (CG)	0.9644	0.3262
Occipital Lobe (CH)	0.9197	0.0159
Cerebellum (CJ)	0.9562	0.1875

Table IV shows the Shapiro-Wilks normality test for all radiation points according to gender. In this test, again 95% significance level has been chosen. The significance value indicate that for males, the dependent variables, CA, CB, CD, CE, CG, CH, CJ, are normally distributed, while the other two dependent variables, CC and CF, are not normally distributed. All points are normally distributed for females, except for CB.

TABLE IV: SHAPIRO-WILK NORMALITY TEST OF HUMAN BODY RADIATION POINTS ACCORDING TO GENDER

Gender	Points	Shapiro-Wilk	
		Statistic	*Sig.*
Male	Base (CA)	0.949	0.443
	Sacral (CB)	0.972	0.858
	Solar Plexus (CC)	0.870	0.022
	Heart (CD)	0.924	0.173
	Throat (CE)	0.909	0.096
	Third Eye/Frontal Lobe (CF)	0.873	0.025
	Crown/Parietal Lobe (CG)	0.927	0.192
	Occipital Lobe (CH)	0.936	0.271
	Cerebellum (CJ)	0.923	0.166
Female	Base (CA)	0.978	0.940
	Sacral (CB)	0.872	0.024
	Solar Plexus (CC)	0.926	0.187
	Heart (CD)	0.984	0.985
	Throat (CE)	0.960	0.626
	Third Eye/Frontal Lobe (CF)	0.930	0.216
	Crown/Parietal Lobe (CG)	0.933	0.242
	Occipital Lobe (CH)	0.973	0.870
	Cerebellum (CJ)	0.947	0.405

Table V shows the standard deviation, skewness and kurtosis for human body radiation points. The maximum standard deviation is 0.19, while the minimum is at 0.08. Hence all points are close to mean. In terms of skewness, all points have a value greater than 0, except for CE. Therefore, CE exhibit negative skew characteristics while others are positively skewed. As for analysis on kurtosis, points CA, CB, CD, CE, CG and CJ have negative values, while others are positive. Therefore, points CA, CB, CD, CE, CG and CJ have lower and broader peak and the tails are shorter and thinner. Meanwhile, the remaining points CC, CF and CH have higher and sharper peak while the tails are longer and fatter.

TABLE V: DISTRIBUTION IN TERMS OF SKEWNESS AND KURTOSIS OF HUMAN BODY RADIATION POINTS

Points	Std. Deviation	Skewness	Kurtosis
Base (CA)	0.1027	0.4368	-0.2946
Sacral (CB)	0.0992	0.1576	-0.0007
Solar Plexus (CC)	0.1075	0.9650	1.0668
Heart (CD)	0.0892	0.2396	-0.6839
Throat (CE)	0.0836	-0.5561	-0.5732
Third Eye/Frontal Lobe (CF)	0.1825	1.1061	2.1255
Crown/Parietal Lobe (CG)	0.1445	0.5567	-0.2164
Occipital Lobe (CH)	0.1916	1.2037	2.0742
Cerebellum (CJ)	0.1326	0.2373	-0.9207

Table VI shows values of standard deviation, skewness and kurtosis for radiation points according to gender. Looking into both male and female, all the points are close to mean. For males, points CE, CG and CJ are negatively skewed, while the others exhibit positive skewness. Analysis on kurtosis shows that points CA, CB, CD, CE, CG and CJ have lower and broader peak, whereas the tails are shorter and thinner. Meanwhile points CC, CF and CH have higher and sharper peak, while the tails are longer and fatter. Analyses on females show that points CB, CD and CE have negative skew, while the others are positively skewed. As for kurtosis, points CC and CD have higher and sharper peak, whereas the tails are longer and fatter. Others have lower and broader peak and the tails are shorter and thinner.

TABLE VI: DISTRIBUTION IN TERMS OF SKEWNESS AND KURTOSIS OF HUMAN BODY RADIATION POINTS ACCORDING TO GENDER

Gender	Points	Std. Deviation	Skewness	Kurtosis
Male	Base (CA)	0.113	0.204	-0.979
	Sacral (CB)	0.103	0.163	-0.759
	Solar Plexus (CC)	0.127	0.977	0.207
	Heart (CD)	0.100	0.428	-1.105
	Throat (CE)	0.074	-0.714	-0.687
	Third Eye/Frontal Lobe (CF)	0.177	1.241	3.699
	Crown/Parietal Lobe (CG)	0.125	-0.095	-1.420
	Occipital Lobe (CH)	0.203	0.972	1.790
	Cerebellum (CJ)	0.116	-0.257	-1.321

Gender	Points	Std. Deviation	Skewness	Kurtosis
Female	Base (CA)	0.115	0.058	-0.326
	Sacral (CB)	0.114	-0.723	-0.892
	Solar Plexus (CC)	0.128	0.427	1.359
	Heart (CD)	0.125	-0.008	0.023
	Throat (CE)	0.136	-0.273	-0.771
	Third Eye/Frontal Lobe (CF)	0.134	0.743	-0.323
	Crown/Parietal Lobe (CG)	0.246	0.679	-0.414
	Occipital Lobe (CH)	0.135	0.247	-0.461
	Cerebellum (CJ)	0.123	0.076	-1.152

Table VII shows the result of Pearson's correlation test between the human body radiation points. Only values that are significantly correlated to each other were displayed. The results show a very strong correlation between points CF and CH and a strong correlation between CB and CC, CB and CE, CF and CJ, and as well as CH and CJ. They also indicate a moderate correlation for points CA and CB, CA and CE, CA and CJ, CB and CH, CC and CD, CC and CE, and CD and CE. The remaining hidden points are weakly correlated.

TABLE VII: PEARSON'S CORRELATION TEST BETWEEN HUMAN BODY RADIATION POINTS

Points	r Value	p Value
CA-CB	0.5593	0.0006
CA-CC	0.3711	0.0307
CA-CE	0.5196	0.0016
CA-CJ	0.4574	0.0065
CB-CC	0.7387	0.0000
CB-CD	0.3558	0.0389
CB-CE	0.6328	0.0001
CB-CH	0.4643	0.0057
CB-CJ	0.3567	0.0384
CC-CD	0.4317	0.0108
CC-CE	0.5456	0.0008
CD-CE	0.5523	0.0007
CE-CG	0.3631	0.0348
CE-CH	0.3637	0.0345
CE-CJ	0.3491	0.0430
CF-CH	0.8587	0.0000
CF-CJ	0.6254	0.0001
CH-CJ	0.6342	0.0001

Table VIII shows the result of Pearson's correlation test between human body radiation points according to gender. Beside points CB and CC, CB and CE also show strong correlation, while points CC and CE are moderately correlated. Other hidden points have weak correlation.

TABLE VIII: PEARSON'S CORRELATION TEST BETWEEN HUMAN BODY
RADIATION POINTS ACCORDING TO GENDER

Points	Male		Female	
	r value	*p value*	*r value*	*p value*
CB-CC	0.751	0.001	0.689	0.002
CB-CE	0.645	0.005	0.617	0.008
CC-CE	0.523	0.031	0.562	0.019

Table IX shows the result of point biserial correlation test between human body radiation points and gender. The obtained coefficients, r, shows that points CD, CE, and CG have moderate correlation, while other hidden points are weakly correlated.

TABLE IX: POINT BISERIAL CORRELATION TEST BETWEEN HUMAN RADIATION
POINTS AND GENDER

Points		Gender
Heart (CD)	r value	0.453
	p value	0.007
Throat (CE)	r value	0.401
	p value	0.019
Crown/Parietal Lobe (CG)	r value	0.497
	p value	0.003

IV. CONCLUSION

The statistical results presented in this study indicate that the data have certain pattern and behaviour. Looking at the distribution of all radiation points, more than 70% are Gaussian distributed. Furthermore, more than 80% of the points are positively skewed and more than 60% have platykurtic distribution. Analysis on all radiation points according to gender show that more than 80% are Gaussian distributed for males and 90% for females. For both genders, 70% of the points are positively skewed. However, 70% of the points have platykurtic distribution for male and 80% for female. In terms of correlation, points CB, CC and CE are found to be significant using Pearson's correlation, while points CD, CE and CG are significant using point biserial correlation.

ACKNOWLEDGMENT

REFERENCES

[1] S. Z. A. Jalil, M. N. Taib, and H. Abdullah, "Investigation of frequency radiations of the human body," in *Biomedical Engineering and Sciences (IECBES), 2010 IEEE EMBS Conference on*, 2010, pp. 377-380.

[2] S. Z. A. Jalil, M. N. Taib, M. Y. M. A. Karim, and H. A. Idris, "Investigation into gender recognition based on body radiation," 2010, pp. 507-510.

[3] S. Z. A. Jalil, M. N. Taib, H. A. Idris, and M. M. Yunus, "Examination of human body frequency radiation," in *Research and Development (SCOReD), 2010 IEEE Student Conference on*, 2010, pp. 4-7.

[4] J. L. Lancaster, M. G. Woldorff, L. M. Parsons, M. Liotti, C. S. Freitas, L. Rainey, P. V. Kochunov, D. Nickerson, S. A. Mikiten, and P. T. Fox, "Automated Talairach atlas labels for functional brain mapping," *Human brain mapping,* vol. 10, pp. 120-131, 2000.

[5] R. Tavakkoli-Moghaddam, S. M. H. Mojtahedi, S. M. Mousavi, and A. Aminian, "A Jackknife technique to estimate the standard deviation in a project risk severity data analysis," in *Computers & Industrial Engineering, 2009. CIE 2009. International Conference on*, 2009, pp. 1325-1329.

[6] L. Al Shalabi and Z. Shaaban, "Normalization as a Preprocessing Engine for Data Mining and the Approach of Preference Matrix," in *Dependability of Computer Systems, 2006. DepCos-RELCOMEX '06. International Conference on*, 2006, pp. 207-214.

[7] K. Potter, "Methods for presenting statistical information: The box plot," *Visualization of Large and Unstructured Data Sets,(LNI),* vol. 4, pp. 97–106, 2006.

[8] B. Guner, M. T. Frankford, and J. T. Johnson, "A Study of the Shapiro-Wilk Test for the Detection of Pulsed Sinusoidal Radio Frequency Interference," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 47, pp. 1745-1751, 2009.

[9] M. Mendes and A. Pala, "Type I Error Rate and Power of Three Normality Tests," *Pakistan Journal of Information and Technology,* vol. 2, pp. 135-139, 2003.

[10] M. Üzümcü, A. F. Frangi, J. H. Reiber, and B. P. Lelieveldt, "Independent Component Analysis in Statistical Shape Models," 2003, pp. 375–383.

[11] D. Joanes and C. Gill, "Comparing measures of sample skewness and kurtosis," *Journal of the Royal Statistical Society: Series D (The Statistician),* vol. 47, pp. 183-189, 1998.

[12] J. Bai and S. Ng, "Tests for Skewness, Kurtosis, and Normality for Time Series Data," *Journal of Business and Economic Statistics,* vol. 23, pp. 49-60, 2005.

[13] W. Wen-Jie and X. Yan, "Correlation analysis of visual verbs' subcategorization based on Pearson's correlation coefficient," in *Machine Learning and Cybernetics (ICMLC), 2010 International Conference on*, 2010, pp. 2042-2046.

[14] J. D. Evans, *Straightforward statistics for the behavioral sciences*: Brooks/Cole Publishing Company, 1996.

BER Performance of LBC coded OFDM in different channels

Vandana B. Malode
Department of Electronics
J.N.E.C.,N-6,CIDCO,Aurangabad
Maharashtra, INDIA

Bhagwat P. Patil
Department of Electronics
Army Institute of Technology, Dighi, Pune,
Maharashtra, INDIA

ABSTRACT: Orthogonal Frequency Division Multiplexing (OFDM) has attracted lot of attention of researchers because of its vital role of it in wireless communication technology. This paper discusses the Bit Error Rate (BER) performance of OFDM system using Differential Phase Shift Keying (DPSK) under various channel conditions. It is seen that coding improves BER performance of OFDM and hence in this work Linear Block coding (LBC) is used for improvement purpose. The reason of selecting differential detection modulation scheme is its receiver structure is simple. The simulation is carried out in MATLAB simulink.

KEYWORD: OFDM, Linear block code, BER, DPSK , Various Channels.

I. INTRODUCTION

OFDM is a very efficient technique for high-speed data transmission used in mobile communication, Digital Audio Broadcasting(DAB) is unidirectional broadcasting systems [1], Digital terrestrial mobile communication, Digital Video Broadcasting terrestrial(DVB-T), wireless asynchronous transfer mode (WATM), Modem/ADSL[2,3]. OFDM has many advantages such as High spectral efficiency, robustness in frequency selective fading channels[4], immunity to inter-symbol interference and capability of handling very strong multipath fading.[5] Recently the advance system such as 3GPP LTE and terrestrial digital TV Broadcasting have been sophisticatedly developed using OFDM technology. Besides the above advantages, however, the bit error occurs in the presence of Doppler frequency shift because of the inter-carrier interference (ICI) in OFDM.

OFDM is preferred to implement due to efficient FFT(Fast Fourier Transform) algorithm. The implementation of FFT engine is proportional to the number of input bit size. However the BER performance is analyzed in AWGN channel. Orlandos Grigoriadis, H. Shrikanta Kamat have given the results of the BER as a function of Multi propagation for each set of carriers[6].In[7] Continuity has to be maintained related to our topic

the high-order modulation like 16-QAM needs more accurate CFO and channel estimation to avoid performance loss. Whereas the improvement of performance elevation criteria of RS coded OFDM transmission over noisy channel analyzed using simulink. Performance of a flat fading channel using Convolution   Interleave that cuts down time delay and memory space increases considerably [8].

In this era the demand of mobile radio communication systems are increasing to provide a variety of high quality services, more features, compactness and other such variety of services to mobile users. Hence modern mobile radio transceiver systems must be able to provide high capacity and variable bit rate information with high bandwidth efficiency, efficient signal power as such[9,10].
Differentially coded modulation enables the receiver to detect the data symbols without knowledge of the references phase therefore no channel estimation required, if differential phase shift keying (DPSK) is used for modulating the subcarriers [11].

This paper is aimed to show that the linear block coding is used in OFDM system for improvement in bit error rate performance under various channels. Section II recalls the basics of channels. The linear block coding is briefly overviewed in section III. Then in Section IV shows proposed simulation model which gives better performance in Bit Error Rate. Simulation Results are compared with approximation in section V.

II. FADING CHANNELS

A signal in mobile communication does not necessarily reach the receiving antenna directly. It gets reflected, diffracted or scattered due to various natural or manmade obstacles in the signal path. Thus the signal undergoes a multipath propagation. This results in infinite number of replicas of the transmitted signal. At receiving antenna the signal reaches as superimposition of the signal which may be constructive or destructive. This affects the amplitude and other such parameters of the signal. Some of the frequencies even get erased due to the

reflections, diffractions and other such effects called as selective frequency fading.

The signal along with its replicas is received at receiving antenna in a scattered manner. A time spread occurs between the first and last of the multipath signal at the receiver called as delay spread. In condition when the transmitter and receiver are in motion relatively in a radio transmission frequency of the signal may increase or the frequency of the signal may increase to certain extent before decreasing depending upon the movement. This is the Doppler shift and collectively these are multipath fading. The signal in LOS or short distance propagation is categorized as Rayleigh fading. When the signal undergoes a Non-LOS propagation or large distance propagation is consider a Rician fading.

III. LINEAR BLOCK CODES

Consider an (n, k) Linear code C with parity-check matrix H, where n is the length and k is the dimension of C. Since $Hc^t = 0$ for any codeword c ∈ C, any vector X ∈ e+c has the same syndrome as e, that is [12, 13]

$$Hx^t = H(e + c)^t = He^t \qquad (1)$$

A binary information sequence is divided into blocks of 4 bits. Each message block is encoded into a codeword C which is 7 bits by a (7, 4) hamming encoder. Hamming codes were designed for correction [14]. The parameters for the family of binary hamming codes are typically expressed as a function of a single integer m≥2 (for m=3, use third gender not necessarily prime, it is any positive integer. A hamming code on GF (2) has code length n=2^m-1, message length k=2^m-1-m , redundancy n-k=m and error connecting capability t=1 bit.

HAMMING CODES

Hamming codes are only single error correcting codes. To improve the error detection and correction capability by adding parity check digit. The resulting code is called the extended binary hamming code. Suppose that c which is 7 bits by a [7,4]. Let $\hat{c}$ be the code obtained by adding a single character to the end of each word in c in such a way that every word in $\hat{c}$ has even weight. The parity check matrix of (8, 4) extended hamming code $\hat{c}$ is $\hat{H}$:

$$\hat{H} = \begin{bmatrix} & & & 0 \\ & & & . \\ & & & . \\ & H & & 0 \\ 1 & 1 & 1 & \end{bmatrix} \qquad (2)$$

According to the formula S= $e\hat{H}^T$, the syndromes which are corresponding to the non-error and one error patterns could be obtained. And other seven two errors patterns could be obtained from the other syndromes. So the standard array of $\hat{c}$ is constructed. The standard array an [n, k] binary linear code C is a M×N array and for extended array an (8,4) for binary linear code $\hat{c}$ is also M×N array. where M=2^{m-K}, N=2^K.

IV. SIMULINK MODEL OF PROPOSED SYSTEM

OFDM system as in Fig.1, which adopts it is DQPSK or DBPSK modulation, 512 sub carriers and linear block codes with extended hamming. During the period of transition the generator matrix G is adopted. The code 512 bit makes 1024 possible values by generator matrix.

Generator Matrix G:

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \qquad (3)$$
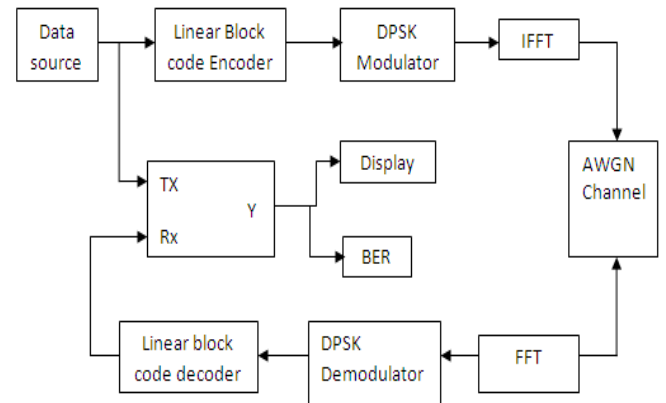
Fig.1 Simulink Model of linear block code of OFDM system

V. SIMULATION RESULTS

This section analyzed the linear block code OFDM system with DPSK modulation technique. The decrease in bit error rate performance of OFDM system is investigated in different channels. Linear block code reduces the effects of both fading and interference. As number of subcarriers increasing the frequency selective fading is reduce. The coherence bandwidth is greater than the subcarrier spacing of the channel, each subcarrier is affected by a flat channel and channel equalizer not needed.

AWGNCHANNEL

The data is subjected to Additive White Gaussian Noise (AWGN) with Signal to noise vector (Es/No) is varied according to the modulation techniques used. In complex signals, the AWGN channel block relates Es/No.
Following algorithm for bit error rate it gives good performance under different channels:

N is number of subcarrier , Input: tx and rx ,Y is output.

```
Algorithm for bit error rate
function y = fcn(tx,rx)
error=0;
for i=1:N
    if(tx(i)~=rx(i))
        error=error+1;
    end
end
ratio=error/N;
y=[error,ratio];
end
```

Specific parameters are used for simulation purpose as shown in table 1. So it gives better performance to reduce bit error rate under various channels.

TABLE 1: Simulation Parameter

Data source	512
Modulation	DPSK
FFT Size	512
Coding technique	Linear block Code
Message Length	4
Codeword Length	8
Channel	AWGN, Rayleigh, Rician
Maximum Doppler Shift	1Hz,2Hz
Rician Parameter	K=1,2,3,4

Fig. 2 is a graph of uncoded and coded of OFDM system under AWGN channel. As compare after LBC coding with extended hamming of OFDM system gives good performance of bit error rate.

It is observed in Fig. 3 that as the number of carriers are increased the bit error rate of OFDM decreases under AWGN channel using linear block code with extended hamming code. For N=64 the bit error rate 10^{-1} dB as the number of carriers is increased the BER decreases to $10^{-2.5}$ dB.
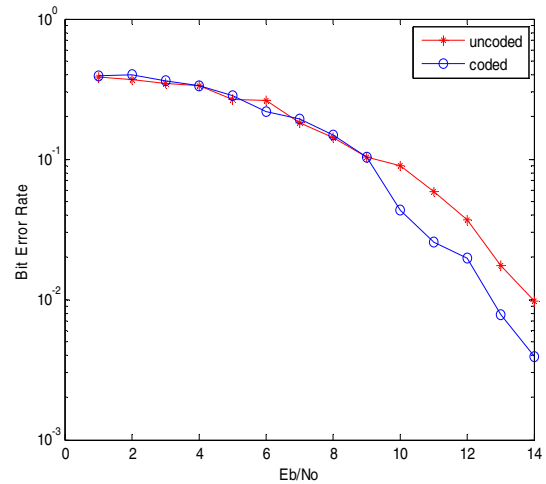
Fig. 2: Bit Error Rate of OFDM under AWGN channel
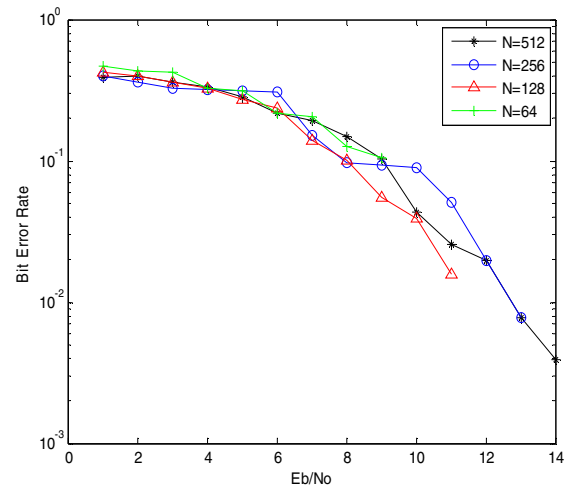
Fig .3: Bit Error Rate of OFDM System for different carriers under AWGN channel using LBC

RAYLEIGH CHANNEL:
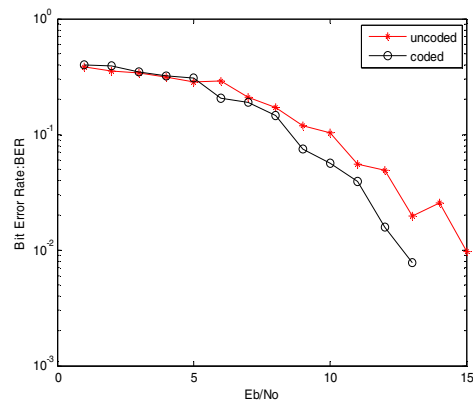
(a)
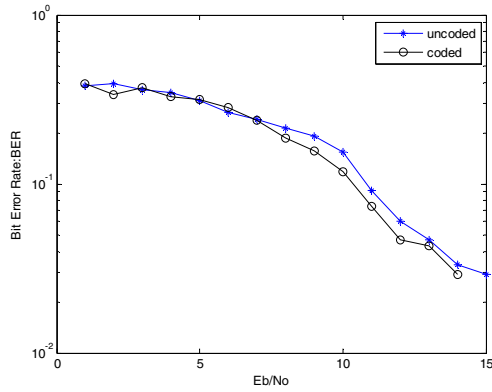
(b)

Fig. 4: Bit Error Rate of OFDM system Under Rayleigh channel using LBC (a)Doppler shift=1hz (b) Doppler shift=2hz
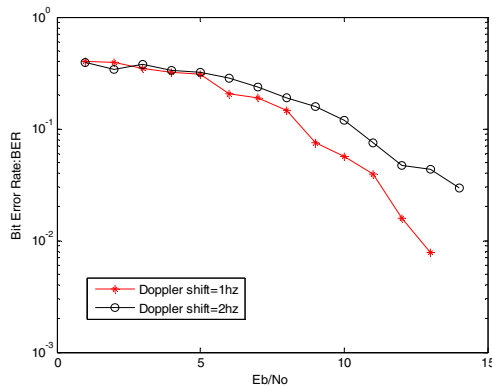
Fig. 5: Bit Error Rate Performance of OFDM system under Rayleigh channel using LBC

The results obtained under the Rayleigh channel using LBC for bit error rate and Eb/No for different values of maximum Doppler shift are showed in Fig.5. Thus it is observed in Fig.4 and Fig.5 that the bit error rate is better for lower values of Doppler shift.

RICIAN CHANNEL:

Fig. 6 exhibits the results of uncoded and coded OFDM system under Rician channel for k=1.Coded graph shows better performance in bit error rate as compare to uncoded OFDM system. The Fig.7 highlights the results for Rician fading channel. It is seen that for various values of Rician parameter 'k' BER varied. The results confirmed that for increase in Rician parameter 'k' factor BER performance got better. Results were obtained by using linear block code with Extended hamming and DPSK Modulation scheme.
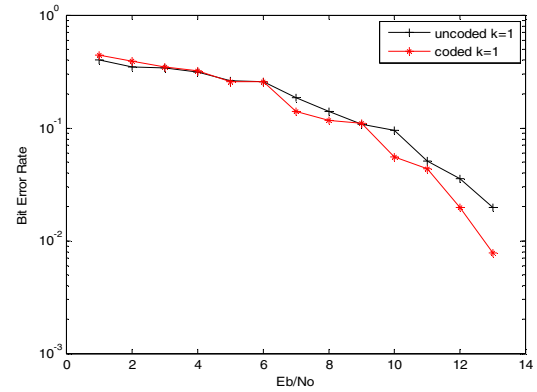
Fig.6: Bit Error Rate of OFDM system under Rician channel For Doppler shift=1hz
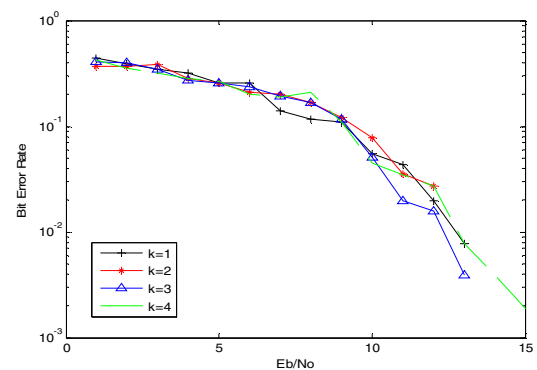
Fig. 7: Bit Error Rate Performance of OFDM system under Rician channel using LBC for Doppler shift=1hz

VI. CONCLUSION

The proposed scheme has used linear block code with extended hamming and differential phase shift keying to improve the BER performance of OFDM system. Linear Block Code with Extended Hamming is used for correcting the two error pattern so as to achieve good BER performance as compare to uncoded OFDM system. In this system Due to DPSK modulation technique and simple structure of the decoder reduce the complexity on the receiver side and no need of equalizer.

REFERENCES

[1] Radio broadcast systems; Digital audio broadcasting (DAB) to mobile, portable and fixed receivers, final Draft pr ETS 300 401, Nov.1994.
[2] J. Bingham, "Multicarrier modulation for data: An idea whose time has come", IEEE communication Mag., pp.5-14, May 1990.
[3] Kyeongcheol Yang and Seok-II Chang, " Peak-To-Average Power control in OFDM using standard Arrays of Linear Block Codes", IEEE communication letters, Vol.7, No.4, pp. 174-176, April 2003
[4] Marvin k. Simon, Mohamed- Slim Alouini, "Digital communication over fading channels", Second Edition, Wiley and Sons Publication

[5] Ying Shan Li and Jin-Soo Park, Nonlinear analysis of the phase noise in the OFDM communications, Vol.52, No.8, August 2004

[6] Orlands Grigoriadis ,H. Shrikanta kamat,"BER Calculation using Matlab simulation for OFDM Transmission",Proceeding of the international Multicarrier of Engineers and Computer Scientists IMECS 2008,Hong kong, vol II, pp. 978-988,19-21 March 2008.

[7] Abhishek katariya, Neha Jain, Amita Yadav,"Performance elevation criteria of RS Coded OFDM transmission over noisy channel'', International Journal of Soft Computing & Engineering (IJSCE) ISSN: 2231-2307, Vol-1, pp.24-28, Issue-3, July 2011.

[8] S.A.Hanna, "Convolutional Interleaving for Digital radio transmissions", ICUPC Conference, pp.443-447,IEEE 1993.

[9] Fumiaki Maehara, " Series Expression of BER Performance for DQPSK/OFDM Signals employing Selection Combing Diversity Reception over Nonlinear Fading Channels" IEEE 2005.

[10]Rappaport, T.S., Wireless Communication Principles and Practice , IEEE Press, New York, Prentice Hall

[11] L. Hamzo, T. Keller, "OFDM and MC-CDMA A Primer, John Wiley and Sons ltd. pp. 43.

[12] Kyeongcheol Yang and Seok-II Chang, "Peak-To-Average Power control in OFDM using standard Arrays of Linear Block Codes", IEEE communication letters, vol.7,No.4,April 2003,PP 174-176.

[13] Ranjan Bose, "Information theory, coding and cryptography", Tata Mc-Graw Hill, Second Edition.

[14] F. J. MacWilliams & N. J. A. Sloane, The Theory of Error Correcting Codes, New York: North –Holand, 1977.

Position Tracking of Automatic Rack and Pinion Steering Linkage System Through Hardware In The Loop Testing

[1]Mohd Zakaria Mohammad Nasir, [2]Abdurahman Dwijotomo, [3]Mohd Zubir Amir

[1,2,3] Department of Automotive,
Faculty of Mechanical Engineering,
Universiti Teknikal Malaysia Melaka,
Ayer Keroh, Melaka
mzakaria@utem.edu.my, arjay.mix88@gmail.com,
zakulive85@gmail.com

[4]Mohd Azman Abdullah, [5]Muhammad Zahir Hassan,
[6]Khisbullah Hudha

[4,5] Faculty of Mechanical Engineering,
Universiti Teknikal Malaysia Melaka,
[6]Faculty of Mechanical Engineering,
Universiti Teknologi Petronas,Perak Darul Ridzuan.
mohdazman@utem.edu.my, zahir@utem.edu.my,
khisbullah@gmail.com

Abstract— **Vehicle handling behavior is much influenced by the performance of steering system and its mechanism. Steering linkage play a very important role in maneuvering of a vehicle. In this paper, a set of kinematic relations of rack and pinion steering linkage system are modeled in MATLAB Simulink environment based on kinematic model equations is presented to study the relationship between steering wheel and tire angle. A Hardware-in-the-loop Simulations (HILS) test rig with actual rack and pinion mechanism has been set up using real time software environment from MathWorks namely xPC Target, LVDT and encoder sensors installed for data measurement at various steering angle. Results from simulation model demonstrate a linear pattern occurred from maximum lock-to-lock steering wheel angle and it is closely follow the sine input trend through HILS experiment with acceptable error.**

Keywords- Rack and pinion, Steering linkage, HILS

I. INTRODUCTION

Steering a car is a driver responsibility as guidance for motor vehicle direction. It involves the driver looking ahead at the intended path relative to the car and somehow processing the preview information and the current position data to yield the steering wheel or control inputs needed to make the car follow the desired path.

Steering linkages play a very importance role in maneuvering of a vehicle. Amongst the steering linkages, the rack and pinion is the most popular and widely used in automotive passenger vehicle [1-4]. Thus it is chosen in this study. This linkage consists of two steering arms (wheel knuckle), two tie rods end as well as a rack and pinion as depicted in Fig. 1. In future automatic steering system, the electric power steering is a vital component for control and improving vehicle handling and stability [5, 6].

The rack and pinion system convert the rotational motion of the steering wheel or driver input into linear motion required to turn the tire, and provide gear reduction or steering ratio, making easier to turn the tire. In most passenger vehicle, approximately three or four complete revolutions of steering wheel is required for the tire maximum turning lock-to-lock from far left to far right. The pinion gear is directly attached to the steering shaft which connected to steering wheel. As a result, when the steering wheel is rotated, the pinion shafts push the rack moving either left or right where lead the tires turning into the desired direction.

In this paper, six bar rack and pinion type is modeled in MATLAB Simulink to predict the correlation between steering wheel angle, steering rack and tire angle.
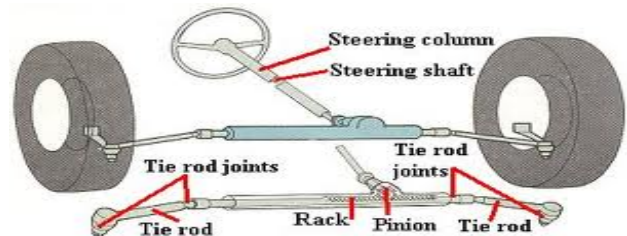
Figure1. Rack and pinion steering linkage in passenger vehicle

II. STEERING LINKAGE KINEMATIC EQUATIONS

The rack and pinion linkages are complex and spatial mechanisms because of the parameter of caster angle and kingpin inclination are in XZ and YZ plane. However, those parameters have little influence on the functionality of the steering linkage [1, 2]. As suggested in literature [7-10], the caster angle and the kingpin inclination provided to correlate suspension and steering system can be neglected. Accordingly, the actual rack and pinion steering linkage can be modeled as a planar linkage as shown in Fig. 2. In this paper, only planar modeling is demonstrated for better understanding of the modeling approach.
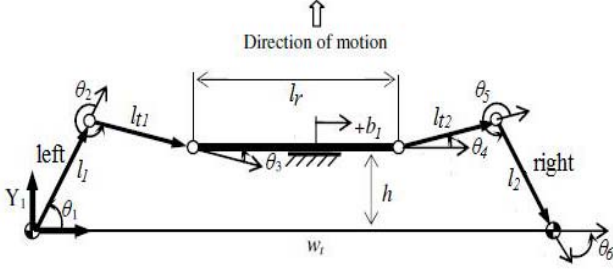
Figure 2. Six-bar rack and pinion steering linkage model. [Hanzaki *et al.* (2007)]

The rack and pinion linkage model is formulated the kinematic equations as a six bar planar linkage [1]. The kinematic equation for the rack and pinion can be written in Eq. (1) where w_t is the wheel track, l_r is rack length and b_1 is the rack travel show in Fig. 2.

$$b = \frac{w_t - l_r}{2} + b_1 \qquad (1)$$

From Fig. 2, it produces c=cosine and s=sine functions in term of scalar components where h is distance from front wheel axis, l_t is tie rod length and l is steering arm length.

$$l_1 c\theta_1 + l_{t1} c\theta_{12} = b \qquad (2)$$

$$l_1 s\theta_1 + l_{t1} s\theta_{12} = h \qquad (3)$$

$$\theta_{12} = \theta_1 + \theta_2 \qquad (4)$$

Equations (1) and (2) can be expressed as $\theta_1 = atan\,2(s\,\theta_1, c\,\theta_1)$, where

$$s\theta_1 = \frac{2z_1}{1 + z_1^2} \qquad (5)$$

$$c\theta_1 = \frac{1 - z_1^2}{1 + z_1^2} \qquad (6)$$

$$z_1 = \tan\frac{\theta_1}{2} \qquad (7)$$

The value of z is the solution of a quadratic equation that obtained from the kinematic equations in (1) and (2). The equation can be written as

$$z_1 = \frac{h + \sqrt{h^2 - k_1^2 + b^2}}{k + b} \qquad (8)$$

$$z_2 = \frac{h - \sqrt{h^2 - k_1^2 + b^2}}{k + b} \qquad (9)$$

$$k = \frac{h^2 + b^2 + l_1^2 - l_2^2}{2l_1} \qquad (10)$$

III. STEERING SYSTEM MODEL

The steering system is modeled in MATLAB Simulink environment based on actual dimension of rack and pinion steering linkage from Malaysian National Car. The steering system model is developed based on the above equations. In this study, a planar model is developed for the standard configuration of the rack and pinion steering linkage by eliminating the kingpin inclination, caster angle and kingpin offset.

The mathematical modeling considered in this study consists of a DC motor model, rack and pinion model and the kinematic model of the steering system. A motor actuator is modeled in this steering system to represented steering input. Fig. 3 shows the pinion as an input and the wheel angle as an output which than function as an input for vehicle model. Some modeling assumptions considered in this study are as follows; the effect of steering inertia is neglected, the front left and front right wheels are assumed to have identical tire angle under excitation from steering input, the efficiency of the DC motor shaft to pinion rotating shaft is neglected.
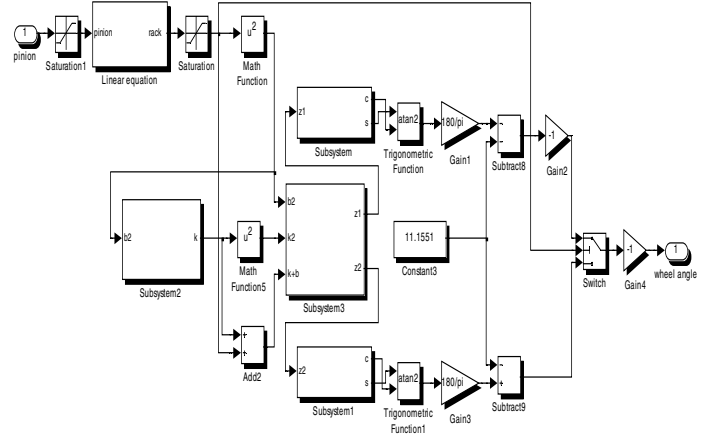
Figure 3. Rack and pinion steering system model in MATLAB SIMULINK

IV. HARDWARE-IN-THE-LOOP SIMULATION (HILS) TEST SETUP

The relation between rotation of pinion and displacement of rack can be defined by perform an experimental on actual rack and pinion steering system through Hardware-in-the-loop simulation (HILS). It can be useful as a development tool since it more repeatable and cost effective than a full in-vehicle test [10, 11]. The front tire is set in normal position, and Linear Variable Differential Transformer (LVDT) and rotary encoder sensors used for measurement. The DC motor is represented as an actuator to generate desired steering wheel angle and it allowed rotating about its rotational axis.

The rotational motion of the pinion is converted into linear motion by the steering rack. The linear motion of the steering rack is allowed to rotate the front tires at its steering axis. The actual steering rack of a compact passenger car is used in HILS testing shown in Fig. 4 using MATLAB xPC Target environment. In HILS testing, rotary encoder was used to measure the rotation angle of the steering wheel and/or DC motor pinion angle. Meanwhile, the LVDT was used to measure the linear displacement of the steering rack which than causes the tire turning into the desired angle.
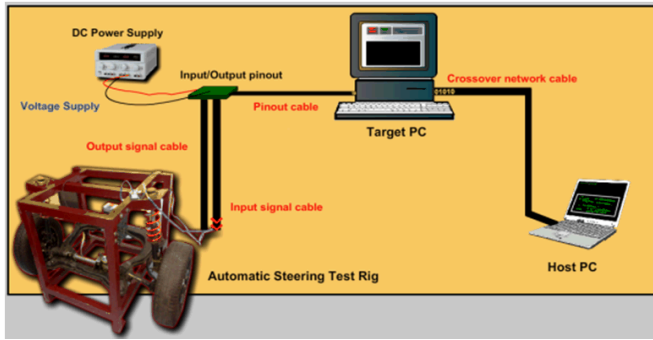
Figure 4. Hardware-in-the-loop- Simulation (HILS) diagram
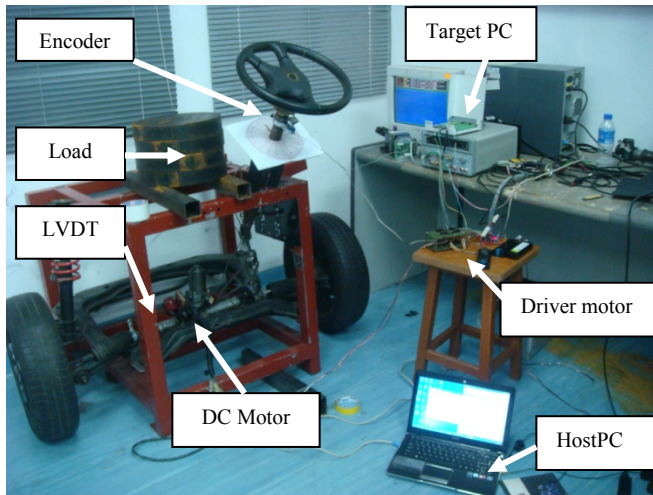
Figure 5. Experimental set up of HILS for rack and pinion steering linkage

The output from both sensors (LVDT and rotary encoder) is transferred into xPC Target in Target PC. The output data from experiment will be compared with steering system model developed in MATLAB Simulink environment.

Position tracking of DC motor is required to perform to ensure the control system design for motor is good enough to follow the desired steering input. Fig. 6 show the controller of DC motor is developed for position tracking in HILS testing. This test is performed on actual rack and pinion steering linkage with both tire contact on ground and carry 150 kg load for engine weight representation.

xPC Target was used to perform model identification for DC motor which must have capabilities to run under various degree of steering input. Signal generator used to represent steering input such as sine 30° mean rotate the steering wheel 30° clockwise and anticlockwise. The desired steering position from the signal input given is compared with the actual angle position of the motor rotating shaft measured by rotary encoder, which results in position deviation error denoted by *e(t)*. The input signal of DC motor is voltage, U(t), which is represented as the required rotational DC motor speed that is commanded to a pulse generator's block. The error is weighted PID controller value via try and error approaches which resulting in current voltage.

$$U(t) = K_p(t)\, e(t) + K_i(t) \int e(t)\, dt + K_d(t) \frac{d}{dt} e(t) \qquad (11)$$
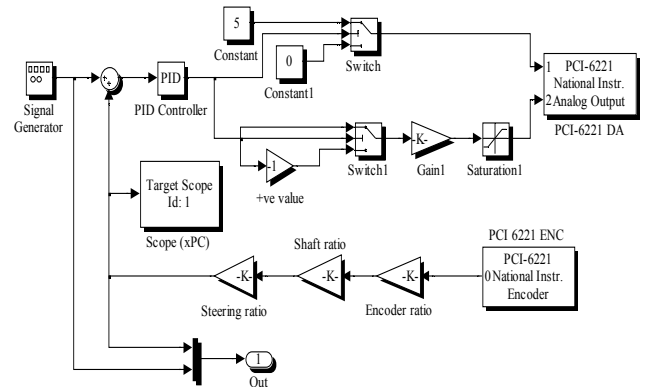
Figure 6. Position tracking controller developed for HILS

V. SIMULATION AND EXPERIMENT RESULTS

The position tracking test for DC motor is performed to ensure the motor could follow the desired steering input. When steering wheel is rotated by giving input signal, the DC motor tend to follow the desired signal by rotating the pinion angle which than pull or push the tie rod end to change the tires angle. Fig. 7 shows the position tracking test for sine wave input for 30°. It demonstrates the actual rack and pinion measured through LVDT sensor and pinion angle closely follow the desired angle of steering input with acceptable error. The control structure therefore can be used to test the six bar rack and pinion linkage system for various steering input.
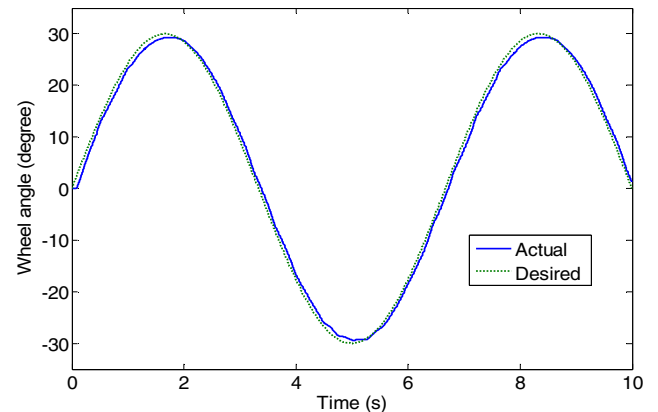
Figure 7. Positioning tracking for DC motor

The result of the pure algorithm simulation of steering system developed in MATLAB Simulink and the Hardware-in-the-Simulations (HILS) test are respectively shown as Fig 8 and Fig 9. The sine steer input parameter at 90° clockwise and anticlockwise at high frequency was used to represent steering wheel rotation angle.

It is demonstrates the steering system model developed with six bar linkage follow the desired angle with acceptable position error. Fig 9 shows the results gathered from HILS testing where the steering system follow the desired trajectory also with less error in average compared with simulations

result. It is evident that the developed HILS model can simulate better for real drive system.
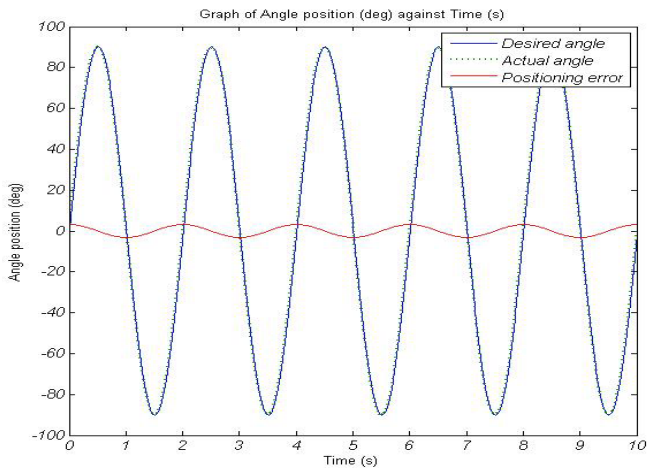
Figure 8. Simulation result of sine input for automatic steering system
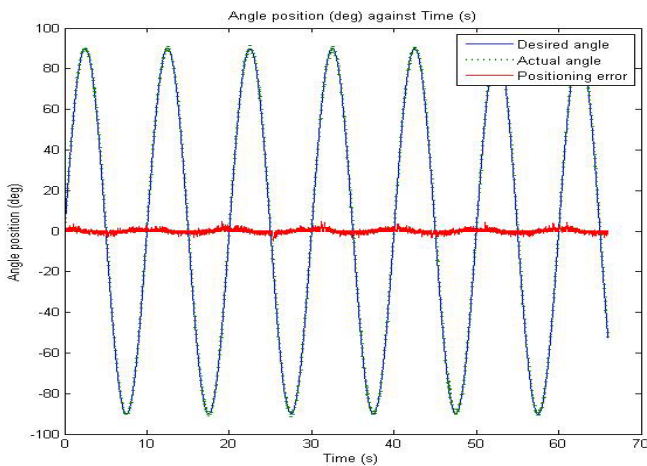
Figure 9. HILS result of sine input for automatic steering system

Figure 10 shows simulation and experimental results for rack and pinion steering linkage system. Y-axis represents the steering rack displacement and X-axis represents the steering wheel angle. The linear trend occurred when the steering wheel is rotate maximum lock-to-lock (610° clockwise and 610° anticlockwise). The dash line in the plotted graph shows a nonlinear behavior from HILS experimental data at 200° anticlockwise. This is due to gear backlash occurred in the steering linkage mechanism. Gear pairs like rack and pinion need clearance or gap between gear teeth matching to avoid wear. When pinion are rotated, slackness will happen at point where it has clearance between pinion and rack. This happen at the range of 22 mm to 25 mm as referred to Fig. 10. To simplify the relationship between the rack and pinion steering mechanisms, the rack displacement is assumed directly proportional to the pinion rotational angle with relationship $Y = 0.00005X - 0.001$.
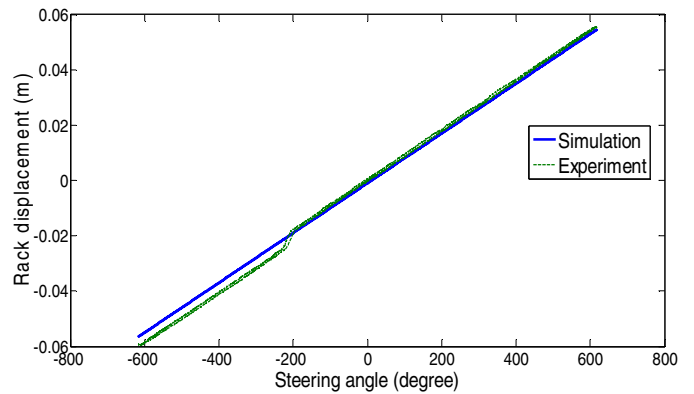
Figure 10. Relationship between rack displacement and pinion angle in MATLAB and HILS Steering System

From steering system model, the correlation between displacement of rack and wheel angle can be defined as shown in Fig. 11. The maximum rack displacement from normal position is 59.5 mm provide maximum tire angle 30.5°. From Fig 10 and Fig 11, the ratio between steering wheel angle and tire angle is 1:20
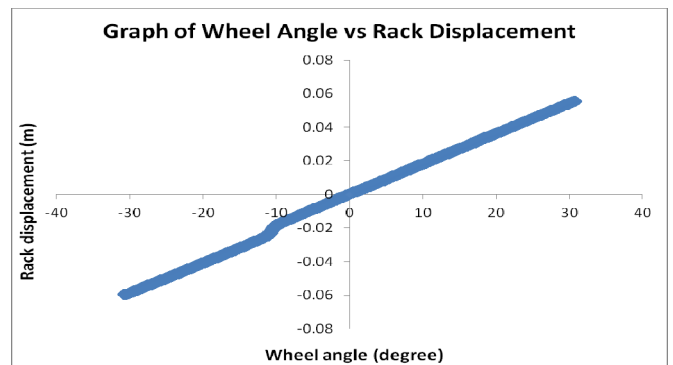
Figure 11. Relationship between rack displacement and tire angle in Steering System

VI.   CONCLUSIONS

The actual dimension of rack and pinion steering linkage system is modeled in MATLAB Simulink environment based on kinematic equations. The relationship of steering rack displacement and pinion angle is achieved by perform simulation and HILS experimental measurement for actual rack and pinion steering linkage system. A result from sine input simulation demonstrates a linear pattern where it's closely follows the trend occurred through HILS experiment with acceptable error. The relation between rack displacement can be assumed directly proportional to the pinion angle and the ratio between steering wheel angle and tire angle is 1:20.

REFERENCES

[1] A. Rahman Hanzaki, S. K. Saha, P. V. M. Rao, "Modeling of rack and pinion steering linkage using Multi-Body Dynamics" 12th International Federation for the Promotion of Mechanism and Machine (IFToMM) Congress, Besancon, France, June 2007.

[2] Naresh Kamble and S. K. Saha, "Developing a virtual prototype of rack and pinion steering system", International Journal Vehicle System Modeling and Testing, Vol. 2, No.1, 2007.

[3] Bishop. R, "A survey of intelligent vehicle applications worldwide", Proceedings of the IEEE Intelligent Vehicle Symposium, pp 25 -30, 2000.

[4] Emm Ping, K. Hudha, Hanif Harun, Hishamudin Jamaludin, "Hardware-in-the-loop simulation of automatic steering", IEEE International Conference on Control Automation, Robotics and Vision, .Singapore, pp. 964-969, 2010.

[5] Hao Chen, Yali Yang, Ruoping Zhang, "Study on electric power system based on ADAMS", Advance in Control Engineering and Information Science, Procedia Engineering 15(2011) 474-478, Elsevier, 2011.

[6] Bishop.R, "Intelligent vehicle applications worldwide", IEEE Intelligent System and their Applications, Vol 15, pp 78-81, No.1, 2000.

[7] Baxter, J., Wou, J. S., Oste, T. D. , "Modeling of mesh friction and mechanical efficiency of rack and pinion steering system", Steering and Suspension Technology Symposium, pp 45-56, 2001.

[8] T. G. Rao, S.K. Saha, I.N. Kar, 'Sensor-actuator based smart yoke for a rack and pinion steering system', SAE International, 2004

[9] H. A Rahmani, S. K. Saha, P. V. M. Rao, 'Analysis of a six-bar rack-and-pinion steering linkage', International Mobility Engineering Congress & Exposition 2005, SAE India Technology for Emerging Markets , 2006.

[10] Youchun Xu, Xiao Wang, Xianbing Zeng Yongsheng Peng Yongjin Zhang, Yi Yuan, "A Hardware-in-the-loop Simulation test- platform for the intelligent vehicle's steering control system", IEEE, 2009.

[11] Mathew Lawson and Xiang Chen, "Hardware-in-the-loop simulation of Fault Tolerant Control for an Electric Power Steering System", IEEE Vehicle Power and Propulsion Conference (VPPC), Harbin , China, September 2008.

Implementation Model Predictive Control (MPC) Algorithm-3 for Inverted Pendulum

Cahyantari Ekaputri

Department of Electrical Engineering
Bandung Institute of Technology
Bandung, Indonesia
cahyantari.ekaputri@gmail.com

Arief Syaichu-Rohman

Department of Electrical Engineering
Bandung Institute of Technology
Bandung, Indonesia

Abstract— **Model Predictive Control (MPC) is based on the idea to produce control input as a solution to real-time optimization problem. Optimization itself is based on the system model. MPC is used to solve multivariable control problem which has more than one variable that may have significant effect on the process. The advantage of MPC is that MPC can works effectively within constraints of the real actuator which are relatively narrow. The disadvantage of MPC lies on its complex algorithm that needs longer time than the other controller. This paper discusses the MPC application for inverted pendulum. The algorithm used is algorithm-3 for floating point. The general guidelines are modeling the system, designing the system using MATLAB, simulating the design using MATLAB, implementing the MPC algorithm on AVR ATmega32 and analyzing the result. After the guideline have been implemented, it can be seen that MPC algorithm-3 with 9 horizon and 1 iteration of quadratic programming can return the pendulum rod at the balance point by the time external force applied.**

Keywords—Model Predictive Control (MPC); algorithm-3; inverted pendulum; AVR ATmega32

I. INTRODUCTION

Model Predictive Controller (MPC) is a control theory which is quite often used to control the complex system domain. MPC has been used in recent years in industrial processes. MPC is very effective to improve the optimization and address the problem of control is usually in the form of the constraints which is owned by the actuator. These constraints are the problem in the design of the system as most systems operate optimally in the working area close to the limit. The main principle of MPC is how to select appropriate control measures to solve the problem of repetition in the optimum control. The goal is the minimization of the performance criteria of the system to a certain horizon, which is usually caused by limitations in the input and output of future behavior which computed based on a process model. The problem that encountered often in the design of MPC is that how to guarantee the stability of the closed loop system, cope with the uncertainty of the model was designed, and reduce the computation performed.

In this paper, the application of MPC will be implemented in microcontroller for controlling the inverted pendulum ECP model 505.

This paper has five sections, introduction section, basic theory section, design and simulation section, implementation and analysis section, and conclusion section.

Section I introduces the general information about this paper. Section II discusses the basic concept of MPC, the computation of MPC tracking problem and MPC algorithm-3 structure. Section III discusses the design and implementation; the plant which is inverted pendulum ECP model 505 will be model and the design of the MPC controller. Section IV discusses the system implementation (including hardware and the algorithm) and the analysis of implemented system. Section V contains the conclusion based on the test from section IV.

II. MODEL PREDICTIVE CONTROL (MPC)

A. The Concept of MPC

MPC will calculate a value which will be used for the input process as a real-time solution of the optimization problem. MPC predict the next output from a process in a certain length of the horizon of a time with using the input and output from previous process. After that the calculation for the control signal in that horizon of time and optimization will be start. Figure 1 show the MPC base diagram of the optimization problem.
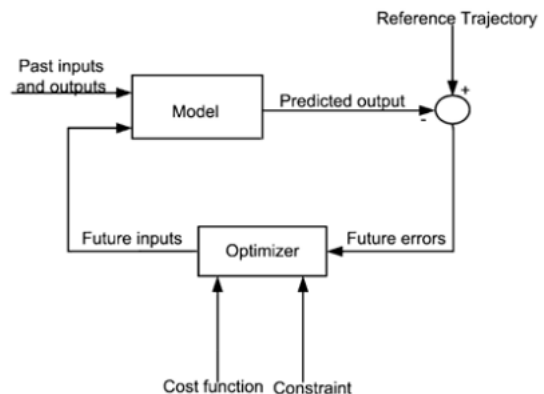
Figure 1. MPC Base Diagram

Figure 2 shows the illustration of MPC base concept which predicts the behavior of plant within M horizons. The prediction's accuracy depends on the accuracy of plant model.

MPC computes the optimal output (control signal) for each input that will concern the constraints. The responses from future control signal computed as problem solvers for n-horizons open loop which determined with first element from optimal control signal responses that used as the next process inputs, meanwhile the other elements will not be used. After that the time will shifted and the input will be computed again so that the computation acquires a new optimal solution that will be used for the next computation. This process always repeated so the output will be optimal.

.
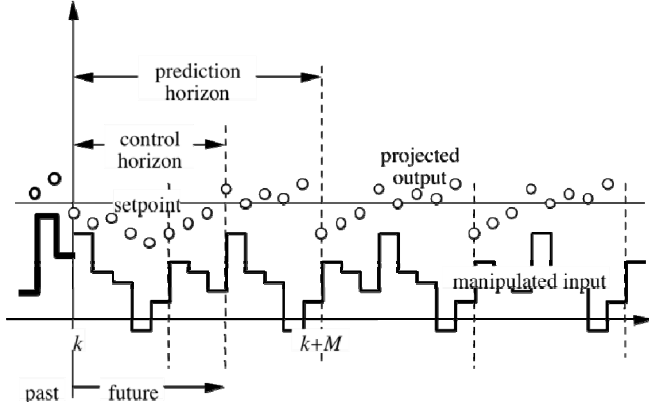
Figure 2. MPC Base Concept

B. MPC Tracking Problem

Consider a detectable plant

$$x_{k+1} = \tilde{A}x_k + \tilde{B}u_k \qquad (1)$$

$$y_k = \tilde{C}x_k \qquad (2)$$

where $x_k \in R^n$, $y_k \in R^l$, and $u_k \in R^m$.

The output y_k is required to track an input reference $w_k \in R^l$. To achieve zero tracking error, an integrator in the form of

$$u_k = u_{k-1} + \Delta u_k \qquad (3)$$

is augmented to the plant so that the augmented plant can be written as

$$\bar{x}_{k+1} = \begin{bmatrix} x_{k+1} \\ u_k \end{bmatrix} = A\bar{x}_k + \Delta Bu_k \qquad (4)$$

$$y_k = C\bar{x}_k \qquad (5)$$

where $A := \begin{bmatrix} \tilde{A} & \tilde{B} \\ 0 & I_m \end{bmatrix}$, $B := \begin{bmatrix} \tilde{B} \\ I_m \end{bmatrix}$, and $C := [\tilde{C} \quad 0]$.

Define the objective function as

$$J = e_{k+Ny}^T \bar{Q}_N e_{k+Ny} + \sum_{t=0}^{Ny-1} e_{k+t}^T \bar{Q} e_{k+t} + \sum_{t=0}^{Nu-1} \Delta u_{k+t}^T \bar{R}\Delta u_{k+t} \qquad (6)$$

where $e_k = y_k - z_k$, N_y = prediction horizon, N_u = control horizon, $N_u \le N_y$, $\bar{Q} = \bar{Q}^T \ge 0$ and $\bar{R} = \bar{R}^T \ge 0$. For the simplicity, assume $N = N_u = N_y$ to define

$$Y = \begin{bmatrix} y_{k+1} \\ y_{k+2} \\ \vdots \\ \cdot \\ y_{k+N} \end{bmatrix}, = \begin{bmatrix} w_{k+1} \\ w_{k+2} \\ \vdots \\ \cdot \\ w_{k+N} \end{bmatrix}, U_d = \begin{bmatrix} \Delta u_k \\ \Delta u_{k+1} \\ \vdots \\ \cdot \\ \Delta u_{k+N-1} \end{bmatrix} \qquad (7)$$

and have

$$Y = \Phi \bar{x}_k + \Gamma U_d \qquad (8)$$

$$J = J_1 + (Y - W)^T Q(Y - W) + U_d^T R U_d \qquad (9)$$

with

$$J_1 = (y_k - w_k)^T \bar{Q}(y_k - w_k) \qquad (10)$$

$Q = \text{diag}\{Q_i\}$, dengan $Q_i = \bar{Q}$ untuk $i = 1,2,\dots,N$

$R = \text{diag}\{R_i\}$, dengan $R_i = \bar{R}$ untuk $i = 1,2,\dots,N$

$$\Phi = \begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^N \end{bmatrix} = \begin{bmatrix} \tilde{C}\tilde{A} & \tilde{C}\tilde{B} \\ \tilde{C}\tilde{A}^2 & \tilde{C}(\tilde{A}+I_n)\tilde{B} \\ \vdots & \vdots \\ \tilde{C}\tilde{A}^N & \tilde{C}\left(\sum_{i=0}^{N-1}\tilde{A}^i\right)\tilde{B} \end{bmatrix} := [\Phi_1 \quad \Phi_2] \qquad (11)$$

$$\Gamma = \begin{bmatrix} CB & 0 & \cdots & 0 & 0 \\ CAB & CB & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ CA^{N-1}B & CA^{N-2}B & \cdots & CAB & CB \end{bmatrix} \qquad (12)$$

If the amplitude limit is considered in the equation, firstly must define

$$U_d = DU - E_1 u_{k-1} := GX + DU \qquad (13)$$

with

$$X = \begin{bmatrix} x_k \\ u_{k-1} \\ -Z \end{bmatrix}, G = [0 \quad -E_1 \quad 0] \qquad (14)$$

$$D = \begin{bmatrix} I_m & 0 & \cdots & 0 & 0 \\ -I_m & I_m & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_m & 0 \\ 0 & 0 & \cdots & -I_m & I_m \end{bmatrix} \qquad (15)$$

they give

$$Y - W = [\Phi_1 \quad \Phi_2]\begin{bmatrix} x_k \\ u_{k-1} \end{bmatrix} + \Gamma(DU - E_1 u_{k-1}) - Z \qquad (16)$$

$$Y - W = \bar{\Phi}X + \bar{\Gamma}U \qquad (17)$$

With these defined variables, the beginning equation can be written to

$$J = J_1 + X^T(\bar{\Phi}Q\bar{\Phi} + G^T RG - F_X^T M^{-1}F_x)X + 2J_{To} \qquad (18)$$

where $\bar{\Phi}$ and $\bar{\Gamma}$ are the functions from A, B, C matrices.

$$X := [x_k^T \quad u_{k-1}^T \quad -Z^T] \tag{19}$$

$$M := \bar{\Gamma}^T Q \bar{\Gamma} + D^T RD \tag{20}$$

$$F_1 := D^T \Gamma^T Q \Phi_1 \tag{21}$$

$$F_2 := D^T \Gamma^T Q \Phi_2 - D^T (\Gamma^T Q \Gamma + R) E_1 \tag{22}$$

$$F := [F_1 \quad F_2] \tag{23}$$

$$F_x := [F \quad H] \tag{24}$$

$$H := D^T \Gamma^T Q \tag{25}$$

$$\bar{\Phi} = [\Phi_1 \quad (\Phi_2 - \Gamma E_1) \quad I_N] \tag{26}$$

$$\bar{\Gamma} = \Gamma D \tag{27}$$

$$E_1 = [1 \quad 0 \quad ... \quad 0] \tag{28}$$

The benefit of computation is to find the U variable which minimize

$$J_{To} := \frac{1}{2} (U + M^{-1} F_X X)^T M (U + M^{-1} F_X X) \tag{29}$$

where $1 u_{min} \leq U \leq 1 u_{max}$

with definition of this equation first

$$U_O := -M^{-1} F_X X = -M^{-1} [F \quad H] \begin{bmatrix} x_k \\ -Z \end{bmatrix} \tag{30}$$

According to the quadratic equivalent programming, MPC tracking can be implemented as the block diagram at Figure 3.
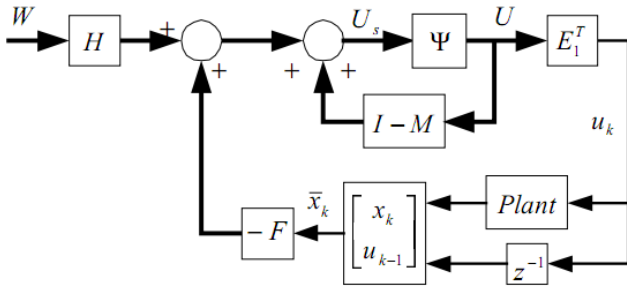
Figure 3.   Block Diagram of the MPC Tracking Implementation

C. Algorithm-3

Algorithm-3 is the iteration of the quadratic programming algorithm. This algorithm can be represented in Figure 4.
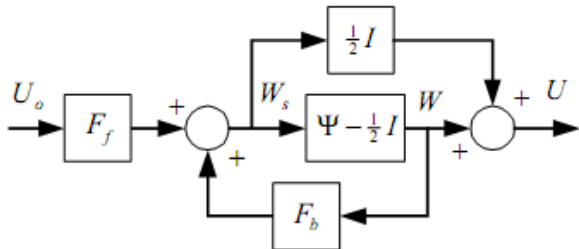
Figure 4.   Block Diagram of the Algorithm-3

The iteration of the algorithm can be seen if there is a feedback. Figure 4 can be written to

$$W_s = U_o + F_b W = U_o + (M^{-1} - I)W \tag{31}$$

$$V = \Psi_b(W_s) \tag{32}$$

The definition of the iteration implementation from that variable is

$$W_{k+1} = \Psi_b(U_o + (M^{-1} - I)W_k) \tag{33}$$

with Ψ_b that a nonlinear function from actuator saturation.

$$\Psi_b(x) := \Psi - \frac{1}{2}I := \Psi_a(x) - \frac{1}{2}x \tag{34}$$

If the iteration equation above has a convergent characteristic, the solution of that loop can be defined as

$$U^* = \frac{1}{2}W^* + U_o + (M^{-1} - I)W^* \tag{35}$$

$$U^* = U_o + M^{-1}W^* - \frac{1}{2}W^* \tag{36}$$

III.   DESIGN AND SIMULATION

A. Inverted Pendulum

The inverted pendulum that will be used is ECP model 505. The design is demanded to be automatically stay on the balance state with a tolerance where close to it. The system will start from the furthest position of pendulum rod from the balance point and will work to keep it on the balance state.
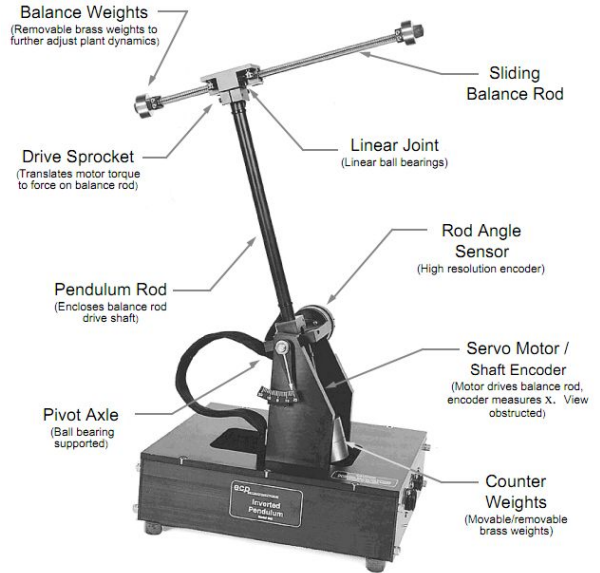
Figure 5.   Inverted Pendulum ECP Model 505

The ECP model 505 has the transfer function shown on (37).

$$\frac{\theta(s)}{F(s)} = \frac{l_0}{J^*} \frac{-s^2 + \frac{g}{l_0}}{s^4 + \left(\frac{(m_1 l_0 - m_2 l_c)g}{J^*}\right) s^2 - \frac{m_1 g^2}{J^*}} \tag{37}$$

118

with

$$J^* = J_0 - m_1 l_0{}^2 \qquad (38)$$

$$J_0|_{x=0} = J_1 + J_2 + m_1 l_0{}^2 + m_2 l_c{}^2 \qquad (39)$$

In this system, there are four state, one input and one output. The input that used in this system is the amount of force which necessary to move the pendulum rod in order to achieve stability when standing vertical $F(t)$, while the output of this system is the large of pendulum angle / position of the pendulum rod to the vertical axis (θ). The states that used in this system written on (40).

$$x = [\theta \quad \dot\theta \quad x \quad \dot x]^T \qquad (40)$$

where

θ = angle / position of the pendulum rod to the vertical axis

$\dot\theta$ = angular velocity while moving the pendulum rod

x = position sliding rod of the pendulum rod

$\dot x$ = shaft speed glide while moving

The parameters of inverted pendulum used for computation are:

l_0 : length of pendulum rod (0.33 m)

l_c : length of mass centre of pendulum rod (0.0281 m)

m_1 : mass of sliding rod (0.213 kg)

m_2 : mass of pendulum rod (1.785 kg)

g : gravitation force (9.81 m^2/s)

J_0 : moment of inertia (0.036 kg.m^2)

State space of the inverted pendulum system can be modeled mathematically in the form of discrete space equation as follows.

$$x_{k+1} = A_{bar}\, x_k + B_{bar}\, u_k \qquad (41)$$

$$y_k = C_{bar}\, x_k + D_{bar}\, u_k \qquad (42)$$

By using the parameter value sampling time 0.04 seconds and compute using MATLAB functions `tf2ss` and `c2d`, the system of mathematical equations in discrete form

$$x_{k+1} = \begin{bmatrix} 1.1303 & 0.6474 & 0 & 0 \\ -0.3354 & 1.1303 & 0 & 0 \\ 0 & 0 & 0.6031 & -0.3473 \\ 0 & 0 & 0.3473 & 0.9092 \end{bmatrix} x_k$$
$$+ \begin{bmatrix} -0.2573 \\ -0.7005 \\ -0.6962 \\ -0.0134 \end{bmatrix} u_k$$

$$y_k = [0.2337 \quad 0.8681 \quad -0.6962 \quad 0.0134]\, x_k + [0]\, u_k$$

The input and output of the equation are trough stabilization realization state space based on Gramian. The parameters that used in this paper define by:

- Sampling time : 0.04 seconds
- $\bar Q = 1$ and $\bar R = 1$
- Set point value : 0
- Simulation time : 16 seconds

B. MPC Designing in MATLAB
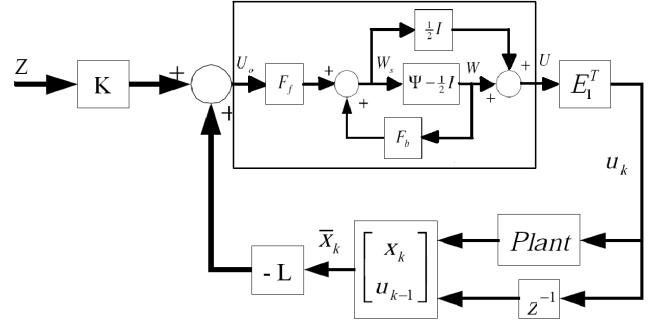
Figure 6.   Block Diagram of the Algorithm-3 Implementation

MPC Designing is calculating the matrices values which become the bases of MPC Algorithm-3 calculation. Calculation of the matrices defines by designing the MPC formula in MATLAB. With simulink in MATLAB, the optimum horizon value calculated for the system with a certain range of iteration value, using the method that analyzing the response time and control signal of the system.

These conditions make the value of 9 horizons has the smallest settling time response system and input system. Table I shows the details.

TABLE I.   SETTLING TIME RESPONSE SYSTEM, INPUT SYSTEM, AND INTERATION NUMBER AT EVERY HORIZON

Horizon	Response Time	Input Control Signal Time	Iteration Number
Horizon 8	13,84 s	12 s	1
Horizon 9	8,92 s	7,2 s	1
Horizon 10	9,2 s	7,48 s	1
Horizon 11	10,36 s	8,36 s	1
Horizon 12	11,76 s	9,44 s	1
Horizon 15	15,2 s	12,2 s	1

The most optimal control signal is reached when the horizon value is 9. The system characteristics shown in Figure 7 and figure 8.

Figure 7 shows the settling time response system which controlled by MPC algorithm 3 with 9 horizons. From the graphic, the response time of system is 8,92 s. Figure 8 shows the input control signal time of system which controlled by MPC algorithm 3 with 9 horizons. From the graphic, the input control signal time of system is 7,2 s.
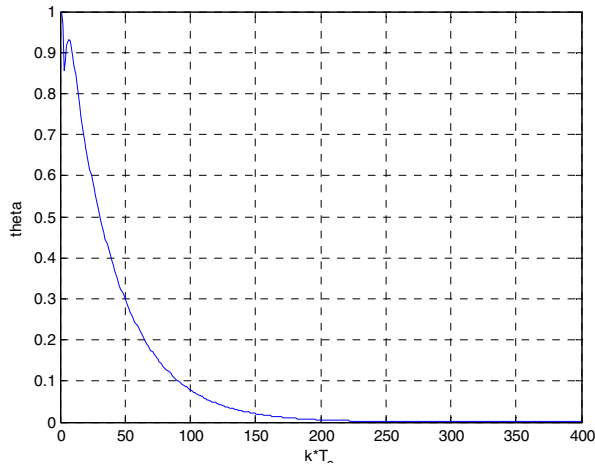
Figure 7.   MPC Algorithm-3 with 9 horizons Response Time
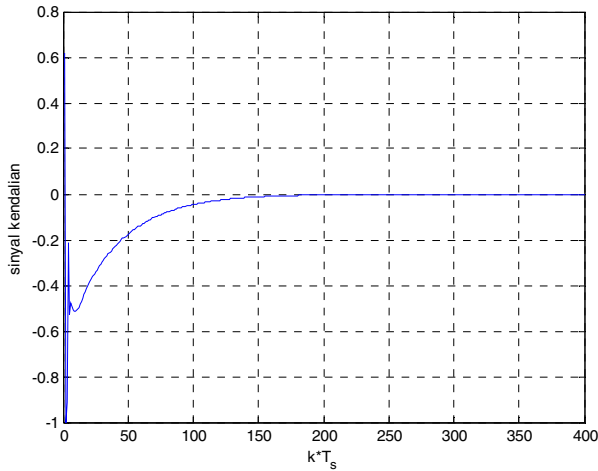
Figure 8.   MPC Algorithm-3 with 9 horizons Input Control Signal

IV.   IMPLEMENTATION AND ANALYSIS

A. Hardware System Design

The hardware system divided into three parts of modules, which are the AVR ATmega32 microcontroller, L298N motor driver IC, and inverted pendulum ECP model 505 as a plant. The block diagram (Figure 9) shows the flow of design.
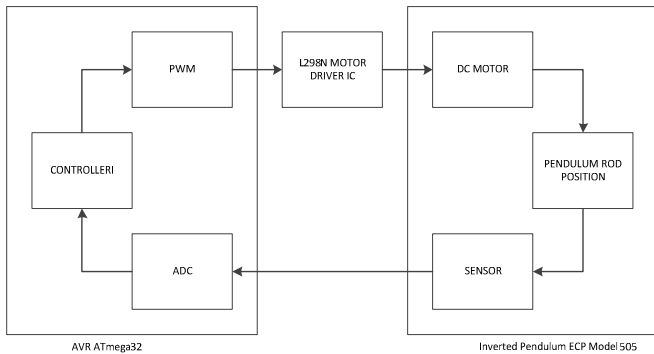
Figure 9.   Block Diagram of the Hardware System

B. System Algorithm Design

The algorithm of the system design by the flowchart in Figure 10. This algorithm will be applied in microcontroller.
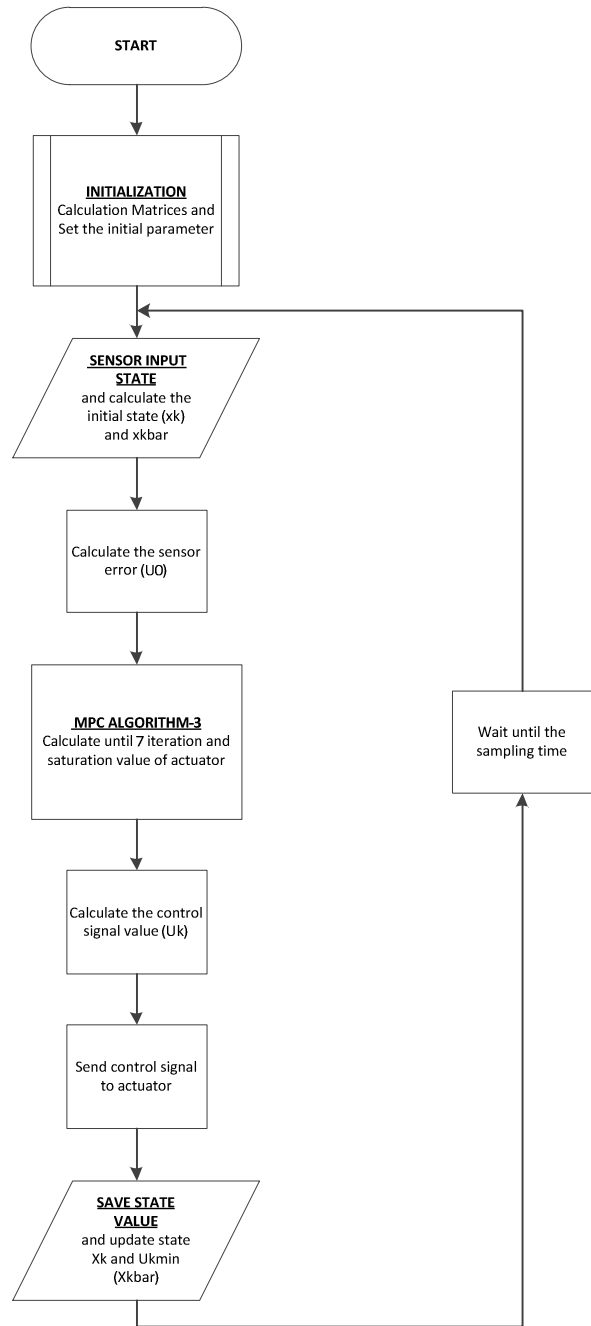
Figure 10. Flowchart of System Algorithm

C. System Implementation

The next step is implementing the design of hardware and algorithm. Then the system will be verified with the result of simulation. The verification step will be divided into three parts, which are input control signal verification, PWM verification and system verification.

The input control signal verification used for verified the input control signal with the input sensor value. From simulation, it's clearly show that if the pendulum rod is around the stable point (ADC conversion value near 0), then the input control signal value will be small (near 0). Otherwise, if the pendulum rod is away from the stable point (ADC conversion value near 1), then the input control signal value will be large. Table II shows the verification.

TABLE II.   INPUT CONTROL SIGNAL VERIFICATION

Sensor Input (ADC)	ADC Conversion Value	Input Control Signal Value
166	-1	-0,66
176	-0,87	-0,57
192	-0,64	-0,42
216	-0,31	-0,2
240	0,03	0,016
267	0,41	0,267
312	1,04	0,68

The PWM verification used for verified the input control signal value with the output PWM from controller (apply to motor DC of the plant). The PWM value is about 1 to 1023. In simulation, if the input control signal value is small, then the output PWM will be small. Otherwise, if the input control signal value is large, then the output PWM will be large. Table III shows the verification.

TABLE III.   PWM VERIFICATION

Input Control Signal Value	PWM Value
-0,66	679
-0,57	585
-0,42	434
-0,2	207
0,016	16
0,267	272
0,68	699

The system verification performed by giving interference of pendulum rod in the stable point, then the pendulum rod will be back to its stable point.

Figure 11 and 12 shows the system verification. In figure 11, the above graphic (CH1) shows the pendulum rod position (angle) and the bottom graphic (CH2) shows the PWM which applied as the input control signal of the system. It's clearly show that if there is an interference in the pendulum rod (pendulum rod position is not in stable point), the MPC controller will calculate the input control signal so that the pendulum rod will be in the stable point. The input control signal will be changed into PWM signal. Then PWM signal will be applied to DC motor in the plant. Figure 12 shows that the settling time of input control signal is 8 seconds (about 200 times with time plotting 0.04 seconds using serial communication in MATLAB).
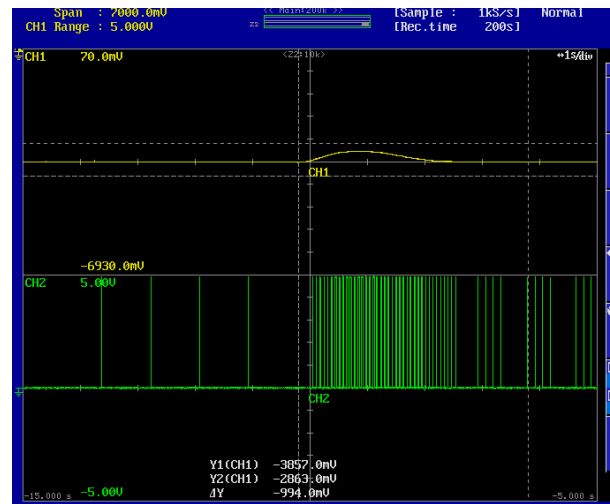
Figure 11. Pendulum Rod Position and PWM Value of System Verification
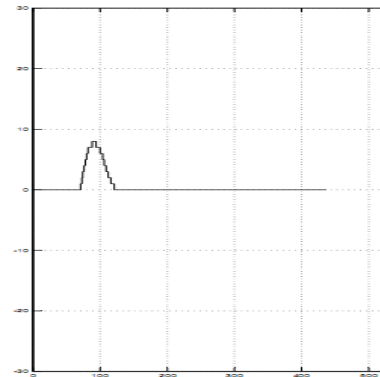
Figure 12. Input Control Signal Plot

D. Analysis

In designing MPC, sampling time is 0.04 seconds. Rule of Thumb says that the ideal sampling of a system is one-tenth of time constant of system (it can be seen from its characteristic equation). Nyquist-Shannon Theorem says the minimal sampling time is half of time constant of system. Based on Nyquist-Shannon Theorem, the chosen sampling time satisfy that theorem.

The settling time of input control signal is 8 seconds. But in the simulation, the settling time is 7.2 seconds. This is caused by delay of system (in actuator, sensor, etc.), the non-ideal actuator (slip, hysteresis, etc.), the accuracy sensor, the modeling system (system is non-linear, but the modeling is linearized) etc.

V.   CONCLUSION

The paper shows that MPC algorithm 3 can be applied to inverted pendulum system ECP model 505 and implemented in AVR ATmega32 microcontroller. The Algorithm-3 calculate the optimal input control system based on the prediction of system's behavior. The algorithm-3 also solve the constraints problem in actuator. The system performed by giving interference of pendulum rod in the stable point, then the pendulum rod will be back to its stable point. For this system, the horizon value set at 9 based on the response time in the

simulation. For further research, the MPC algorithm can be implemented in more complex processor for complex computation.

REFERENCES

[1] Franklin, G. F. dan J. D. Powell, and A. Emami-Naeni (1997), *Feedback Control on Dynamic Systems, 3rd ed*, Addison Wesley, New York.

[2] Gorinevsky, Dimitry. *Model Predictive Control: Industrial MPC,* Stanford University. www.stanford.edu/class/ee392m/Lecture15_MPC.pdf.

[3] Nikolaou, Michael. *Model Predictive Controllers: a Critical Synthesis of Theory and Industrial Needs.* www.chee.uh.edu/faculty/nikolaou/MPCtheoryRevised.pdf.

[4] Ogata, Katsuhiko (1997), *Modern Control Engineering, 3rd ed,* Prentice-Hall, New Jersey

[5] Orukpe, P.E. *Basics of Model Predictive Control*, Imperial College, London. www3.imperial.ac.uk/portal/pls/portallive/docs/1/50918.PDF.

[6] Syaichu-Rohman, Arief dan Richard H. Middleton - *A Multivariable Nonlinear Algebraic Loop.* http://www.nt.ntnu.no/users/skoge/prost/proceedings/ecc03/pdfs/001.pdf

[7] Syaichu-Rohman, Arief, and Middleton, Richard H. *Convergence Study of Some Fixed Point Iteration QP Algorithms*. In Proceedings of the 43rd Conference on Decision and Control. Pradise Island, Bahamas.

[8] Syaichu-Rohman, Arief. *Plant Performance Optimization*. The 6th Asian Control Conference. July 18-21, 2006. Inna Grand Bali Beach Hotel, Sanur, Bali, Indonesia.

[9] Syaichu-Rohman. 2006. *Introduction to Model Predictive Control.* Indonesia: Workshop of the 6th Asian Control Conference.

[10] Taylor, Rod Jason. 2008. *Model Predictive Control dengan Algoritma-2 Quadratic Program pada dsPIC30F4011*. Institut Teknologi Bandung, Indonesia.

Challenges in High Accuracy of Malware Detection

Muhammad Najmi Ahmad Zabidi

Kulliyyah of Information and Communication Technology,
International Islamic University Malaysia,
53100 Gombak,
Kuala Lumpur,
Malaysia
najmi@kict.iium.edu.my

Mohd Aizaini Maarof[†] and Anazida Zainal[‡]

Information Assurance and Security Research Group,
Faculty of Computer Science and Information Systems,
Universiti Teknologi Malaysia,
81310 Skudai, Johor Bahru,
Malaysia
aizaini@utm.my[†], anazida@utm.my[‡]

Abstract—Malware is a threat to the computer users regardless which operating systems and hardware platforms that they are using. Microsoft Windows is the most popular operating system and the popularity also make it the most favourite platform to be attacked by the adversaries. Current detection for Windows relies on the signature based detection which is fairly fast although suffers undetected binaries. Here, we propose a method to increase the detection rate of malware by manipulating machine learning methods. Our focus is on the Microsoft Windows binaries.

Index Terms—malware; feature selection; machine learning

I. Introduction

The work on computer virus formally traced to Fred Cohen's work in 1986[1]. The term "virus" was given by Cohen's supervisor Dr Adleman which the public use until currently. Nowadays, the term malware most likely being used, which covers computer worm, virus, dialers, trojans, rootkits and so many others threat. Hence the term malware which means *MAL*icious soft*WARE* being used. Machine learning methods has been widely used in the area of finance, for example in credit card fraud detections, patient's drug prescriptions and other areas. In the area of malware research, machine learning plays role in several phases, dimension reductions with feature selection for reducing the number of features; without reducing the accuracy rate. It is also able to classify unknown data based on clustering(unsupervised methods). The room of improvements depends on the number or which feature are selected; which means the quality of the chosen features.

A. Problem statement

In order to detect the malicious features within a malware, two ways of detection methods being used; static analysis and dynamic analysis. Static analysis deals with parsing the malware binaries so that the malicious strings could be find as well as by reverse-engineer method; disassembling the malware. Dynamic analysis in the other hands, monitoring the activities of malware by executing them in a safe environment; for example in a virtual machine. Each method has its own strengths and weaknesses and most of the time it is advisable to use both methods to analyze a malware[2]. The rich sets of malicious features could be reduced for malware detection without sacrificing the accuracy. This could save some of the

researcher's time for analysis. Our concern is there is no obvious necessity to detect a malware with many features when it could be reduced to several strong features which could perform the same. In order to begin with the feature selection step for malware, the mechanism or algorithm candidates must be reviewed into. The problem to be solved is how to reduce the number of features in malware detection without reducing the accuracy rate and secondly, how to detect malware which never been seen before by comparing commonalities with the previous malware.

II. Methodology

We obtained a dataset of malware samples from a Malaysian government agency CERT (Computer Emergency Response Team), CyberSecurityMalaysia (CSM) which consists of 2GB size of malware, roughly around 30,000 malware. In order to extract the features, pre-processing methods has to be done prior to the next stage. We use static analysis to parse the visible Application Programming Interface (API) calls and dynamic analysis to capture the obfuscated API calls. The most popular way for malware to invoke commands is by using calls within Dynamically Linked Library (DLL). Two commands within "kernel32.dll" which are "LoadLibrary" and "GetProcAddress" are required to be invoked to perform the DLL calls[3]. General ideas about the process can be referred from [4], [5] and [6]. In order to test the accuracy of the detection, we mix the testing sample with benign software in order to know whether there is false positive or false negative in the detection. Roughly our methods consist of :

1) Feature Selection(Ranking/Pruning)
2) Supervised Classification
3) Unsupervised Classification

Item 2) and 3) above also could be combined to a method known as "Semi Supervised Classification".

A. Feature Selection as a Dimension Reduction Method

We review several methods, for example by using Information Gain(IG) as in [7], [8], [9], [10], [11]. The formula of IG in [10] and [11] is the same except [11] added an error correction method after the result to reduce the error rate, which is basically the area under the bell curve selection.

The amount by which the entropy of X *decreases* reflects additional information about X provided by Y is called *information gain*, given by

$$IG(X|Y) = H(X) - H(X|Y) \qquad (1)$$

[11] introduced the following algorithm to "correct out" error the results.

$$IG(X)' = IG(X) \pm \frac{\sum_{i-0}^{n} IG(X_i)}{n} \qquad (2)$$

While [12] used:

$$IG(t) = \sum_{c \in \{c_i, c_i\}} \sum_{t' \in \{t, t\}} P(t', c) log \frac{P(t', c)}{P(t')P(c)} \qquad (3)$$

We expect to look in the said IG algorithm above, due to the interest by the previous researchers in applying the algorithm for feature ranking purpose.

We already evaluated several malware in [4] by using the following features; GetSystemTimeAsFileTime, SetUnhandledExceptionFilter, GetCurrentProcess, TerminateProcess, LoadLibraryExW, GetVersionExW, GetModuleFileNameW, GetTickCount, SetLastError, GetCurrentProcessId, GetModuleHandleW, LoadLibraryW, InterlockedExchange, UnhandledExceptionFilter, FreeLibrary, GetCurrentThreadId, QueryPerformanceCounter, CreateFileW, InterlockedCompareExchange, UnmapViewOfFile and GetProcAddress:.

Our next target is to evaluate the fitness and quality of the chosen features so that only the most relevant features could be used for representing malicious features. This, theoretically could be done by using distance measurement and data clustering by plotting the details on a graph.

[13] for example argues on the credibility of detection by using API calls. This is true, since API calls is a legit method to call operating system functions as it interfacing operating systems with applications. API calls detection, although very popular in term usage in research papers, is not the only method for measuring the maliciousness level. It could also being examined along with the entropy level, malicious strings existence, anti debugger and anti virtual machine detector strings, RedPill detector, XORed strings and many more.

B. Supervised Learning for Malware Detection

Supervised learning needs an instance label so that it can sort of "predict" the upcoming results based on the rules fed before. This could be efficient for relatively known malware, but has has some problems if the analyzed features were never found.

C. Unsupervised Learning for Malware Detection

Due to the nature of malware rapidness growth, the prospect of using the unsupervised method is high, due to the nature of the unsupervised learning process which does not need any instance label prior to the categorization process. The separation of clusters can be done hence the malware can be segregated according to groups and measuring distance between them.

III. CHALLENGES

Malware writers are usually being regarded as "lazy", which means usually the "new" malware in the wild are actually derived from the previous malware, hence if there is large similarities the new malware will be regarded as the new variant. In addition to that, malware writers also used to insert garbage calls in order to confuse the analyst with fake API calls. This however could be avoided by using different ways of analyzing the malware. Also, malware writers encrypts the important details within the malware body, however this partly already being addressed with our tool in [4], if the malware writer used XOR as the encryption method. Malware also usually being packed, some use well known packer and some rarely used packer. "Packing" is a method to compress a Windows executables without having the user needs to manually decompress them. The purpose is basically legit since it is used by benign Windows executables as well but it has already being exploited by the criminals for compressing the size plus encryptions. Ideally, malware could be analyzed better by unpacking them. However the process of unpacking without the existence of the packing software is difficult and time consuming, hence methods has been developed to analyze malware at the surface level by using entropy analysis [14], [15], [16], [17]. We tested entropy test on the recent sample of Duqu malware[18], [19], which is known as an offspring of the infamous Stuxnet, however Duqu malware seems unsuspicious to the entropy test.

There is one important aspect to consider, according to [20], the process of cyber security experiments are considered impossible to be replicated, due to the nature of sensors and data involved. In malware research for example, usually the usually the malware samples were obtained from VXhaven website[1], a repository for malware enthusiasts. However the VXhaven portal is currently down due to the directive by the Ukranian government[21]. There are two other education institutions that provided their datasets, those are University of Mannheim[2], Germany and Nexginrc[3], which is based in Pakistan.

There are issues in term of analyzing the malware too. Safe environment for malware analysis is needed, in order to ensure the malware that is being analyzed will not infecting elsewhere. One way is by disconnecting the network connection. This method however could work for static analysis process. For dynamic analysis, the connection is required hence several works being done on sandboxes or internet simulator[22], [23], [24]. The virtual environment could also guarantee the safety, but many malware now have the feature of virtual machine detector. This problem has already being addressed and could be detected by our tool in [4].

IV. CONCLUSION

Malware detection is a problem which the researchers have tried to solve for so many years by using enormous types

---

[1]http://vx.netlux.org/index.html
[2]http://pi1.informatik.uni-mannheim.de/malheur/
[3]http://nexginrc.org/Datasets/Default.aspx

of methods. Each of the proposed methods has its own weakness and later being improved by they themselves or other researchers. In the mean time, the malware writers out there are also improving their malware codes from time to time and work more organized in the underground world. This "cat and mouse" game is like vicious circle, as the process of improvement need to be done from time to time.

ACKNOWLEDGMENT

We would like to thank to the Ministry of Higher Education and International Islamic University Malaysia for supporting the main author of this paper. Also to Universiti Teknologi Malaysia for the second and third authors for providing facilities and advises for the contents of this paper.

REFERENCES

[1] F. B. Cohen, "Computer viruses," Ph.D. dissertation, University of Southern California, Los Angeles, CA, USA, 1986, AAI0559804.

[2] D. Glynos, "Packing heat!" http://census.gr/media/packing-heat.pdf, May 2012, (Accessed at 12 May 2012).

[3] J. Dai, "Detecting malicious software by dynamic execution," Ph.D. dissertation, University of Central Florida, 2010.

[4] M. N. A. Zabidi, M. A. Maarof, and A. Zainal, "Malware analysis with multiple features," *Computer Modeling and Simulation, International Conference on*, vol. 0, pp. 231–235, 2012.

[5] ——, "Ensemble Based Categorization and Adaptive Model for Malware Detection," *2011 7th International Conference on Information Assurance and Security (IAS)*, vol. 978-1-4577-2153-3, pp. 80–85, 2011.

[6] M. N. A. Zabidi, "Compiling Features for Malicious Software," http://conference.hitb.org/hitbsecconf2011kul/materials/D1\%20SIGINT\%20-\%20Muhammad\%20Najmi\%20Ahmad\%20Zabidi\%20-\%20Compiling\%20Features\%20for\%20Malcious\%20Binaries.pdf, October 2011, (Accessed at January 11, 2012).

[7] B. Zhang, J. Yin, J. Hao, S. Wang, and D. Zhang, "New malicious code detection based on n-gram analysis and rough set theory," Y. Wang, Y.-M. Cheung, and H. Liu, Eds. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 626–633.

[8] S. Sabnani, "Computer security: A machine learning approach," Master's thesis, 2008.

[9] R. Perdisci, A. Lanzi, and W. Lee, "McBoost: Boosting Scalability in Malware Collection and Analysis Using Statistical Classification of Executables," in *Computer Security Applications Conference, 2008. ACSAC 2008. Annual*, Dec. 2008, pp. 301–310.

[10] A. Altaher, S. Ramadass, and A. Ali, "Computer Virus Detection Using Features Ranking and Machine Learning," *Australian Journal of Basic and Applied Sciences*, vol. 5, no. 9, pp. 1482–1486, 2011.

[11] P. Singhal and N. Raul, "Malware detection module using machine learning algorithms to assist in centralized security in enterprise networks," *International Journal of Network Security & Its Applications*, vol. 4, 2012.

[12] Q. Jiang, X. Zhao, and K. Huang, "A feature selection method for malware detection," in *2011 IEEE International Conference on Information and Automation (ICIA)*, june 2011, pp. 890–895.

[13] S. M. Abdulla, M. L. M. Kiah, and O. Zakaria, "Minimizing errors in identifying malicious api to detect pe malwares using artificial costimulation," in *International Conference on Emerging Trends in Computer and Electronics Engineering (ICETCEE'2012)*, 2012.

[14] X. Ugarte-Pedrero, I. Santos, B. Sanz, C. Laorden, and P. Bringas, "Countering Entropy Measure Attacks on Packed Software Detection," in *Proceedings of the 9th IEEE Consumer Communications and Networking Conference (CCNC2012)*, 2012, in press.

[15] I. Sorokin, "Comparing files using structural entropy," *Journal in Computer Virology*, pp. 1–7, 2011, 10.1007/s11416-011-0153-9. [Online]. Available: http://dx.doi.org/10.1007/s11416-011-0153-9

[16] R. Lyda and J. Hamrock, "Using entropy analysis to find encrypted and packed malware," *Security & Privacy, IEEE*, vol. 5, no. 2, pp. 40–45, 2007.

[17] I. Briones and A. Gomez, "Graphs, entropy and grid computing: Automatic comparison of malware," in *Proceedings of the 2004 Virus Bulletin Conference*, 2004.

[18] B. Bencsáth, G. Pék, L. Buttyán, and M. Félegyházi, "Duqu: A stuxnet-like malware found in the wild," http://www.crysys.hu/publications/files/bencsathPBF11duqu.pdf, 2011.

[19] ——, "Duqu: Analysis, detection, and lessons learned," *ACM European Workshop on System Security (EuroSec) 2012.*, 2012.

[20] T. Dumitras and P. Efstathopoulos, "The provenance of wine," http://www.ece.cmu.edu/~tdumitra/public_documents/dumitras12wineprovenance.pdf, 2012, (Accessed at 12 May 2012).

[21] J. Kirk, "Ukraine shuts down forum for malware writers," http://www.computerworld.com/s/article/9225693/Ukraine_shuts_down_forum_for_malware_writers, March 2012, (Accessed at 12 May 2012).

[22] T. Holz, F. C. Freiling, and C. Willems, "Toward Automated Dynamic Malware Analysis Using CWSandbox," *IEEE Security & Privacy*, vol. 5, no. 2, pp. 32–39, 2007.

[23] K. Yoshioka, T. Kasama, and T. Matsumoto, "Sandbox analysis with controlled internet connection for observing temporal changes of mal-

ware behavior," in *The Fourth Joint Workshop on Information Security (JWIS 2009)*, 2009.

[24] K. Aoki, T. Yagi, M. Iwamura, and M. Itoh, "Controlling malware http communications in dynamic analysis system using search engine," in *Cyberspace Safety and Security (CSS), 2011 Third International Workshop on*. IEEE, 2011, pp. 1–6.

Data Mining Development in "Actual Mobile guide for Tourist" Application

Rifki Wijaya[*1], Ary Setijadi Prihatmanto[*2], Mochamad Vicky Ghani Aziz[*3]

Control System and Computer Laboratory, School of Electrical Engineering and Informatics
Bandung Institute of Technology, Ganeca Street 10, Bandung, Indonesia. Postal code: 40132

[1]rifkiwijaya@gmail.com
[2]asetijadi@lskk.ee.itb.ac.id
[3]vickyaziz@live.com

Abstract—IT and tourism is a new innovative combination, the tourist can get the information of the object they want by only get the picture of the object by using their camera in mobile phone, and then the tourist will get the detail information.
Actual Mobile Guide for Tourist (AMGT) is a Markerless Augmented Reality application for Tourist to guide them about information of the sites. This application required a server to save a picture of tour object as data. Application needs more available data to cover more area. User generated content is suitable for this application. Data Mining from all available source is need to be set up. Application server is needed to crawl more data from around the world. Huge data for this application can make a tourist more satisfied.
Application server, database server, and content management application is stored on the cloud. The purpose is to make more flexible in storage and computing process, this is the most effective and efficient way for the AMGT cases. Data mining on AMGT can make Business success probability on tourist site can be calculated.

Keywords—data mining, crawl data, augmented reality, tourist guide, mobile guide.

I. INTRODUCTION

Mobile applications have an amazingly growth in this last century, includes innovative operating system, hardware and software in line with the progress of the internet. Tourism area have prospect of economy value to be gain, combining with the mobile applications. Indonesia as a developing country which has thousands of islands, have business opportunity when analyzed over the country will increase foreign exchange. Viewed from the eyes of technology, the behavior of foreign tourists is the community to the needs of high Information Technology (IT) [1].

IT and tourism is a new innovative combination and it still rare in the world. The tourist can get the information of the object they want by only get the picture of the object by using their camera in mobile phone, and then the tourist will get the detail information as soon as possible (depends on the data speed of their cellular operator).

Actual Mobile Guide for Tourist (AMGT) on the firstly design is for satisfied all tourist in Indonesia. This application has few data about the interesting spot in Indonesia. This application is a markerless augmented reality-based and running on the Android platform. By using SIFT algorithm on the server to process the image that captured by client (the tourist).

II. ARCHITECTURE OF AMGT

Figure 2.1 show the Architecture of the Actual Mobile guide for Tourist. There are about sub-system of the AMGT like client application, server application, database and content management server.
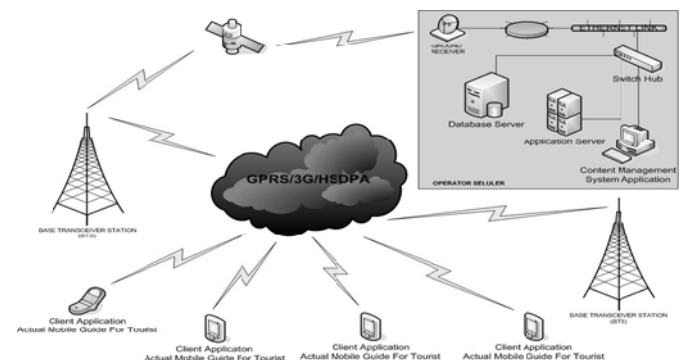
Figure 2.1 – Architecture of the Actual Mobile Guide for Tourist

Based on Figure 2.1, there are four sub-systems that integrated in this application, including:
1. Mobile Phone (Client Application).
   This client will be on the tourist side. The data will be captured using this device, and then the data sent to the server via data cellular of the mobile phone.
2. Database Server.
   The data will be collected on the server. The picture that be sent from the client will be stored in server using BLOB (binary large object).
3. Application Server.
   The data processing will be held in the server side. This application consists of Artificial Neural Network for the matching process. And then will be combining with the SIFT algorithm.
4. Content Management System Application.
   User interface of the application server will be.

Application server, database server, and content management application is stored on the cloud. The purpose is

to make more flexible in storage and computing process, this is the most effective and efficient way for the AMGT cases.

III. CLIENT APPLICATION DESIGN SPECIFICATION

Application on client (mobile phone) did not change much (Figure 3.1). Otherwise on server there is more change since it is uploaded into the cloud.

There is some data that transferred from mobile client application into the server, there are:
1. LBS Parameter.
2. Image Data.
3. Key point Descriptor.
4. Network and Client parameter (such as Cell ID phone, phone number).
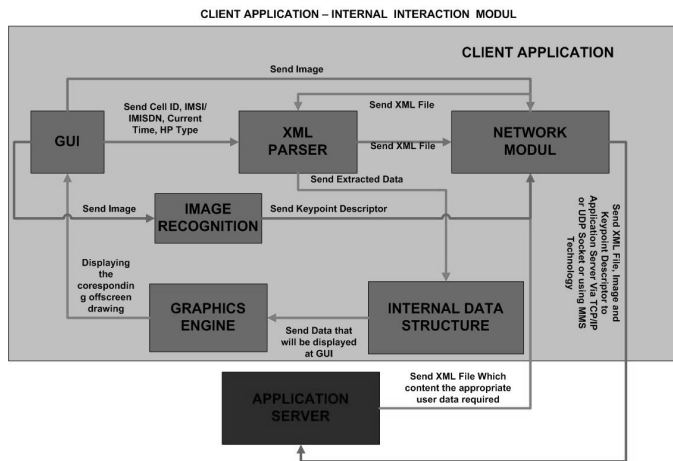
Figure 3.1 – Block Diagram for the Client Application

Figure 3.1 shows the block diagram of client application of the AMGT.

IV. SERVER APPLICATION DESIGN SPECIFICATION

This is data flow and subsystem component on server application (Figure 4.1).
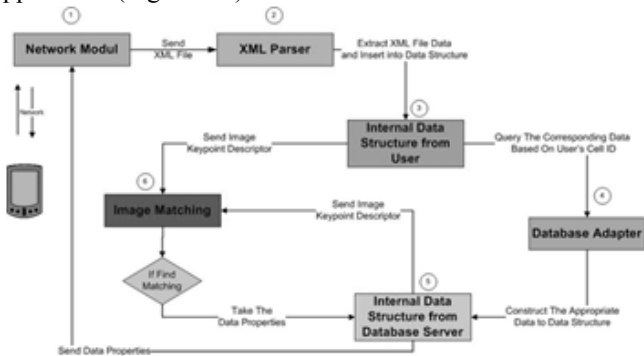
Figure 4.1 – Block Diagram for the Server Application

Server receive all data and do a matching keydescriptor in application server. If server find matching data, server will send data Properties(information) to mobile client. Image matching module in this system spend much computing power.

Application server need to be evolve, since the database have little data. Application server need more image database

and information based on image downloaded. If server download a pisa tower, server must download an information about that too. Adding one different module for filling database with information is required.

Server will send information data to mobile client such as an information about image captured by user, hotel, restaurant and gas station around the image captured.

This is AMGT prototype.
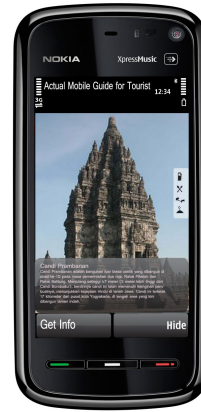
Figure 4.2 AMGT Prototype

V. DATA GATHERING USES CLOUD

Data mining is based on keyword of object tourist such as place of beautiful scenery like beach, mountain, 7 wonder, or unique place.
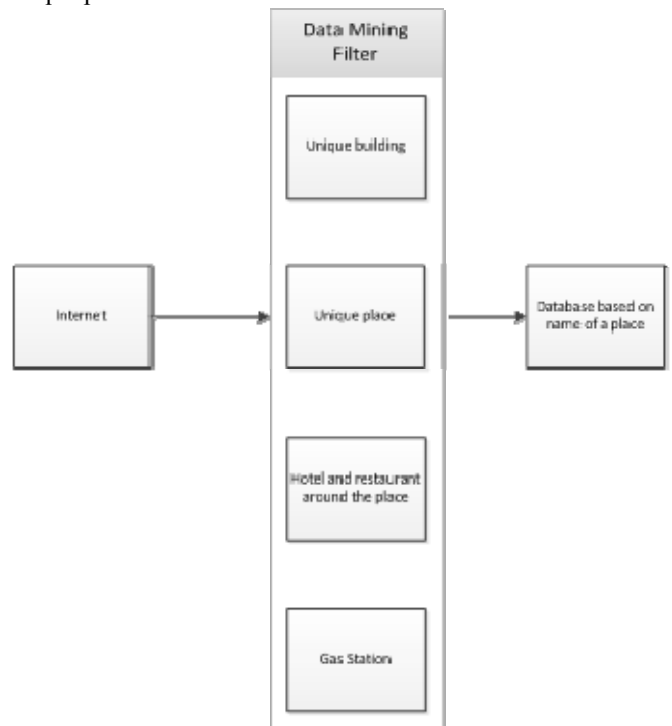
Figure 5.1 – Data Mining Filtering Process

Database server have to update their data since building or places might be change cause by natural disaster or renovation. Data gathering working 24/7 to get more data. Data mining or web mining in this case need to be gain correct data.

There is a data specification for web mining to be searched.

1. Image
2. Name of image(in AMGT case is building)
3. Information about the building
4. Information about gas station, unique building, unique place, hospital, ATM, and restaurant.

Knowledge Discovery in database(KDD) is appropriate for this problem. The traditional method of turning data into knowledge relies on manual analysis and interpretation. For example, in the health-care industry, it is common for specialists to periodically analyze current trends and changes in health-care data, say, on a quarterly basis. The specialists then provide a report detailing the analysis to the sponsoring health-care organization; this report becomes the basis for future decision making and planning for health-care management [2].
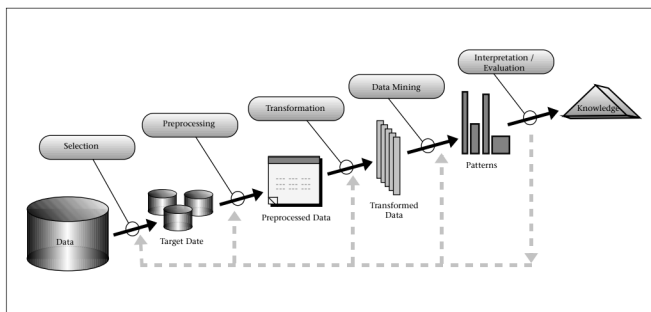
Figure 5.2 – Steps of Composing the KDD

Figure 5.2 show an overview of the steps that compose the KDD process.
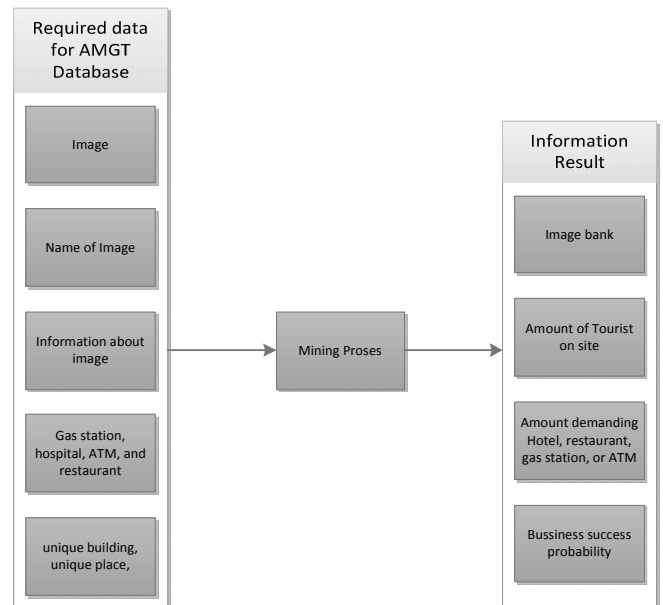
Figure 5.3 Data mining proses result.

Data gathering will show information result that will valuable for government or some people to make a Business(as describe in Fig 5.3). Business man more convinced to deal and to build restaurant, hotel or Business related.

Government will know amount of tourist come to the site. Government will also know when to make renovation of the site to make tourist more satisfied. Government will know when to turn to another strategy to make the site be discovered by the tourist.

Image bank and the information in huge database can be build a wiki for tourist. Website will be build to make this information growth and making more value to the application. Website will contain travel guide to the site, how do the tourist get there, and how much does it spend your money, and several testimonial of site, hotel, and restaurant who has he came.

Each of mining source will create a lot of data that would make our database increase rapidly. Cloud is appropriate solution for AMGT case.

VI. Conclusion

Cloud is important in this case, because it will unite all data into one place. Flexibility in the cloud make this application evolving. Computation and storage are really important since data will be burst to database and application server.

Data mining provides support in other areas. Business success probability can be calculated as data increase to database. Government can take data from application and get more information like number of tourist every month and year or view the site. More data can convince data validity.

References

[1] I W Aciek, S P Ary, "Gathering Information and Anywhere (GIRA)", International Conference on Electrical Engineering and Informatics (2011).

[2] U Fayyad, G P-Shapiro, and P Smyth, "From Data Mining toKnowledge Discovery in Databases," AI Magazine Volume 17 Number 3 (1996).

[3] Jonathan Mooser, Lu Wang, Suya You and Ulrich Neumann, "An Augmented Reality Interface for Mobile Information".

[4] Thomas Bakken, "An evaluation of the SIFT algorithm for CBIR," in Telenor R&I Research Note 30/2007.

[5] Braitner, Michael.," Market Development of Mobile Device Classes and Operating Systems", Seminar On Mobile Business, Institute for Information Systems Research LV 71407 Winter term 2006/2007.

[6] Lowe, David G.,"Object Recognition Using SIFT", Proc. of the International Conference on Computer Vision, Corfu (Sept. 1999).

FPGA Implementation of High Speed Pipelined JPEG 2000 Encoder

Vijay Kumar Sehrawat[1], Amit Gupta[2], Mamta Khosla[3]

Department of Electronics and Communication Engineering

Dr. B.R. Ambedkar National Institute of Technology

Jalandhar, Punjab, India

[1]vijaykumarsehrawat@gmail.com, [2]amit.gupta326@gmail.com, [3]khoslam@nitj.ac.in

Abstract---.This paper presents the design and implementation of a high speed JPEG 2000 encoder that employs the Distributive Arithmetic (DA) technique for the complex computation of Discrete Wavelet Transform (DWT). DWT is the most popular transformation technique adopted for image compression in the previous works, which, however, is computationally complex. In this paper, DA-DWT technique is employed for reducing the complex computations, thereby increasing the speed and throughput. The co-processor also exploits the principles of pipelining and parallelism to the maximum extent in order to obtain high speed and high throughput. The design is modeled using Verilog HDL and implemented on Virtex2 Pro FPGA that consumes 7% resources of FPGA and shows the clock latency of 34 cycles at 250 MHz clock frequency. The proposed encoder can encode color as well as gray image formats. The performance parameters PSNR, MSE and SNR for some standard test images have been tabulated.

Keywords- Image Compression, Distributive Arithmetic, Pipelined JPEG, DWT.

I. INTRODUCTION

With the rapid advancements in image processing domain, the image processing systems need to store High Definition (HD) images into storage devices.Despite the rapid progress in the technology of data storage devices, the demand of data storage capacity and bandwidth requirement of HD image processing systems outstripped the capabilities of available mass storage technologies. So, there is an urgent need to develop such systems that can be effectively used for compressing the HD images without losing the prime information. This work presents a very high speed, high throughput, low power dedicated hardware co-processor to compress the originalimages that exploits the benefits of distributive arithmetic (DA) and pipelined architecture for JPEG 2000[14] encoder.

JPEG 2000 standards [13] were evolved from the efforts of Joint Photographic Expert Group for compression of still images. It is based on DWT that decomposes the image into time and frequency domain and hence requires less data bits to store the same image. Due to this type of decomposition JPEG 2000 is better than JPEG image. DWT has inherent benefits of high compression ratio, improved *PSNR* (Peak Signal to Noise Ratio) and is free from blocking artifacts at low bit rates [5]. There are complexities in DWT due to large number of arithmetic operations that affect the speed and performance of the JPEG encoder. To overcome these shortcomings DA technique is used to decrease this processing time of DWT. To further compress the size of image, DWT bits are quantized and encoded using the Huffman encoding. The FPGA implementation of the proposed encoder further accelerates its performance.

The paper is organized in the following sections: In section II previous related works for image compression are discussed. Section III presents the proposeddesign of high speed JPEG 2000encoder; DA-DWT technique and pipelined architecture are explained in this section. The performance parameters of the system are discussed in section IV. In Section V, FPGA implementation is covered and finally conclusions have been drawn in section VI.

II. RELATED WORKS

This section gives a brief review of the work reported by various researchers in this field.

In [1]-[3] high speed image compression system that are based on the Discrete Cosine Transform (DCT) based JPEG codec are discussed. The systems proposed by the authors also exploited the benefits of pipelining to reduce the clock latency. But the use of DCT in image compression produces the blocking artifacts at high bit rates and degrades the quality of image.

In [5] the use of DA in the computation of DWT is proposed that is suitable for FPGA implementation. The compression of the system can be further improved by including the entropy encoding [3].

The FPGA implementation of 2D-DCT with zig-zag and variable length coding is implemented in Zhang's paper[3]. The system can be modified by the use of 2D-DWT and Huffman encoding for better performance.

The image encoder proposed in this paper addresses the shortcomings of the earlier reported techniques. The proposed image encoder is implemented on FPGA and exploits the use of both DA and pipelining for faster results, and Huffman encoder for high compression ratio.

III. PROPOSED DESIGN OF HIGH SPEED PIPELINED JPEG 2000 ENCODER

The architecture of the proposed JPEG 2000 image encoder is shown in Fig. 1. The JPEG 2000 compression involves three main blocks: DA-DWT, Quantization and Huffman encoding.
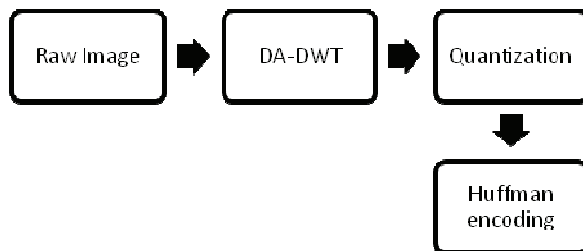
Figure. 1 Proposed DA-DWT based JPEG 2000 image encoder.

A. Distributive Arithmetic –Discrete Wavelet Transform (DA-DWT)

An image consists of pixels that are arranged in two dimensionalmatrix. In spatial domain adjacent pixels are very highly correlated and hence redundant. This redundant information can be removed without affecting the quality of the image by exploiting the special characteristic of human vision. Human eye is very much sensitive to low frequencies, so the image is decomposed into frequency sub-bands of which lower sub-bands are selected and transmitted and higher sub-bands are rejected with the usage of DWT. The basic architecture of a 2 level DWT is shown in Fig. 2.The raw image consisting of rows and columns is first passed through high pass and low pass filters.The coefficients for these filters are computed using Matlab. Multiple levels of decomposition can be done to compress the image. Each successive

decomposition divides the image into low frequency and high frequency components and the low frequency components are passed to next decomposition level.

H - High Pass Frequency Components
L - Low Pass Frequency Components
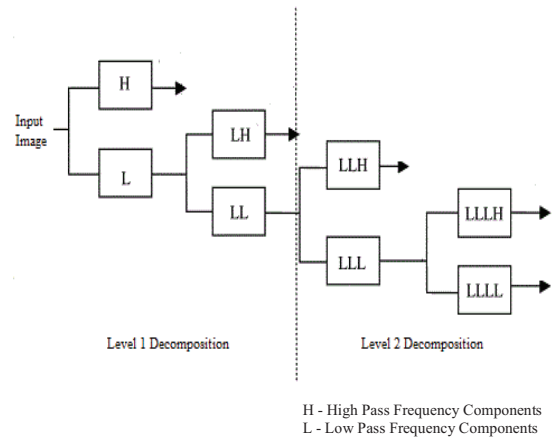
Figure 2. Architecture of 2 level DWT.

However, in this work, speed of the system is accelerated by employing DA memory architecture as shown in Fig. 3. The DA memory architecture replaces the multipliers with the look up tables in a very efficient way so that the DWT block attains high throughput, high speed, and lessarea.

To further speed up this process, parallelism is realized on this technique in which the input samples are divided in even and odd samples. The memory size is approximately reduced to half due to the symmetric property of filter coefficients. This allresults in increased throughput because input samples are used in parallel to read data from LUT. The process is shown in Fig. 4.[12].

B. Quantization

After the wavelet transform, the coefficients are scalar-quantized to reduce the number of bits to represent them, at the expense of quality. The output is a set of integer numbers which have to be encoded bit-by-bit. The parameter that can be changed to set the final quality is the quantization step: the greater the step, the greater is the compression and the loss of quality. With a quantization step that equals 1, no quantization is performed.The step size used in the quantization of DWT coefficients is 2. The quantization process is mathematically described in (1).

$$q(n) = round(y) \left\lfloor \frac{|y|}{\Delta_b} \right\rfloor \qquad (1)$$

Where, $q(n)$ is the quantized DWT coefficient, $\Delta_b$is the step size for quantization.

In hardware implementation, the quantization is performed by using separate registers for holding

integer and fractional part of the DWT coefficient, the integer part is retained and the fractional part is discarded.
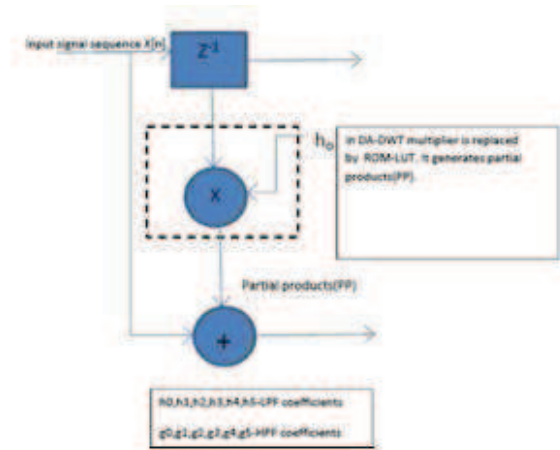
Figure 3.Scheme of DA-DWT used in proposed image encoder.
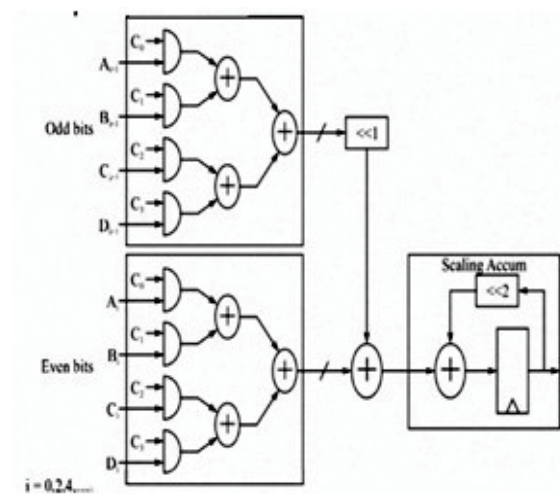
Figure 4.Parallel architecture for faster DA-DWT implementation [2]

C. Huffman encoder

Huffman coding is an entropy encoding algorithm used for lossless data compression. The term refers to the use of a variable-length code table for encoding a source symbol where the variable-length code table has been derived in a particular way based on the estimated probability of occurrence for each possible value of the source symbol. In the proposed system three different Huffman tables are derived for different components of color image. It helps in

further compressing the image for better compression ratio.

D. Pipelining

In this proposed work, pipelined architecture is employed to improve the performance of JPEG 2000 encoder. The pipelining of different modules makes the design faster with little increase in area. In the process of pipelining, consecutive computations are performed in parallel holding them in buffer close to the processor until each instruction operation can be performed [3]. Also the pipelining gives increased throughput in a given period of time.

IV. Performance Parameters

The performance of the proposed encoder is estimated by the following performance parameters:

A. Mean Square Error (MSE)

MSE is the cumulative squared error between the compressed and the original image

$$MSE = \frac{1}{M \times N} \sum_{y=1}^{M} \sum_{x=1}^{N} [I(x,y) - I'(x,y)]^2 \qquad (2)$$

Where $I(x,y)$ is the original image and $I'(x,y)$ is the approximated image. x, y represent the dimensions of the image. A lower value of MSE means lower error.

B. Peak Signal to Noise ratio (PSNR)

$PSNR$ is derived from MSE and it is the measure of peak error in the compressed image.

$$PSNR = 20 \times log_{10}(255/\sqrt{MSE}) \qquad (3)$$

A higher value of $PSNR$ is good because it means ratio of signal to noise is higher. Here the signal is the original image and noise is the error in reconstruction.

C. SNR (Signal to Noise ratio)

SNR is usually taken to indicate an average signal to noise ratio, as it is possible that instantaneous signal to noise ratios will be considerably different in an image.

V MATLAB AND FPGA IMPLEMENTATION RESULTS

The scheme for image compression proposed here is tested for different images using Matlab and corresponding results are shown in Fig. 4. compression of different images with different

resolution is shown in Table.1. The performance of the system is estimated by the parameters given in Table. 2.

In applications that require real-time compression, a dedicated hardware is required. The proposed system in this work has been modeled in Verilog HDL and simulated using Xilinx ISE and targeted for Virtex2 Pro. The results of the FPGA implementation are given in Table. 3. The results obtained after FPGA implementation demonstrate that system has clock latency of 34 clock cycles at 100 MHz clock frequency, that makes the system suitable for real-time applications.

Original image   Compressed image

Figure4. Original and compressed images (a) Lena, (b) Mig, (c) India

Table1   Size of different images after compression

Image	Resolution	Original image size	Compressed image size
Lena	1024*1024	2Mb	40Kb
Mig	256*256	512Kb	15Kb
India	512*512	1.2Mb	30Kb

Table2   Comparison of factors used in image compression

Image	PSNR	MSE	SNR (dB)
Lena	34.8064	21.5	15.1395
Mig	34.9229	20.9309	15.6019
India	33.4436	29.4255	14.1748

Table 3   Parameters for hardware implantation used

FPGA	Virtex2 Pro
Area	7 %
Clock Latency	34
Maximum Resolution	1024*1024
Clock	250 MHz
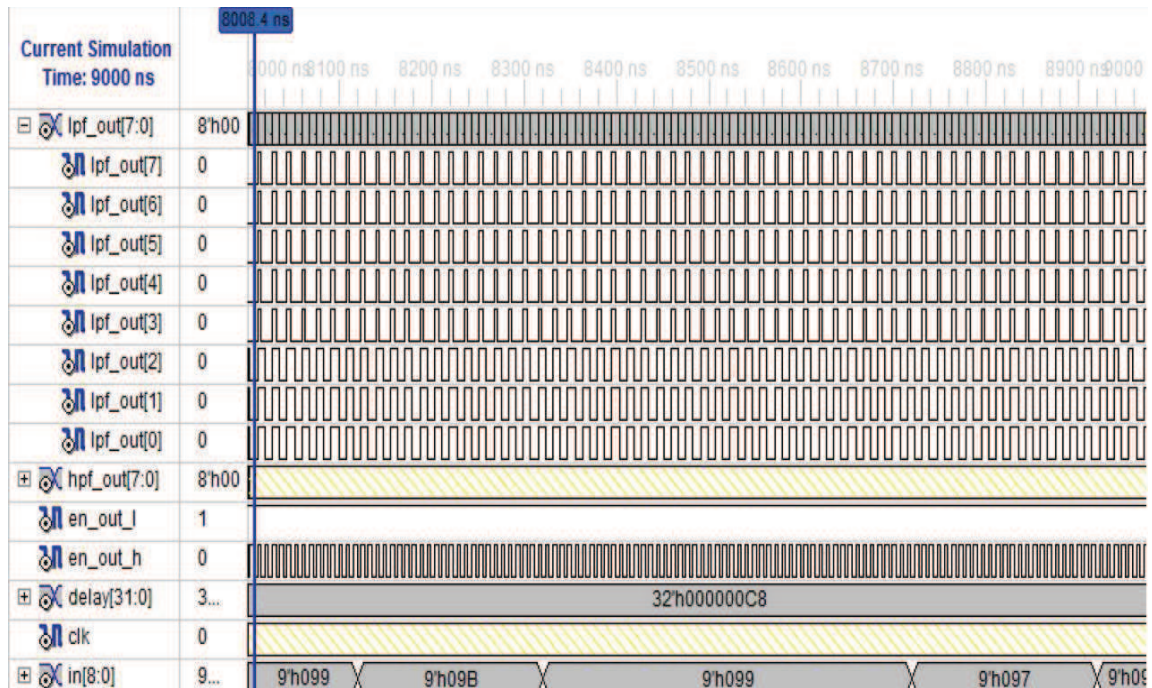
Figure5. Simulation of pipelined JPEG2000 encoder

VI CONCLUSIONS

The coprocessor design presented in this paper gives faster results and can be used in real time applications like digital cameras. *PSNR, MSE, SNR* are compared with both colored and gray images of different resolution. As compared to the previous counterparts, processing time of the proposed system is reasonably less. Decoder part also lies in the same flow. Presently the authors are working on the design of DA-IDWT (inverse DWT) based decoder with an objective to complete the design of high speed JPEG 2000 codec.

REFRENCES

[1] N. Tiwari, S.C Reddy, "Performance measurement of a fully pipelined JPEG codec on emulation platform," *IEEE 2nd International Advance Computing Conference* (IACC), pp.167-171, 19-20 Feb. 2010.

[2] Zheng Wei; Hou Zhisong, "Analysis of JPEG encoder for image compression," *International Conference on Multimedia Technology (ICMT- 2011),* pp.205-208, 26-28 July 2011

[3] Zhang Qihui; Chen Jianghua; Zhang Shaohui; Meng Nan, "A VLSI implementation of pipelined JPEG encoder for grayscale images," *International Symposium on Signals, Circuits and Systems, ISSCS 2009*, pp.1-4, 9-10 July 2009.

[4] S.Sanjeevannanavar, A.N.; Nagamani, "Efficient design and FPGA implementation of JPEG encoder using verilog HDL," *International Conference on Nanoscience, Engineering and Technology (ICONSET)*, pp.584-588, 28-30 Nov. 2011.

[5] M.Nagabushanam; P.C.P. Raj; S.Ramachandran , "Design and FPGA implementation of modified Distributive Arithmetic based DWT-IDWT processor for image

compression," *International Conference on Communications and Signal Processing (ICCSP), 2011*, vol., no., pp.1-4, 10-12 Feb. 2011

[6] Lynn Abbott, "Pipelined Architectures for Morphological Image Analysis" Machine Vision and Applications, *Springer Verlag* New York 1988.

[7] Jain, A.K., *Digital Image Processing*, Prentice Hall of India 2004.

[8] M. Vishwanath, R.M. Owens and MJ. Irwin, "VLSI Architecture for the Discrete Wavelet Transform", *IEEE Trans*. Circuits and Systems, vol. 42, pp. 305-316, May 1996.

[9] David S. Tabman and Michael W. Marcelliun, "JPEG 2000 – ImageCompression, Fundamentals, Standards and Practice", KulwerAcademic Publishers, Second printing 2002.

[10] Cyril Prsanna Raj and Cittibabu, "Pipelined DCT for image compression*", SASTech Journal*, Vol. 7, pp. 34-38, 2007

[11] M.Kovac; N.Ranganathan; M. Zagar; "A prototype VLSI chip architecture for JPEG image compression," *European Design and Test Conference,. ED&TC 1995*, Proceedings, pp.2-6, 6-9 Mar 1995.

[12] Qihui Zhang; Nan Meng; , "A low area pipelined 2-D DCT architecture for JPEG encoder,". *MWSCAS '09*. 52nd IEEE International Midwest Symposium on Circuits and Systems, pp.747-750, 2-5 Aug. 2009.

[13] M. Nagabushanam, Cyril PrasannaRaj , S. Ramachandran, "Design and Implementation of parallel and pipelined Distributive Arithmetic Based Discrete Wavelet Transform IP Core", EJSR, Vol. 35, No. 3, pp. 378-392, 2009.

[14] *The International telegraph and telephone consultative committee (CCIT) Information Technology* – Digital Compression and coding of continuous-Tone Still images – Requirements and GuidelinesRec.T.81,1992

Modified Nevanlinna – Pick Interpolation Theory for Control System Design

Aruna B

Department of Electrical and Electronic Engineering
Hindustan Institute of Technology and Science
Chennai, India
arunarajan@hindustanuniv.ac.in

Devanathan R

Department of Electrical and Electronic Engineering
Hindustan Institute of Technology and Science
Chennai, India
deanes@hindustanuniv.ac.in

Abstract –**In this paper we present a new modification of the classical Nevanlinna-Pick (N-P) scalar interpolation problem and provide a solution to the modified N-P problem. The modified N-P result can be used to place the closed loop poles in a desired region in the left half of the complex plane in control system design. Numerical examples illustrate the theory.**

Keywords—Interpolation, Robustness, Sensitivity, Closed loop design.

I. INTRODUCTION

The Nevanlinna-Pick (N-P) interpolation describes a class of complex interpolation problems. The problem was originally studied by Pick in 1916[1] and independently by Nevanlinna in 1919[2]. The classical N-P interpolation problem is to find (if one exists) an analytic function $f: D \rightarrow \bar{D}$ such that $f(a_i) = b_i$, $a_i \in D$, $b_i \in \bar{D}$, i = 1,2,...,q, where D and $\bar{D}$ correspond to open and closed unit disks respectively. An interpolating function f exists if and only if the N-P matrix, $N = \left[\frac{1-b_i\bar{b}_j}{1-a_i\bar{a}_j}\right]_{i,j=1,2,...,q}$ is positive semidefinite. The N-P problem has found numerous applications in model approximation, robust stabilization, the model matching problem in $H\infty$ control, circuit theory etc.

Khargonekar and Tannenbaum [3] in their work, describe a modification of the classical N-P interpolation problem. i.e. they are interested in finding an analytic function $f: D \rightarrow \bar{D}$ such that $f(a_i) = \alpha b_i$, $a_i \in D$, $b_i \in \mathbb{C}$, i = 1,2,...,q and $\alpha \geq 0$ *is in* $\mathbb{R}$. The function $f(a_i)$ exists only for $\alpha \leq \alpha_{max}$, where $\alpha_{max} = \frac{1}{\sqrt{\lambda_{max}}}$ and λ_{max} is the largest eigenvalue of $A^{-1}B$ where

$$A = \left[\frac{1}{1 - a_i\bar{a}_j}\right]_{i,j=1,2,...,q}, B = \left[\frac{b_i\bar{b}_j}{1 - a_i\bar{a}_j}\right]_{i,j=1,2,...,q}$$

The modified theory is applied to robust stabilization problem.

Kimura's paper [4] deals with the problem of robust stabilization for single-input single-output plants with a prescribed uncertainty band in the frequency domain. Classical N-P interpolation theory is used to derive a condition for robust stabilizability. Here the problem of robust stabilization is reduced to an interpolation problem of bounded real functions called Schur-class function (i.e. a holomorphic function mapping the unit disk into itself). The Nevanlinna-Pick problem is solvable for Strongly Bounded Real (SBR) functions, if and only if the matrix P is positive definite.

Doyle, Francis, and Tannenbaum [5] used the N-P interpolation theory to solve the model-matching problem. The NP problem is to find a function $G \in \mathcal{S}$ satisfying the two conditions. $\|G\|_\infty \leq 1$, G (a_i) = b_i, i = 1... n. Here $\mathcal{S}$ stands for the space of stable, proper, a complex-rational function. The latter equation says that G is to interpolate the value b_i at the point a_i and must be stable, proper, and satisfy $\|G\|_\infty \leq 1$. The N-P problem is said to be solvable if such a function G exists.

J.A. Ball, et al [6] in their paper allowed the interpolant to be matrix-valued and elaborate on the analogy with the theory of Nevanlinna-Pick interpolation on a finitely-connected planar domain.

Dorato P, Yunzhi L[7] presented a modification of the classical Nevanlinna-Pick interpolation algorithm for bounded-analytic functions which permits interpolation with bounded-real functions. This is an essential modification when the Nevanlinna-Pick theory is applied to problems such as robust stabilization, since otherwise compensator transfer functions will in general have complex coefficients when complex interpolation points are present. The results are applied to a numerical problem in robust stabilization.

I. Byrnes, et al [8] presented a generalized entropy criterion for solving the rational Nevanlinna–Pick problem for n+1 interpolating conditions and the degree of interpolants bounded by n. The primal problem of maximizing this entropy gain has a very well-behaved dual problem. This dual is a convex optimization problem in a finite-dimensional space and gives rise to an algorithm for finding all interpolants which are positive real and rational of degree at most n. The criterion requires a selection of a monic Schur polynomial of degree n. It follows that this class of monic polynomials completely parameterizes all such rational interpolants, and it

therefore provides a set of design parameters for specifying such interpolants. The algorithm is implemented in state-space form and applied to several illustrative problems in systems and control, namely, sensitivity minimization, maximal power transfer and spectral estimation. However, in all these applications, it is important that the interpolating functions be rational with a degree which does not exceed some prescribed bound.

I. Pendharkar, et al [9] address the N-P problem using concepts from behavioural systems theory and quadratic differential form (QDFs). The NP problem is solved using a certain "dualization of data". They address system theoretic motivations for this dualization and the advantages gained in this process. Finally the problem of constructing interpolating functions that satisfy a "frequency dependent" norm condition is explained.

Hari Bercovici, et al[10] carried out their work on interpolation problems where the interpolating functions are bounded not in norm but in spectral radius or structured singular values. These problems arise naturally in the robust control of systems with structured uncertainty. Here they extend their structured matrix Nevanlinna-Pick interpolation theory to the (one-sided) tangential case. With these results, they have laid the ground work for a rigorous analytic procedure for μ-synthesis.

Closed loop design using N-P theory, such as, Khargonekar and Tannenbaum [3] seeks to place the closed loop poles in the left half of the complex plane. But, usually, there is no attempt to control the region in which the closed loop poles can be placed. Our new result on the modified N-P theory provides a systematic methodology to find the smallest circle in the left half of the complex plane within which the closed loop poles can be placed such that a solution to the design problem exists. Also, the method allows us to place the closed loop poles inside any given circle beyond the minimal circle. Our modified N-P theory caters to open loop systems with stable poles and zeros while that proposed in [3] assumes that the open loop poles and zeros lie in the right half of the complex plane. The methodology, as given in [3], involves finding the sensitivity function first followed by the controller. Also, the method allows the robust control problem (such as uncertainty in gain) to be considered in a fundamental way at the design stage itself. In contrast, the other robust control methods address the problem typically by inserting a compensator into the system while control system design itself cannot take the robustness into account.

To summarise the rest of the paper, Section II describes the approach taken and states the modified N-P scalar interpolation problem. Section III provides a solution to the modified N-P problem. Section IV gives numerical examples to illustrate the theory. Section V concludes the paper.

II. STATEMENT OF THE PROPOSED APPROACH

The classical N-P interpolation theory has been successfully applied to robust stabilization problems in control [3]. Typically the closed right half of the complex plane is mapped by a stable sensitivity function to a connected region in the complex plane which excludes regions corresponding to parameter uncertainty. The sensitivity function interpolates zeros and poles of open loop plant transfer function in the closed right half plane to points 1 and 0 respectively. Once the sensitivity function is determined, the controller can be derived easily from it[11]. The above design will ensure that the closed loop system will not have any closed loop characteristic equation zeros in the closed right half of the complex plane even under the effect of parameter uncertainty.

Most of the engineering systems, for example electrical motors, thermal systems etc. are open loop stable. In order to apply the above N-P theory to these systems, we can proceed as follows. First choose a circle well inside the left half of the complex plane such that the plant stable poles and zeros lie outside the circle. The complement of the circle area in the complex plane can be redefined as the closed right half plane with the circumference of the circle being the redefined imaginary axis (through conformal mapping). One can then proceed to apply N-P theory of [3] to the resulting system.

While the circle inside the left half of the complex plane was arbitrarily chosen except that the stable plant poles and zeros lie outside the circle, one can ask what is the smallest circle one can choose (corresponding to tightest control) before the sensitivity function (and correspondingly the controller function), for a given set of open loop poles and zeros of the plant, ceases to exist. The problem of existence of an interpolating function mapping the region outside the minimal circle in the left half of the complex plane to a connected region in the complex plane containing 1 and 0 can be transformed through conformal mapping into the existence of an interpolating function mapping closed unit disk into the open unit disk as given below. The is the modified N-P problem addressed in this paper.

Given that a function $f: \bar{D} \to D$ exists such that $f(a_i) = b_i$, i $= 1, 2, \dots, q$, $a_i \in \bar{D}$, $b_i \in D$, what is the minimum value of k, $0 < k < 1$, such that $f'(ka_i) = b_i$, i $= 1, 2, \dots, q$ exists.

III. SOLUTION TO THE MODIFIED NAVANLINNA-PICK PROBLEM

Theorem 1: Let $a_i \in \bar{D}$, $b_i \in D$, i $= 1,2, \dots, q$. Let a function f exist such that $f(a_i) = b_i$ such that the Pick matrix

$$P = \left[\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j} \right]_{i,j=1,2,\dots,q} > 0 \qquad (1)$$

Let $0 < k < 1$ be in $\mathbb{R}$. Consider the function $f': \bar{D} \to D$ such

that $f'(ka_i) = b_i$, i $= 1, 2, ..., q$. Then with $k \cong 1$ there exists a minimal k, k_{min} given by

$$k_{min} = \sqrt{\frac{\lambda_{max} - 1}{\lambda_{max}}} \qquad (2)$$

where λ_{max} is the largest eigen value of $A^{-1}B$, where

$$A = \left[\left(\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j} \right) \right]_{i,j=1,2,...q}$$

and

$$B = \left[\left(\frac{a_i \bar{a}_j (1 - b_i \bar{b}_j)}{(1 - a_i \bar{a}_j)^2} \right) \right]_{i,j=1,2,...q}$$

such that the function $f'(ka_i)$ exists for $k \le k_{min}$.

Proof: For the function f' such that $f'(ka_i) = b_i$ to exist, it is required that the Pick matrix

$$P' = \left[\frac{1 - b_i \bar{b}_j}{1 - k^2 a_i \bar{a}_j} \right]_{i,j=1,2,...q} \ge 0 \qquad (3)$$

ie P' should be positive semi definite.

Now P' can be expressed as

$$P' = \left[\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j + a_i \bar{a}_j - k^2 a_i \bar{a}_j} \right]_{i,j=1,2,...q}$$

$$= \left[\frac{1 - b_i \bar{b}_j}{(1 - a_i \bar{a}_j) + a_i \bar{a}_j (1 - k^2)} \right]_{i,j=1,2,...q}$$

$$P' = \left[\frac{1 - b_i \bar{b}_j}{(1 - a_i \bar{a}_j) \left(1 + \frac{a_i \bar{a}_j (1 - k^2)}{(1 - a_i \bar{a}_j)} \right)} \right]_{i,j=1,2,...q} \qquad (4)$$

Let

$$\epsilon = \frac{a_i \bar{a}_j (1 - k^2)}{(1 - a_i \bar{a}_j)}$$

Then

$$P' = \left[\left(\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j} \right) \left(\frac{1}{1 + \epsilon} \right) \right]_{i,j=1,2,...q} \ge 0 \qquad (5)$$

Since with

$$k \cong 1, \quad |\epsilon| \ll 1, \quad then \quad \frac{1}{1 + \epsilon} \cong 1 - \epsilon$$

Then

$$P' \cong P'' = \left[\left(\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j} \right) (1 - \epsilon) \right]_{i,j=1,2,...q}$$

$$P'' = \left[\left(\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j} \right) - \epsilon \left(\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j} \right) \right]_{i,j=1,2,...q} \qquad (6)$$

. Substituting for ϵ, we get

$$P'' = \left[\left(\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j} \right) - \left(\frac{a_i \bar{a}_j (1 - k^2)(1 - b_i \bar{b}_j)}{(1 - a_i \bar{a}_j)^2} \right) \right]_{i,j=1,2,...q}$$

$$P'' = A - (1 - k^2)B \qquad (7)$$

where

$$A = \left[\left(\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j} \right) \right]_{i,j=1,2,...q}$$

and

$$B = \left[\left(\frac{a_i \bar{a}_j (1 - b_i \bar{b}_j)}{(1 - a_i \bar{a}_j)^2} \right) \right]_{i,j=1,2,...q}$$

It is required to find k_{min} for $P'' \ge 0$.

Since $A = P > 0$, A^{-1} exists by assumption.

Also, since $0 < k < 1$, $0 < (1 - k^2) < 1$

Let

$$\lambda = \frac{1}{1 - k^2}$$

Then $P'' = \lambda A - B \ge 0$ implies

$$x^*(\lambda A - B)x \ge 0 \qquad \forall \ x \qquad (8)$$

Then there exists a transformation $x = Z\xi$ [12] that reduces simultaneously

$$x^*Ax \quad to \quad \sum_k |\xi_k|^2 \qquad and$$

$x^*Bx \quad to \quad \sum_k \lambda_k |\xi_k|^2$, where eigen values $\{\lambda_k\}_{k=1}^q$ are all real.

Thus

$$\lambda \sum_k |\xi_k|^2 - \sum_k \lambda_k |\xi_k|^2 \ge 0 \quad \forall \ \xi_k \qquad (9)$$

$$\lambda - \lambda_k \ge 0 \quad \forall \ k$$

$$\lambda \ge \lambda_{k \ max}$$

$$\frac{1}{1-k^2} \geq \lambda_{k\,max}$$

$$1-k^2 \leq \frac{1}{\lambda_{k\,max}} \tag{10}$$

where $\lambda_{k\,max}$ is the largest eigenvalue of $\{\lambda_k\}_{k=1}^q$.
The above condition(10) can be written for $\lambda_{max} > 1$ as

$$k^2 \geq \frac{\lambda_{max}-1}{\lambda_{max}} \tag{11}$$

$$k \geq \sqrt{\frac{\lambda_{max}-1}{\lambda_{max}}} = k_{min}, \quad \lambda_{max} > 1 \tag{12}$$

For $\lambda_{max} \leq 1$, condition(10) is always satisfied, since $0 < (1-k^2) < 1$.

Now it is required to check whether $P' \geq P'' \geq 0$.

For this it is sufficient to prove that $P' - P'' \geq 0$

$$P' - P'' = \left[\left(\frac{1-b_i\bar{b}_j}{1-a_i\bar{a}_j} \right) \left(\frac{1}{1+\epsilon} - (1-\epsilon) \right) \right]_{i,j=1,2,\ldots q}$$

$$P' - P'' = \left[\left(\frac{1-b_i\bar{b}_j}{1-a_i\bar{a}_j} \right) \left(\frac{\epsilon^2}{1+\epsilon} \right) \right]_{i,j=1,2,\ldots q} \tag{13}$$

Substituting for ϵ, we get

$$P' - P'' = \left[\left(\frac{1-b_i\bar{b}_j}{1-a_i\bar{a}_j} \right) \frac{a_i^2\bar{a}_j^2(1-k^2)^2}{(1-a_i\bar{a}_j)(1-k^2a_i\bar{a}_j)} \right]_{i,j=1,2,\ldots q}$$

i.e. $P' - P''$
$$= \left[\frac{1-b_i\bar{b}_j}{1-a_i\bar{a}_j} \frac{(1-k^2)}{(1-k^2a_i\bar{a}_j)} \frac{a_i^2\bar{a}_j^2(1-k^2)}{1-a_i\bar{a}_j} \right]_{i,j=1,2,\ldots q} \tag{14}$$

$$= P_1 \circ P_2 \circ P_3$$

where the symbol $\circ$ stands for Hadamard product [13]. P_1, P_2 and P_3 are given by

$$P_1 = \left[\frac{1-b_i\bar{b}_j}{1-a_i\bar{a}_j} \right]_{i,j=1,2,\ldots q} \tag{15}$$

$$P_2 = \left[\frac{1-k^2}{1-k^2a_i\bar{a}_j} \right]_{i,j=1,2,\ldots q} \tag{16}$$

and

$$P_3 = \left[\left(\frac{a_i^2\bar{a}_j^2(1-k^2)}{1-a_i\bar{a}_j} \right) \right]_{i,j=1,2,\ldots q} \tag{17}$$

Then $P' - P'' \geq 0$, if each one of these matrices P_1, P_2 and P_3 is positive semi definite as per Schur Product Theorem [13].

$$P_1 = \left[\frac{1-b_i\bar{b}_j}{1-a_i\bar{a}_j} \right]_{i,j=1,2,\ldots q} > 0 \tag{18}$$

By assumption in the theorem statement, $P_1 = P$ corresponds to the Pick matrix of the function $f(a_i) = b_i$, $i = 1,2,\ldots q$ which is assumed to exist.

$$P_2 = \left[\frac{1-k^2}{1-k^2a_i\bar{a}_j} \right]_{i,j=1,2,\ldots q} \geq 0 \tag{19}$$

Since P_2 is the Pick matrix of the constant function $f_2: \bar{D} \to D$ such that $f_2(ka_i) = k$, $(k < 1)$, which obviously exists.

$$P_3 = \left[\left(\frac{a_i^2\bar{a}_j^2(1-k^2)}{1-a_i\bar{a}_j} \right) \right]_{i,j=1,2,\ldots q} \tag{20}$$

can be written as the product of three matrices as follows:

$$P_3 = diag[a_1^2\,a_2^2\,\ldots a_q^2] \left[\frac{1-k^2}{1-a_i\bar{a}_j} \right]_{i,j=1,2,\ldots q} diag[\bar{a}_1^2\,\bar{a}_2^2\ldots\bar{a}_q^2] \tag{21}$$

Since

$$\left[\frac{1-k^2}{1-a_i\bar{a}_j} \right]_{i,j=1,2,\ldots q}$$

is the Pick matrix of the constant function $f_3: \bar{D} \to D$ such that $f_3(a_i) = k$, $i = 1,2\ldots q$, which obviously exists.

Hence

$$\left[\frac{1-k^2}{1-a_i\bar{a}_j} \right]_{i,j=1,2,\ldots q} \geq 0 \tag{22}$$

.

Since for any complex number Z
$$(\bar{Z})^2 = \overline{Z^2} \text{ and } \overline{(\bar{Z})^2} = Z^2$$

P_3 can be written as

$$P_3 = Q^*P_4Q \tag{23}$$

where

$$P_4 = \left[\frac{1-k^2}{1-a_i\bar{a}_j} \right]_{i,j=1,2,\ldots q}$$

and

$$Q = diag[\bar{a}_1{}^2 \ \bar{a}_2{}^2 \ \bar{a}_3{}^2 \ ..... \ \bar{a}_q{}^2]$$

Since Q is non-singular, P_3 given by (23) is congruent to $P_4 \geq 0$ as above. Hence $P_3 \geq 0$

Since $P_1 \geq 0$, $P_2 \geq 0$ and $P_3 \geq 0$, their Hadamard product is given by

$$P' - P'' = P_1 \circ P_2 \circ P_3 \geq 0 \qquad (24)$$

by Schur Product Theorem.

Therefore

$$P' \geq P'' \geq 0 \qquad (25)$$

Hence

$$P' \geq 0 \qquad (26)$$

This is the required result.

IV. NUMERICAL EXAMPLE

A. Interpolating points are all real

Let $a_i \in \bar{D}$,$b_i \in D$, i = 1,2,3 be defined as follows.

a₁ = -1/3 = -0.3333   ; b1 = 0.2

a₂ = 1/3 = 0.3333   ; b2 = 0

a₃ = 3/5 = 0.6000   ; b3 = 0

$$A = \left[\left(\frac{1 - b_i \bar{b}_j}{1 - a_i \bar{a}_j} \right) \right]_{i,j=1,2,...q}$$

$$= \begin{bmatrix} 1.0800 & 0.9000 & 0.9000 \\ 0.9000 & 1.1250 & 1.2500 \\ 0.9000 & 1.2500 & 1.5625 \end{bmatrix} > 0$$

Hence as per Theorem 1, the function $f: \bar{D} \rightarrow D$ such that $f(a_i) = b_i$, i = 1, 2, 3 exists.

$$B = \left[\left(\frac{a_i \bar{a}_j (1 - b_i \bar{b}_j)}{(1 - a_i \bar{a}_j)^2} \right) \right]_{i,j=1,2,...q}$$

$$= \begin{bmatrix} 0.1350 & -0.0900 & -0.1389 \\ -0.0900 & 0.1406 & 0.3125 \\ -0.1389 & 0.3125 & 0.8789 \end{bmatrix}$$

Notice that the eigen values of the pencil $(\lambda A - B)$ are the eigen values of $A^{-1}B$.

$$A^{-1} = \begin{bmatrix} 3.3069 & -4.7619 & 1.9048 \\ -4.7619 & 14.8571 & -9.1429 \\ 1.9048 & -9.1429 & 6.8571 \end{bmatrix}$$

$$A^{-1}B = \begin{bmatrix} 0.6104 & -0.3720 & -0.2733 \\ -0.7102 & -0.3393 & -2.7315 \\ 0.1276 & 0.6857 & 2.9051 \end{bmatrix}$$

Eigen values of $A^{-1}B$ are 0.0115, 1.0203, 2.1445

Hence $\lambda_{max} = 2.1445$

As per (12),

$$k_{min} = \sqrt{\frac{\lambda_{max} - 1}{\lambda_{max}}} = 0.7305$$

Consider the function $f'(ka_i) = b_i$, i = 1,2,3 where $k = k_{min} = 0.7305$. For the function f' to exist, it is required that the Pick matrix

$$P' = \left[\frac{1 - b_i \bar{b}_j}{1 - k^2 a_i \bar{a}_j} \right]_{i,j=1,2,...q}$$

$$= \begin{bmatrix} 1.0205 & 0.9440 & 0.9036 \\ 0.9440 & 1.0630 & 1.1195 \\ 0.9036 & 1.1195 & 1.2378 \end{bmatrix} \geq 0 \text{ , which is true.}$$

Also eigen values of P' are 0.0039, 0.2230 and 3.0945 which are positive. Thus Theorem 1 is verified.

B. Interpolating points include complex conjugate pair

Let $a_i \in \bar{D}$,$b_i \in D$, i = 1,2,3 be defined as follows:

a₁ = 0.6   ; b1 = 0.2

a₂ = 0.4 + j0.7 ; b2 = 0

a₃ = 0.4 - j0.7 ; b3 = 0

$$A = \begin{bmatrix} 1.500 & 1.008 - 0.557j & 1.008 + 0.557j \\ 1.008 + 0.557j & 2.857 & 0.639 + 0.269j \\ 1.008 - 0.557j & 0.639 - 0.269j & 2.857 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.844 & -0.302 - 0.566j & -0.302 + 0.566j \\ -0.302 + 0.566j & 5.306 & -0.303 + 0.075j \\ -0.302 - 0.566j & -0.303 - 0.075j & 5.306 \end{bmatrix}$$

Eigen values of $A^{-1}B$ are 0.3658, 2.4849 and 3.9142

Here $\lambda_{max} = 3.9142$ and $k_{min} = 0.8629$

For the function $f'(ka_i) = b_i$, $i = 1,2,3$ to exist, it requires that the Pick matrix must be positive semi definite.

$$P' = \begin{bmatrix} 1.312 & 1.063 - 0.405j & 1.063 + 0.405j \\ 1.063 + 0.405j & 1.938 & 0.722 + 0.242j \\ 1.063 - 0.405j & 1.063 - 0.405j & 1.938 \end{bmatrix}$$

This is true in this case also as the eigen values of P' are 3.7398, 0.2067 and 1.2405 and are positive. Thus the result of Theorem 1 is again verified.

V. CONCLUSION

In this paper, we have stated and proved a modification of the classical Nevanlinna–Pick (N-P) result on scalar interpolation of a function mapping $\overline{D}$ to D. The modified result can be used by employing conformal mapping in a straight forward way as per[3,12] to design a closed loop control system such that the closed loop transfer function poles are limited to a desired circular region strictly inside the left half of the complex plane thus providing tight control. The result can also be extended to the robust stabilization of open loop stable or unstable systems.

REFERENCES

[1] G. Pick, "Uber die beschr ankungen analytischer funktionen, welche durchvorgegebene funktionswerte bewirkt warden", Math Ann. 77 (1916), pp. 7–23.

[2] R. Nevanlinna,"Uber beschr¨ankte funktionen die in gegebenen punkten vorgeschrieben werte annehmen.", Ann. Acad. Sci. Fenn. ,Sel A 13,1919.

[3] Promod P Khargonekar, and Allen Tannenbaum," Non –Euclidian Metrics and the Robust Stabilization of Systems with Parameter Uncertainty", IEEE Transactions on Automatic Control, vol., No.10, October 1985, pp 1005-1013.

[4] Hidenori Kimura, "Robust Stability for a Class of Transfer Functions" IEEE Transactions on Automatic Control", vol AC 29, No.9, September, 1984, pp 788-793.

[5] John Doyle, Bruce, Francis, Allen Tannenbaum,"Feedback Control Theory ", isbn-10: 0486469336, 2009, www.e-booksdirectory.com.

[6] J.A. Ball, V. Bolotnikov, S. ter Horst,"A constrained Nevanlinna-Pick interpolation problem for matrix-valued functions" www.math.vt.edu/people/terhorst/Papers/BBtH1.pdf

[7] Dorato, P, Yunzhi L," A modification of the classical Nevanlinna-Pick interpolation algorithm with applications to robust stabilization", IEEE Transactions on Automatic Control, Volume: **31,** pp 645 – 648, July 1986.

[8] Christopher I. Byrnes, Tryphon T. Georgiou, Anders Lindquist, " A Generalized Entropy Criterion for Nevanlinna- Pick interpolation with degree constraint " IEEE Trans. Automat. Control, vol 38, issue 3, May-June 2003, pp 645-656.

[9] I. Pendharkar, H. K. Pillai and P. Rapisarda," A behavioral view of Nevanlinna-Pick interpolation", Proceedings of the44th IEEE Conference on Decision and Control, and the European Control Conference 2005,Spain, December 2005, pp836-841.

[10] Hari Bercovici, Juan C. Cockburn, Ciprian Foias and Allen Tannenbaum, "On structured tangential Nevanlinna-Pick interpolation," Proceedings of the 32nd IEEE Conference on Decision and Control, 1993, Vol. 4,pp. 3020-3021.

[11] B Aruna Rajan, R Devanathan, "Robust control of Permanent Magnet Synchronous Motor: Nevanlinna – Pick Approach", Proceedings 2008 Joint International conference on Power System Technology and IEEE Power India Conference, New Delhi, October 2008, pp 985-988.

[12] Gantmacher,"The Theory of Matrices", vol.1, pp337-338, Chelsea Publication Company, New York, USA, 1960.

[13] Roger.A.Horn, Charles.R.Johnson,"Matrix Analysis",Cambridge university press, N*ew York.,* 1990.

Implementation of Interval Type-2 Fuzzy Systems with Analog Modules

Mamta Khosla, R K Sarin

Department of Electronics and Communication
Engineering
Dr B R Ambedkar National Institute of Technology
Jalandhar- 144011, India
khoslam@nitj.ac.in

Moin Uddin

Delhi Technological University
Delhi

Prof_moin@yahoo.com

Abstract – **The objective of this paper is to present the implementation of Type-2 Fuzzy Logic System (T2 FLS) with analog modules. Specifically, this implementation focuses on Interval T2 FLSs (IT2 FLS) since they are computationally much simpler than general T2 FLSs and hence the most widely used in almost all the applications. The realization of IT2 FLS has been done by using two Type-1 Fuzzy Logic Systems (T1 FLS) which in turn have been implemented through analog modules built around operational-amplifiers (op-amps) and discrete components. The suggested implementation approach is based on rule-by-rule architecture and hence truly captures the inherent parallel nature of fuzzy systems.**

Keywords- Type-2 fuzzy logic systems, Interval type-2 fuzzy logic systems, analog modules, footprint of uncertainty.

I. INTRODUCTION

Type-2 fuzzy sets (T2 FSs) introduced by Zadeh [1] as an extension to the concept of fuzzy sets, now known as Type-1 Fuzzy Sets (T1 FSs) [2], are now well established and are gaining more and more popularity. T2 FSs allow us to handle linguistic uncertainties, as typified by the adage, *"words can mean different things to different people* [3]". Therefore, fuzzy logic systems based on T2 FSs are expected to perform better as they are capable of handling higher order of uncertainties, but at the same time are computationally intensive due to the involvement of a third dimension [4][5]. The third dimension in T2 FSs and footprint of uncertainty (FOU) provide an additional degree of freedom for T2 FLS to directly model and handle uncertainties.

Implementation of fuzzy systems depends on the application towards which it is addressed. Non-real time fuzzy systems are implemented through software with a high level programming language and this approach provides a high flexibility. On the other hand, software implementation is not adequate for applications demanding small size, low power consumption and high inference speed. A wide range of domains such as control, robotics, automotive etc. requires real-time operation and thus calls for hardware realization, The approaches for implementing T1 FLSs cover technologies like microcontrollers, Field Programmable Gate Arrays (FPGAs) among others [6]. The three basic processing stages of a T1 FLS are viz. fuzzification, inference and defuzzification. Since the general-purpose processors are exclusively sequential, so that these operations have to be performed serially that results in a low inference speed. The requirement of low inference speed or high response time can only be met by using dedicated hardware. The hardware realization of T1 FLSs is a well known and established area. On the contrary, the hardware realization of T2 FLSs in a relatively nascent research area. Their design and realization is more challenging due to high computational complexity, difficulty in visualization and non-availability of development tools. A few digital implementations reported in literature have been around microcontrollers, FPGAs etc. [7][8].

To the best of our knowledge, there is no report of a parallel hardware implementation of an IT2 FLS. This paper presents a simple arrangement for implementing IT2 FLS based on the averaging of two T1 FLSs, which have been realized with some standard analog modules built around op-amps and discrete components [9].

The paper is organized as follows. Section II briefly presents T2 FLS. The implementation of IT2 FLS presented in this paper is based on the use of two T1 FLSs and is explained in Section III. Simulation results are also presented in this section that validate the proposal. Section IV presents the various analog modules and their organization for IT2 FLS implementation. Finally, Section V draws conclusions.

II. INTRODUCTION TO T2 FUZZY SYSTEMS

Type-2 fuzzy system is a rule based fuzzy system that uses T2 FSs to describe its linguistic variables. A T2 FS can be informally defined as a fuzzy set that is characterized by a fuzzy or non-crisp membership function. This means there is uncertainty in the primary membership grades of a T2 membership function (MF) which introduces a third dimension to the MF, defined by the secondary membership grades [4].

A T2 FS denoted by Ã is characterized by a T2 MF $\mu_{\tilde{A}}(x,u)$, where $x \in X$ and $u \in J_x \subseteq [0,1]$. Ã can be expressed mathematically in the following form

$$\widetilde{A} = \{(x,u), \mu_{\widetilde{A}}(x,u) \mid \forall x \in X, J_x \subseteq [0\,1]\} \qquad (1)$$

Uncertainty in the primary memberships of a T2 FS set consists of a bounded region that we call the *footprint of uncertainty* (FOU). It is the union of all primary memberships [4] as in (2)

$$FOU(\tilde{A}) = \bigcup_{x \in X} J_x \qquad (2)$$

All the embedded FSs of FOU are T1 FSs and their union covers the entire FOU. FOU provides a very convenient verbal description of the entire domain of support for all the secondary grades of a T2 MF. It also lets us depict a T2 FS in two-dimensions instead of three dimensions which is easier to draw and makes it easy to understand the uncertainties inherent in that specific MF. The 2-D shape of the FOU depends on the type of uncertainty in the primary MF. However, because of the computational complexity of using a general T2 FLS, most designers only use IT2 FSs in a T2 FLS, the result being an IT2 FLS.
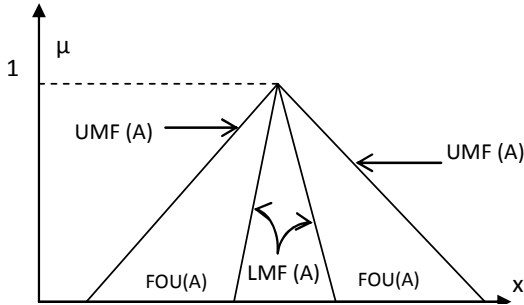
Figure 1.   FOU of an IT2 FS

IT2 is a special case of a T2 FS where all the secondary membership grades equal one. IT2 FS is completely characterized by its 2-D FOU that is bound by a Lower MF (LMF) and an Upper MF (UMF) as shown in Fig. 1. IT2 FSs are the most widely used T2 FSs to date used in almost all applications because all calculations are easy to perform. LMF and UMF together are popularly used in most of research papers to represent IT2 FLSs [10].

III.   IMPLEMENTATION METHODOLOGY FOR IT2 FLSS

The implementation of IT2 FLS presented in this paper is based on the use of two T1 FLSs to emulate a T2 FLS. The first T1 FLS is constructed using LMFs and the second one with the UMFs so as to emulate the FOU in a T2 FLS. The fuzzification, fuzzy inference and defuzzification are done as traditionally for two T1 FLSs and the outputs are averaged as shown in the Fig. 2.
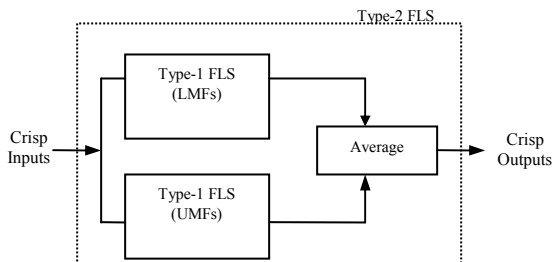
Figure 2.   Implementation methodology for IT2 FLS

This methodology has been validated though two case studies by the authors [11]. The advantage of using this approach is that all the existing realization methods for T1

FLSs can be adopted for T2 FLSs. This realization methodology has been utilized by Sepúlveda *et al.* [12, 13] in their works and they implemented the IT2 FLS on FPGA. The authors have also designed an IT2 fuzzy controller chip in $0.18\mu m$ CMOS technology based on this methodology [16]. This realization methodology reduces the computational burden while preserving the advantages of IT2 FLSs as it eliminates the requirement of type-reduction.

IV.   ORGANIZATION OF ANALOG MODULES FOR IT2 FLS IMPLEMENTATION

From architectural point of view, fuzzy systems are parallel systems and parallelism can be built into the system by adopting rule-by-rule architecture, which provides a data path for each rule [6]. The design of T1 FLS based on rule-by-rule architecture is shown in Fig. 3.

Each T1 FLS realizes a type of fuzzy inference where the rule consequents are constant values called singletons and each rule has the following format.

| *IF (x is A) AND (y is B)* | (3) |
| *THEN z = c* | |

In the above, *x* and *y* are input variable, *A* and *B* are fuzzy labels of *x* and *y*. Furthermore, *z* is an output variable and *c* is some constant. As compared to general case where the consequents are fuzzy sets, this type of fuzzy inference requires much less complex hardware [14].

Although each block in the figure can be implemented in analog or digital domain, but in this paper the focus is only on analog implementation. Each module in Fig. 3 can be realized using op-amps and other discrete components. Analog modules based design has been reported in literature [9] for T1 FLS realization that has been used for different applications [15]. We have arranged these analog modules to build two T1 FLSs and designed an additional *Averager* analog module that averages the defuzzified outputs of two T1 FLSs so as to give the overall output of IT2 FLS. The membership function generators (MFGs) are used for fuzzification i.e. to find the membership grades corresponding to the input variables.

For the purpose of simplicity, a fuzzy system with only two inputs and single output is considered. Further each input variable is represented by three IT2 FSs. Analog modules for implementing membership functions of different shapes viz. Z, triangular/trapezoidal and S are shown in Figs. 4, 5 and 6 respectively. The input variables are scale-mapped that transfer the range of input variables in the corresponding universe of discourse [0 +Vcc] and output voltage will be zero for all negative voltages, since Vee=0.

The fuzzy operator AND can be implemented using MIN module, which computes the rule firing strength. Scalar circuit calculates the weighted rule strength and MAX module is used for defuzzification. Analog modules for MIN operation (2 inputs), scaling operation for a single input, and for performing the MAX operation (9 inputs) are shown in Figs. 7, 8 and 9 respectively. The *Averager* circuit that averages the defuzzified outputs of the two T1 FLSs is shown in Fig 10. The parameters of the MFGs viz. slope and position are adjustable through the feedback resistor values RF and the

reference voltages Vref. Also the scaling factor of the *Scalar* and the averaging ratio of the *Averager* can be adjusted through the resistor values RF and R1. The simulations of these analog modules are shown in Figs. 11 to13.
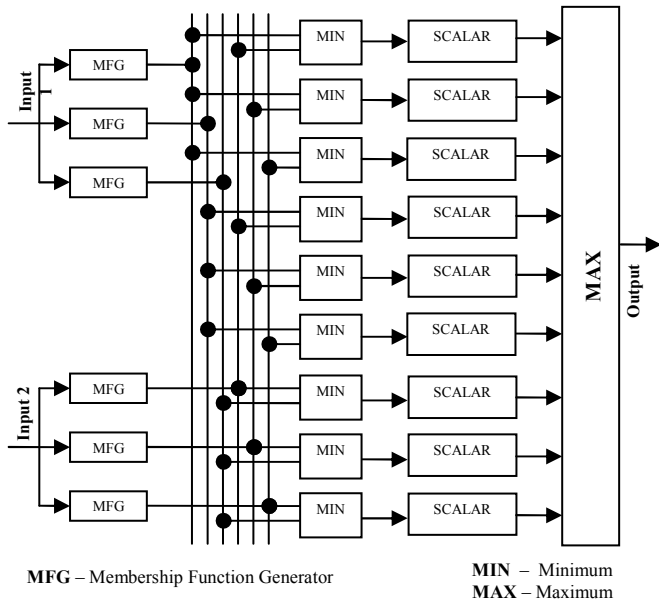
Figure 3.   T1 FLS based on Rule-by-Rule Architecture

MFG – Membership Function Generator

MIN – Minimum
MAX – Maximum

The values of the control feedback resistors and Vref are so chosen as to obtain the desired shapes of the fuzzy sets. The variation of slope of MF with different values of FB resistors is shown in Fig. 14. And the variation of the position of the MF with Vref values is shown on Fig. 15.
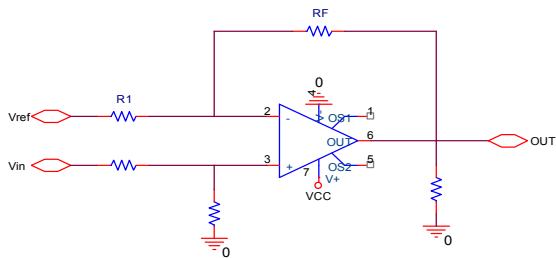
Figure 4.   Analog Module for Z MFG
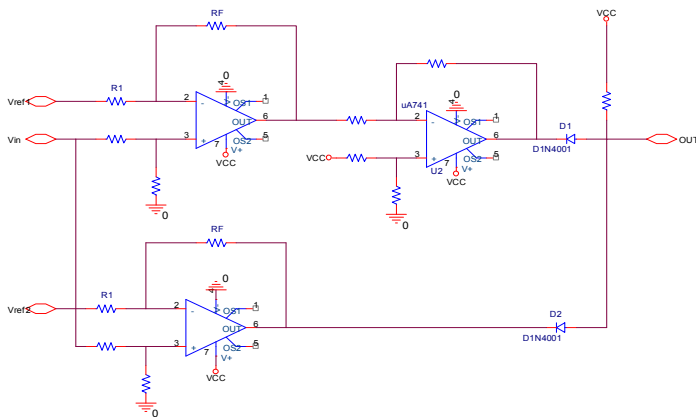
Figure 5.   Analog Module for Triangular/Trapezoidal MFG
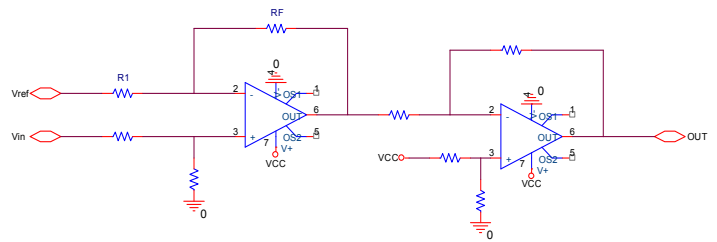
Figure 6.   Analog Module for S MFG
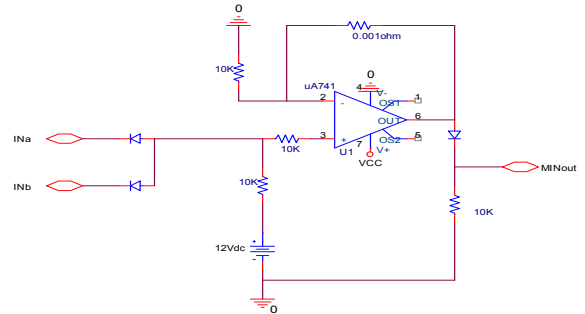
Figure 7.   Analog Module for MIN Operation (2 inpits)
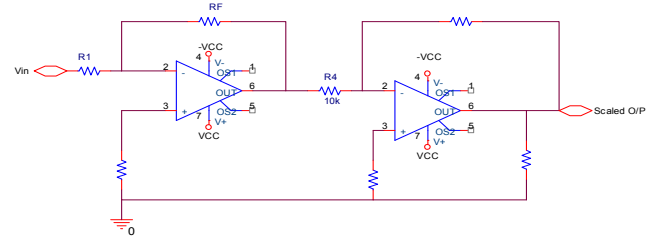
Figure 8.   Analog Module for *Scalar*
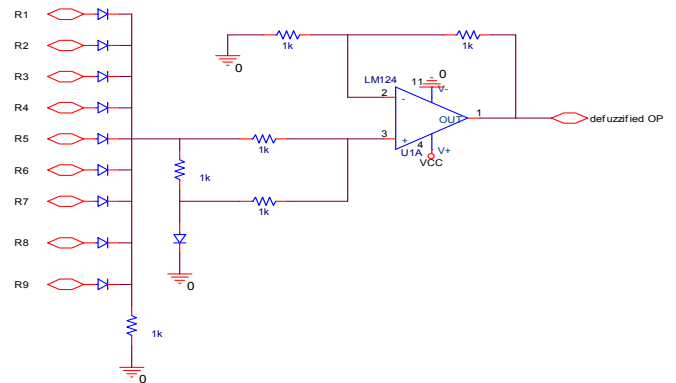
Figure 9.   Analog Module for MAX Operation (9 inputs)
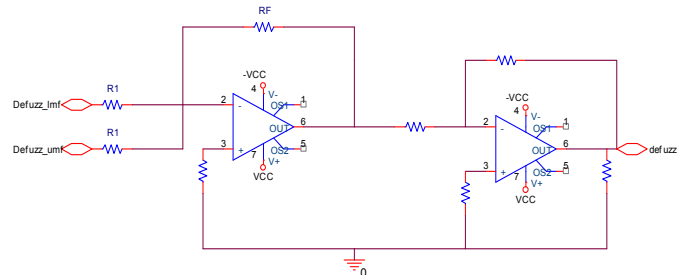
Figure 10. Analog Module for *Averager* (2 inputs)

An IT2 FLS based on the implementation methodology described in Section III and using various analog modules described in this section was created in OrCAD Spice simulator and the arrangement is shown in Fig. 16. The response of fuzzifier corresponding to the sweep of one of the two input variables that generate LMFs and UMFs is shown in Fig. 17. The load resistors shown in MIN, MAX and *Scalar* modules are connected only for measuring the individual responses. However, these load resistors are not required in the complete IT2 FLS realization as per the arrangement shown in Fig. 16. The number of discrete components required for implementation of each analog module is given in Table I. And the hardware budget for the realization of IT2 fuzzy processor is given in Table II.
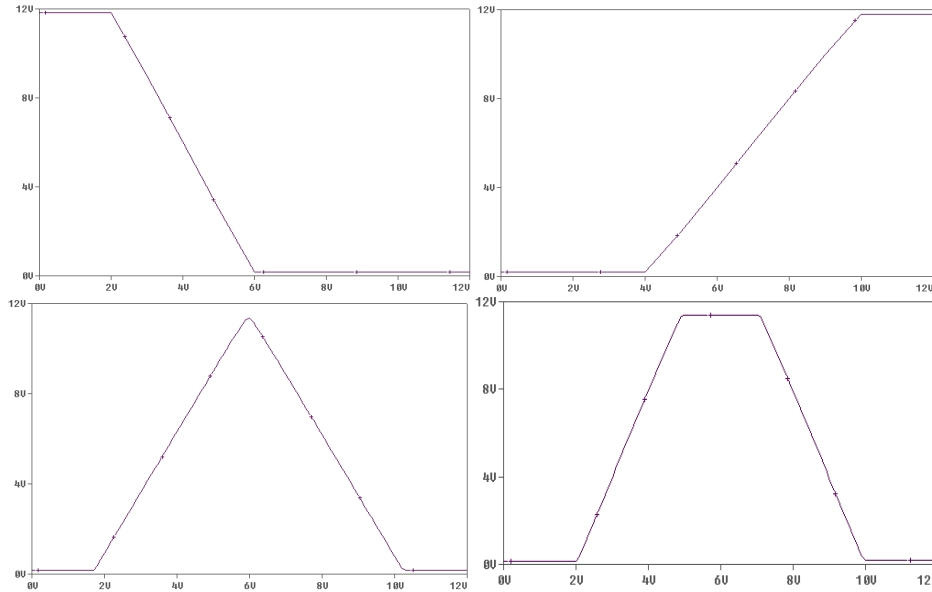
Figure 11. Simulation Results for (a) Z MF (b) S MF (c) Triangular MF (d) Trapezoidal MF
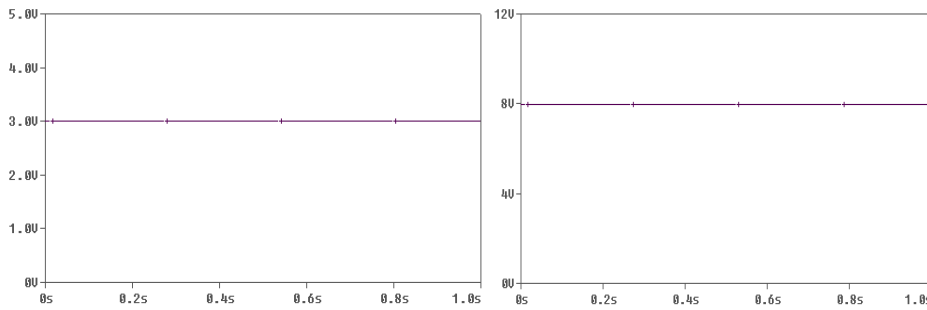
Figure 12. Simulation Results for (a) MIN (INa=8V, INb=3V) (b) MAX (2V, 3V, 4V, 5V, 6V, 7V, 8V, 1V, 4.5V)
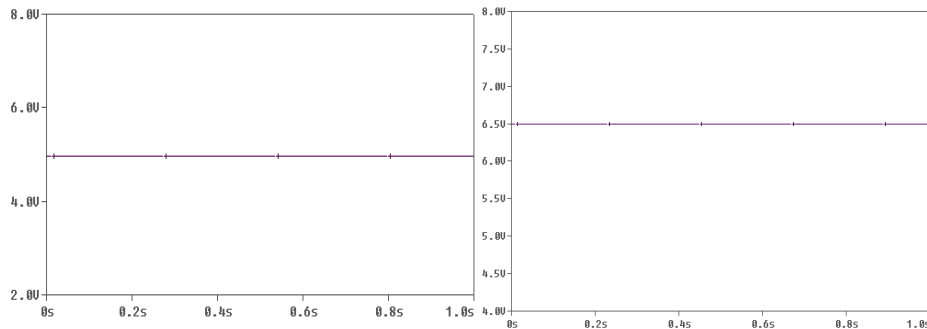
Figure 13. Simulation Results for (a) *Scalar* (input voltage=0.5V, scaling factor=10) (b) *Averager* (average of 8V and 5V)
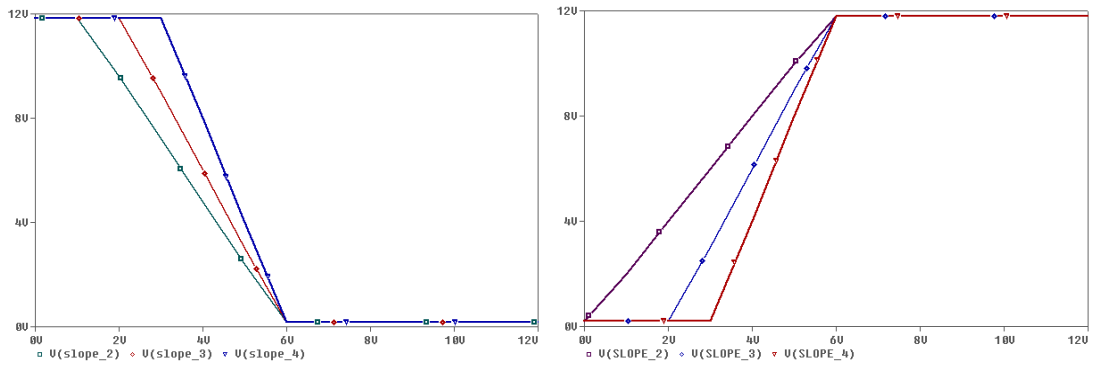
Figure 14. Slope variation (a) Z MF with slope (RF/R1)=2, 3, 4, (Vref=6V) (b) S MF with slope (RF/R1)=2, 3, 4, (Vref=6V)
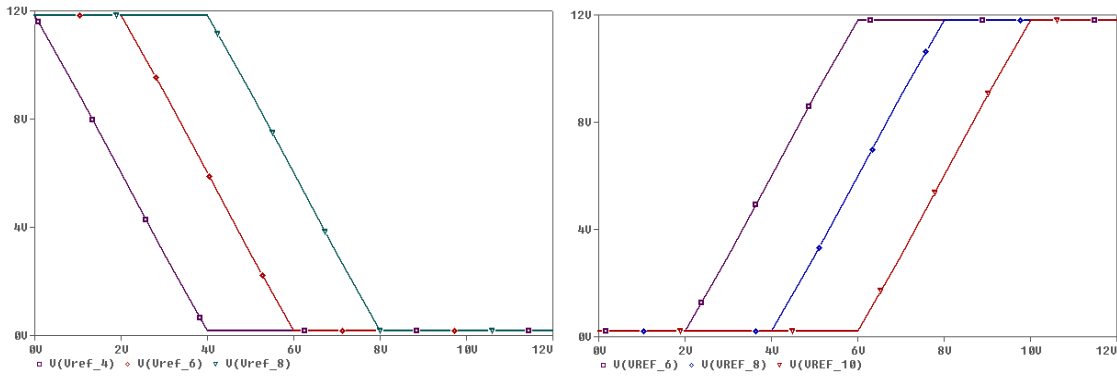
Figure 15. Position Variation (a) Z MF with Vref=4, 6, 8V {slope (RF/R1)=3}, (b) Z MF with Vref=6, 8, 10V {slope (RF/R1)=3}
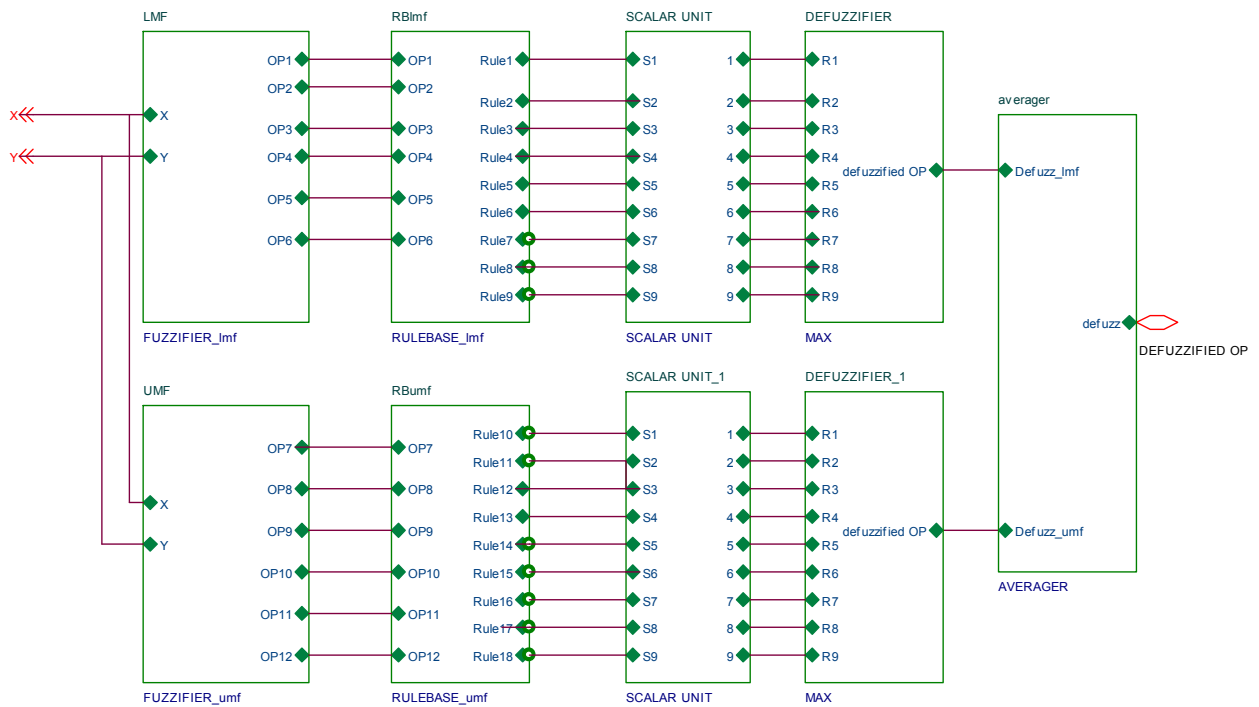
Figure 16. IT2 FLS Implementation created in OrCAD
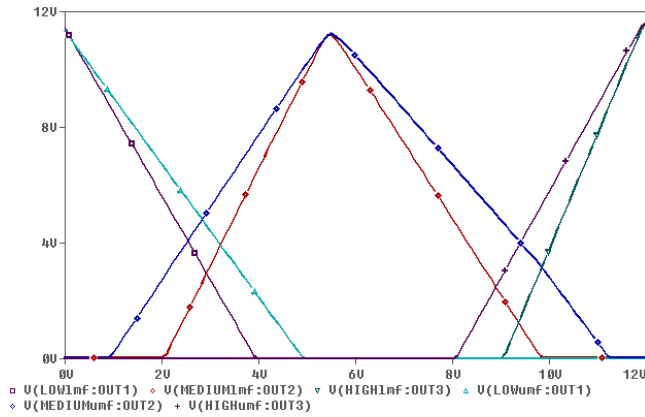
Figure 17. Simulation Results of Fuzzifier corresponding to one input variable

TABLE I.   HARDWARE BUDGET FOR EACH ANALOG MODULE

Analog Module	OP-Amp	Diode	Resistor
Z MFG	1	0	4
Triangular MFG	3	2	13
S MFG	2	0	8
MIN	1	3	4
MAX	1	10	6
Scalar	2	0	6
Averager	2	0	8

TABLE II.   HARDWARE BUDGET FOR IT2 FLS

Analog Module	No.	OP-Amp	Diode	Resistor
Z MFG	2	2	0	8
Triangular MFG	2	6	4	26
S MFG	2	4	0	16
MIN	18	18	54	72
MAX	2	2	20	12
Scalar	18	36	0	108
Averager	1	2	0	8
Total no. of Components		**70**	**78**	**250**

V.   CONCLUSIONS

This paper presents a simple, easy and effective implementation of IT2 FLS with analog modules built around op-amps and discrete components. The realization is based on use of two T1 FLSs. The implementation follows rule-by-rule architecture and would provide a very high inference speed and hence good performance. Even people with non-electrical background and having little understanding of basic electronic components can realize IT2 FLS for their applications.

However, the limitation of this arrangement is lack of flexibility and excessive wiring.

REFERENCES

[1] L. A Zadeh, "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning – 1", Information Sciences 8, pp. 199–249 1975.

[2] L. A Zadeh, "Fuzzy Sets", Information and Control, vol. 8, pp. 338–353, 1965.

[3] J. M. Mendel, "Computing with Words, When Words can Mean Different Things to Different People", Proceedings of 3rd International ICSC Symposium on Fuzzy Logic and Applications, Rochester University, Rochester, NY. 1999.

[4] J. M. Mendel and R.I. John, "Type-2 Fuzzy Sets Made Simple", IEEE Transaction on Fuzzy Systems, vol. 10, no. 2, pp. 117–127, 2002.

[5] N. N. Karnik and J.M. Mendel, "Applications of Type-2 Fuzzy Logic Systems to Forecasting of Time-Series", Information Sciences 120, pp. 89–111, 1999.

[6] Baturone, Barriga, Carlos Jimenez-Fernandez and Diego, R. Lopez, Microelectronic Design of Fuzzy-Logic Based Systems, CRC Press, 2000.

[7] J. Bulla, G. Sierra and M. Melgarejo, "Implementing a Simple Microcontroller-Based Interval Type-2 Fuzzy Processor", Proceedings of 51st Midwest Symposium on Circuits and Systems (MWSCAS), pp. 69-72, 2008.

[8] M. A. Melgarejo and C.A. Pena-Reyes, "Hardware architecture and FPGA implementation of a type-2 fuzzy system", Proc. ACM GLSVLSI, Boston, MA, p. 458–461, 2004.

[9] A. Sanz, "Analog Implementation of Fuzzy Controller," Proceedings of 3rd IEEE Conference on Fuzzy Systems, pp. 279-283, 1994.

[10] Q. Liang, J.M. Mendel, "Interval type-2 fuzzy logic systems: Theory and design", IEEE Trans. Fuzzy Syst., vol. 8, no. 5, pp. 535–549, 2000.

[11] Mamta Khosla, Rakesh Kumar Sarin, Moin Uddin, Satvir Singh, Arun Khosla, *Realizing Interval Type-2 Fuzzy Systems with Type-1 Fuzzy Systems,* Book Chapter for the Book Cross-Disciplinary Applications of Artificial Intelligence and Pattern Recognition: Advancing Technologies, Eds.: Vijay Mago and Nitin Bhatia, IGI Global, USA, 2012, pp. 412-427.

[12] Roberto Sepúlveda, Oscar Castillo, Patricia Melin, Oscar Montiel: An Efficient Computational Method to Implement Type-2 Fuzzy Logic in Control Applications. Analysis and Design of Intelligent Systems using Soft Computing Techniques 2007, pp. 45-52.

[13] Roberto Sepúlveda, Oscar Montiel, Gabriel Lizárraga, Oscar Castillo: Modeling and Simulation of the Defuzzification Stage of a Type-2 Fuzzy Controller Using the Xilinx System Generator and Simulink. Evolutionary Design of Intelligent Systems in Modeling, Simulation and Control 2009, pp. 309-325.

[14] T. Yamakawa, "A Fuzzy Inference Engine in Nonlinear Analog Mode and its Application to a Fuzzy Logic Control", IEEE Transanction on Neural Network, Vol. 4, pp 496-522, May 1993.

[15] Satvir Singh and J.S. Saini, "Captive Power Management using Op-amp based FLS", Proceedings of Asian Conference on Intelligent Systems and Networks, Jagadhari, February 24-25, 2006,India, pp. 160-166.

[16] Mamta Khosla, Rakesh Kumar Sarin and Moin Uddin, "Design of an Analog CMOS based Interval Type-2 Fuzzy Logic Controller Chip", International Journal of Artificial Intelligence and Expert Systems, vo. 2, no.4, pp. 167-183, 2011.

Development of MATLAB GUI for Multivariable Frequency Sampling Filters Model

Muhammad Hilmi R. A. Aziz[1], Lee Kim Huat[2] & Rosmiwati Mohd-Mokhtar[3]

School of Electrical & Electronic Engineering

Universiti Sains Malaysia, Engineering Campus

14300, Nibong Tebal, Pulau Pinang, Malaysia

mhilmiee87@gmail.com[1]; jaysonlee208@hotmail.com[2] & rosmiwati@ieee.org[3]

Abstract—**This paper presents a MATLAB GUI platform for multivariable frequency sampling filters algorithm. The developed GUI is aimed to deliver information and summarization of analysis in terms of graphical representation for the user. The GUI interface allow SISO, SIMO, MISO and MIMO data systems to be employed directly by user in order to be processed using frequency sampling filters model. From FSF GUI, results based on step and frequency response estimates of the tested system are illustrated. The analysis of rise time, settling time, maximum overshot and peak time of the system is also calculated based on the response obtained. By performing frequency sampling filters, the compressed, cleaned and unbiased data will be obtained, in which, it can be further used to develop a model of the system, conduct time and frequency response analysis and also plan for next control system design.**

Keywords-MATLAB GUI; frequency sampling filters; data compression; system identification

I. INTRODUCTION

Frequency sampling filters (FSF) structure was first introduced to the areas of system identification and automatic control by [1] and [2]. Research in the signal-processing field is looking at designing modified frequency-sampling filters that account for finite word-length effects [3]. In 1997, FSF model has been used to estimate a step and frequency response model along with its corresponding time domain confidence bounds [4]. This method is also referred as data compression process as the raw data will be analyzed in order to obtain only an important and meaningful parameter that describes the empirical model of the analyzed data. From FSF, the step response estimates and frequency response estimates with reduced number of data are obtained. Both of these responses can be used to identify a mathematical model of a system. There are many successful results demonstrated for implementations of FSF method in the identification field. One of the most recent publications discussed the development of FSF into multivariable platform process [5,6].

The FSF approach perform a frequency decomposition of the identified model, separating low and medium frequency parameters from high frequency parameters and then choose to ignore these high frequency parameters in the final model

structure. This frequency decomposition is a linear transformation of the FIR model [7,8]. Figure 1 illustrates the SISO FSF model structure.
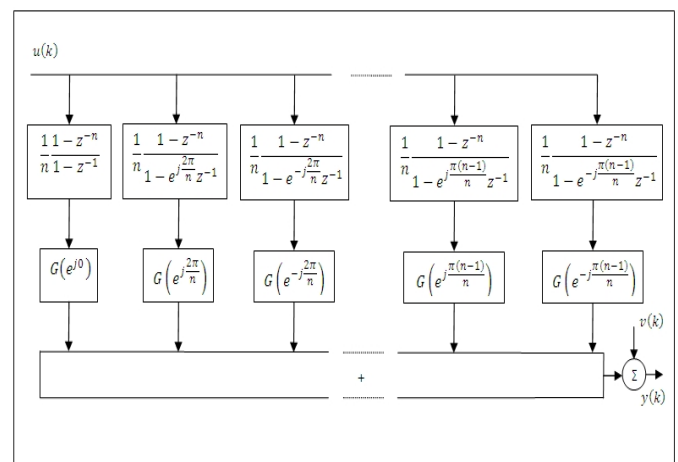
Figure 1. FSF model structure [8]

In this figure, it is shown that the process input first passes through the set of frequency sampling filters arranged in parallel. Then, the output of each filter is weighted by the discrete-time process frequency response evaluated at the corresponding centre frequency. Finally, the weighted filter outputs are summed to form the noise-free process output. The FSF filters are narrow band limited around their respective centre frequencies. All the filters have identical frequency responses except for the location of their centre frequencies.

In this paper, a graphical user interface (GUI) is developed for multivariable frequency sampling filters model by using MATLAB software. GUI can be easily understood as the interface between user and computer program. The history of computer technology shown that computer in the early history just consists of the blank screen with only a prompt that gives no indication of user input. User unable to clearly see the desired output based on the input key in. To solve this problem, GUI is being implemented. At more understanding level, GUI can be easily understand as interface between computer and human by which people and computers communicate with each other [9]. A good GUI design

removes the impediment of communication with the computer system and allows the user to work directly on the problem at hand.

In terms of computer science and engineering, the GUI is a visual operating display that the monitor presents on the monitor to the computer [10]. Almost all GUI have common characteristic such as windows, icons, menus, and push-buttons. GUI allows users to issue commands to the other computer applications. GUI usually has three major components; those are windowing system, imaging model and application program interface (API). The windowing system builds the windows, menus, and dialog boxes that appear on the screen. The imaging model defines the fonts and graphics that appear on the screen. Lastly, the API is the means in which the user specifies how and what windows and graphics appear on the screen.

In this paper the GUI for multivariable FSF model is developed as to facilitate the user with user friendly domain for data compression and data analysis especially for the large amount of data collected from real plant/system. The developed platform will aid any ordinary user who wishes to process the raw data without needing true knowledge to understand the real algorithm that runs within it. This will become the main contribution of this paper.

In brief, this paper is described as follow. Section II describes and elaborates about FSF process. Details on development of MATLAB GUI to integrate with FSF will be elaborated in Section III. Then, the simulations based on implementation of developed GUI will be illustrated in section IV. Finally, Section V concludes the paper.

II. FREQUENCY SAMPLING FILTERS FRAMEWORK

FSF approach involves the usage of FIR model and the ML (maximum likelihood) method which play role in eliminating the bias and noise effects of the data collected from the systems [11,12]. The FIR model is widely used in the field of process identification because it requires no prior knowledge about the process other than its settling time. In order to obtain a proper FSF parameter optimization, the least squares model estimates using the orthogonal decomposition algorithm based on Predicted Residual Sum of Squares (PRESS) computation is used [13]. This computation act to minimize the prediction error as well as determine the most suitable candidate for dynamic model structure. Figure 2 shows the clear illustration for the FSF process.
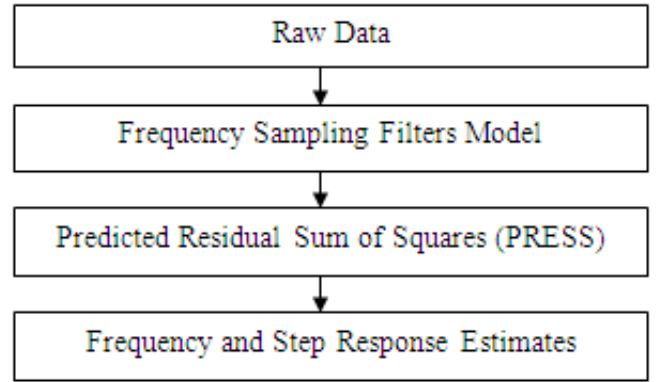
Figure 2. Flow of FSF process

A. FSF Model

The FSF approach used in this paper is originally obtained from [7,8]. The FSF model can be identified as

$$y(k) = G(z)u(k) + v(k) \qquad (1)$$

where $u(k), y(k)$ is input and output signal respectively. $v(k)$ is the zero mean disturbance term. $G(z)$ is described as

$$G(z) = \sum_{i=0}^{n-1} \frac{1}{N} \sum_{m=-\frac{n-1}{2}}^{\frac{n-1}{2}} G(e^{j\frac{2\pi m}{n}}) e^{j\frac{2\pi mi}{n}} z^{-i} \qquad (2)$$

Interchanging the summations in (2) gives the transfer function in its FSF model form

$$G(z) = \sum_{m=-\frac{n-1}{2}}^{\frac{n-1}{2}} G(e^{j\frac{2\pi m}{n}}) \frac{1}{N} \frac{1-z^{-n}}{1-e^{j\frac{2\pi m}{n}} z^{-1}} \qquad (3)$$

where

$$\sum_{i=0}^{n-1} e^{j\frac{2\pi mi}{n}} z^{-i} = \frac{1-z^{-n}}{1-e^{j\frac{2\pi m}{n}} z^{-1}} \qquad (4)$$

Define a set of transfer functions extracting from (3)

$$H^m(z) = \frac{1}{N} \frac{1-z^{-n}}{1-e^{j\frac{2\pi m}{n}} z^{-1}} \qquad (5)$$

for $l = 0, \pm 1, \pm 2, ..., \pm \dfrac{n-1}{2}$, the above equation is referred as the m-th FSF with the centre frequency of the m-th filter is at $\dfrac{2\pi n}{n}$ radians. Let $z = e^{j\omega}$, Equation (3) will become

$$G(e^{j\omega}) = \sum_{m=-\frac{n-1}{2}}^{\frac{n-1}{2}} G(e^{j\frac{2\pi n}{n}}) \frac{1}{N} \frac{1 - e^{-j\omega n}}{1 - e^{j\frac{2\pi n}{n}} e^{-j\omega}} \qquad (6)$$

At $\omega = \omega_l$, the following condition holds

$$H^a(e^{j\omega}) = 0 \qquad \text{for} \qquad a \neq l$$
$$H^a(e^{j\omega}) = 1 \qquad \text{for} \qquad a = l$$

where a is an integer like l in the range $[-\dfrac{n-1}{2}, \dfrac{n-1}{2}]$. In this case, the value of the process frequency response in (6) reduces to the value of the process frequency response coefficient $G(e^{j\frac{2\pi n}{n}})$.

Define the parameter vector as

$$\theta = \begin{bmatrix} G(0) \\ G(e^{j\frac{2\pi}{n}}) \\ G(e^{-j\frac{2\pi}{n}}) \\ \vdots \\ G(e^{j\left(\frac{(n-1)\pi}{n}\right)}) \\ G(e^{-j\left(\frac{(n-1)\pi}{n}\right)}) \end{bmatrix} \qquad (7)$$

and its corresponding vector regressor vector as

$$\phi(k) = \begin{bmatrix} f_0(k) * u(k) \\ f_1(k) * u(k) \\ f_{-1}(k) * u(k) \\ \vdots \\ f_{\frac{n-1}{2}}(k) * u(k) \\ f_{-\frac{n-1}{2}}(k) * u(k) \end{bmatrix} \qquad (8)$$

where $f_m(k)$ is defined according to (5). Thus, (1) can be rewritten as

$$y(k) = \theta^T \phi(k) + v(k) \qquad (9)$$

For N data measurements, (9) can also be written in matrix form as

$$Y = \Theta\Phi + V \qquad (10)$$

Equation (10) can then be used to solve the least squares estimate of Θ given by

$$\hat{\Theta} = (\Phi_N^T \Phi_N)^{-1} \Phi_N^T Y_N \qquad (11)$$

which minimize the performance index of the form

$$J(N, \hat{\Theta}) = \sum_{k=0}^{N} |Y - \Pi\Phi|^2 \qquad (12)$$

The matrix $(\Phi_N^T \Phi_N)$ is called the correlation matrix and the invertibility condition on this matrix is sometimes called the sufficient excitation condition for parameter estimation. In order to obtain a proper FSF parameter optimization, the least squares model estimates based on PRESS (Predicted Residual Sum of Squares) computation is used. The PRESS criterion will ensure that the FSF model has the greatest predictive capability among all its candidate models [13].

B. The PRESS Criterion

The idea of PRESS is to set aside each data point, estimate a model using the rest of the data, and then evaluate the prediction error at the point that was removed. The PRESS statistic can be applied as a criterion for model structure detection in dynamic system identification.

Define the prediction error as

$$\begin{aligned} e_{-k}(k) &= y(k) - \hat{\theta}^T \phi(k) \\ &= y(k) - \hat{y}_{-k}(k) \end{aligned} \qquad (13)$$

where $e_{-k}(k), k = 1, 2, ..., N$ are called the PRESS residuals and $\hat{\theta}$ has been estimated according to (11) without including $\phi(k)$ and $y(k)$. The PRESS residuals $e_{-k}(k)$ represent the true prediction errors and can be calculated according to the following equation

$$e_{-k}(k) = \frac{e(k)}{1 - \phi(k)^T (\Phi^T \Phi)^{-1} \phi(k)} \qquad (14)$$

The PRESS statistic is defined as

$$PRESS = \sum_{k=1}^{N} e_{-k}(k)^2 \qquad (15)$$

The average PRESS is calculated as

$$PRESS_{av} = \sqrt{\frac{\sum_{k=1}^{N} e_{-k}(k)^2}{N-1}} \qquad (16)$$

Both (15) and (16) provide measures of the predictive capability of the estimated model. In terms of model structure selection, the chosen structures are based on the smallest PRESS value.

C. Step and Frequency Response Estimate Using FSF Model

In the estimation of step response, the description of the system using FSF can be described as follow

$$y(k) = \sum_{m=-\frac{n-1}{2}}^{\frac{n-1}{2}} G(e^{j\frac{2\pi m}{n}})H^m(z)u(k)+v(k) \qquad (17)$$

where for a suitable choice of $G(e^{j\frac{2\pi m}{n}})$ and $H^m(z)$ define as in (2) and (5) respectively. Upon obtaining the estimate of the frequency response parameters (according to FSF model and PRESS criterion), the estimate of the step response at sampling instant, m can be expressed by

$$\hat{g}_m = \sum_{i=0}^{m-1} \hat{h}_i \qquad (18)$$

where the estimated impulse response coefficients $\hat{h}_0, \hat{h}_1, \hat{h}_2,...,\hat{h}_{m-1}$ are related to frequency response via

$$\hat{h}_i = \frac{1}{N} \sum_{m=-\frac{n-1}{2}}^{\frac{n-1}{2}} \hat{G}(e^{j\frac{2\pi m}{n}})e^{j\frac{2\pi m i}{n}} \qquad (19)$$

By substituting (19) into (18), the estimated step response coefficient can be rewritten as

$$\hat{g}_m = \sum_{m=-\frac{n-1}{2}}^{\frac{n-1}{2}} \hat{G}(e^{j\frac{2\pi m}{n}}) \frac{1}{N} \frac{1-e^{j\frac{2\pi m}{n}(m+1)}}{1-e^{j\frac{2\pi m}{n}}} \qquad (20)$$

Although the FSF approach is cast in the discrete-time domain and the corresponding z-transform domain, the resultant model can be used to obtain continuous time step response. The system impulse response $\hat{g}(t)$ can be approximately computed using the continuous time equivalent as

$$\hat{g}(t) \approx \hat{g}_{fsf}(t) = \sum_{m=-\frac{n-1}{2}}^{\frac{n-1}{2}} \hat{\theta}_m \hat{h}_m(t) \qquad (21)$$

and

$$\hat{h}_m(t) = \frac{1}{T} e^{j\frac{2\pi m i}{T}} \quad \text{for} \quad t < T\Delta \qquad (22)$$

where T is a sampling period. The step response is determined as

$$y_s(t) = \int_0^t \hat{g}(\tau)d\tau \qquad (23)$$

In the other hand, the frequency response estimates is obtained over a transformation of $z = e^{j\omega\Delta}$. Thus, in frequency domain form, the equations can be rewritten as

$$G(j\omega) \approx G_{fsf}(j\omega) = \sum_{k=-\frac{n-1}{2}}^{\frac{n-1}{2}} \theta_k H_k(j\omega) \qquad (24)$$

$$H_k(j\omega) = \overline{H}_k(e^{j\omega\Delta}) \quad \text{for} \quad \omega < \aleph\Omega \qquad (25)$$

where n is odd and the frequency sampling interval $\Omega = \frac{2\pi}{T}$, $H_k(z)$ is the k-th FSF and θ_k is the corresponding (complex) parameter for the frequency range of $0 \le \omega \le \aleph\Omega$.

III. CREATING GRAPHICAL USER INTERFACES

Creating a GUI involves two basic tasks those are; (i) Laying out the GUI components and (ii) Programming the GUI components. Three main principal elements required to create a MATLAB Graphical User Interface are [14]:

- Components: such as pushbuttons, labels and edited boxes. The types of components include graphical controls (pushbuttons, edit boxes, lists, sliders, etc.), static elements (frames and text strings), menus, and axes.
- Figures: Each and every components of the GUI must be arranged in an area within a figure, which is the layout of the GUI in computer screen. Figures also can be automatically created to represent the data in the form of graph.
- Call-backs: Call-backs function performs desire action whenever users click a mouse on a button or insert information to the application program by using keyboard. A good MATLAB program for graphical user interface must be able to respond to each event which represent by mouse click or key press. For instance, if a user clicks on a button, that event must cause the MATLAB code that implements the function of the button to be executed. Call back is actually the process where the code executed as response to an event.

MATLAB allow user to design graphical user interface via a tool called GUIDE. These tools greatly simplify the process of laying out and programming a GUI [15]. GUIDE is a set of layout tools and also generates an M-file that contains code to handle the initialization and launching of the GUI. This M-file provides a framework for the implementation of the call-backs. Figure 3 shows a flowchart of the developed GUI.

IV. SIMULATIONS

Figure 4 illustrated the MATLAB GUI which is developed in order to integrate with the multivariable FSF. This GUI consists of four main portions which are type of data, data menu, option menu and step response analysis.

Each portions of this MATLAB GUI has been separated by using different panels. There are four different types of push buttons which represents four different types of data. User can choose which type of data they wish to process by clicking on the respective push button. This will enable the data menu for the selection data type and disable the others data menu in one time. The limitation numbers of input/output for SIMO, MISO and MIMO data due to the limitation of the developed multivariable FSF algorithm.
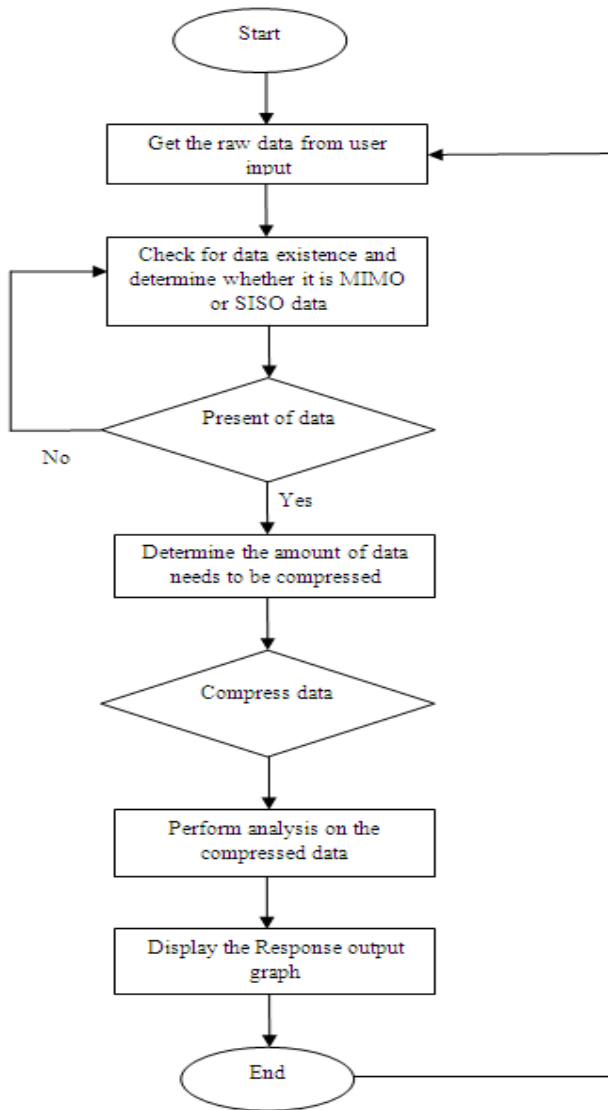
Figure 3. GUI design flowchart

With this MATLAB GUI, user can directly load their data and set the number of compression data that they need. Moreover, the analysis of rise time, settling time, maximum overshot and peak time of the tested system will be calculated and shown in this GUI. In addition, the user can save the step or frequency response estimates obtained so that it can be used to developed mathematical model of the systems or for further control analysis.

In this paper, it only involves one windows of GUI. If the number of input/output for multivariable FSF can be added for the next research work, each type of data can be considered to be displayed in the separate windows. In this case, only one window has been used because of loading time factor other than to make it user friendly. The functionality of this developed MATLAB GUI to integrate with multivariable FSF has been tested. The following example demonstrated the results obtained.

Example – MISO system (pH Neutralization Data)

Details description of this data can be referred in [6]. The results obtained from this running process also can be compared with the results have been published in this paper. This example only want to show the functionality of this developed MATLAB GUI integrate with multivariable FSF algorithm. Figure 4 and Figure 5 illustrated the step and frequency response estimates obtained respectively. The analysis of rise time, settling time, maximum overshot and peak time of this data can be seen in Figure 6.
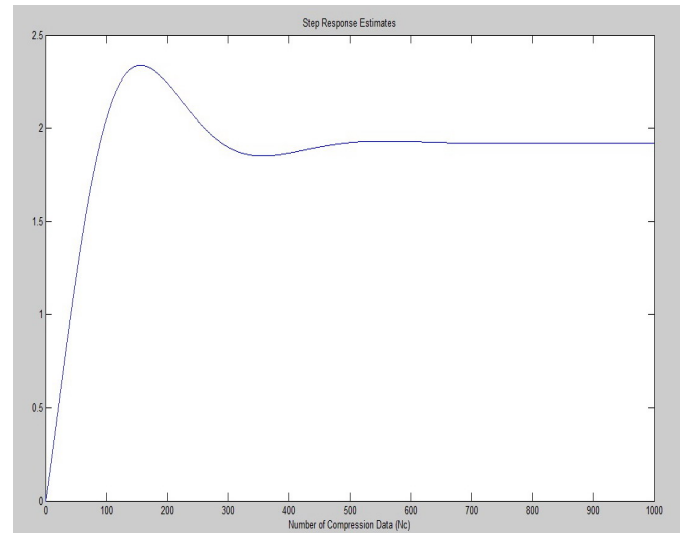
Figure 4. Step response obtained from MISO FSF model
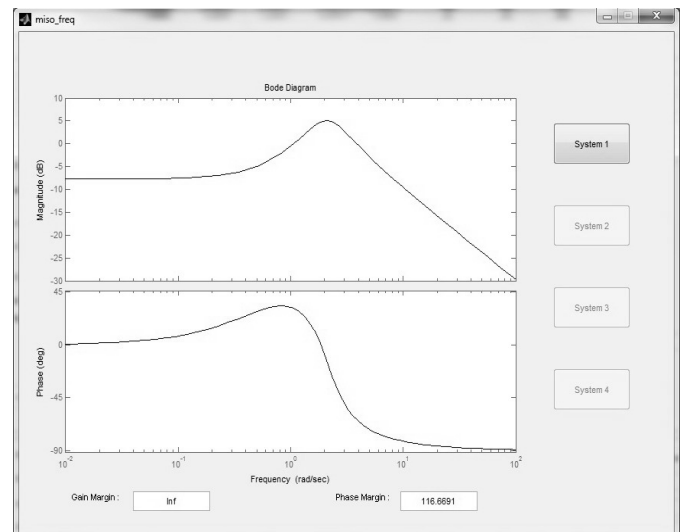
Figure 5. Bode diagram obtained from MISO FSF model

V. CONCLUSION

The developed MATLAB GUI software which is user friendly tool for multivariable FSF algorithm has been presented in this paper. All the operating conditions are mathematically modelled using MATLAB instruction in GUI environment.
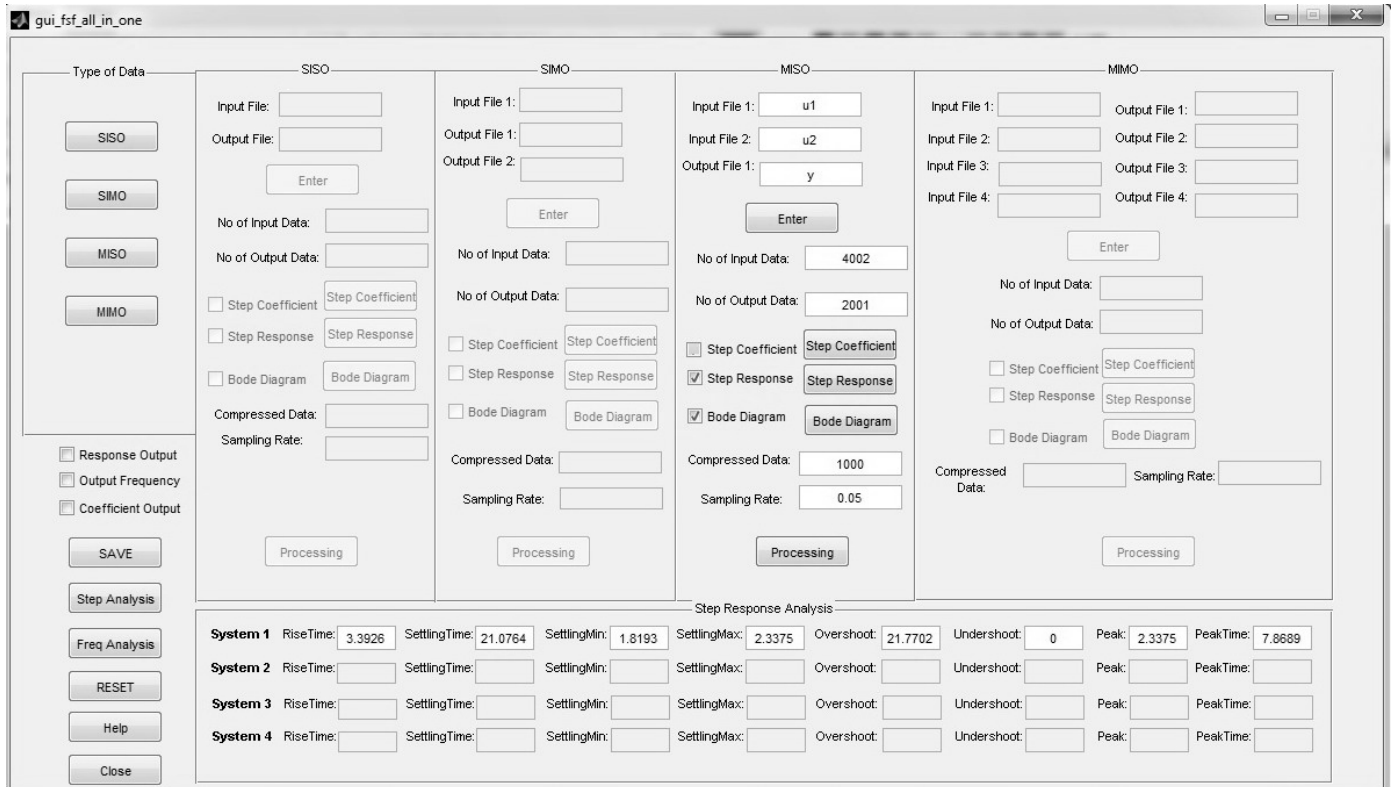
Figure 6. MATLAB GUI for multivariable FSF

REFERENCES

[1] R. R. Bitmead and B. D. O. Anderson, "Adaptive Frequency Sampling Filters," IEEE Trans. On Circuits and Systems, vol. 28(6), pp. 524-534, 1981.

[2] P. J. Parker and R. R. Bitmead "Adaptive Frequency Response Identification," Paper presented at the 26th IEEE Conf. on Decision and Control, Australia,1987.

[3] P. A. Stubberud and C. T. Leondes "A Frequency Sampling Filter Design Method Which Accounts for Finite Word Length Effects," IEEE Trans. on Signal Processing, vol. 42(1), pp. 189-193, 1994.

[4] W. R. Cluett, L. Wang and A. Zivkovic, "Development of quality bounds for time and frequency domain models: application to the Shell distillation column," J. Process Control, vol. 7(1), pp. 75-80, 1997.

[5] M. H. R. A. Aziz and R. Mohd-Mokhtar, "Identification of MIMO Magnetic Bearing System Using Continuous Subspace Method with Frequency Sampling Filters Approach," Paper presented at the 37th Annual Conference of the IEEE Industrial Electronics Society, Melbourne, Australia, 2011.

[6] M. H. R. A. Aziz and R. Mohd-Mokhtar, "Multi Input Multi Output Frequency Sampling Filters for Real System Applications," Paper presented at the International Conference on Electrical Engineering and Informatics, Bandung, Indonesia, 2011.

[7] L. Wang and W. Cluett, "Frequency-sampling Filters: An improved model Structure for Step-response Identification," Automatica, vol. 33(5), pp. 939-944, 1997.

[8] L. Wang and W. Cluett, "From Plant Data to Process Control: Ideas for Process Identification and PID Design," London: Francis & Taylor, 2000.

[9] B. Gui, "Interpretations of Human User Interface," Visible Language, vol. 24(3), pp. 262-285, 1990.

[10] B. A. Harding, "Windows, Icon & Mice, Oh My! The Changing Face of Computing," Paper presented at the 19th Annual Conf. of Frontiers in Educatioon, 1989.

[11] L. Wang, P. Gawthrop and C. Chessari, "Indirect approach to continuous time system identification of food extruderm," Journal of Process Control, vol. 14(6), pp. 603-615, 2004.

[12] R. Mohd-Mokhtar, M. H. R. A. Aziz, M. R. Arshad and N. A. A. Hussain, "Data compression for underwater glider system using frequency sampling filters," Indian Journal of Geo-Marine Sciences, vol. 40(2), pp. 227-235, 2011.

[13] L. Wang and W. R. Cluett, "Use of PRESS Residuals in Dynamic System Identification." Automatica, vol. 32(5), pp. 781-784, 1996.

[14] S. J. Chapman, "MATLAB Programming for Engineers," Brooks Cole, 2001.

[15] G. D. Petropol-Serb, I. Petropol-Serb, A. Campeanu and A. Petrisor, "Using GUI of Matlab to create a virtual laboratory to study an induction machine," Paper presented at the International Conference on "Computer as a Tool", Warsaw, 2007.

Multiobjective Quantum-Inspired Evolutionary Programming for Optimal Load Shedding

Zuhaila Mat Yasin
Faculty of Electrical Engineering,
Universiti Teknologi Mara
Shah Alam, Malaysia
e-mail: yzuhaila@hotmail.com

Titik Khawa Abdul Rahman
Faculty of Engineering
Universiti Pertahanan Nasional Malaysia
Kuala Lumpur, Malaysia
e-mail: titik.khawa@upnm.edu.my

Zuhaina Zakaria
Faculty of Electrical Engineering,
Universiti Teknologi Mara
Shah Alam, Malaysia
e-mail: zuhainaz@salam.uitm.edu.my

Abstract— **This paper present new technique namely Quantum-Inspired Evolutionary Programming (QIEP) to determine the optimal location and optimal amount of load to be shed for undervoltage load shedding schemes. This approach is based on the concept of quantum mechanics in the Evolutionary Programming (EP). Quantum-Inspired is implemented according to three levels defined by quantum individuals, quantum groups and quantum universe in order to improve the speed of the algorithm. The QIEP is employed to search for the best location and amount of load to be shed based on multiobjective functions which are power loss minimisation, voltage profile improvement and power interruption cost minimisation. The effectiveness of multiobjective QIEP optimisation technique is illustrated in IEEE 33-bus distribution test system, IEEE 69-bus distribution test system and 141-bus distribution system. The results were also compared with other techniques in terms of fitness values and computation time.**

Keywords- **Quantum-Inspired Evolutionary Programming, Optimal Load Shedding, Multiobjective Optimisation**

I. INTRODUCTION

Power systems are currently being operated closer to their limits with the increasing in demand of electric power due to economic and environmental constraints. This has led to a risk of voltage instability. Among different countermeasures for the prevention of voltage instability, load shedding is the last line of defence when there is no alternative to stop an approaching voltage collapse [1]. During an emergency situation, system operators are required to make load shedding decisions based on system security concerns such as voltage, current, power and frequency to alleviate constraints and maintain system stability [2]. Various types of load shedding schemes have been studied by researchers in the past. Underfrequency load shedding strategy has been presented in order to identify amount of load to be shed using the rate of change of frequency [3-4]. Another type of load shedding scheme is undervoltage load shedding.

Undervoltage load shedding is an effective measure to prevent system from voltage collapse due to large disturbance. Many researchers have conducted extensive studies about undervolatge load shedding (UVLS). A new approach of undervoltage load shedding based on energy methods to obtain a security region and stability margin was proposed in [5]. Reference [6] also presented a method for undervoltage load

shedding in power system. The analysis takes into account the generator and load dynamics. A UVLS criterion is developed using a dynamic load model. The criterion is tested on a two-bus system and a nine-bus system. Another study that used a dynamic load model to calculate the amount of load to be shed also presented in [7]. The parameters of the dynamic model were estimated online using non-linear optimisation technique. A steady-state load shedding approaches was presented in [8] to determine minimum amount of load to be shed due generation deficits and overload conditions. Another study on overload conditions also carried out by Roberto et al. [9]. Its combines load shedding and demand side management (DSM) to provide reliable power system operation under normal and emergency conditions. DSM is specifically devoted to peak demand shaving and to encourage efficient use of energy.

The research on combining evolutionary algorithms and quantum computing has shown a great performance in solving several problems. Reference [10] have presented a novel evolutionary computing method based on the concept and principles of quantum mechanics called Quantum-inspired Genetic Algorithm. Basic terminology of quantum mechanics was introduced and comparative study between the proposed algorithm and classical genetic algorithm was also presented for solving travelling salesperson problem. Han and Kim [11] presents a novel evolutionary algorithm inspired by quantum computing called quantum-inspired evolutionary programming (QEA). The algorithm is based on the concept and principles of quantum computing such as quantum bits and superposition of states. Instead of using binary, numeric or symbolic representation, QEA uses a probabilistic representation. The results show that QEA performed better than conventional genetic algorithm in solving knapsack problem. The algorithm was further improved [12] by changing the update procedure in order to prevent the case of irreversible choice and the hitchhiking phenomena.

The application of QEA for real and reactive power dispatch has been presented in [13]. The QEA algorithm determined the optimal settings of control variables, such as generator outputs, voltage magnitudes, transformer taps and the setting of shunt VAR compensation devices, which are considered in a quantum or Q-bit individuals of QEA. The results have been compared with other techniques such as ant colony system, enhanced GA and simulated annealing (SA). Another application of QEA in solving power engineering problem is

unit commitment [14]. The results show the ability of the algorithm to solve a large-scale of unit commitment problems. This study will propose a new technique namely Quantum-Inspired Evolutionary Programming (QIEP) based on multiobjective model to determine the optimal location and amount of load to be shed.

II. QUANTUM-INSPIRED EVOLUTIONARY PROGRAMMING

Quantum-Inspired Evolutionary Programming (QIEP) is combining evolutionary programming (EP) with quantum mechanics concepts such as interference and superposition. QIEP is adapted from Quantum-Inspired Evolutionary Algorithms (QEA) in [11] . It consists of three levels, quantum individuals, quantum groups and quantum universe. The overall structure of the proposed algorithm is shown in Fig. 1.
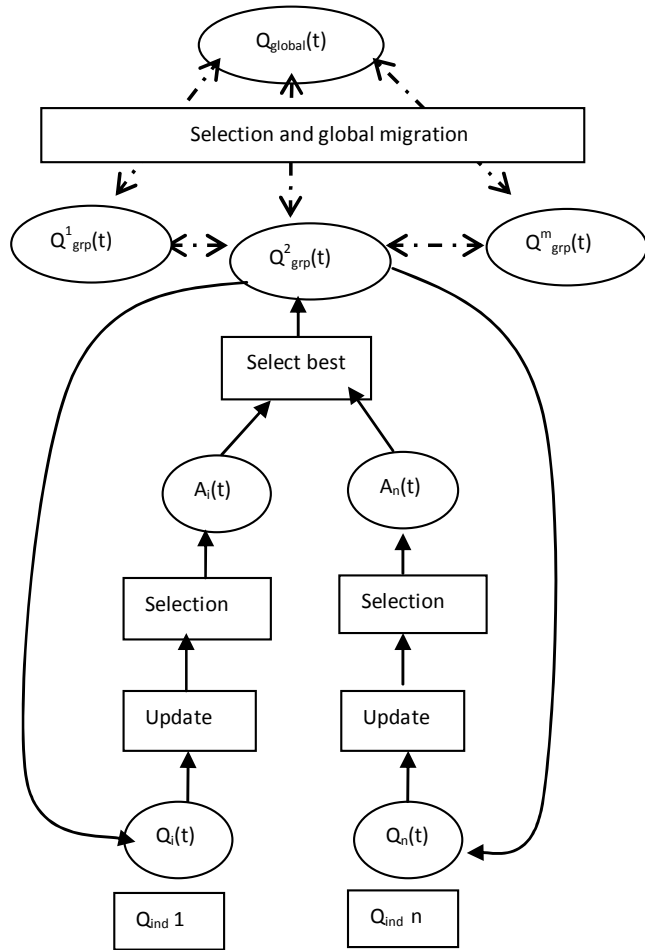
Fig. 1. Overall Structure of QIEP

The lowest level corresponds to quantum individuals (Q_{ind}). The initial value of each Q_{ind} is generate randomly and satisfy the constraint such as losses must less than initial losses and voltage profile must be in the range of 0.90 p.u. to 1.05 p.u. Then, the Q_{ind} is update through mutation process as shown in (1-3).

$$\eta'_{i,j} = \eta_{i,j} \exp\left(\tau'N(0,1) + \tau N_j(0,1)\right) \tag{1}$$

$$L'_{i,j} = L_{i,j} + \eta'_{i,j}\left(N_j(0,1)\right) \tag{2}$$

$$L'_{oi,j} = L_{oi,j} + \eta'_{i,j}\left(N_j(0,1)\right) \tag{3}$$

where:

$$\tau = \sqrt{\frac{1}{\sqrt{2n}}}$$

$$\tau' = \frac{1}{\sqrt{2n}}$$

L_i and L_{oi}, $\eta_{i,j}$ and $\eta'_{i,j}$ is i^{th} components of the respective vectors. $N(0,1)$ is a normally distribution one dimensional random number with mean 0 and 1. $N_j(0,1)$ indicates the new random number for each value of j.

The offspring population generated from mutation process are merged (combination process) with the parent. All the members were sort according to their fitness. The best population in quantum individuals will be selected to the second level *i.e.* quantum group (Q_{grp}). Quantum group consists of m Q_{grp}, each containing n Q_{ind}. The set of all n x m Q_{ind} forms the quantum universe and defines as the top level of QIEP. The best population among all Q_{grp}, noted Q_{global}, is stored every generation. The termination criterion is specified by the difference between the maximum and minimum fitness to be less than 0.001.

In this paper, bus i and amount of load to be shed at bus i (P_{shed}) are the variables to be optimised in the QIEP optimisation technique. The reactive power of the load at bus i was determined using (4) and the power factor of the system is set to be 0.85.

$$Q_{shed} = P_{shed} \; x \; tan^{-1} \; \phi \tag{4}$$

where ϕ is power factor angle and $Cos \; \phi = 0.85$,

III. PROBLEM FORMULATION

In this paper, the optimal location and amount of load to be shed are identified by new optimisation technique namely Quantum-Inspired Evolutionary Programming (QIEP). This problem has three objective functions which must be simultaneously optimised. These three components are power loss minimisation, voltage profile improvement and power interruption cost minimisation. Power losses are minimised based on (5).

$$Min \sum_{i=1}^{n} P_{loss,i} \tag{5}$$

where, n = number of lines in the system
voltage constraint, $V_{i,min} < V_i < V_{i,max}$

The second objective function in this study is voltage profile improvement. The third objective function is cost of interruption load minimisation. Cost of interruption load can be defined as the total cost of load to be shed in order to maintain the voltage at appropriate level. The cost items can be formulated as (6).

$$Min \quad f_{\cos t} = \sum_{i=1}^{n} C_T P_{shed,i} \qquad (6)$$

where f_{cost} is the cost objective function ($), n is the total number of load bus to be shed, $P_{shed,i}$ is the amount of load to be shed at bus i (MW) and C_T is the electricity market price ($/MWh). The value of C_T is $25/MWh.

In multiobjective problem, each objective function would be considered for the optimal solution. The multiobjective index (F_T) for the performance calculation of location and amount of load to be shed considers all previously mentioned objective function by strategically giving a weight. This can be performed since all the objective functions were normalised (values between 0 and 1). The normalised value (f_n) for each objective function can be calculated using (7).

$$f_{ni} = \frac{\max(f_i) - f_i}{\max(f_i) - \min(f_i)} \qquad (7)$$

where f_i is the fitness value for objective function i

In this paper, the final solution is selected based on maximum value of F_T as described in (8).

$$Max \quad F_T = \sum_{i=1}^{k} (\alpha_i \times f_{ni}) \qquad (8)$$

where $$\sum_{i=1}^{k} \alpha_i = 1$$

k is numbers of objective function.
α_i is weighting factor for i^{th} objective function

The optimisation also took into consideration the following constraints:

Load curtailment limits:
$$0 \le P_{shed,i} \le P_{L,i} \qquad (9)$$
$$0 \le Q_{shed,i} \le Q_{L,i} \qquad (10)$$

where $P_{shed,i}$ is the amount of load to be shed at bus i in MW
$P_{L,i}$ is the amount of load at bus i in MW
$Q_{shed,i}$ is the amount of load to be shed at bus i in Mvar
$Q_{L,i}$ is the amount of load at bus i in Mvar

Voltage limits:
$$V_{i,min} < V_i < V_{i,\,max} \qquad (11)$$

where V_{min} is the minimum voltage 0.90 p.u. and V_{max} is the maximum voltage 1.05 p.u.

IV. RESULTS AND DISCUSSION

The proposed algorithm was tested on IEEE 33-bus distribution test system, IEEE 69-bus distribution test system and 141-bus distribution system at Caracas, Venezuela. The initial data for IEEE 33-bus test system, IEEE 69-bus test system and 141-bus distribution system are given in [15], [16] and [17] respectively. Table 1 presents the initial data from the loadflow for the distribution system used in this study.

TABLE 1
INITIAL DATA FOR THE TEST SYSTEM

	P_{loss}(kW)	Q_{loss}(kVar)	V_{min}(p.u)
33-Bus	0.3411	0.2241	0.8565
69-Bus	0.2195	0.0999	0.9093
141-Bus	0.6133	0.4740	0.9264

Load shedding scheme was performed in the initial data of IEEE 33-bus distribution test system in order to improve the voltage profile. The results obtained for IEEE 33-bus distribution test system from the proposed technique are compared with previous technique developed in [18] as shown in Table 2. Multiobjective approaches have been presented in [18] to determine the optimal load shedding in IEEE 33-bus distribution test system. The objective of the optimisation is to minimise the sum of curtailed load and the voltage deviation index (VDI) using genetic algorithm (GA).

TABLE 2
COMPARISON OF FITNESS FUNCTION FOR IEEE 33-BUS TEST SYSTEM

Technique	Pshed (MW)	Ploss (MW)	Vmin (p.u.)	Cost ($/h)
Base case		0.3411	0.8565	
QIEP	0.9057	0.0998	0.9187	22.6430
Ref [18]	1.9315	0.1133	0.9180	48.2875

From the results obtained in Table 2, it could be observed that total power loss, amount of shed load and cost of power interruption cost for QIEP are much lower as compared to [18]. The total power loss is reduced 70.74% from the base case using QIEP technique. Whereas, in [18] the power loss is reduced 66.78% from the base case. QIEP technique also provide better performance in voltage profile than genetic algorithm optimisation technique used in [18].

The proposed study was also conducted on the IEEE 69-bus distribution test system and the 141-bus distribution system. However, since the minimum voltage of IEEE 69-distribution test system and 141-bus distribution system are greater than 0.90 p.u, the load shedding scheme only applied to both system with load increase 160%. The data for the load increase 160% are given in Table 3.

TABLE 3
INITIAL DATA FOR LOAD INCREASE

	P_{loss}(kW)	Q_{loss} (kVar)	V_{min} (p.u)
69-Bus	0.6233	0.2821	0.8452
141-Bus	1.6808	1.3002	0.8765

The performances of QIEP optimisation technique were copared with those obtain from conventional evolutionary programming (EP) optimisation technique in terms of fitness values and computation time for single objective and multiobjective function. The location of load shed is randomly optimised and it is set to five buses for case 1 and three buses for case 2 in IEEE 69-bus distribution test system. The best results for IEEE 69-bus distribution test system from 10 execution times of both techniques are tabulated in Table 4 and Table 5 for single objective optimisation and multiobjective optimisation respectively. The number of groups in Q_{grp} is set to three, the population size is 20 and the maximum number of generation is 300 for QIEP technique. The population size for EP is also set to 20 with maximum 300 generations.

TABLE 4
SINGLE OBJECTIVE OPTIMISATION OF CASE 1

Objective	Technique	Fitness	Time (s)
Ploss (kW)	QIEP	0.0607	7478
	EP	0.0636	24915
Vmin (p.u)	QIEP	0.9638	9416
	EP	0.9616	27891
Cost ($/h)	QIEP	22.97	12464
	EP	23.67	27695

TABLE 5
MULTIOBJECTIVE (MO) OPTIMISATION FOR CASE 1

		Ploss (kW)	Vmin (p.u)	Cost ($/h)	time (s)
MO	QIEP	0.0963	0.9527	54.31	13323
	EP	0.0969	0.9523	54.43	32008

From Table 4, it can be seen that best results of fitness value of the objective function for QIEP are better than EP. Although the different of fitness values between QIEP and EP are small, QIEP technique has shown a tremendous improvement in terms of computation time For multiobjective conditions in Table 5, QIEP shows better performance in terms of power losses and power interruption cost as compared to EP. EP shows a slower convergence rate than QIEP due to its single population size. The convergence time for QIEP technique is 3.7 hours which is much faster than EP optimisation technique that converge after 8.89 hours. Table 6 presents the best results obtained from single objective optimisation for case 2 within 300 generations over 10 runs. In case 2, the number of load bus to be shed is three. Table 6 shows the value of all fitness function proposed by QIEP technique are much better than those obtained from EP technique. As compared to QIEP technique in case 1, the load interruption cost for case 2 are much lower. However, the power losses and voltage profile for case 1 are better than case 2 due to bigger number of load bus to be shed. The convergence time of QIEP technique is smaller than EP technique in case 2.

TABLE 6
SINGLE OBJECTIVE OPTIMISATION OF CASE 2

Objective		Fitness	Time (s)
Ploss (kW)	QIEP	0.0776	13872
	EP	0.0836	30222
Vmin (p.u)	QIEP	0.9542	13391
	EP	0.9541	22081
Cost ($/h)	QIEP	23.51	12705
	EP	23.52	28592

TABLE 7
MULTIOBJECTIVE (MO) OPTIMISATION FOR CASE 1

	Ploss (kW)	Vmin (p.u)	Cost ($/h)	time (s)
QIEP	0.1030	0.9501	52.68	13136
EP	0.1063	0.9504	51.91	29634

The best results from 10 times simulation for multiobjective optimisation is tabulated in Table 7. Similar pattern can be seen in Table 7 where power interruption cost in case 2 is better than case 1 and power losses and voltage profile for case 1 are better than case 2.

Similar analysis also carried out using 141-bus distribution system to verify that the proposed technique also applicable to bigger system. The network data for 141-bus distribution system is presented in Appendix A. Two cases are considered. The number of buses to be shed is ten for case 3 and seven for case 4. The best results of single objective optimisation and multiobjective optimisation for case 3 are presented in Table 8 and Table 9 respectively.

TABLE 8
SINGLE OBJECTIVE OPTIMISATION OF CASE 3

Objective		Fitness	Time (s)
Ploss (kW)	QIEP	0.6666	12859
	EP	0.6985	22173
Vmin (p.u)	QIEP	0.9303	7757
	EP	0.9301	17226
Cost ($/h)	QIEP	45.01	12346
	EP	48.69	23204

TABLE 9
MULTIOBJECTIVE (MO) OPTIMISATION FOR CASE 3

	Ploss (kW)	Vmin (p.u)	Cost ($/h)	time (s)
QIEP	0.8953	0.9207	99.65	11843
EP	0.9044	0.9203	99.09	17713

Table 8 revealed that the maximum impact on the improvement of power losses, voltage profile and load interruption cost are also given by QIEP technique with less computation time. The computation time for single objective optimisation using QIEP technique is less than 3.6 hours, whereas, EP technique takes more than 6 hours to converge. According to the results obtained in Table 9, multiobjective QIEP also provide better results in terms of fitness value and

computation time although in a bigger system. The power loss is reduced by 46.74% from the pre-optimisation value which measured before implementing load shedding. The minimum voltage in the system improved to 0.9207 p.u. With the reduction of the number of bus to be shed, the cost of interruption load is reduced. The same concept of obtaining result in Table 8 and Table 9 are used to tabulate Table 10 and Table 11. However, this case the number of load bus to be shed is reduced to seven buses.

TABLE 10
SINGLE OBJECTIVE OPTIMISATION OF CASE 4

Objective	Technique	Fitness	Time (s)
Ploss (kW)	QIEP	0.7958	8903
	EP	0.8082	18519
Vmin (p.u)	QIEP	0.9231	12222
	EP	0.9205	21210
Cost ($/h)	QIEP	43.20	9967
	EP	44.12	22498

TABLE 11
MULTIOBJECTIVE (MO) OPTIMISATION FOR CASE 4

	Ploss (kW)	Vmin (p.u)	Cost ($/h)	time (s)
QIEP	0.9571	0.9207	85.80	10178
EP	0.9904	0.9148	87.02	23352

Based on the results tabulated in Table 10 and Table 11, the results achieved by QIEP with multiobjective function has the most significant results on the reduction of power losses, improvement of voltage profile and reduction of interruption cost. For multiobjective optimisation, the power loss is reduced 49.57% from the pre-optimisation value with 160% load increase. This technique also capable to improve the voltage profile to 0.9207p.u with lower cost as compared to EP technique. The computation time for QIEP technique is much faster than EP technique for all objective function.

V. CONCLUSION

This paper proposes an efficient method for optimal location and sizing of DG in radial distribution system namely Quantum-Inspired Evolutionary Programming (QIEP). QIEP is an algorithm that employs the concept of quantum mechanics in the Evolutionary Programming (EP). Quantum-Inspired is implemented according to three levels defined by quantum individuals, quantum groups and quantum universe. The problem has been formulated as the multiobjective function considering power loss minimisation, voltage profile improvement and interruption cost minimisation. The results of multiobjective QIEP optimisation technique were compared with genetic algorithm optimisation technique developed previously in [18]. A comparative study was also conducted to

those obtained by conventional EP optimisation technique in term of fitness value and computation time. The proposed method was tested on three distribution system consisting of IEEE 33-bus distribution system, IEEE 69-bus distribution test system and 141-bus real distribution system. The result shows better improvement in term of fitness function and computation time.

REFERENCES

[1] Y. Wang, *et al.*, "Strategy to Minimise the Load Shedding Amount for Voltage Collapse Prevention," *IET on Generation, Transmission & Distribution* vol. 5, pp. 307-313, 2011.

[2] A. R. Malekpour and A. R. Seifi, "An Optimal Load Shedding for DIstribution Networks with DGs Considering Capacity Deficiency Modelling of Bulked Power Supply," *Modern Applied Journal (CCSE),* vol. 3, pp. 143-151, May 2009.

[3] P. Mahat, *et al.*, "Underfrequency Load Shedding for an Islanded Distribution System With Distributed Generators," *IEEE Trans on Power Delivery,* vol. 25, April 2010.

[4] Y.-Y. Hong and P.-H. Chen, "Genetic-Based Underfrequency Load Shedding in a Stand-Alone Power System Considering Fuzzy Loads," *IEEE Trans on Power Delivery,* vol. 27, pp. 87-95, Jan 2012.

[5] A. Nanda and M. L. Crow, "An energy based approach to undervoltage load shedding," *Electric Power Systems Research,* vol. 32, pp. 11-18, 1995.

[6] S. Arnborg, *et al.*, "On undervoltage load shedding in power systems," *Int Journal of Electrical Power & Energy Systems,* vol. 19, pp. 141-149, 1997.

[7] R. Balanathan, *et al.*, "Undervoltage Load Shedding to Avoid Voltage Instability," *IEE Proc. on Gen., Trans. and Distr.,* vol. 145, pp. 175-181, 1998.

[8] M. A. Mostafa, *et al.*, "A computational comparison of steady state load shedding approaches in electric power systems," *Power Systems, IEEE Transactions on,* vol. 12, pp. 30-37, 1997.

[9] R. Faranda, A. Pievatolo and E. Tironi, "Load Shedding: A New Proposal," *IEEE Trans. on Power Systems,* vol. 22, pp. 2086-2093, 2007.

[10] A. Narayanan and M. Moore, "Quantum-inspired genetic algorithms," *Proc. of IEEE Int. Conf. on Evolutionary Computation,* pp. 61-66, 1996.

[11] K.-H. Han and J.-H. Kim, "Quantum-Inspired Evolutionary Algorithm for a Class of Combinatorial Optimization," *IEEE Trans. on Evolutionary Computation,* vol. 6, Dec. 2002.

[12] M. D. Platel, S. Schliebs and N. Kasabov, "A Versatile Quantum-Inspired Evolutionary Algorithm," *IEEE Congress on Evolutionary Computation,* pp. 423-430, 2007.

[13] J. G. Vlachogiannis and K. Y. Lee, "Quantum-Inspired Evolutionary Algorithm for Real and Reactive Power Dispatch," *IEEE Trans. on Power Systems,* vol. 23, pp. 1627-1636, 2008.

[14] T. W. Lau, *et al.*, "Quantum-Inspired Evolutionary Algorithm Approach for Unit Commitment," *IEEE Trans. on Power Systems,* vol. 24, pp. 1503-1512, 2009.

[15] M. E. Baran and F. F. Wu, "Network Reconfiguration in Distribution Systems for Loss Reduction and Load Balancing," *IEEE Trans. on Power Delivery,* vol. 4, pp. 1401-1407, 1989.

[16] M. E. Baran and F. F. Wu, "Optimal Capacitor Placement on Radial Distribution Systems," *IEEE Trans. on Power Delivery,* vol. 4, Jan 1989.

[17] H.M. Khodr, *et al.*, "Optimal Cost-Benefit for the Location of Capacitors in Radial Distribution Systems," *IEEE Trans. on Power Delivery,* vol. 24, April 2009.

[18] M. H. Moradi and M. Abedini, "Optimal load shedding approach in distribution systems for improved voltage stability," *2010 4th Int Power Engineering and Optimization Conf. (PEOCO), ,* pp. 196-200, 23-24 June 2010.

APPENDIX A

Network data for 141-bus distribution systems

i	j	R	X	kVA	i	j	R	X	kVA	i	j	R	X	kVA	i	j	R	X	kVA
2	2	0.0577	0.0409	0	70	72	0.07	0.0495	150	29	30	0.0342	0.0248	0	30	100	0.0033	0.0008	300
3	3	0.1725	0.1223	0	42	73	0.0231	0.0164	300	30	31	0.0128	0.0091	0	91	101	0.0231	0.0164	15
4	4	0.0009	0.0006	0	73	74	0.003	0.0064	300	31	32	0.0347	0.0245	150	101	102	0.0578	0.0409	0
5	5	0.0092	0.0065	0	43	75	0.0379	0.0268	45	2	33	0.0443	0.0314	0	102	103	0.0889	0.0217	125
6	6	0.0068	0.0049	0	44	76	0.0552	0.0391	75	33	34	0.002	0.0009	150	103	104	0.0629	0.0153	0
7	7	0.0469	0.0625	0	46	77	0.0516	0.0436	150	5	36	0.2274	0.0554	300	104	36	0.117	0.0285	300
8	8	0.0736	0.0981	75	76	78	0.0167	0.011	0	5	2	0.1265	0.1565	150	104	2	0.0114	0.0026	150
9	9	0.0649	0.0459	10	78	79	0.0415	0.0101	502	6	3	0.0055	0.0073	50	92	3	0.0849	0.0207	502.5
10	10	0.0507	0.0359	0	79	80	0.1003	0.0244	750	37	4	0.2036	0.144	0	94	4	0.0612	0.026	0
11	11	0.0116	0.0082	0	79	81	0.1513	0.037	0	38	5	0.0938	0.0663	20	108	5	0.0452	0.0192	750
12	12	0.1291	0.0913	25	81	82	0.0033	0.0008	150	39	6	0.0347	0.0245	0	94	6	0.0033	0.0008	750
13	13	0.1227	0.0866	75	47	83	0.0085	0.0062	75	40	7	0.0918	0.065	75	7	7	0.0719	0.0509	25
14	14	0.0488	0.0345	0	49	84	0.0517	0.0449	225	41	8	0.2318	0.164	0	10	8	0.107	0.0261	500
15	15	0.0957	0.0677	0	50	85	0.0147	0.0036	0	42	9	0.1207	0.0854	0	11	9	0.0347	0.0245	75
16	16	0.086	0.0609	0	85	86	0.0037	0.0016	500	43	10	0.0443	0.0314	50	13	10	0.0623	0.0441	0
17	17	0.0398	0.0282	150	86	87	0	0	150	44	11	0.0405	0.0288	0	114	11	0.0668	0.0473	0
18	10	0.0828	0.0566	0	7	88	0.0174	0.0231	75	45	12	0.016	0.0127	0	114	12	0.004	0.001	300
16	19	0.0186	0.0132	0	88	89	0.0469	0.0625	65	46	13	0.0636	0.045	0	14	13	0.0506	0.0366	65
19	20	0.0559	0.0395	75	89	90	0.0299	0.0398	0	47	14	0.0417	0.0295	125	15	14	0.0161	0.0114	0
20	21	0.0365	0.0246	75	90	91	0.0212	0.0283	0	48	15	0.0732	0.051	150	118	15	0.0462	0.0327	110
21	22	0.0573	0.0307	0	91	92	0.0315	0.042	0	49	16	0.0828	0.0556	0	119	16	0.0424	0.03	0
22	23	0.0263	0.0191	75	92	93	0.028	0.0373	0	50	17	0.0398	0.0282	125	120	17	0.0507	0.0359	0
23	24	0.0683	0.0497	0	93	94	0.0206	0.0274	110	51	18	0.0225	0.0159	75	121	18	0.0732	0.0518	0
24	25	0.0398	0.0282	0	94	95	0.0206	0.0274	0	38	19	0.0841	0.0595	100	122	19	0.0584	0.0414	100
27	26	0.0729	0.053	150	89	96	0.0687	0.0486	150	42	20	0.0161	0.0114	0	19	20	0.061	0.0432	125
26	27	0.0335	0.0244	75	96	97	0.097	0.0686	0	54	21	0.0527	0.0373	0	20	21	0.0783	0.0554	0
27	23	0.0584	0.0414	0	97	98	0.0902	0.0196	300	55	22	0.0893	0.0632	25	21	22	0.0834	0.0607	0
28	29	0.0655	0.0463	75	97	99	0.0033	0.0008	0	56	23	0.0867	0.0613	0	22	23	0.0347	0.0245	75
61	62	0.0411	0.0291	200	131	132	0.0347	0.0245	75	57	24	0.0674	0.0477	300	23	24	0.057	0.042	75
60	63	0.0353	0.025	0	131	133	0.092	0.0669	45	58	25	0.0469	0.0332	150	24	25	0.0585	0.0425	110
63	64	0.1047	0.0741	300	121	134	0.0841	0.0612	35	55	26	0.0334	0.0236	0	25	26	0.0103	0.0073	112.5
64	65	0.0674	0.0477	150	16	135	527	0.0373	25	60	27	0.0327	0.0232	300	119	27	0.0355	0.0253	0
65	66	0.0302	0.0214	225	16	136	0.0302	0.0214	75	63	69	0.0366	0.0259	300	25	139	0.095	0.0673	50
66	67	0.0456	0.0323	50	16	137	0.0584	0.0414	55	66	70	0.0231	0.0164	0	30	140	0.0519	0.0377	150
67	63	0.0218	0.0154	100	23	138	0.0769	0.0559	50	70	71	0.012	0.0029	300	31	141	0.0584	0.0414	75

Power Consumption Saving During Charging Period for Thermal Energy Storage System

A.H. Kassim

Faculty of Electrical Engineering
Universiti Teknologi Mara
Shah Alam, Selangor Darul Ehsan
abd.hafiz@yahoo.com

Z.M. Zain

Faculty of Electrical Engineering
Universiti Teknologi Mara
Shah Alam, Selangor Darul Ehsan

Abstract- **Thermal Energy Storage (TES) system is already well established and one of the most important methods of demand to reduce peak-cooling loads. This system is divided by two periods known as charging period and discharging period. Discharging period is starting from 7 a.m until 10 p.m. Charging period to produce ice is 10 hours from 10 p.m until 7 a.m with capacity rated 24 refrigerant ton (RT) of cooling capacity for each tank. Nowadays, the issue of cost in term of power consumption becomes an obstacle to implement this system. This paper presents the power consumption saving during charging period for TES system in Engineering Complex, UiTM Shah Alam Selangor. The data obtained were observed and analyzed in order to saving the power consumption. In conclusion, there are some calculations that including with Tenaga Nasional Berhad (TNB) tariff to show the comparison in term of power consumption during charging period and electricity bill for a month between actual design, current design and propose design.**

Keywords- Thermal energy storage, charging period, discharging period, refrigerant ton, cooling capacity, peak-cooling loads

I. INTRODUCTION

Thermal energy storage (TES) for cooling system, also known as cool storage, chill storage, or cool thermal storage, is a relatively mature technology that developed to reduce the electricity charge by charging ice during off-peak hours when electricity rates are lower. TES system has a potential to shift load from peak to off-peak utility periods [1-4].

In practice, many research and development activities related to energy have been concentrated on the efficient energy use and energy savings, leading to energy-conservation. In this regard, TES appears to be one of the most attractive thermal applications and exergy as the best tool in analyzing their performances [5, 6].

Previous study shows that some authors have been discussed about the power consumption. For example, in Taiwan, more than 30% power consumption for air conditioning system are use [7] and remain a major cause of

the increasing in peak load. Refer to the B.A. Habeebullah [8], economic feasibility of thermal energy storage systems state that, air conditioning (A/C) systems are major contributors to buildings' energy consumption especially in hot climate locations about 60% of the energy is consumed for A/C.

During charging period, the chiller will produce the ice or chilled water and stores it in the insulated storage tanks. During the day, the TES system can then either assist or replace the chiller for meeting the day-time cooling load to the buildings to reduce the cost of chiller, especially during peak hours which the amount of on-peak energy used and peak demand charges can be saved. In another word, TES system is the solution of the latest technology for solving peak load problems [9-11].

As TES systems are truly system optimizers for existing building chiller systems and not stand-alone units connected directly to the electric load, comparing their efficiency directly to other energy storage technologies is difficult. On average, TES systems lose through thermal loss only approximately 1% of their energy during inactive use. This efficiency is comparable to or much higher than other technologies.

The objectives of this paper are to save the power consumption during charging period and reducing the electricity bill in Engineering Complex UiTM Shah Alam.

A. Background of TES System in Engineering Complex UiTM Shah Alam

TES system in Engineering Complex UiTM Shah Alam was operated starting year 2002 and among the earliest uses of this system in Malaysia. The application of this system is covering the entire Engineering Complex of Block 1, Block 2, Block 3, Block 4, Block 5, Tower 1 and Tower 2. TES system is supported by 4 main sub-system called chiller, ice cell, heat exchanger and cooling tower.
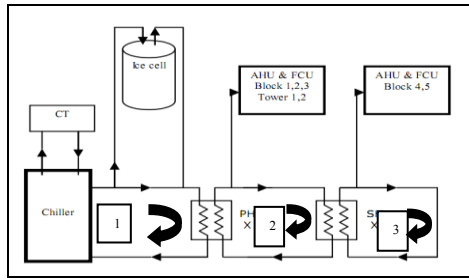
Figure. 1   Operation schematic diagram of the plant

Fig. 1 above shows the diagram of the plant where AHU is air handling unit, FCU is fan coil unit, CT is cooling tower, PHX is primary heat exchanger and SHX is secondary heat exchanger. The plant at the Engineering Complex consists of 3 closed pumping loops namely primary, secondary tertiary loops with the TES system in the upper stream.

In the Engineering Complex, TES system is operated by using partial storage system. Partial storage system is divided by two strategies namely load-levelling strategy and demand-limiting strategy. Strategy that used in Engineering Complex UiTM Shah Alam is load-levelling strategy as shown in Fig. 2. Load-levelling is most effective where the peak cooling load is much higher than the average load.
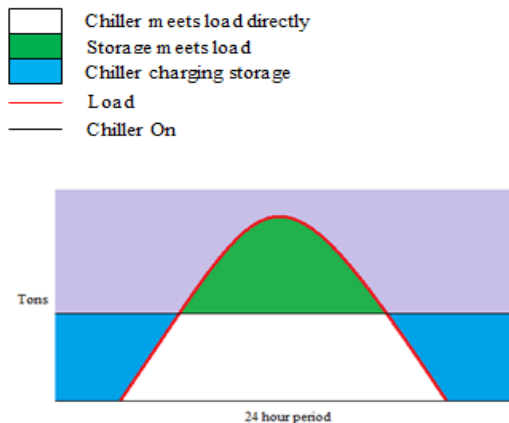
Figure. 2   Partial storage; load-levelling

By using Dunham Bush Ice Cell Technology, process of producing of ice is occurred in the ice cell. Total of the ice cell in the Engineering Complex is 73 tanks, with 28 tanks is located at the ground of car park area and remain 45 tanks is located underground at car park area as shown in Fig. 3 and Fig. 4 respectively. Currently TES system in Engineering Complex is using 45 tanks that located underground at car park area because remain of 28 tanks is not ready to use.

Figure. 3   Ice storage tank at the outdoor (Parking lot)

Figure. 4   Ice storage tank at the indoor (Underground)

Based on Operation & Maintenance Manual And As-Built Drawings Volume 1-3 from Bahagian Pembangunan Universiti Teknologi Mara Shah Alam stated that charging period to produce ice is 10 hours with capacity rated 24 refrigerant ton (RT) of cooling capacity for each tank. Time for charging period is starting from 9.00 pm until 7.00 am when the cooling requirement is almost zero as the occupancy rate is very minimal while discharging period is happen during day time from 7.00 am until 9.00 pm when the occupants start to enter the building as shown in Table I.

TABLE 1.   CHARGING AND DISCHARGING PERIOD

MODE	TIME
Charging period	9.00 p.m. until 7.00 a.m.
Discharging period	7.00 a.m. until 9.00 p.m.

However, to avoid from on-peak tariff from Tenaga Nasional Berhad (TNB), charging and discharging period need to be reschedule. Table II below is the current charging and discharging period for TES system in Engineering Complex.

TABLE II.   CHARGING AND DISCHARGING PERIOD

MODE	TIME
Charging period	10.00 p.m. until 7.00 a.m.
Discharging period	7.00 a.m. until 10.00 p.m.

B. *Electricity Bill by Tenaga Nasional Berhad (TNB)*

Electricity bill during charging period need to be refer TNB pricing and tariff. Table III below shows the commercial tariff (tariff C2) that effective on 1 June 2011. In this tariff, TNB stated that TES system consumers will enjoy higher discount for off-peak electricity consumption. In addition, Firm Standby charge for Co-generators will be reduced to encourage Green Technology, Energy Efficiency and Demand Side Management (DSM) initiatives [12].

TABLE III.   COMMERCIAL TARIFF (TARIFF C2)

Tariff C2 – Medium Voltage Peak/Off-Peak Commercial Tariff		
For each kilowatt of maximum demand per month during the peak period	RM/kW	38.60
For all kWh during the peak period	sen/kWh	31.2
For all kWh during off-peak period	sen/kWh	19.2
The minimum monthly charge is RM600		

II.   METHODOLOGY

Based on actual design in Engineering Complex UiTM Shah Alam, charging period required 10 hours to fully freeze the water in the tanks. From specification of chiller RT during charging period that state in Operation & Maintenance Manual And As-Built Drawings Volume 1-3 from Bahagian Pembangunan Universiti Teknologi Mara Shah Alam, chiller will support 1250 RT to produce the ice in ice cell tank.

Table IV below, shows the Engineering Complex district cooling design. During discharging and charging period, TES system require 2 chillers to be operate which is 1800 RT for discharging and 1250 RT for charging period. Total ice storage that install is 45 tanks and capacity for each tank is 24 RT. Maximum cooling load for Engineering Complex is 3600 RT per hour.

TABLE IV.   DISTRICT COOLING DESIGN

Chiller Installed	2 x 1800 RT Base Mode 2 x 1250 RT Ice Mode
Ice storage Capacity	45 x 240 RTh Ice Cell
Building Cooling Load Requirement	Block 1 550 RT Block 2 580 RT Block 3 380 RT Block 4 850 RT Block 5 660 RT Tower 1 290 RT Tower 2 290 RT
Maximum Cooling Load	3600 RT

All of the equations below are the calculation for this paper where it is used to compare between actual, current and proposed designs. It is also involved time for charging and discharging period.

$$H_{opt} = \frac{RT_{req} * H_{act}}{CH_{support}} \qquad (1)$$

Where H_{opt} is the optimize of hours for charging period, RT_{req} is RT that required to freeze the water for 1 hour , H_{act} is actual hours to fully freeze the water in the tank of ice cell and $CH_{support}$ is actual chiller support to produce the ice in ice cell tank.

From the equation (1) above, optimum hours that we get is 8.64 hours (8 hour and 38.4 minute). So, we assume proposed hours for charging period is 8.5 hours (8 hour and 30minute).

Maximum building load equation;

$$RT_b = TotCL_{RT} * 0.8 \qquad (2)$$

$$MaxRT_{tcd} = RT_b * t_{discharging} * 0.75 \qquad (3)$$

Where RT_b is the RT of the building load, $TotCL_{RT}$ is total cooling load in building, $MaxRT_{tcd}$ is RT for maximum total cooling demand and $t_{discharging}$ is total time for discharging period in working hours.

Maximum cooling supply by chiller and ice;

$$MaxRT_{exp} = [(CH_{RT} * 2) * t_{discharging}] + [t_{charging} * IC_{t1}] \qquad (4)$$

Where $MaxRT_{exp}$ is the maximum expected cooling supply to building, CH_{RT} is the cooling capacity of chiller in RT, $t_{discharging}$ is the total time for discharging period in working hours, $t_{charging}$ is the total time for charging period and IC_{t1} is an ice cell storage capacity in 1 hour.

Maximum cooling remains after 10 hours;

$$Max_{cr} = MaxRT_{exp} - MaxRT_{tcd} \qquad (5)$$

Max_{cr} is the maximum cooling remains.

From equation (4), calculation of cooling supply by chiller and ice are expand by depending on the percentage of opening valve of chiller (V_1) and ice cell (V_2). The equation below is the expanded from equation (4).

$$RT_{exp} = [\{\% \text{ of chiller valve opens } V_1 \text{ in } t_1 \text{ hour} * CH_{RT} * 2\}$$
$$+ \{\% \text{ of chiller valve opens } V_1 \text{ in } t_2 \text{ hour} * CH_{RT} * 2\} + ... +$$
$$\{\% \text{ of chiller valve opens } V_1 \text{ in } t_n \text{ hour} * CH_{RT} * 2\}] +$$
$$[t_{charging} * IC_{t1}] \qquad (6)$$

From the equation above, comparison the value of RT_{exp} between actual design, current design and optimum design could done.

III. RESULTS AND DISCUSSION

This paper shows the comparison between actual design, current design and proposed design of power consumption during charging period.

From equation (2) and (3) will get the total cooling demand in building (Table V). Total maximum cooling demand for the Engineering Complex after considers all of the losses with 10 hours discharging period is 21600 RT.

TABLE V. TOTAL COOLING DEMAND IN BUILDING

$TotCL_{RT}$	RT_b	$t_{discharging}$	$MaxRT_{tcd}$
3600 RT	2880 RT	10 hours	21600 RT

Table VI shows the comparison between actual design, current design and proposed design for maximum cooling supply by chiller and ice that refer to the equation (4). From that table shows that actual design can be supply 28800 RT cooling to the building. For other designs that are current and proposed design, it can be supply 27720 RT and 27180 RT cooling to the building respectively.

TABLE VI. COMPARISON BETWEEN ACTUAL DESIGN, CURRENT DESIGN AND PROPOSED DESIGN FOR MAXIMUM COOLING SUPPLY BY CHILLER AND ICE

	Actual Design	Current Design	Proposed Design
CH_{RT}	900 RT	900 RT	900 RT
$t_{discharging}$	10 hours	10 hours	10 hours
$t_{charging}$	10 hours	9 hours	8.5 hours
IC_{t1}	1080 RT	1080 RT	1080 RT
$MaxRT_{exp}$	28800 RT	27720 RT	27180 RT

Table VII below shows the result of the cooling remains of actual design, current design and proposed design using equation (5). Cooling remains after 10 hours for proposed design is 5580 RT which is less than actual and current design. Cooling remains for actual design is 7200 RT and current design is 6120 RT.

TABLE VII. COOLING REMAINS BETWEEN ACTUAL DESIGN, CURRENT DESIGN AND PROPOSED DESIGN

	Actual Design	Current Design	Proposed Design
Max_{cr}	7200 RT	6120 RT	5580 RT

From the equation (6), we compare the value of RT_{exp} between actual design, current design and optimum design (Table VIII).

TABLE VIII. VALUE OF THE RT_{EXP} FOR ACTUAL DESIGN, CURRENT DESIGN AND OPTIMUM DESIGN BY CALCULATION

	Actual Design	Current Design	Optimum Design
RT_{exp}	17100 RT	16020 RT	15480 RT

The value of maximum cooling remain in proposed design is 5580 RT. In Malaysia, working per day is from 8 am until 5 pm (9 hours). That's means, proposed hours for charging period (8 hours and 30 minutes) is very sufficient to supply cooling to the building load after considered total cooling demand in the building. Cooling remain should be release in the building before start charging period.

Both of the figures below (Fig. 5 and Fig. 6) demonstrate the monthly cooling load and average cooling load in 7 months respectively from April 2011 until October 2011. From the graph, we can deduce that the maximum cooling load is at 9 a.m. and the average maximum cooling load is from 8 a.m. to 4 p.m. This may be caused by too many consumers at that time, including students and staffs. We can see where cooling load starts decreasing after 4 pm and will continue the descent.
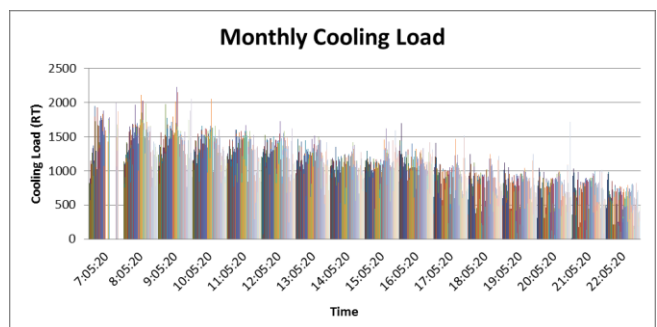
Figure 5. Typical monthly cooling load profile of a building
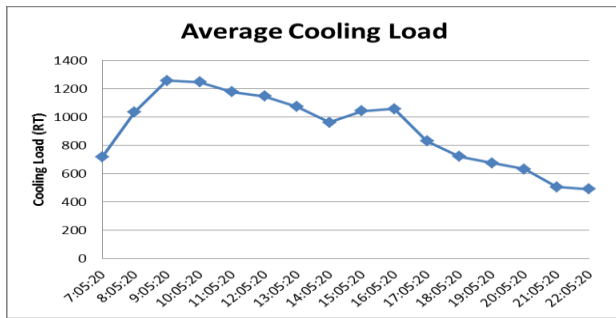
Figure 6.   Average cooling load profile of a building

IV.   CONCLUSION

In this paper, we have proved that power consumption can save by reschedule the charging period from 9 hours charging to 8.5 hours charging. With shorten the charging period, remainder 30 minutes can saved power consumed by the chiller. In addition, chiller also can get more rest compared with 10 hours and 9 hours charging.

Charging period by using proposed design can fulfill the need to the building in 10 hours because maximum cooling supply by chiller and ice is greater than maximum cooling remain ($MaxRT_{exp} > Max_{cr}$).

However, in the theoretical method, ice will be fully exhausted depend on the opening valve of ice cell. Then, chiller will be only supply the chill water to the building for remain hours that ice can't supply. Cooling supply from ice cell for proposed design is quicker exhausted compare than actual design and current design. However, due to the environment temperature after 6 p.m. and decreasing the number of consumer in the building, cooling that supply from chiller and remain cooling from ice cell in the pipe is believed to provide cooling to the building.

V.   RECOMMENDATION

Thermal energy storage (TES) system is needed to manage the power consumption in order to reduce the cost of electricity consumption. Maintenance must be done continuously to ensure that it can be utilized optimally.

VII.   REFERENCES

[1]   Opportunities in thermal storage R&D, *EPRI Rep. EM-3159-SR*, Electric Power Research Institute, Palo Alto, CA, USA, 1983.

[2]   Commercial cool storage design guide, *EPRI Rep. EM-3981*, Electric Power Research Institute, Palo Alto, CA, USA, 1985.

[3]   Commercial cool storage presentation material, *EPRI Rep. EM-4405*, Electric Power Research Institute, Palo Alto, CA, USA, 1986.

[4]   Cool storage marketing guidebook, *EPRI Report EM-5841*, Electric Power Research Institute, 1988.

[5]   M.A. Rosen, Appropriate thermodynamic performance measures for closed system for thermal energy storage, ASME Journal of Solar Energy Engineering 144 (1992) 100-105.

[6]   I.Dincer, S. Dost, X. Li, Performance analyses of sensible heat storage system for thermal application, International Journal of Energy Research 21 (10) (1997) 1157-1171.

[7]   Huei-Jiunn Chen, Wang David WP, Sih-Li Chen. Optimization of an ice-storage air-conditioning system using dynamic programming method. Appl Therm Eng 2005:461–72.

[8]   B.A, H., Economic feasibility of thermal energy storage systems. Energy and Buildings, 2007. **39**(3): p. 355-363.

[9]   T. Lehman, D.D. Jones, D.R. Vogel, Off-peak HVAC is once again hot, Journal of Consulting Specifying Engineering (November) (2001) 46 49.

[10] ASHRAE 1995. Thermal Energy Storage, Handbook of Applications, American society of heating ventilation and air conditioning engineers, Atlanta, Georgia.

[11] G. Zhou, M. Krarti, G.P. Henze, Parametric analysis of active and passive building thermal storage utilization, ASME Journal of Solar Engineering 127 (2005) 37–64.

[12] http://www.tnb.com.my/residential/pricing-and-tariff.html

Extended Kalman Filter (EKF) Prediction of Flood Water Level

Ramli Adnan [#1], Fazlina Ahmat Ruslan [#2], Abd Manan Samad [*3], Zainazlan Md Zain [#4]

[#]Faculty of Electrical Engineering
[*]Dept. of Surveying Science and Geomatics, Faculty of Arc., Planning and Surveying
Universiti Teknologi MARA 40450, Shah Alam, Selangor, Malaysia
[1]ramli324@salam.uitm.edu.my
[2]fazlina419@salam.uitm.edu.my

Abstract— **This paper addresses Extended Kalman Filter (EKF) algorithm that is uses to predict and estimate flood water level. In this respect, good estimates of the flood water level are needed to enable the filter to generate accurate forecasts. The EKF is the best predictor of the flood water level as it is the extended of the basic Kalman Filter algorithm that is only able to solve linear problems. EKF is developed to solve nonlinear problems and flood phenomenon suite well as the water level fluctuates highly nonlinear. This theory is also supported with the simulation results that produce small value of Root Mean Square Error (RMSE) which is close to zero.**

Keywords — *Kalman Filter; Extended Kalman Filter (EKF); Tracking and Prediction*

I. INTRODUCTION

Over the past decade, rapid increases in development and deforestation have led to flood inundation problems. The key factor that contributes to flood problems is the increasing of the water level. So, the prediction of flood water level is very important as an early warning system to the public. For the purpose of flood water level prediction, parameter updating methods are applied to update the parameters of simulation models based on observed errors. For example, Lees et al. [1] applied single-input single-output transfer function in downstream water level prediction by using upstream information along the River Nith at Dumfries. Parameter prediction algorithm also was proposed to estimate real-time downstream observations [2].

The state-parameter estimation using standard Kalman Filter algorithm was introduced in [3] and later, state augmentation technique in [4]. However, both techniques were limited to linear dynamics system. Since many problems are nonlinear cases, the modified Kalman Filter was introduced. The most straight forward extension of standard Kalman Filter is Extended Kalman Filter (EKF). EKF has been implemented as a state observer for induction motor drives in various configurations [5, 6, 7].

For nonlinear dynamic system, Extended Kalman filter is the best predictor and it was used to track manoeuvring targets in [8]. This manoeuvring target do not have straight motion and constant velocity, thus causes the process equation to be nonlinear. Same goes to flood water level because it fluctuates highly nonlinear. Therefore, Extended Kalman Filter was used in this study.

This paper was organized in the following manner: Section II describes the theory of EKF algorithm; Section III describes methodology; Section IV is on results and discussion; and finally, Section V is the conclusions.

II. KALMAN FILTER

A. Theory of Discrete Kalman Filter

In 1960, Kalman successfully solved filtering problem using state-space approach [9]. The solution is now widely known as Kalman Filter (KF). Two main features in KF: the vector modelling of random processes; and the process noise must be Gaussian [10]. However, the general KF algorithm is limited for linear systems. Therefore, EKF is considered as an extension of KF in solving nonlinear problems. As EKF is the extension of KF, the basic theory of KF algorithm will be discussed first.

KF used feedback control to estimate a process. First, the filter estimates the process state and then obtains the feedback in the form of noisy measurements. Because of that, the equation for the KF falls into two types: time update equations; and measurement update equations. The time update equations are responsible to obtain the priori estimates for the next step by projecting the current state and error covariance estimates, whereas, the measurement update equations are responsible for the feedback. Indirectly, the time update equations act as the predictor equations while the measurement update equations act as the corrector equations. Indeed, the KF estimation algorithm resembles the predictor-corrector algorithm as shown in Figure 1.
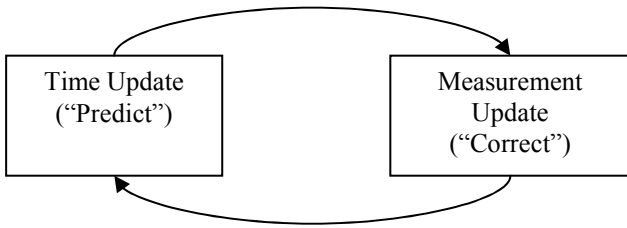
Figure 1. The ongoing discrete Kalman Filter cycle

B. Theory of Extended Kalman Filter (EKF)

It is known that the KF is able to estimate the state of a discrete-time controlled process that is governed by a linear stochastic differential equation. When a process to be estimated is nonlinear, KF is not able to do the estimates. Then the KF extension is introduced that is able to linearize the current mean and covariance. The extension KF is referred to as Extended Kalman Filter (EKF).

The EKF algorithm has two main stages: prediction step; and update (filtering) step. In prediction steps, previous state estimate of the system is being used and in the update step, the predicted state is corrected by using feedback correction step. The feedback correction step contains the weight of the measured and estimated output signals. In EKF, by calculating the stochastic properties of the noise, the initial values of the matrices can be arranged correctly. Furthermore, a critical part of EKF is to apply correct initial values for various covariance matrices. Covariance matrices have an important effects on converge time and filter stability. The computational origin of the EKF is explained below.

$$x_k \approx \widetilde{x}_k + A(x_{k-1} - \hat{x}_{k-1}) + W_{w(k-1)}$$
(1)

$$z_k \approx \widetilde{z}_k + H(x_k - \widetilde{x}) + V_{v(k)}$$
(2)

where:

- $x_k \ \& \ z_k \ \rightarrow$ actual state and measurement vectors
- $\widetilde{x}_k \ \& \ \widetilde{z}_k \ \rightarrow$ approximate state and measurement vectors from (1) and (2)
- $\hat{x}_k \qquad \rightarrow$ a posteriori estimate of the state at step k
- $w_k \ \& \ v_k \ \rightarrow$ process and measurement noise
- A $\ \rightarrow$ Jacobian matrix of partial derivatives of f with respect to x
- W $\ \rightarrow$ Jacobian matrix of partial derivatives of f with respect to w

- H $\ \rightarrow$ Jacobian matrix of partial derivatives of h with respect to x
- V $\ \rightarrow$ Jacobian matrix of partial derivatives of h with respect to v

Prediction error:

$$\widetilde{e}_{xk} \equiv x_k - \widetilde{x}_k$$
(3)

Measurement Residual:

$$\widetilde{e}_{zk} \equiv z_k - \widetilde{z}_k$$
(4)

Using Eq. (3) and Eq. (4), the equation for process error is:

$$\widetilde{e}_{xk} \approx A(x_{k-1} - \hat{x}_{k-1}) + \varepsilon_k$$
(5)

$$\widetilde{e}_{zk} \approx H\widetilde{e}_{xk} + \eta_k$$
(6)

where ε_k and η_k represent new independence variables with zero mean and covariance matrices.

A posteriori state estimate for the original nonlinear process are defined as below:

$$\hat{x}_k = \widetilde{x}_k + \hat{e}_k$$
(7)

Eq. (5) and Eq. (6) have the following probability distributions:

$$p(\widetilde{e}_{xk}) \approx N(0, E[\widetilde{e}_{xk}\widetilde{e}_{xk}^T])$$

$$p(\varepsilon_k) \approx N(0, WQ_kW^T)$$

$$p(\eta_k) \approx N(0, VR_kV^T)$$

By setting the predicted values of $\hat{e}_k$ be zero, Kalman filter equation for $\hat{e}_k$ is;

$$\hat{e}_k = K_k\widetilde{e}_{zk}$$
(8)

Substituting Eq. (8) in Eq. (7) and using Eq. (4),

$$\hat{x}_k = \widetilde{x}_k + K_k\widetilde{e}_{zk}$$
$$= \widetilde{x} + K_k(z_k - \widetilde{z}_k)$$
(9)

Eq. (9) is now can be used for measurement update in EKF with $\widetilde{x}_k$ and $\widetilde{z}_k$ and the Kalman gain, K_k obtained from previous equations.

B.   Complete Picture of Operation of EKF Algorithm

Time Update ("Predict")
(1) Project the state ahead
$$\hat{x}^-_k = f(\hat{x}_{k-1}, u_{k-1}, 0)$$
(2) Project the error covariance ahead
$$P^-_k = A_k P_{k-1} A^T_k + W_k Q_{k-1} W^T_k$$

Measurement Update ("Correct")
(1) Compute the Kalman Gain
$$K_k = P^-_k H^T_k \left(H_k P^-_k H^T_k + V_k R_k V^T_k \right)^{-1}$$
(2) Update estimate with measurement z_k
$$\hat{x}_k = \hat{x}^-_k + K_k \left(z_k - h(\hat{x}^-_k, 0) \right)$$
(3) Update the error covariance
$$P_k = (I - K_k H_k) P^-_k$$
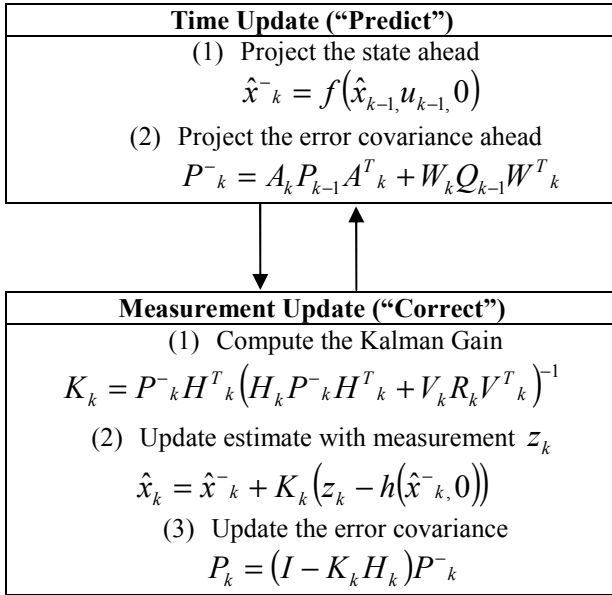
Figure 2. Extended Kalman Filter (EKF) Loop

Figure 2 summarizes the overall operation of EKF algorithm. The prediction stage starts by projecting the state and error covariance ahead using the present value of Jacobian matrix and prior value of error covariance. Then, based on the projected error covariance value, the Kalman gain is calculated and measurement $\hat{x}_k$ is estimated. Finally, the error covariance value is updated. The loop will continue until the desired error covariance value is achieved.

III.   METHODOLOGY

A. Instrumentation and Data

The data used in this paper is from Sg. Batu Pahat at Batu Pahat located in the state of Johor, Malaysia. The water level data is in meters, starting from 0.00 a.m June 1, 2012 until 12.00 p.m June 4, 2012. This real-time data is available online from website www.water.gov.my. The water level data is measured using Supervisory Control and Data Acquisition System (SCADA) by the Department of Irrigation and Drainage Malaysia.

B. Simulation Studies

MATLAB version 2009a is used as software platform for simulation studies and the numbers of estimated data are 119 since the data is taken on hourly basis. The RMSE of estimation is applied to show the accuracy of filtering, which is defined as follows:

$$RMSE(\hat{x}_k) = \left[\frac{1}{N_{data}} \sum_{i=1}^{N_{data}} \left(x_k^{(i)} - \hat{x}_k^{(i)} \right)^2 \right]^{\frac{1}{2}} \qquad (10)$$

IV.   RESULTS AND DISCUSSION

The simulation result of flood water level prediction using Extended Kalman filter algorithm is given in Figure 3. From the figure, it can be observed that predicted value tracking nicely the true value with small RMSE of 0.9236 m. Tracking fail to follow high frequency movement or sharp edges. In this study, all parameters which is the Jacobian matrix was set to unity. Hence, the parameters value effect may contribute to the overshoot problem. However, the small value of RMSE indirectly explains that the result is still acceptable.
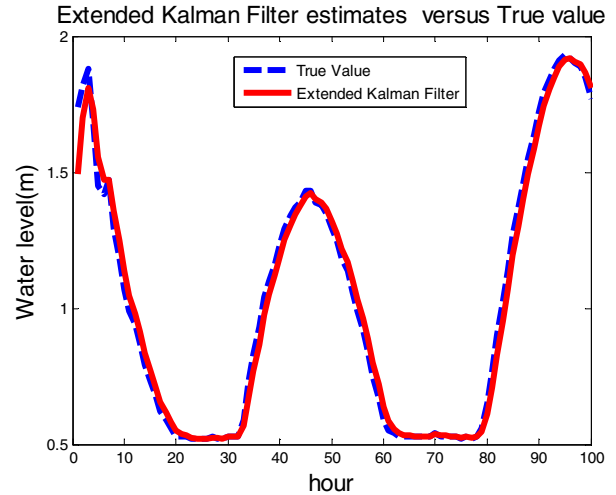
Figure 3. Water level prediction using Extended Kalman Filter (EKF) algorithm

VI.   CONCLUSIONS

This study has demonstrated that Extended Kalman Filter can be employed successfully to flood water level prediction phenomenon. This is because EKF is purposely developed to solve nonlinear problems and flood is one of the nonlinear cases involved. Even though EKF can estimate flood water level well, there are many more types of Kalman Filter that are able to do that. Hence, for future works, other methods can be used and the performance result can be compared and analyzed.

REFERENCES

[1] Lees et al, "An adaptive flood warning scheme for the River Nith at Dumfries," In:White, W.R., Watts, J. (Eds), 2nd International Conference on River Flood Hydraulics, J. Wiley & Sons, Chichester, pp. 65-77, 1994.

[2] Young, P.C., "Recursive estimation and time series analysis," Springer-Verlag, Berlin, 1984.

[3] Todini, E., "Mutually interactive state-parameter (MISP) estimation, application of Kallman filters to hydrology, hydraulics and water resources," In: Chiu C-L, editor, proceedings of AGU Chapman Conference, 1978.

[4] Bras RL, Rodriguez-Iturbe , "Random functions and Hydrology," Reading, MA,USA: Addison WesleyF. Li-hua and L. Jia, "The practical research on flood forecasting based on artificial neural networks," ELSEVIER Expert Systems with Applications, vol. 37, pp. 2974-2977, 2010., 1985, pp. 559, 1985

[5] M. Barut, S. Bogosyan and M. Gokasan, " Speed-sensorless estimation for induction motors using extended Kalman filters," IEEE Trans. Ind. Electronics, vol. 54, no 1, pp. 272-280, Feb 2007.

[6] N. Salvatore et. al.,"Optimization of delayed-state Kalman filter based algorithm via differential evolution for sensorless control of induction motors," IEEE Trans. Ind. Electron., vol. 57, no 1, pp. 385-394, Jan 2010.

[7] Y.R. Kim, S.K. Sul, M.H. Park, "Speed sensorless vector control of induction motor using Extended Kalman Filter," IEEE Trans. Ind. Applic., vol.30, pp.1225-1233, 1994.

[8] Sadeghi, B., Moshiri, B., " Second order EKF and Unscented Kalman Filter fusion for tracking maneuvering targets," IEEE International Conference on Information Reuse and Integration, pp. 514-519, 13-15 August 2007.

[9] R.A. Best and J.P. Norton, "A new model and efficient tracker from target with curvilinear motion," IEEE Trans. Aerosp. Electron. Syst., vol.33, no. 3, pp. 1030-1037, Jul 1997.

[10] R.G. Brown and P.Y.C. Hwang," Introduction to random signals and applied kalman filtering," New York:Wiley, 1997.

Air-Conditioning Energy Consumption of an Education Building and it's Building Energy Index: A Case Study in Engineering Complex, UiTM Shah Alam, Selangor

M.B.A.Aziz [1] , Z.M.Zain [1], S.R.M.S.Baki [1], and R.A.Hadi [2]
[1] Faculty of Electrical Engineering
[2] Facility Management Office
Universiti Teknologi MARA,
40450 Shah Alam,
Selangor Darul Ehsan

Abstract – **This paper presents the Air Conditioning energy consumption of an education building and it's Building Energy Index at Engineering Complex formerly Science and Technology (S & T) Complex building, University Teknologi MARA Shah Alam Selangor. A fact has shown that the buildings are responsible for using a vast amount of energy. The energy consumption in buildings is increased by air conditioning system. To implement actions to save energy in buildings, it is necessary to measure and study its energy efficiency. This study describes the result of user attitudes towards air conditioning usage, electric energy consumption pattern and building energy index in building. Through the site survey, the results show that, the operation of air conditioning system was unsatisfying. In conclusion, the air conditioning energy consumption can be reduced by several proposed measures. The proposed measures are categorized into two ways: by improving user attitude towards using air conditioning through awareness and technical improvement.**

Keywords – *Air Conditioning, Energy Consumption, Building Energy Index (BEI)*

I. INTRODUCTION

Buildings consume about one-third of the world's energy and the worldwide energy consumption for buildings is expected to grow from 45% from 2002 to 2025[1], [2]. At present, commercial buildings consume well over one-third of all electricity in the ASEAN region and will account for more than 40% of the demand for additional generating capacity in the near future [1], [3].

The energy consumed in Malaysia is 90% in the form of electricity. If these trends continue, buildings will consume almost as much as industry and transport combined [1], [4]. This is quite a concern as Malaysia has become one of the fastest growing building industry in the world lately. Although this maybe a cause for anxiety, but it does provide opportunities for the development of sustainable energy technology.

The typical distribution of energy consumption in Malaysian buildings is shown in Table I, where shopping complexes and offices used more than 90% of the energy for lighting and air conditioning. Residential buildings used the least energy for both. Here again there is ample opportunities for energy efficiency measures and use of renewable energy source [1].

TABLE 1
DISTRIBUTION OF ENERGY CONSUMPTION IN MALAYSIAN BUILDINGS (%)

	Residential	Hotels	Shopping Complexes	Offices
Lighting	25.3	18.0	51.9	42.5
Air-Conditioning	8.3	38.5	44.9	51.8
Total	33.6	56.5	96.8	94.3

The following building energy efficiency potentials are attainable [5]:

a) 40% to 50% reduction of energy consumption of new buildings
b) 15% to 25% in reduction in energy consumption of existing buildings
c) Shift of electricity demand for buildings from day to night, thus improving the load factor on electricity generating equipment for some ASEAN countries

The main aim of the paper is to address the issues related to the field of air conditioning energy consumption and thermal comfort in Engineering Complex, University Teknologi MARA Shah Alam Selangor. A case study will be carried out to investigate the current performance of the operation of Air conditioning system to provide thermal comfort at this Complex.

A. The Plant at Engineering Complex

Air conditioning system in Engineering Complex was built with a centralised chiller water plant with Thermal Energy Storage (TES) scheme in year 2002. The plant was commissioned in June 2002 with capacity of 3600 RT installed Chiller capacity and 10,800 RTh Thermal Energy Storage capacity. In principle, the design of the TES scheme is meant to shift load from peak to off-peak. Table II below, shows the Engineering Complex district cooling design [10].

TABLE II
DISTRICT COOLING DESIGN

Chiller Installed	2 x 1800 RT Base Mode
	2 x 1250 RT Ice Mode
Thermal Storage Capacity	45 x 240 RTh Ice Cell
Building Cooling Load Requirement	Block 1 550 RT
	Block 2 580 RT
	Block 3 380 RT
	Block 4 850 RT
	Block 5 660 RT
	Tower 1 290 RT
	Tower 2 290 RT
MaxCooling Demand	3600

The TES plant is using Dunhambush Ice Cell technology with 45 numbers of Ice cell stored under parking lot. The plant was designed to have one unit chiller 900RT x 2 compressors to charge the Ice Cell during off-peak hours between (10.00 p.m. to 8.00 a.m.) every day. This will give a low production cost benefits to UiTM where the night tariff is much cheaper compared to the on-peak hours. Not only the charge is lower but the Maximum Demand Charge can be avoided by almost 40% [10].
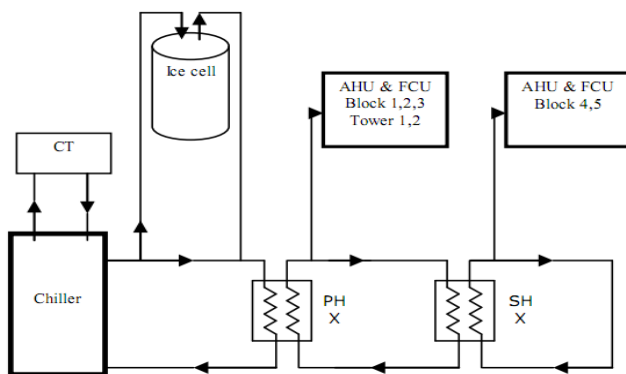
Fig. 1 Operation schematic diagram of the plant [10]

As shown in Figure 1, the whole system consists of 3 closed circuit namely primary, secondary, and tertiary loop. The liquid flowing in the primary loop is ethylene glycol and chilled water flows in the secondary and tertiary loops. This design of multiple loops allowed the system to operate in all 5 operating modes; ice build, ice build with cooling, cooling with ice, cooling with chiller, and cooling with ice and chiller [10].

The plant utilizes partial storage operating strategy in which during discharge mode, the ice storage and the chiller running day mode work together to cover the building load. This strategy is quite critical, where controlling the contribution of chiller and ice storage are critical to the system economy and comfort. Modulating valves are used to manage the relative contribution of storage and chiller [10].

Since cooling load reflected by numbers of occupant and chiller production, aiming to achieve and maintaining 44°F/6.67°C is important as to achieve thermal comfort. As the load and chiller contribution varies, the modulating valves will automatically allow sufficiently flow though the storage system to maintain 42°F/5.56ºC of ethylene glycol to be supplied to the primary HX. In between ethylene glycol and chiller water, the process of heat transfer is occurred, to gain the desired temperature of 44°F/6.6°C, chilled water will be supplied to the Air Handling Units (AHU) and Fan Coil Unit (FCU) for the entire building and return at 54ºF/12.2°C to 58ºF/14.5°C.

The plant is using water cooled chillers regarding of its beneficial contribution in Engineering Complex. This system usually embedded inside the building and heat from chiller is carried by regulating water to outdoor cooling tower. Its potential advantages appear in term of energy efficiency referring to operational time of chiller in Engineering Complex [10].

B. Standards and Building Energy Index

The energy consumption in buildings is normally in term of the Building Energy Index (BEI). The South East Average BEI is 233 kWh/m^2/yr whereby the Malaysian and Singaporean average are 269 kWh/m^2/yr and 230 kWh/m^2/yr respectively [1]. In Malaysia not much published data is available yet on the Energy Performance or actual annual energy consumption of buildings. Pusat Tenaga Malaysia - PTM is still in the process of auditing Government Office Buildings and Center for Environment, Technology & Development, Malaysia[6], [7], [8].

The Malaysian Ministry of Energy, Water and Communication (MEWC) or previously known as the Ministry of Energy, Communication and Multimedia (MECM) before the year 2004; introduced the Guidelines for Energy Efficiency in Non-Domestic Buildings in 1989. In the first version, the BEI was recommended to not be more than 135 kWh/m^2/yr [9]. In 2001, it was improved to include revised equations based on the latest research work on local buildings. As the energy efficiency awareness was beginning to set in at the time, the BEI of several new buildings began to demonstrate a decrease in value.

The guidelines now were renamed as the Malaysian Standard MS 1525:2001. The purpose of the MS was as follows:

a) To encourage the design of new and existing buildings so that they may be constructed, operated and maintained in a manner that reduces the use of energy without constraining the building function, nor the comfort or productivity of the occupants and with appropriate regard for cost considerations.

b) To provide the criteria and minimum standards for energy efficiency in the design of new buildings, retrofit of existing buildings and methods for determining compliance with these criteria and minimum standards

c) To provide guidance for energy efficiency designs that demonstrates good professional judgment and exceeds minimum standards criteria.

The standards were later improved in 2007 to include the following:

a) provide guidance for energy efficiency designs that demonstrate good professional judgment and exceeds minimum standards criteria; and

b) Encourage the application of renewable energy in new and existing buildings to minimize non-renewable energy sources, pollution and energy consumption whist maintaining comfort, health and safety of the occupants.

The most important inclusion was in the additional objective statement:

1. To encourage the application of renewable energy in new and existing buildings to minimize non-renewable energy sources, pollution and energy consumption whist maintaining comfort, health and safety of the occupants.

2. The MS 1525:2007 sets out minimum standards for designers to design and select equipment above those stipulated in the code in the following areas:
   a) Architectural and passive design
   b) Building Envelope
   c) Lighting
   d) Electric power and distribution
   e) Air conditioning and mechanical ventilation (ACMV) system
   f) Energy management control system

II. METHODOLOGY

In order to achieve the objective if this study several measures have been selected to gain information and data for the study, which are:

A. Questionnaire survey

User attitudes towards electrical energy usage in Engineering Complex were conducted by means of questionnaires. A total of 100 respondents were chosen for this survey. The survey was conducted via questionnaires and interviewing the users.

B. Energy consumption pattern

Basically the electricity bill consists of 2 sections covering electrical system utilizing tariff C1 and air conditioning system utilizing tariff C2. The trend of electric consumption in Engineering Complex was studied based on the monthly electricity bills data by Tenaga Nasional Berhad (TNB) and Building Energy Index was calculated.

C. Comfort level in building

HOBO data logger was used as measurements of the temperature (in °C) and the relative humidity (in RH %) for the data collection process at selected rooms.

III. RESULTS & DISCUSSION

A. Site survey- Users behaviour study

From the questionnaire study the occupancy pattern of Engineering Complex was established

Duration of occupancy
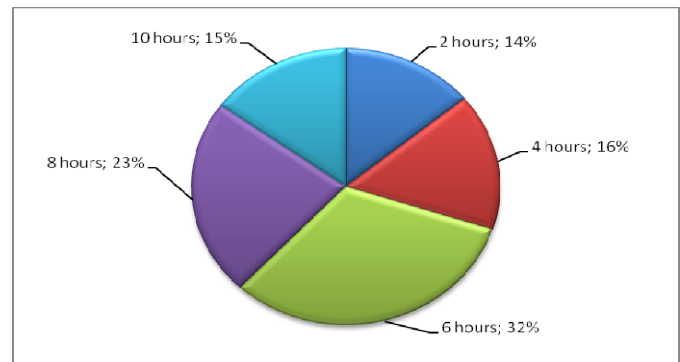
Fig. 1 Duration of user occupancy in Engineering Complex

According to Figure 1, basically, most of the users spent more than 3 hours in Engineering Complex daily. Only 30% spent less than 5 hours per day in the building. Majority of them spent between 6 to 8 hours daily in the building. It can be concluded that, about 70% users spent their time (working, attend classes, assignment, lab, research and act...) in the building between 6 to 10 hours daily.

Occupants' attitude towards switching off air conditioning system
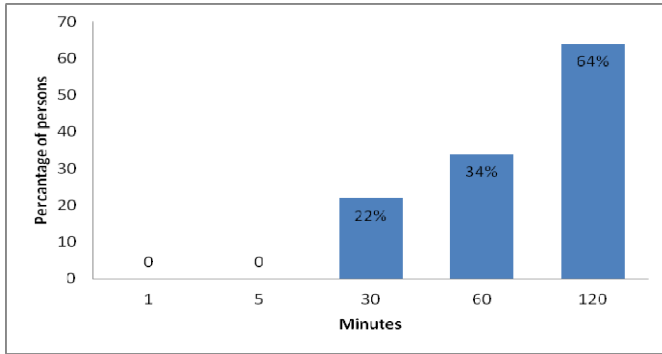
Fig. 2 Number of person's switching off the air conditioning system when leaving their rooms and class

One of question in the survey form is to find out whether the users turn off or not the air conditioning systems when they left their rooms and class. As the result shown in Figure 2, the survey found that nobody switched off the air conditioners when they went out for less than five minutes. Only 22% of users switched off the air conditioner when they are not in their rooms and class for 30 minutes and 34% of users switched off air conditioner for 1hour, when they are not in their rooms and class. The good practice is that majority of the users switch off their air conditioner when they return home.

Occupants' response to coolness level provided by air conditioning system
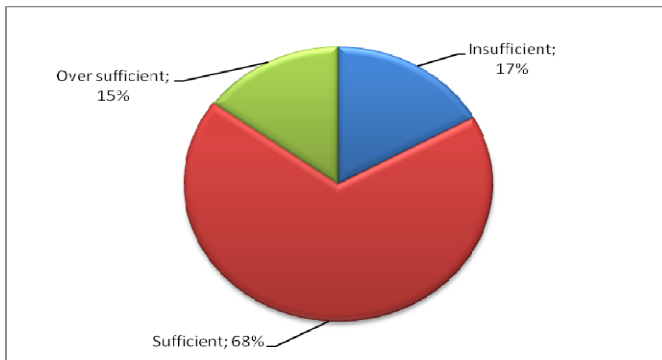
Fig. 3 Users respond on their air conditioning system

The users were also asked whether the coolness level provided by air conditioning system is insufficient, sufficient, or excessive. As shown in Figure 3, 17% of the users replied that the air conditioning system was insufficient, 68% replied that it was sufficient and 15% claimed that the air conditioning system was excessive.

Occupant's satisfaction with the comfort level of the existing Air-Conditioning system
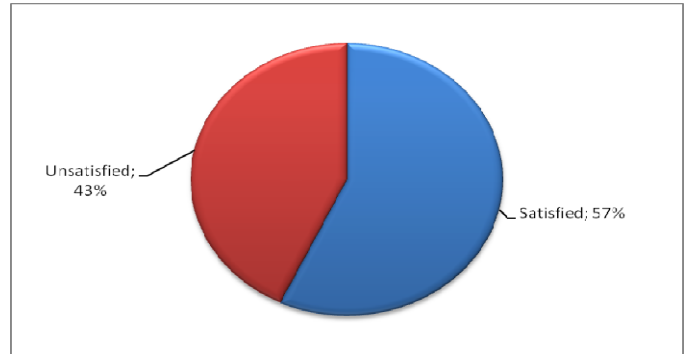
Fig. 4 Users respond on the comfort of the existing air conditioning system

The last question of the questionnaire is about comfort. The question was given to the users is to evaluate the comfort of the existing air conditioning system, either it is comfortable or uncomfortable. As shown in the Figure 4, the result shown about 57% of the users was satisfied and 43% claimed that the air conditioning system was unsatisfied.

B. Trend of Energy Consumption

The analysis of electricity consumption was based on monthly summary report for year 2010. According to summary, the total annual energy consumption for Engineering Complex was 11,532,465 kWh amounting to RM 3,712,519,20 with average monthly electricity bill RM 309,376.60. From the electricity bill 2010, it was observed that the monthly electricity bill from August till October were reduced but it was showed an increment on the November and then followed by a reduction at the end of the year (as shown in Figure 5). The highest energy consumption was in March 2010 with 1,196,882 kWh amounting to RM 349,269.87 which was the beginning of the new academic session year.
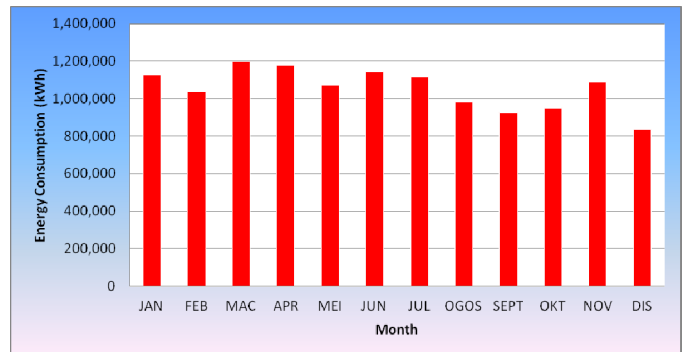
Fig. 5 Montly electricity consumption for Engineering Complex in the year of 2010

TABLE III
MONTHLY ELECTRICITY CONSUMPTION FOR ENGINEERING
COMPLEX IN 2010

Month	C2 Tariff (kWh)	Other-C1 Tariff (kWh)	Total (kWh)	BEI (kWh/m2.yr)
JAN	1,127,661	830,227	1,957,888	196.51
FEB	1,037,261	748,585	1,785,846	179.24
MAC	1,196,882	863,138	2,060,020	206.76
APR	1,179,176	807,567	1,986,743	199.40
MEI	1,074,206	763,375	1,837,581	184.43
JUN	1,144,155	718,879	1,863,034	186.99
JUL	1,115,736	800,115	1,915,851	192.29
OGOS	981,783	821,980	1,803,763	181.04
SEPT	924,977	689,367	1,614,344	162.03
OKT	949,273	824,477	1,773,750	178.03
NOV	1,087,151	705,655	1,792,806	179.94
DIS	835,865	682,466	1,518,331	152.39
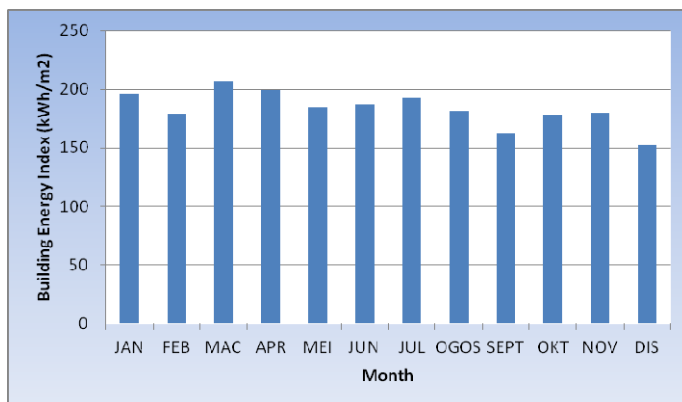
Fig. 6 Monthly BEI of Engineering Complex in 2010

Calculation of Building Energy Index:

Average of Energy, kWh in 1 year = 1,825,830
Estimated total Energy, kWh for 1 year,
    = 1,825,830 x 12
    = 21,909,960

Total Energy (kWh) ÷ gross floor area based on an air-conditioned
    = 21,909,960 ÷ 119,561.27 m²
    = 183.25 kWh/ m²/year

Hence, the Building Energy Index;
    = 183.25 kWh/ m²/year
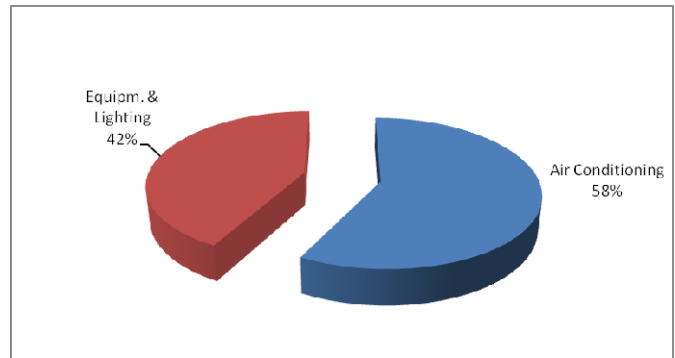
Fig. 7 Breakdown of energy consumption in the Engineering Complex

The monitored building energy index (BEI) based on total energy consumed in the building is 183.25 kWh/ m^2 per year, based on an air-conditioned area of 119,561.27 m^2. The energy for cooling represents 58% and equipment & lighting 42% of the total energy consumption of the building (as shown in Figure 7). In comparison, results from a study conducted by Pusat Tenaga Malaysia in 2003 on government buildings based on conventional design reveal that energy consumption is typically 62% air-conditioning, 18% lighting and 20% plug loads.

Building energy index (BEI) values computed from monthly energy consumption data are shown in Figure 6. The monitored energy consumption and BEI in August, September and October were lower than the other months due to the implementation of EE strategies on AC system followed by November and December were lower than the other months due to the chiller problems (lower occupancy rate due to semester break), resulting in the reduced energy consumption.

Currently, from calculation the Building Energy Index (BEI) of Engineering Complex is 183.25 kWh/ m^2/year. The best BEI practice and recommended by Malaysian Standard is 135 kWh/ m²/year. Therefore, Engineering Complex BEI is 46.6% above the recommended value. Meanwhile, the recommended achievable target is 160 kWh/m²/year which is 28% saving.

Nevertheless, this reference new suggestion for the value especially the concept of the green building has been forwarded now, where the Building Energy Index as high as 183.25 kWh/m²/year still acceptable. The energy audits carried out by Pusat Tenaga Malaysia, PTM, of Office buildings in Malaysia revealed that the majority of Malaysian Office building had BEI in the range of 200 to 250kWh/m²/year.

C. Thermal comfort evaluation

From the collected data on the studies activities conduct, the following comments were made on thermal comfort evaluation as tabulated in Table IV as shown below:

TABLE IV
TEMPERATURE AND RELATIVE HUMIDITY AT SELECTED ROOMS

No	Room	Temperature °C	Relative Humidity (RH),%	Comment
1	Artificial Intelligence Research Group Lab	24.3	70	Temp. and RH comply to Malaysian Standard
2	Pejabat FKE, Level 9	24.5	63	Temp. and RH comply to Malaysian Standard
3	Bilik Seminar Ibn Al-Haytham, Level 8	24.0	66	Temp. and RH comply to Malaysian Standard
4	CISCO Lab, Level 8	27.0	72	Temp. and RH not comply to MS but still acceptable
5	Bilik Persatuan FKE, Level 8	24.3	67	Temp. and RH comply to Malaysian Standard
6	B1-A13-3	25.3	80	Temp. comply to MS, but RH unacceptable
7	B2-A5-7	29.9	69	Temp. unacceptable to MS, but RH comply to MS
8	B3-A12-6	26.6	71	Temp. and RH not comply to MS but still acceptable
9	B4-A11-3	25.0	80	Temp. comply to MS, but RH unacceptable
10	B5-A12-1	27.9	72	Temp. and RH not comply to MS but still acceptable

TABLE V
THERMAL COMFORT INDICATOR

	Relative Humidity (RH %)	Temperature (°C)	Comment
A	50 ± 10	Humidity Ratio = < 12 g vapor per kg dry air < 25	Comply to ASHRAE and MS
B	< 70	24.5 ± 2.5	Comply to MS
C	< 75	27.5	Not comply to MS but still acceptable
D	< 80	29.5	Unacceptable
E	Not complied to D		Extremely unacceptable

IV CONCLUSION

In conclusion, the demand for electricity in Engineering Complex is growing annually at the total annual energy consumption year by year and for year 2010 with average monthly electricity bill RM 309,376.60 because of the continual renovation in building, increasing electrical tariff, increasing population in building and the reliance on air conditioning systems for the cooling of building.

Based on the survey study, 15% of the respondents state that the air-conditioners are too high in cooling level and this leads to the waste of energy and electricity. As much as 17% of the respondent said that the air-conditioners are not cooling enough and this might be so because some room have malfunctioning air-conditioners. In conclusion from the questionnaire about the coolness level, respondents have been asked about the comfort level of the air-conditioner in the building. 47% state their dissatisfaction because based on the comments stated out in surveys, among them are in some lecture rooms, the air-conditioners are not functioning well enough, the air-conditioning system in buildings frequently shuts down and that the rooms are too cold.

V RECOMMENDATION

Based on the results of the present study, the recommended thermal comfort indicator for air conditioning system has been introduced to provide thermal comfort with energy efficiency at Engineering Complex (as shown in Table V).

Based on current energy consumption trend in Engineering Complex, there are many rooms to improve the present electric energy consumption. Several energy saving opportunities were proposed in this study. Proposal of electric energy saving opportunity is categorized into the two, one is saving opportunity by improving occupants' behaviors and another one is saving opportunity by means of technical measures. Improving human behavior is possible through training, poster campaigns, sign, published, tip, guideline, monitoring and control and technical measures which is occupancy sensor should be used in classroom and lecturer office to avoid wastage of cooling energy. The temperature set point and flow rate should be according to the design specification and comply with Malaysia Standard.

VI ACKNOWLEDGMENT

VII REFERENCES

[1] Azni Zain Ahmed, Integrating Sustainable Energy in Buildings: A Case Study in Malaysia, FAU conference, Copenhagen, Denmark, 14-15 May 2008

[2] Klee, H. *Energy Efficiency in Buildings (EEB) Project*. 2007

[3] Ministry of Energy, Communication and Multimedia , National Seminar on Energy Efficiency in Buildings. Kuala Lumpur, 2001

[4] ABCSE, *Renewable Energy in Asia: Malaysia Report*. Australia, 2005

[5] Kannan, K, *Energy Efficiency in Buildings*, National Seminar on Energy Efficiency. Kuala Lumpur, 2001

[6] "MECM LEO SEMINAR Advances on Energy Efficiency and Sustainability in Buildings" Palace of Golden Horses Kuala Lumpur 21-22 January 2003.

[7] "Making Malaysian Homes Energy Efficient – Stakeholder Workshop organized by CETDEM" Impiana Hotel Kuala Lumpur 12 June 2004

[8] Ar Chan Seong Aun, Energy Efficiency, Designing Low Energy Building Using Energy 10, CPD seminar 7th August 2004.

[9] *Code of Practice on Energy Efficiency and Use of Renewable Energy for Non-Residential Buildings*, Department of Standards Malaysia, MS 1525: 2007.

[10] Aziz, M.B.A.; Zain, Z.M.; Baki, S.R.M.S.; Muslam, M.N.; , "Review on performance of Thermal Energy Storage system at S & T Complex, UiTM Shah Alam, Selangor," Control and System Graduate Research Colloquium (ICSGRC). 2010 IEEE , pp.49-54, 22-22 June 2010

Rectangular Spiral Microstrip Antenna for WLAN Application

Aiza Mahyuni Mozi, Dayang Suhaida Awang Damit, Zafirah Faiza

Faculty of Electrical Engineering ,

Universiti Teknologi MARA MALAYSIA,

13500 Permatang Pauh, Pulau Pinang, MALAYSIA

Abstract - **Microstrip patch antennas represent one family of compact antennas that offer a conformal nature and the capability of ready integration with communication system's printed circuitry. This paper highlights a constructed of 2.4 GHz of a coaxial – fed rectangular spiral microstrip patch antenna for WLAN application where it's mounted on a high frequency laminates Rogers 5880 Duroid with 2.2 dielectric constant and 0.508 mm of thickness. The radiation pattern of this antenna illustrates an omni-directional pattern with voltage standing wave ratio of less than two, return loss less than -10 dB, the line impedance of 50 Ohm. The results show the consistency in the simulated and fabricated antenna.**

Keyword – **rectangular spiral microstrip antenna, WLAN, Rogers 5880 Duroid, omni-directional**

I. INTRODUCTION

The evolution of modern wireless communications systems has increased dramatically together with the demand for antennas which serve a wireless land mobile or terrestrial-satellite network. With time and requirements, these devices become smaller in size and also to be smaller and lightweight which results microstrip antenna. The advantage of microstrip antenna makes them popular in many applications requiring a low profile antenna and it's promising to be a good candidate for the future technology due to the flexibility of the structure as it is easy to fabricated, have low cost and can be easily incorporated into the communication equipments.

There are many types of antennas used in WLANs which includes wire antennas, aperture antennas, microstrip antennas, array antennas, reflector antennas, and lens antennas [1]. Many researchers have studied different structure [2], [3] and different techniques to increase the radiation efficiency in single element antenna by using double PIFA [4], U-slot [5] and other structures. One of the requirement for WLANs antenna is that it should have omni-directional properties with circular polarization [6].

A spiral antennas offer (due to their relatively simple geometry) additional attractive features like light weight, easy fabrication as well as integration with microwave and millimeter-wave circuit. Furthermore, spiral antennas have characteristics that are vital to wireless communications in many applications. Their wide – band characteristics accompanied with their low profile as well as their low cost make these antennas prime candidates for use in sensors and some mobile applications compare to other band characteristic such as dual band characteristics[7]. Figure 1 shows two types of the spiral antennas that are used in WLANs application [8].

Due to the circular polarization characteristic and the frequency independent characteristics at a relatively small size for WLANs applications, this paper aims to investigate on the design of rectangular spiral microstrip antenna to achieve higher directive gain and radiation efficiency.

In fact, this type of antenna can be used to avoid the usage of multiple antennas for different services since an excessive number of access points in one area can complicate the access due to the interface with other access point [9]. The prototype was fabricated on Rogers 5880 Duroid with 2.2 dielectric constant and 0.508 mm of thickness and is fed by a standard 50 Ohm coaxial fed. The centre frequency is 2.45 GHz and the bandwidth is 22 MHz. This rectangular spiral antenna was fed by using coaxial cable at the centre.
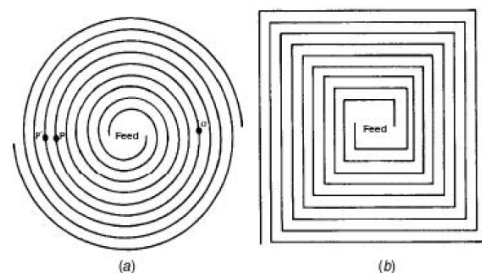
Figure 1 : Archimedean spirals (a) round (b) square

II. METHODOLOGY

A. Design Procedure

The design and simulation of proposed rectangular spiral microstrip patch antennas was conducted using CST Studio Suite software with the resonant frequency of the antenna at 2.45 GHz. The specifications and the substrate properties as in Table I and Table II respectively. It's used for theoretical and simulated study where the dimensions are in mm range. The layout of the antenna demonstrate in Table 3 with dimension of 0.058mm (h), 4mm (W), 4mm (inner radius), 3mm (gap) and 4 turns of spiral as in Table III.

TABLE I : SPECIFICATION OF THE ANTENNA

Center frequency (fc)	2.45 GHz
Bandwidth (BW)	22 MHz
Radiation Pattern	Omni-directional
Voltage Standing Wave Ratio (VSWR)	≤ 2
Return Loss (S_{11})	$\leq$-10 db
Line Impedance (Z_o)	50Ω

TABLE II : SUBSRATE PROPERTIES

Properties	Rogers 5880 Duroid
Dielectric constant, ε_r	2.2
Mue	1
El Tand	0.025
Thickness (mm), h	0.508
Metal Thinckness (mm)	0.03
Resistivity	1
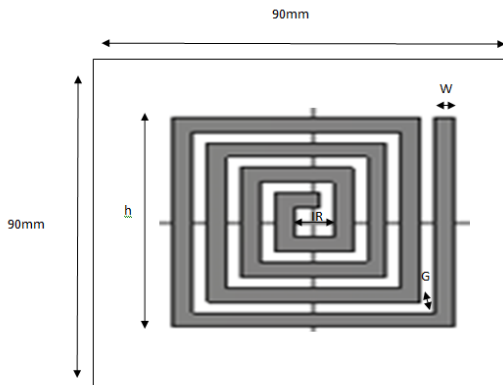
Figure 2 : Dimension of Rectangular Spiral Microstrip Antenna

TABLE III: DIMENSION OF RECTANGULAR SPIRAL MICROSTRIP PATCH ANTENNA

Height, h (mm)	0.058
Width, W(mm)	4
Gap, G	3
Inner radius, IR (mm)	4
No. Of Turns	4

B. Feeder network

This antenna has a simple structure fed by 50Ω microstrip line. There are several methods that can be used to feed the antenna which can be classified into two categories: contacting and non-contacting. Contacting Coaxial feeding method is used in [9]. Rectangular Spiral Microstrip Patch Antenna are fed from underneath via a probe as shown in Figure 3. The outer conductor of the coaxial cable is connected to the ground plane, and the center conductor is extended up to the spiral antenna. The position of the feed can be altered to control the input impedance. The coaxial feed introduces an inductance into the feed that may need to be taken into account if the height h gets large. In addition, the probe will also radiate, which can lead to radiation in undesirable directions [10].
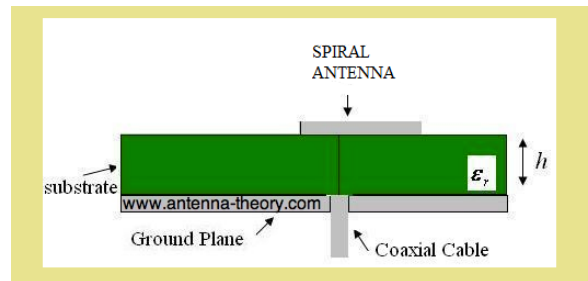
Figure 3 : Coaxial cable feed for Rectangular Spiral Microstrip Patch Antenna

C. Simulation

Simulation of Rectangular Spiral Microstrip Patch Antenna inclusive of a coaxial feed was carried out. The dimensions of the antenna was kept constants at 9 mm x 9 mm. Figure 4 shows the simulated Rectangular Spiral Microstrip Patch Antenna. Here, CST Microwave Studio is used to obtain the return loss, the radiation patterns and others needed parameter. Figure 5 shows return loss at centre frequency while Figure 6 shows line impedance; Figure 7 shows voltage standing wave ratio, and Figure 8 shows radiation pattern. As for this paper an omnidirectional was designed, the actual antenna gives an isotropic radiation pattern in a single plane.
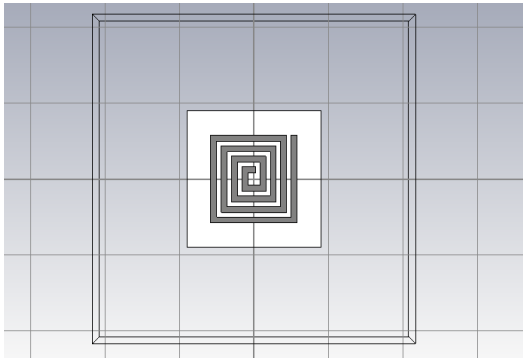
Figure 4: Simulated Rectangular Spiral Microstrip Patch Antenna
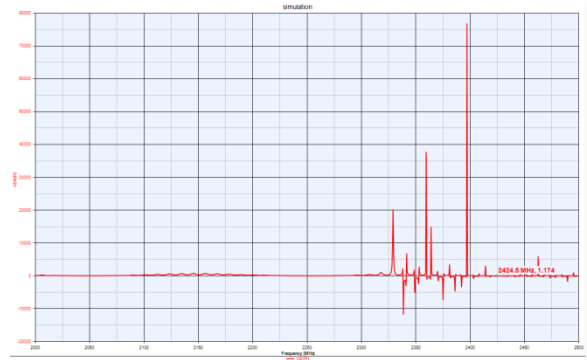
Figure 7 : Voltage Standing Wave Ratio (VSWR)
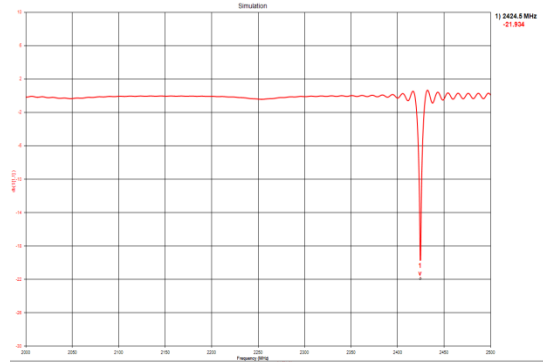
Figure 5 : Return Loss, S_{11} at center frequency

Frequency   = 2.45
Main lobe magnitude = -0.1 dB
Main lobe direction = 2.0 deg.
Angular width (3 dB) = 77.2 deg.
Side lobe level = -18.9 dB
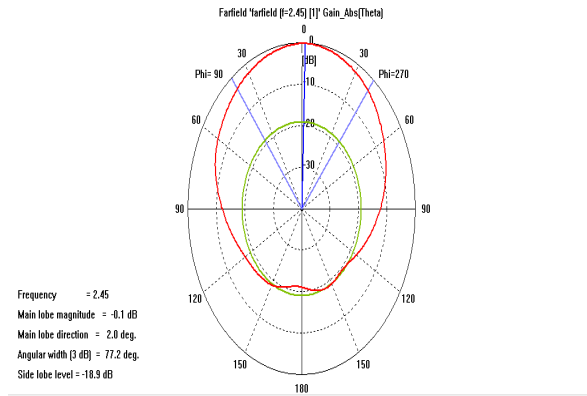
Figure 8 : Radiation Pattern

The parameters for the simulated antenna were tabulated in Table 4.

As presented in Table IV, it can be observed that the simulation values are very close to the specified value for the antenna.

Figure 6 : Line Impedance, Z_o

TABLE IV : COMPARISON ON SIMULATED VALUES AND SPECIFICATIONS

Parameter	Specification	Simulted
fc	2.45 GHz	2.42 GHz
Radiation Pattern	Omni-directional	Omni-directional
VSWR	≤2	1.174
S_{11}	≤ -10 dB	≤ -21.294
Z_0	50Ω	50.1Ω

III.   RESULT AND DISCUSSION

In Figure 9, the antenna has designed to cover specific operating frequency which is 2.45 GHz. The S- parameters for the fabricated antenna was measured with the aid of Rohde-Schwarz ZVA40 vector network analyzer. The measurement was done at Communication Lab, UiTM. Prior to the measurement, the VNA needs to undergo

calibration process. The calibration utilizes the OSM (Open-Short-Match) standards.

Figure 9 : Fabricated circuit

Based on the result shown in Figure 10, the simulation result gives 21.934dB of return loss (S11) at 2.4245GHz. However, this value changes to 23.796dB at 2.4825GHz from the measurement process. The frequency is being shifted by 2.4% to the right.
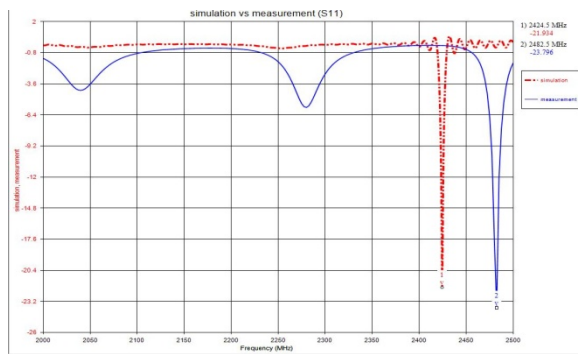
Figure 10 : Measured and simulated values of S_{11}

Figure 11 shows the value for both simulated and fabricated antenna. For simulated, the VSWR is 1.174 at 2.4245 GHz while for the fabricated,the measured VSWR is 1.138 at 2.4825 GHz. The differences in term of the frequency and VSWR between the measured and simulated circuit are 2.4% and 0.036 respectively.
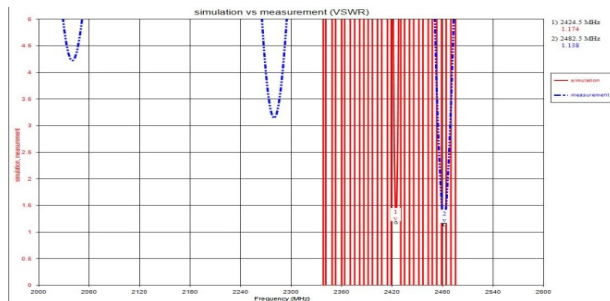
Figure 11 : Measured and simulated values of VSWR

The radiation pattern for both simulation and fabrication is being depicted in Figure 12. The software simulation radiation pattern shows that a 3dB beam width of 77.2° at 2.45 GHz. The measured radiation pattern at 2.45GHz has similar radiation pattern as the one from the simulation.
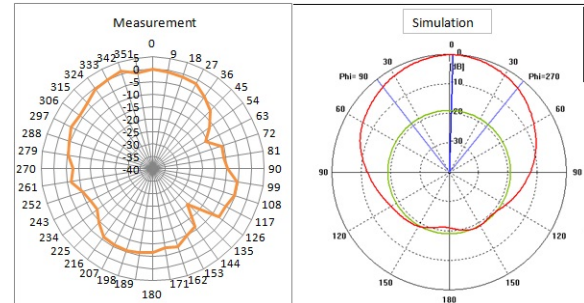
Figure 12 : Measured and simulated radiation pattern

TABLE V : COMPARISON ON SIMULATED VALUES AND MEASURED

Parameter	Simulation	Measured	Different
f_c (GHz)	2.45GHz	2.42GHz	2.4%
S_{11} (dB)	-21.294	-23.796	-2.502
VSWR	1.174	1.138	

Due to the fact that error exists during the fabrication process, table V summaries all the parameters taken into consideration in this antenna analysis.

The discrepancies of the simulated and fabricated results that cause losses in antenna performances may be due to several for example the measurement process does not take into the vacuum state as in the simulation. The best place to measured radiation pattern is anechoic chamber.

III.   CONCLUSION

The unique feature of this antenna is its simplicity to get higher performance. This paper presents a geometric configuration for the Rectangular Spiral Microstrip Patch Antenna for WLAN Application, which provides a mean to get higher directive gain and maximum radiation efficiency without using special techniques. The responses are considered to be acceptable as the measured results are quite similar to the simulated circuit.

REFERENCES

[1]   Electronic Communications Systema, Fifth Edition, by Wayne Tomasi, ISBN 981-244-642-7 Prentice-Hall Inc.
[2]   Dinisio Raony Ribeiro, Leonardo Augusto de Santana, Marlen Carneiro Alves and Jose Felipe Almeida, Carlos Leonidas da S.S. Sobrinho,"Spiral Microstrip Antenna", IEEE MTT-S International Microwave & Optoelectronics Conference (IMOC 2007), pp. 104-106,2007.
[3]   A. Mehta, D. Mirshekar-Syallkal, and H.Nakano, "Single Element Beam Steerable Antanna – From Rectangular Spiral to Star", EMTS

2007 International URSI Commission B-Electromagnetic Theory Symposium, Ottawa, ON, Canada, July 26-28, 2007.

[4] P. S. Hall, E.Lee, C. T. P. Song, Planar Inverted-F Antennas Chapter7, "Printed Antennas for wireless Communications" Edited by R. Waterhouse, John Wiley & Sons, Ltd., 2007.

[5] H. F. AbuTarboush, H. S. AI-Raweshidy, R. Nilavalan, "Triple Band Double V-Slots Patch Antenna for WiMAX Mobile Applications", 14th Asia-Pacific Conference on Communications, Japan, pp. 1-3,2008.

[6] Nikolaos L. Tsitsas, Constantinos A. Valagiannopoulos, "Mathematical Modeling of Spherical Mircrostrip Antennas and Applications", National Technical University of Athens, Greece.

[7] Ugur Saynak, "Novel Rectangular Spiral Antennas", School of Engineering and Sciences of Izmir Institute of Technology.

[8] N.H. Abdul Hadi, K.Ismail, S.Sulaiman, M.A Haron,"Design of a Rectangular Spiral Antenna for Wi – Fi Application",Faculty of Electrical Engineering Universiti Teknologi Mara.

[9] Salman Haider, "Microstrip Patch Antennas for Broadband Indoor Wireless Systems",Faculty of Engineering University of Aukland,2003

[10] D. Misman, M.Z.A Abd. Aziz, M.N Husain, P.J Soh, R.B Ahmad, "Design and Analysis of an UteM Antenna",Faculty of Engineering Electronic and Computer Engineering University Technical Malaysia Malacca.

Index on Fabric Wrinkle Using Image Processing Technique

Anis Diyana Rosli, Hadzli Hashim, Noor Ezan Abdullah, Nadiah Ismail, Zuraidah Halil

Faculty of Electrical Engineering

Universiti Teknologi Mara

Malaysia

Abstract—Image processing has been gradually growing on demand as a technique for evaluating and analyzing the gradation of fabric wrinkle. In this work, 150 samples of the Wrinkle Recovery Replica was captured digitally, processed and statistically analyzed to produce respective ranking index based on the ratio of black and white pixel elements, firstly introduced as Wrinkle Pixel Index (WPI_{hz}). There are five classes of fabric wrinkle gradation generated (WPI_{hz-1}, WPI_{hz-2}, WPI_{hz-3}, WPI_{hz-4} and WPI_{hz-5}) which are based on group clustering of the acquired measured indices during image processing. In this experiment, the scope is limited to produce wrinkle index representing only for one tone color of wrinkled fabric samples. Finally, validation is done by comparing the experimental results with the ones that were already being classified by an expert observer prior before doing the image processing. The outcome has shown that this novice technique has achieved similarity result of 86.67% with the expert's classification.

Keywords-component; Image processing, Matlab, Wrinkle rating

I.   INTRODUCTION

Conventionally, the morphological method of classifying ranking for any wrinkled fabric is by observing the wrinkle lines and compared them to several references of Wrinkle Replicas (WR). This method requires the service of an expert observer, who sometimes has diverse evaluation depending on their own judgment, vision ability and so on [1]. The American Association Textile Chemists and Colorist (AATCC) and American Society for Testing and Materials, ASTM D-123 are two techniques used extensively in the industry for the evaluation of fabric wrinkle [2].

The protocol for ascertaining smoothness grade of a fabric is outlined in the American Association Textile Chemists and Colorist (AATCC) Test Method, numbered as Method 124 [3]. This standard test was designed to evaluate the smoothness of fabric after undergoing finishing process which is the process of reducing wrinkle using chemical by Textile Researcher. Three test specimens, with each cutting of 15x28 cm per fabrics, have gone through three standard steps. Preliminary step involves the specimen to be weighted by 3500g with AATCC Wrinkle Tester as shown in Figure 1.

Figure 1: AATCC Wrinkle Tester

The weight then be removed after 20 minutes and the specimen is placed on the cloth hanger for 24 hours. The expert observer will rate each specimen under lighting equipment as demonstrated in Figure 2.
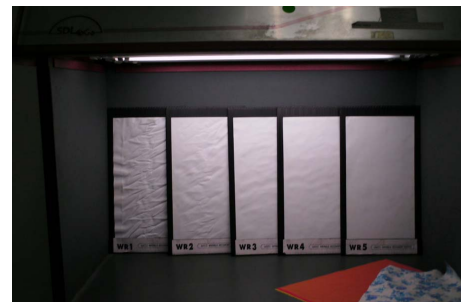
Figure 2: Lighting equipment for viewing test specimens

Figure 3 illustrates a set of Wrinkle Recovery Replicas, arranged ascendingly, from rating number 1 to rating number 5. The smoothest and best retention of original appearance indicate by rating number 5 which also corresponds to WR-5 replica. The decrease in rating number signifies the decrease in appearance retention of original appearance. Therefore, rating number 1, also known as WR-1 replica, has the poorest appearance retention of original appearance [4].

Figure 3: Images of the Wrinkle Recovery Replicas

The conventional method seems easy to be carried out but it is unreliable and time ineffective assessment of fabric wrinkle [2]. Consequently, image processing technique can play significant role not only in defining fabric wrinkle but also on various works relating to the drape behavior of fabrics, seam puckering and pilling [5].

The research which involves image processing technique has been proposed by Kang et al. [6]. Their research employs a 3D-projecting grid technique to assess the level of fabric wrinkling on American Association Textile Chemists and Colorist (AATCC) replicas. The level of fabric wrinkle is classified based on parameters acquired from image processing of the AATCC replica. Those parameters are roughness ratio, surface area ratio, wrinkle density and power spectrum density of the fast Fourier transform.

II. METHODOLOGY

In this work, the process of WPI_{hz} classification, which consists of five steps-data collection, image acquisition, pre-processing, range index extraction and classification- was used. Figure 4 described the workflow for fabric wrinkle grading.
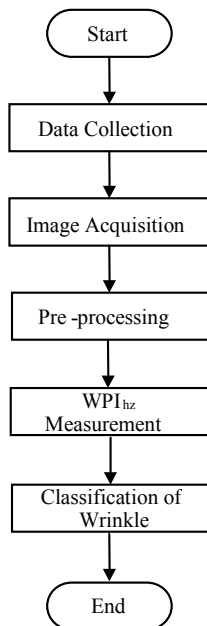
Figure 4: Workflow of fabric wrinkle grading

A. Data Collection

A set of Wrinkle Recovery Replicas was borrowed from the Faculty of Applied Science, Universiti Teknologi Mara Shah Alam. This set includes five Fabrics Smoothness Ratings. Each number of ratings corresponds to the same number of Wrinkle Replica. For instance, rating number 5 represents an appearance equivalent to the Wrinkle Replica5, rating number 4 represents an appearance equivalent to the Wrinkle Replica 4 and so on. 30 samples were taken from each replicas used in the research which total up to 150 samples.

B. Image Acquisition

The luminance of lighting was controlled within 1420.02 ± 22.13 lux during the morning session and 1416.87 ± 12.756 lux for the evening session. This is due to the significant of p-value for morning-evening is 0.177 as tabulated in Table I.

Table I: Paired Sample Test for Morning, Evening and Night Session

No	Paired sample test	P-value
1	Morning-evening	0.177
2	Morning-night	0.035
3	Evening-night	0.000

The Wrinkle Replicas was placed on the table. The light source (Digicolor Sportlight K-250C) was positioned 1.25 meters from the Wrinkle Replicas by angle of 70° and 2 meters height as demonstrated in Figure 5.
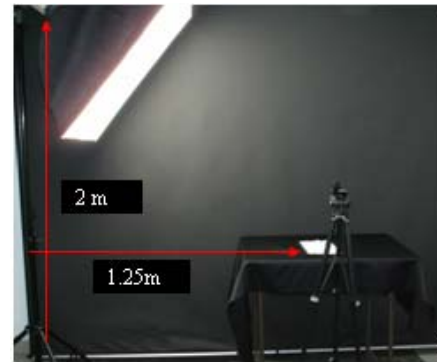
Figure 5: Position of spotlight from the Wrinkle Replica

The camera was set 0.35 meters from the surface of the table while the angle was set to 35° as shown in Figure 6. All the settings were set up to get the best image. The captured image was analyzed using the algorithm developed in MATLAB 7.0 with Image Processing Toolbox.
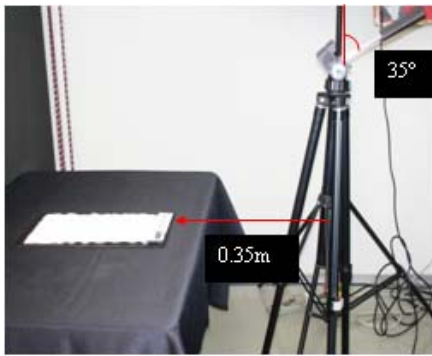
Figure 6: Distance and angle of camera from the Wrinkle Replica

C. Pre-processing

The preliminary step involved in pre-processing is the image be resized to a dimension of 256 x 256 pixel area as shown in Figure 7a. As in Figure 7b, the image then was converted to grayscale since most of the processing steps need the image to be in grayscale form. The edge detection was used to identify the edge for each image. In this research, the Canny method was applied to assimilate the information about the edge. The Canny method displayed the wrinkle in white color and the surface in black color as illustrated in Figure 7c. Subsequently the Canny image was transformed using Imcomplement to get the wrinkle appearance in black color and the surface in white color as indicate in Figure 7d.

Figure 7a: Load resized image

Figure 7b: Grayscale image

Figure 7c: Edge detection image
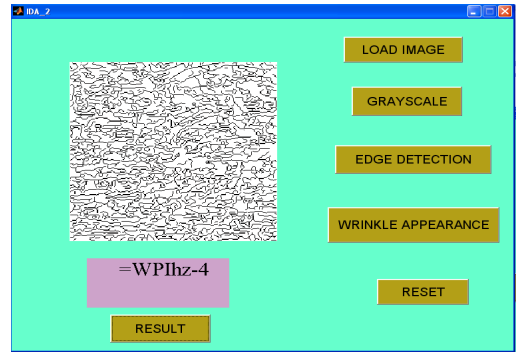
Figure 7d: Final wrinkle appearance result

D. Measurement

The number of white and black pixel was figured out to extract the percentage of wrinkle appeared on the sample. Equation (1) was used to calculate the percentages of black pixel in Figure 7d. Equation (2) was used to compute the accuracy between both methods.

$$k = \% \text{pixel of black} = \frac{\sum \text{pixel of black}}{(\sum \text{pixel of black} + \sum \text{pixel of white})} \quad (1)$$

$$\text{Accuracy} = \frac{\sum \text{Testing samples ok}}{(\sum \text{Testing samples})} \text{X100}\% \quad (2)$$

E. Classification

In this study, condition operator was used to make decision in classification of WPI_{hz} based on the percentage of black pixel (k). The classification of five ranges of WPI_{hz} was described in Table II below.

Table II: WPI_{hz} coding for five classification groups

Range	Coding
k ≤0.12204	for index 1 or WPI_{hz-1}
0.12205≤ k ≤0.1523	for index 2 or WPI_{hz-2}
0.1524≤ k ≤0.1805	for index 3 or WPI_{hz-3}
0.1806≤ k ≤0.1921	for index 4 or WPI_{hz-4}
k≥0.1922	for index 5 or WPI_{hz-5}

These five ranges of WPI_{hz} are demonstrated in Figure 8.
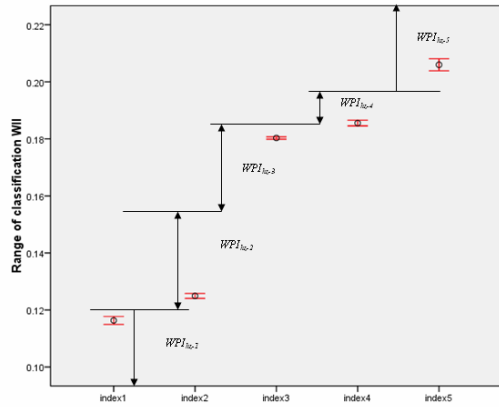
Figure 8: Range of WPI_{hz} classification

III. RESULTS AND DISCUSSION

This novice technique in fabric industry had been tested on 30 fabric specimens randomly selected from 10 different types of fabric. Only this amount is available from the resource lab during the validation process. However, this size is still qualified and the minimum number for many practical situations that allow the usage of normal distribution as an approximation for sampling distribution [7]. Nevertheless, it is expected that any results should improve if more samples are considered which is recommended in the future work. Take note that the 30 fabric specimens were evaluated by two expert observers prior to using the WPI_{hz} technique for automatic classification. Table III shows the result of comparison between the expert observers' decision and the WPI_{hz} technique.

Table III: Comparison between expert observers and WPI_{hz} technique results

Number	Expert Observer (WR)	WPI_{hz}	Remark
1	3	3	OK
2	3	3	OK
3	3	3	OK
4	1	1	OK
5	1	1	OK
6	1	1	OK
7	5	5	OK
8	5	5	OK
9	5	5	OK
10	3	3	OK
11	3	3	OK
12	2	3	Not OK
13	3	3	OK
14	3	3	OK
15	3	3	OK
16	3	3	OK
17	3	3	OK
18	3	3	OK
19	2	4	Not OK
20	2	3	Not OK
21	2	4	Not OK
22	5	5	OK
23	5	5	OK
24	5	5	OK
25	1	1	OK
26	1	1	OK
27	1	1	OK
28	5	5	OK
29	5	5	OK
30	5	5	OK

From Table III, the accuracy of WPI_{hz} technique is calculated to be 86.67% compared to the result obtained from expert observers.

IV. CONCLUSION AND FUTURE WORK

A novel technique has being introduced in this research where gradation of fabric wrinkle can be processed using image processing technique instead of the classical grading done by observation from an expert person. This technique has proposed that for a one tone fabric wrinkle factor can be measured by an index known as WPI_{hz} where eventually, there are five classes of fabric wrinkle gradation generated (WPI_{hz-1}, WPI_{hz-2}, WPI_{hz-3}, WPI_{hz-4} and WPI_{hz-5}) which are based on group clustering of the acquired measured indices during image processing. Validation process has proved that this technique is in agreement with the expert observers' decision with similarity of 86.67%. Therefore, this novel technique of computing WPI_{hz} can be adopted by any inexpert and novice observer for fabric wrinkle classification grading with reliable accuracy.

For future study, the capability and consistency of the technique should be tested with more than 30 fabric specimens. In addition, a new algorithm can be explored for finding WPI_{hz} specifications for multi-tone fabric colors. Also, the parameters produced by the WPI_{hz} technique can be used as input parameter in designing an intelligent system such as Fuzzy Logic or Neural Network for classification of wrinkle factor.

REFERENCES

[1] S. A. Mirjalili, E. Ekhtiyari, "*Wrinkle assessment of fabric using image processing*," Fibres & Textiles in Eastern Europe, vol. 18, pp. 60–63, 2010.

[2] B. K. Behera, R. Mishra, "*Measurement of fabric wrinkle using digital image processing*," Indian Journal of Fibre & Textile Research, vol. 33, pp. 30–36, March 2008.

[3] *AATCC Technical Manual*, (American Association of Textile Chemists & Colorists, USA), pp. 208, 1991.

[4] *AATCC Technical Manual*, (American Association of Textile Chemists & Colorists, USA), pp. 214, 2005.

[5] M. Mitsuo , "*Fabric handle and its basic mechanical properties*," Journal of Textile Engineering, vol. 52, pp. 1–8, 2006.

[6] T. J. Kang, D. H. Cho, H. S. Whang, "*A new objective method of measuring fabric wrinkles using a 3-D projecting grid technique*," Textile Research Journal, vol. 69, pp. 261–268, April 1999.

[7] G. Keller, B. Warrack, "*Inference About the Comparisons of Two Populations*," Statistics for Management and Economics, vol. 5th., ed California: Duxbury Thompson Learning, pp. 394-470, 2000.

A Development and Challenges of Grid-Connected Photovoltaic System in Malaysia

M. Z. Hussin, N. Hasliza, A. Yaacob, Z. M. Zain,
A. M. Omar
Faculty of Electrical Engineering
Universiti Teknologi MARA
40450 Shah Alam, Selangor Darul Ehsan
mohamadzhafran@gmail.com

S. Shaari
Faculty of Applied Sciences
Universiti Teknologi MARA
40450 Shah Alam, Selangor Darul Ehsan
Malaysia

Abstract—**This paper presents the update status of a grid-connected photovoltaic (GCPV) system installed in Malaysia. There are 125 sites with total PV capacity power is approximately 1137.21 kWp of GC BIPV systems until the end of June, 2011 was monitored by Photovoltaic System Monitoring Centre (PVSMC). After 5 years of operation under MBIPV project, two main problems have been identified due to technical and environmental, which contributes to 60% and 40% respectively. This factor may contribute the decrement on the PV system performances in terms of field yield. By identifying the potential problem occurred, the knowledge and experiences in a GCPV system will used as the input and the potential problem will be avoided, in order to achieve the maximize the energy production. All issues are identified, which affects the system performances based on 5 years experience through MBIPV project by conducting short-term investigation.**

Keywords- Grid-Connected (GC); Photovoltaic (PV).

I. Nomenclature

SREP   Small Renewable Energy Programme
UNDP   United Nations Development Programme
GEF   Global Environment Facility
MBIPV   Malaysia Building Integrated Photovoltaic
PVSMC   Photovoltaic System Monitoring Centre
TNB   Tenaga Nasional Berhad
TNBR   Tenaga Nasional Berhad Research
SETM   Success Electronics and Transformer Manufacturer
EETC   Environmental & Energy Technology Centre

II. Introduction

RENEWABLE energy at national level was first stated in the Eighth Malaysia Plan (2001-2005) through introduction of the Small Renewable Energy Programme (SREP), and in the Third Outline Perspective Plan 2001-2010. Under SREP, small power generation plants which utilizes of all sources types of Renewable Energy, including biomass, biogas, municipal waste, solar, mini-hydro and wind power, are allowed to sell electricity to the utility grid through the Distribution Grid system. Renewable energy (RE) is again explicitly addressed in the Ninth Malaysia Plan (2006-2010), by launched the MBIPV project starting from July 2005 and was planned to end around December, 2010 [1]. The MBIPV project will specifically focus on the market development for Building Integrated Photovoltaic (BIPV) technology, and building the national capacities on three major parts: (a) policy and education; (b)

technical skill and market implementation; (c) technology development support. Subsequently from Ninth Malaysia Plan, Government of Malaysia (GoM) through Tenth Malaysia Plan (2011-2015), has introduced the Renewable Energy Act 2010 and taking the MBIPV project as a significant input toward RE development in the country. In the earlier of December, 2011, a feed-in-tariff scheme for renewable energy electricity has started and supervised by Sustainable Energy Development Authority (SEDA) [2].

The Malaysia Building Integrated Photovoltaic (MBIPV) Project is a national development programme implemented by the Ministry of Energy, Green Technology and Water Malaysia (KeTTHa) with support from United Nations Development Programme (UNDP) and Global Environment Facility (GEF) with a total fund of USD 24.5 million allocated over a 5-year period. The GoM has decided to support building integrated photovoltaic (BIPV) technology in view of identifying BIPV's technical potential and long-term viability for widespread application in Malaysia, and in other tropical such as ASEAN countries. An initial pilot study on grid-connected PV application that was conduct during 1998-2002 established that the technology is technically viable and would produce higher electricity output in comparison to similar systems installed in temperate climate countries. Nevertheless, the relative high initial costs, as well as the lack of enabling environment have prevented BIPV from penetrating the local market. In addition, the purposes of the MBIPV project as a input to verify the technical potential and the market status of BIPV technology in Malaysia's context. At the end of the MBIPV project, the experiences and lessons learned with adequate information will give significantly to determine the value and cost effectiveness of BIPV technology market in Malaysia's perspectives.

Photovoltaic System Monitoring Centre (PVSMC) is a body that monitors the performance of PV and BIPV installations under MBIPV project. The MBIPV project team agreed and awarded a contract to the Universiti Teknologi MARA (UiTM) on 29th November 2006 to classify PVSMC as a PV monitoring centre [3]. After 5 years of operation, PVSMC has officially closed on June, 2011.

This paper presents the statistics and issues of the grid-connected system installation in Malaysia until the end of June, 2011. In order to know clearly about statistics and critical issues of the Grid-Connected Photovoltaic (GCPV) systems in the Malaysian context are discussed.

III. DELEOPMENT OF GCPV SYSTEM IN MALAYSIA

A. Update of Development of Grid-connected PV System

In Figure 1 illustrates the cumulative growth in PV capacity since 1998 within the two primary statuses for PV in Malaysia. About 468.2 kWp of total PV capacity were installed from 1998 until 2006 including baseline of GCPV installation. After MBIPV project was launched in July 2005, the increasement is more than 60% over the previous year, which brought the total PV capacity installed up to 1605.4 kWp by the end of December, 2011. To date, the largest PV installation with 361.9 kWp is the Enterprise Four Building at Technology Park Malaysia (TPM) as illustrated in Figure 2, whereas the second largest PV installation situated at the Malaysian Green Technology Corporation (GreenTech Malaysia) with PV capacity of 92.00 kWp, was commissioned in June 2007 [4]. Other than that, the third largest commissioned in 2009 with 71.4 kWp GCPV system located at Energy Commission Building (EC Building), Putrajaya using First Solar FS273 modules (CdTe) as shown in Figure 4.
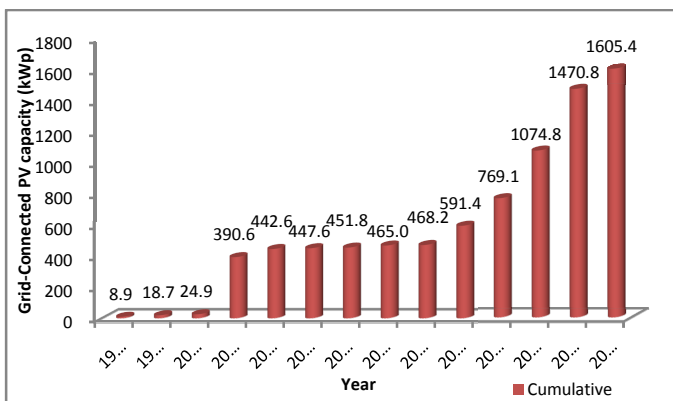
Figure 1. Cumulative of GC BIPV system installed capacity in Malaysia (1998 - June, 2011).

Figure 2. A largest of GCPV system with 361.9 kWp at Enterprise Four Building, Technology Park Malaysia (TPM), Malaysia.(Sources: TPM)

Figure 3. A second largest of GCPV system with 92.00 kWp at Malaysian Green Technology Corporation, Bandar baru Bangi , Malaysia [1].
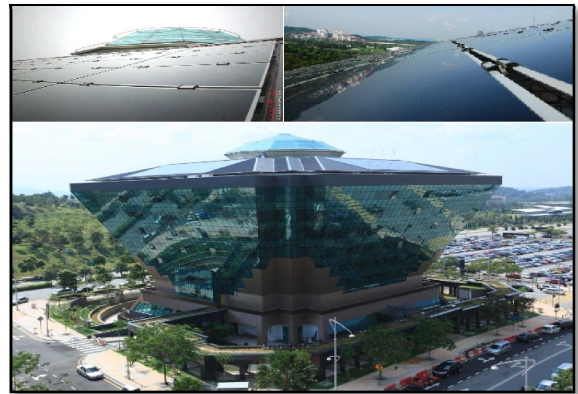
Figure 4. A third largest CdTe thin-film of GCPV system with 71.4 kWp at Energy Commission Building , Putrajaya, Malaysia [1].

B. History of Development of Grid-Connected (Baseline)

In Malaysia, the first installations were in 1998 when three of GC BIPV systems with a combined total capacity of 16.86 kWp were installed. One of the first GCPV system has been installed in August 1998 on the rooftop of College of Engineering, Universiti Tenaga Nasional (UNITEN), Kajang, Selangor Darul Ehsan [5,6]. A 3.15 kWp of GCPV installation was initiated by TNB Research Sdn. Bhd. (TNBR) as part of a pilot for "Demo Renewable-Energy" and the research project supported by Malaysia Electricity Supply Industry Trust Account (MESITA) and Tenaga Nasional Berhad (TNB). Under MESITA and TNB as co-funding, six pilots GCPV systems were installed in between 1998 until 2001. In the same year, two further GCPV systems have been installed by BP Malaysia and Universiti Kebangsaan Malaysia (UKM). A 7.95 kWp GCPV system has installed at BP petrol station along the KESAS highway [5], and a 5.76 kWp PV at the Solar Energy Research Park, Universiti Kebangsaan Malaysia. A 5.76 kWp with single phase installation was operated by the Solar Energy Research Group in Universiti Kebangsaan Malaysia which is funded by Intensification of Research in Priority Areas (IRPA) [7]. In addition, the BP Petrol's PV system of 7.95 kWp at the BP petrol station was officiated by HM Queen of England during her royal visit to Malaysia for the Commonwealth Games in 1998. However, both BP PV systems have been removed in 2001 and 2006.

In October 1999, a bungalow used as a PV research centre was built with 3.60 kWp of GC PV system used as the roof of the bungalow in TNB Research Centre, Kajang, Selangor.

Other than that, a family of a TNB's senior officer became the first Malaysian family to experience BIPV system at their home in August, 2000 when the 3.15 kWp BIPV system was installed at their house in Port Dickson, Negeri Sembilan [6]. The PV installation increase drastically in the year 2001 when a one-off installation of 361.9 kWp of total capacity was installed by Technology Park Malaysia (TPM). This installation at Enterprise Four Technology Park Malaysia, Bukit Jalil demonstrated Malaysian capability to handle and manage large PV installations and this project was funded by TPM is the biggest PV installation in South-East Asia Pasific region at that time [8]. It comprises of 4824 fixed mounted roof modules and the system is coupled with a UPS battery bank and a generator as backup. In early 2002, another bungalow was built by an architect firm (NLCC Architect) to demonstrate PV integration as a roof. A collaboration between NLCC Architects Sdn Bhd and Fraunhofer Institute of Solar

Energy (ISE Fraunhofer), Germany had built of 3.19 kWp GCPV project is called "Prototype Solar House" in SIRIM with three PV systems had integrated into three different roofs consisted of Monocrystalline BP275 module – 1.05 kWp, triple junction Thyssen Solartec® a-Si PV module / thin film – 1.02 kWp, and Polycrystalline module type Lafarge Braas SRT 35 – 1.12 kWp [9]. In 2004, BIPV's have also been installed at Centre for Environment, Technology and Development Malaysia (CETDEM) office's in Petaling Jaya, whereas MEGTW-LEO building's at Putrajaya. By the end of November, 2005, a roof of the private bungalow had integrated with a 5.25 kWp BIPV system and located in Semenyih, Selangor, belongs to a private company, Lucaswork Sdn Bhd (formerly known as Soon Hoe Technology Sdn Bhd). TABLE I presents the development of GCPV systems (baseline) in Malaysia from 1998 until 2006.

TABLE I.   OVERVIEW OF DEVELOPMENT FOR GRID-CONNECTED PV SYSTEM (BASELINE) INSTALLATION IN MALAYSIA [6,7,10].

No	Site	Nominal power (kWp)	Commissioning Date
1*	BP Petrol Station, KESAS Highway	7.95	January 1998
2	Solar Energy Research Park, Universiti Kebangsaan Malaysia	5.76	June 1998
3	College of Engineering , Universiti Tenaga Nasional (UNITEN), Kajang, Selangor	3.15	August 1998
4	SIRIM Berhad, EETC, Shah Alam, Selangor	2.39	January 1999
5	Solar PV Research & Monitoring Centre (Solar House) , TNB Research, Kajang	3.60	October 1999
6	TNB House, Port Dickson, Negeri Sembilan	3.08	August 2000
7	Public House, Seksyen 11, Shah Alam, Selangor	3.12	November 2000
8#	BP Petrol Station, PLUS KL- SEREMBAN Highway	16.5	March 2001
9	Enterprise Four Building , Technology Park Malaysia (TPM), Bukit Jalil	361.9	March 2001
10	Sekolah Sri Cempaka, Bukit Damansara, Kuala Lumpur	0.99	August 2001
11	TNBR's Nursery, Kajang, Selangor	3.78	October 2001
12	Residential House, Subang Jaya, Selangor	2.80	November 2001
13	SIRIM II : Prototype Solar House, Section 2, Shah Alam, Selangor	3.19	March 2002
14	Universiti Teknologi Malaysia, Skudai, Johor	3.00	June 2002
15	Computer Integrated Farming TPM, Bukit Jalil, Kuala Lumpur	30	September 2002
16	Universiti Islam Antarabangsa, Gombak, Kuala Lumpur	15.84	October 2002
17	Solar Hydrogen Eco-House, Universiti Kebangsaan Malaysia (UKM)	5.04	December 2003
18	MEGTW-LEO Building, Putrajaya	3.30	May 2004
19	CETDEM, SS2/53, Petaling Jaya, Selangor	0.90	November 2004
20	SETM Sdn Bhd's Solar Park, Sg. Buloh, Selangor	7.89	August 2005
21	Private Bungalow, Taman Tasik Kesuma, Semenyih, Selangor	5.25	November 2005
22	Faculty of Electrical , Universiti Malaya, Kuala Lumpur	3.20	2006

*7.95 kWp system at BP petrol station on KESAS Highway dismantled in 2001 [10].

#16.5 kWp system at BP petrol station on KL-Seremban Highway dismantled in May 2006 [10].

IV.   STATUS OF GCPV SYSTEM UNDER MBIPV PROJECT

A. BIPV Pricing Trends
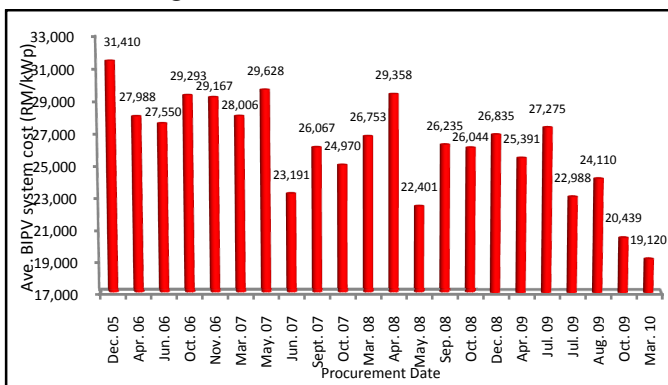
Figure 5.   Local BIPV pricing trend based on market survey by MBIPV [1].

According to the price cost in Figure 5, a "trendline" shown a significant decline in the average price per kWp BIPV installation in Malaysia. Before MBIPV launched, the average BIPV cost price was approximately RM 31,410 in December, 2005 and showed the decrement in March, 2010 with RM 19,120 per kWp [1]. It can be seen that the price reduction achieved more than 60% until the end of March, 2010.

B. PV's Site Monitored By PVSMC

The total capacity of GCPV systems project installed and commissioned in Malaysia is around 1,605.4 kWp until June, 2011 which includes 468.2 kWp baseline of GCPV systems installed before the MBIPV project commenced and some have been dismantled [1]. However, only 125 sites were monitored by PVSMC from July 2005 until the end of June, 2011 with total PV capacity of 1137.2 kWp, Selangor has the highest PV site installation among the other states with total power of 469.9 kWp recorded. In contrast, the lowest nominal

power by Perak with one site of 5.28 kWp as depicted in Figure 6.
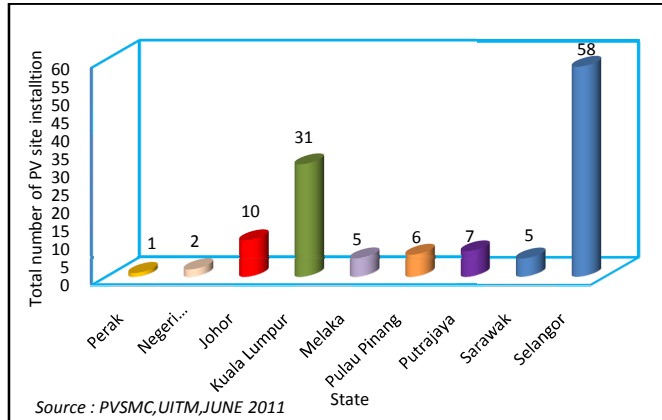
Figure 6.   Selected locations in Malaysia monitored by PVSMC [3].

C. *Market Share of PV Module Technologies Used Under MBIPV Project.*

Figure 7 below presents the total share percentage of PV site installed using different types of PV module technologies under MBIPV project.
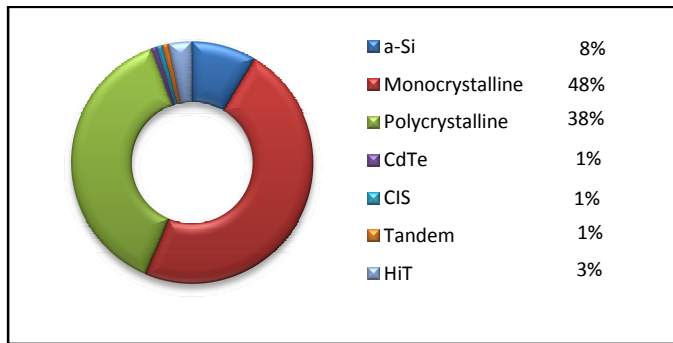
Figure 7.   Market share percentage of PV module technologies installed in Malaysia under MBIPV project [3].

The total of PV system installed and monitored in Malaysia is around 125 sites with various different types of PV module technologies installed. Mostly installation were

used monocrystalline module technologies about 43%, whereas the polycrystalline module as a second dominate the market share under MBIPV project in Malaysia around 38%. For the thin-film technologies, 10% (a-Si, CdTe, CIS) were used and others 4% using Heterojunction with Intrinsic Thin-layer (HiT) and Tandem (a-Si/µc-Si) modules based on the data monitored by PVSMC, UiTM.

V.   ISSUES AND CHALLENGES OF RENEWABLE ENERGY DEVELOPMENT BASED MALAYSIAN EXPERIENCE.

A. *Environmental Problem Under MBIPV Project*

After 5 years monitoring process done by PVSMC, several problems were identified such as environmental problem, e.g. shading, low energy due to orientation of PV array, lightning strikes due to SPD defect and technical problem which influences the PV system performance. The data were extracted from [3].

Shading is an unavoidable presence on the PV array in many PV systems. It is always advisable to eliminate shading to the best of the installer's ability and keep away from excessively shaded on the PV array surfaces. Other than that, shading problem that occurs even in one PV module will give a strong influence to the overall PV performances which contributes to the power output wastages.

Over monitored period, 10 sites were influenced by shading effect in Malaysia's field as tabulated in TABLE II. In each case for 10 sites were identified, a reduction in the energy harvested in terms of final yield performance occurred. Mostly, this problem occurred in the residential house. The environmental factor such as shading effect is identified as one of the problem occurred in the GC BIPV systems installed in Malaysia during short-term investigation.

Futhermore, GC BIPV system also facing lower energy problem at certain sites. About 6 sites facing lower energy were reported due to orientation and one site was detected caused by Surge Protection Device (SPD) damage on DC side as summarized in TABLE II.

TABLE II.   OVERVIEW OF PV'S SITE THAT POSSIBILITY AFFECTED BY ENVIRONMENTAL PROBLEMS [3].

State	Site	Shading	Low Energy	SPD Defect
Kuala Lumpur	Private Bungalow House, Jalan Maarof, Bangsar.	√		
	Private Bungalow House (2), Kota Damansara.	√		
	Private Bungalow House, Bangsar.	√		
	Private Bungalow House, Bukit Damansara.	√		
	Private Bungalow House, Taman Desa		February 2009	
Selangor	BORID Energy, seksyen U1, Glenmarie, Shah Alam	√		
	Private Bungalow House, Bangi,	√		
	Private Bungalow House, Salak Tinggi, Sepang	√		
	Private Bungalow House Subang Jaya, Subang Jaya	√		
	Monash University, Sunway	√	Sept & Oct 2008, Feb 2009	
	Private Bungalow House, Bandar Kinrara.		February 2009	√
	Unit Bungalow House Z1B (COLUMBA) , Setia Eco Park		Jan 2010 & May 2010	
	Unit Bungalow House Z1A (AQUILA), Setia Eco Park		June 2008	
	Unit Bungalow House Z1E (LEO), Setia Eco Park.		October 2009	
Melaka	Private Bungalow House, Bukit Beruang.	√		

B. *Technical Problem Under MBIPV Project*

Short-term study was conducted at 125 sites to identify the technical problem occurred. During the survey, 12 sites identified facing due to technical problem such as inverter shutdown/damage, datalogger/sensor failure and PV meter which not accessible are the most frequently occurred on the GC BIPV system in Malaysia's real environmental [11].

In the past 5 years, approximately 5 sites having problems attributed to the inverter malfunction as tabulated in TABLE IV. It can be seen that Fronius inverter's type and PV module technology using monocrystalline were more dominant compared to the other four manufacturer inverters and PV modules as the most frequent result of inverter failures.

Other than that, one case has been detected with inappropriate sizing between inverter and PV array. The appropriate of PV array power needs to be match with respect to the inverter nominal power to prevent the inverter clamping the power generated by the PV array. If the PV array output is larger than the inverter's nominal power, the inverter will decrease the output power from PV array to the inverter's nominal power rated. If inverter is oversized than PV array output, the inverter will operate at a lower efficiency rate when a low irradiance occurred. One site has been taken as a case study that affected by sizing of PV array and inverter problem due to the sizing ratio as shown in Figure 8 below.
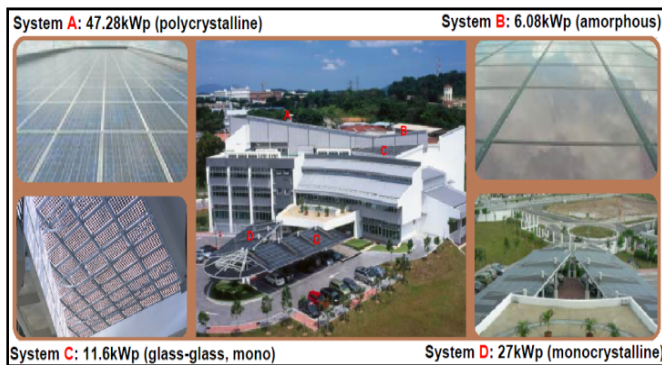
Figure 8. GreenTech Malaysia-Pack B, Bandar Baru Bangi, Selangor, Malaysia [3].

TABLE III.   CASE :GREENTECH –PACK B MALAYSIA'S SYSTEM SPECIFICATION.

Subject	Specification
Site	GreenTech Malaysia– Pack B
Type of PV system	Grid-connected
Mounting type	Building Integrated
Nominal power	6.08 kWp (95 units)
PV module type	a-Si / Kaneka GPA 064 (64W)
Inverter brand	Fronius IG 60

Malaysia Green Tech Corp (formerly known as Malaysia Energy Centre) has equipped of 92 kWp with four different GC PV systems by using various PV module technologies [4]. One of the PV module technologies has been installed with 6.08 kWp using amorphous-Silicon (a-Si) as tabulated in TABLE III. This system has operated and still monitored from June 2007 until today. In 2010, about 6 months facing unavailability of data due to inverter failures after 42 months

of operation. The red colour indicates a "warning" and the inverter replacement had been made on 10th December, 2010 as shown in Figure 9 below.
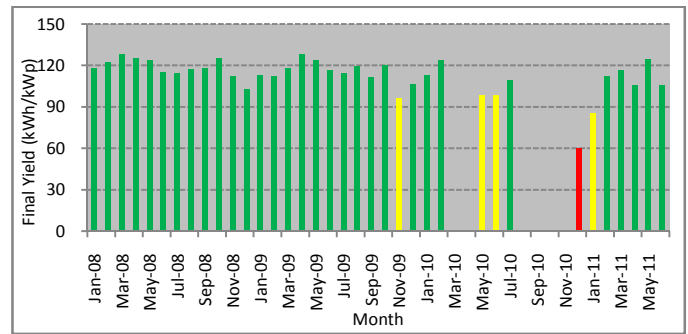
*Green –Good , *Yellow –Required Attention, *Red- Warning,

Figure 9.   Monthly profiles of final yield generated at GreenTech Malaysia-Pack B using a-Si modules.

In addition, it can be seen that the inverter damage in Selangor's state at GreenTech Malaysia – Pack B due to the inappropriate combination of the PV array and inverter sizing as tabulated in TABLE IV. The under-sizing of the inverter around 25% of the PV array power rated has caused "clamping" phenomenon by the inverter due to the power has exceeded the maximum DC power to the inverter rated was detected in this system. Typically, thin-film modules exhibit initial higher to the power level in between 18% to 38% in the first installation before reach their stabilized period based on the module manufacturer's data sheet [12-13]. So, the system designer needs to take consideration on the thin-film module's behavior and characteristic, especially in installation at equatorial and tropical climate.

Other than that, several issues also occurred in GCPV system in Malaysia's field condition as illustrated in Figure 10 and Figure 11 below.

Figure 10. Environmental problem occurred on the PV modules in Malaysia.

Figure 11. Technical problem occurred on BOS component in Malaysia.

TABLE IV.   INVERTER AND SYSTEM FAILURE ISSUES OF GRID-CONNECTED UNDER MBIPV PROJECT IN MALAYSIA [3].

State	Site	Specification	Inverter			System
			Shutdown	Defect	Sizing	Failure
Kuala Lumpur	Private Bungalow House, Taman Duta	Sharp ND130TJ (mc-Si) /SMA SMC 10000TL	May 2009.			
	Private Bungalow House, Taman Desa	Solarworld 175 (mc-Si) / Fronius IG 40	March 09, June - Sept 10.			
Selangor	GreenTech Malaysia -Pack A1, Bangi	Mitsubishi PVMF120EC3 (p-Si) /Fronius IG 500	July 2010			Mac , Apr 10
	GreenTech Malaysia -Pack B, Bangi	Kaneka GPA 064 (a-Si) / Fronius IG 60	Aug – Nov 10	10th Dec 2010	√	Mac , Apr 10
	GreenTech Malaysia- Pack D, Bangi.	Sharp NUSOE3E (mc-Si) / Fronius IG 300	June 09 , Oct & Dec 10, Feb 11.			Mac , Apr & Sept 10
	Monash University , Sunway	Kaneka GPA 064 (a-Si) / Solarmax (2 units)				March 09
Putrajaya	Show Unit Bungalow House Lot 2422	Sharp NUS0E3E (mc-Si) / Fronius IG 60	Sept - Dec 10			August 09

C. Summary Of Problem Identification

There are 10 problems show the influence on the GCPV system performances as can be seen in Figure 12 below. About 40% of environmental problem occurred due to the shading and low energy crisis issues. In contrast, 8 problems recorded in technical issue in terms of inverter shutdown was approximately 15%, as well as PV meter not accessible was 25%, and other failures about 20% involved inverter defect, datalogger/sensor, Surge Protection Device (SPD), system failures, site renovation and inverter sizing during short-term investigation had done.
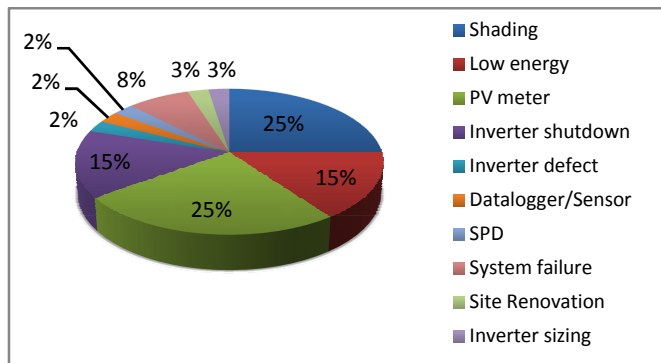
Figure 12. The percentage of technical and environmental issues under MBIPV project based PVs site installation survey (January 2007- June 2011).

VI. CONCLUSION

Malaysia is fully gifted with RE sources and is environment-friendly in nature. However, RE capacities are grossly under-utilized, particularly from biomass and solar energy. The Malaysian Government recognizes the potential of RE as an alternative to ensure the sustainability of energy resources. Thus, the green technology policy was launched to promote the utilization of low carbon energy and technology of which RE has been identified as the promising green energy option for Malaysia. The development of GCPV system applications shows the installation of PV increases from year to year after MBIPV launched. In addition, the problems and challenges in handling of the GCPV system in Malaysia are recognized. After 5 years experience in monitored the performances, the short-term investigation found the majority issue occurred about 60% of technical whereas 40% of portion

came from the environmental problems over the monitoring period. All issues are identified through monitoring process had been done by PVSMC.

REFERENCES

[1] (2012) Malaysia Building Integrated PV (MBIPV) Project portal, [Online]. Available: http://www.mbipv.net.my

[2] (2012) Sustainable Energy Development Authority (SEDA) portal, [Online]. Available : http://www.seda.gov.my

[3] (2012) Photovoltaic System Monitoring Centre (PVSMC), [Online]. Available: http://pvmc.uitm.edu.my/solarman

[4] (2012) Malaysian Green Technology Corporation Official Site. [Online]. Available: http://www.greentechmalaysia.my/

[5] A. H. Haris, "Grid-Connected and Building Integrated Photovoltaic: Application Status & Prospect for Malaysia", Master Builders Journal, vol. 3, pp. 91-95, 2006.

[6] S. Sairan, Electricity Utility Research And Development In Renewable Energy. Presented at Asean Residential School in Electric Power Engineering (ARSEPE07), 2007.

[7] K. Sopian, M.Y. Othman, B. Yatim and W. R. W. Daud, "Future Directions in Malaysian Environment Friendly Renewable Energy Technologies Research and Development". ISESCO Science and Technology Vision , Vol. 1 pp.30-36, May 2005.

[8] D. Ruoss, MBIPV – Malaysian Building Integrated PV Application Technology, Final Report GEF PDF-B (Project Development Facility – Block B). 2nd August 2004.

[9] L. C. Haw, "PV In Buildings For Malaysia: Prototype Solar House". 5th International Seminar in Sustainable Environment and Architecture (SENVAR), 2004.

[10] Pusat Tenaga Malaysia , "Market Status Report on BIPV in Malaysia". January 2007.

[11] M. Z. Hussin, A. Yaacob, Z.M. Zain, S. Shaari, A.M. Omar. "Status of a Grid-connected MBIPV Project in Malaysia", 2011 3rd International Symposium & Exhibition in Sustainable Energy & Environment (ISESEE), Melaka, Malaysia, pp. 4-11, 2011.

[12] Kaneka Corporation, Specification and installation manual for Kaneka Thin Film PV: GPA type. (issue 06/2004).

[13] Kaneka Corporation, Specification and installation manual for Kaneka Thin Film PV: GSA type. (issue 04/ 2006).

Hybrid Approach using Correlation and Morphological Approaches for GFDD of Plain Weave Fabric

V.Jayashree

Electrical Engineering Department

Textile and Engineering Institute

Ichalkaranji, India

jayashreevaddin@gmail.com

Shaila Subbaramn

Electronics Engineering Department

Walchand College of Engineering

Sangli, India

Shaila.subbaraman@walchandsangali.ac.in

Abstract—This paper presents the Hybrid Approach (HA) using correlation followed by Morphological Approach (MA) to detect the micro natured defects in fine grey plain weave fabrics. The fundamental method of suspected object detection in an image is achieved using Correlation Approach (CA) to determine correlation of defect template with defect image to detect Region of Interest (ROI). The FFT based correlation technique assisted by appropriate threshold was followed by the MA using suitable Structuring Element (SE) to filter and retain only the defect region and to facilitate the reduction of False Alarm Rate (FAR). Three methods viz., MA, CA and HA were tested and compared for Grey Fabric Defect Detection(GFDD) of plain weave defects of varying sizes. Performance comparison of the three algorithms was carried out by adopting simple binary based Defect Search Algorithm (DSA) as a last step in the experimentation to detect defects. The progressive improvement observed in terms of % Overall Detection Accuracy (ODA) for these three methods was from 22 to 79 for benchmark defect samples and from 25 to 93 for normal samples. Experimentation results on HA showed 3.5 fold increase in ODA for defect samples and about 3.7 fold increase in ODA for normal samples. Though HA method could detect different defects, the drawbacks of this are the selection of the appropriate defect template and selection of SE which is tedious. The details of the experimentation and the results thereof are presented in this paper.

Keywords- Correlation, Template, Texture Periodicity, Morphological, Structuring Element, Hybrid Approach, GFDD.

I. INTRODUCTION

The in process inspection of the grey fabric is carried out either manually or opto-electronically. Both methods fail to detect subtle micro-natured fabric defects. Hence the computer based automatic inspection systems have gained widespread acceptance for quality checks and control in recent years. Fabric surface is characterized by regular textural features. So occurrence of any fabric defect over regular texture background can be monitored and quantified by state-of-the-art methods using Image Processing (IP). From the literature survey, it is seen that, Correlation Approach (CA) has been used extensively especially for pattern matching in various fields by many researchers [1-7] but not so for Fabric Defect

Detection (FDD). Bodnarova et al. have used the Correlation Coefficient (CC) from multiple templates to generate a correlation map for defect declaration and have achieved correlation sensitivity of 95% for fifteen defective fabric images with five different defects [8, 9]. However the application of this method to detect microstructure defects and multiple grey fabric faults seem to be rarely addressed and suggest its rigorous testing against multiple flaw types. All these prompt that; CA can be used to locate features within an image corresponding to imperfections in regularly textured backgrounds. Motivated by this fact, the authors have used this approach to detect the various grey fabric defects in fabrics of different specifications. It was shown that correlation (CR) method can support the detection of variety of microstructure defects of different sizes for plain grey fabric with different specifications [11].

Spatial domain Morphological Approaches (MA) was proposed by different researchers to address the problem of FDD assisted by different methods for selection of structuring element (SE) [12-16]. Their experimental results on MA indicated the suitability of MA for FDD [12, 16], no obvious improvement over other available approaches [13] and reported detection rate of 60% in [14] and 97% in [15].Though correlation technique can detect variety of defects, its dependency on the right defect template, template size, thresholding and false detection of normal region as defective puts limitation on its use [11]. This indicates that there is a need of further investigation to reduce False Alarm Rate (FAR). Hence MA was tested on plain weave for FDD. However it did not produce satisfactory results. This led us to the intuition that hybridization of CA and MA may produce better results. Hence hybridization of MA and CA was carried out as one possible solution that takes the advantage of the individual methods. The CA followed by MA can filter the ROI. These two methods used in cascade provide double surety of defect being recognized as defect and normal as normal and thus reduce False Alarm Rate (FAR). Therefore The Hybrid (CA+MA) method was experimented for plain weave GFDD which is not cited in literature so far. The experimental details and results thereof on MA, CA and HA are discussed further.

II. THEORITICAL BACKGROUND

The work related with this paper is based on the idea of applying the combined approach viz., HA to test its usefulness for texture inspection. Hence, brief theory on fabric structure, types of defects, correlation theory and morphological filter operations for detection of localized microstructure defects in plain grey fabrics is explained in the following sections.

A. Fabric Patterns and Plain Weave Fabric Defects

The Woven fabrics is composed of longitudinal or warp threads (ends) and transverse or weft threads (picks), interlaced with one another according to the class of structure and form of design that are desired. In simple normal plain weave pattern, the ends and the picks intersect at right angles in the cloth. Any weave repeats on a definite number of ends and picks. Hence one repeat unit of weave pattern is important [19]. The kind of defects considered for this study are the plain weave micro natured defects such as thick place, loose-weft, double pick and warp break of varying size. Thick place occurs due to the yarn uniformity variation by more than 50%, loose-weft is a single filling yarn under insufficient tension where as the double pick defect is more specifically a unintentional weaving of two weft wise threads and warp break defect indicates break in warp thread for a small length of warp [20].

B. Template Matching using Correlation

Correlation can be used to locate features within an image using templates. Hence it is capable of concentrating to smaller local feature variation which can occur due to loss of fabric texture regularity and change in the intensity levels in the faulty region of the image. Therefore CA can be used for addressing the problem of detecting the dissimilar defect region from a fine regularly repeating structure of grey fabric. Thus, we were motivated to use CA as a tool to detect microstructure defects. Given an image $f(x, y)$, the correlation problem is to find all places in the image similar to the given sub image $h(x, y)$ or template $w(x, y)$. The 2-D spatial correlation is given as below,

$$f(x,y) \circ h(x,y) \Leftrightarrow F^*(u,v)H(u,v) \qquad (1)$$

In other words the correlation method involves the comparison of replica of an object of interest to all unknown objects in an image field [17]. To obtain good accuracy in template matching, a common procedure is to produce difference of image template and image field at all points of the image field. An object is deemed to be matched when the difference is smaller than some established level [18]. The usual mean square error difference is defined as,

$$D(m,n) = \sum_j \sum_k \left(F(j,k) - T(j-m,k-n) \right)^2 \qquad (2)$$

Here $F(j, k)$ represents the image field to be searched and $T(j,k)$ denotes the template. Limitation search is restricted to the overlap region between the translated template and the image field. A template match is said to exist at coordinate by comparing if,

$$D(m,n) < L_D(m,n) \qquad (3)$$

Here is $L_D(m, n)$ threshold value. Equation (2) is expanded to yield,

$$D(m,n) = D_1(m,n) - 2D_2(m,n) + D_3(m,n) \qquad (4)$$

where,

$$D_1(m,n) = \sum_j \sum_k [F(j,k)]^2 \qquad (5)$$

$$D_2(m,n) = \sum_j \sum_k [F(j,k)]T(j-m,k-n) \qquad (6)$$

$$D_3(m,n) = \sum_j \sum_k [T(j-m,k-n)]^2 \qquad (7)$$

Here $D_1(m, n)$ is image energy over window which varies slowly across the image field, while term $D_2(m, n)$ is cross correlation between the image field and the template and $D_3(m, n)$ represents the summation of the template energy which is constant valued and independent of the coordinate (m, n)[19]. At the coordinate location of template match, the cross correlation becomes very large to yield small difference as per (2) [17].

C. Marphological Approach

Morphological Operation (MO)s can find the image geometry using the SEs of certain shapes. Like spatial convolution, the SE is moved across the input image pixel by pixel and by placing the resulting pixels in the output image. Each pixel in the input image, is logically compared with its neighbors against SE to determine the logical value of output pixel. Most of the MOs are carried out considering image as 2-D integer space Z^2, where each pixel is considered as tuple (2-D vector) whose coordinates are (x, y) [18]. Let Z^2 denote the set of all ordered pairs of elements and A denote the set of pixels positions of the fabric image in the whole Euclidean space E. Generally used MOs are:

- Dilation: Dilation of A by B is the set of all the displacements, z, such that B and A overlap at least by one pixel element. Dilation can be used for bridging the gaps in the image is given by,

$$A \oplus B = \{z \mid [(\hat{B})_z \cap A] \subseteq A\} \qquad (8)$$

Where, φ denotes the empty set. This equation is based on obtaining the reflection of B about the origin, $\hat{B}$ and shifting the reflection by z.

- Erosion: It can be used for removing the irrelevant details in the binary image and is expressed as,

$$A \Theta B = \{z \mid [(B)_z \subseteq A\} \qquad (9)$$

- Open and close: Combination of operations e.g. erosion followed by dilation is called as morphological opening while dilation followed by erosion is called as closing operation and are given as below.

$$\text{open: } (A, B) = A \circ B = (A \Theta B) \oplus A \qquad (10)$$

Jathar Textiles, Ichalkaranji, Maharastra, India.

close: $(A, B) = A \bullet B = A \oplus B) \Theta A$ (11)

- Boundary Extraction: The boundary of a set, A denoted by _(A) can be obtained by first eroding A by suitable structuring element, B and then taking the set difference between A and its erosion as given by,

$$\beta(A) = A - (A \Theta B) \qquad (12)$$

- Region Filling: It is based on a set of dilations involving complementation and intersection. The filling operation is established using 4/6-connected background neighbors for 2-D/3-D MO. Adopting the convention of '0' for non background points and '1' for the seed point, the region filling is expressed as,

$$X_k = (X_{k-1} \oplus B) \bigcap A^c, \text{ for } k = 1, 2, 3.... \qquad (13)$$

Where, X0 = p is the seed point and B is the SE. The algorithm is iterated till convergence is achieved. The set union of X_k and A contains the filled set and its boundary limits the result inside the region of interest [18]. For all the above MO expressions A and B refer to the fabric image A and structuring element respectively.

III. IMAGE ACQUISITION

To carry out the FDD, both defective and normal samples of plain weave shirting fabric with different specifications as shown in Table I were collected from the textile industry. Motorized Micro-stereoscope with optical magnification was used as an imaging device. The images of the fabric samples were taken after standardizing the Motorised Zeiss Sterio microscope for magnification of 12,0x, focus of 19.0 and depth z of -1.45mm. The resolution of the image obtained was 1280x1040 pixels which were then reduced to 512x512 pixels using Microsoft Digital image editor.

TABLE I.   DETAILS OF PLAIN WEAVE FABRIC SPECIFICATIONS

Sr.No	Specifications	Defect Type
S_1	$\dfrac{132 \times 120}{60 \times 60}$	1. Normal 2. Warp break
S_2	$\dfrac{124 \times 64}{30^s \times 150 D}$	1. Normal 2. Double pick 3. Loose weft
S_3	$\dfrac{144 \times 144}{80 \times 80}$	1. Normal 2. Double pick 3. Thick place

IV. MA, CA AND HA FOR FDD

To arrive at an appropriate method for GFDD, two traditional approaches viz., MA and CA were verified. Based on these results and the intuition developed, a modified method viz., hybrid approach was tested. Accordingly the experimental procedure and observations using MA, CA and HA for GFDD are presented in this section.

A. Morphological Approach for Plain Weave Defects

The experimental procedure using MA only was followed on benchmarking plain weave fabric defects such as warp-

break(S_1), thick place (S_3) shown (columsC_1/C_2 in Figure1.a/Figure1.b) and double pick(S_2) (column C_3 in Figure1.b). Two sets of observations were taken on the fabric images on the histogram equalized image shown in row R_l of Figure1.a and Figure1.b.
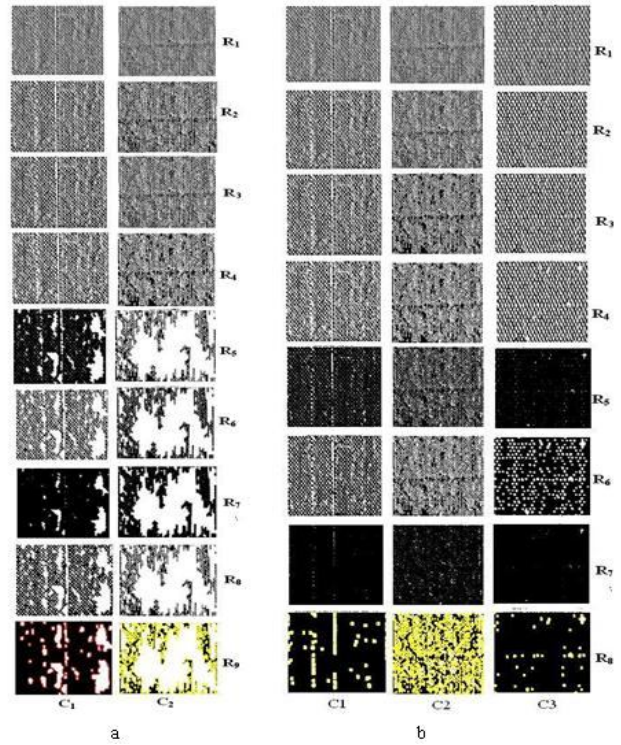
Figure 1.   Results of MA Applied on Defect Samples

a. With image filling operation added in sequence
b. Without image filling operation

TABLE II.   RN /RN-1- A PARTICULAR MO ON IMAGE OF ROW RN WITH IMFILL AND RN-1 WITHOUT IMFILL TABLE TYPE STYLES

Rn /Rn-1- A particular MO on image of row Rn with imfill and Rn-1 without imfill.		
Fig a	Fig b.	MA Operations
R_1	R_1	Thresholded Gray image
R_2	R_2	Opening Image after thresholding using disk SE
R_3	R_3	Closing image of row R_2
R_4	-	Filling small holes in row R_3
R_5	R_4	Erosion of image in row R_4/R_3 with SE eye
R_6	R_5	Dilation of images in row R_5/R_4 with SE eye
R_7	R_6	Erosion of image in row R_6/R_5 with SE rectangle
R_8	R_7	Dilation of images in row R_7/R_6 with SE rectangle
R_9	R_8	ROI detection of images in row R_8/R_7

The sequence of morphological operations followed to obtain Figure1.a and Figure1.b are as shown in Table II. The first set consisted of all morphological operations with image filling (imfill) operation while the second set consisted all steps except image filling operation respectively.

Following observations are made for the MA application for normal and defective plain grey fabric samples;

- Observing Figure1.a for warp break and thick place defect, it is clear that, MO operation with imfill operation fails to detect ROI for plain weave pattern. This is due to the fact that, plain shirting fabric being made out of very fine yarn count makes the repeat unit of weave pattern very small in area and very dense in texture nature. This makes the image filling operation to remove texture information more than expected causing too high degradation of ROI (Refer row images from R_5 to R_7 of Figure 1.a.)
- Referring to Figure 1.b it is seen that, though elimination of imfill operation from MO sequence results in improvement in detection of ROI compared to Figure 2.a but it does not give satisfactory results (Refer row R8 of Figure1.b).

Comparing images in row R_8 of Figure1.b for the warp break, thick-place and double pick, it is seen that detection of ROI for warp break is better than that for double pick while it is the worst for thick-place samples. The above observations of MA on plain weave defects suggest that, MO operation fails for plain weave defects in detection of normal region and defect region. Image filling (imfill) operation does not work effectively for noise removal. This is attributed to the fact that the noise addition due to protruding yarn for defective region of plain weave fabric made of fine count is almost feeble even in the defect region. Accordingly the steps of MA need modification to achieve better results for detection of ROI and hence to achieve good ODA for both normal and defective samples. Looking at the poor performance of MA, other traditional IP method viz., CA was experimented for plain weave FDD.

B. Correlation Approach for Plain Weave Defects

For plain weave FDD using CA, four different kinds of plain weave defects including defects of subtle nature were considered as shown in Figure2. With reference to Figure2, column C1 depicts images of CA on warp break (S_1), double pick (S_2), thick-place (S_3) and loose-weft (S_2) defects respectively from top to bottom. The defect templates corresponding to these defects are as shown in column C_2 while the result of correlation convolution of the defect images with their respective defect template are shown in column C_3. The defect area for each kind of defect obtained after thresholding of Correlation Coefficients (CC)s are as shown in column C_4 of Figure 2. The threshold which best approximates the actual defect location and area was obtained after many trials and was used further for experimentation on other samples with the similar kind of defects.

Following are the observations on CA.

- Observing the various results of CA in Figure 2 it is seen that, the results are not very satisfactory, since there is false identification due to Normal Region Identified as Defective (NRID) at many places of an image. This is quite obviously seen in all result images in column C_4 of Figure2.
- There is a need of different kind of template for each. There is a need of different kind of template for each kind

of defect considered. Also it is been found that, the defect and its area identified were dependent on the kind and size of the defect template which makes GFDD a tedious and time consuming process.
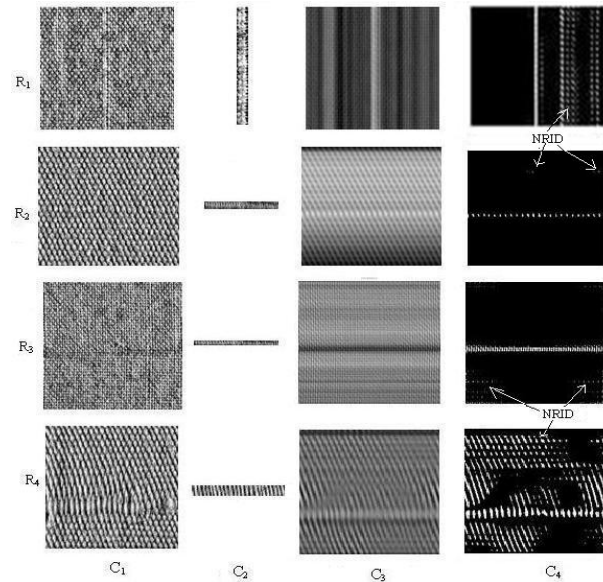
Figure 2. Results of Correlation Approach

Though correlation technique can detect variety of defects, its dependency on the right defect template, its size, thresholding and false detection of normal region as defective, puts limitation on its use. The other disadvantage is that fabric without defect is detected as defective which is objectionable. All these are indicative of further exploration on CA for GFDD and this issue demands further research work in this area.

C. Hybrid Approach(HA) using CA and MA

From careful observation of the results on FDD on plain weave, it seen that both, MA and CA can identify the defect region i.e. ROI when applied independently, but there is more probability of the fabric normal region also getting identified as the defective as reported in [1,2] and also verified in [4] for micro natured defects. To overcome this drawback and to utilize the advantages of correlation property of CA and filtering property of MA to detect ROI, the next drive in this research was a hybrid approach for GFDD. In this direction hybrid algorithm using combination of CA followed by MA was developed and experimented on plain defects. The results of this HA on fabric plain weave are presented below.

1) Hybrid FDD Algorithm

In HA, after preprocessing of the fabric image, each image was subjected to CA first using a defect dependent template for each kind of defect. Further suitable MOs as explained further were applied after thresholding the CCs. The various steps followed for HA are as shown in Figure 3. After converting fabric RGB image to gray image, it was subjected to the adaptive histogram equalization for contrast adjustment.

After choosing the proper defect template from histogram adjusted image, its size was made equal to the size of the image to make CA more efficient.

In CA algorithm, the correlation of the template image with the original image was computed by rotating the template image by 180^0 and then using the FFT based convolution technique described by fast convolution. This is for the reason that, if the convolution kernel is rotated by 180^0 then the convolution is equivalent to correlation. CCs were then thresholded using the simple statistics of the coefficients such as mean and maximum value to retain the defect region. It is necessary to keep the threshold at optimal value to extract ROI from all similar defect images of one class of fabric and to avoid any defect region being missed. But this increases FAR for normal region. Hence filtering property of MA was used on the thresholded CC images to retain only defect region. Suitable area of SE for opening operation of MO was selected after many trials to retain the seed of the pattern which was subsequently closed by the same SE. The grey fabric consists of protruding yarn resulting in the form of small pepper noise which needs (which gets exaggerated in fault region) to be removed by filling operation. The next steps followed for extracting ROI were similar to steps used in MA. The implementation results of this hybrid approach are discussed in the following subsection.

2) Experimental Results

For experimentation on HA, the defects that were used for GFDD using CA were considered for study with additional double pick defect of class S_3. The columns C_1, C_2, C_3 and C_4 of Figure 4 depict the defect gray image, results of MA, CA and the hybrid CA and MA approach respectively. Focusing the discussion to column C_2 and C_3 of this figure 4 on detected ROI for all plain weave defects, it is seen from images that, neither MA nor CA can independently identify the defective region correctly. Normal region is identified as defect at several places indicating more FAR. This is for the reasons that MA cannot catch purely ROI for plain weave defects when count of the fabric becomes fine. Also in CA as stated earlier, defect area identification is mainly dependent on the thresholding value of CCs. The threshold that could best detect the defect region was chosen and used for further experimentation to avoid any chance that defect free region could be detected as defect region. The problem of NRID i.e. identifying the normal region as defective at several places as seen in images of columns C_3 and C_4 of Figure4 is almost eliminated by the successive CA followed by MA (Refer to result images in column C_4 of Figure 4). Similar results were observed for the other kind of defects of plain weave defect samples.

3) Testing of Data Set using Defect Search Algorithm

To test the performance of the MA,CA and HA a simple binary based Defect Search Algorithm (DSA) to find presence or absence of defects was used with zero value indicating the true normal sample and one value indicating the true defect sample. Testing was carried out on 115 benchmark normal samples and 75 warp break samples belonging to class S_1.

DSA was applied on the MA result images in column C_3 of Figure1.b, CA result images in column C_4 of Figure 2, and HA result image in column C_4 of Figure 3. Table III shows result summary of the experiment conducted.
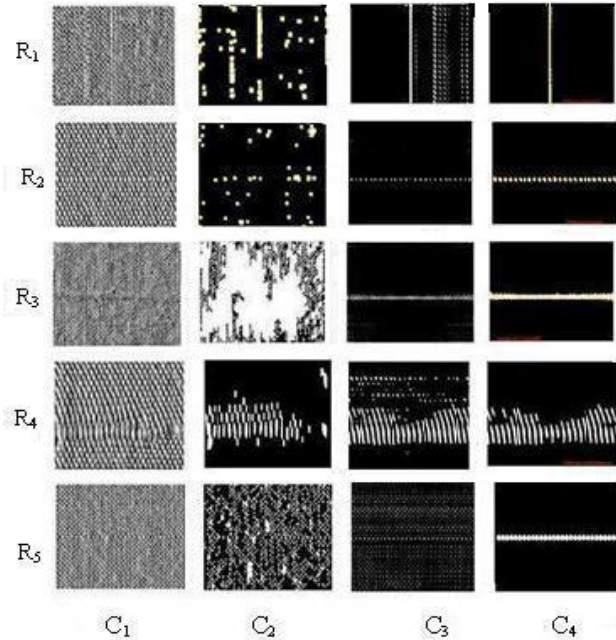
Figure 3.   Result Images Showing Deffect Region Detected by MA,CA and HA Algorithms respectively along C_2, C_3 and C_4 columns.

TABLE III.   DEFECT IMAGES AND IP ALGORITHM APPLIED (COLUMNWISE) ON DIFFERENT CLASS SAMPLES

Weave Type	Defect name (Row wise)	IP operation (Column wise)
Plain	R_1: S_1-Warp break	C_1:Gray Image
	R_2:S_2 -Double pick,	C_2:Result of MA
	R_3:S_3-Thick place	C_3: Result of CA
	R_4:S_2-Looseweft	C_4:Result of HA
	R_5:S_3-Double pick	

The performance of the MA, CA and proposed HA were evaluated in terms of Overall Detection Accuracy (ODA) using DSA.

$$ODA\ \% = \frac{N_{CD}}{N_T} \times 100 \tag{14}$$

Where, N_{CD} is number of correctly detected samples and N_T is the total number samples of that class. Results of MA,CA and HA Method of Plain Weave FDD for Class S_1.

TABLE IV.   RESULTS OF MA,CA AND HA METHOD OF PLAIN WEAVE FDD FOR CLASS S_1

S_1 Class Defect /Normal	*Number of Samples*	MA	CA	HA
		ODA%	*ODA%*	*ODA%*
Warp break	78	28.2	55.1	79.1
Normal	115	25.3	45.9	93.4

The results of this procedure are presented in Table IV. The results showed progressive improvement in ODA% of 28.7, 55.1 and 79.1 for warp break defect samples and 25, 46.2 and 93 for normal samples subjected to MA, CA and HA approaches respectively. Hence it is evident from these results that, HA is more encouraging than MA only and CA only.

V.   RESULTS AND DISCUSSION

The overall summary of observations is as below.

- Size of SE for different stages of MA operation needs to be selected by many trials. So there is a need for establishing some relation for selection of SE size with pattern.
- Correlation approach using template matching can identify defects in different kind of fabric of plain weave patterns consisting of single and multiple defects in a single image as well as similar multiple defects located at different locations of the image.
- The exact area of fault indicated by this method depends on how best the template for a given type of fault is chosen and also on the threshold value.
- From the performance of HA, it is found that HA works well for identifying the defect and normal region of plain weave pattern more accurately than MA and CA. Also 3.5 fold increase in % ODA results obtained compared to only MA and CA is promising.

Thus the hybrid approach of CA and MA combined has resolved the problem of only CA or MA misidentifying the normal region as defect region. The results of reduced False Alarm Rate indicate that HA is a promising method for GFDD. Such hybrid method does not seem to have been reported in the literature so for. Therefore we claim this as the contributions of our research work.

VI.   CONCLUSION AND FUTURE SCOPE

Hybrid Approach of combining CA and MA resulted into reducing FAR considerably. This is attributed to the fact that the benefits of CA and MA doubly take care of retaining only ROI due to the inherent property of CA, detecting template matched region and MA aiding to filter any remaining repetitive fabric pattern after CA and also aiding the detection of defect region of interest with the help of appropriate SEs. Further the information such as number of defects detected in a given area and the area occupied by the defect can be obtained as needed for grading standards of grey fabric. It is further recommended that a modified approach to HA should incorporate automatic selection of size of SE for MA while defect independent template for CA.

REFERENCES

[1] Toshiyuki Tanaka and Yoko Murase " Classification of Sarcomas using Textural Features" Proc. of the 22nd Annual EMBS international conference, July 23-28,2000, Chicago IL, pg.682 -685.

[2] Asker M. Bazen, Gerben T.B Verwaaijen, Sabila H.Gerez, Leo P.J Veelenturf and Bernad Jan van der Zwaag, "A Correlation Based Fingerprint Verification System," ISBN:90-73461-24-3, Nov 30-Dec-1 2000, pg.205-213.

[3] Cunjian Yang1,2, Jiyuan Liu2, He Huang, Shanshou Cao, "The Correlation Analysis of the LANDSAT TM Data and Its Derived Data with the Biomass of the Tropical Forest Vegetation," 0-7803- 7929- 2/03(C) 2003 IEEE, pg. 2583-2585.

[4] Jong-Un Won, Yun-Su Chug', In-Soo Kim', Jae- Gark Choi2, and Kil- Houm Park, "Correlation Based Video-Dissolve Detection" IEEE 0- 7803- 772 9103/2003, pg.104-107.

[5] Yi Li, Zhiyan Wang, Member, IEEE, and Haizan Zeng, "Correlation Filter: An Accurate Approach to Detect and Locate Low Contrast Character Strings in Complex Table Environment," IEEE Transations on Pattern Analysis And Machine Intelligence, VOL. 26,NO.12,pg.1639- 1644, Dec.2004.

[6] K. V. Arya , P. Gupta , P.K. Kalra , P.Mitra, "Image Registration using Robust M-Estimators," 0167- 8655/2007 Elsevier doi:10.1016/j.patrec, 2007-05-006., pg .1958-1967.

[7] Minglei Sun, Guohui Qiao, Rong Zhang, Guanghua Zong, " Characteristics of Independence on Image Gray Level in NCCO Applications," International Conf.on Information Technology and Computer Science IEEE, DOI 978-0-7695-3688-0/09 2009.

[8] Bodnarova A., Bennamoun M.,and Kubik K.K. "Defect Ddetection in Textile Materials Based on Aspects of the HVS" Proc. IEEE SMC'98 Conf.,San Diego (U.S.), pp. 4423-4428.

[9] Bennamoun M. and Bodnarova A. "Automatic Visual Inspection and Flaw Detection in Textile Material: Past, Present and Future'" Proc. IEEE Conf. SMC 1998, pp. 4340-4343.

[10] A.Brodnarova, M.Bennamoun and K.K Kubik, "New Suitability Analysis of Techniques for Flaw Detection in Textiles using Texture Analysis," Pattern Analysis and applications: 2000 Springerverilog London Limited pp.254-266.

[11] V.Jayashree and Dr. S.Subbaraman, "Correlation Approach for Grey Fabric Defect Detection and Defect Area Measurement," Proc. of National Conf. on Recent Trends in Electronics Engg. 2-4 Dec 2011,pg.420-425.

[12] Y. F. Zhang and R. R. Bresee, "Fabric Defect Detection and Classification using Image Analysis," Text. Res.J., vol. 65, pp. 1-9, Jan. 1995.

[13] B.Goswami and A. K. Datta, "Detecting Defects in Fabric with Laser- Based Morphological Image Processing," Text.Res.J.,vol.70, pp.758- 762,Sep 2000.

[14] Che-Seuing Cho, Byeong-Mock Chung, Member, IEEE, and Moo-Jin Park.,"Development of Real Time Vision Based Fabric Inspection System," IEEE Transactions on Industrial Electronics, VOL 52,No.4,August 2005.

[15] K.L Mak,P.Peng and H.Y.K.Lau, "Optimal Morphological Filter Design for Fabric Defect Detection", 0-7803-9484-4/05,2005 IEEE, pp.799-804.

[16] Jayanta Chandra,Pradipta K.,Banerjee Asit, K. Datta, "Morphological Reconstruction Operation for the Detection of Defects inWoven Fabric" 2008,IEEE R-10 Colloquium and the 3rd ICIIS, Kharagpur, India, Dec.8-10.

[17] William K Pratt , Digital Image Processing,Wiley 3rd Edition ,pp. 625- 631,2003.

[18] Rafael Gonzalez and Richard E.Woods, Digital Image Processing,2nd Edition, Pearson Education, 2004, pp.519-537.

[19] Z.J.Grosciki. Watson's Textile design and Colour Elementary Weaves and Figured Fabrics, 7th Ed.,WP,2004.

[20] ASTM standard Designation D 3990-93, Standard Terminology Relating to Fabric Defects, pp.205- 211.

Web-based Student Attendance System using RFID Technology

Murizah Kassim, Hasbullah Mazlan, Norliza Zaini, Muhammad Khidhir Salleh
Faculty of Electrical Engineering
Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia
murizah@salam.uitm.edu.my, hasbullahmazlan@gmail.com, drnorliza@salam.uitm.edu.my, khidhir193@gmail.com

Abstract— **This paper describes the development of a student attendance system based on Radio Frequency Identification (RFID) technology. The existing conventional attendance system requires students to manually sign the attendance sheet every time they attend a class. As common as it seems, such system lacks of automation, where a number of problems may arise. This include the time unnecessarily consumed by the students to find and sign their name on the attendance sheet, some students may mistakenly or purposely signed another student's name and the attendance sheet may got lost. Having a system that can automatically capture student's attendance by flashing their student card at the RFID reader can really save all the mentioned troubles. This is the main motive of our system and in addition having an online system accessible anywhere and anytime can greatly help the lecturers to keep track of their students' attendance. Looking at a bigger picture, deploying the system throughout the academic faculty will benefit the academic management as students' attendance to classes is one of the key factor in improving the quality of teaching and monitoring their students' performance. Besides, this system provides valuable online facilities for easy record maintenance offered not only to lecturers but also to related academic management staffs especially for the purpose of students' progress monitoring.**

Keywords-RFID, students attendance, online system, web, sensor

I. INTRODUCTION

The most common means of tracking student attendance in the classroom is by enforcing the students to manually sign the attendance sheet, which is normally passed around the classroom while the lecturer is conducting the lecture. There are numerous disadvantages of using such system. For instance, lecturers with a large class may find the hassle of having the attendance sheet being passed around the class and the manual signing of attendance by students are burdensome and most likely distract them from teaching and getting full attention from the students [1]. Besides, as the attendance sheet is passed around the class, some students may accidentally or purposely sign another student's name. The first case leads to a student missing out their name, while the latter leads to a false attendance record. Another issue of having the attendance record in a hardcopy form is that a lecturer may lose the attendance sheet [2]. As a consequence of that, lecturer can no longer trace the students overall attendance record throughout the particular semester. Apart from that, a lecturer also has limited access to the single-copy record, e.g. only at the work

place. In terms of attendance analysis, the lecturer also has to perform manual computation to obtain the students' attendance percentage, which normally consume a lot of time.

Having said the limitations imposed by the conventional attendance recording system, we propose a solution in the form of an attendance tracking system based on RFID technology. The main idea behind the system is to capture student attendance in a semi-automated way where the students are required to flash their student card at the RFID reader upon entering the classroom. This way, the student ID is instantly captured by the reader, after which the data is sent to the online server for recording purpose [3]. Such system promotes a more organized and systematic student attendance recording. Having the attendance data instantly uploaded to the online server prevents data loss, while allowing the data to be available and accessible to the lecturers or other academic staffs as long as they are connected to the internet [4]. The system also helps to automatically compute the percentage of attendance for each individual student. One related project was referred [5], where an attendance system using the RFID technology was developed in the university environment. This work was taken as guidance in developing our web-based student attendance system by using the RFID technology.

II. SYSTEM DESIGN

The web-based student attendance system using Radio Frequency Identification (RFID) technology promotes semi-automated recording of student attendance, developed especially for the university. In drawing the system design, a number of web-based methods and frameworks have been reviewed and referred [6-7]. Some comparisons performed on the different technologies being used in the similar systems were also learnt [8]. Related research on the system-development and testing were also reviewed to provide guidance in customizing and evaluating the process, while learning the disciplined approach in designing the proposed system [9-13].

Our proposed system consists of three main modules namely the RFID reader module, Data Reporter module and Web server module. Those modules are integrated together in order to allow its full functionality. Each module carries its own functions and special features. The general process flow is illustrated in Figure 1, where all of these components are involved and operational.
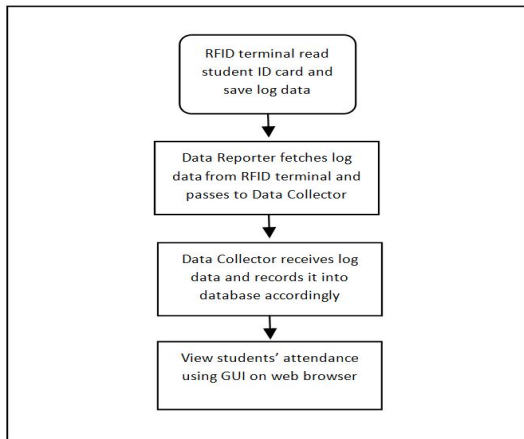
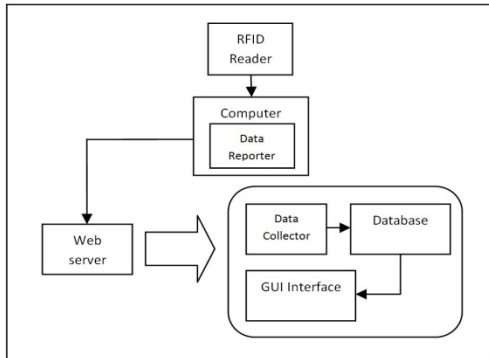Figure 1: The overall student attendance recording process

Figure 2: System Architecture of the web-based student attendance system using RFID

The overall architecture of the system is illustrated in Figure 2, where the three main components are shown. Each of these components will be described in the following sub-sections.

A. RFID Reader and Tag

RFID reader is the device capable of reading and retrieving information stored inside the RFID tags. There are two types of RFID reader, which are the active and passive RFID readers. Active RFID reader can detect an active RFID tag while passive RFID reader can only detect passive RFID tag at a few centimeters away from the reader. The RFID reader being used in the system is a low cost reader for reading passive RFID tags. It operates at 0~ 400C temperatures, 20~80% of humidity, 125 kHz frequency and 12V power supply [7]. The effective detection range of the reader is around 5-8cm.

Each RFID tag has a unique serial number or ID. There are three types of RFID tags which are active, semi-passive and passive. The main difference between these RFID tags is that active and semi-passive RFID tags require internal battery while passive RFID tags do not use any internal battery. Adapted to our scope of work, the student cards being used to identify each individual student are the RFID cards that consist of passive RFID tag, which do not require internal battery [6]. When such cards are passed through the field generated by a compatible Reader, they transmit information back to the Reader [7]. Figure 3 illustrates how data transmission is performed between an RFID reader and a student card.
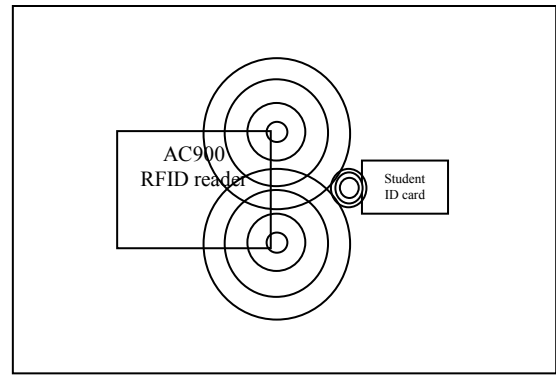
Figure 3: Data Transmission between an RFID reader and a student card

B. Data Reporter

Data Reporter is a component that fetches all logging data from the RFID reader such as the captured student ID, time and date for every 30 minutes interval. The collected data are then passes to the online server, which will record the data into the database. This component should always be kept up and running and needs to be automatically restarted each time the operating system reboots.

C. Web Server

The web server here refers to either hardware (computer) or software (application) that helps to deliver content publicly accessible through the Internet. It provides the web site functionality by accepting requests from the user's browser and responds by sending back HTML documents (Web pages) and files. To enable the system dynamic functionalities, the web server hosts the data collector component, a database and the graphical user interface (GUI) pages enabling online interaction with the system users.

D. Data Collector

The role of the online data collector is to continually listen to incoming data sent by the Data Reporter component. The received log data will then be inserted to the database for recording purpose.

E. Database

A database is defined as an organized collection of data and tailored to our system, our database is employed to mainly store the data captured by the RFID reader. Secondarily the database is also used to store data gathered from the online web-interface, such as class schedule and students personal information. In offering more features to the users, our online-system can manipulate the recorded student attendance record by querying the database for complex data retrieval [14]. This includes automated operation, such as summarizing an individual student attendance by calculating the attendance percentage for a particular course.

F. Graphical User Interface (GUI)

The GUI component of the system is purposely developed for friendly interaction with the users. Both types of users, namely the students and academic staffs are given unique

access to their individual member area, where the students can access their personal information, while the academic staffs can monitor their students information. The developed GUI is in the form of dynamic web pages, which are database driven. This signifies that the information displayed on the web pages are constructed based on the data extracted from the database. The web pages are categorized into four modules, namely the User List, Log, Timetable and Attendance. The pages are developed using the Hypertext Preprocessor (PHP) scripting language and compatible with all major web browsers [15].

III.   SYSTEM IMPLEMENTATION

Based on the system design presented earlier, the system implementation was carried out. The overall process flow of system implementation is illustrated in Figure 4, which composed of 5 main steps.   The deployment of networked RFID readers is also illustrated in Figure 5.
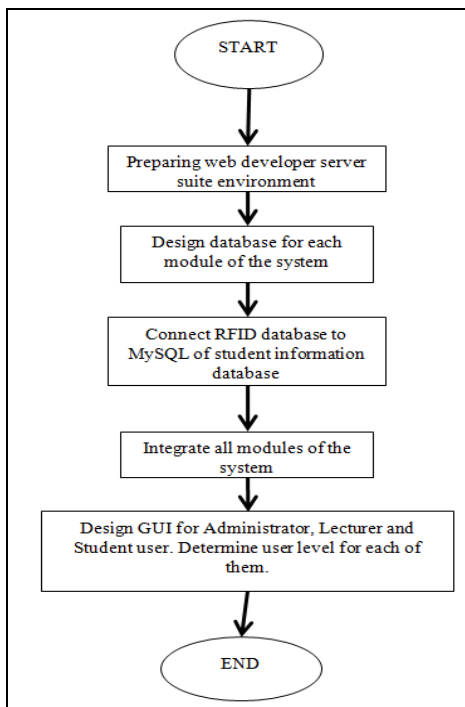
Figure 4: Overall process flow in developing the Web Based Student's Attendance System

A.   RFID System Device

The system employs RFID-Mifare terminals as the readers, which can be installed across all classrooms or laboratories, each denoting a specific venue [15-16]. These readers are connected to a hub or a switch through RJ45 port or also known as LAN port. In this setting, the hub or switch is linked to a router for internet access, which is required in order to connect the RFID readers to the Data Reporter component via TCP/IP protocol.
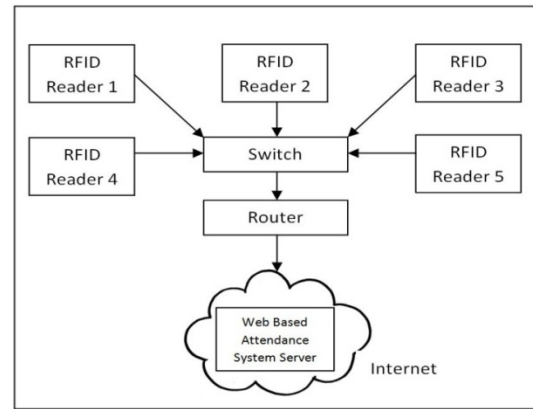
Figure 5: Network connection multiple RFID Reader to Web Based Attendance System server

B.   Web Development

The web application was initially developed on a local host that runs the XAMPP server. Having such web server hosted on the local host allows the PHP scripts to be developed and locally tested and debugged. Besides the web server, a database server is also run on the local host allowing the scripts' connectivity with the database. The database design will be briefly described next.

C.   MySQL Database

The web-based Student Attendance system is connected directly to a MySQL database. In general, there are four main tables created to store the collected data. A table called the Log table is used to store all users' logging data that include the check-in date and time, paired with the student ID. These data are initially captured by the Data Reporter component from the RFID reader, which are then submitted to the Data Collector, which is responsible in storing the data onto the database. The User table is another table in the database that is used to record the students' details that include their profile information such as name, address, email, phone number and etc. Courses information such as the course code, scheduled class venue, start and end time class are loaded into the Timetable table. Meanwhile the analyzed students' attendance information is recorded in the Attendance database. A set of pre-defined SQL queries are written into common functions allowing easy data insertions, update and retrievals for web-page displaying purpose.

D.   Web Graphical User Interface (GUI) Design

A user is allowed to enter the member area by logging in to the system via a login form. This type of authentication is important in order to prevent access by unauthorized users. The system grants access to 3 different types of users namely the Administrator, Lecturer and Student. According to the user type, each user is given the specific level of access. For instance, the Administrator level is given an access not only to view but to moderate and modify data in the system. In contrast, other user levels will be given only limited access to the system.

For easy reference, the GUI is designed with the horizontal style for navigation menu. For administrators, a navigation

menu is given that provides links to the four main functionalities of the system. These include the "User List" function for adding and removing the students' personal details. "Log" function is provided for the purpose of monitoring the students' check-in details. The "Timetable" function is used to manage the classes' schedule, while the "Attendance" function is to monitor the students' attendance. Figure 6 shows the navigation-menu design of the system.
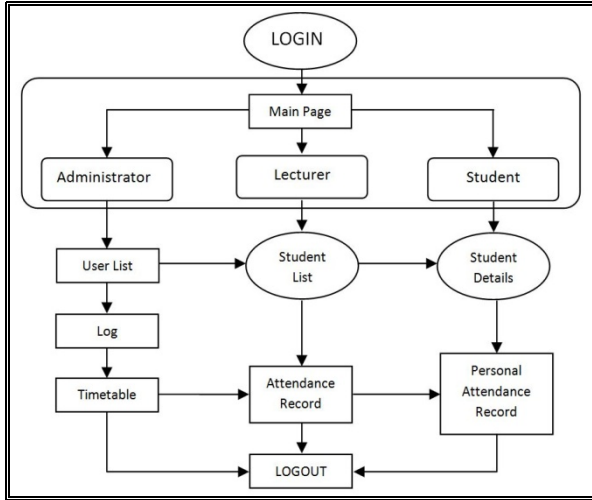
Figure 6: GUI navigation menu

IV. SYSTEM TESTING AND RESULT

To ensure the system correctness and completeness, system testing has been performed across the system environment that includes the client-side application, server-side application and the hardware.

A. Hardware and Application Integration Testing

RFID Reader was tested by having it to capture the ID and current timestamp for several RFID student cards in order to ensure the device is not malfunction and operates as desired. Firstly, a new student card is flashed to the RFID Reader to capture and recognize the card's unique ID. The captured ID is then used to register the mapped student ID into the system. Once registered, the student card can then be used by flashing it to the RFID Reader to record the student attendance. For this purpose, the log data, which include the current timestamp, is also recorded.

The next step was to set up the Data Reporter component so that it is able to fetch log data from RFID Reader and pass it to Data Collector. Both Data Reporter component and the log table (in the database) are then monitored for new log data specifically to detect whether Data Reporter has successfully fetched and passed the log data to Data Collector, which will store the data into the log table. If no new log data was recorded, Data Reporter and Data Collector components would be debugged and reconfigured until they successful perform the required function. Figure 7 illustrates the process flow of the hardware and application testing.
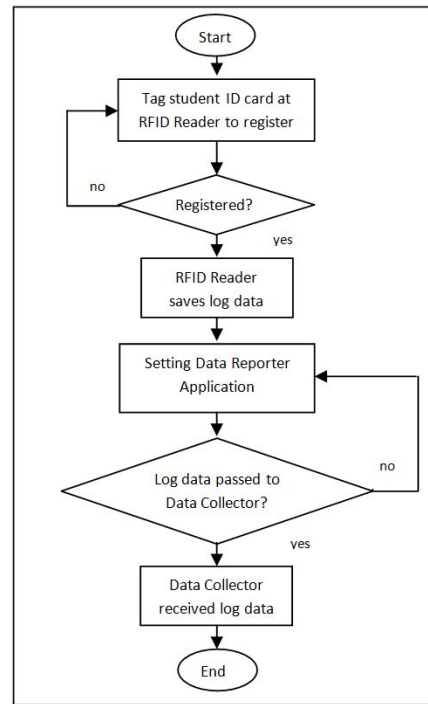
Figure 7: The hardware and application testing process flow
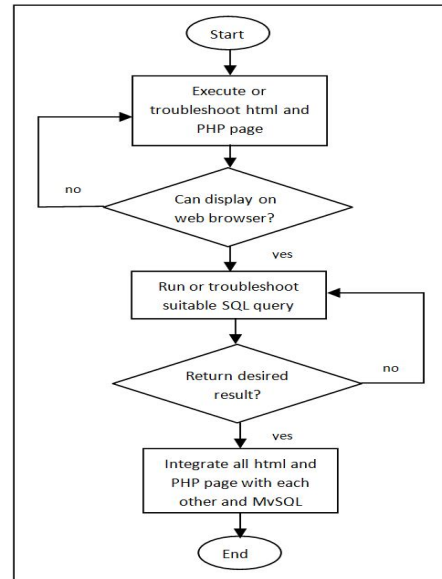
Figure 8: The online system testing process flow

B. Online System
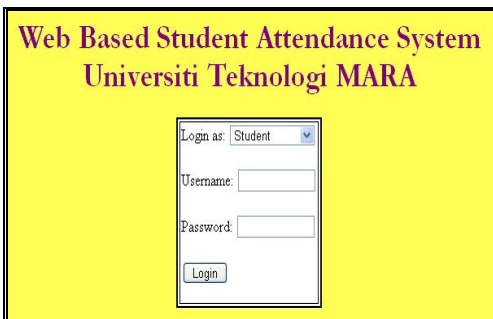
Unit testing for each online component was performed by monitoring the outputs they produced through a web browser. Such testing was done while coding each of the defined component. In detailed, after coding a particular function of a particular online component in PHP, the result or output is observed through the web browser. If there was no display or an error was displayed, the recently coded function will be debugged through a PHP editor. The similar test was performed to validate SQL queries performed onto the database. If an SQL query was found to return a wrong result, then the query would be debugged to attain the desired result.

Once all online pages are correctly displayed, an integrated system testing would be performed between the RFID reader and the online system together to complete the overall system testing. Figure 8 illustrates the process flow of the online system testing.

C. System Testing Results

In general, our aim to develop a prototype of a Web Based Student Attendance System was successful. The user-login authentication process was successfully tested, which prevents an unauthorized access into the system. Once a user is successfully logged in, the user is given the access to the main page that displays a menu listing a set of features offered to the user. Figure 9 illustrates the Web Based Student Attendance System login interface.

Figure 10 shows the system Web Based Student Attendance System Users List page, viewable only by the system administrator. On this Users List page, all users' details are displayed. In addition the administrator may choose to display a list of users according to their role, namely administrator, lecturer or student role. From the displayed list, the administrator is allowed to add new users or to delete existing users by clicking the delete or add user button. Figure 11 shows a page displaying the student's detailed information. This page is called the Student Attendance System User details page. For the similar feature, an administrator may view details of all users, while a lecturer may only view details of users identified as student. A particular student may only view details of their own account. On such page, an administrator can edit details of any user and a lecturer may edit details of a student but a student is not allowed to do any modification to their data.

Figure 9: Web-Based Student Attendance System login menu

Figure 10: Attendance List menu

V. CONCLUSION

The developed Web-Based Student Attendance System using Radio Frequency Identification technology will significantly improve the current manual process of student attendance recording and tracking system, especially in a university or school environment. The system promotes a semi-automated approach in capturing the student attendance, i.e. by having the students to flash their student cards to the RFID reader.

Figure 11: Web Based Student Attendance System User details page

In addition, a number of other advantages are gained by having an online web-based system, acting as a central repository of student attendance record. Firstly all processes of managing the student attendance record are performed online, allowing administrators and lecturers to view or modify the users' data through any computer via the web browser, as long as they are connected to the Internet. This way, no specific software installation is required. The captured student attendance data are also processed and analyze automatically with less risk of data loss, compared to a manual filing approach. Specific to lecturers or teachers, they can easily monitor their student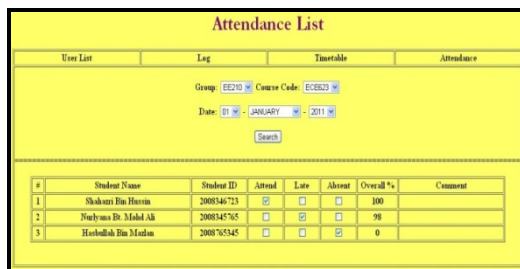s' attendance online and this could improve the quality of teaching since less time is needed to manage the student attendance record. The developed system can be improved and upgraded further, e.g. by extending the system with new features and modules or by improving the web-interface layout with new display style. Better yet the system can be enhanced further to offer another significant enhancement where the system can be extended to monitor staff attendance record.

REFERENCES

[1] Gatsheni, B.N., R.B. Kuriakose, and F. Aghdasi. Automating a student class attendance register using radio frequency identification in South Africa. in Mechatronics, ICM2007 4th IEEE International Conference on. 2007.

[2] Mahyidin, M.F. Student Attendance Using RFID System. 2008 December 2012]; Available from: http://umpir.ump.edu.my.

[3] Tokiwa, Y., K. Nonobe, and M. Iwatsuki. Web-based tools to sustain the motivation of students in distance education. in Frontiers in Education Conference, 2009. FIE '09. 39th IEEE. 2009.

[4] He, Z. and J. Zheng. Design and Implementation of Student Attendance Management System Based on MVC. in Management and Service Science, 2009. MASS '09. International Conference on. 2009.

[5] Company, C. Web Based RFID System. http://www.chiyu-t.com.tw/ 2011 [cited 2011 April 2011].

[6] Kassim, M. and S. Yahya. A case study: Reliability of smartcard applications and implementation in university environment, Malaysia. in 2009 International Semiconductor Device Research Symposium, ISDRS '09. 2009. USA.

[7] Lim, T.S., S.C. Sim, and M.M. Mansor. RFID based attendance system. in Industrial Electronics & Applications, 2009. ISIEA 2009. IEEE Symposium on. 2009.

[8] Kassim, M., C.K.H.C.K. Yahaya, and M.N. Ismail. A prototype Of Web Based Temperature Monitoring system. in 2010 2nd International Conference on Education Technology and Computer (ICETC), . 2010. Shanghai, China.

[9] Yahaya, C.K.H.C.K., M.N. Ismail, and M. Kassim. A study on automated, speech and remote temperature monitoring for modeling Web based temperature monitoring system. in 2010 2nd International Conference on Education Technology and Computer (ICETC). 2010. Shanghai, China.

[10] Yahaya, C.K.H.C.K., et al., A framework on halal product recognition system through smartphone authentication, in (2011) Lecture Notes in Electrical Engineering. 2011. p. 49-56.

[11] Sulaiman, N.A. and M. Kassim. Developing a customized software engineering testing for Shared Banking Services (SBS) System. in 2011 IEEE International Conference on System Engineering and Technology, ICSET 2011. 2011. UiTM, Shah Alam, Malaysia.

[12] Sulaiman, N.A., M. Kassim, and S. Saaidin. Systematic Test and Evaluation Process (STEP) approach on Shared Banking Services (SBS) System identification. in 2010 2nd International Conference on Education Technology and Computer (ICETC). 2010. Shanghai, China.

[13] Sulaiman, N.A. and M. Kassim. An approach using RUP Test Discipline process for Shared Banking Services (SBS) system. in (2010) 2nd International Conference on Computer Research and Development, ICCRD 2010. 2010. Kuala Lumpur, Malaysia.

[14] Qaiser, A. and S.A. Khan. Automation of Time and Attendance using RFID Systems. in Emerging Technologies, 2006. ICET '06. International Conference on. 2006.

[15] Aljawarneh, S., et al. A Web Client Authentication System Using Smart Card for e-Systems: Initial Testing and Evaluation. in Digital Society, 2010. ICDS '10. Fourth International Conference on. 2010.

[16] Wahab, M.H.A., et al. Web-based laboratory equipment monitoring system using RFID. in Intelligent and Advanced Systems (ICIAS), 2010 International Conference on. 2010.

An FPGA Implementation of Shift Converter Block Technique on FIFO for RS232 to Universal Serial Bus Converter

Nurul Fatihah Jusoh, Muhammad Adib Haron, Fuziah Sulaiman

Faculty of Electrical Engineering
Universiti Teknologi MARA, 40450 Shah Alam
Selangor Darul Ehsan, Malaysia
E-mail: Isku_nurul03@yahoo.co.uk

Abstract— **To meet the standard modern system communication demands, the paper represents the implementation of bidirectional shift converter technique for the embedded converter RS232 to Universal Serial Bus circuit block within FPGA using Verilog HDL language to be applied in a system wireless communication within Zigbee protocol. Utilizing the ModelSim-Altera, RTL model of the shift converter was developed and synthesized then stimulated using TimeQuest Timing Analyzer to observe its functionality.**

 Keywords- UART, Shift Converter, FIFO, USB, FPGA.

I. INTRODUCTION

Universal Serial Bus (USB) is a new personal computer interconnection protocol, developed to make the connection of peripheral devices to a computer easier and more efficient [1-3]. It has unique ability specification and performance in order to reduce a cost of end user [1-4], improved communication speed [1-4]. It was also able to connect more than computers and peripheral's devices [5] such as mice, modems and keyboards where an interconnection interface to the computer as main host via one device USB host ports that were been the most particular reasons it becomes widespread as an example of modern commonly peripheral interface bus nowadays. Base on that the sector industries and a telecommunication leader have been developed to implement of the Plug and Play technology.

Therefore, the interconnection of multiple devices (up to 127) and solves the problem like resource conflicts, IRQs and DMAs [6], easing the devices' attachment to PCs. It allows and supports the transfer rates up to 12Mbs, comprehending the low speed and mild speed data ranges. This gives more advantages compared RS232 and conversion with serial communication UART controller system as applied throughout the system of RS232 and used widely. The communication services in such networks can be high-speed and low-speed data, video, and many others with different performance and traffic requirements [7].

Normally, in a certain complex system, communication for RS232 between master and slaver controller divided into two ports connection, which is serial and parallel port. However, RS232 extensively used in serial communication because it

only interfaces and used with a simple structure [8] and have a long transmission distance [8]. Besides that, serial communication reduces the distortion into the signal, therefore makes conversions data transfer between two separated in great distance possible [8]. It takes bytes of data and transmits it in sequential pattern transmission form. As knowledge, UART is an abbreviation form for universal asynchronous receiver/ transmitter integrated circuit of serial communication protocol. This serial transmission of UART however, depends on the two major factors on the modes applied in the hardware, whether in synchronous or asynchronous operation. On the other hand, it means that the outputs of asynchronous machine are functions either the current state only, or of the current state and the current input, corresponding to Moore and Mealy machine, respectively [9]. Current state is directly related to the proceeding sequence of input and state.

A Field Programmable Gate Array (FPGA) is a programmable logic device that supports implementation of relatively large logic circuits. [10]. FPGAs are distinguished on the basic of several features: architecture, number of gates, mechanism for programming, program volatility, the granularity and robustness of a functional/logical unit, physical size (footprint), pin out, time-to-prototype, speed, power, and the availability of internal resources for connectivity [11-12]. It always became one of the recommended devices suitable used to implement simple interface circuit or complex state machine to satisfy different system requirement either in medium or high density applications. The performance and the architecture of operational unit configurable reprogrammed with no practical limit [13] of FPGA make it special advantage features in improved the system integration [14], reliability [14] and in reducing power consumption [13-14]. Moreover, it can configure into the logic blocks to perform complicated combinational functions in ordinary logic gates such as AND and XOR [15].

This paper presents design of converter development for an RS232 protocol to Universal Serial Bus (USB) protocol focusing on the sum added implementation of Shift Converter Block technique onto and interfaces with FIFO and FPGA. The Altera board EP2C35F672C6 Cyclone II FPGA family

has been used as FPGA device and Verilog HDL as a main language code. Normally, when the deal with UART structure code with FPGA device, it was connecting straight via an RS232 port compared with connects directly via USB host port to transmit and receive information. This technique is much easier because without needed to think how to create and learn more in order to design behavior converter coding structure protocol language needed when a deal with the USB host of FPGA. This specification of the converter must be able and responsible for receiving data from peripheral device serial interface and then sending it to the computers USB interface.

However, as knowledge the application of performance and special features of USB host port have been described above give more advantage compared with RS232. Therefore, the sum added implementation of shift converter block technique were formed to control package bits of data, whether for transmit and received operands from via USB protocol engine between FIFO technique design and UART protocol engine of RS232. It is because, as knowledge, UART protocol engine takes bytes of data and transmit/receive the data individual bit in sequential pattern. In contrast, the FIFO designs were transmit/receive in same taken bytes of data. Then, this architecture combination of all the design can be able implemented practical this converter RS232 protocol engine to USB engine onto wireless communication via Zigbee Standard Protocol on FPGA as future work.

II. BACKGROUND OF RESEARCH

A. USB (Universal Serial Bus)

The USB Host System basically focusing on the specification features into it such are descriptions of bus attributes, protocol definition, programming language code interface and other features for each different device type that required to design and build the system and peripheral compliant with the USB standard. Nowadays, there are two common versions of the USB specification which are USB1.1 and USB2.0 specification supports low-speed (1.5Mbps) and full-speed (12Mbps) devices, respectively. The newer USB 2.0 version specification is backwards- compatible with the USB1.1 version. However, it can also support high max speed devices 480Mbps (deliver packet). This EP2C35F672C6 Altera board family used supports the USB 2.0 protocol engine that was be integrated into a peripheral device. Hence, there are four types of transaction supported, which are first is Control (used by USB system software to configurable the device). Next types are Bulk and Isochronous as a reliable transfer who includes the Handshake phase and to configurable the real time, saves an overhead by excluding handshake phase, respectively. It is interrupt as a limited latency transfer to or from the device. Moreover, in order to achieve the communication requirement speeds of USB device varies from 10 to 100Kbps, it is important to verify and considerable to use interrupt transfers into the converter design. USB work as a master/slave bus depends on the types where USB Host (typically PCs) is the Master while the devices are the Slaves. The Host is responsible to receive and transmit an I/O request packet to perform the task of scheduling the transactions over the USB (which is composed as a series of USB packets). The number

of register to store the data and memory needed to depend on the types the device connected. This called Endpoint. An Endpoint is uniquely identifiable portion of a USB device as the final stop of communication flow between host and peripheral device (as buffer) as illustrate in Fig. 1 below.
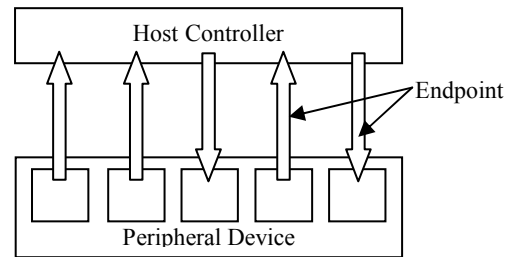
Fig. 1: Structure of Device Endpoint

Firstly, the USB operation considered packet transfer protocol. During the data transfer, the USB systems function as in the following overall manner. When there are, some interface occurs into the system (via an IN/OUT request packet). This transaction of request packet will be handling by the Host Controller based on the contents of the Interface transfer descriptor built by the USB Host Controller Driver. Then Endpoint Descriptor will run followed with Packet ID (PID) field to identify subfield of type's categories packet data. These types of package's categories were depending on usage. Then, the variable length of packet specific information field will identify PID Types and PID Check categories Subfield values for Token, Data, Handshake and Special packages before being sent. Although 4 bits are required from 11 basic different package type (such are SOF, SETUP, DATA0, ACK, etc.), to run this operation, 8-bits are used in order to improve efficiency of error checking by using technique first compliment form.

Next, Cyclic Redundancy Check Field or known as CRC will take over as a method function for checking binary data for presence of errors in the Package Specific Information Field. There are two types CRC used are CRC5 and CRC16. It created based on the implementation of each both types CRC as given below for general polynomial and residual, where Equation (1) and (2) for CRC5 and Equation (3) and (4) for CRC16, respectively.

$$G(x)_{CRC5} = x^5 + x^2 + x^0 \qquad (1)$$
$$\text{Residual, } G(x)_{CRC5} = x^4 + x^3 \qquad (2)$$

$$G(x)_{CRC16} = x^{16} + x^{15} + x^2 + x^0 \qquad (3)$$
$$\text{Residual, } G(x)_{CRC16} = x^{15} + x^3 + x^2 + x^0 \qquad (4)$$

Finally, all the packages were terminating except PRE package with function End-of-Package, EOP indicator.

B. UART (Universal Asynchronous Receiver Transmitter)

A frame in an asynchronous serial protocol for RS232 protocol engine is a non-divisible packet of a bit that embeds one start bit, seven or eight data bits, one parity bits (depend on the width data either even parity or odd parity), and stop bits. However, most of the bits in a frame are self-explanatory. Fig. 2 below, showed the component block in UART. Referring to this Fig. 2 baud rate generator used to generate a

suitable output clock for transmitter and receiver module. Bit cell duration applied for baud rate needed is 104.16μs. It is as default baud rate to be applying for future work. The percentage tolerance occurred in the baud rate generator were also determined to check whether it is acceptable or not in order to minimize lost signal and getting a good transmission. It were done by considered applied value master clock and baud rate needed by using a specific formula for getting percentage accuracy as shown in Equation (5) below. Therefore, using the given equation below the percentage tolerance occurs is 0.124%. It is within the acceptable tolerance for transmission signal as in theoretical expectation.

$$\text{Percentage Accuracy} = \frac{(X - Y)}{X} \times 100 \qquad (5)$$

Where,

X = Baud rate needed
Y = Actual baud rate
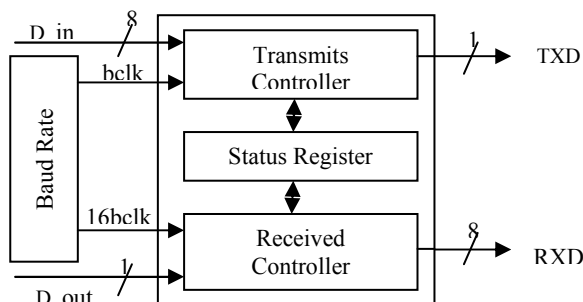
Fig. 2. Structure of UART block

C. Shift Converter

A shift register is well-known as the sequential logic module constructed from flip-flops that manipulates the bit positions of binary data by shifting data bits to the left or right depend on the instruction setting in the module. There are two methods involve in this shift converter, which is serial to parallel converter bit and parallel to serial converter bits. The rotation modes depend on the type of packet data from current transmission when transmitting or receiving process. In this paper, there are five majors components involve implemented in designing shift converter which are controller, status register, parallel to serial converter bit, current state, and serial to parallel converter bit as illustrated in Fig. 3 below.

When, the shift converter in parallel-in, serial-out (PISO) register it will accept binary input data onto parallel and generates binary output data in a serial form. The register in this module is set for left-shift device. Then, three conditions will determine the value of the bits shifted onto the vacated position on the right. After that, zeroes will fill that position upon the operand shifting. The serial output data generated from the output of the internal load register. However, when the shift converter is in serial-in, parallel-out (SIPO) register, then another typical synchronous iterative network will be identical as a cell. Data will enter the register form from the right and shift serially to the left through all the stages involved inside depend on the width of packet data. 1-bit positions will do operand every clock pulse. After that, the register will be in a state fully loaded, and the bits were transfer to the destination. Lastly, a new packet byte data will replace the position and continues the operands. This SIPOs

operation is a very useful application to generate the sequence of non-overlapping pulses for system timing [16].

All both register will do operation shift at each active clock transition. A converter is implementing with synchronous reset. It means that when the reset is detected, all the operation and register will be in the state idle and clear. Otherwise, it will operate as instruction given in the module.
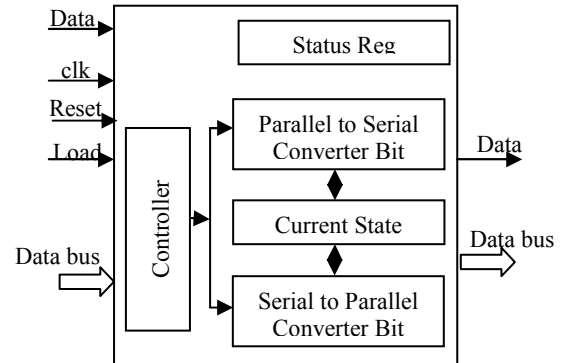
Fig. 3. Structure of shift converter block

D. Asynchronous FIFO

An Asynchronous FIFO refers to FIFO design that used to buffer data in a digital system. FIFOs always used for data cache, storing different of frequency or phase of asynchronous signal [2]. The requirement of FIFO will arise when read function are slower than the writes. It interfaces with two clocks domains for writing and reading operations that asynchronous to each other. It means that, data is writing into the FIFO from one clock domain, and it is reading from another clock domain. Because of this, it required a memory architecture wherein two port memories are available one for write input operation and other one for read input operation. Moreover, FIFO can be use to complete parallel or series data. Generally, to facilitate error-free operations of FIFO there are two types of signal generated corresponding with the clock, which are FIFO Full Flag and FIFO Empty Flag. When the system is in write operand, and it is a full flag, write system will not accept and write data to the read system until the system is in an empty flag again. It is to prevent the data from lost. Commonly, two types of counter used as FIFO pointers, which are binary counter and Gray counter. No matter under what circumstance, these two methods have merits and demerit. The decisive factor in choosing the right counter design as a pointer must be referring on the pitfalls between the read and write clock domain and its synchronization advantages. For that reason, the counter who chooses as FIFOs pointer in this paper is in binary counter. This counter is easy to design and implement in addressing FIFOs. To avoid metastability between the transactions of the binary bits of the pointer the widths of counter bits are set suitable with the addressing location of memory address and produce some holding register this holding register can synchronize signal.

III. PROJECT PLAN/DESIGN METHODOLOGY

The implementations of bidirectional shift converter block technique in a converter for UART to transmit and receive data to USB Protocol Engine via FIFO was designed using

EP2C35F672C6 FPGA device as illustrate in Fig. 4. Firstly, based on Fig. 4 below, the behaviors of these entire blocks are characterized using Verilog HDL language. In this programming derivation stage, design specification such as input and output elements was determined depending on the functionality of each created module.

Then, each code for the module was compiled in order to check the syntax of the design and synthesized to update or convert the designed from Verilog HDL code into a symbol logic block. After this stage, the simulations waveforms of each module had presented to ensure the designed output waveform matched with theoretical expectation. These simulations had done using ModelSim-Altera. Next, all these blocks are connected and combined together into one as a main symbol block logic. Lastly, it will compile and analyze the simulation output waveform again via the same tool.
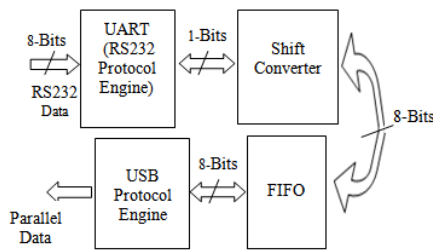
Fig. 4. Block diagram of the proposed converter via FIFO

A. Symbol Logic Block

The first part is designing UART of RS232 Protocol Engine to transmit and received signal from bidirectional Shift Converter. It considered with three main parts, which are Transmit Controller, Received Controller and Current State Register. In this Transmit Controller part, it consists of two major components, which are Transmit Buffer and Shift Register. Transmit Buffer is a function to load and transmit data from local CPU into Load Register. Then, Shift Register will load and shift the data from Transmit Buffer before send it through TXD output pin one by one bit to the shift converter. While, Receive Controller also have two main components, which are Receive Shift Register that functions to receive the data from the shift converter from RXD pin one by one bit and Receive Buffer to load and accept the data from Shift Register to local PC read. Based on this, UART had a special function register to analyze and control the register to transmit or receive data.

The second part is designing the bidirectional shift converter that has two functions as illustrated in Fig. 5 below. First function is to receive data from RS232 Protocol Engine one by one bit using port "datain_tx" before load and shift bit into the special internal register to convert it into eight bits data width. After it had done, the data were transmits to FIFO using port "dataout_tx" to read and write the data. Then, FIFO will transmit in byte data width to the USB protocol engine. Protocol engine for USB type ISP1362 device used in EP2C35F672C6 Altera board family. In this ISP1363 USB Protocol Engine, for all the operation of the sub modules, involve into it as describe in above created.

Second functions are to receive data ISP1362 USB Protocol Engine in the parallel bytes of data width to FIFO. Afterwards, after it interfaces operation in FIFO, byte data bus was sent to the bidirectional Shift Converter using port "data_in" before load and shift bit into a special internal register that function to convert from eight-bit data width into one-bit data width. Then

the data will be transmitted to RS232 Protocol Engine using port "data_ rx." In order to check and analyze the bit operation "cs_tx" and "cs_rx" stand for current state for transmit data to UART (RS232protocol engine) and four transmit data to FIFO are created, respectively.

Besides, from this Fig. 5 the structure of the bidirectional shift converter is name as "conv_bit" in symbol logic block. This symbol logic block was created using Verilog HDL from two directional transmit and receive sub-module converter. Then, it compiled, synthesized and stimulated before download together into this symbol logic block onto FPGA to verify it functionality. It consists of four input ports, including the 'clk', 1-bit 'datain_tx', 'reset', 'load', 'store' and 'data_in' with 8-bit length package data. These 2 input ports which are 'datain_tx' was connected to UART while 'data_in' was connected to FIFO. The output ports are labeled as 'cs_tx' and 'cs_rx' with 4 four bus widths each act as a current state for transmit signal to FIFO and UART respectively. Furthermore, data out ports that are 'data_rx' as output port connected to UART and 8-bit widths bus "dataout_tx" as output port connected to FIFO.
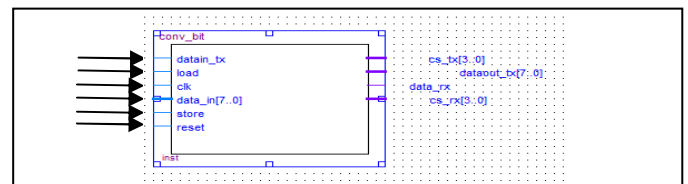
Fig. 5. The symbol block logic structure of shift converter

In addition, the methodology in this designing an operation of the shift converter was also summarized it into two phases. Firstly, the converter for transmitting package data bus from UART (RS232 Engine Protocol) to USB Protocol Engine via FIFO and second receiving package data bus from USB Protocol Engine via FIFO to Universal Asynchronous Receiver and Transmitter (RS232 Protocol Engine) as illustrated in Fig. 6 and Fig. 7 respectively. Each phase involved with four main modules, which are RS232 Protocol Engine, Bidirectional Shift Converter, FIFO and USB Protocol Engine.
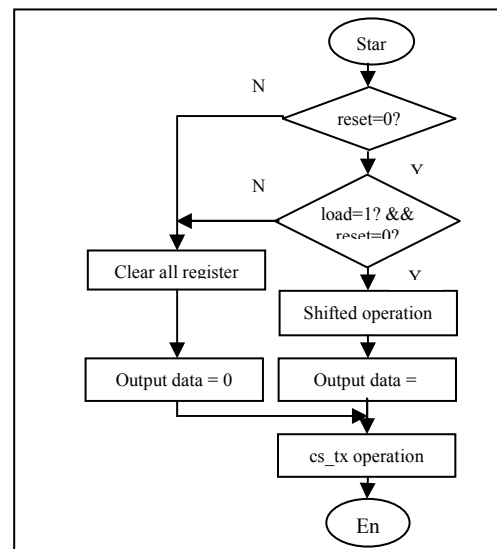
Fig. 6. Flow chart shift converter for transmit/receive signal to FIFO

Therefore, based on Fig. 6 above, from the starting point, if "reset" is at logic '1', then all the register will be cleared into '0' including "load" and output "dataout_tx" bit logic. Otherwise, if "load" is at logic '1' and "reset" is '0' the "shift"

process will operate using count operation technique once it detects. Finally, outputs "dataout_tx" will transmit data to FIFO in eight widths package data bus. In order to check the bit's operation to FIFO output register output pin "cs_tx" stand for current state inserted. All the Verilog HDL code created either main or sub-module were compiled and synthesized using Modelsim-Altera device in order to obtain and analyzed the results.

Fig. 7 below illustrates the summarized operation shift converter receiving package data bus from FIFO to UART (RS232 Protocol Engine). From the starting point, if "reset" is at logic '1', then all the register will be cleared into '0' including "data_hold" and output "data_rx" bits logic. This 'data_hold' will act as special internal register and decoded by sampling each frame received. Otherwise, if "store" and "reset" at bit '1' and '0' logic respectively the "shift" operation will on and "data_in" as bit input is load into register "data_hold". Next, if "shift" is at logic '1' will start shifting operation using 'count' concept operation. Finally, outputs "data_rx" will transmit bit to RS232 Protocol Engine. In order to checking the bits operation to UART output register "cs_rx" stand for current state inserted. The Verilog HDL code were compile and synthesize in simulation using Modelsim-Altera device in order to obtain and analyze the results.
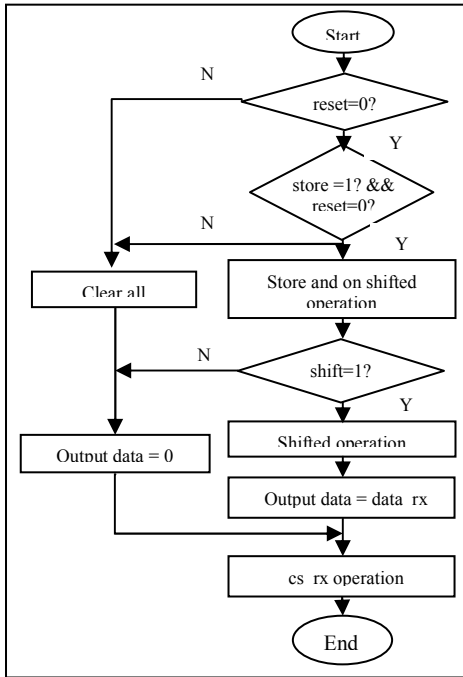
Fig. 7. Flow chart shift converter for transmit/receive signal to UART

The third part in the design was the asynchronous FIFO where data values are written sequentially into a FIFO Buffer using a one clock domain, and the data values are sequentially read from the same FIFO Buffer using another clock domain. These two 'clk' domains are asynchronous to each other. The FIFO depth used is 16. In order to facilitate error-free operation, FIFO Full Signal and FIFO Empty signals were used. FIFO Full Signal used a 'wr_clk' while FIFO Empty Signal is driven by 'r_clk'. When the FIFO is in-state Full Signal, it will stop writing into FIFO, however if the is in state Empty Signal it will stop reading from it.

Lastly, the last parts were USB Protocol Engine for ISP1362. This is quite tough because it considered so many

things in order to relate with their advantage used of full speed mode where it peak data rate able to reach 12Mbps. There are such Endpoint descriptor transfer transaction, declaration of four packets type (Token, Data, Handshake and Special), PID field, CRC (such are CRC5and CRC16), End-of-Package, (EOP), etc.

IV. RESULTS AND DISCUSSIONS

The architecture was modeling in Verilog HDL. The functionality and characteristic all designs were compiled and synthesized using Quartus II software. Then, the results of the simulation in ModelSim-Altera had been compare and analyzed with theoretical expectation. The "conv_bit" applied with 1ns\1ns timescale. Final simulations results in the Shift Converter bit carried out and divided into two. First waveform result from the converter bit "conv_bit" analyzed between RS232 Protocol Engine to USB Protocol Engine via FIFO and second waveform result between USB Protocol Engine to RS232 Protocol Engine via FIFO.

All the operations involved were design based on their characteristics. After synthesized using ModelSim-Altera the architecture is run and compiled using TimeQuest Timing Analyzer to checking the summary of timing clock transmit and receive data through the shift converter. From this it is an ability to show the summary report and create timing Netlist using SDC language for all input and output timing clock and so on including the specific device reports for metastability.

A. Simulation Waveform Result

The result as shown in Fig. 8 below indicates the transmitting waveform carried out for the bidirectional shift converter bit "conv_bit" between RS232 Protocol Engine to FIFO using 1ns/1ns timescale. Two packages 1-byte bus input data '11010010' and '10100100' respectively were applied that receive from TXD port in RS232 Protocol Engine one by one bit. While the rest of the input's port "reset" and "load" were forced at certain logic to be implement in embedded converter from RS232 to USB. However, as shown in Fig.8 Below, input's port "reset" and "load" was forced at logic '0' and '1' respectively. The input's data in "datain_tx" one by one bit were shifted serially every posedge "clk" to the 1-byte bus at output data "dataout_tx" starting '10000000', '11000000' and so on from 24900ns until the last byte first package '11011110' at 25700ns then followed 25750ns to 26510ns per byte data bus for the second package. Furthermore, the output port "cs_tx" clearly show the currently state operation.

Fig. 8. Simulation waveform result for transmit signal to FIFO

The result shown in Fig. 9 below depicts the receiving waveform carried out for the converter bit between FIFO to RS232 Protocol Engine. Four packages 1-byte widths bus input data '10110101', '10001101', '10010001' and '10001101' respectively transmit serially one by one to input port FIFO in 1-byte bus package data each time. The rest of the input's port "reset" and "store" were forced at certain logic to check it functionality in order to be implement in embedded converter from RS232 to USB. However, as shown in Fig. 9 below inputs port "reset" and "store" was forced at "0" and "1" respectively. The package bus inputs data "data_in" were loaded into "data_hold" before shifted serially to one by one bit at output port "data_ rx." After it had done, it will continue with the next package bus data as describe above until the last package. Therefore, based on the two output waveform the output data was totally depending on the input data before totally transmit to UART shifted serially one by one all the receive package input data from FIFO.
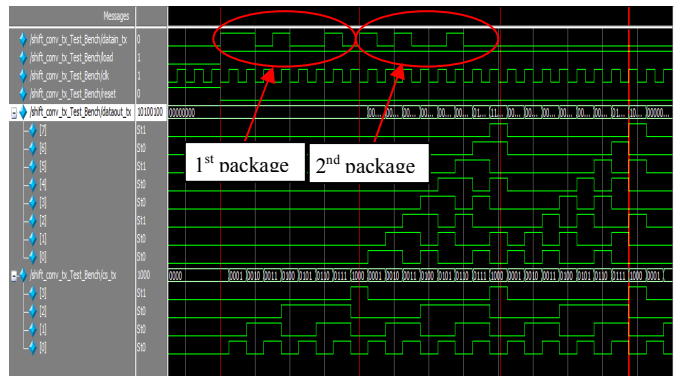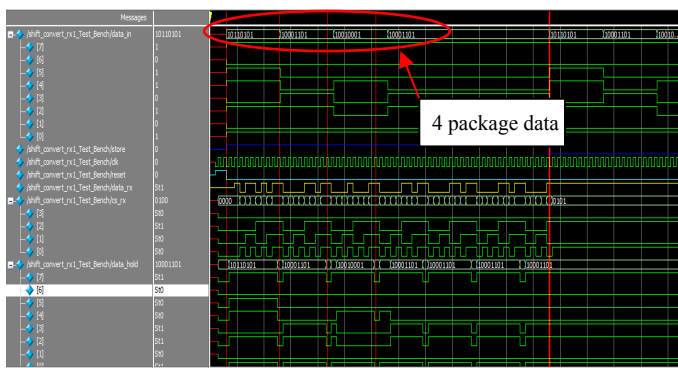

Fig. 9. Simulation waveform result for receive signal to UART

Then, the architecture was compiled and stimulates using TimeQuest Timing Analyzer simulation to check the timing summary of the database. Based on Fig. 10 below, it has shown the transmission-line clock delay in the data arrival is 2.685ns with zero metastability of the device. It is acceptable for reliable communication as theoretical expected. Besides that, a small delay occurred during the output data after rising clock edge, which is the most probably caused by net delay pins on EP2C35F672C6 FPGA.
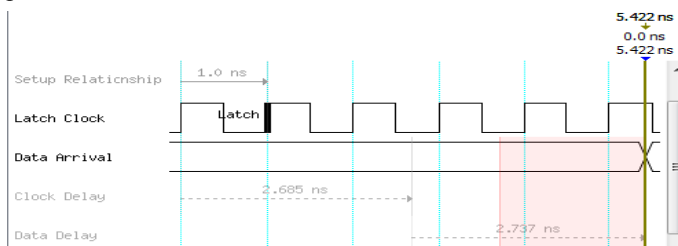
Fig. 10. Simulation waveform using TimeQuest Timing Analyzer
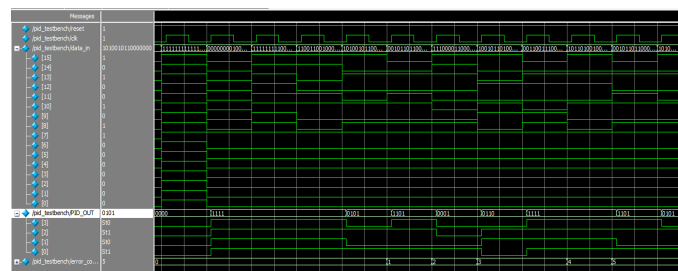
Fig. 11. Simulation waveform PID for USB

Fig. 11 above depicts the produced and displays the output and test for the error when the SYNC detected for most all types of PID packet data. When the valid pid code observed in sync pattern in serial USB input data stream.

V. CONCLUSION

The results and discussions discussed in the previous section gives a clear perspective that the objective of this study been successfully achieved. The entire physical layer design of implementation's bidirectional Shift Converter Block technique for converter UART on FIFO to USB and FPGA developed using Verilog HDL and simulated using ModelSim-Altera. The simulation results show that the converter bit model follows the theoretical development as a part implemented in embedded converter RS232 to USB. The model of the shift converter intended for embedded converter RS232 to USB design to Zigbee Protocol Engine will implement as a future work on this study.

VI. ACKNOWLEDGEMENT

.

REFERENCES

[1] Vijaya,V;Valupadasu,R;Chunduri, B.R; Rekha, C.K; Sreedevi. B., "FPGA Implementation of RS232 to Universal Serial Bus Converter", IEEE Symposium on Computers & Informatics (ISCI), pp 237-242, 2011.

[2] Ana Luiza de Almeida Pereira Zuquim, Claudionor JosC Nunes Coelho Jr, Antanio Ot6vio Fernández, Marcos PCgo de Oliveira, AndrCa labrudi Tavares, "An Embended Converter From RS232 to Universal Serial Bus", IEEE.

[3] Jan Axelson, "USB Complete, Everything you need to develop custom USB Peripherals", Penram Intl. Publishing (India), 1999.

[4] Sunggu Lee, "Advanced Digital Logic Design Using Verilog, State Machines and Synthesis for FPGAs", in Proc. 1st Edition, pp 46-107, ISBN 0-534-55161-0, 2006.

[5] http://www.usb.org.

[6] Mindshare, Inc., D. Anderson, "Universal Serial Bus System Architecture", Addison-Wesley Developers Press, U.S., 1997.

[7] R.Wyrmas,W. Zhang, M.J. Miller, and R.Anjaria, "Multiple Access Option for Multimedia Wireless System", in Proc. 3rd Workshop on Third Generation, pp 289-294, Apr.1992

[8] L. K. Hu and Q.CH. Wang, "UART-based Reliable Communication and performance Analysis", Computer Engineering, Vol 32 No. 10, pp15-21, May 2006

[9] Cavanagh, Joseph J.F, "Digital Design and Verilog HDL Fundamentals", in Proc. 1st Edition, pp 566- 713, ISBN 978-1-4200-7415-4, 2008.

[10] Brown, Stephen D., "Fundamental of digital logic with Verilog design", in Proc. 2nd Edition, pp77-166, ISBN 978-0-07-338033-9, 2008.

[11] Chan PK, Mourad S., "Digital Design Using Field Programmable Gate Arrays", Upper Saddle River, NJ: Prentice-Hall, 1995

[12] Oldfield JV, Dorf RC., " Field programmable Gate Arrays", New York: Wiley Inter-science,1995

[13] Ciletti, Michael D, "Advance Digital System with the Verilog HDL", in Proc. 1st Edition, pp103-132, ISBN 0-13-089161-4.

[14] Shouqian, Y., Y. Lili, et al. (2007). Implementation of a Multi-channel UART Controller Based on FIFO Technique and FPGA. 2nd IEEE Conference on Industrial Electronics and Applications, 2007. ICIEA 2007.

[15] http://www.altera.com.

[16] Cavanagh, Joseph J.F, "Digital Design and Verilog HDL Fundamental",in Proc.1stEdition, pp620-622, 2008, ISBN 978-1-4200-7415-4

Integration of Low Altitude Aerial & Terrestrial Photogrammetry Data in 3D Heritage Building Modeling

Khairil Afendy Hashim[1], Anuar Ahmad[2], Abd. Manan Samad[1], Khairul NizamTahar[1] & Wani Sofia Udin[3]

[1]*Pixelgrammetry & Al-Idrisi Research Group (Pi_ALiRG)*
Centre of Studies Surveying Science & Geomatics,
Faculty of Architecture, Planning & Surveying,
Universiti Teknologi MARA,
40450 Shah Alam, Selangor MALAYSIA

[2]*Institute for Science and Technology Geospatial (INSTEG),*
Faculty of Geoinformation & Real Estate,
Universiti Teknologi Malaysia,
81310 Skudai, Johore MALAYSIA

[3]*Faculty of Earth Sciences,*
Universiti Malaysia Kelantan, Lock Bag 100, 17600 Jeli, Kelantan, MALAYSIA

email: khairilafendy@perlis.uitm.edu.my

Abstract—**Aerial photogrammetry is one of the geomatic technologies that can be used to produce topographic map while terrestrial photogrammetry has been used to produce 3D model of objects. Both methods could be used for many applications including for recording heritage building. This study is carried out to investigate the combination of two photogrammetric methods i.e aerial photogrammetry which utilizes light weight rotary-wing unmanned aerial vehicle (UAV) and terrestrial photogrammetry for the production of 3D heritage building. In this study, a building is simulated as a heritage building located in the precinct of Universiti Teknologi Malaysia. The UAV was used to acquire the digital aerial images of the building and its' surrounding where the UAV is equipped with high resolution digital camera. For terrestrial photogrammetry, the digital images of the building facades are acquired from the ground using the similar high resolution digital camera. For accurate measurement, the high resolution digital camera is calibrated and later the recovered parameters are used for digital image processing. Surrounding and on the façades of the building, ground control points (GCP) and control points (CP) are established respectively using total station. The GCP is used in image processing of aerial photogrammetry to produce the orthophoto and foot print of the building. For the CP it is used in image processing of terrestrial photogrammetry to produce 3D model of the building. Then the results from both methods are combined together where the 3D model of the building is placed on the footprint of the building in the orthophoto. In conclusion, product of geomatic has been used by archeologist for decades, however, product such as 3D building could assist archeologist to produce better recoding of a heritage building.**

Keywords-Geomatic, aerial photogrammetry, terrestrial photogrammetry, closerange photogrammetry and 3D building

I. INTRODUCTION

With the rapid development in surveying hardware, computer and software, the term surveying has changed and widely known as geomatic. It is well known that the output of geomatic such as topographic map, plan, 3D coordinates, cross-section etc has benefited various professions such as engineers, architect, planner, administrator, scientist, researcher etc. In this paper, more discussion about photogrammetry whether the photograph is acquired from the air or ground surface is discussed. Other geomatic technologies will also be discussed.

Photogrammetry is the technique that can be used for 3D object measurement and modeling using photographs/imageries. In photogrammetry, the basic concept involves the acquisition of two successive photographs which overlapped to one another that are necessary to allow the object image projection applies for the purpose of 3D measurement and modeling as shown in **Figure 1**. Photogrammetry can be divided into two main categories that are aerial photogrammetry and terrestrial photogrammetry (or widely known as close range photogrammetry).

Figure 1: Two successive photographs (left and right photograph) for 3D measurement and modeling. Source: [22]

II. AERIAL PHOTOGRAMMETRY

A. Manned Aircraft

This platform has been used since the last decade, especially in the work of preparation of topographic maps. In this method, photographs are taken using a digital metric cameras special mounted on the manned commercial aircraft. **Figure 2** shows the aerial camera mounted on the aircraft, aircraft used for photographic mission and the aerial photography. The costs of aerial photogrammetry are usually very costly or very expensive in relation to the limited extension of the surveyed areas.

Figure 2: (a) Aerial camera system mounted in an aircraft, (b) Aircraft used for aerial photography, (c) Aerial photography. Source: [17].

B. Unmanned Aerial Vehicle

According to [5], UAVs are to be understood as uninhabited and reusable motorized aerial vehicles. These vehicles are remotely controlled, semi-autonomous, autonomous, or have a combination of these capabilities.

Figure 3: Example of different types of civilian unmanned aerial vehicle (UAV) (a) Rotary (b) Fixed-wing

Some fixed-wing platforms and mini-helicopters (rotary) as shown in the above **Figure 3** can be considered UAVs, as they are "aircraft which are designed or modified, not to carry a human pilot and are operated through electronic input initiated by the flight controller or by an onboard autonomous flight management control system that does not require flight controller intervention" [4]. Nowadays, UAVs with off-the-shelf digital cameras, panoramic cameras and a large number of diverse customer still video cameras and camcorders [18]. This is improved by GPS/INS systems for precise navigation and positioning [24, 25]. **Figure 4** shows the ideal UAV system for mapping purposes.
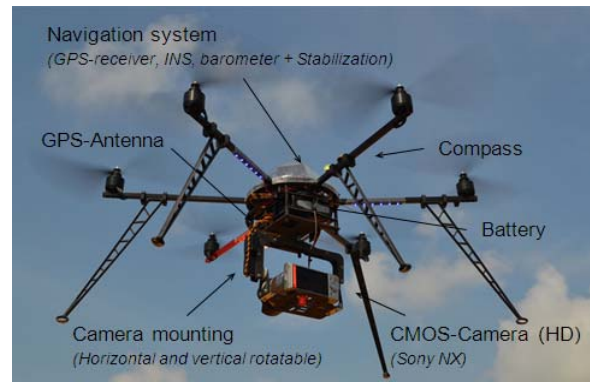
Figure 4: Unammaned Aerial Vehicle (UAV) system for mapping application-Hexakopter^{TM}. Source: [2]

III. TERRESTRIAL PHOTOGRAMMETRY

Terrestrial photogrammetry refers to applications where the camera is supported on the surface of the earth. Typically the camera is mounted on a tripod and the camera axis is oriented in or near the horizontal plane as shown in **Figure 5**.
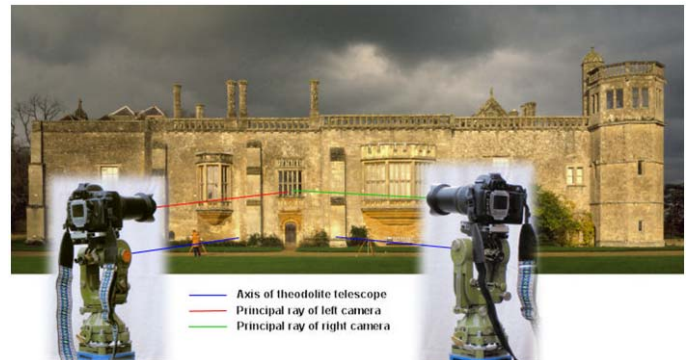
Figure 5: Terrestrial photogrammetry: Camera mounted on a theodolite tripod while taking heritage building photograph. Source: [11]

A. CloseRange Photogrammetry

Terrestrial photogrammetry is often called close range photogrammetry when the distance to the object is less than approximately 300 m [3]. **Figure 6** shows an example of close range photogrammetry work begin from planning until the production of 3D model.
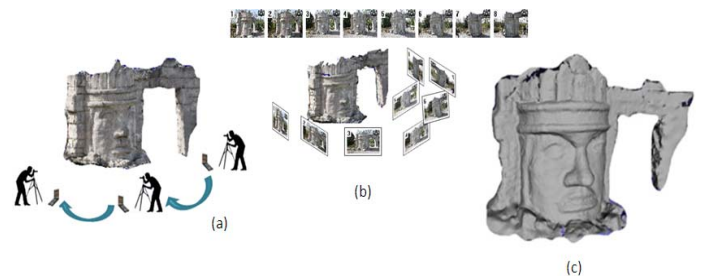
Figure 6: Several examples of close range photogrammetry work: (a) Taking pictures (b) Images captured in sequence (c) 3D model Source: [23]

IV. UAV FOR CULTURAL HERITAGE

Let consider for example the completion of the 3D survey of a construction, where the roof is to be capture at high resolution as well, or when a high-rising building has to be imaged with its' roof top. The availability of a micro or mini UAV equipped with a high resolution digital camera enables one to capture images from unconventional point-of-views and to respond to the above mentioned critical requirements. UAVs are successfully applied to map archaeological sites, where other surveying techniques are not suitable [6]

A. Unesco Interest

As can be seen from the UNESCO World Heritage List (whc.unesco.org), many recent additions can actually be classified in terms of culture and nature. In a press release (No. 2002-77: "For UNESCO, Space Technologies should be Harnessed for Sustainable Development") UNESCO has stressed the use of aerial sensor technology for monitoring World Heritage sites [21].

V. RESEARCH METHODOLOGY

A. Research Background

In this study, two main hardware are used which include the light weight rotary wing UAV and high resolution digital camera. Low altitude UAV is preferable in this study because it focuses on large scale mapping which involved small area only. **Figure 8** shows the UAV and digital camera used in this study.
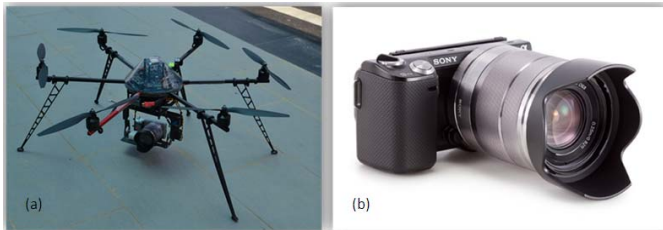
Figure 8: Instrument used in this study (a) Light weight rotary wing UAV HexaKopter (b) digital camera

In this study, high resolution Sony Alpha NEX-5N digital camera has been used in acquiring simulation heritage building model. This type of camera has 16.1MP CMOS sensor and 1080p 60p HD movies in AVCHD format. In this study, micro UAV also known as HexaKopter has been used in acquiring images of simulation model. HexaKopter has 6 blades where 3 blades rotate clockwise direction and 3 others blade turn to counter clockwise. The Sony digital camera is attached at the bottom of HexaKopter to capture aerial images during flight operation. In this study, aerial photogrammetry data is integrated with close range photogrammetry data to construct 3D heritage building model. In another word the data will cover from bottom to the top of heritage building. For the purpose of the study, one of the Universiti Teknologi Malaysia buildings is simulated as a heritage building. **Figure 9** shows an aerial view and side view of the new building of Faculty of Geoinformation & Real Estate (FGRE), UTM.

Figure 9: Simulated heritage building (a) Main entrance side view and aerial view (b) Front left side view and aerial view

B. Data Acquisition

Aerial photography using aerial platform (rotary wing UAV)

Flight and data acquisition is planned in the lab/site with UAV flight planning software, starting from the area of interest (AOI), the required ground sample distance (GSD) or footprint (download from virtual earth), and knowing the intrinsic parameters of the mounted digital camera. Thus fixing the image scale and camera focal length, the flying height is derived. The camera exposure station in space or known as 'waypoint' is computed by fixing the longitudinal and transversal overlap of strips, while the presence of GNSS/INS onboard is usually exploited to guide the image acquisition. During flight, the platform is normally observed with a control station which shows real-time flight data such as position, speed, attitude and distances, GNSS observations, battery or fuel status, rotor speed, etc. This system allows image data acquisition following the computed 'waypoint'. **Figure 10** depicts the operation/procedure involved in acquiring digital aerial photograph of the study area.

Figure 10: Procedure of data acquisition sequence using rotary wing UAV (a) Planning waypoint, (b) Planned data linked to the UAV, (c) Take off autonomously and (d) Aerial photography (e) Landing Autonomously

Aerial photography using ground platform

For the implementation of the digital photographic works, the digital images of object structure must be captured. There are some condition that must be considered for taking photograph such as situation of environment, distance from camera to object/structure, angle of photography, photography configuration and number of. The placing of camera station is very important stage. This is because the captured images will be processed using digital close range photogrammetry software known as PhotoModeler[TM]. In this software, there are certain things that must be consistent with the rules of the PhotoModeler[TM] and the principle of close-range photogrammetry

C. Camera calibration

For highest accuracy, it is necessary in close range photogrammetry to account the variation of lens distortion within the photographic field. The camera needs to be calibrated for the purpose of eliminating image distortion. The process of camera calibration will provide the recovery of camera parameters such as focal length, coordinates of principal point and lens distortion. **Figure 11** shows the calibration plat used in camera calibration process, camera position during image acquisition and camera intersection ray during calibration data processing.
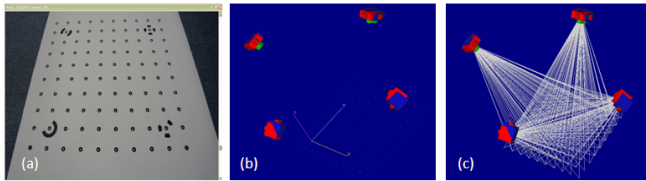
Figure 11: Camera calibration (a) Calibration plat (b) Camera position (c) Intersection bundle of ray

The results of camera calibration are shown in **Table 1**. There are 10 parameters recovered in the camera calibration process which include the focal length (c), principal point (XP and YP), format width (FW), format height (FH), radial lens distortion (K1, K2, K3) and decentering lens distortion (P1, P2).

Table 1: Camera calibration results

	Sony NEX-5N
c (mm)	18.717038 mm
X_P	11.843107 mm
Y_P	7.895478 mm
F_W	24.009243 mm
F_H	15.947883 mm
K_1	$5.505e^{-004}$
K_2	$-6.732e^{-007}$
K_3	$0.000e^{+000}$
P_1	$4.739e^{-005}$
P_2	$-5.980e^{-005}$

D. Digital images processing

Aerial photogrammetry

All acquired digital aerial images were processed using photogrammetric software known as Erdas Imagine[TM]. This software allows triangulation is carried out if more than two photographs exist. After the triangulation is carried out stereoscopic model could be formed and vectoring or digitizing could be performed to produce digital map. The software also allows automatic digital terrain model (DTM) and orthophoto be generated where these products could be used later in GIS environment.

Close range photogrammetry

At this stage, the acquired digital images of the building facades are processed to build the 3D model. This process is performed in the processing software known as PhotoModeler[TM]. This window based software is used to extract measurements and 3D models from digital images. By using a camera as an input device, PhotoModeler[TM] lets the user to capture accurate detail in a short period and generated 3D model. A 3D model is a set of connected 3D points, edges, curves and cylinders or shapes which represent an object. Three dimensional points have coordinate values for each of the Cartesian axes (X, Y, and Z). The points and edges in the 3D model can be connected by lines and/or by patches, called surfaces. This connection is used to visualize in three dimensions when the model is projected onto a flat surface such as a computer monitor or a printed page.

VI. RESULT & DISCUSSION

A. Aerial Photogrammetry using UAV

As mentioned in Section *V (E)*, the digital aerial images are processed using Erdas Imagine[TM] software together with GCPs. The triangulation is performed and later the DTM and subsequently the orthophoto are generated. The original purpose of generating the orthophoto is to obtain the footprint of the new FGRE building and later the 3D model of the building is generated. Subsequently, both the footprint and the 3D model of the building is combined together or integrated on the same coordinate system. **Figure 12** shows the orthophoto of the study area generated using Erdas Imagine[TM] software. These also conclude that data acquired from the rotary wing UAV can be used to produce orthophoto.

New FGRE Building

Figure 12: Orthophoto of study area

B. Close Range Photogrammetry using UAV & Ground Platform

The main objective of this study is to integrate between aerial photogrammetry and close range photogrammetry to develop simulated 3D model of heritage building using the new FGRE building as case study. For aerial photogrammetry, the digital aerial images were acquired and processed successfully **(Figure 12)**. The development of the 3D model of the new FGRE building is still in progress and it takes sometimes to produce it due to its' complex structure. At present a small Tenaga Nasional Building (TNB) within the precinct of new FGRE building was successfully modeled in 3D as shown in **Figure 14.**

Table 2 shows the point ID, RMS residual (pixels), largest residual (pixels), X precision, Y precision, Z precision and angle in PhotoModeler^{TM} software. The RMS residual must not larger than 5 pixels to ensure the measurement project is accurate and always in control. If the RMS value larger than 5 pixels then the point will be marked on the wrong place because of the error especially in image referencing using epipolar matching tool. PhotoModeler^{TM} software contains tools for detecting epipolar traces, and after defining new connecting point, calibration is iteratively upgraded for every imported point or angle. If the user defines a connecting point too far from epipolar trace, that point will not be used for calibration and the point will be marked as potentially wrong. **Figure 13** shows how this software defined a connecting point and detected epipolar trace on edge.

Next the visual output/result process that has been processed using PhotoModeler^{TM} is shown in **Figure 14**. The output started with line, followed by wireframe and shading. This software allows the product exported into AutoCAD for further editing and process depend on application involve. **Figure 15** shows the camera position of UAV platform and close range photogrammetry platform. This remarked that these two platforms can be integrated together in produce 3D model of the small TNB building. By using the same procedure, it is anticipated that the new FGRE building could be modeled successfully too, however, more time is required to achieve the objective of the study.

Figure 13: Epipolar tool in image referencing process using PhotoModelerTM
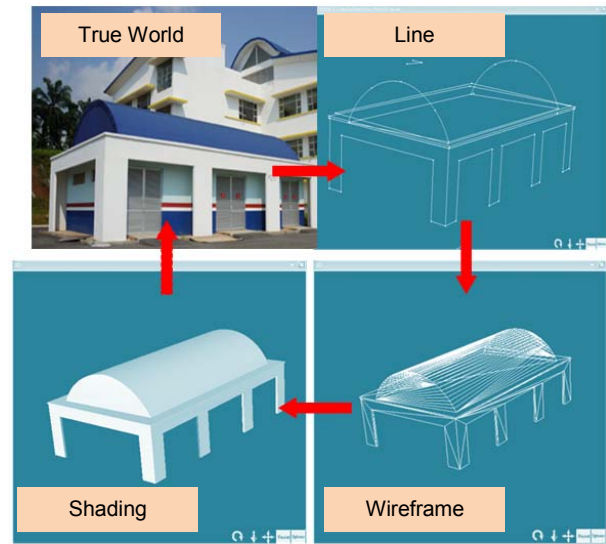
Figure 14: Building structure produced using PhotoModeler^{TM}

Figure 15: Camera position shown in PhotoModeler ^{TM} software

Table 2: Point Table-Quality Assessment

Id	Name	RMS Residual (pixels)	Large Resid (pi...	Photo Largest Residual	Photos	X Precision	Y Precision	Z Precision	Tightness ()	Angle (deg.)	Use In Proces...	Frozen	Ref Che Tag
75		2.9729...	3.3175...	2	1,2,3	0.0010...	0.0008...	0.0014...	0.0025...	69.248...	yes	no	
73		1.7569...	1.9668...	1	1,2,3	0.0010...	0.0008...	0.0013...	0.0014...	67.254...	yes	no	
15		1.4058...	1.9465...	1	1,2,3,4	0.0004...	0.0004...	0.0012...	0.0014...	83.600...	yes	no	
119		1.7160...	1.9179...	4	3,4	0.0007...	0.0005...	0.0011...	0.0005...	28.715...	yes	no	
71		1.3394...	1.3854...	1	1,2,3	0.0012...	0.0010...	0.0015...	0.0010...	65.111...	yes	no	
168		1.2372...	1.3072...	3	3,4	0.0009...	0.0006...	0.0012...	0.0005...	24.166...	yes	no	
17		1.0691...	1.0716...	2	1,2	0.0006...	0.0007...	0.0042...	0.0013...	11.696...	yes	no	
144		0.8015...	0.9027...	4	3,4	0.0006...	0.0006...	0.0010...	0.0002...	26.763...	yes	no	
76		0.7264...	0.8488...	1	1,2,3	0.0006...	0.0006...	0.0012...	0.0006...	74.252...	yes	no	
107		0.7929...	0.8217...	3	3,4	0.0006...	0.0005...	0.0010...	0.0003...	29.306...	yes	no	
110		0.7973...	0.8064...	3	3,4	0.0006...	0.0006...	0.0009...	0.0002...	35.548...	yes	no	
108		0.7173...	0.7415...	3	3,4	0.0007...	0.0004...	0.0009...	0.0002...	29.348...	yes	no	
16		0.4863...	0.6992...	2	1,2,3,4	0.0005...	0.0002...	0.0008...	0.0003...	73.434...	yes	no	
72		0.4951...	0.6922...	3	1,2,3	0.0010...	0.0011...	0.0016...	0.0003...	70.756...	yes	no	
78		0.6918...	0.6922...	1	1,2	0.0007...	0.0009...	0.0040...	0.0008...	11.499...	yes	no	
152		0.6507...	0.6885...	3	3,4	0.0009...	0.0006...	0.0012...	0.0006...	24.194...	yes	no	
103		0.6441...	0.6842...	3	3,4	0.0010...	0.0006...	0.0013...	0.0003...	23.190...	yes	no	
13		0.4431...	0.6083...	4	1,2,3,4	0.0004...	0.0006...	0.0010...	0.0002...	87.135...	yes	no	
115		0.5823...	0.6031...	4	3,4	0.0006...	0.0003...	0.0009...	0.0001...	46.247...	yes	no	
112		0.5779...	0.5789...	3	3,4	0.0006...	0.0002...	0.0008...	0.0001...	37.461...	yes	no	
56		0.3581...	0.5746...	1	1,2,3	0.0006...	0.0006...	0.0008...	0.0003...	70.283...	yes	no	
57		0.4105...	0.5263...	1	1,2,3	0.0007...	0.0004...	0.0008...	0.0001...	64.520...	yes	no	
12		0.5123...	0.5179...	1	1,2	0.0008...	0.0008...	0.0049...	0.0006...	11.287...	yes	no	
19		0.4694...	0.4820...	1	1,2	0.0015...	0.0006...	0.0043...	0.0005...	10.563...	yes	no	
154		0.4544...	0.4702...	1	3,4	0.0006...	0.0005...	0.0010...	0.0001...	29.213...	yes	no	
4		0.4514...	0.4569...	1	3,4	0.0017...	0.0016...	0.0055...	0.0006...	9.6111...	yes	no	

VII. CONCLUSION

The main advantage of photogrammetric modeling is the economy of procedure. The terrain work is much shorter compared to on site measuring dimensions. Time needed for modeling is about the same as time needed for drawing object. Greater quantity of information is obtained using 3D modeling compared to drawing in 2D. Model can be exported in other modeling or animation software which allows further treatment (element editing, adding lighting, creating dynamic view, etc.). Depending on the software quality, accuracy and precision of model can be higher. Important advantage of photogrammetric modeling is obtaining information about inapproachable structures or some of its elements. This is essential for heritage objects or buildings which are located on rough terrain, or structure elements having high and inapproachable facades. In contrast to traditional architectural drawings of heritage building or objects which contain only dimensions, model also contains detailed information about used materials. This is especially important if damaged parts of structure appear which are then dimensioned altogether with the model.

The 3D building structure of this simulation building had successfully been produced using digital close range photogrammetry (DCRP) approach integrated with images captured using rotary wing UAV which is based on aerial photogrammetric technique. Our next task is to model the entire new FGRE building based on DCRP and subsequently combine with the footprint of the building on a common reference coordinate system and GIS environment. In geomatic, many type of products have been used by archeologist for decades such as topgraphic map, thematic map, digital map, orthophoto, topographic database, engineering plan, certified plan (CP) and others. However, product such as 3D model of building could assist them to produce better recording especially for heritage building.

REFERENCES

[1] Anuar Ahmad, 2011. Digital Mapping Using Low Altitude UAV. *Pertanika Journal of Science & Technology, Vol. 19(S), Oct. 2011, pp 51-58.*

[2] Anuar Ahmad, Khairul Nizam Taha and Wan A Aziz, 2011. Unmanned Aerial Vehicle Technology for Large Scale Mapping. *Buletin GIS, Jabatan Ukur & Pemetaan Malaysia. Vol 2, pp 98-108*

[3] Anuar Ahmad & Zulkarnaini Mat Amin, 1998. *Unsur-unsur Fotogrametri dengan Pentafsiran Fotoudara dan Penderiaan Jauh.* Universiti Teknologi Malaysia.

[4] American Institute of Aeronautics and Astronautics (AIAA), 2004. *Committee of Standards "Lexicon of UAV/ROA Terminology".*

[5] Blyenburgh, P. V. 1999. UAVs: An overview. *Air & Space Eur.*1(5): 43-47.

[6] Bendea, H., Chiabrando, F., Giulio Tonolo, F., Marenchino, D., 2007. Mapping of archaeological areas using a low-cost UAV. The Augusta Bagiennorum Test Site. In: *ISPRS Archives Vol. XXXVI-5/W1, pp. 117-122.*

[7] Chiabrando, F. Nex, D. Piatti, F. Rinaudo., 2011. UAV and RPV systems for photogrammetric surveys in archaelogical areas: two tests in the Piedmont region (Italy). *Journal of Archaeological Science 38 (2011) 697-710*

[8] Colomina, I., Blázquez, M., Molina, P., Parés, M.E., and M. Wis, 2008. Towards a New Paradigm for High-Resolution Low- Cost Photogrammetry and Remote Sensing. *IAPRSSIS, 37(B1): 1201-1206.*

[9] Eisenbeiss, H., 2005. Applications of photogrammetric processing using an autonomousmodel helicopter. *Proceedings of theXXCIPASymposium,Torino(Italy),1e6.*

[10] Eisenbeiss, H., 2008. The Autonomous Mini Helicopter: a Powerful Platform for Mobile Mapping. *IAPRSSIS*, 37(B1): 977-983.

[11] Enderson, H., 2012. Photo Intersection Overview. *http://www.hugha.co.uk/index.html*

[12] Everaerts, J., 2008. The use of unmanned aerial vehicles (UAVS) for remote sensing and mapping. In: *ISPRS Archives Vol. XXXVII. ISPRS Congress, Beijing, China, pp. 1187-1192.*

[13] Faez M Hassan, Lim, H.S and Mat Jafri, M.Z., 2011. Cropcam UAV for Land USe/Land Cover Mapping oevr Penang Island, Malaysia. *Pertanika Journal of Science & Technology, Vol. 19(S), Oct. 2011, pp 69-76.*

[14] Grenzdörffer, G.J., Engel, A., and B. Teichert, 2008. The Photogrammetric Potential of Low-Cost UAVs in Forestry and Agriculture. *IAPRSSIS*, 37(B1): 1207-1213.

[15] Kerle, N., Heuel, S., and N. Pfeifer, 2008. Real-time data collection and information generation using airborne sensors. In: Zlatanova, S., and J. Li (ed.s), *Geospatial Information for Emergency Response.* Taylor & Francis, London (UK), pp. 43-74.

[16] Khairul Nizam Tahar and Anuar Ahmad, 2012. A simulation tudy on the capabilities of rotor wing unmanned aerial vehicle in aerial terrain mapping. *International Journal of Physical Sciences, Vol. 7(8), pp 1300-1306.*

[17] Kimball, L.R., 2011, Aerial photographic equipment. *CDI Infrastructure, LLC.*

[18] Nebiker, S., Annen, A., Scherrer, M., and D. Oesch, 2008. A Light-weight Multispectral Sensor for Micro UAV - Opportunities for Very High Resolution Airborne Remote Sensing. *IAPRSSIS*, 37(B1): 1193-1200.

[19] Püschel, H., Sauerbier, M., Eisenbeiss, H., 2008. A 3d model of castle Landenberg (CH) from combined photogrammetric processing of terrestrial and UAV-based Images. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 37, 93-98.*

[20] Skarlatos, D., Theodoridou, S., Glabenas, D., 2004. Archaeological surveys in Greece using radio-controlled helicopter. *In: Proceedings of FIGWorkingWeek, Athens (Greece), pp. 1e15.*

[21] UNESCO, 2002. For UNESCO, Space Technologies should be Harnessed for Sustainable Development. *Unesco Press release No. 2002-77*

[22] Vesna Stojaković, 2008. Terrestrial Photogrammetry and Application to Modeling Architecture Objects. *Facta Universitatis, Architecture and Civil Engineering Vol 6, pp. 113-125*

[23] Vosselman, 2010 Guided real-time close-range photogrammetry for 3D modelling, *PhD Project, Faculty of Geo-Information Science and Earth Observation, ITC*

[24] Wang, J., Garratt, M., Lambert, A., Wang, J.J., Han, S., and D. Sinclair, 2008. Integration of GPS/INS/Vision Sensors to Navigate Unmanned Aerial Vehicles. *IAPRSSIS*, 37(B1): 963-969.

[25] Ying-Chih, L and Shau-Shiun J., 2011. Altitude estimation based on fusion of Gyroscopes and Single Antenna GPS for small UAV under the influence of vibration. *GPS Solution, 15, 67-77*

[26] Zongjian, L., 2008. UAV for Mapping – Low Altitude Photogrammetric Survey. *IAPRSSIS*, 37(B1): 1183-11

Web-based Adaptive Audio-Therapy Recommender System

Muhammad Khidhir Salleh, Norliza Mohamad Zaini
Faculty of Electrical Engineering
Universiti Teknologi Mara
Shah Alam, Malaysia
khidhir193@gmail.com, drnorliza@salam.uitm.edu.my

Abstract—**This paper presents an online system that aims to recommend suitable audio-therapy to improve mental states of a user. It is designed and implemented with a capacity to analyse user's current mental states and based on the acquired result, recommend a suitable audio-therapy for the user. The web based system was developed with a number of components. The core engine of the system is an online questionnaire component used to determine the user's current mental states and an audio-therapy matching component. The latter is driven by an algorithm created to match the user's score derived from the answered questionnaire to suitable therapy audios, recommended for the purpose of improving the user's mental states. The system was implemented using database-driven PHP server script, enabling dynamic functionalities and features offered to both end-users and psychiatrists.**

Keywords- audio therapy, questionnaire, recommender, improve mental states, online application, depression

I. INTRODUCTION

The world's activities become more challenging from day to day and consequently the number of people with mental illness keeps increasing considerably. According to the World Health Organization (WHO), in year 2020 depression will be the second major disability leading to various diseases in the world. This fact becomes our main motive where the difficulty in obtaining treatment for mental illness is a major issue that we are focusing on. For most people, treatments for mental illness are considered too troublesome, which require sacrifice in terms of efforts, time and cost. Acknowledging this reality, this project has tried to come out with a convenient system, specifically designed for the purpose of improving users' mental states. Different brainwave patterns are associated with different mental states or consciousness. In this regards, brainwaves are normally categorized by five main power spectrums, namely delta, theta, alpha, beta and high beta. Table 1 summarizes the association between these different brainwave frequency-groups to the associated mental states.

Our main aim was to produce a system that is accessible almost anywhere and anytime. Based on this requirement, the best choice was to implement a web-based system. The main rationale behind this decision was due to the fact that web-browser application is one of the most common applications supported by all operating systems and platforms such as on personal computers, hand phones and other portable electronic devices. To accomplish this objective, we have implemented the system using PHP scripting language, offering a dynamic online–based application for the users.

Table 1. Brainwave Frequencies associated with different mental states [1]

Brainwave Frequency Range	Associated Mental States
Delta frequencies (1-4 Hz)	deep sleep
Theta frequencies (4-8 Hz)	light sleep creativity insight
Alpha frequencies (8-12 Hz)	a calm and peaceful yet alert state
Beta frequencies (13-21 Hz)	Thinking focusing state
High beta frequencies (20-32 Hz)	intensity or anxiety

In order to make this system more beneficial, it is designed to enable it to capture and collect useful data entered by the users. Manipulating such data, this system is applicable as a useful tool for psychiatrists to view the user records. This is particularly helpful in a situation where a psychiatrist is needed to analyse the user's data further. In order to improve the current mental states of a user, this system will recommend a therapy audio for the user to listen. This relates to works done on brainwave entrainment as therapy audio, e.g. binaural beats entrainment that leads to alterations of brainwave states, aiming to create a positive effect in the physiological and psychological state of human subjects, i.e. by restoring 'normal' healthy brain wave patterns [1,2,3]. The recommendation on such therapeutic audio is made based the score acquired by the user from the mental- state analysis process, i.e. by using questionnaire. Such recommendation was specifically made with the aim of treating the user's depression (if present) or improving his or her mental state.

II. SYSTEM DESIGN

The system consists of several components, namely the Login Form Component, Member Registration and Scoring Component, One-time Scoring Component, Questionnaire Component, User Contact Profile Component, Audio Matching Component, Diagnostic Component, Patient Record Component, User Database Component, Edit Question Component and Database Component. Fig. 1 shows the architecture of our online adaptive system. There are about eleven components that dependently work with each other. As shown by the figure, most components require access to the

database component, which can be regarded as the backbone of the system. The only component that does not interact with the database component is the one-time scoring component. Three ways a user may access the questionnaire component is either via the login form, one-time scoring or member registration and scoring component. Only a user who has successfully passes through the login component with required authentication will have access to the whole system. Details for each of these components will be described in the following section.
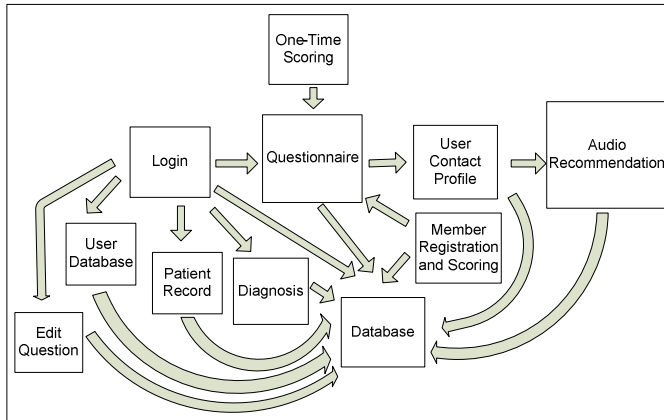
A. System Architecture

Fig. 1 Architecture of online adaptive system

B. System Component

1) Login Form Component: The first component of this system is the login form as shown in Fig. 2. This login form will be able to authenticate any user into the system, given that the user has provided correct login and password. For security purpose, the authentication process is also useful to avoid any attempt of SQL injection. Once a user is authorized into the system, the user will be assigned to a particular access level, which indicates an authorization category that allows user access to specific components of the system, while denying access to other components. Adopting this approach can avoid misuse of system functionalities by categorizing users into administrator and patient groups. Each of these groups will be assigned several accessible and inaccessible components. Upon successful login, the system will display the logout function for the user to exit and invalidate their session from the system. The reason for this action is to ensure that once the user has finished completing their task or decided to end their session, they are able to end their session immediately. This is enabled mainly for security purpose, which will prohibit other users of the same machine (PC) from misusing another user's session.

Fig. 2 Login form component

2) Member Registration and Scoring Component: Next component in this system is the member registration and scoring form. This form performs two functions at once. The first function is to register a new user into the system, while the second function is to record other necessary profile details such as the user's gender, age, and employment status. This information is important to be tailored with the questionnaire component. For the user registration function, a user must provide three inputs, which are their unique username, password and its confirmation. Upon reception of these inputs, the system will first make sure that the username provided has never been used before and also the password and its confirmation are matched. If these requirements are not satisfied, the data will not be recorded by the system and the user will be redirected back to the previous page with an error notification due to incorrect inputs entered by the user.

The second functionality of this member registration and scoring component is to record necessary user's profile details. As mentioned earlier, although this function may not be important to the system's core functionality itself, but it is greatly important for data analysis and statistics purpose, e.g. for depression analysis. So, this function is set as compulsory that requires a user to select appropriate options to proceed to other components in this system. Failure to obtain these additional inputs from the user will result in difficulty to generate statistic and analysis for depression. Fig. 3 below illustrates the registration component's interface.

Fig. 3 User registration

3) One-time Scoring Component: The one-time scoring component (see Fig. 4) basically skips any user registration or profile recording, which allows the user to directly proceed to the next component of the system, without requiring any authorization process. In this case, the system will view the user as a guest who is allowed a limited access to the system's components. As reflected by its name, this component is mainly used by the user to access the questionnaire component and proceed to other components that are related to the questionnaire component.

Start The Test

I just want to take the test

Just Take The Test

Fig. 4 *One-time Scoring* Component

4) Questionnaire Component: The next and most important component of this system is the questionnaire component. It consists of two parts, which are the question and results parts. The question part will display all questions required to be answered by the user. The set of questions displayed by the system are dynamic in the sense that they are changeable over time when necessary. The questions are in the form of objective or multiple choice questions. For this type of questions, the system is able to accept only one selected answer from a set of optional answers. Unanswered question is unacceptable by the system and this is based on the questionnaire answering procedure that does not allow a user to skip any of the questions. These questions are used to screen the user's current mental states, e.g. to detect whether they have symptoms of depression. Thus they do not have the right or wrong answer, but each of the given answers leads to a different score.

The result part will display the total score counted for the user based on the answers given for the whole questionnaire set. Besides the score, the interpretation or meaning of the achieved score is also presented to the user, which indicates the user's level of depression, e.g. normal, depressed, etc. The process of calculating the score was made to exactly follow the scoring scale adopted by the screening questionnaire set. This way, the automated scoring calculation should not be different from the manual calculation. The system is also capable of determining the user's answering process, e.g. whether the user has answered all questions or has stopped in the middle of the question-answering process. Fig. 5 illustrates a sample question displayed by the questionnaire component.

Questions 1:

I was bothered by things that usually dont bother me.

⊙ Rarely or none of the time (less than 1 day)

⊙ Some or a little of the time (1- 2 days)

⊙ Occasionally or a moderate amount of time (3-4 days)

⊙ All of the time (5-7 days)

Next Question

Fig. 5 Questionnaire component's interface

5) User Contact Profile Component: Another component of our online adaptive audio-therapy system is a component used to update the user's contact profile. Its interface is shown in Fig. 6. The difference between the interface generated by this component with the interface generated by the member registration and scoring component is on the details requested

from the user. The user's details requested by the form before the questionnaire are the user's background information. Such information is necessary to analyse the characteristics of the user, in relation to the questionnaire given. On the other hand, the user's contact profile basically provides the information on how a particular user can be contacted or reached.

First Name :

Last Name :

Address :

Phone Number :

Update Profile

Fig. 6 User contact profile component

The user registration interface was specifically designed in such a way that it does not look too complicated to prevent the user from feeling difficult to initiate the test at the first place. Once a user has filled in all requested details and completed the questionnaire, only then the user will be provided with this user contact profile to be completed. However, this form is set as an optional form to be filled in by the user and the user may straight away proceed to another page if they choose not to provide their contact details in the system.

6) Audio Matching Component: The main role of the audio-therapy matching component is to recommend a suitable audio-therapy to the user. The recommendation is made based on the score obtained by the user from the answered questionnaire. Different therapy-audios are recommended to different users depending on their scores, which basically indicate the different mental states or level of depression (if present). The main aim of this component is to recommend a suitable therapy-audio specifically to improve the user's mental states or level of depression, e.g. by using an entrainment audio. In general, the whole system is coined to be an adaptive system based on this particular feature it offers. The correct pair-matching between the scores acquired by users and the audio-therapy deemed as suitable should be based on the judgements made by experts in the relevant fields, e.g. brainwave entrainment experts. Otherwise, this can be a research issue own its own. Having said that, the decision process in matching the user-scores and audio-therapy pairs, is beyond the scope of this paper, where the proposed system only provides the platform in enabling such matching or mapping process to occur.

7) Diagnostic Component: The diagnostic component does not differ much from the questionnaire component in terms of their operations. However the main difference between these two components is that the diagnostic component is used by a system administrator, who most likely is the psychiatrist. This diagnostic component is referred to by the psychiatrist to analyse the user or patient's condition when they meet. This component is divided into two parts. The first part is the

questionnaire set, which is changeable to a different set if necessary, while the second part is the result part. To view and analyze the previous record for an individual patient, the psychiatrist may choose a particular patient's username from the users' list. For diagnostic purpose, based on initial interactions with the patient, the diagnostic questionnaire will be filled in by the psychiatrists to help them diagnose the patient's mental condition. Once a questionnaire is completed, the result will be displayed indicating the patient's condition, e.g. either the patient is experiencing major depressive episode or recurrent major depressive episode. This result will be captured for patient record purpose.

Fig. 7 Patient record component

8) Patient Record Component: The main function of the patient record component (see Fig. 7) is for the psychiatrist or administrator to view the records of a particular patient. This important feature will help the psychiatrist to analyze the patient's mental condition in a long run. The administrator or psychiatrist just needs to enter a patient's username to view the patient's past records. This component is strictly offered only for administrator or psychiatrist use since patients' data need to be kept private and confidential. The patient record includes the patient's score obtained from the filled-in questionnaire and also their profile and contact details.

Fig. 8 User database component

9) User Database Component: To promote research and statistical analysis, which may be necessary in the future, this system includes an important component which is the user database (see Fig. 8). The User database component and Patient Record component have quite similar functionality but with different purpose. The user database component will not display user contact details as in the Patient Record component. However this component will display all users' record in terms of the user profile details such as age, gender and employment status and also the questionnaire score obtained by each patient with the recorded date and time of when the questionnaire was submitted. In addition, this

component offers a filtering mechanism where it allows the administrator or psychiatrist to choose users with specific criteria. For example, the administrator can choose to list only male users, who are around 25 years of age and have completed the questionnaire. Basically this system allows the administrator, based on their preference to get a list of users who satisfy a set of preferred criteria of any combination. This feature is really helpful in statistical data collection and will help to improve the efficiency of related research work.

10) Edit Question Component: Changes in the questionnaire sets are sometimes required and necessary at some point of time, e.g. there is a new questionnaire set to be adopted or some corrections are required to be made to the existing questionnaires. Acknowledging this requirement, the Edit Question component is developed to allow the administrator to perform necessary changes to the questionnaire set. The types of changes allowable include modifications to the content, i.e. the questions, answers and scores being defined for the selected answers. Additions of new question-answer-score sets are also allowed to be performed to the existing questionnaire set.

The system is made to be more dynamic in adapting the questionnaires to different sets where the system allows the number of multiple choice answers of each question to vary between two to four optional answers. This way, the administrator is able to insert and manage the number of multiple choice answers provided to the user. A question with more than 4 multiple choice answers is believed to be too confusing or difficult for users especially those with depression to answer and thus we limit the choice up to the maximum of four answers. Any modifications performed to the questionnaires are directly reflected in the database and thus will take effect immediately. As a step of caution, the administrator is only allowed to make any changes to the questionnaire when no user is answering the questionnaire. Without this caution, activities concurrently performed by the administrator and users may lead to unreliable data being recorded in the database.

11) Database Component: The database component is an important component that becomes the backbone of the system. It keeps all of users' details that include users' profile and contact information, authentication details and the obtained score. All components of the system are database-driven except for the one-time scoring component since it does not require any data to be captured or recorded. The database was specifically designed in such a way that the stored data are well organized to allow easy data retrieval.

C. Questionnaire

The questionnaire component handles two types of questionnaires. The first type is known as the screening questionnaire such as the Depression Anxiety Stress Scales (DASS), Hospital Anxiety and Depression Scale (HADS) and Geriatric Depression Scale [4]. This type of questionnaire is patient-rated type of questionnaire where the questionnaire should be answered by the users or patients themselves. For

our system, the screening questionnaire is specifically used to test or measure the current mental state or depression level of an individual. The results obtained from this screening questionnaire will be used by the audio-matching component to determine and recommend suitable audio for the user as a therapy. Such screening questionnaire is capable of detecting users' short term depression level, e.g. for a period of one week. This detection will be very useful to treat or improve mental states of the user with such condition but not reliable to determine if the user has recurrent depressive episode.

To determine if a user is having recurrent depressive episode, the second type of questionnaire is adopted, which is known as the diagnostic questionnaire. Some examples of diagnostic questionnaire are M.I.N.I, Montgomery-Asberg Depression Scale (MADRS) and Hamilton Rating Scale for Depression [5]. The main difference between diagnostic questionnaires with screening questionnaires is that the diagnostic questionnaire can only be rated by psychiatrist. Only psychiatrist will have access to this type of questionnaire, which will be used to examine user's mental condition. There will be some cases, where the users themselves meet the psychiatrist after knowing that they are having depression or tendency of being depressed based on the answered screening questionnaire. For such cases, the psychiatrist will be able to retrieve the patient's screening records prior to analysing the patient's mental condition by using the diagnostic questionnaire, i.e. to determine if the user has recurrent depressive episode.

As a whole, the system can help users to determine their current mental states, while helping the psychiatrist to examine users to determine the type of depression being experienced. From these offered features, this system can be seen as a complete set or tool that can be used to control or improve users' mental states and treating the user depression if necessary.

III.   SYSTEM IMPLEMENTATION

A.   Programming Language, Scripting and Database

1)   Several factors have been considered in developing and implementing the online adaptive system for depression such as compatibility, scalability, platform, usability, cost, performance and speed. Considering all these factors, this system will be built on a web server using PHP and HTML as the main scripting language and MySQL as the database server to fulfil the design.

2)   Being a dynamic PHP-based system, the system is more useful by integrating it with a database. PHP may support many types of databases such as MySQL, Informix, Oracle, Sybase, Solid, PostgreSQL, Generic ODBC and other type of database. For this project, MySQL server is used as its database server since MySQL is suitable either for a small or large applications and it supports standard SQL. Besides, MySQL is free to use, which makes it more preferred than the other type of databases.

Column	Type	Null
id	int(11)	No
username	varchar(20)	No
password	varchar(20)	No
type	varchar(20)	No

Fig. 9 Member table

The database is designed to consist of three tables. The first table is named as "member" (see Fig. 9), which is used to keep records on registered username, password and the user's authorization level. This table is accessed by the login-form, member-registration and scoring components. The login-form component access to the database is specific to authorization process, while the member-registration and scoring component access to the database is more to perform insertion of data. The users' authorization level recorded in this table indicates the types of allowable users' access to the different features and components of the system.

Column	Type	Null
id	int(11)	No
username	varchar(20)	No
gender	varchar(10)	No
age	varchar(10)	No
employ_status	varchar(20)	No
start_time	datetime	No
end_time	datetime	No
mark	int(100)	No

Fig. 10 Patient table

Patient table as shown in Fig. 10 is another table stored in the database. This table contains eight fields that hold users' details and questionnaire results. The member-registration, scoring component and the questionnaire component are ones that perform data insertion and update onto this table. On the other hand, the user-database and patient-record components are ones that fetch data from this table.

Another table is the Profile table, which keeps record of users' contact details. The components related with this table are the User Contact Profile and Patient Record component. User Contact Profile component performs insertion, update and retrieval from this table, while the Patient Record component will only fetch data from this table.

B.   Relation Page and Components

Each of the components described in the system design are implemented as dynamic PHP pages, viewable on users' web browsers. The index page is populated by four components which are the login form, one-time scoring, member registration and scoring component. The index page has several dynamic views or appearances depending on the user's level authorized into the system. By default, a new incoming user or guest is able to see all components' interface on the index page. However, for users who have registered or logged in as a patient, the one-time scoring component's interface (with username and password input box inside the registration form) will not appear on the index page. On the other hand, for a user logged in as the administrator, interfaces of all other components will be hidden except for the login component's interface and a link to the administrator page.

The test page is driven by the questionnaire component, which captures a specific user-session for a particular user. Once a user has answered all questions displayed by the system, the result will be presented, after which the user will be given access to the update profile page. This page is driven by the user contact profile component. Based on the unique user-session created earlier for a particular logged-in user, the system only allows the user to update their own profile. Once the user has done with the user contact profile or chooses to skip the component, the user will be guided onto the audio page, which will recommend audio-therapy especially to improve the user's mental states. The audio recommendation page is driven by the audio matching component.

An administrator user is given a different authorization level compared to an end user. The administrator will be able to access the diagnostic page, which is used to diagnose a patient. This page is driven by the diagnostic component. Another page accessible by the administrator is the patient record page, which displays a particular patient's record. This page is driven by the patient record component. As for the database page, which is also accessible by the administrator, the page is displays all users' record filtered by the criteria chosen by the administrator.

IV. CONCLUSIONS

This Online Adaptive System has been designed and implemented specifically aiming to address the issues of obtaining easier and portable treatments for mental illness. By using this online system, the users are given access to get the recommended treatment, which is in the form of audio-therapy, particularly aimed to improve the users' mental states at their own convenience. The system is also designed to be a database-driven system, specifically to allow centralized storage and maintenance of users' records. Moreover, appropriate therapy-audios will be recommended by the system based on the score acquired by the user from the answered screening questionnaire. The mechanism used to match a user's score with suitable therapy audios is one of core engine and the main contribution of the system. Although the suggested therapy-audio is not particularly aiming to control the user's mental state in a long run, but the recommended therapy-audio is expected to be able to instantly reduce the user's depression level and to improve the user's mental states. As a whole, the system can be seen as a great system-combination with a mechanism to analyze the user's current mental condition, providing results to be employed by another mechanism, which recommends suitable therapy audio as a treatment to the user. Another mechanism especially offered to the administrator or psychiatrist is one that keeps track of the patients' profile and analysis records. Having said that, the system is proposed as a package offering benefits for both the end-user patients and the psychiatrists.

ACKNOWLEDGMENT

REFERENCES

[1] T.L. Huang, C. Charyton. "A comprehensive review of the psychological effects of brainwave entrainment". Alternative Therapies In Health And Medicine. 2008; Vol. 14, no. 5. (September 1): 38.

[2] J. Leeds. "The Power of Sound: How to Manage Your Personal Soundscape for a Vital, Productive, and Healthy Life". Rochester, Vermont: Healing Arts Press; 2001.

[3] F. Atwater. "The Monroe Institute's Hemi-sync Process, A Theoretical Perspective". The Monroe Institute of Applied Sciences, 1988. pp. 4, 7.

[4] Maratos AS, Gold C, Wang X and Crawford MJ, "Music therapy for depression (Review)," in The Cochrane Library, 2008, Issue 1.

[5] D. Sheehan, J. Janavs, R. Baker, K. Harnett-Sheehan, E. Knapp, M. Sheehan Y. Lecrubier, E. Weiller, T. Hergueta, P. Amorim, L. I. Bonora and J. P. Lépine, "MINI International Neuropsychiatric Interview," M.I.N.I. 5.0.0, July 1, 2006.

Research in Characterization of Cyclic Load Parameters using Fiber Bragg Gratings System

Mohamad Adli Bin Adam

Faculty of Mechanical Engineering
Universiti Teknologi MARA
Shah Alam, Malaysia
adliadam@gmail.com

Wahyu Kuntjoro[1], Mohd Kamil Bin Abdul Rahman[2]

Faculty of Mechanical Engineering[1], Faculty of Applied Science[2]
Universiti Teknologi MARA
Shah Alam, Malaysia
wkuntjoro@yahoo.com[1], drkamil@salam.uitm.edu.my[2]

Abstract— **Fatigue is the primary cause of structural failure. Fatigue failures are caused by fatigue loadings in the form of cyclic loads. To avoid fatigue failures, structures need monitoring. This paper reports a study of optical fiber Bragg grating responses under variable amplitude fatigue load. Use of optical fiber in monitoring structural condition becomes extensive research nowadays due to its advantageous characteristics than conventional method such as using strain gauges. A standard specimen (ASTM E606) is used for initial study. The specimens were attached with fiber Bragg grating sensor and test configuration were set up. All the data and information were analyzed and comparisons with strain gauge sensor were performed. It was observed that response from FBG sensor is consistent with applied cyclic loads.**

Keywords-fatigue; structural monitoring; fiber Bragg gratings; strain gauge

I. INTRODUCTION

Fatigue is a beginning for a catastrophic structural failure. When repeated stresses are applied to the structure, fatigue failure will start to happen. Fatigue failures happen in a sudden and total nature, thus it is very dangerous. Continuous monitoring on structure can avoid the structure from fatigue failure.

Structural health monitoring (SHM) is a process of gaining information of the current condition of a structure. Early detection of damage can be spotted in this method; hence the life of structure can be extended. The prediction of structural deterioration over time can be made so remaining life can be calculated by analyzing the gathered history data through SHM methodology [1].

In analyzing fatigue failures by using SHM, fatigue parameters must be known so the prediction of remaining life time of the structure can be done. To obtain these parameters, a good sensor shall be attached to the structure. A sensor must be sensitive to the measured parameters, do not influence the behavior of structure and also not responding to other constraints. Sensor also must be durable, sensitive to small changes, non-electric, and immune to intensity fluctuation.

Fiber Bragg grating (FBG) provides advantages compared to conventional sensor. FBG sensor is sensitive to the change of physical parameters in the form of wavelength shift of the reflected light. Since the reflected light has a very narrow bandwidth (typically 0.1-0.5 nm), it is ideally suited for sensing applications [2].

A research on the application of optical fiber Bragg grating sensor for structure monitoring was reported by M. Kamil Abd Rahman et.al. They were able to show that the Fiber Bragg Grating (FBG) sensors show matching strains with applied forces to the vibrating wire strain gauge (VWSG) sensor had a small error. The experiment done using three FBG sensors and vibrating wire strain gauge (VWSG) were mounted on 650mm (R12.5) mild steel rod. The tensile stress test was conducted. It can be concluded that the sensitivity of FBG sensor is substantially high, the measurement of tensile stress and other related parameters on structures and environment can be accurately and precisely determined [3].

Development of fatigue monitoring to fighter aircrafts were reported by W Kuntjoro et.al [4]. Local strain approach has been used for strain estimation techniques on life prediction. Neuber's rule was used to predict the notch geometry effect. Rain flow cycle counting method was employed to extract cycle from random loading spectrum. Mean stress correction factor was based on Smith-Watson-Topper (SWT) equation and Palmgren-Miner rule was used to conduct summation of the cycle damage. The monitoring was based on the accelerometer and strain-gauges installed on the airplanes.

A study of FBG sensor and electrical strain gauge for strain measurement by B. A. Tahir et.al was conducted to make a comparison of both sensor performance and advantages. It alleged that FBG sensor data correlates better with the theoretical calculation than the electrical cantilever gauge value [5]. FBG sensor also has good qualities over electrical strain gauge. FBG are immune to electromagnetic interface, can measure very high strain (>10 000μm/m), and having good corrosion resistance [3][5]. On a single fiber, many FBG sensors can be located in a row and it will ease the installation [6][7].

In this paper, a research with the general objective of studying the response of optical fiber Bragg grating under the variable of amplitude load is reported. Based on the response of optical fiber, the characterization of fatigue load parameters can be eventually performed. In this research, a standard specimen (ASTM E606) is used for initial study. The specimens were attached with fiber Bragg grating sensor and test configuration were set up. All the data and information were analyzed and a comparison with strain gauge sensor was performed. It is planned for the follow on study, the structure to be attached by fiber Bragg grating system is a wing-box aircraft structure..

II. FATIGUE FAILURE

In static, when yield strength point was exceeded a large deflection on structure occurs then failure will happen. Thus many static failures can be detected with naked eye. But fatigue failure is contradictory to that. An unexpected and total failure will occur. A fatigue failure has a similarity with a brittle fracture where the fracture surfaces are flat and perpendicular to the stress axis with the absence of necking [8]. The initial fatigue failure occurs when one or more microcrack(s) initiate due to cyclic plastic deformation followed by crystallographic propagation extension. Then microcrack(s) become macrocrack(s) and will reach a critical size of cracks. The structure will suddenly fracture.

### A. Characterizing Fluctuating Stresses.

Fluctuating stress is often in the form of a sinusoidal pattern. Is has been found that in periodic patterns exhibiting a single maximum and a single minimum of force, the shape of the wave is not important, but the peaks on the high side (maximum) and the low side (minimum) are important [8]. There are three basic factors necessary to cause fatigue failure. These are maximum tensile stress of sufficiently high value, large enough variation or fluctuation in the applied stress, and sufficiently large number of cycles of the applied stress. Below is figure 1 which illustrates some of the various stress-time relations.
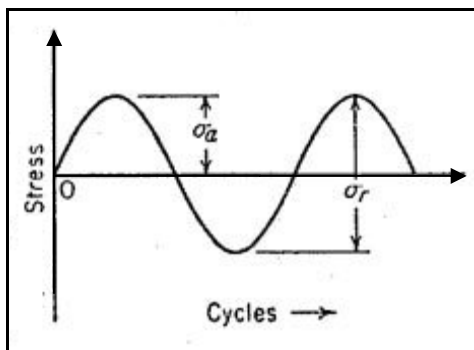
(a)
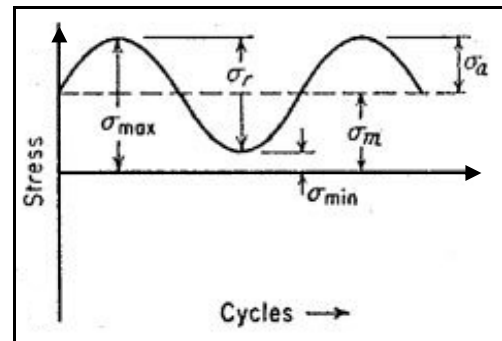
(b)
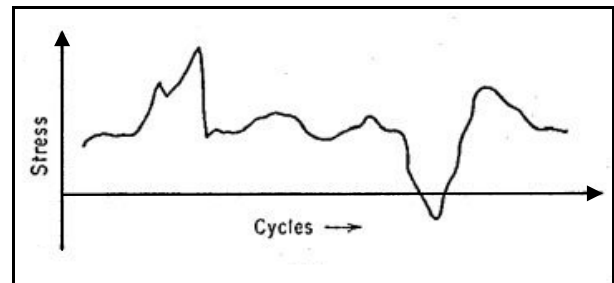
(c)

Fig. 1.Illustration of various cyclic loading

Fig. 1(a) illustrates a completely reversed cycle of stress of sinusoidal form. For this type of stress cycle the maximum and minimum stresses are equal. Tensile stress is considered positive, and compressive stress is negative. Fig. 1(b) illustrates a repeated stress cycle in which the maximum stress σ_{max} (R_{max}) and minimum stress σ_{min} (R_{min}) are not equal. In this illustration they are both tension, but a repeated stress cycle could just as well contain maximum and minimum stresses of opposite signs or both in compression. Fig. 1(c) illustrates a complicated stress cycle which might be encountered in a part such as an aircraft wing which is subjected to periodic unpredictable overloads due to gusts [9].

Fatigue loading normally is random loads. For fighter airplanes, the load spike can be very high and approaching the yield limit of material. FBG is appropriate for handling this fluctuating random signal and can deal with a very high strain value.

III. FIBER BRAGG GRATING (FBG)

A. History

In 1978, Hill *et al.* demonstrated a first formation of permanent gratings in an optical fiber. Initially, the gratings were fabricated using a visible laser propagating along the fiber core. In 1989, Meltz and colleagues demonstrated the modern transverse holographic technique from the side of the fiber utilizing the interference pattern of ultraviolet light [10].

In early 1990s, fiber Bragg grating (FBG) sensors have been intensively developed due to their desirable advantages such as the small size, absolute measurement capability, immunity to electromagnetic interference, wavelength multiplexing, and distributed sensing possibilities [11].

B.   Theory

The fundamental of FBG operation is when light travelling between medium of different refractive indices may both reflect and refract at the interface. The periodic perturbation of the core refractive index along the fiber length gives to the FBG an ability to work as a wavelength selective mirror. Bragg's wavelength (λ_B) is when the reflectivity of the grating is maximum. λ_B is defined by the relationship

$$\lambda_B = 2n_{eff}\Lambda \tag{1}$$

where n_{eff} is the effective index of the grating in the fiber core and Λ is the grating period.
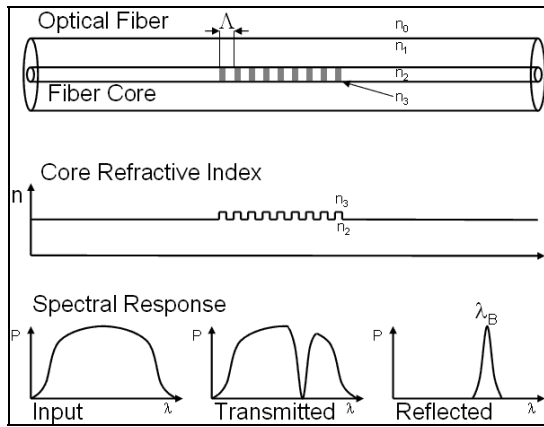

Fig. 2. A FBG structure, with refractive index profile and spectral response

When FBG is extended or compressed, the grating spacing is changed, thus it causes a shift of the Bragg wavelength. The wavelength shift due to the strain can be defined as

$$\Delta\lambda = (1 - p_e)\lambda_B\varepsilon \tag{2}$$

where ε is the applied strain, and p_e is the photoelastic coefficient defined as

$$p_e = \left(\frac{n_{eff}^2}{2}[p_{12} - v(p_{11} + p_{12})]\right) \tag{3}$$

where p_{1j} coefficients are the Pockel's coefficient of the strain-optic tensor, and v is Poisson's ratio [2][12].

IV.   METHODOLOGY

A.   Characterization of FBG sensor

FBG sensor must be tested to identify the performance of sensor. Result obtained was compared to the data sheet from supplier. Fig. 5 and Fig. 6 are schematic diagram of FBG sensor for transmission and reflectance light characteristics respectively. Central wavelength, bandwidth and reflectivity percentage of FBG sensor are among an important to be determined. The optical fiber sensing equipment belong to the Faculty of Applied Science, UiTM Shah Alam is used.

Fig. 5. Schematic diagram for FBG determine transmission light characteristic
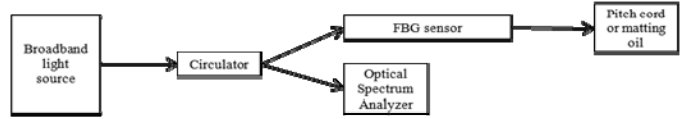
Fig. 6. Schematic diagram for FBG determine reflectance light characteristic

B.   Design and Fabricate the Standard Specimen and Box-Structure

A standard specimen according to ASTM Standards E606 will be prepared. A FBG sensor will be attached to the specimen and standard practice for strain-controlled fatigue testing will be done. This process will provide an initial study or information of the response of FBG sensor and to familiarize with the arrangement of FBG system before conducting the test/monitoring to more complicated structure. Designing the box-wing aircraft structure will be done and it is planned to have a structure with bending/transverse type of loading. This is similar to the load nature in the wing structure of aircraft. A certain dimension and loading are set according to the capacity of INSTRON Universal Testing Machine at the Faculty of Mechanical Engineering, UiTM Shah Alam. Finite element will be done to analyze the static response of box-wing structure. After that, the specimen is fabricated in general workshop at Faculty of Mechanical Engineering, UiTM Shah Alam.

C.   Testing with Variable Amplitude Loading

Optical fiber sensing configuration was attached on the specimen. The test is conducted by applying variable amplitude loading. Athough currently applied on standard specimen, later on the FBG system will be applied as fig. 7 on wing-box configuration. The mechanical test set up can be conducted in the Faculty of Mechanical Engineering, UiTM Shah Alam. Data logger of FBG sensor belongs to the Photronix Technologies (M) Sdn. Bhd. is used for capturing response from fiber optical sensor.
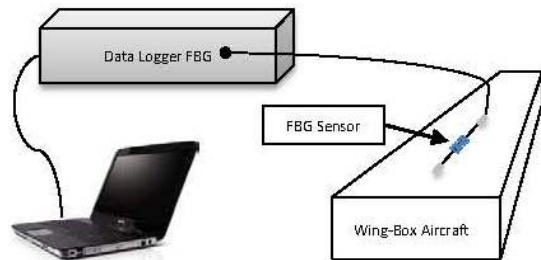
Fig. 7. Set up for FBG system on wing-box aircraft structure

D.   Optical Fiber Equipment

Below is the component of optical fiber system.
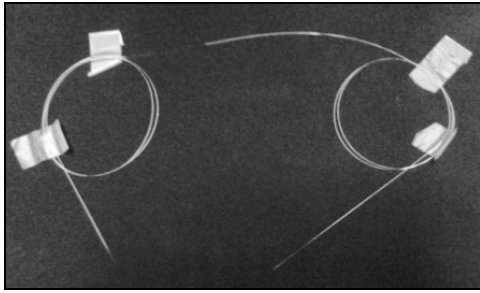
Fig. 8. Fiber Bragg grating sensor

Fig. 8 is one of the fiber Bragg grating sensors. There are three different reflectivity wavelength FBG sensors will be used, 1548nm, 1550nm, and 1552nm wavelengths.

Fig. 9. Laser Diode

Fig. 10. Laser pump (at top) and Optical Spectrum Analyzer (at bottom)

Laser pump (Fig. 10) will drive the high density laser to laser diode (Fig. 9). Then the laser diode will mount the light to through the optical fiber. Optical spectrum analyzer (OSA), Fig. 10, will be used for monitoring the center wavelength of FBG sensor.

V.   PRELIMINARY RESULT AND DISCUSSION ON FBG SENSOR RESPONSE

Variable loading has been applied to standard specimen (ASTM 606) to study the response from FBG sensor.
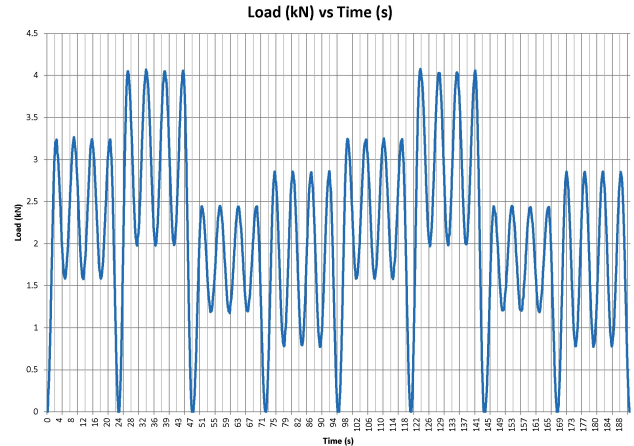
Graph. 1. Variable loading applied to standard specimen
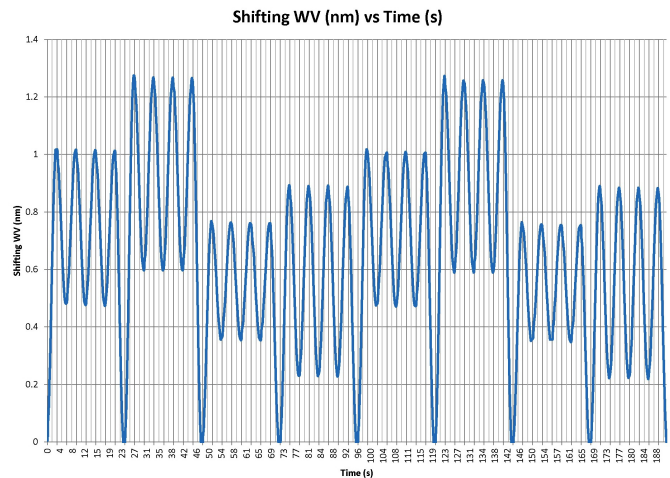
Graph. 2. Response from FBG sensor through data logger FBG

Both graphs above observe that the data response of FBG sensor has a same trending from data variable loading applied to the specimen. It means that response from FBG sensor is consistent with a variation loading applied. The relation between response of FBG sensor and strain will be determined.

VI.   CONCLUSSION

FBG is proposed for monitoring the fatigue load condition of the structure. Gaining a good knowledge of FBG sensor use in SHM to characterize the fatigue parameters is important to implement it in the actual structural health monitoring. A research of studying the response of optical FBG under the variable of amplitude load is reported. Test set-up for standard

specimen is presented in this paper. It is planned that, based on the response of optical fiber, the characterization of fatigue load parameters will be performed. With all the fatigue loading parameters are obtained, a calculation to predict a life time of structure can be done.

ACKNOWLEDGMENT

REFERENCES

[1] Daniel Balageas, "Structural Health Monitoring", ISTE Ltd, 1st South Asian Edition, 2007, pp. 13-43.

[2] Gerard F. Fernando, "Fiber Optic Sensor Systems for Monitoring Composite Structure", RP Asia 2005 conference, Bangkok, 25-26 August 2005

[3] M. Kamil Abd Rahman, M. Nasir Taib, Azmi Ibrahim, "Application of Optical Fiber Bragg Grating Sensor for Structure Monitoring", Materials Science Forum Vols. 517, pp. 202-206, June 2006

[4] W Kuntjoro, M S Ashari, M Y Ahmad, A M Mydin, "Development of Fatigue Life Monitoring of RMAF Fighter Airplanes", Book Chapter in "ICAF 2009 – Bridging the Gap between Theory and Operational Practice", Springer

[5] B. A. Tahir, J. Ali Saktioto, M. Fadhali, R. A. Rahman, and A. Ahmed, "A Study Of FBG Sensor And Electrical Strain Gauge for Strain Measurements", Journal Of Optoelectronics And Advanced Materials, Vol. 10, No. 10, October 2008, p. 2564 – 2568

[6] Alan D. Kersey, Michael A. Davis, Heather J. Patrick, Michel LeBlanc, K. P. Koo, C. G. Askins, M. A. Putnam, and E. Joseph Friebele, "Fiber Grating Sensors", Journal Of Lightwave Technology, Vol. 15, No. 8, August 1997.

[7] Hill, K.O. and Meltz, G., "Fiber Bragg Grating Technology Fundamentals and Overview", Journal of Lightwave Technology, Vol. 15, No. 8, August 1997.

[8] Richard G. Budynas and J. Keith Nisbett, "Fatigue Failure Resulting from Variable Loading", *Shigley's Mechanical Engineering Design Book,* McGraw Hill, 8th edition in SI Units, 2008, pg 257-339.

[9] Cyrus K. Hagigat, "Using Commercially Available Finite Element Software for Fatigue Analysis", Proceedings of the 2005 American Society for Engineering Education Annual Conference & Exposition, U.S.A., American Society for Engineering Education, 2005

[10] Meltz, G.; et al. "Formation of Bragg gratings in optical fibers by a transverse holographic method". Opt. Lett., 1989, 14: 823

[11] J. Ang, H.C.H. Li, I. Herszberg, M.K. Bannister, and A.P. Mouritz, "Tensile Fatigue Properties of Fibre Bragg Grating Optical Fibre Sensors", International Journal of Fatigue, March 2009, Australia.

[12] Li Sheng and Jiang De-sheng, "Structural Large Strain Monitoring Based on FBG Sensor", Symposium on Photonics and Optoelectronics, 2009.

Terrain Slope Analyses between Terrestrial Laser Scanner and Airborne Laser Scanning

Wan Aziz, W.A.[1], Syahmi, M.Z.[1], Anuar, A [1]

[1] Department of Geoinformatics, Faculty of Geoinformation Science & Engineering
Universiti Teknologi Malaysia
81310 UTM Johor Bahru, Johor, Malaysia
wanaziz@utm.my; anuarahmad@utm.my

Khairul Nizam T.[2]

[2] Department of Surveying Science & Geomatics, Faculty of Architecture, Planning & Surveying
Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia
nizamtahar@gmail.com

Abstract--Landslide is a geohazard phenomenon of mass movement of terrain. In order to prevent landslide, understanding the behavior of the landslide is of importance. In such a case, many landslide monitoring rely on accurate digital terrain models (DTM). Typically, DTM can be generated by point based measurements such as GPS, leveling, total station and laser distancemeters. However, these methods will raise obvious measurement problems because of the discontinuous character since they only allow measuring discrete points, and therefore it can be difficult to ensure that problems are recorded and handled in a consistent manner. In recent years, the 'Light Detection and Ranging – LiDAR' technology in the mode of Terrestrial Laser Scanning (TLS) and Airborne Laser Scanning (ALS) find rapidly growing interest in remote sensing and photogrammetry field as efficient tools for fast and reliable three-dimensional (3D) point cloud data acquisition. These laser techniques obtains the 3D reconstruction of objects; the measure of the distance from the platform (eg. ground-based for TLS, airplane or helicopter for ALS) derives generally from the time employed by the laser beam to go and to come back (time-of-flight principle). This paper describes the capabilities of TLS and ALS in slope mapping studies. The study has been conducted at two test sites, namely Tanah Rata and Habu, Cameron Highland, Malaysia. Several data acquisitions on selected cut slopes were performed using the TLS instrument: Topcon GLS1500. Ground control points were established using Real Time Kinematic GPS to provide a local coordinate system on laser scanning data. The ALS raw data was provided by a private mapping company. This study shows the ability of TLS and ALS to obtain reliable 3-D slope information over unstable area.

Keywords: Terrestrial Laser Scanner; Airborne Laser Scanner, Point Cloud, Digital Terrain Model, Slope Gradient.

I. INTRODUCTION

Landslide hazards annually cause a lot of fatal destruction due several factors such as uncontrolled infrastructure developments, continued deforestation and an increased regional precipitation caused by climate change. Its activities endanger inhabitants, jeopardize infrastructures (highways, railways, buildings etc.), and also can obstruct river channels, consequently dam breaking (by overtopping and erosion) would cause catastrophic floods. In the last few years there is a great increasing interest in reliable methods for landslide studies and it has led to intensive study efforts to automatically detect, reconstruct and modeling the landslide characteristics. There are many comparable surveying techniques to generate - with or without object contact and with a precision commensurate with scale - Digital Terrain Models (DTMs), a fundamental tool to detect, classify and monitoring landslides. For example, Global Positioning System (GPS) static and rapid-static modes, digital aerial and terrestrial photogrammetry, airborne/terrestrial laser scanning and satellite-based/ground-based radar interferometry had been recognized as trustworthy methods for the acquisition of accurate and reliable 3D-spatial data distributed in the studied area [23], [26], [2], [8], [39], [24], [21].

The possibility of acquiring datasets of the terrain surface with a high accuracy and high spatial resolution, either using laser, optical or radar sensors mounted on terrestrial, aerial or satellite equipment are currently opening up new ways to visualize, 3D-model and interpret mass displacement processes. Recently, there has been an increasing interest in exploiting the 'Light Detection And Ranging (LiDAR) of surveying methodologies: 'Terrestrial laser scanner (TLS)' and 'Airborne Laser Scanner (ALS)' [29], [17], [18], [1], [32], [30], [3]. TLS is a ground-based sensor; obtain their maximum resolution on surfaces perpendicularly oriented to the beam incidence angle, such as mountainous rock faces, hillside slopes, marine cliffs, etc. It is one of the most promising remote sensing techniques for slope gradient characterization and monitoring because of its capability to accurately acquire dense three-dimensional (3-D) coordinates of the terrain. The new technology of airborne LiDAR can provide high-precision DEM data, facilitate detection of minor changes in elevation and reveal subtle geomorphic features within an entire area. The airborne LiDAR technique includes the use of the GPS, Inertial or Navigation Measuring Units (IMU), and a laser distance scanner. Through combined use of laser transmitter with high repeating pulse frequency (RPF) and high speed scanning system, dense measurements of the terrain surface are obtained, giving high resolution of DTM. DTM is including any features and object that are available during scanning operation. One of the major factors in landslide deformation is a slope gradient of the DTM. Slope gradient played an important role in landslide risk analysis. The aim of this study

is to determine the effect of TLS and ALS data capture in slope mapping at landslide prone area.

II.   REVIEWS ON LIDAR METHOD

Light Detection And Ranging (LiDAR) is a modern surveying method conceived more than ten years ago that in a short time can generate good quality DTM as a result of elaborations with specific procedures that apply specific filtering and interpolation algorithms; only nowadays this technique is having growing development and applications thanks to information technology improvements. A telemeter laser obtains the 3D reconstruction of objects; the measure of the distance from the ground derives generally from the time employed by the laser beam to go and to come back (time-of-flight principle). Terrestrial laser scanning (TLS) is a ground-based technique for collecting high-density 3D geospatial data – see Fig. 1. Many TLS is designed for very large-scale, complex 3D scanning, such as entire buildings, factories, rooms, landscapes, and other big structures and locations, both inside and out.

Figure 1. Terrestrial Laser Scanner
(Source: © FARO Technologies Inc.)

TLS technology is also referred to as ground-based LiDAR or tripod LiDAR. It is an active imaging system whereby laser pulses are emitted by the scanner and observables include the range and intensity of pulse returns reflected by the surface or object being scanned. TLS measurements, combined with the orientation and position of the scanner, produce a 3-dimensional "point cloud" dataset. The primary capability of TLS is the generation of high resolution 3D maps and images of surfaces and objects over scales of meters to kilometers with centimeter to sub-centimeter precision. It is a powerful geodetic imaging tool ideal for supporting a wide spectrum of user applications in many different environments. The incorporation of GPS measurements provides accurate georeferencing of TLS data in an absolute reference frame. The addition of digital photography yields photorealistic 3D images. The main advantage of TLS and ALS over other geodetic techniques is therefore its property of direct, rapid and detailed capture of object geometry in 3D.

While airborne laser scanning - ALS (using airplanes or helicopters) constitutes one well established method landslide surveying [10]. The ALS measures the height of the ground surface and other features in large areas of landscape with a resolution and accuracy hitherto unavailable. ALS operates by using a pulsed laser beam which is scanned from side to side as the aircraft/helicopter flies over the survey area, measuring between 20,000 to 100,000 points per second to build an accurate, high resolution 3D-model of the ground and the features upon it. The ALS scanning device is mounted at the bottom or sides of an aircraft flying along selected parallel flight paths (Fig. 2). The instrument is bundled with an onboard differential GPS and an inertial navigation system to locate accurately the position of the aircraft. The information, together with the orientation of the scanning mirror for firing the laser, can compute the geodetic reference of the objects scanned. The digital elevation model can typically be constructed with an accuracy of ±0.15 m in vertical height. The ALS mapping can provide more comprehensive and precise topographic features than other conventional methods, such as aerial photographs interpretation and digital photogrammetry.
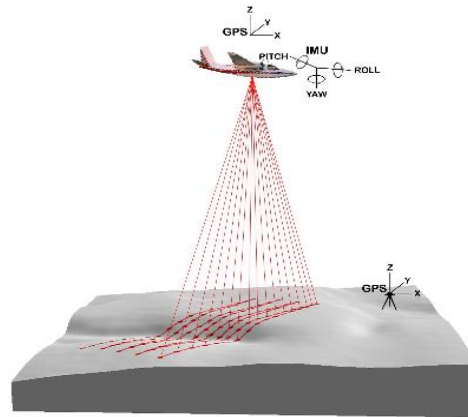
Figure 2. Airborne LiDAR (Source : USDA)

Other advantages are the dramatic reduction in costs and much faster project completion (within several days), possibility to survey remotely very complex, inaccessible and hazardous objects and areas, where the traditional techniques fail, the data is largely independent of ambient illumination (one may scan even at night), completeness and comprehensiveness of scanning: "everything" in the scene is captured at once. The accuracy of the scanner is to a great extent limited by the systematic instrumental errors, which means that it has to be properly calibrated

III.   STUDY AREA

The study was conducted at two sites of the landslide prone areas in Cameron highland, namely Tanah Rata and Habu. Fig. 3 shows the location study area and the condition of slope at the landslide study areas in Cameron Highlands. The reason of these slopes has been selected for this study due to the landslide incident occurs in July, 2011.
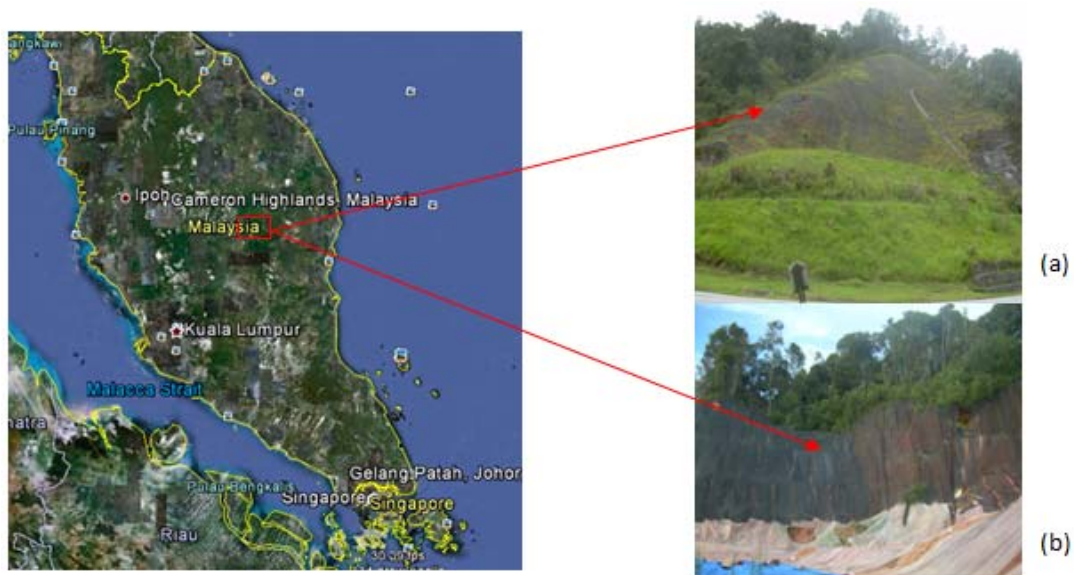
Figure 3. Study Area (a) Habu and (b) Tanah Rata, Cameron Highlands. (Source: Google earth map)

IV. THE METHODOLOGY

In this paper, we utilize the end-product of airborne LiDAR data and terrestrial laser scanning data. Both measurements were conducted at two slope areas, Habu and Tanah Rata, in which located at Cameron Highland. The selection of study area is based on the slope characteristic which is suitable to meet the objective of this study. Fig. 4 describes the methodology flowchart of this study. The main result in this study is digital terrain model at the slope areas. Later, both DTMs were used to generate slope map at the slope area. In order to avoid conflicts in coordinate system, all data obtained from TLS and LiDAR were converted into RSO coordinate system. In term of elevation, LiDAR measurement provides point clouds with MSL elevation based on the vertical control via bench mark around the area, meanwhile TLS measurement is based on RTK-GPS observation with geoid model correction provided by MyRTKnet infrastructure.

A. LiDAR Data

Taking into account, LiDAR data are commercial-processed obtained by a firm that provides LiDAR measurement service. As the end-user product, we have request LiDAR data with 50 cm point resolution and it is already filtered from vegetations and non-ground points. Therefore, the result of DTMs at both slope areas can be obtained.

B. TLS Operation

TLS operation consists of survey planning, fieldwork campaign and point clouds data processing. It is essential to outlines a survey planning before the fieldwork campaign is carried out to avoid repeat observation. In this study, site investigation, identification of registration and georeferencing approach, scanner position, number of scans, datum and desired resolution were concerned.

Figure 4. Methodology Flowchart

During the fieldwork campaign, the equipment used is TOPCON GLS-1500 3D Laser Scanner which based on time of flight (TOF) measurement. The advantage of this instrument is the integration of Phase shift algorithm that utilized in the TOF measurement to resolve the fluctuation in the pulse waveform

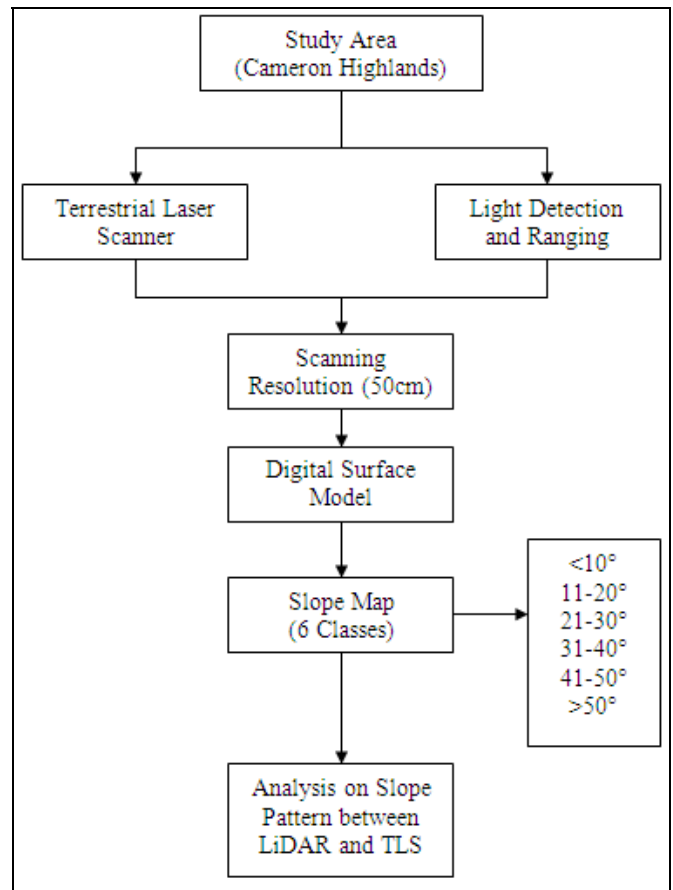thus provides clean and ultra-low-noise data. This type of instrument can measure up to 500m in range, emits laser beam 30,000 times per second, 360° horizontal and $\pm35^o$ vertical FOV and integrated with 2 megapixels digital camera. The accuracy of distance measurement is 4mm for every 150 m and 6 second for horizontal and vertical angle measurement.

In order to obtain high degree of coverage and avoid shadow zones, two scans were made from two standing points with 50 cm of point resolution. Data were collected approximately 30m and 50m from the slope area at Tanah Rata and Habu respectively. The time taken for each scans is approximately about 30 minutes, however it's depend on the FOV of observation. Larger areas need more time of observation as done at Habu (dimension 100 m height x 180 m width) compared to Tanah Rata (50 m height 100 m width). In this study, direct georeferencing approach was implemented where the scanner was set-up on known point and observation of backsight known point was done in order to obtain the orientation, almost similar as conventional total station measurement. The scanner also was accurately centered and leveled as well as its height was measured. The control points used for georeferencing were established before the TLS survey campaign by using RTK-GPS measurements with post processed (correction) mode derived from MyRTKnet infrastructure. The control points were converted in RSO coordinate system. The densities of points acquired were processed using ScanMaster software.

In the processing stage, the georeferenced point clouds from two scanner standing points were combined into one cohesive space. Since the direct georeferencing approach was implemented, there is no additional task needed in this combination task. The point clouds are then cleaned from any undesired laser returns such as vegetation, power lines, vehicle, and the area that is not related. As preparation for analysis purpose, DTMs were generated which is based on Triangulated Irregular Network (TIN).

V. RESULTS AND ANALYSIS

The analysis conducted in this study is mainly based on the slope gradient attribute derived from DTMs generated in which the data were acquired using LiDAR and TLS measurements. Since both measurements were based on different platforms, the differences between these results were compared. To realize it, DTMs from both LiDAR and TLS measurements were export into ArcGIS software to derive slope gradient attribute. The slope map were classified into 6 classes for example below than <10°, 11-20°,21-30°,31-40°,41-50° and more than >50° (See Fig. 5). As emphasis of this analysis, the comparison between slope map from TLS and LiDAR have been done to identify slope characteristic at Tanah Rata and Habu areas. Based on the slope results (Fig. 5), it was found that, scanning platform did affect the slope map products. Although both measurements utilize same resolution, the slope map generated was clearly different.
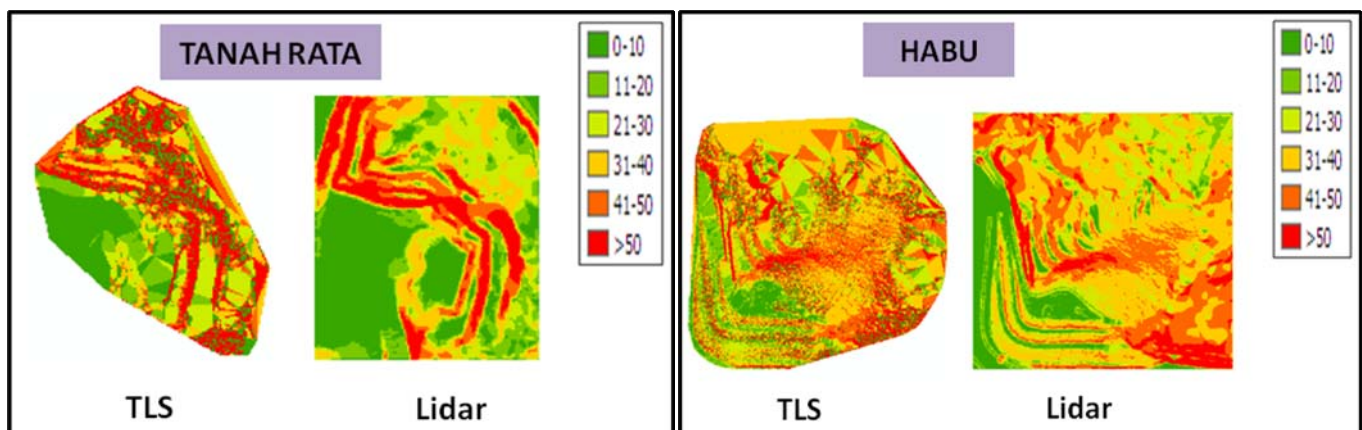
Figure 5. The slope map of Tanah Rata and Habu generated from TLS and LiDAR data

There are pro and con from the results obtained in this study. TLS has advantage in measures vertical surface since the normal of the surface is parallel with the laser direction. Therefore, the details of the steep slope can be precisely obtained. In contrast, LiDAR measures the slope surface from aerial platform, in which the normal of the surface is perpendicular with the laser direction. As the result, the point distribution at the vertical surface was lower than TLS point distribution. LiDAR measurement however has advantages in term of coverage area. Compared with TLS data in which consist of "shadow zone" occurred from the vegetation and complex slope geometry. In addition, LiDAR measurement permits the recording of full waveform or multiple returns so that the last return pulse can be obtained in order to indicate the true ground points [11]. As the result, LiDAR data permits higher interpolation of whole area compared with TLS data in which the triangles of the interpolation can be clearly seemed. Three dimensional models for both areas were demonstrated in Fig. 6. Fig. 6 shows the three dimensional models at Tanah Rata and Habu slope using Cyclon software. Based on this figure, in graphical view, a model generated from LiDAR and TLS data do not give a huge different. However TLS data give the better result because it is contain a detail of slope compared with LiDAR data (red circle).

Figure 6. DTM generated from TLS and LiDAR measurements, (a) Tanah Rata; (b) Habu

Figure 7. The percentage of the slope gradient to each different measurement, (a) Tanah Rata; (b) Habu

It is because TLS is ground based and all slope detail can be obtained while LiDAR data is airborne system which acquired data from certain altitude and aerial angle view. The differences between TLS and LiDAR data in producing slope map can be quantified as shown in Fig. 7. The results show the differences were relative with the area complexity. In this case Tanah Rata is more complex in geometry compared to Habu. Therefore at Tanah Rata, LiDAR data has lea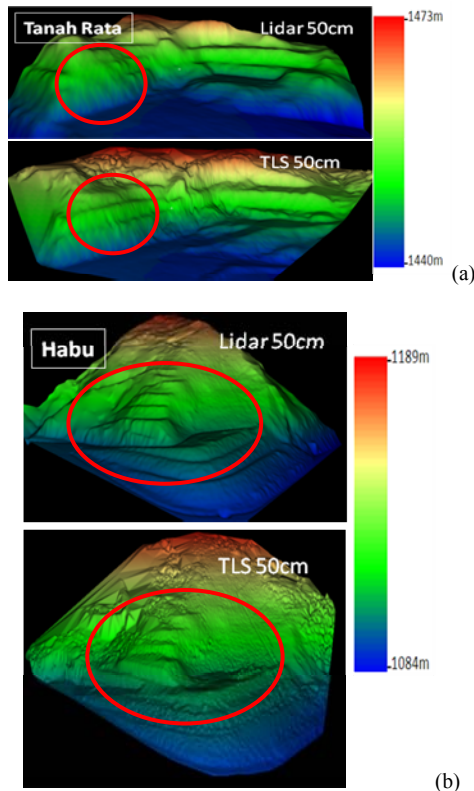ding in measures the horizontal surface. Meanwhile, for steeper slope (>50°), the percentage of TLS data shows much higher than LiDAR therefore proved that TLS has advantage in measures steep slopes. The same result also obtained at Habu area but the differences between both data are small due to the lower geometry complexity of the slope. From the results also shows that, for slope gradient 41° > 50°, TLS and LiDAR data percentage were quite similar for both areas.

VI. CONCLUSIONS
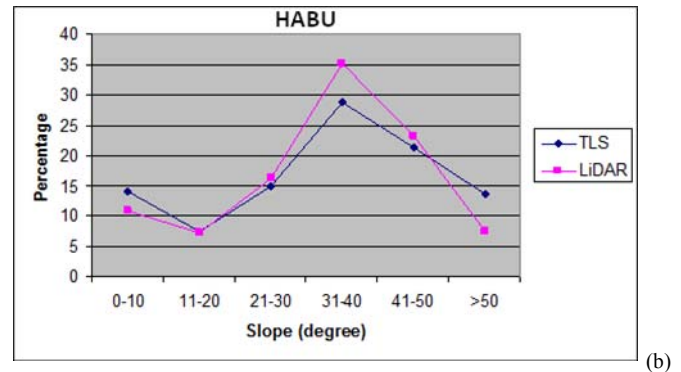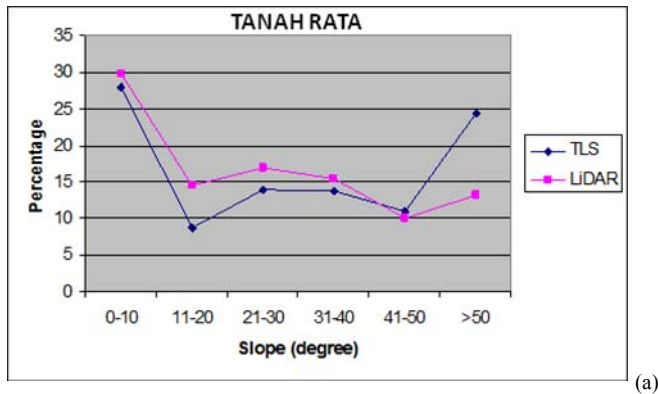
This paper has discussed the comparison results of slope gradient using TLS and ALS. In conclusion, it was found that scanning resolution effected slope map result. Therefore, the scanning resolution must be set according to the application needs. The TLS and ALS are the faster data collections in slope area and it provided reasonable accuracy data. The density of point cloud also depends on the resolution of point cloud. Scanning resolution might cause time delay in slope observation using the TLS and ALS methods.

ACKNOWLEDGEMENTS

REFERENCES

[1] A. Abellán, J. Calvet, J. Vilaplana, and J. Blanchard, "Detection and spatial prediction of rockfalls by means of terrestrial laser scanner monitoring," Geomorphology, Vol. 119, Issues 3–4, pp. 162–171, 2010.

[2] A. Pesci, P. Baldi, A. Bedin, G. Casula, N. Cenni, M. Fabris, F. Loddo, P. Mora and M. Bacchetti, "Digital elevation models for landslide evolution monitoring: application on two areas located in the Reno river valley (Italy)," Annals of Geophysics, Vol. 47, pp. 1339–1353, 2005.

[3] A. Pocsai, A. Zámolyi, B. Székely, G. Molnár, A. Roncat, and P. Drexel. "Change detection on the Doren landslide, using geological field measurements and laser-scanned data (Vorarlberg, Austria)," Geophysical Research Abstracts, Vol. 13, EGU2011-938,2011.

[4] A. Wehr, and U. Lohr, "Airborne laser scanning – an introduction and overview," ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 54, pp. 68 – 82, 1999.

[5] B. Collins, and N. Sitar, "Application of high resolution 3D laser scanning to slop stability studies," 39th Symposium on Engineering Geology and Geotechnical Engineering. Butte:Montana, 2004.

[6] B. Koska, T. Kremen, M. Stroner, J. Pospisíl, and M. Kaspar, "Development of Rotation Scanner, Testing of Laser Scanners," In Proceedings of INGEO 2004 and FIG Regional Central and Eastern European Conference on Engineering Surveying, Bratislava, Slovakia, November 2004. URL: http://www.fig.net/pub/bratislava/papers/ts_02/ts_02_koska_etal.pdf.

[7] C. Hapke, "Estimation of Regional Material Yield from Coastal Landslide Based on Historical Digital Terrain Modelling," Earth Surface Processes and Landforms, Vol. 3,. Issue 6, pp. 679-697, 2005.

[8] C. Squarzoni, C. Delacourt and P. Allemand, "Differential single-frequency GPS monitoring of the La Valette landslide (French Alps)," Engineering Geology, Vol. 79, pp. 215–229, 2005.

[9] C. Tucker, "Testing and Verification of the Accuracy of 3D Laser Scanning Data," In Proceedings of Symposium on Geospatial Theory, Processing and Applications. Ottawa, Canada, July 2002. URL:http://www.isprs.org/commission4/proceedings02/pdfpapers/537.pdf.

[10] D. Barber, J. Mills, and P. Bryan, "Laser Scanning and Photogrammetry: 21st century metrology," Proceedings of 18th International Symposium CIPA, Potsdam, Germany, pp. 360 – 366, September 2001.

[11] D. Harding, "Pulsed laser altimeter ranging techniques and implications for terrain mapping, Chap 5.In: Shan J, Toth CK (eds) Topographic laser ranging and scanning: principles and processing, " CRC Press, Taylor & Francis, pp 173–194, 2008.

[12] D. Lichti, M. Stewart, M. Tsakiri, and A. Snow, "Benchmark Tests on a Three- Dimensional Laser Scanning System," Geomatics Research Australasia, No. 72, pp. 1 – 24, 2000.

[27] D. Lichti, and B. Harvey, "An investigation into the effects of reflecting surface material properties on terrestrial laser scanner measurements," Geomatics Research Australasia, No. 76, pp. 1 – 21, 2002.

[28] D. Lichti, and S. Jamtsho, "Angular resolution of terrestrial laser scanners," Photogrammetric Record, Vol. 21, pp. 141–160, 2006.

[29] D. Tarchi, N. Casagli, R. Fanti, D. Leva, G. Luzi, A. Pasuto, M. Pieraccini and S. Silvano, "Landslide monitoring by using ground-based SAR interferometry: an example of application to the Tessina landslide in Italy," Engineering Geology, Vol. 68, pp. 15–30, 2003.

[30] D. Yue, J. Wang, J. Zhou, X. Chen, and H. Ren, "Monitoring slope deformation using a 3-D laser image scanning system: a case study," Mining Science and Technology (China), Vol. 20, Issue 6, pp. 898–903, November 2010.

[13] E. Huising, and L. Gomes Pereira, "Errors and accuracy estimates of laser data acquired by various laser scanning systems for topographic applications," ISPRS Journal of Photogrammetry & Remote Sensing, Vol. 53, Issue 5, pp 245–261, October 1998.

[14] G. Cheok, S. Leigh, and A. Rukhin, "Calibration Experiments of a Laser Scanner, "NISTIR 6922, National Institute of Standards and Technology, Gaithersburg, MD, September 2002. URL: http://www.bfrl.nist.gov/861/CMAG/publications/NISTIR calibration 6922.

[15] G. Cheok, (Ed.), "Proceedings of the LADAR Calibration Facility Workshop, NISTIR 7044, National Institute of Standards and Technology, Gaithersburg, MD, October 2003. URL: http://www.bfrl.nist.gov/861/CMAG/publications/index.html.

[16] G. Godin, J. Beraldin, M. Rioux, M. Levoy, L. Cournoyer, and F. Blais, "An Assessment of Laser Range Measurements of Marble Surfaces. In Optical 3-D Measurement Techniques V, A. Grün / H. Kahmen (Eds.)," pp. 49 – 56, 2001. URL: http://iit-iti.nrc-cnrc.gc.ca/iitpublications-iti/docs/NRC-44210.

[17] G. Teza, A. Galgaro, N. Zaltron, and R. Genevois, "Terrestrial laser scanner to detect landslide displacement fields: a new approach," International Journal of Remote Sensing , Vol. 28, Issue 16, pp 3425-3446, 2007.

[18] G. Teza, A. Pesci, R. Genevois, and A. Galgaro, "Characterization of landslide ground surface kinematics from terrestrial laserscanning and strain field computation," Geomorphology, Vol. 97, Issues 3–4, pp. 424–437, 2008.

[19] H. Bernhard, and P. Norbert, "Correction of laser scanning intensity data: Data and model-driven approaches," Journal of Photogrammetry & Remote Sensing, Vol. 62, Issue 6, pp 415-433, 2007.

[20] H. Ingensand, A. Ryf, and T. Schulz, "Performances and experiences in terrestrial laserscanning," In Optical 3-D Measurement Techniques VI, A. Grün / H. Kahmen (Eds.), 2003. URL: http://www.geometh.ethz.ch/.

[21] J. Cardenal, J. Delgado, E. Mata, A. González-Díez, J. Remondo, J. de Terán, E. Francés, L. Salas, J. Bonachea, I. Olague, A. Felicisimo, J. Chung Ch, A. Fabbri, and A. Soares, "The Use of Digital Photogrammetry Techniques in Landslide Instability," In Geodetic Deformation Monitoring: From Geophysical to Geodetic Roles, (Gil Cruz and Sanso, Eds.), IAG Springer Series, pp. 259-264, 2006.

[22] J. Clark, and S. Robson, "Accuracy of measurements made with a Cyrax 2500 laser scanner against surfaces of known color," Survey Review, Vol. 37, 294, pp. 626 – 638, 2004.

[23] J. Gili, J. Corominas, J. Rius, "Using Global Positioning System Techniques in Landslide Monitoring," Science Direct Engineering Geology, Vol. 55, pp. 167-192, 2000.

[24] J. Li, Z. Ouyang, X. Wei, and H. Zhou, "Earthquake Application of reflectorless total station measure technique in landslide deformation monitoring," The Chinese Journal of Geological Hazard and Control Series, issue 01, 2010.

[25] J. Mills, and D. Barber, "An Addendum to the Metric Survey Specifications for English Heritage – the collection and archiving of point cloud data obtained by terrestrial laser scanning or other methods," 2003. URL: http://www.heritage3d.org/.

[26] J. Moss, "Using the Global Positioning System to monitor dynamic ground deformation networks on potentially active landslides," International Journal of Applied Earth Observation and Geoinformation, pp. 24–32, 2000.

[31] Kayen, "Ground-LIDAR visualization of surface and structural deformations of the Niigata-ken Chuetse," Earthquake Spectra, Vol. 22, pp. 147-162, 2005.

[32] K. Johan, D. Christophe, A. Pascal, P. Pierre, J. Marion, and V. Eric, "Application of a Terrestrial Laser Scanner (TLS) to the Study of the Séchilienne Landslide (Isère, France)," Remote Sens. Issue 2, pp. 2785-2802, 2010.

[33] K. Kazu, M. Tomonori and T. Masataka, " Accurate Evaluation of Laser Scanner Data Depending on Location of GCPs for Monitoring Landslide," Proceedings of the 26th Asian Conference on Remote Sensing, 2007.

[34] K. Rowlands, L. Jones and M. Whitworth, "Landslide Laser Scanning: a new look at an old problem," Quarterly Journal of Engineering Geology and Hydrogeology, Vol. 36, pp. 155–157, 2003.

[35] L. Sui, X. Wang, D. Zhao and J. Qu, "Application of 3D laser scanner for monitoring of landslide hazards," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B1. Beijing, 2008.

[36] M. Balzani, A. Pellegrinelli, N. Perfetti, and F. Uccelli, "A terrestrial laser scanner: accuracy tests," In Proceedings of 18th International Symposium CIPA, Potsdam, Germany, pp. 445 – 453, September 2001.

[37] M. Johansson, "Explorations into the behaviour of three different high-resolution groundbased laser scanners in the build environment," In Proceedings of International Workshop on Scanning for Cultural Heritage Recording – Complementing or Replacing Photogrammetry, Corfu, Greece, September 2002. URL: http://www.isprs.org/commission5/workshop/part2.pdf.

[38] M. Tsakiri, D. Lichti and N. Pfeifer, "Terrestrial Laser Scanning For Deformation Monitoring," 3rd Iag / 12th Fig Symposium, Baden, 2006.

[39] O. Al-Bayari, and B. Saleh, "Geodetic monitoring of a landslide using conventional surveys and GPS," Vol. 39, No. 306, pp. 352-356, October 2007.

[40] O. Monserrat, and M. Crosetto, "Deformation measurement using terrestrial laser scanning data and least squares 3D surface matching," ISPRS Journal of Photogrammetry & Remote Sensing, Vol. 63, pp. 142–154, 2008.

[41] P. Baldi, S. Bonvalot, P. Briole, and M. Marshella, "Digital Photogrmmetry and Kinemtic GPS Applied for the Monitoring of Volcano Island, Aerolin Arc, Italy," Geophysical Journal International, Vol. 142, Issue 3, pp. 801-811, 2000.

[42] P. Mora, P. Baldi, G. Casula, M. Fabris, M. Ghirotti, E. Mazzioni and A. Pesci, "Global Positioning Systems and digital photogrammetry for the monitoring of mass movements: application to the Ca' di Malta Landslide (northern Apennines, Italy)," Engineering Geology, Vol. 68, pp. 103–121, 2003.

[43] T. Kremen, "Testing of terrestrial laser scanners," In Optical 3-D Measurement Techniques VII, A. Grün / H. Kahmen (Eds.), Vol. II, pp. 329 – 334, 2005.

[44] T. Miyazaki, K. Kinoshita, and M. Takagi, "Accurate Geometric Transformation of Laser Scanner Data for Landslide Monitoring," Infrastructure Systems Engineering, Kochi University of Technology Tosayamada-city, Kami-shi, Kochi, pp. 782-8502, 2008.

[45] S. Gordon, D. Lichti, M. Stewart, and M. Tsakiri, "Metric performance of a highresolution laser scanner," SPIE Proceedings, Vol. 4309, pp. 174 – 184, 2000.

[46] S. Gordon, D. Lichti, and M. Stewart, "Application of a high-resolution, ground-based laser scanner for deformation measurements," In Proceedings of the 10th FIG International Symposium on Deformation Measurements, Orange, California, USA, pp. 23 – 32, March 2001.

Development of Geographical Based Library Information System (GeoLIS)

Muhammad Jamaluddin Bazlan & Abdul Rauf Abdul Rasam

Centre of Studies Surveying Science and Geomatics
Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA (UiTM), Malaysia
tag2707@yahoo.com, rauf@salam.uitm.edu.my

Abstract— In Malaysia, the implementation of GIS in the LIS is in early stages and still no extensive studies are done in this field yet. The current problem of the UiTM Library System (OPAC) is to find the exact location of library materials particularly books. GIS is an application with the power of tools which enable a user to manipulate, store, analyze the data or information. The aim of this study was to design and develop the Geographic-Based Library Information System (GeoLIS) using GIS technology and other information system tools. The objectives of this study were, i) to carry out user requirements analysis for development of GeoLIS , ii) to design the GeoLIS, iii) to develop the GeoLIS and to test the performance of the GeoLIS. The methods used in this study were adapted from standard System Development Lifecycle (SDLC). Network Analyst in ArcGIS, Microsoft Visual Basic 6.0, Map Object, and Microsoft Access were software used in developing the GeoLIS. The user requirements analysis was distributed to the selected users using interviews, observation and questionnaire. The result of user requirements showed that the current system used by UiTM (OPAC) had several limitations, particularly it could not facilitate the student to find exact location of the book and need to be improved. Therefore, the proposed system (GeoLIS) allowed the user to perform database query or book searching, provide user with map direction of the book and detail information about the book such as image, description and review. The GeoLIS could also be used by the PTAR's organization and users as a searching reference for any future library management and development.

Keywords- GeoLIS, Geographical information system (GIS), UiTM Library System (OPAC), library information system

I. INTRODUCTION

A good library management information system in Malaysia is essential to be developed as the system can help librarians and users manage books systematically. Many libraries in the country use the Integrated Library Management Utility (ILMU) to utilize their respective books and facilities. The ILMU is a library information and knowledge management software developed by Universiti Teknologi MARA (UiTM) and Paradigm Systems Sdn Bhd to assist librarians and library staff for daily activities and operation. Although the ILMU represented one modern library systems toward true digital library, it does not provide a location or map direction of spatial book and library facility information. This study was to design and develop a Geographical Based Library Information System (GeoLIS) in Perpustakaan Tuanku Abdul Razak 1 (PTAR 1) UiTM using GIS technology and other spatial information technologies.

II. GIS FOR DEVELOPING LIBRARY INFORMATION MANGEMENT SYSTEM

A. Geographical Information System (GIS)

GIS can be described as a system for capturing, storing, checking, integrating, manipulating, analyzing and displaying data which are spatially referenced to the earth. This normally involves a spatially referenced computer database and appropriate applications software. A GIS envision the geographic aspects of a body of data. Basically, it lets users query or analyze a database and receive the results in the form of some kind of map. Since many kinds of data have important geographic aspects, a GIS can have many uses such as weather forecast, sales analysis, population forecast and land use and building planning to name a few.

B. UiTM Library: Perpustakaan Tuanku Abdul Razak (PTAR)

A library is a collection of sources, resources, and services; organized for use by a public body. Library in UiTM, PTAR was formed in 1960 under the Rural and Industrial Development Authority (RIDA) in Petaling Jaya. In 1972 this library was re-established in Shah Alam after a new library building was completed and named as Perpustakaan Tun Abdul Razak 1 (PTAR 1). It is not only a repository of research data and learning materials, but also central to the student experience, encouraging learning, dialogue, discussion and reflection. PTAR 1 is the main library for network, with the same name, at all UiTM, such as Perpustakaan Tun Abdul Razak 2 (1986), Perpustakaan Tun Abdul Razak 3 (2003), and Perpustakaan Tun Abdul Razak 4 (2004).

ILMU (Integrated Library Management Utility) is library information and knowledge management software developed by Universiti Teknologi MARA (UiTM) and Paradigm

Systems Sdn Bhd. ILMU was implemented at PTAR in January 1999 that represents one modern library systems toward a true digital library. It is a library information system designed to assist librarians and library staff for daily activities and operation. Online Public Access Catalog (OPAC) is a one system providing ILMU modules in info track. Info track serves as Info-center such as library map, library messages, library information, library calendar, looking for particular book (Figure 1).

Figure 1. Electronic Library Applying in PTAR UiTM
(*http://library.uitm.edu.my/*)

C. GIS Application in Library Information Management

The strength if GIS lies in its capability to carry operation of integrating both spatial (map) and non-spatial (attribute) data to devise solutions for technical and strategic decision-making. Other unique features of GIS include data conversion or exchange, data analysis, data simulation or 3D modeling, decision querying, database connectivity or sharing, retrieval of data, dynamic presentation, user friendly and multi user operations. A few examples of using GIS in various facility planning and management include infrastructure, town planning, and building. After seeing this type of GIS application, it is suitable to implement in library. Applications of GIS in library have 3 possibilities such as [4]:

- Use of GIS in mapping the library collections.
- Use of GIS in providing information service based on use of GIS and spatially related database.
- Use of GIS in mapping and networking library resources.

Book searching system in library is still in a new phase for GIS applications. Development of GIS application in library system in Malaysia is still in the preliminary stage. Library system in Malaysia was not so exposed to the GIS application. Most of the library system in Malaysia is developed using other applications like flash and text-based systems, and it was developed by particular bodies or organization to fulfill their public or private libraries. Overseas library system using GIS application was developed in advance either by GIS application or others compare with Malaysia. The development of library systems are not addressed by any related bodies or organizations other than systems that use GIS applications such as in transportation field.

III. METHODOLOGY

Development of the Geographic-Based Library Information System (GeoLIS) was conducted based on methodology flowchart as shown in Figure 2. After conducting the study objectives and reviews, data collection were collected by using PTAR 1 layout plan, library bookshelf plan, book details, and local plan coordinates. Then, the system was developed based on standard system development life cycle (SDLC)

The SDLC involved four stages i.e. feedback of user's requirement, design and development of the system and test the performance of the developed system as shown in result and analysis stage. The selected user requirements analysis conducted using the informal method includes observation of the existing system, survey questionnaire and semi structure interview. The selected respondents were chosen based on their professional skill, knowledge and opinions as consumers.

The individual perspectives information on the other hand supplies information regarding the users' perception and expectation through interview, observations and questionnaire. These groups provide useful information for the design of the proposed system user interface. The survey results are analyzed using Microsoft Office Excel 2007. ArcGIS, Microsoft Visual Basic (VB) and MapObjects Windows were used to process data and develop the system such as map scanning, registration, creating geodatabase, digitizing process, designing user interface, and programming language. The selected users and experts from PTAR 1 were appointed to evaluate the performance of the system.
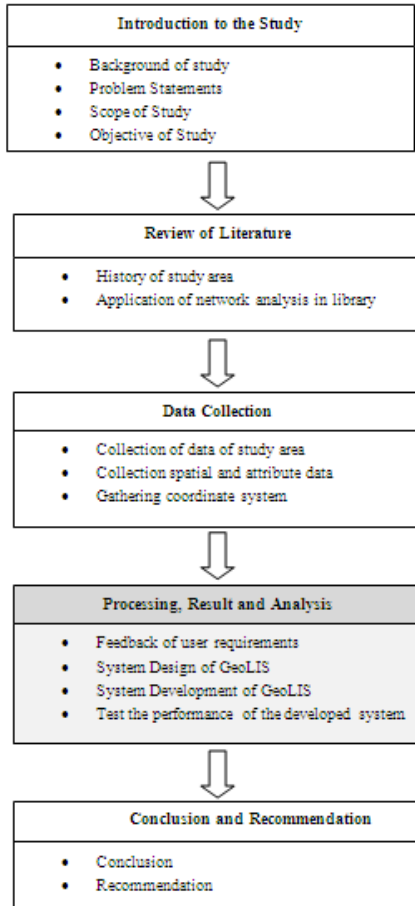
Figure 2: Flowchart of the Study Methodology

IV. RESULT AND DISCUSSION

This section focuses on the respondents of user requirement, the functions of developed system and the evaluation of the developed system.

A. Respondents of user requirement on the current and the proposed system

The respondents agreed that the current system did not facilitate the student with the map which can easily find the exact location of the books and it also can be a directory for users. Besides that, the system was not user-friendly especially for beginners (Figure 3). So, 80% of the respondents agreed that the current system should be improved with dynamic functions such as query and searching, mapping direction, and graphically visualization (Figure 4). They need a system in order to get information and all matters about the library management when it is needed. In addition, with the availability of the system, library management becomes more organized and efficiently when it is controlled by the system.
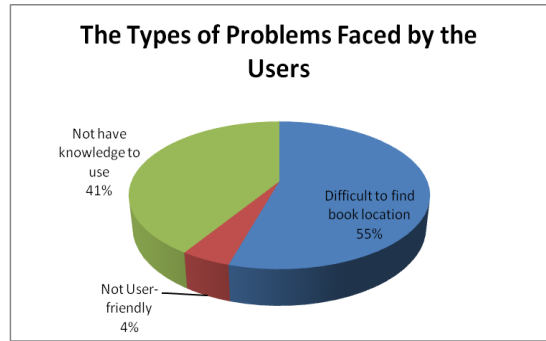
Figure 3: Types of Problem Faced by the Users Using the Existing System
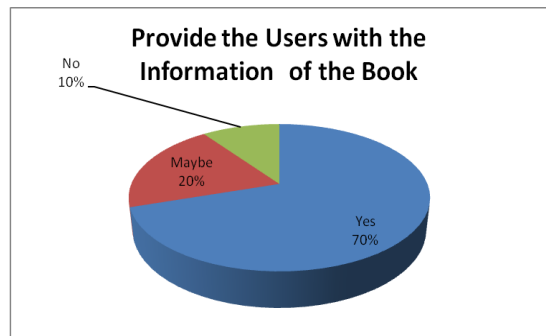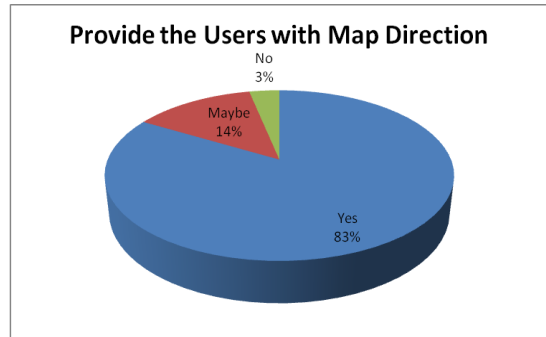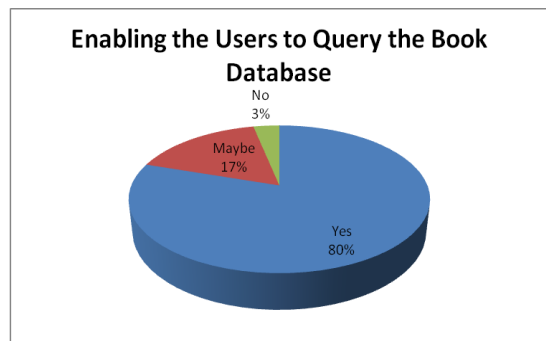
Figure 4. The Proposed Functions Suggested by the Users

B. System Design and Development

The system development of the GeoLIS involves 2 phases. The first phase was data preparation and second was design of Graphical User Interface for the GeoLIS. Three major steps involved in the first phase were data acquisition, extract layer and post processing. Designing graphic user interface for the proposed GeoLIS was the second phase in system development of the proposed system. This design was based on the user requirements analysis that had been carried out before. System interface is a vital asset in any system which must meet the user's requirement and make it easier to use. Visual Basic 6 is a medium in development system interface for the proposed GeoLIS. Generally, 2 processes in designing GUI using Visual Basic are for planning the project and creating the project [1].

Generally, GeoLIS starts with the users choosing a profile. Then, the floor directory window appears and user can select main function of GeoLIS which allows the users to perform query by book call no, title, author and price. Then, the query result will show and users can call map direction or book details. There are three functions that can be classified for this system. Firstly, it enables users to query the book database to get the information of the books. Secondly, it provides users with map direction to find the exact location of the book. Lastly, it provids users with details information of the book such as image, book description and book reviews. Figure 5 shows the main functions of GeoLIS.
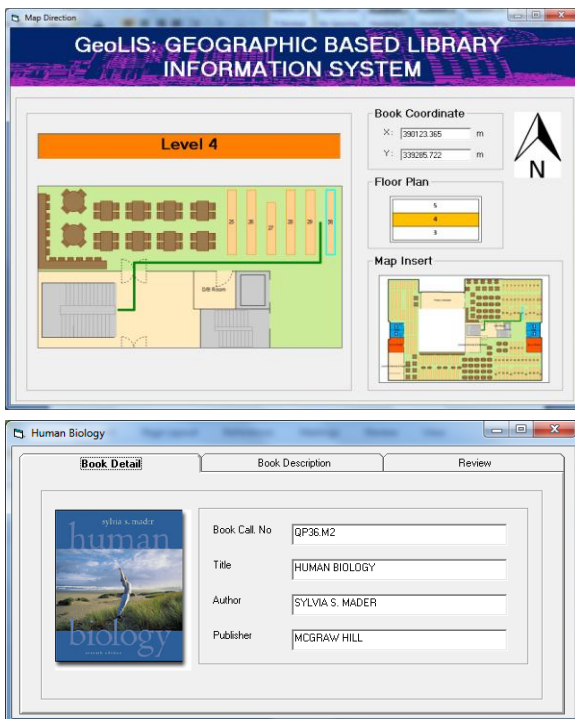
Figure 5. Main Functions of The GeoLIS; Introduction Window, Main Interface Window, Search and Query, Map Direction Window and Book Databases.

C. System Testing and Evaluation

The function of system testing and evaluation is to analyze the proposed system to fulfill the user's needs. The GeoLIS was tested and evaluated by two groups of users, regular user and professional. Regular users tested the two elements of GeoLIS; system interface and system functions.

The users were satisfied with the interface design and functions of the systems as the GeoLIS interface was easy to understand. and has good concept applications in the GeoLIS i.e map direction and book details. Professional users were selected to test and evaluate three components (System interface, system function and database). They were satisfied with the proposed functions of the system, but there are some suggestions that need to be considered, particularly the use of pop-up window and adding suitable programming for books database.

V. CONCLUSION

This study was to develop a Geographic-Based Library Information System (GeoLIS) with the integration of ArcGIS, Visual Basic application, Map Objects Windows, SDLC in PTAR 1, UiTM. Generally, library users spent most of their time to find the location of the book they want. But with GIS based information system, this problem can be solved with its implementation in library management. GIS not only can represent spatial data but also attributes data. GIS Application is able to store, manipulate and do analysis. Although there are some limitations in the GeoLIS, the concept framework of the system can be implemented in the real library system in Malaysia to facilitate the library user in finding the book location.

REFERENCES

[1] A. C. Bradley and Millspaugh A. C, *Programming in Visual Basic 6.0* (Update Edition). The McGraw-Hill Companies, Inc, 2004.

[2] A. M. Irfan, "Application of Geographic Information Systems in Transportation for Road Network Analysis" Retrieved on 1 March 2011 fromhttp://www.efka.utm.my/thesis/thesis/images/4MASTER/2005/3JGP/Transportation1/irfanalimemonma0141071d05ttp.pdf. 2005

[3] A. Peter, Burrough and A. M. Rachael, "Principles of Geographical Information Systems" Retrieved on 20 February 2011 from http://ukcatalogue.oup.com/product/9780198233657.do. 1998.

[4] D.N. Phadke, "Geographical Information Systems (GIS) in library and information services" Concept Publishing Company, New Delhi, 2006.

[5] E. Nancy, "Finding the Fastest Route" Retrieved on 15 February 2011 from http://proceedings.esri.com/library/userconf/proc03/p0320.pdf. 2003.

[6] H. S. Abdullah, *Research Method Handbook 2005/2006: Student's Edition*, UiTM, Shah Alam M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[7] K. Chang, Introduction to Geographic Information Systems (Fifth Edition) International Edition 2010. The McGraw-Hill Companies, Inc, 2010.

[8] M-S. Eirini "User Requirements Specification" Retrieved on 13 October 2011 from http://www.ict-papyrus.eu/files/Papyrus-D2.2-v02.1.pdf. 2008.

[9] UiTM, Universiti Teknologi MARA. Perpustakaan UiTM Malaysia. Retrieved 5 Mac 2012 from http://library.uitm.edu.my. 2012.

[10] The Linux Information Project, "Graphical User Interface Definition" Retrieved on 3 October 2011 from http://www.linfo.org/gui.html. 2004.

Height Discrepancies Based on Various Vertical Datum

Saiful Aman Hj Sulaiman[1], Kamaluddin Hj Talib[1], Mat Akhir Md Wazir[1], Othman Mohd Yusof[1] & Shafiq Azwan Zalil[2]

[1]Center of Study Surveying Science and Geomatics
Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA, MALAYSIA
Shah Alam, MALAYSIA
saifulaman@salam.uitm.edu.my

[2]Geodesy Section
Department Survey and Mapping Malaysia (DSMM)
shafiqazwan@jupem.gov.my

Abstract— **A vertical datum is the core information needed in deriving height or elevation of points on the surface of the earth. This information used as a starting point or zero value of height. Currently, various vertical datum used as a zero level in determination of height and depth in Peninsular Malaysia. Considering the current needed in seamless geospatial information across land and sea area, deep understanding on the discrepancies of vertical datum is necessary. In this paper, height of points referred to Peninsular Geodetic Vertical Datum (PMGVD), Local Mean Sea Level (LMSL) and Chart Datum (CD) are investigated. The discrepancies of height referred to the various datum also analysed.**

Keywords-Vertical Datum;PMGVD ;Local Mean Sea Level (LMSL);Chart Datum (CD)

I. INTRODUCTION

Height or depth on the Earth's surface is the vital component in three (3) dimensional coordinate system. Generally, any value of height or depth will always refers to the specific reference surface , known as a vertical datum.

Traditionally, vertical datum can be divided into 2 main categories: those based on the mean sea level (MSL) and lowest astronomical tide (LAT) [1]. Surveying and mapping activities on the land area usually used MSL as a vertical datum. Contrary, for the charting activities at the sea area, LAT is used. Both of this reference datum started at difference zero level. Therefore, height or depth of the same position have a different value if referred to different vertical reference datum.

Currently, surveying and mapping activities in Peninsular Malaysia always referred to Peninsular Geodetic Vertical Datum (PMGVD) via First Order Precise Levelling Network (FOPLN) as a vertical datum. While for charting activities, Local Chart Datum (CD) via Lowest Astronomical Tide (LAT) are used. However, there are also surveying, mapping and charting activities used Local Mean Sea Level (LMSL) as a vertical datum in preparing map, plans and chart.

At the coastal area, the information about the discrepancy of height or depth becomes extremely crucial due to rapid development. Therefore, investigation on the discrepancy of height when using different reference datum is necessary. The result of this investigation is expected to assist the authorities in planning and development of coastal areas more sustainable.

This paper aim to discuss on the vertical datum currently used in Malaysia for surveying, mapping and charting activities. Finally, the analysis on the discrepancy of height referred to various vertical datum in Malaysia presented numerically and graphically.

II. MALAYSIA VERTICAL REFERENCE DATUM

Surveying and mapping activities in producing maps, plan or charts were stated since 1900's. All the observation usually referred to specific reference surface either vertical or horizontal surface. Currently, in Malaysia, there a numbers vertical datum were used for land area and sea area in the determination of height or depth. However, all the reference datum were derived based on tidal observation.

A. Land Survey Datum 1912 (LSD12)

Land Survey datum 1912 also known as LSD12, is the first official national vertical datum [2]. This datum derived based on Mean Sea Level (MSL) of tidal observation data for eight (8) months between 1st September 1911 until 31st May 1912 by HMS Waterwich at Port Swettenham (Currently known as Port Kelang)[3]. The first precise levelling network known as First Order Levelling Network 1967 (FOLN67) referred to LSD12 as a vertical datum.. FOLN67 established to provide vertical control via benchmark for whole Peninsular Malaysia included Pulau Pinang, Pulau Langkawi and Pulau Tioman. All activities regarding to surveying and mapping on the land area referred to this vertical datum. Until 1994, new vertical datum established to replace LSD12. However, there are areas that still remain used LSD12 as a vertical datum such as Pulau Pinang, Pulau Langkawi and Pulau Tioman.

B. Peninsular Malaysia Geodetic Vertical Datum (PMGVD)

Peninsular Malaysia Geodetic Vertical Datum (PMGVD) established to replace old vertical datum (LSD12). This datum derived based on tidal observation data for ten (10) years between 1989 until 1993at the same point of LSD12 derived. Result from this epoch, new derived MSL, is 3.624m above zero tide gauge. It shows the new MSL is lower that the LSD around 6.5cm [2]. Furthermore, the new value of MSL adopted as a new zero level.

Based on the new zero level, transferring datum value from Port Klang to a memorial monument of PMGVD performed using combination precise levelling and gravity survey. Furthermore, height value of each vertical control around Peninsular Malaysia excluded Pulau Pinang, Pulau Langkawi and Pulau Tioman started from PMGVD monument.

Figure 1 Peninsular Malaysia Geodetic Vertical Datum Monument

C. Chart Datum (CD)

Modern charting activities in Malaysia stated since 1868 by The Royal British Naval Hydrographic Services. The main objective of that activities is to publish Admiralty nautical charts. Hydrographers around the world agreed that the chart datum is the extreme low water which the sea level seldom fall below it. Currently National Hydrographic Center (NHC) adopted Lowest Astronomical Tide (LAT) as a vertical reference datum. The level of this datum can predicted to occur under certain astronomical conditions. However, this datum is various depends on the location. To date, NHC established chart datum at 29 standard port around Peninsular Malaysia with a minimum period of tidal observation for 3 month s [4]

III. DATA USED AND METHOD

Data in this study collected from 12 tide gauges station operated by Department Survey and Mapping Malaysia (DSMM). The geographical distribution of tide gauges around Peninsular Malaysia are shown in Figure 2. The longest tidal observation exist at Johor Bahru and Port Klang Station, starting in 1984 and the shortest from Tanjung Sedili and Geting Satation, starting in 1987.
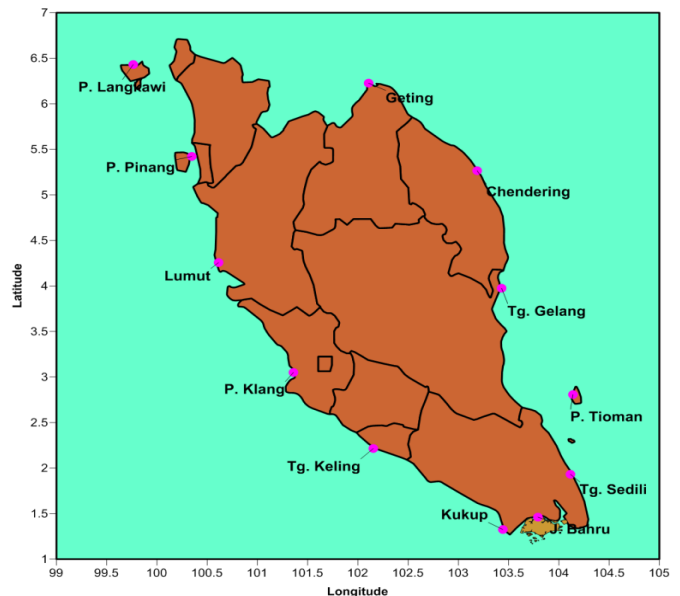
Figure 2 Geographical Distribution of Peninsular Malaysia Tide Gauges

In this study, tidal observation data since 1984 from all tide gauge stations is used. Before all this data can be used in computation, identification of missing data during the observation need to be done. This needed to be carried out to make sure no missing data more than three (3) months consecutively. If tidal observation data is missing more than three (3) months consecutively, yearly mean sea level will not be calculated for that years. The result, yearly mean sea level for Port Klang in year 2006 and 2009 , Tg. Keling in year 2005, Johor Bahru in year 1994 and 2009 and finally Pulau Tioman in year 2003 is not computed due to missing data. Finally, Local Mean Sea Level at the tide gauges station is determined based on averaging the yearly mean sea level (Table 1).

TABLE 1 LOCAL MEAN SEA LEVEL SINCE 1984 -2010

Location	LMSL(1984-2010)
P. Langkawi	2.220
P. Pinang	2.689
Lumut	2.202
P. Klang	3.641
Tg. Keling	2.855
Kukup	4.008
J. Bahru	2.858
Tg. Sedili	2.411
P. Tioman	2.838
Tg. Gelang	2.807
Chendering	2.214
Geting	2.299

*All height in meters

The height value of bench mark (BM) near to tide gauges station were obtained from Department Survey and Mapping

Malaysia (DSMM). According to DSMM the height of BM was carried out from PMGVD via precise levelling and gravity survey.

TABLE 2 BENCHMARK NUMBER AND HEIGHT NEAR TO TIDE GAUGE STATIONS

Location	BM No.	H_{BM_PMGVD}
P. Langkawi	K0172	3.417
P. Pinang	P0379	2.427
Lumut	A0401	3.517
P. Klang	B0169	3.87
Tg. Keling	M0311	3.668
Kukup	J5328	3.007
J. Bahru	J0416	3.422
Tg. Sedili	J0801	2.257
Tg. Gelang	C0331	3.835
Chendering	T0283	2.604
Geting	D0354	3.852

* All height in meters

A. BM height refers to Local Mean Sea Level (LMSL).

As mention above, existing height value of benchmark near to tide gauges station carried out from PMGVD via precise levelling and gravity survey. That means reference datum for that value is zero level at Port Klang.

To determine the height value of same benchmark, but refers to the LMSL datum, assumption is made that the LMSL obtained by tidal observation at that location is zero. The relationship between zero tide gauge and LMSL in determination of BM height refers to LMSL given in Figure 3.
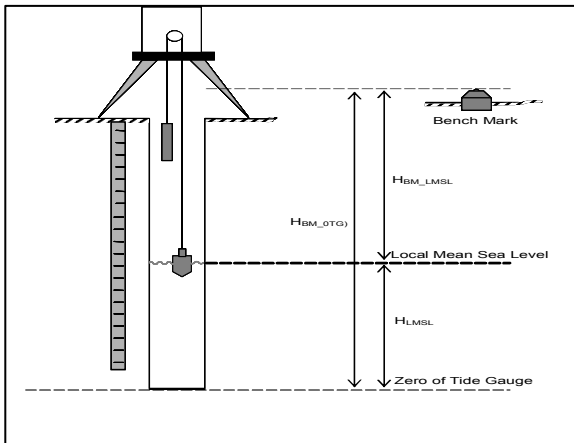
Figure 3 Relationship between zero tide gauge and local mean sea level in deriving Height referred to Local Mean Sea Level (LMSL)

In numerical computation, the equation (1) can be used in determination of BM height refers to LMSL.

$$H_{BM\_LMSL} = H_{BM\_0TG} - H_{LMSL} \tag{1}$$

where H_{BM_L} is the height of BM above LMSL, H_{BM} is the height of BM from zero tide gauge, and H_L is the height of mean sea at tide gauge from zero tide gauge.

B. BM height refers to Chart Datum (CD).

In determination of BM height refers to CD, same methodology in determination of GM heights refers to LMSL is applied but differ in terms of reference datum used. Figure 4 shows the relationship between zero tide gauge at the specific location and
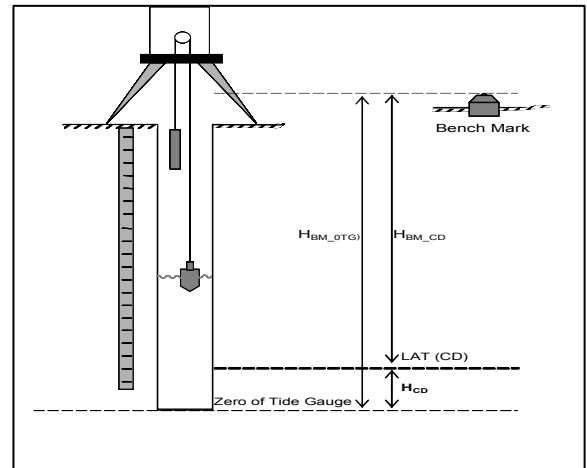
Figure 4 Relationship between zero tide gauge and lowest astronomical tide in deriving Height referred to Chart Datum (CD)

Equation (2) shows the numerical solution to calculate the height of BM refers to Chart Datum (CD).

$$H_{BM\_CD} = H_{BM\_0TG} - H_{CD} \tag{2}$$

where H_{BD} is the height of BM above CD, H_{BM} is the height of BM from zero tide gauge, and H is the height of predicted lowest astronomical tide from zero tide gauge.

C. Discrepancies of BM height referred to different vertical datum

The main objective of this study to identify the discrepancy of height value if refers to Peninsular Malaysia Geodetic Vertical Datum (PMGVD), Local Mean Sea Level (LMSL) and Chart Datum (CD). As mention above that all the datum commonly used in Peninsular Malaysia in determination of height of certain point.

Equation (3) and equation (4) shows the simple formula for calculating the discrepancy of height refers to specific reference surface.

$$\partial H_{PMGVD\_LMSL} = H_{BM\_PMGVD} - H_{LMSL} \tag{3}$$

where ∂H_{PMGVD_L} is the height discrepancy between PMGVD and LMSL, H_{BM_PM} is the height of BM transferred by precise levelling from PMGVD monument and H_L is the computed height of BM referred to LMSL.

$$\partial H_{PMGVD\_CD} = H_{BM\_PMGVD} - H_{CD} \qquad (4)$$

where ∂H_{PMGVD_L} is the height discrepancy between PMGVD and LMSL, H_{BM_PM} is the height of BM transferred by precise levelling from PMGVD monument and H_L is the computed height of BM referred to LMSL.

IV. RESULT AND ANALYSIS

The height value of BM for each selected BM has been calculated referred to the specific reference datum using equation (1) and (2). The computed height at selected BM refers to specific reference datum is given in Table 3 and 4

TABLE 3 COMPUTED HEIGHT AT SELECTED BM REFERS TO LMSL

Location	BM No.	$H_{BM\ 0TG}$	H_{LMSL}	$H_{BM\ LMSL}$
P. Langkawi	K0172	5.545	2.220	3.325
P. Pinang	P0379	4.962	2.689	2.273
Lumut	A0401	5.685	2.202	3.483
P. Klang	B0169	7.494	3.641	3.853
Tg. Keling	M0311	6.427	2.855	3.572
Kukup	J5328	6.88	4.008	2.872
J. Bahru	J0416	6.079	2.858	3.221
Tg. Sedili	J0801	4.459	2.411	2.048
Tg. Gelang	C0331	6.496	2.807	3.689
Chendering	T0283	4.688	2.214	2.474
Geting	D0354	5.964	2.299	3.665

All height in meters

TABLE 4 COMPUTED HEIGHT AT SELECTED BM REFERS TO CD

Location	BM No.	$H_{BM\ 0TG}$	H_{CD}	$H_{BM\ CD}$
P. Langkawi	K0172	5.545	0.400	5.145
P. Pinang	P0379	4.962	0.980	3.982
Lumut	A0401	5.685	0.350	5.335
P. Klang	B0169	7.494	0.610	6.884
Tg. Keling	M0311	6.427	1.660	4.767
Kukup	J5328	6.88	2.300	4.580
J. Bahru	J0416	6.079	0.650	5.429
Tg. Sedili	J0801	4.459	0.710	3.749
Tg. Gelang	C0331	6.496	0.880	5.616
Chendering	T0283	4.688	0.760	3.928
Geting	D0354	5.964	1.560	4.404

All height in meters

In computation the height of BM refers to LMSL or CD, things that should be given attention is where is the zero level. For height of BM refers to LMSL, the zero level is the MSL derived based on tidal observation data at each tide gauge station since the station started observed until 2010. While, the height of BM refers to CD, zero level of CD derived from predicted lowest astronomical tide (LAT). From both zero level, the height of BM is computed.

Table 5, shown the numerical discrepancies of height referred to PMGVD, LMSL and CD. It should keep in mind the height of BM refers to PMGVD is carried out from zero level of tide gauge at Port Klang and that value derived from 10 years tidal observations data.

TABLE 5 HEIGHT DISCREPANCIES BETWEEN PMGVD, LMSL AND CD

Location	BM No.	Discrepancies (meter)	
		$\partial H_{PMGVD\ LMSL}$	$\partial H_{PMGVD\ CD}$
P. Langkawi	K0172	0.092	-1.728
P. Pinang	P0379	0.154	-1.555
Lumut	A0401	0.034	-1.818
P. Klang	B0169	0.017	-3.014
Tg. Keling	M0311	0.096	-1.099
Kukup	J5328	0.135	-1.573
J. Bahru	J0416	0.201	-2.007
Tg. Sedili	J0801	0.209	-1.492
Tg. Gelang	C0331	0.146	-1.781
Chendering	T0283	0.130	-1.324
Geting	D0354	0.187	-0.552

All height in meters
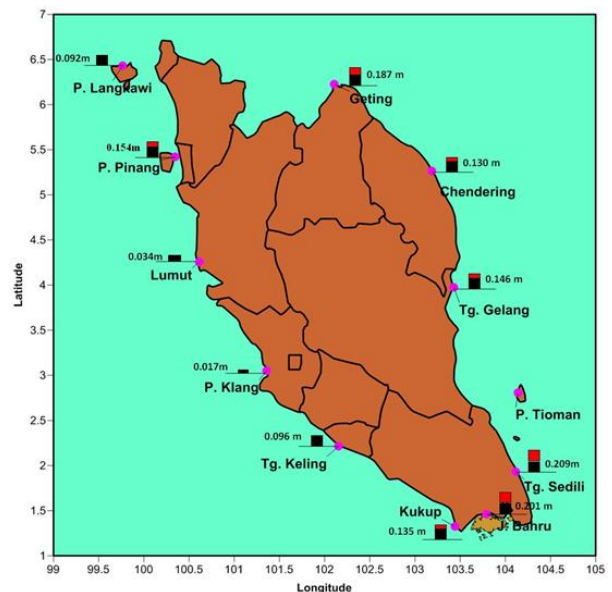
Figure 5 Geographical Distribution of Height Discrepancies Between BM Refers to PMGVD and LMSL.

The value of height discrepancies is computed based equation (3) and (4), in which data from Table 3 and 4 as an input value. Figure 5 and 6 illustrated the geographical distribution of Peninsular Tide Gauges with numerical

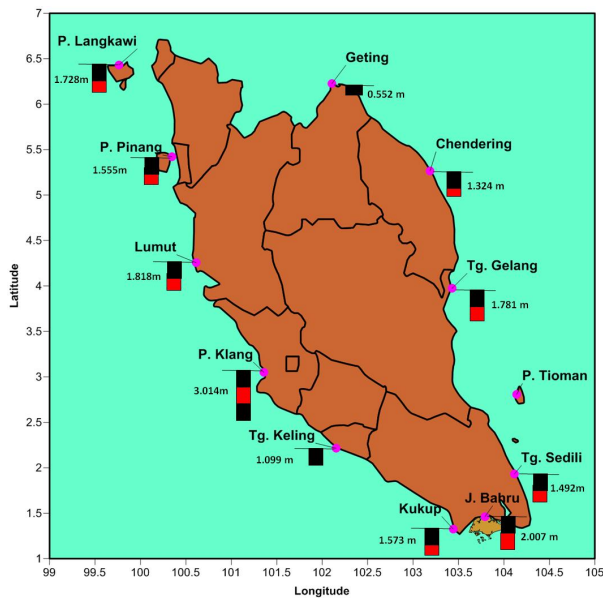discrepancy between PMGVD versus LMSL and PMGVD versus CD.

Figure 6 Geographical Distribution of Height Discrepancies Between BM Refers to PMGVD and CD.

Based Table 5 and Figure 5, the range of height discrepancies between PMGVD and LMSL at the west coast of Peninsular Malaysia is 1.7 to 13.5cm. While, at the east coast area, the range of height discrepancies is from 18.0 to20.9cm. The maximum height discrepancy is at Tg. Sedili and the minimum discrepancy is at Port Klang. Overall height discrepancies between PMGVD and LMSL is not too large.

In Table 5 and Figure 6, the maximum height discrepancies between PMGVD and CD is located at Port Klang around 3.014m. While, the minimum height discrepancies are at Geting around 0.552m. Overall result in height discrepancies between PMGVD and CD is too large.

V. CONCLUSION

The main objective in this paper is to investigate the height discrepancies occurs when using a different reference datum in surveying, mapping and charting activities. Currently in Malaysia there are three (3) common datum widely used: Peninsular Geodetic Vertical Datum (BM), Local Mean Sea Level (LMSL) and Chart Datum (CD). All these three (3) datum shows different zero levels.

Based on the result of this investigation found that the discrepancies between height referred to PMGVD and LMSL is relatively small and will not effect on the accuracy of the work especially for construction and small scale mapping.

Contrast with the results shows in discrepancies between height referred to PMGVD and CD. In this case, the discrepancies are too large and strictly not acceptable in surveying, mapping and charting activities.

As a conclusion, the knowledge of discrepancies of height when refers to various vertical reference datum is very important especially when dealing with coastal area. This is needs to ensure that geospatial information produced is accurate and reliable.

REFERENCES

[1] B. Parker, D. Milbert, and K. Hess, "Integrating bathymetry, topography, and shoreline, and the importance of vertical datums," *OCEANS 2003.*, pp. 758-764, 2003.

[2] A. Mohamed, "An Investigation Of The Vertical Control Network Of Peninsular Malaysia Using A Combination Of Levelling, Gravity, GPS and tidal Data.," Universiti Teknologi Malaysia (UTM), 2003.

[3] H. Jamil, "GNSS Heighting and Its Potential Use in Malaysia," in *GNSS Processing And Analysis*, 2011, no. 5410.

[4] Z. Hasan, *Tide Table Malaysia 2012*. National Hydrographic Centre, Royal Malaysian Navy, 2012.

[5] F. Baart, M. van Koningsveld, and M. J. F. Stive, "Trends in Sea-Level Trend Analysis," *Journal of Coastal Research*, vol. 280, no. 2, pp. 311-315, Mar. 2012.

[6] R. M. Ramli, "Tide and sea Level Variation of The Peninsular Malaysia Region," University College of North Wales, Bangor, 1990.

[7] K. M. Omar, S. Ses, and A. Mohamed, "Development of Rapid Monitoring System for The National Geodetic Vertical Datum (HGVD)," Johor Bahru, Johor, 2005.

[8] R. V. Donner, R. Ehrcke, S. M. Barbosa, J. Wagner, J. F. Donges, and J. Kurths, "Spatial patterns of linear and nonparametric long-term trends in Baltic sea-level variability," *Nonlinear Processes in Geophysics*, vol. 19, no. 1, pp. 95-111, Feb. 2012.

[9] A. Unnikrishnan, "Are sea-level-rise trends along the coasts of the north Indian Ocean consistent with global estimates?," *Global and Planetary Change*, vol. 57, 2007.

Contribution of GIS and Remote Sensing Technologies for Managing Foodborne Diseases in Malaysia

Abdul Rauf Abdul Rasam & Abdul Malek Mohd Noor

Centre of Studies Surveying Science and Geomatics
Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA (UiTM), Malaysia
rauf@salam.uitm.edu.my, raufrasam@yahoo.com

Abstract— Foodborne diseases (e.g. cholera, dysentery, typhoid, and Hepatitis A) are still considered a threat to human health and economy of a country. The cooperation among governments, the food industries and consumers are essential in order to control the diseases effectively. These emerging diseases can be controlled by implementing an effective technology application and innovation. Geographical information system (GIS) and satellite remote sensing are not new technologies in the world, but their applications in Malaysia particularly in the field of health are not used widely although the technologies have potential in providing a holistic system for public health care and epidemiology. This study was to demonstrate practically the implementations of GIS and remote sensing technologies in managing cholera control plan as a foodborne disease in the country. Initially, it showed that the technologies could be applied in three operations; i) interactive mapping of cholera disease, ii) analyzing dynamic changes for cholera transmission risk system, and iii) a framework for public healthcare management system. These potential contributions of technologies in health hopefully could be beneficial to the epidemiologist and public health staffs of MOH institutions toward creating an efficient MyGeoHealth (Malaysian Geospatial and Health) approach.

Keywords- Foodborne Disease, Geographical information system (GIS), Satellite Remote Sensing, MyGeohealth

I. INTRODUCTION

World Health Organisation (WHO) defines foodborne diseases the result of ingestion of foodstuff contaminated with microorganisms or chemicals from environmental contamination, including pollution of water, soil or air. It includes cholera, dysentery, typhoid, Hepatitis A, food poisoning, salmonellosis, Escherichia coli and so on. Although the situation of the diseases in Malaysia is under control, certain high risk areas are still unpredictable occur [1,2] and sporadically like cholera epidemics and food poisoning in schools are easily detected, diffuse outbreaks often go unreported [7,8] Geographical information system (GIS) and satellite remote sensing are geospatial technology; to collect, process, analyze data spatially and temporally especially in spatio-temporal infectious disease monitoring and management. The techniques have also better analysis combined with epidemiological and statistical approaches for effective foodborne disease control plan and management.

II. GEOSPATIAL TECHNOLOGIES IN PUBLIC HEALTH APPLICATIONS OF MALAYISIA

A. Geographical Information System (GIS) and Remote Sensing Technology

Generally, a geospatial technology is the information technology field of practice that acquires, manages, interprets, integrates, displays, analyzes and uses the geographic, temporal, and spatial data such as GIS, satellite remote sensing, GPS and so on. GIS is a computer system for capturing, storing, querying, analyzing, and displaying geospatial data [6], while remote sensing is a technology involving the use of sensors placed on a platform moving at a far distance from earth's surface and it can be used to collect data of the earth for the purpose of inventorying and monitoring [11]. These geo-innovative techniques have advantages in terms of spatial data collection, geo-mapping, geo-visualization, geo-analysis, geo-modelling, and the management of geospatial health features and phenomena [2] particularly for effective foodborne disease control plan.

B. Foodborne Outbreaks in Malaysia

The outbreaks of foodborne disease are defined as the occurrence of more than two cases of a similar illness. According to WHO (2012), foodborne diseases have high levels of mortality in the general risk-groups of population. In Malaysia, there are still reported incidents of food and waterborne diseases such as cholera, dysentery, typhoid, Hepatitis A and food poisoning. Previous local studies have shown that the main contributing factor of the disease was identified as insanitary food handling procedures [8] especially in schools' and academic institutions' food preparation premises and inappropriate food handling practices [7].

There are many intervention methods introduced by the government to control the disease systematically; including surveillance and monitoring system, training for preventative control of food safety management systems. The Foodborne Diseases Network (MyFoodNet) and Food Safety Information System of Malaysia (FoSIM) were established to monitor the surveillance system of foodborne diseases in the country. Training Programme in 1996 was also conducted to ensure hygienic practices during handling, preparation and sale of food. The Epidemic Intelligence Programme (EIP) was also initiated in 2002 to train public health practitioners. The implementation of Hazard Analysis Critical Control Point (HACCP) in food regulation as a requirement has been considered to have a positive influence on food safety and governments have mandated the use of HACCP system [7]

Although the government has supplied clean water, appropriate training and facilities to control the foodborne disease, effective ways to prevent the disease by using a multisectoral approach among Ministry of Health (MOH) and other government or private agencies and individual need to be conducted in the management of the systematic foodborne diseases in Malaysia [4,7].

C. GIS and Remote Sensing Applications in Foodborne Disease Management

GIS is not a new tool in the world, but its application in Malaysia particularly in the field of health is not used widely although this technology has potential in providing a holistic system for public health care and epidemiology [1,2,12,13,14]. Abdul Rasam *et al* (2011) introduced GIS-based approach for managing foodbone disease (cholera) in Malaysia (Figure1). The general functions of GIS in health studies are disease mapping and modeling, spatio-temporal changes analysis and risk assessment, public health care and hospital management. This study explores spatial analysis tool set in ArcGIS software for mapping and analyzing cholera diseases combined with remote sensing techniques for environmental or ecological risk assessment using spatio-temporal changes analysis and statistical correlation.

An increasing number of health studies have used remotely sensed data for monitoring, surveillance, or risk mapping, particularly vector-borne diseases [5] and non-borne diseases. Since no single spatial, temporal, or spectral resolution is universally appropriate for understanding the transmission risk for any disease, the National Aeronautics and Space Administration's (NASA) Center for Health Applications of Aerospace Related Technologies (CHAART) evaluated current assessed and planned satellite sensor systems in enabling human health to determine data relevant for the epidemiologic, entomologic, and ecologic aspects of their research.

Figure 1. Malaysian Geographical/Geospatial Health (*My*GeoHealth): Innovative method for monitoring environmental risk factors on human health (e.g. GIS-Based Cholera Information System in Sabah. [1]

III. GIS AND REMOTE SENSING APPLICATIONS IN MANAGING FOODBORNE DISEASE OF MALAYSIA

Based on the study conducted in Sabah on Cholera disease management using geospatial technology, the majority local selected respondents (80%) strongly agreed that these technologies have potentials to be applied in monitoring and managing the disease outbreaks. Practically, the study was carried out with three main GIS-based operations in controlling cholera disease; i) interactive mapping of foodborne disease, ii) analyzing dynamic changes for foodborne transmission risk factor, and iii) framework of a foodborne disease management system:

A. Interactive mapping of foodborne disease

The existing static or pin map of cholera distribution of Jabatan Kesihatan Negeri (JKN) Sabah could be improved using dynamic cholera disease mapping (Figure 2). Cholera map can help JKN staff in decision-making process showing the geographical distribution of the diseases in communities of the district, help elucidate the cause of the disease, and may be used to generate hypotheses of disease causation [15].
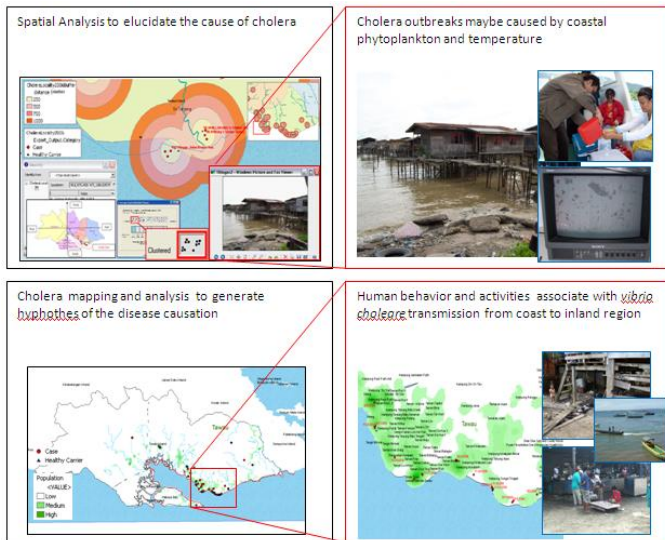
Figure 2: Dynamic Functions of Cholera Disease Maps in JKN, Tawau Using ArcGIS Software and Statistical Analysis

B. Dynamic analysis for foodborne transmission risk factors

Cholera disease mapping in Tawau can also display descriptive spread of the spatial disease, and generating hypotheses about disease causation, but they cannot establish the precise cause of multi-factorial disease (triggered) that need specific information about disease etiology, and local knowledge for interpretation of demographic and temporal aspects [15]. Therefore, for ecological or epidemiological risk assessment of SST and Chl-*a* on cholera distribution, data of satellite spatio-temporal changes and GIS and statistical correlation or interpretation are utilized together (Figure 3).
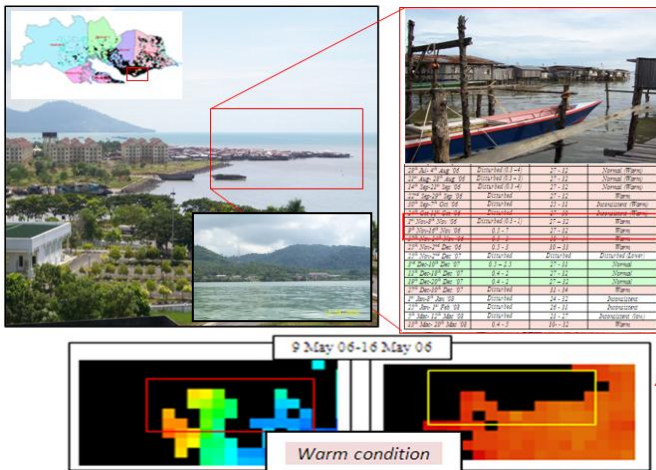
Figure 3. Analysis of Cholera High Risk Area Condition during the Warm SST Condition from May to September

C. Framework of a foodborne disease management system

Figure 4 demonstrates the proposed framework of Cholera Information System (CHIFS) in JKN, Tawau. The triggered environmental variables and other influential environments could be identified and integrated in related GIS software and other animation system. Layers of cholera incidents, SST and Chl-*a* are combined with other relevant spatial database (e.g. healthcare facilities) using GIS, statistical model, and other ICT (e.g. multimedia and virtual) tools for effective and interesting predictive model and early warning surveillance or system.
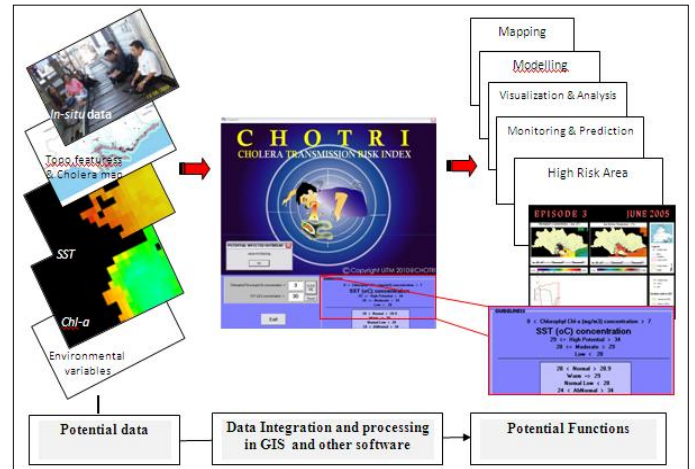
Figure 4: Proposed framework of a foodborne disease management system (e.g. Cholera)

IV. CONCLUSION

This study was to demonstrate the potential implementations of GIS and remote sensing technologies in managing foodborne disease in Malaysia especially cholera disease. There are three main GIS-based operations used in managing foodborne disease; i) interactive mapping of foodborne disease, ii) analyzing dynamic changes for foodborne transmission risk factor, and iii) framework of a foodborne disease management system. These potential contributions of the technologies in health hopefully could be beneficial to the epidemiologist and public health staff from JKN and other MOH institutions for:

i.   Monitoring the spread of infection diseases in order to identify the cause of the infection

ii.   Monitoring health services usage such as the uptake of vaccination or the use of community cares services

iii.   Mapping of the non-infectious diseases is valuable for generating hypothesis about disease causation or for identifying clusters of disease.

iv.   Mapping exercises can incorporate ecological analyses and can combat the disease effectively

ACKNOWLEDGMENT

REFERENCES

[1] A.R. Abdul Rasam, A.M. Mohd Noor, N. Ahmad & Ghazali, R. "*My*GeoHealth: GIS-Based Cholera Transmission Risk System in Sabah, Malaysia" Proceedings - 2011 IEEE 7th International Colloquium on Signal Processing and Its Applications, CSPA 2011;Penang;4 March 2011through6 March 2011;Code84971. Retrieved 28 May 2012 from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=05759925. 2011a

[2] A.R. Abdul Rasam, A.M. Mohd Noor, N. Ahmad "A Review of Geospatial Applications in Public Health and Health Care in Malaysia" 10th International Symposium & Exhibition on Geoinformation 2011 (ISG) & ISPRS 2011, 27-29 Sept 2011, Shah Alam Convention Centre Selangor, Malaysia. 2011b

[3] A.R.Abdul Rasam, A.M. Mohd Noor, W.M.N. Wan Mohamad, and, J.R. Abdul Hamid, "An Initial study of the relationship among cholera, sea surface temperature, and phytoplankton in Tawau, *Resources & Environmental Management and Environmental Safety & Health 2009* (NREM & ESH 2009) Sabah." *Proceeding in International Conference on Natural. 2010*

[4] A. Ismail, " An update on diarrhoeal diseases in Malaysia" Department of Medical Microbiology and Parasitology, School of Medical Sciences, University Sains Malaysia, Penang. Med J Malaysia. 2002

[5] B. Lobitz, L. Beck, A. Huq, B.Wood, G. Fuchs, A.S.Faruque & Colwell, R. "Environment & Health, Climate And Infectious Disease: Use Of Remote Sensing For Detection Of Vibrio Cholerae By Indirect Measurement" Proc. Natl. Acad. Sci. Usa 97, 1438–1443. 2002.

[6] K.T. Chang. *Introduction to Geographic Information Systems (4th Ed)*. The MaGraw Hill Companies, Inc, New York, Singapore, 1-418. 2008.

[7] J.M. Soon, H. Singh, & R. Baines, "Foodbornediseases in Malaysia: A review". Journal of Food Control, Volume 22, Issue 6, June 2011, Pages 823–830,

http://www.sciencedirect.com/science/article/pii/S0956713510004391. 2010

[8] Ministry Of Health Malaysia [MOH] Garis Panduan Umum Pengurusan Wabak Penyakit-Penyakit Bawaan Makanan Dan Air Di Malaysia Bahagian Kawalan Penyakit. (2nd Ed), Moh/Epi/23.00 (Gu). Fwbd/Umu/Gp/001. 2006

[9] WHO, World Health Organisation "Foodborne Diseases" Retrieved 28 May 2012 From Http://Www.Who.Int/Topics/Foodborne_Diseases/En/. 2012

[10] WHO, World Health Organisation, Geneva . Retrieved 20 May 2010 from http://www.who.int/whr/2007/en/index.html. 2007

[11] ARSM, Malaysian Remote Sensing Agency. Retrieved 20 May 2012 from http://www.remotesensing.gov.my/index.cfm

[12] P.R.. Hunter, M. Waite & E. Ronchi. *Drinking water and infectious disease: Establishing the links.* CRC press LLC and IWW, UK, Acid free paper, USA, 1-10, 143-153. (2003)

[13] D. Alan. *Managerial Epidemiology: Practice, Methods and Concepts* USA. Jones & Bartlett. (2006)

[14] C. Corvalan, D. Briggs & G. Zielhuis. *Decision-Making in Environmental Health, (1st Ed.).* E & FN spon, London, USA and Canada, *Tj International Ltd Padstow, Cornwall Great Britain*. 133-154. (2000)

[15] A.B. Lawson & F.L.R. Williams. *An Introductory Guide To Disease Mapping.* John Wiley & sons, LTD UK (1-14) & 53-75. (2001)

Reforming Health Care Facility Using Geographical Information System

Mohamad Taufik Lokhman, Abdul Rauf Abdul Rasam & Abd Malek Mohd Noor

Centre of Studies Surveying Science and Geomatics
Faculty of Architecture, Planning and Surveying
Universiti Teknologi MARA (UiTM), Malaysia
taufikzet_88@yahoo.com, raufrasam@yahoo.com

Abstract— The current application of geographical information system (GIS) in health care has been widely practised to channel information, particularly in developed countries. However, in Malaysia, the use of GIS application in healthcare is limited because of lack of exposure in delivering information to users or patients who need government health care center services. The problem that can be detected is that the crucial information needed by the patients or users such as the availability of expertise in every government hospital is not yet available in the system. This study tries to solve the problem by creating a GIS-based information analyses and system so that patients can find out the information about necessary treatment in a hospital based on specialist services provided. The finding showed the successfully integration of information on the actual position of all government health centers in Johor Bahru with the services and facilities provided. In addition, this study demonstrated the function of the shortest path in network analysis for finding the nearest health care facilities from patients' location. GIS technology not only can benefit patients or users, but it also can better management for all organizations if the database is fully integrated with all the government hospitals in all the districts of Johor towards creating an efficient MyGeoHealth (Malaysian Geospatial and Health) approach.

Keywords- GIS; mapping and location; network analysis; health care facility; MyGeohealth

I.   INTRODUCTION

Best services provided by each health care center will improve the quality of communities' health care. It includes the preparation of an information system that helps users to reach each health care center for treatment. In Johor Bahru, there are two government hospitals operating in the community. Therefore, by using GIS, it provides a set of guidelines which health information system can be refocused as a monitoring tool to improve the, quality, access and use of Health Management Information. This study also shows that an alternative way of improving the flow of health information is to dedicate resources specifically to co-ordinate access, use and ongoing development of relevant information to users.

As recommended by the World Health Organization (WHO), Malaysia should not leave behind the use of GIS in the health field even though it is still new in our country. It is the policy of the Ministry of Health (MOH) to expand the use of modern technology in their efforts to improve health in all countries, especially the use of computers and comprehensive management information system. The Municipal Council of Johor Bahru has initiated the use of GIS for diseases prevention and control in 1996 towards creating healthy community.

Traditionally, the location of each health care center in this country is based on addresses, without giving details of the actual position of the health center. There is no information system developed to describe the actual situation in spatial health care centers. With the help of GIS applications, the actual position of the health center can be mapped accurately and information such as coordinating each health care center can be obtained and pinpointed. The combination of spatial information, expertise available in the healthcare centers and road network will be able to give more information on travel distance and travel time from users' home to the nearest medical center.

II.   GIS APPLICATIONS IN HEALTH CARE IN MALAYISIA

Generally, a geographic information system is a computer-based information system that enables capture, modeling, storage, retrieval, sharing, manipulation, analysis, and presentation of geographically referenced data. The term GIS is near generic in usage, however, variant acronyms describe similar information system. For example, health information system (HIS) [1,5] describe geographic information system processing health utilization and resources database, respectively.

GIS integrates spatial and other kinds of attribute data within a single system. This offers the flexibility to combine data from a variety of disparate sources from many different databases. The process of converting maps and other kinds of spatial information into a digital form, via GIS, enables new and innovative methods for the manipulation and display of geographic data and knowledge to be formulated. GIS makes connections between activities based on geographic proximity these connections are often missing without GIS, but can be vital to understanding and managing activities and resources.

For example, we can link toxic waste records with school locations through geographic proximity.

A Health care Information System (HIS) includes a wide array of applications and information systems that are tied. It supports the provision of care to patients and the business aspects of the health care organization by communicating information. The gathering of information systems and the architecture by which they are integrated is unique for each organization. Early manual systems evolved to automated stand-alone systems, while current systems are well integrated. Two or more components that are merged together into one system are often referred to as an integrated system. The expansion of integrated systems is more pronounced today as information and its efficient communication have become vital to the survival of health care organizations in today's economy and regulatory world [7].

Organizational strategic aims and business plans determine the composition of the applications, hardware, and network structure that make up an individual HIS. A health facility is any location at which medicine is practised regularly. It ranges from small clinics and doctor's offices to urgent care centers and large hospitals with elaborate emergency rooms and trauma centers. The number and quality of medical facilities in a country or region is one common assess of that area's prosperity and quality of life. Types of health care facilities are based on the criteria determined by Ministry of Health.

  i.   Specialist hospitals
  ii.   Non specialist hospitals
  iii. Healthcare clinics
  iv.   Dental clinics
  v.   Rural clinics

According to previous studies, a GIS accessibility model based on least cost path analysis is a critical resource that can be used by health service planner. Poor access to primary health services, such as a general practitioner or doctor, can result in people with simple health problems not being advised by general practitioners or doctor and subsequently developing more complex problems with considerable discomfort that can be expensive to treat. An important consideration for such decision-making is where people are living, and their travel distance and time to closest hospital or clinic. Geographical access models developed using network analysis function in a GIS have enormous potential for informing policy development and grounding debate on how to achieve social equity of primary health care access.

III. METHODOLOGY

This section explains the implementation of this project. It covers the process of project planning, data collection, data processing, data analysis as presented in Figure 1. The first stage of this study is project planning which includes the

selection of the study areas and the software which is needed to implement this project. Johor Bahru is selected as study area which is located on the southern tip of the peninsular Malaysia and the state capital of Johor Darul Ta'zim.
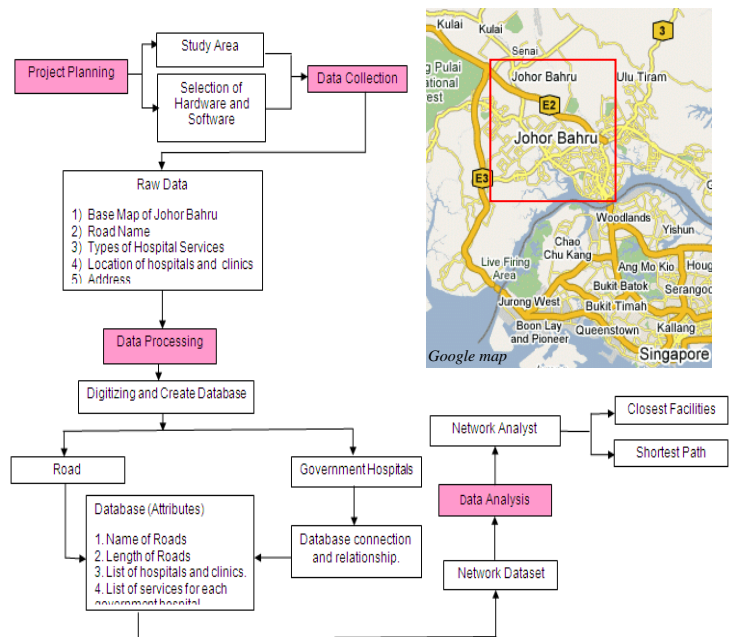
Figure 1. Research methodology

In the data collection, the data required is in the form of spatial and attribute. Spatial data that is required is road network and government hospitals. This spatial data is obtained through the process of digitizing base map of the Johor Bahru district. Road networks are categorized into three categories: highway, primary road and secondary road. For government hospitals spatial data, it must be connected to the information services provided at each government hospital. Attribute data that need to be stored in the database are roads' name and length and name of hospital and their facilities.

In the data processing, spatial data and attribute data are merged to build the GIS database that contains information required in data analyzing stage. For example, the road's name and its distance information, while hospitals with position information and the name of each hospital. In georeferencing basemap or also known as rectifying map, this is the process of registering map on the earth's surface and it is a very critical stage to obtain high accuracy in measurements. The digitizing process creates spatial data and allows the attribute data to be stored in together.

Creating Personal Geodatabase, is the core of the data processing stage. It is a place for storing all the information about the projection of maps, spatial data and attribute data. Database connection and relationship is done to join external database with an existing personal geodatabase that has been created. A network dataset is created from the feature source

or sources that participate in the network. It incorporates an advanced connectivity model that can represent complex scenarios, such as multimodal transportation networks. Preparation of the network analyst such as the Shortest Route using Shortest Path tool: the closest facilities using closest facility tool.

IV. RESULT AND ANALYSIS

The analysis focuses on the creating of geo-database and network analysis of health care services.

A. Creating map and geo-database of healthcare facilities

To integrate information on the actual position of all government hospitals in Johor Bahru with the services and facilities provided. After all the information about government hospitals has been integrated, the result is shown in Figure 2. It describes the graphical view of government hospitals on the earth's surface. From this information, the viewer can imagine the location of the government hospitals and its surrounding. This information gives a better understanding to users to find the hospitals from their location rather than having only the address of government hospitals as reference.
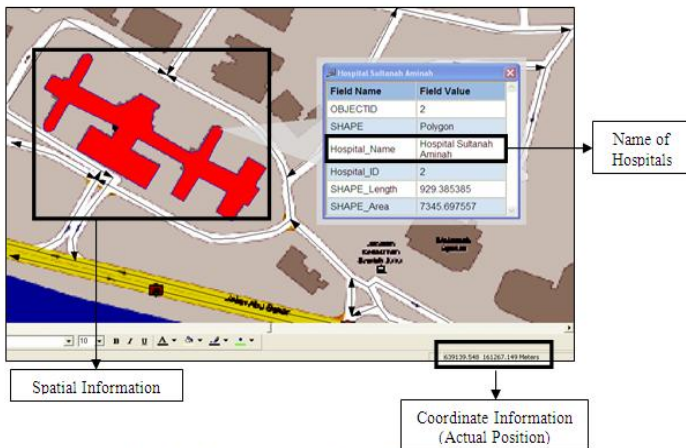
Figure 2. Spatial and related information about Hospital Sultanah Aminah, Johor

By making database connection and relationship with another database, the delivering of information is increased. In this project, problems arise when the relationships among the database in ArcGIS software permits only one-to-one and many-to-many relationship which only one spatial data can have one record. This database relationship allows many tables or other database related to each spatial data (Figure 3).
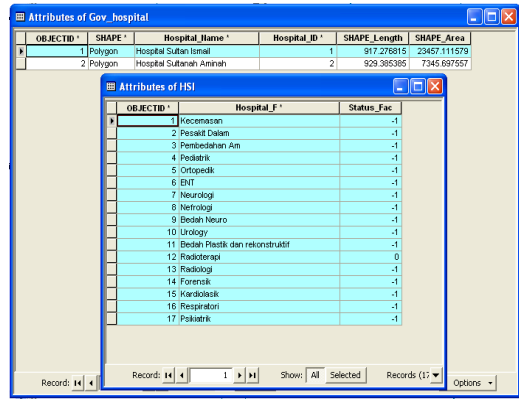
Figure 3. Database and services provided for Hospital Sultan Ismail, Johor

B. Analyses of network analysis for healthcare facilities

This part is to analyze the shortest path to find the nearest health care centers from patients' location or incident area and closest facility services.

1) The Shortest Path Analysis

The shortest path analysis in Network Analyst tool used in this project is to identify the shortest path to reach the health centers in terms of the shortest distance from the patient's or user's location. The result from the shortest path analysis shows the shortest route from the location of patients to selected government hospital in the shortest distance criteria. The shortest path analysis will find the direction starting from the patients' location (Figure 4).
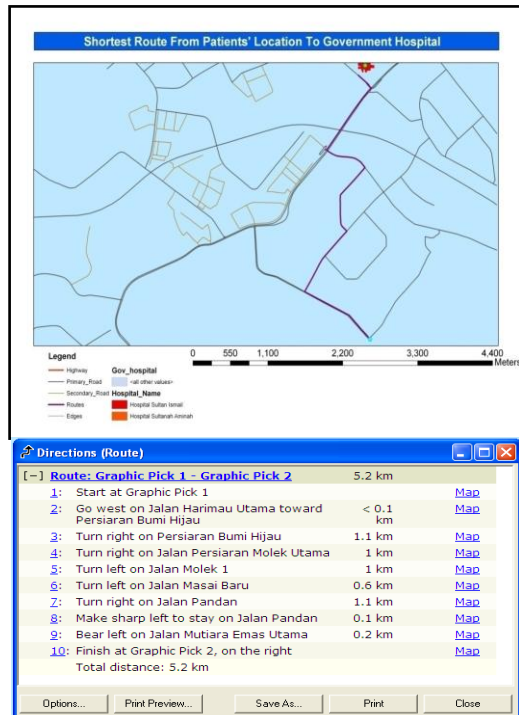
Figure 4. Shortest route analysis result & Direction for shortest path

2) The Closest Facility Analysis

In this project, the closest facility analysis will be conducted when there is more than one facility chosen from the patient's location. For example, Hospital Sultanah Aminah and Hospital Sultan Ismail provide an emergency service. But for patients to find the nearest hospital and the shortest route for emergency cases, the closest facility analysis is necessary to be carried out (Figure 5).
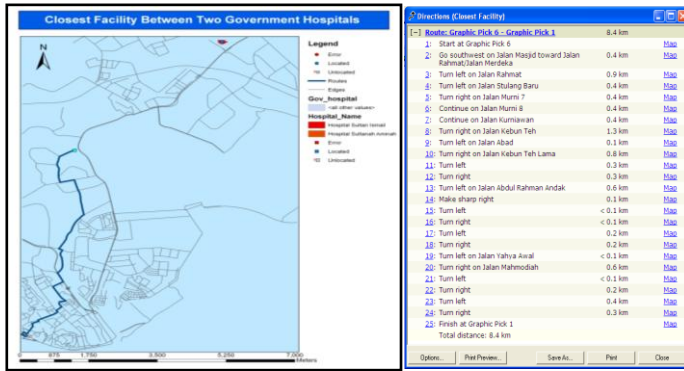
Figure 5: Result of closest facility analysis & Direction for closest facilities analysis result

V. CONCLUSION

The objective of integration of information on the actual position of all government hospitals in Johor Bahru with the services and facilities provided is to put the GIS element in delivering health information. Health care facilities such as government hospitals are a focal point for communities to seek treatment. With the use of GIS in the delivery of information, it can help in solving the problems faced by the public. Meanwhile, Network analysis which is carried out in this project is a process of manipulating road network data and hospital locations to generate new information to facilitate public's obtain direction to reach the hospital. The direction will give the shortest route from the patients' location to hospital. GIS application in health care facilities initiates a new era in delivering information to the public. It not only can benefit patients or users, but it also can give better management for all organizations if the database is fully integrated with all the government hospitals in the districts of Johor.

REFERENCES

[1] J. Nicol. "Geographic Information System within the national Healths, the scope for implementations, planning out look" vol. 34, No. 1, pp. 37-42. 1991

[2] Jabatan Kesihatan Negeri Johor. 2007

[3] K. Omar. *Geographic Information Systems and Health Applications*, 289-305. 2003

[4] K.T. Chang. *Introduction to Geographic Information Systems (4th Ed)*. The MaGraw Hill Companies, Inc, New York, Singapore, 1-418. 2008

[5] L. Twigg. "Health based geographical information systems: their potential examined in the light of existing data sources". Soc Sci Med. 1990;30(1):143-55. 1990

[6] S. P. Englebart,   & R.. Nelson,. *Health Care Informatics an Interdisciplinary Approach*, 150. 2002

[7] V.K. Saba & K.A. McCormick. *Essentials of Computers for Nurses (2nd Ed.)*. NewYork: McGraw-Hill. 1996

MONITORING CELLULAR ACTIVITIES USING CELLTRACK WITH GPS ENABLE

Mohd Saufi Hj Nasro Ali, Azita Laily Yusof, Melati Ismail, Norsuzila Ya'acob

Faculty of Electrical Engineering
Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia
saufinasroali@yahoo.com, laily012001@yahoo.com, melati_mona@yahoo.com, norsuzilayaacob@yahoo.com

Abstract - **The challenge that the modern technology is facing today is to provide a quality and better signal strength to users. Since the most widely deployed and used mobile computing device today is the mobile phone, so there is a need to provide good services to the customers. The purpose of this project is to present the study on the impact of Received Signal Strength (RSS) and signal quality between two network providers which are Celcom (013) and Digi (010) in Shah Alam area. Both of these providers are the main telecommunication service providers in Malaysia. This paper focused on the collecting and analyzing data by monitoring cellular activities using the CellTrack software. These processes involving in the use of two NOKIA phone with the Global Positioning System (GPS) that enable the data on coordinate of every checkpoint along the route to be recorded. The data was collected in three days and three times per day for fine day and one data collection for raining day in order to see any variation in the Received Signal Strength (RSS) and signal quality at different time and different weather condition. The result shown that the Digi provider is better in received signal strength (RSS) compared to Celcom provider.**

Keywords- Receive Signal Strength (RSS); CellTrack; quality signal; cellular; GPS enable

I.   INTRODUCTION

Global System for Mobile communication (GSM) is a digital mobile telephony system that is widely used in Europe and other parts of the world [1]. In Malaysia, we are know that there are several company that provide mobile services such as Celcom Axiata Berhad, Maxis Communication Berhad, Digi Telecommunication Sdn Bhd and U Mobile. This entire mobile network has their ongoing investments in network coverage, capacity and performance, and intends to maintain their technology leadership and position as the country's best mobile service provider. Although all of the companies use the same system, but there are differences in the quality signal. This is because the signal normally varies with environmental factors such as building barriers, trees, traffic flows and how far a mobile phone is located from the Base Transceivers Station (BTS).

The Mobile Station (MS) is carried by the subscriber. The Base Station Subsystem (BSS) controls the radio link with the MS [2]. In a cellular radio system, hexagonal shaped cells were used in a land area to be supplied with radio service. Each of these cells is assigned multiple frequencies which have corresponding radio base stations. Cells may vary in radius in the ranges of 1 km to 30 km. The boundaries of the cells can overlap between adjacent cells and large cells can be divided into smaller cells [3]. The smaller the cell size and the faster the movement of the MS through the cell, the more handover of going call required. In additional, the location update is needed in the handover procedures.

Extended study on previous similar researches shows that there are several methods that can be used to perform signal strength and quality testing such as Test Mobile System (TEMS) [4], and Reliability and System Architecture Testing (RSAT) [5]. For TEMS tools, it basically consist of two TEMS mobile phone that must been control by software running on a personal computer. When both of a mobile phone has connected to laptop, the software then can be used to monitor parameters and to control the transmitting of message between the mobile and the cellular network. For RSAT tools, it must be connected to personal computer. The signal strength data were recorded and stored in the laptop. Both of the methods used were slightly same in term of measuring signal strength. According to author [6], the site simulations are obtained by using software named Asset3g by AIRCOM International. However, strong foundation is needed to use the software as it is complex software.

Thus, we bring out this research in order to study and analyze the different in Received Signal Strength (RSS) and signal quality between Celcom and Digi networks service provider in Shah Alam area. The content of this paper is as follows. Section 2 describes the software description and GPS visualize. Section 3 focuses on software installation, data collection and drive test study area. Section 4 covered the results for case study area and discussion for the results. We conclude our work in section 6.

II. SOFTWARE

A. Software Description

CELLTRACK is a program to collect some phone information about the cell you are connected to - like the net monitor and this software only support mobile phone that using Symbian Operating System. Figure 1 shows the main screen of the CELLTRACK91 software and the version for Symbian 3rd Generation is only fit for S60 operating system.
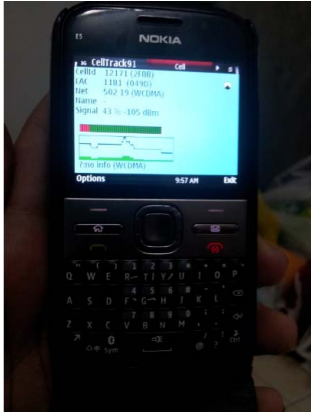

Figure 1: Software CELLTRACK91 views.

The GSM information and functions showed as below [7].

- **Cell id** – the id from the actual cell

- **LAC** – the location Area Code from mobile station

- **NET**- the net we are moving example, 262 for Germany and 02 for Vodafone. If WCDMA/UMTS and GSM cell is present is shown here. The number has still 4 digits.

- **Name** – the name of cell. This is shown by an algorithm when we select 'Cell Name from Id' in the setting or from the cell broadcast when we have turned 'Cell Name from CBS' on and have defined a service number. Otherwise one a '-' is shown. If both are turn on CBS overwrites the generated name.

- **Signal** – shows the quality of receive signal. The percentage value is taken from the Nokia phone API (Application Programming Interface). The dBm values are covered at command (CELLTRACK62) or the newer API (CELLTRACK70).

- **Battery** – shows the capacity of the battery.

- **Description** – shows the information from the database.

- **Diagram** – shows the dBm values as green and the percentage as black line. The red lines a cell change.

B. GPS Visualizer

23 different locations are chosen through Shah Alam area. The coordinate at every location is measured using mobile phone with GPS enable. The data from mobile phone can be transfer to software name GPS visualizer. This software is simplified and flexible in terms of input, and its enormous number of options regarding the output. When we upload the GPS data, GPS Visualizer will automatically detect what kind of file it is and process accordingly [8]. In this research, the output will be present in form of Google Earth and 3D view. This software used to get the information and displayed along driving routes in Shah Alam area.

III. METHODOLOGY

A. Flow Chart

Special software program which called as CellTrack is used to collect the data. This software is very specializing program because it can collect detailed information about the carrier network and GSM cell that you are connected to, just like the net monitor. This software only can be used for the hand phone that have Symbian platform. Symbian is a mobile Operating System (OS). Figure 2 illustrates the procedure of the project.
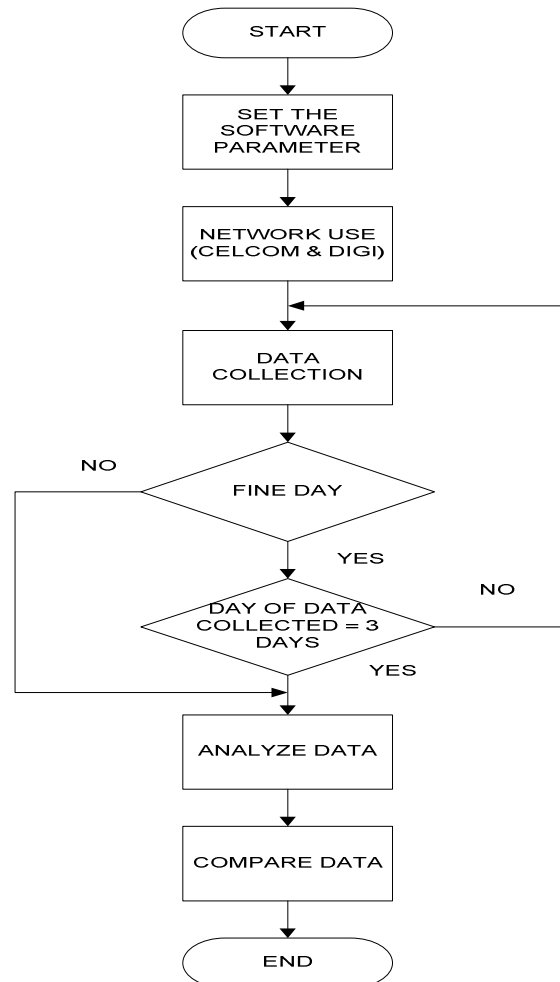

Figure 2: General flowchart represents of the project

B. Software Installation

CELLTRACK software can be downloading through the website [7]. User was reminded to download suitable version of software depending on mobile phone model. In this project, software Celltrack91 - Version for S60 Symbian 3rd Generation was been download is used.

Once the software was downloaded, user must get the certificate to sign in the software [9]. The user required to enter International Mobile Equipment Identity (IMEI) number.

C. Drive Test Study Area

Data collection has been taking by using two models NOKIA E5 phone. This model is used because it is already equipped with Global Positioning System (GPS) and enable navigate to the target location. In addition, GPS from the phone was also used to get the coordinate at 23 throughout Shah Alam area as shown in Figure 3 and Table 1 shows the coordinate system for the target location.

The data collection has being taken at several locations which is in the specific target area. This process was conducted over four days, where the three days run on the fine day while one day for raining day. For data collection in fine weather, three sessions will be conducted, during the early morning (8.00am to 9.00am), afternoon (12.00pm to 1.00pm) and evening (5.00pm to 6.00pm). While for data collection on the rainy day, only one session has been done. In this study, the data was taken during the afternoon (12.00pm to 1.00pm).This data collection was carried out in urban areas involving around Shah Alam.

The data information was collected while the vehicle moving along the study area. To simplify this process, at least two people are required which mean one to drive the vehicle while another is required to record all data required. Vehicle is driven at speed 30 km/h to 50 km/h depending on traffic flow. After that, the data will be automatically saved in the mobile phone memory and it can be viewed later via notepad or Microsoft Excel when the data has been transferred to the computer.

IV. RESULT AND DISCUSSION

The drive test studies between two network providers which are Celcom and Digi have be done in urban areas involving around Shah Alam.
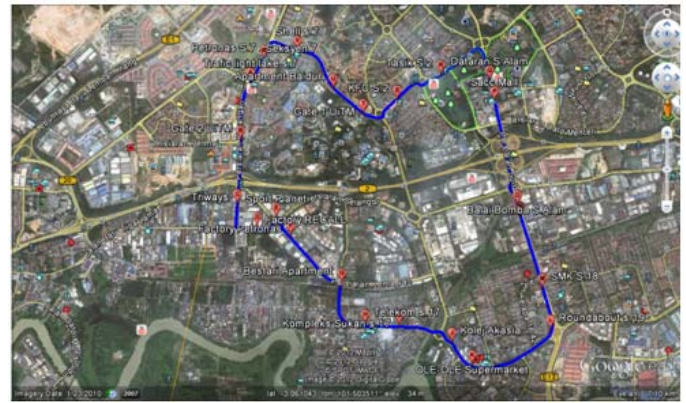

Figure 3: The blue line represents the route .The red dots represents the coordinate of the checkpoints

There are 23 different locations was chosen throughout Shah Alam area. The coordinate at every location is recorded by mobile phone and then transferred into Google Earth. Table 1 shows the coordinate system for the target location.

TABLE 1: Coordinate System for Selected Location

PT	DISCRIPTION	COORDINATE		
		LATITUDE	LONGITUDE	ELEVATION(m)
1	Traffic Light Section 7	3.0738988°	101.4924804°	11
2	Gate 2 UiTM	3.0676330°	101.4914157°	7
3	Triways	3.0607268°	101.4910750°	7
4	Brambles Corporation Sdn. Bhd	3.0582690°	101.4932038°	7
5	Sport Planet	3.0592726°	101.4953000°	12
6	Petronas Factory	3.0573066°	101.4968446°	18
7	Bestari Apartment	3.0521883°	101.5023553°	13
8	Telekom Section 17	3.0477131°	101.5048947°	5
9	Sport Complex Section 17	3.0475791°	101.5085631°	5
10	College Akasia	3.0457504°	101.5142799°	7
11	Ole-Ole Supermarket	3.0433209°	101.5164534°	8
12	Roundabout Section 19	3.0471337°	101.5249270°	9
13	SMK Section 18	3.0516029°	101.5240472°	9
14	Police Station Section 15	3.0593532°	101.5216457°	9
15	Fire Station Section 15	3.0606190°	101.5213295°	10
16	SACC Mall	3.0718759°	101.5188347°	13
17	Dataran Shah Alam	3.0741089°	101.5183150°	24
18	Taman Tasik Shah Alam	3.0746958°	101.5130558°	8
19	KFC Section 2	3.0719646°	101.5083305°	14
20	Gate 1 UiTM	3.0704814°	101.5046804°	26
21	Baiduri Apartment	3.0731699°	101.5014515°	44
22	Shell Section 7	3.0773615°	101.4975588°	13
23	Petronas Section 7	3.0763386°	101.4939660°	12

The drive test study of the target location on the tracking performance was carried on 1st, 3rd and 5th May 2012 for a fine weather. For raining day, the drive test study has been done on 2nd May 2012 at afternoon. In this research, 3 different days and times are taken for both operators.

    1) Morning: 8 am – 9 am
    2) Afternoon: 12 pm – 1 pm
    3) Evening: 5 pm – 6 pm
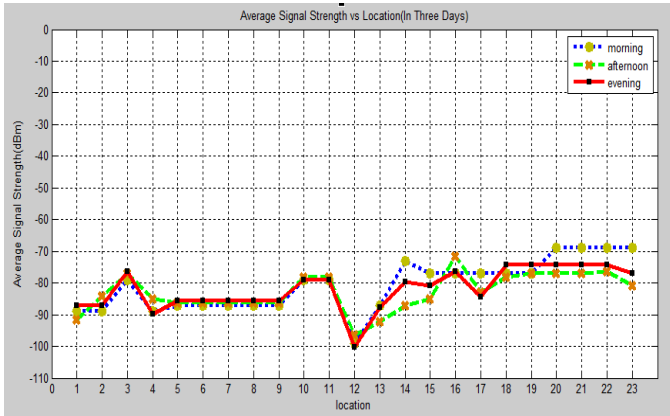
A. CELCOM and DIGI

Figure 4: Graph of Average Signal Strength (CELCOM) versus Location (in three days)

The standard range of the RSS is between -110dBm to -10dBm. For practical purpose, the range of RSS is considered good if the reading is greater than -90dBm. If the reading is lower than -90dBm, the signal strength is considered bad [4]. Figure 4 is the combination of the average signal strength to show the comparison between morning, afternoon and evening for Celcom provider. The signal strength was quite similar for each time. From the graphs, it can be shown that signal strength in the morning is mostly higher compared to other condition.
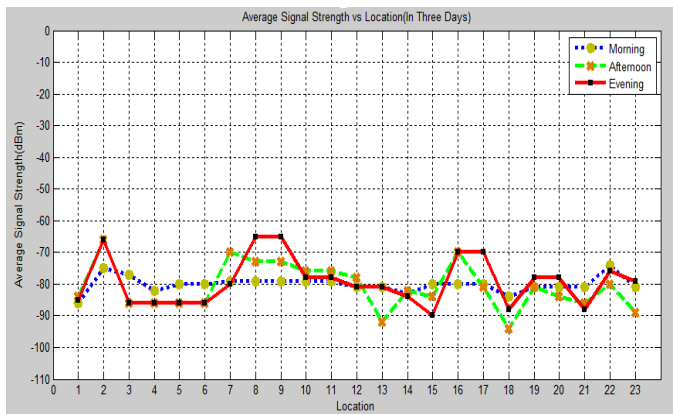
Figure 5: Graph of Average Signal Strength (DIGI) versus Location (in three days)

Figure 5 shows the average of three readings for Digi provider. The signal shows the comparison of signal strength in the morning, afternoon and evening.

From previous provider, it showed that the signal strength was quite similar for three days. But for DIGI provider, it shows that the signal strength distribution pattern is quite different for each time. This is because the signal strength continuously fluctuates due to noise and interference. However, the average signal strength in the morning for three days is quite flat pattern and the signal strength along the trip in the range of -86dBm until -74dBm. This result is considered

good and this is because the signal strength is captured from the nearest BTS.

From the figures, we can see the average receive signal strength in several location is lowest for three days in the afternoon. This is because of the attenuation of path loss and surrounding environment due to atmospheric effect. Besides that, the weather condition was the effect of received signal strength. This is because the temperature in the afternoon is quite hot and in some study showed that signal strength may perform better during the day or at night [10].

B. Comparison Average Signal Strength for Fine Days

Figure 6: Graph of Average Signal Strength versus Location (Morning)
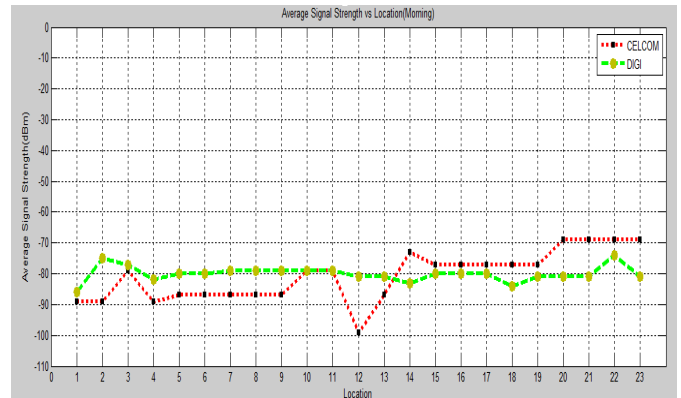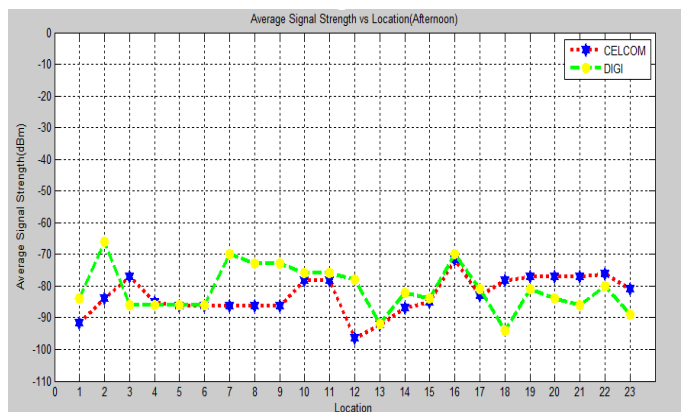

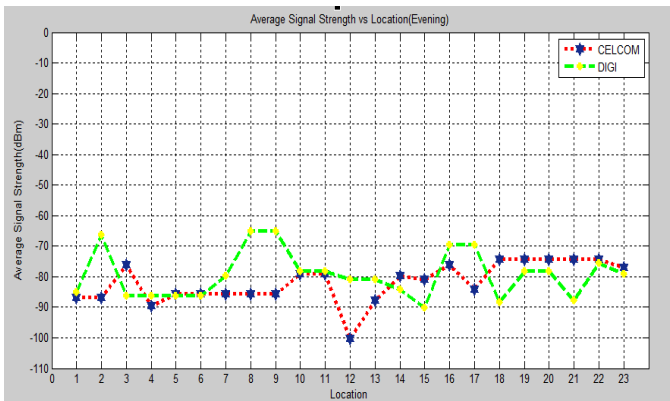Figure 7: Graph of Average Signal Strength versus Location (Afternoon)

Figure 8: Graph of Average Signal Strength versus Location (Evening)

Figure 6 – 8 show the graph of average signal strength versus location (Latitude) in different times between two service providers which are Celcom and Digi. From the graph, we can see that the signal strength fluctuates because of interference.

The graph shows that the signal strength for Digi is mostly higher than Celcom. This is because in urban area, Digi provider has more repeater of antenna than Celcom provider. Signal strength can be maintained when there are lots of antenna repeaters present in that particular location. The maximum signal strength for Digi collected is -65.0 dBm in the evening and located area Telekom S.17,(8) and Sport Complex S.17,(9) while the minimum signal strength collected is -94 dBm in the afternoon and located at Taman Tasik Shah Alam S.2,(18). For Celcom, maximum signal strength collected is -69 dBm in the morning and located between Gate 1 UiTM,(20) until Petronas S.7,(23) while the minimum signal strength collected is -100.3 dBm in the evening and located at Roundabout Section 19,(12). The minimum signal strength for Celcom is constantly low at this area for each time. Even these locations are the near away area to the BTS, but in reality the mobile station will received the signal strength from another BTS. For short distance which 40 meter or less, this is not strictly correct and less approximation between the base and the mobile because of the dependence of the antenna patterns on elevation angle. This value only can be excellent approximation when the distance greater than 40 meter [11].

C. Signal Strength on Raining Day

Basically the signal strength will be decrease on the raining condition. Figure 9 shows Signal Strength versus Location on Raining Day between two providers. The data has been collected on 2nd May 2012 in the afternoon around 12pm to 1pm. From the graph, it shows that the signal strength for DIGI provider is good in raining as compared to CELCOM provider. This is because the BTS for Digi is close compared to Celcom even in certain location the BTS is sharing between these two providers. The other reason is Digi provider was less interference such as building or trees. When raining condition, the signal strength for mobile phone slightly different for both

providers due to this weather condition has less effect interference to the signal.
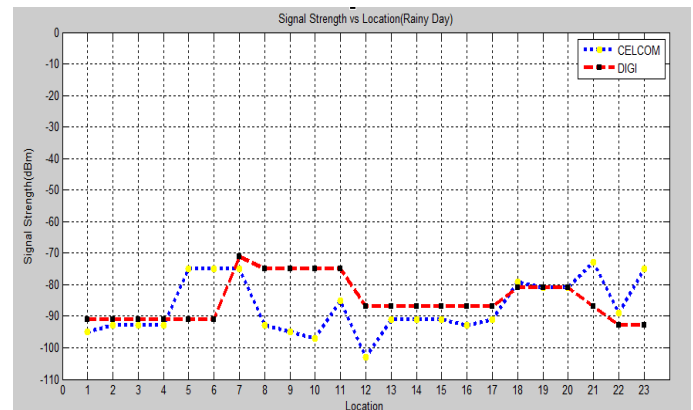
D. Comparison between Fine Day and Rainy Day

Figure 9: Graph of Signal Strength versus Location (Raining Day)

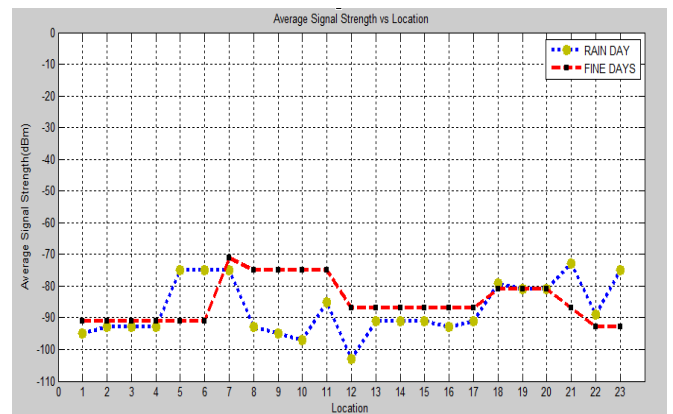
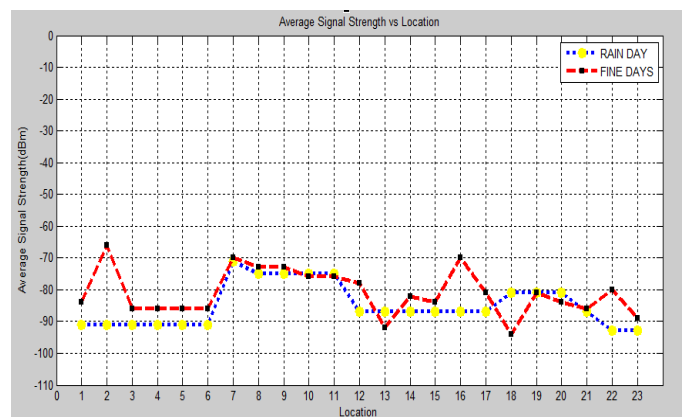Figure 10: Graph of Average Signal Strength versus Location Fine days and Rainy day (CELCOM)

Figure 11: Graph of Average Signal Strength versus Location Fine days and Rainy day (DIGI)

Figure 9 – 11 show the comparison of signal strength between fine and rainy day. The result from both providers in the fine day time was being compared with the result on raining time. From the graph, we found that the signal strength decrease in almost locations in Shah Alam area when rainy day. One of the factor effect the signal strength is because of rain attenuation. The precipitation in the atmosphere cause signals to be absorbed and weakened. In Malaysia General Packet Radio Service (GPRS) operates in low frequency band in the ranges of 900MHz to 1.8GHz and is not affected much by rain attenuation [12].

VI.   CONCLUSION

As a conclusion, based on data obtained from monitoring cellular activities, different area has different signal strength for both Celcom and Digi network providers. The comparison of the signal strength between the two networks providers has been analyzed and we found that in Shah Alam area, Digi has better signal strength as compared to Celcom. This showed the impact of the RSS of the target on the tracking performance for both network providers.

Future work will be focused on the simulation work between test bed and simulation environment by considering the prediction of the channel and use's behavior. The further work is needed to find the most optimize handover procedure by comparing the test bed and algorithm with the handover decision criteria specific by the standard.

ACKNOWLEDGMENT

REFERENCES

[1]   Timothy Sohn, Alex Varshavsky, "Mobility Detection Using Everyday GSM Traces," University of California, 2006, pp 3.

[2]   John Scourias, "Overview of Global System for Mobile Communication," University of Waterloo, 1995, pp. 4.

[3]   Ahmed Hefny, "Speed of different cellular networks," Carnegie Mellon University, 2012, pp. 4–5.

[4]   N. A. B. Mattar, *et al.*, "Measuring and analyzing the signal strength for Celcom GSM [019] and Maxis [012] in UiTM Shah Alam campus," in *Research and Development, 2002. SCOReD 2002. Student Conference on*, 2002, pp. 489-493.

[5]   J. F. Wagen, "Signal strength measurements at 1881 MHz for urban microcells in downtown Tampa," in Global Telecommunications Conference, 1991. GLOBECOM '91. 'Countdown to the New Millennium. Featuring a Mini-Theme on: Personal Communications Services, 1991, pp. 1313-1317 vol.2.

[6]   S. Ting and C. Tiong Teck, "WCDMA Network Planning and Optimisation," in Telecommunication Technologies 2008 and 2008 2nd Malaysia Conference on Photonics. NCTT-MCP 2008. 6th National Conference on, 2008, pp. 317-322.

[7]   www.afischer-online.de/sos/celltrack

[8]   http://www.gpsvisualizer.com/about.html

[9]   http://www.s60certkey.com

[10] Kenneth Bannister, et al, "Sensor networking for "Hot" application: Effects of temperature on signal strength,data collection and localization," Arizona Atate University, 2010, pp. 1 – 5.

[11] S. Aguirre, *et al.*, "Signal strength measurements at 915 MHz and 1920 MHz in an outdoor microcell environment," in *Universal Personal Communications, 1992. ICUPC '92 Proceedings., 1st International Conference on*, 1992, pp. 01.04/1-01.04/7.

[12] I. Chan, *et al.*, "Understanding the effect of environmental factors on link quality for on-board communications," in *Vehicular Technology* Conference, 2005. VTC-2005-Fall. 2005 IEEE 62nd, 2005, pp. 1877 -1881.

[13] I. Kostanic, *et al.*, "Measurement based QoS comparison of cellular communication networks," in *Communications Quality and Reliability, 2009. CQR 2009. IEEE International Workshop Technical Committee on*, 2009, pp. 1-5.

Investigation of Common Compounds in High Grade and Low Grade A*quilaria Malaccensis* using Correlation Analysis

Nurlaila Ismail, Mohd Hezri Fazalul Rahiman,
Rozita Jailani, Mohd Nasir Taib
Faculty of Electrical Engineering
UiTM Shah Alam, Selangor, Malaysia
nrk_my@yahoo.com, hezrif@ieee.org,
rozita@ieee.org, dr.nasir@ieee.org

Nor Azah Mohd Ali
Forest Research Institute Malaysia
52109, Selangor, Malaysia
norazah@frim.gov.my

Saiful Nizam Tajuddin
Faculty of Industry, Science and Technology,
University Malaysia Pahang, Malaysia
saifulnizam@ump.edu.my

Abstract—**This paper presents a correlation analysis between high grade and low grade of *Aquilaria Malaccensis (AM)*. Seven selected samples of *AM* were identified using Gas Chromatograph – Mass Spectra (GC-MS) and several statistical techniques have been employed. The samples were separated into two grades such as high grade and low grade. The analysis involves of normality test such as Shapiro Wilk, standard deviation, skewness, kurtosis and correlation analysis. All the analysis was done using SPSS ver. 15.0. The result shows that there are no significant correlation between high grade and low grade and it is a high correlation among compounds in similar grade. Three common compounds and highly correlated found in both grade and they are hexadecanoic acid, valencene and α-guaiene.**

Keywords-correlation, high grade, low grade, Aquilaria malaccensis, agarwood oil

I. INTRODUCTION

The *AM* is one of the species of agarwood oil. The agarwood oil is extracted from the highly valuable wood, agarwood or gaharu. The wood was found in *Aquilaria* species, also known as aloeswood, oudh, kanankoh or eaglewood [3-5]. In 2011, researchers found 15 species of agarwood [6]. The *Aquilaria Malaccensis* has a big population in Malaysia forest [7] and *Aquilaria Agallocha Roxb* is originately found in Japan [8]. In Japan, the wood is called according to their quality, Kanankoh for high quality and Jinkoh for low [8-10]. The agarwood oil is much appreciated in various application such as in medicine, religious prayer, as incense, perfumery and as a symbol of wealthy [5]. Thus, the market demand for agarwood oil is increasing [7, 11]. Not to be surprised if 1 tola (12cm^3) of oil is cost RM400 to RM2000 [12].

Based on human sensory panel, conventionally the agarwood oil is graded according to its colour, lasting aroma, high fixative and consumer perception [2, 13]. However, standard agarwood oil quality is limited due to human nose cannot accept many samples in one time and easily get fatigues especially when dealing with continuous production [14]. The human sensory panel also limited in terms of subjectivity, poor reproducibility, time consumption and large labor expense [15]. These are constraining factors to increase agarwood oil trade and penetrate the market [6].

Laboratory technique, such as GC-MS analysis [6, 8, 10, 14, 16, 17] was preffered to solve the conventional problems. GC-MS analysis identifies the significant existence of common compositions such as valerianol [2, 3], pentadecanoic, tetradecanoic, dodecanoic and hexadecanoic acid [3], elemol [2, 7, 13], β-agarofuran, α-agarofuran, epi- -eudesmol and agarospirol [2] and α-guaiene [8]. GC-MS discriminates the volatile compounds of agarwood oil [4, 5] to different quality and showing difference in oil compositions [5, 8]. For example, agarospirol and β-agarofuran as Type A and α-agarofuran, 10- epi- -eudesmol and oxo-agarospirol as Type B [18].

The GC-MS helps to relate the compounds and odour based on odoriferous properties in agarwood oil [8, 16]. For example, (-)-guaia-1(10),11-dien-15-al has a pleasant β-damascenone-like woody note with a touch of camphor and (-)-selina-3,11-dien-9-one has a fresh and sweet odour to represent gorgeous and elegant character of agarwood oil [8].

 Instead of represent on odour, GC-MS also described colour by its profiles. Noteworthy, different colour corresponds to different quality of oil [8, 10, 17]. Both colour and odour has a complex mixture of sesquiterpenes and chromone derivatives, as indexed by GC-MS [19]. This analysis extracted major volatile components such as sesquiterpenes and it is popular in market [9]. The studies on colour in oil volatile components has been carried out and they found that different colour has different chemical constituents [6, 14]. For example, GC-MS chromathographed that oxidoagarochromone as a pale yellow in oil component [19].

 In this study, *Aquilaria Malaccensis* identified using GC-MS for analysis. Several statistical techniques were employed including boxplot and normality test which is consists of Shapiro-Wilk, standard deviation, skewness and kurtosis. Then the study was end up by correlation test between high grade and low grade, and among similar grade. Lastly, a conclusion is made.

II. METHODOLOGY

A. Data preparation

The GC-MS data is obtained by the seven selected *Aquilaria Malaccensis* from Forest Research Institute Malaysia (FRIM) [20]. The data comprises of seven different samples with 44 chemical compounds. GC-MS analysis was performed by referring the method in [21]. Extending from the previous work [1], six significant compounds extracted using PCA were reported and will be used in this study. They are dihydro collumellarin, hexadecanoic acid, α-guaiene, elemol, valencene and thujopsenal.

B. Data analysis

A box plot is used to have an ideal/general impression on the data under study and represent the standard deviation for the entire sample under measurement [22]. From the box plot median, the skewness of sample can be defined by looking at the whiskers [23]. The skewness is relates to the distribution of the data either it is normal distribution or not. The outliers is defined as data points or scores that are different from the remainder of the scores [23]. Normality test is part of data pre-processing was performed to inspect the histograms of scores on each variables [24]. There are many ways that can be used to check the normality of data and four of them are Shapiro_Wilk, standard deviation, skewness and kurtosis. Shapiro-Wilk test is well known as an effective test to detect Gausian or Non-Gaussian of any signal [25]. The test proved that the result provide reasonable performance and well suggested for the normality test [25]. Standard deviation shows that the data deviation from the mean of sample [24]. Another method for normality test is skewness, it is measured by plus minus twice of its standard error. Skewness is applied to explain the scores of the histogram in a way such as positive skewness and negative skewness [24, 26]. The fourth normality technique is kurtosis. Kurtosis is used to know either the data distribution is fat-tailed or thin-tailed. Fat-tailed means the data scores in each tail has a greater than normal proportion and vice versa for the thin-tailed [27]. In other word, kurtosis is plus minus twice of its standard error. Kurtosis test was performed better in a study of pulsed-sinusoidal radio frequency interference [25]. The performance of Shapiro-Wilk Normality test is measured if sig. (p-value) is greater than 0.05, then the data is Gauss (Normal) distributed. Otherwise, the data is not normal. Standard deviation, skewness and kurtosis are applied using the equation in (1), (2) and (3), respectively. The distribution is normal when the skewness and kurtosis are within the range of plus minus twice of their standard errors.

$$ s = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^{n} (xi - \overline{x})^2} \tag{1} $$

$$ G_1 = \frac{\left(\sqrt{n(n-1)}\right) g1}{n-2} \tag{2} $$

$$ G_2 = \frac{(n-1)[(n+1)g2 + 6]}{(n-2)(n-3)} \tag{3} $$

where, s is a standard deviation, N is number of scores in sample, $\overline{x}$ is sample mean, Σ is sum of .., G_1 is skewness, and G_2 is kurtosis.

Data grading is done by separating six compounds to two grades, high grade and low grade. As in previous research published by Ali et. al [2], four of seven samples are graded as high grade. The rest, the samples are graded as low grade.

In this study, differences between 2 groups analysed using two assessments:

i) if the sig. (p-value) is equal or less than 0.05, there is a significant difference in the mean score between High quality and Non High quality, and

ii) if the value is above 0.05, there is no significant difference between High quality and Non-High quality.

Correlation test is used to explore the strength or the relationship between variables and groups. In a simple term, it will indicate whether the concerned groups or variables are correlate or not, and the degree of their correlation [28].

III. RESULT AND DISCUSSION

Fig. 1 and Fig. 2 show the distribution for high grade and low grade compounds, respectively. No outliers are detected in both figures. However, it is noticed that, every compound has different median, minimum and maximum values. For high grade, the median is the minimum value for three compounds such as hexadecanoic acid, elemol and valencene. For low grade, two compounds are identified to have same result i.e. the median is minimum value. The compounds are valencene and thujopsenal. Graphically, two different pattern of median value are detected between high grade and low grade. Their correlation is presented in Table III and will be explained later.
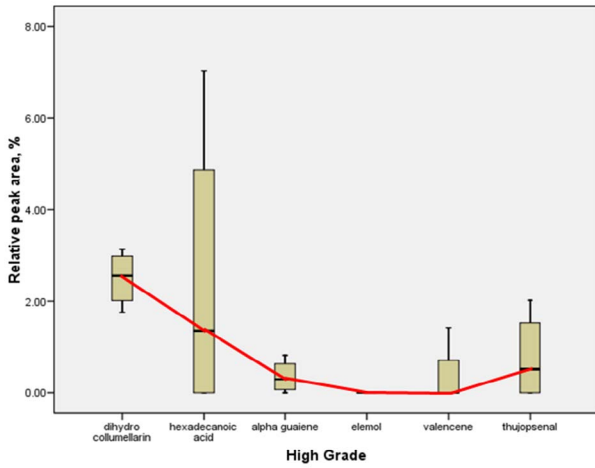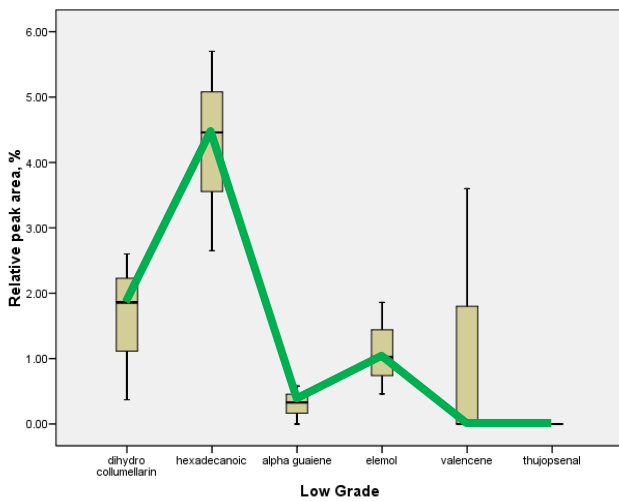
.

Fig. 1Distribution for high grade compounds

Fig. 2Distribution for low grade compounds

Table 1 shows the normality test for on normalized GC-MS data of *AM* using Shapiro-Wilk Normality test. The test resulted four compounds have sig. values greater than 0.05. The compounds are dihydro collumellarin, hexadecanoic acid, α-guaiene and elemol with 0.442, 0.553, 0.591 and 0.014, respectively. The other two compounds are not normal distributed.

TABLE I.   TEST OF NORMALITY

Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
dihydro collumellarin	0.916	7	0.442
hexadecanoic acid	0.930	7	0.553
alpha guaiene	0.935	7	0.591
elemol	0.754	7	0.014
valencene	0.628	7	0.001
thujopsenal	0.647	7	0.001

Table II shows the descriptive statistic for six compounds involved of standard deviation, skewness and kurtosis value. Among all compounds, the most deviated compound was hexadecanoic acid with 2.69338 and the least deviated compound was α-guaiene with 0.30846. It was observed that three compounds has the value within the range ±1.588 (refer (1)) such as hexadecanoic acid, alpha guaiene and elemol with 0.062, 0.428 and 1.476, respectively. It means the skewness not seriously violated, and then the distribution for these compounds are normal distributed. From this technique, the other three compounds are not normal. For kurtosis, all compounds have the values in the range ±3.174 (refer to kurtosis calculation in (2), except valencene. The value is not within the range, such as 3.769. From the kurtosis result, all the compounds are found to be normal distributed except valencene which is not normal distributed or leptokurtic. Generally, hexadecanoic acid, α-guaiene and elemol had passed both tests to qualify them as normal distributed.

TABLE II.   THE DESCRIPTIVE ANALYSIS

	N	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
dihydro collumellarin	7	2.1229	.91564	-1.197	.794	1.772	1.587
hexadecanoic acid	7	3.2214	2.69338	.062	.794	-1.267	1.587
alpha guaiene	7	.3314	.30846	.428	.794	-.943	1.587
elemol	7	.4771	.72087	1.476	.794	1.428	1.587
valencene	7	.2029	.53671	2.646	.794	7.000	1.587
thujopsenal	7	.4357	.79712	1.738	.794	2.235	1.587
Valid N (listwise)	7						

Table III shows the correlation between high grade and low grade. It is noticed that the correlation value is 0.567, which means high grade and low grade is not correlated. The sig. level (p-value) is 0.241, so, the correlation between high grade and low grade is not significant.

TABLE III.   CORRELATION BETWEEN HIGH GRADE AND LOW GRADE

Correlations

		High grade median	Low grade median
High grade median	Pearson Correlation	1	.567
	Sig. (2-tailed)		.241
	N	6	6
Low grade median	Pearson Correlation	.567	1
	Sig. (2-tailed)	.241	
	N	6	6

Table IV results the correlation among high grade compounds. Among all, hexadecanoic acid and valencene is

highly correlated, 0.923 (almost 1). The high correlate is also for α-guaiene and thujopsenal with the value of 0.880. However, no significant correlation is detected since all sig. levels are more than 0.05. Elemol is found not correlate with any compound since no variable (relative peak area is 0%) is detected. In general, four compounds are highly correlated and they are hexadecanoic acid, valencene, α-guaiene and thujopsenal.

TABLE IV. CORRELATION AMONG HIGH GRADE

Correlations

		dihydro collumellarin_h	hexadecanoic acid_h	alpha guaiene_h	elemol_h	valencene_h	thujopsenal_h
dihydro collumellarin_h	Pearson Correlation	1	.400	-.013	a	.690	-.043
	Sig. (2-tailed)		.600	.987	.	.310	.957
	N	4	4	4	4	4	4
hexadecanoic acid_h	Pearson Correlation	.400	1	.030	a	.923	-.481
	Sig. (2-tailed)	.600		.970	.	.077	.519
	N	4	4	4	4	4	4
alpha guaiene_h	Pearson Correlation	-.013	.030	1	a	.179	-.880
	Sig. (2-tailed)	.987	.970		.	.821	.120
	N	4	4	4	4	4	4
elemol_h	Pearson Correlation	a	a	a	a	a	a
	Sig. (2-tailed)	.	.	.		.	.
	N	4	4	4	4	4	4
valencene_h	Pearson Correlation	.690	.923	.179	a	1	-.525
	Sig. (2-tailed)	.310	.077	.821	.		.475
	N	4	4	4	4	4	4
thujopsenal_h	Pearson Correlation	-.043	-.481	-.880	a	-.525	1
	Sig. (2-tailed)	.957	.519	.120	.	.475	
	N	4	4	4	4	4	4

a. Cannot be computed because at least one of the variables is constant.

Table V shows the correlation among low grade compounds. Among all, hexadecanoic acid is high correlated with dihydro collumellarin (0.956) and valencene is high correlated with alpha guaiene (0.903). Also for valencene and hexadecanoic acid, is high correlated with the value of 0.915. However, no significant correlation is detected since all sig. levels are more than 0.05. It is found that Thujopsenal has no correlation with any other compounds since no variable is detected i.e. relative peak area is 0%. The total compounds for high correlated compounds are hexadecanoic acid, valencene, α-guaiene and dihydro collumellarin

TABLE V. CORRELATION AMONG LOW GRADE

Correlations

		dihydro collumellarin_l	hexadecanoic acid_l	alpha guaiene_l	elemol_l	valencene_l	thujopsenal_l
dihydro collumellarin_l	Pearson Correlation	1	-.956	-.400	.565	.755	a
	Sig. (2-tailed)		.191	.738	.618	.455	.
	N	3	3	3	3	3	3
hexadecanoic acid_l	Pearson Correlation	-.956	1	.652	-.297	-.915	a
	Sig. (2-tailed)	.191		.548	.808	.265	.
	N	3	3	3	3	3	3
alpha guaiene_l	Pearson Correlation	-.400	.652	1	.530	-.903	a
	Sig. (2-tailed)	.738	.548		.644	.283	.
	N	3	3	3	3	3	3
elemol_l	Pearson Correlation	.565	-.297	.530	1	-.115	a
	Sig. (2-tailed)	.618	.808	.644		.927	.
	N	3	3	3	3	3	3
valencene_l	Pearson Correlation	.755	-.915	-.903	-.115	1	a
	Sig. (2-tailed)	.455	.265	.283	.927		.
	N	3	3	3	3	3	3
thujopsenal_l	Pearson Correlation	a	a	a	a	a	a
	Sig. (2-tailed)	.	.	.	.	.	
	N	3	3	3	3	3	3

a. Cannot be computed because at least one of the variables is constant.

Fig. 3 shows the comparison for the highly correlated compounds found in both high grade and low grade. The common compounds are hexadecanoic acid, valencene and α-guaiene. The result demonstrates that there are big different between high grade and low grade in term of compounds' mean in hexadecanoic and valencene.
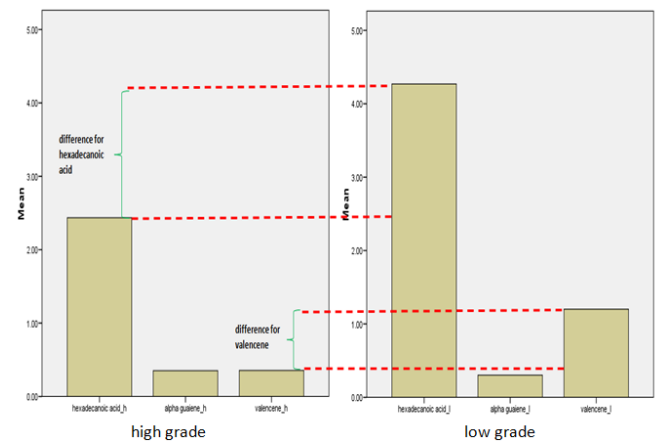
Fig. 3 Comparison for highly correlated compounds in high grade and low grade

IV. CONCLUSION

This paper has presented a correlation analysis on selected *Aquilaria Malaccensis*. The result shows that there are no significant correlation between high grade and low grade. Among compounds in similar grade, there is a high correlation. Three common compounds and highly correlated found in both high grade and low grade. They are hexadecanoic acid, valencene and α-guaiene. The histogram for these compounds marked large differences for high grade and low grade in hexadecanoic acid and valencene.

REFERENCES

[1] N. A. M. Ali, N. Ismail, and M. N. Taib, "Analysis of Agarwood oil (Aquilaria Malaccensis) based on GC-MS data," in *2012 IEEE 8th International Colloquium on Signal Processing and its Applications (CSPA)*, 2012, pp. 470-473.

[2] N. A. M. Ali, A. M. Jamil, J. Mailina, A. Abu Said, S. Saidatul Husni, H. Nor Hasnida, M. N. M. Arip, and Y. S. Chang, "Profiles of selected supreme Agarwood oils from Malaysia," *Herbal Globalisation: A New Paradigm for Malaysian Herbal Industry (MAPS 2008),* pp. 393-398, 2009.

[3] P. Tamuli, P. Boruah, S. C. Nath, and P. Leclercq, "Essential oil of eaglewood tree: a product of pathogenesis," *Journal of Essential Oil Research,* vol. 17, pp. 601-604, 2005.

[4] M. N. I. Bhuiyan, J. Begum, and M. N. H. Bhuiyan, "Analysis of essential oil of eaglewood tree (Aquilaria agallocha Roxb.) by gas chromatography mass spectrometry," *Bangladesh Journal of Pharmacology,* vol. 4, pp. 24-28, 2008.

[5] R. Naef, "The volatile and semi-volatile constituents of agarwood, the infected heartwood of Aquilaria species: a review," *Flavour and Fragrance Journal,* vol. 26, pp. 73-87, 2011.

[6] J. Pornpunyapat, P. Chetpattananondh, and C. Tongurai, "Mathematical modeling for extraction of essential oil from Aquilaria crassna by hydrodistillation and quality of agarwood oil," *Bangladesh Journal of Pharmacology,* vol. 6, pp. 18-24, 2011.

[7] S. N. Tajuddin, "Resolution of complex sesquiterpene mixture in Aqualaria malaccensis (Gaharu) volatile oils using gas chromatography methods," PhD, Faculty of Industry and Science Technology (FIST), Universiti Malaysia Pahang (UMP), 2011.

[8] M. Ishihara, T. Tsuneya, and K. Uneyama, "Guaiane sesquiterpenes from agarwood," *Phytochemistry,* vol. 30, pp. 3343-3347, 1991.

[9] M. Ishihara, T. Tsuneya, M. Shiga, and K. Uneyama, "Three sesquiterpenes from agarwood," *Phytochemistry,* vol. 30, pp. 563-566, 1991.

[10] M. Ishihara, T. Tsuneya, and K. Uneyama, "Fragrant sesquiterpenes from agarwood," *Phytochemistry,* vol. 33, pp. 1147-1155, 1993.

[11] A. Barden, N. A. Anak, T. Mulliken, and M. Song, "Heart of the matter: agarwood use and trade and CITES implementation for Aquilaria malaccensis," *TRAFFIC International, Cambridge, UK,* 2000.

[12] K. H. K. Hamid, *Professorial Lecture: Enhancing The World of Fragrance Through Malaysian Bio-Gaharu,* 1st ed. UiTM Shah Alam: University Publication Centre (UPENA), 2011.

[13] N. A. M. Ali, C. Y.S., M. J., A. S. A., A. M. J., S. H. S., N. H. H., and N. Y. Y., "Comparison of Chemical Profiles of Selected Gaharu Oils from Peninsular Malaysia," *The Malaysian Journal of Analytical Sciences,* vol. 12, pp. 338-340, 2008.

[14] S. N. Tajuddin and M. M. Yusoff, "Chemical Composition of Volatile Oils of Aquilaria malaccensis (Thymeleaceae) from Malaysia," *Natural Product Communications,* vol. 5, pp. 1965-1968, 2010.

[15] P. E. Keller, "Mimicking biology: Applications of cognitive systems to electronic noses," 1999, pp. 447-451.

[16] P. Pripdeevech, W. Khummueng, and S. K. Park, "Identification of Odor-active Components of Agarwood Essential Oils from Thailand by Solid Phase Microextraction-GC/MS and GC-O," *Journal of Essential Oil Research,* vol. 23, pp. 46-53, 2011.

[17] M. Ishihara, T. Tsuneya, and K. Uneyama, "Components of the agarwood smoke on heating," *Journal of Essential Oil Research,* vol. 5, pp. 419-423, 1993.

[18] T. Nakanishi, E. Yamagata, K. Yoneda, and T. Nagashima, "Three fragrant sesquiterpenes of agarwood," *Phytochemistry,* vol. 23, pp. 2066-2067, 1984.

[19] T. Yagura, N. Shibayama, M. Ito, F. Kiuchi, and G. Honda, "Three novel diepoxy tetrahydrochromones from agarwood artificially produced by intentional wounding," *Tetrahedron Letters,* vol. 46, pp. 4395-4398, 2005.

[20] H. F. Lim, A. M. A. Rashid, S. S. Lee, P. A. Fauzi, Y. S. Chang, Y. A. Zuhaidi, M. M. Parid, E. H. Lok, A. M. Farid, and M. A. N. Azah, "Penanaman Aquilaria dan Penghasilan Gaharu di Malaysia," Forest Research Institute Malaysia (FRIM), Kepong, Malaysia 93, 2010.

[21] N. A. M. Ali, Y. S. Chang, M. J., A. S. A, A. M. J., S. H. S., N. H. H., and N. Y. Y., "Comparison of Chemical Profiles of Selected Gaharu Oils from Peninsular Malaysia," *The Malaysian Journal of Analytical Sciences,* vol. 12, p. 340, 2008.

[22] N. C. Schwertman, M. A. Owens, and R. Adnan, "A simple more general boxplot method for identifying outliers," *Computational Statistics & Data Analysis,* vol. 47, pp. 165-174, 2004.

[23] M. Hubert and E. Vandervieren, "An adjusted boxplot for skewed distributions," *Computational Statistics & Data Analysis,* vol. 52, pp. 5186-5201, 2008.

[24] J. Pallant, *SPSS Survival Manual: A step by step guide to data analysis using SPSS*, 4th ed. Australia: Allen and Unwin, 2011.

[25] B. Guner, M. T. Frankford, and J. T. Johnson, "A study of the Shapiro–Wilk test for the detection of pulsed sinusoidal radio frequency interference," *Geoscience and Remote Sensing, IEEE Transactions on,* vol. 47, pp. 1745-1751, 2009.

[26] A. P. Field, *Discovering statistics using SPSS*: SAGE publications Ltd, 2009.

[27] L. Kin Keung, Y. Lean, and W. Shouyang, "Mean-Variance-Skewness-Kurtosis-based Portfolio Optimization," in *First International Multi-Symposiums on Computer and Computational Sciences, 2006. IMSCCS '06. ,* 2006, pp. 292-297.

[28] S. Zani, M. Riani, and A. Corbellini, "Robust bivariate boxplots and multiple outlier detection," *Computational Statistics & Data Analysis,* vol. 28, pp. 257-270, 1998.

Implementation of Real –Time Steam Temperature Control for High Efficient Induction- Based Extraction Essential Oil using Hybrid Fuzzy-PD plus PID Controller

Zuraida Muhammad[#1], Zakiah Mohd Yusoff[#2], Mohd Noor Nasriq Nordin[#3], Mohd Hezri Fazalul Rahiman[#4], Mohd Nasir Taib[#5]

[#]Faculty of Electrical Engineering, UiTM Shah Alam,Selangor, Malaysia
[1]muhammmad_zuraida@yahoo.com
[2]zakiah_yusoff@yahoo.com
[3]nasriq85@yahoo.com
[4]hezrif@ieee.org
[5]dr.nasir@ieee.org

Abstract— **Steam temperature is well known as one significant parameter in process of extraction essential oil that can contribute to increase the output yield and quality of the product. In this study, induction heating system was adopted in steam distillation pot to replace the conventional heating element. Induction heating system was preferred is due to high efficient, clean and save energy. In term of temperature control, this study implemented hybrid fuzzy PD plus PID at steam distillation pot with induction heating system for essential oil extraction system. Initially, the system was modeled to represent the system using ARX structure with PRBS was triggered to the system as an input. The parameter estimation and tuning is derived by simulation for HFPP control scheme. The performance of HFPP with 25 and 49 rules is compared by step response test. From experiment of real time control, demonstrate the proposed HFPP using 49 rules giving a better performance based on settling time, rise time, percent overshoot and RMSE value.**

Keywords— **High efficient inductiom –based extraction essential oil , Hybrid Fuzzy PD and PID.**

I. INTRODUCTION

Steam distillation column is a well known as process that steam produce to enhanced essential oil extraction process[1]. Steam produce in distillation column makes the fraction of hydrocarbons to be distilled at temperature at lower than their boiling point. Steam and the vaporized essential oil will be flow into condenser region and condense become water and pure essential oil. Since the steam continuously produces, the process will be repeated until end of the process.

Nowadays, demand on essential oil increase and countless research efforts have been reported in improving the extraction techniques such as superheated water extraction[2-4] , supercritical fluid extraction[3], supercritical Co2[5], soxhlet[6], subcritical water extraction[7], solid-phase micro extraction[7] and solvent extraction[7].

In term of quality of oil yield produced by steam distillation, several studies made to compare with other extraction methods. According to [8], a quality chemical composition of volatile compound obtained by steam distillation and subcritical water extraction (SWE) were relatively similar. Steam distillation also compared with supercritical fluid extraction (SFE) where both extracts were similar in terms of quality. [3]were determined that steam distillation was the most effective method compare with SFE and SWE. [5]stated that the distillation extraction technique produces more oils as compared with SFE supercritical Co^2 extraction. Those methods were compared and the steam distillation method produced more yields and more economical to compare with others.

Temperature plays an important rule to improve quality and rate of output yield. For that reason, temperature is chosen as a process variable to the proposed research. The prototype development is divided into two categories. They are the steam distillation pot itself and the induction heating system. The distillation pot development covers the setting up of plant with addition of induction heating system. Meanwhile, the induction heating system were be modified to covers the sensing system, software development and system interfacing.

Conventional methods of steam distillation mostly are not control at all. Since the steam distillation are lacking in term of system robustness towards parameters change[9], some controller need to introduce to overcome these shortcoming and improve the performance of steam distillation plant. In this study, HFPP control is proposed. Real time control was introduced to steam distillation plant by send the control signal to induction heating system by right to maintain the steam temperature inside distillation pot at set. From the experiment, controller is expected to execute and giving better

control of steam temperature during distillation process in order to promises the best quality and output yield. In this paper it demonstrated how HFPP control can be implemented at the steam distillation plant to control the steam temperature. Fuzzy logic controller associated with Propotional, Integral and Derivative is most popular among researcher to extend their work[10-17]. For temperature control, demonstrate Fuzzy plus P, I, and D give better performance which improve system response and uncertainties [12, 14]. Fuzzy logic controller has detected produces oscillations in the process response[14]. Implementation of adaptive and hybrid control schemes reduces severe oscillation and the hybrid control scheme enhance control precision. In the design of PID-fuzzy control[11],knowledge obtained from the process reaction curve procedure is employed to determine proper membership function[15, 16].

The rest of this paper was organized as follows. Section II presented steam distillation pot and data collection in section III. Section IV is about modelling strategy, V Hybrid fuzzy PD plus PID, In section VI, real time HFPP implementation and finally, conclusion were included in section VII.

II. STEAM DISTILLATION POT

The steam distillation that was developed in this work is intended for extracting the essential oil by steam distillation technique. Nowadays, there are various extraction techniques implemented ranging from conventional to advanced methods for the purpose of achieving better quality and high quantity of essential oil. Even though steam distillation is considered as a conventional method, there are a few advantages associated with it such economical in terms of operational and maintenance, most oil extraction for botanical materials can be performed via steam distillation, resulting in cleaner production and easy to operate with considerably high production and quality as well

The steam distillation pot that was developed in this work is intended for extracting the essential oil by steam distillation technique. Heating system for steam distillation usually use electrical heater and so on but for distillation pot the induction heating system was introduced. The steam pot actually has a special design using stainless still material to make it working well with induction heating system. Therefore, the pot is equipped with a material bed and attached to a condenser build from copper material and water oil separator. Steam produced in pot works as a transportation medium for enhance essence particles, containing oil when passing through the botanical material. Steam and oil mixture evaporate and condensed flow to condenser then produce a layer of oil and hydrosol (mixture of oil and water) at collecting container. Figure 1 show sketch of steam distillation pot with induction heating system and on figure 2 show the diagram of experiment setup for this system.
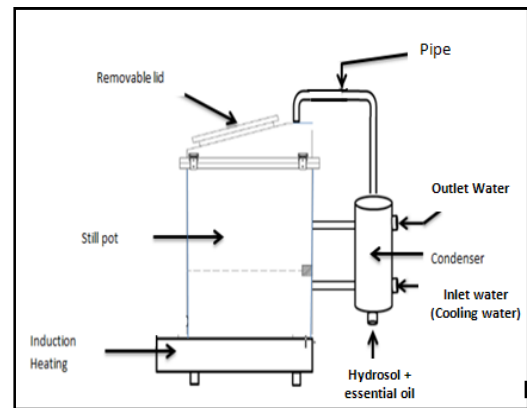

Figure. 1 Distillation pot system prototype sketch

Induction heating system with power range between 800W to 2000W was use in this system to heat the water in pot to produce steam for extraction essential oil purpose. An induction heating system is a powerful, high-frequency electromagnet, with the electromagnetism generated by sophisticated electronics in the "element" under the unit's ceramic surface. When a good-sized piece of magnetic material--such as, for example, a cast-iron skillet--is placed in the magnetic field that the element is generating, the field transfers ("induces") energy into that metal. That transferred energy causes the metal of pot become hot. By controlling the strength of the electromagnetic field, we can control the amount of heat being generated in the pot and we can change that amount instantaneously.Because of all induction heating criteria, it was choose as a new heating element in this study.

Induction heating system will set in range power from 800W to 2kW and the steam temperature will be measured by a resistive temperature detector (RTD PT100) with advantages such high accuracy, low drift, wide operating range and suitable for precision applications where that located upper material bed and conditioned by a signal conditioning circuit (SENECA K109PT) before send the signal to DAQ PC LabCard Advantech PCI-1716 through port AI1. Data gathered from plant will be stored at PC via Matlab programming for analysis purpose. By using MatLab programming, sampling time was set one data per second. The dynamic is as anticipated since the system is not equipped with any cooling system**.**
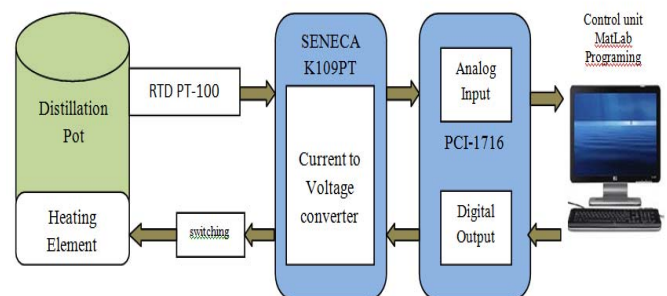

Figure. 2: Block diagram of steam distillation pot with induction heating system, computer and interface

III. DATA COLLECTION

Data of steam temperature from steam distillation pot were carried out and transfer to DAQ PC LabCard Advantech PCI-1716 to store in Matlab software. Using Matlab version 2009 make the process of data collected become easier when it provide a sufficient tools that use for data acquisition purpose. The data stored in Matlab programming will display in workspace and use for analysis purpose. In this study, the steam temperature data was collected with pseudorandom binary sequence (PRBS) as the input signal. During experiment, the input signal PRBS was generated using matlab software with 0.2 proportional band. The PRBS was set 1, high and 0, low correspond 1200W and 0W of induction heating system. The measured output from system must be informative to ensure the model present is accurate. Based on the PRBS triggering, the output steam response of steam temperature output is visualized in figure 3.

The collected data 5000 (input-output samples) plant steam distillation pot are divided into two parts for estimation and validation by using interlacing technique where odd data for estimation and even data for validation.
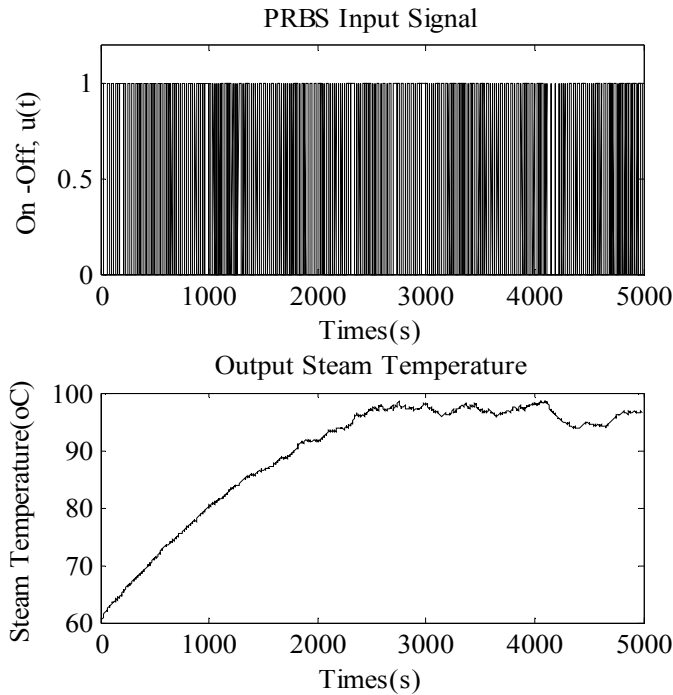

Figure. 3: PRBS input and measured temperature

IV. MODELING STRATEGY

A. Model structure selection ARX

The linear system identification is obtained by use the data collected from the experiment that comprises s of input, output and disturbance.

ARX model structure can be written as:

$$A(q)y(t) = B(q)u(t) + e(t) \quad (1)$$

Where the polynomials A(q) and B(q) defined by:

$$A(q) = 1 + a_1 q^{-1} + \ldots \ldots + a_{n_a} q^{-n_a}$$
$$B(q) = b_0 + b_1 q^{-1} + \ldots \ldots + b_{n_b} q^{-n_b} \quad (2)$$

Where A(q) and B(q) represent system dynamic and q^{-1} as time shift operator and this q description in equivalent to the Z-transform form[18]

V. HYBRID FUZZY − PD PLUS PID CONTROLLER

The hybrid fuzzy-PD plus PID controller is a combination of fuzzy-PD controller with PID control structure to improvised the performance of the system. Without the present of control scheme, the steam temperature response is originally very dynamic and uncertain therefore smoother output is required for temperature regulation. In hybrid fuzzy PD plus PID, there are eight gain parameters consists as shown in figure 4 that is Kp1, Kd1, Gf1, Gf2, and Gf3 for fuzzy-PD gains, whereas, Ge, Ki2 and Kd2 are PID gains. All parameter should be tuned to optimize the performance of controller before attempting real time application. To realize this process, simulation work should be done first in order to guarantee the system stability and get the appropriate value for each gain for HFPP.
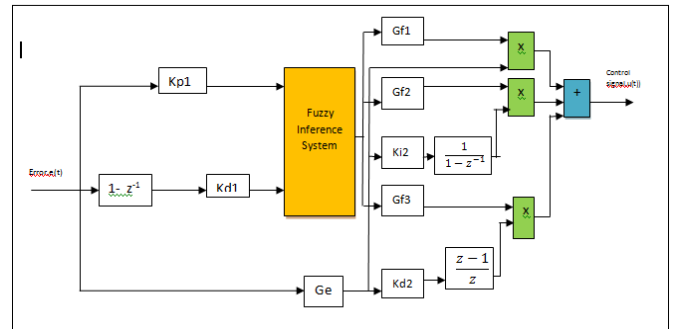

Figure 4 : Structure of hybrid Fuzzy-PD plus PID(HFPP) controller

The block diagram of HFPP as a control module was illustrate in figure 5 which gives a control signal to induction heating system. For robustness check, the system was introduced with load disturbance and set point tracking.
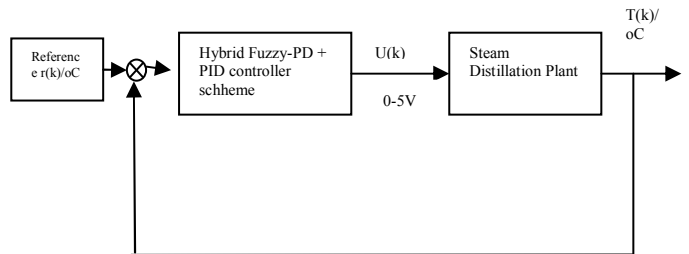

Figure 5: Diagram of hybrid fuzzy-PD plus PID controller

The algorithm of HFPP is stated as in equation (3) and (4). The control signal, u is calculated based on summation of all products from fuzzy components and PID components. Whereas n signal is derived from fuzzy operation, $F \{\}$, that consist of two inputs from error and differential error as in (4). The advantage of HFPP is that the output from fuzzy, n can have better fine tune of desired output as the presence of PID. If error occurred, the double controller (Fuzzy-PD and PID) is forced the response to minimize the error signal. If error is zero the controller setting will maintained at present condition.

$$n = F\{kp_1|kd_1e\} \quad\quad (3)$$

$$u = ((Gf_1.n)x\,(Ge.e)) + \left((Gf_2.n)x\left[Ge.Ki_2\int_0^r e(\sigma)d(\sigma)\right] \right) + ((Gf_3.n)x\,(Ge.Kd_2.e(\tau))) \quad (4)$$

In control of steam temperature, fuzzy with 49 rules was employed by constructed from 7 conditions of membership function for input error and differential error. In this paper, fuzzy with 25 rules also executed to compare output steam response with fuzzy 49 rules. Figure 6, 7, and 8 shows input error controller, input change of error and output membership function respectively.
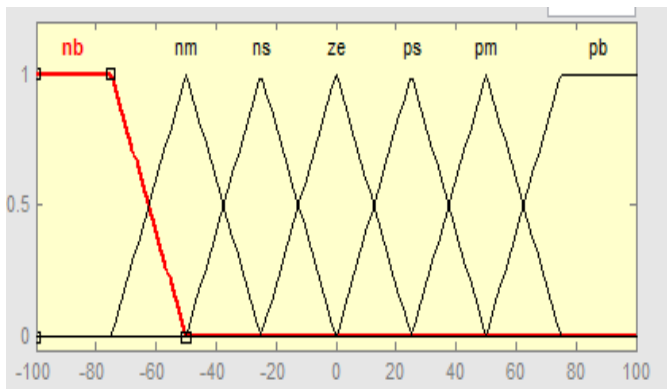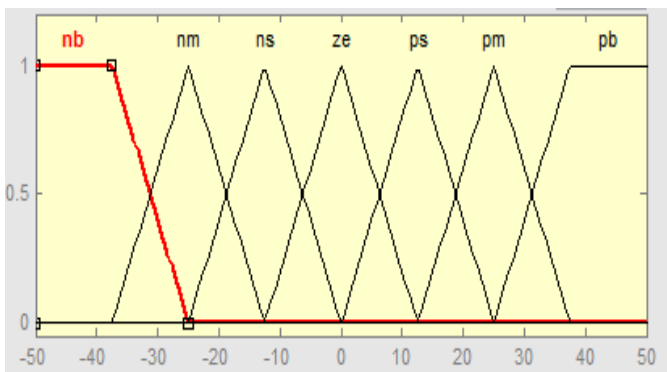

Figure 6: Input error controller membership function

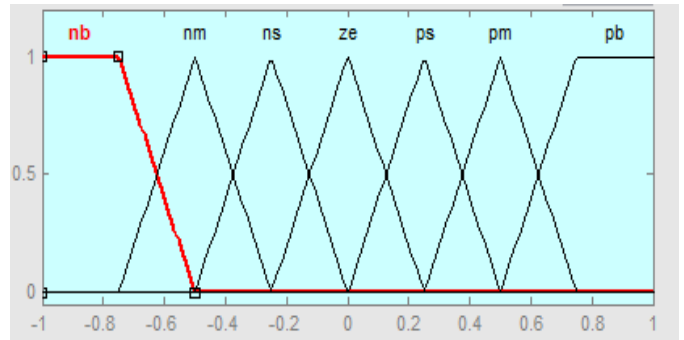
Figure 7: Input change of Error membership function

Figure 8: Output membership function

VI. REAL TIME HYBRID FUZZY PD PLUS PID IMPLEMENTATION

Some tuning can be applied by adjusting fuzzy rules, membership function and gain parameters. However, according to real time experience, adjusting the membership function and change rule not really gives much affect to the response if compared by adjusting the gain parameters. The heating process well known is a slow process[14]. Since the slow process, output response is supposedly to execute with no overshoot, rise time was reduced with minimizing steady state error[9]. The fluctuation at transient stage need to eliminate because during higher temperature(above boiling point) the response is highly dynamic and uncertain behaviour. Table 1 was tabulated of tuned parameter gains for HFPP.

TABLE 1 HFPP controller parameters values

HFPP Parameters	Gains
Kp1	1.0
Kd1	-1.0
Ge	6
Ki2	0.2
Kd2	2.0
Gf1	0.1
Gf2	0.1
Gf3	1.0

The design of HFPP controller for plant high efficient induction- based essential oil extraction system is to provide smooth control for steam temperature also robust to changes of operation condition during extraction process. The selection of fuzzy rule is very important in order to achieve smoother response and less oscillation at the steady state. In this study, fuzzy rules with 49 rules give a better performance compared to fuzzy 25 rules. Figure 9 and figure 10 shows the steam temperature response using HFPP with 25 and 49 rules. The performance of controller was measured based on rise time, settling time, percent of overshoot and RMSE value. Table 2 was tabulated the summary of steam temperature response for HFPP 25 and 49 respectively. Figure 11 visualize the

comparison between output steam response with fuzy 25 and 49 rules.

Figure 9: Output temperature steam response using HFPP with 25 rules
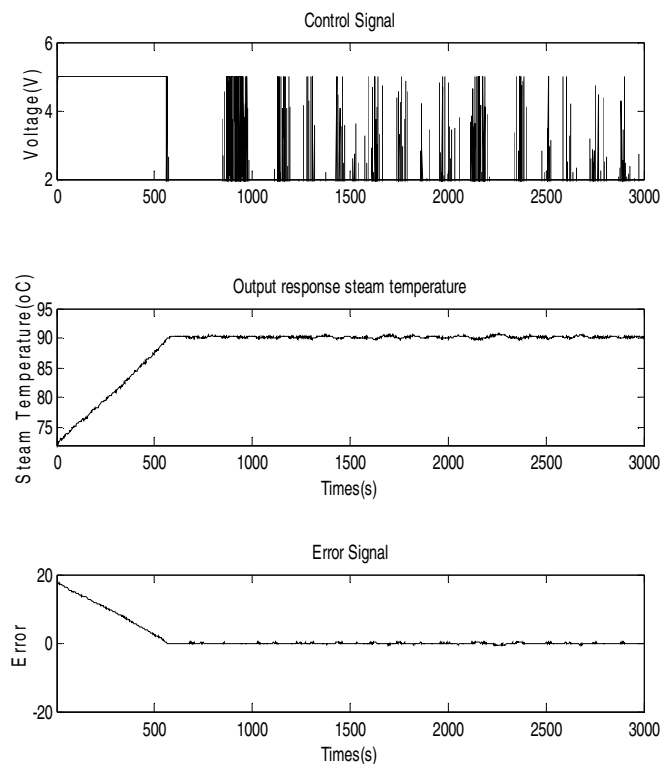
Figure 10: Output temperature steam response using HFPP with 49 rules
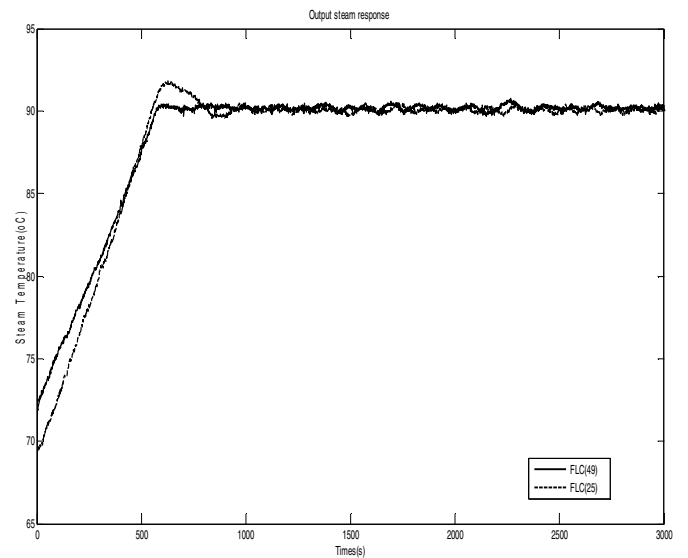
Fig 11 Output Steam temperature response for HFPP using 25 and 49 fuzzy rules

Table 2 Summary of performance HFPP with fuzzy 25 and 49 rules

	Rise time(s)	Settling Time(s)	% Overshoot	RMSE value
HFPP 25 rules	381	2986	2.1142	0.4176
HFPP 49 rules	318.76	536.2	0.4118	0.23

VIII CONCLUSION

In this work, the proposed real time HFPP controller has been developing for control the steam temperature of high efficient induction- based extraction essential oil system. The aim of this study to implement HFPP in real time control was achieve when HFPP with fuzzy 49 rules gives a good performance in control the steam temperature at set point with small overshoot, fast rise time and shorter settling time. It proves that HFPP was capable in order to minimize the output error and eliminate the fluctuation of temperature response at transient state during higher temperature where the response is highly dynamic and uncertain behavior.

ACKNOWLEDGMENT

VIIII REFERENCE

[1]   M. T. Vafaei, R. Eslamloueyan, and S. Ayatollahi, "Simulation of steam distillation process using neural networks," *Chemical Engineering Research and Design,* vol. 87, pp. 997 - 1002, 2009.

[2] J. M. Roldan-Gutierrez, J. Ruiz-Jimenez, and M. D. L. d. Castro, "Ultrasound-assisted dynamic extraction of valuable compounds from aromatic plants and flowers as compared with steam distillation and superheated liquid extract," *Talanta,* 2008.

[3] A. Ammann, D. C. Hinz, R. S. Addleman, C. M. Wai, and B. W. Wenclawiak, "Superheated water extraction, steam distillation and SFE of peppermint oil," *Fresenius Journal Chem,* vol. 364, pp. 650-653, 1999.

[4] M. Z. Ozel and H. kaymaz, "Superheated water extraction, steam distillation and Soxhlet extraction of essential oils of origanum onites," *Anal Bional Chem,* vol. 379, pp. 1127-1133, 2004.

[5] H. Ebrahimzadeh, Y. Yamini, F. Sefidkon, M. Chaloosi, and S. M. Pourmortazavi, "Chemical composition of the essential oil and supercritical CO2 extracts of Zataria multiflora Boiss," *Food Chemistry,* vol. 83, pp. 357-361, 2003.

[6] P. Richter, B. Sepulveda, K. C. R. Oliva, and R. Seguel, "Screening and determination of pesticides in soil using continuous subcritical water Chromatography-mass spectrometry," *Journal of Chromatography A,* pp. 169-177, 2003.

[7] G. Song, D. W. C. Deng, and Y. Hu, "Comparison of Headspace Solid-Phase Microextraction with solvent Extraction for the analysis of the Volatile Constituents of Leaf Twigs of chinese Arborvitae," *Chromatographia,* pp. 769-774, 2003.

[8] M. Z. Ozel, F. Gogus, J. F. Hamilton, and A. C. Lewis, "Analysis of volatile components from Ziziphora taurica subsp. Taurica by steam distillation, superheated-water extraction, and direct thermal desorption with GCxGCTOFMS," *Anal Bioanal Chem,* vol. 382, pp. 115-119, 2005.

[9] N. Kasuan, Z. Yusuf, M. N. Taib, M. H. F. Rahiman, N. Tajuddin, and M. A. A. Aziz, "Robust Steam Temperature Regulation for Distillation of Essential Oil Extraction Process using Hybrid Fuzzy-PD plus PID Controller," *World Academy of Science, Engineering and Technology,* vol. 71, 2010.

[10] K. K. Ahn and D. Q. Truong, "Online tuning fuzzy PID controller using robust extended Kalman filter," *Journal of Process Control,* vol. 19, pp. 1011-1023, 2009.

[11] A. J. B. Antunes, J. A. F. R. Pereira, and A. M. F. Fileti, "Fuzzy control of a PMMA batch reactor: Development and experimental testing," *Computers & Chemical Engineering,* vol. 30, pp. 268-276, 2005.

[12] Hisbullah, M. A. Hussain, and K. B. Ramachandran, "Design of a Fuzzy Logic Controller for Regulating Substrate Feed to Fed-Batch Fermentation," *food and bioproducts processing,* vol. 81, pp. 138-146, 2003.

[13] C.-H. Jung, C.-S. Ham, and K.-I. Lee, "A real-time self-tuning fuzzy controller through scaling factor adjustment for the steam generator of NPP," *Fuzzy Sets and Systems,* vol. 74, pp. 53-60, 1995.

[14] N. Threesinghawong, P. Sanposh, N. Chinthaned, and P. Srinophakun, "Temperature control of Koji process by using fuzzy PD controller," in *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008. ECTI-CON 2008. 5th International Conference on,* 2008, pp. 669-672.

[15] A. Visioli, "Tuning of PID controllers with fuzzy logic," *Control Theory and Applications, IEE Proceedings -,* vol. 148, pp. 1-8, 2001.

[16] Z.-W. Woo, H.-Y. Chung, and J.-J. Lin, "A PID type fuzzy controller with self-tuning scaling factors," *Fuzzy Sets and Systems,* vol. 115, pp. 321-326, 2000.

[17] Z. Yusuf, Z. Janin, and M. Nasir Taib, "Application of Fuzzy PD controller with Smith Predictor for glycerin bleaching temperature control," in *Intelligent and Advanced Systems (ICIAS), 2010 International Conference on,* pp. 1-5.

[18] L.Ljung, "System identification.Theory for the user," *Prentice hall,* vol. Englewood cliffs, 1987.

Real Time PID Control for Hydro-diffusion Steam Distillation Essential Oil Extraction System Using Gradient Descent Tuning Method

Zakiah Mohd Yusoff[#1], Zuraida Muhammad[#2], Mohd Noor Nasriq Nordin[#3], Mohd Hezri Fazalul Rahiman[#4], Mohd Nasir Taib[#5]

#Faculty of Electrical Engineering, UiTM Shah Alam,Selangor, Malaysia
[1]zakiah_yusoff@yahoo.com
[2]muhammmad_zuraida@yahoo.com
[3]nasriq85@yahoo.com
[4]hezrif@ieee.org
[5]dr.nasir@ieee.org

Abstract—**This paper present the implementation of real time PID controller for hydro-diffusion steam distillation essential oil extraction system based on comparison of Gradient Descent (GD) and Ziegler Nichols (ZN) tuning methods. The first order of Auto Regressive Exogenous (ARX) model was used to describe the behavior of the temperature system and will be use in the controller design. A PID controller is expected to execute a robust response towards parameters changes and better control of steam temperature during distillation process. The system had been evaluated based on rise time, % overshoot, settling time and root mean square error (RMSE). The temperature control was achieved by controlling the voltage fed to the heater ranging from 0V to 5V via digital-to-analogue converter (DAC). Robustness test of the PID controller are based on: i) introduce disturbance and ii) set-point tracking. From the result, the performance of PID controller using GD tuning method reveals that this controller can be adapt for the system because being sensitive to parameters changes and robust as the response can compensate the load disturbance and set point change.**

Keywords— **Gradient descent, Ziegler Nichols, ARX model, hydro-diffusion, distillation process, set-point tracking.**

I. Introduction

Essential oils (EO) can be defined as the volatile aromatic compound by extraction process which can be found mostly from plant's flowers, bark, wood, leaves, root, seed or resin [1, 2] and it stored in the pockets of the botanical material which can be extracted by either breaking, heating or stirring the pockets in boiling water[3]. The usage of essential oil covers in aromatherapy, food and beverages, perfume, antibacterial activity, and spa practices [2, 4, 5]. Generally, an analytical procedure for essentials oils can be categories into two steps: extraction process (such as steam distillation, hydro-distillation, simultaneous distillation-extraction and solvent extraction), and analysis process gas chromatography (GC), gas chromatography coupled to mass spectrometry (GC-MS) [1, 4-6].

Steam distillation is the earliest and popular extraction technique for most botanical materials[7] compared to others techniques such as solvent extraction, expression and critical fluid extraction[1, 6]. The proportion of essential oils extracted by steam distillation is 93% compared to the remaining only 7% of essential oil extracted[1]. Steam distillation usually applied on the fresh and dried material[8]. Based on the review above, the author already developed new system that uses steam which is hydro-diffusion steam distillation system.

There are several factors affect the extraction yield such as temperature, pressure, distillation time, chemical composition and particle size [9]. This paper concentrating on steam temperature as a controlled variable since it is one of the significant factors in the extraction process. Controlling and maintaining the temperature at the desired is a bit challenging tasks due to various factors such as slow dynamic response and process have lag or time delay[10, 11]. In order to make sure the quality of essential oil, a good temperature controller is needed. The PID controller have been well developed for industrial and process control today [11, 12, 13] due to simplicity, ease of design, low cost and effectiveness for most linear systems [12]. The application of PID controller in hydro-diffusion steam distillation essential oil extraction system is expected to execute a robust response towards parameters changes and better control of steam temperature during distillation process.

In later part of the paper is organized as follows: In section 2, hardware set up and system configuration is presented. Section 3 describes the plant's modeling and followed by PID controller design in section 4. The results of experimental evaluation of the proposed technique on the hydro-diffusion steam distillation are presented in Section 5, which is followed by the concluding remarks.

II. METHODOLOGY

i. *Hydro-diffusion steam distillation essential oil extraction system*

Hydro-diffusion steam distillation essential oil extraction system is visualized as in Figure 1 and installed at Distributed Control System Laboratory (DCS) in University Teknologi Mara, Malaysia. The distillation column with 10 litre water is heated by immersion heater to produce steam. Pressurized steam is supply from the top of the plant and passed through the botanical materials. Pressurized steam is used to vaporize the volatile oils in the plant material. The steam and oil vapor mixture is then going down through the small tube to the condenser. The implementation of heating process in pilot plant of essential oil extraction system is described in Figure 1.
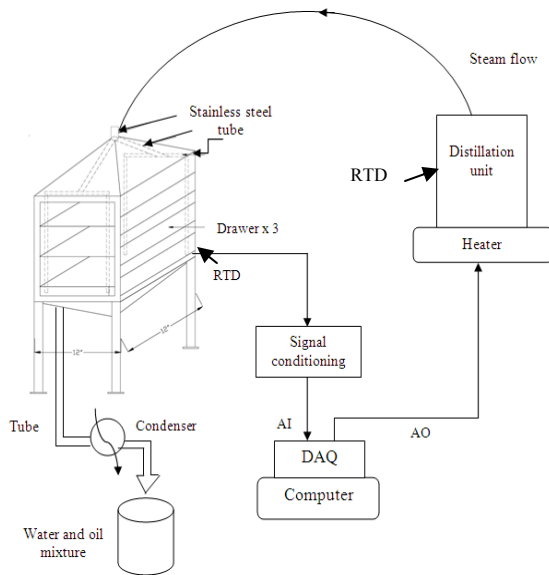

Figure 1 The hydro-diffusion steam distillation essential oil extraction system

ii. *System Configuration*

The system is divided into 3 main parts which is process plant, data acquisition using PCI-1711 card and control unit as in Figure 2 below.
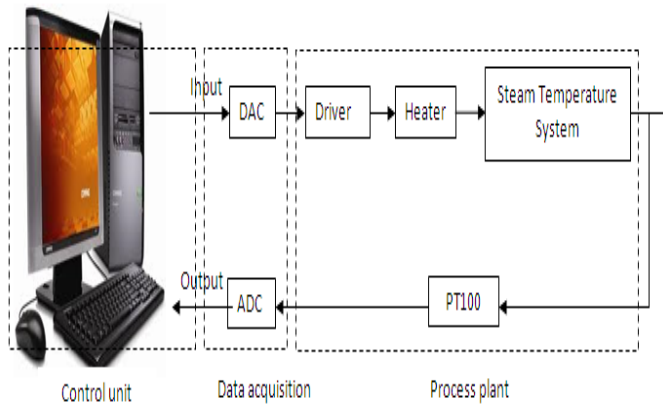
Figure 2 Block diagram of system configuration

The system consists of three phase 240 Vac immersion type heater with a power rating of 1.5kW. The sensor is a platinum sensing element, PT-100 3 wires type placed in the material tray. PT100 connected with the signal conditioning circuit varies in terms of resistance value of 100ohm at 0 ^{o}C. The temperature control was achieved by controlling the voltage fed to the heater ranging from 0V to 5V via digital-to-analogue converter (DAC). Computer was used as a control unit. PCI-1711 Advantech card was used to interface between hardware and control unit. Software for the system was developing using MATLAB programming to monitor the signal response and for real-time implementation, MATLAB Real Time Workshop (RTW) was employed.

III. PLANT'S MODELING

Modeling approach basically starts with the design of experiment to obtain input and output data, followed by model structure selection, then model estimation and lastly is model validation. The Auto Regressive Exogenous (ARX) model was used to describe the behavior of the global model based on the relationship between input and output data. The system is perturbed with Pseudo Random Binary Sequence (PRBS) as an input and temperature as an output signal. The PRBS amplitude is selected to adhere at maximum of 100% from the full scale (5V) when PRBS signal is 1 and 50% of full scale (2.5V) when the PRBS signal at 0. The output of steam temperature response and PRBS input as in Figure 3.

(a)

(b)

Figure 3 Input-output data of heating process, (a) output, (b) input.

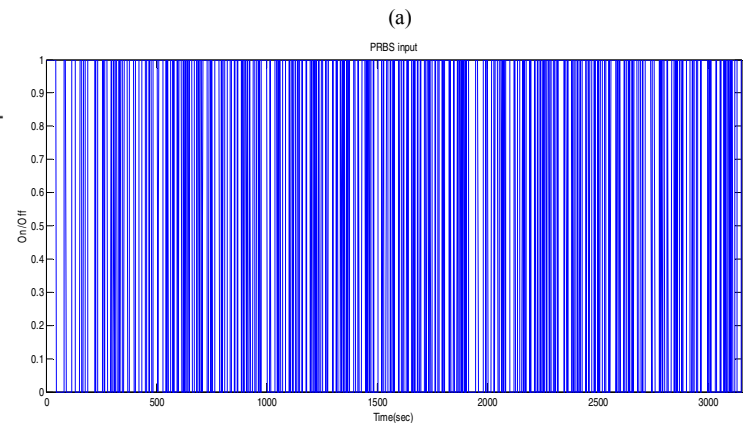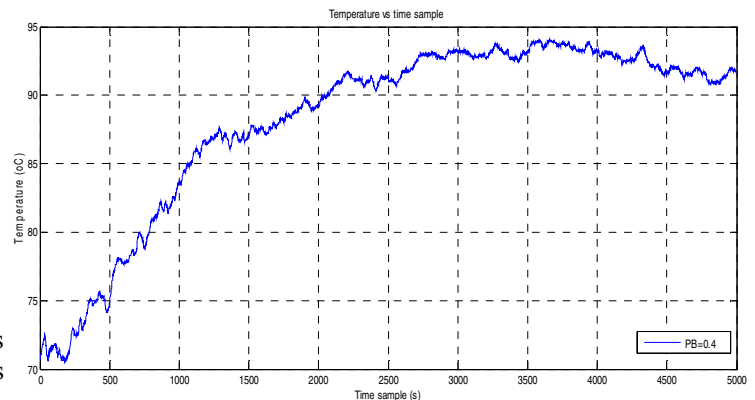There are 5000 input-output data collected in this experiment. The PRBS signal is generated with PB=0.4 and clock period factor Nb=2.5 (Nb=1/PB). All data have been sampled at Ts=1. The initial temperature at steam tray is 70.8 ^{o}C and was gradually increased up to 94.1 ^{o}C.

i. ARX Modeling

The fundamental of ARX model structure is given by equation (1) below [14]

$$y(t) = q^{-nk}\frac{B(q)}{A(q)}u(t) + \frac{1}{A(q)}e(t) \tag{1}$$

where the polynomials $A(q)$ and $B(q)$ are given by

$$A(q) = 1 + a_1 q^{-1} + \cdots + a_{na} q^{-na}$$
$$B(q) = b_o + b_1 q^{-1} + \cdots + b_{nb} q^{-nb} \tag{2}$$

where q^{-1} is the delay operator and $e(t)$ represent the white noise.

The input-output data are divided equally into 2 sets: odd number samples data (1, 3, 5...) are used for estimation, the even number samples (2, 4, 6...) are used for validation purposed by using interlacing technique. The modeling is done by using model order 1. The obtained ARX model was employed as in equation (3)

$$A(q) = 1 - q^{-1}$$
$$B(q) = 0.0082\ q^{-1} \tag{3}$$

In represent the system, good model should be representative in all condition of the system. The validation on ARX model shows the performance criteria using fitness, R^2 is 98.41%. The performance criteria based on RMSE and mean square error (MSE) are 0.1037 and 0.0108 respectively. This model accepted to perform further for controller design.

IV. PID CONTROLLER DESIGN

Figure 4 illustrates the block diagram of PID module for controlling steam temperature for hydro-diffusion steam distillation plant.
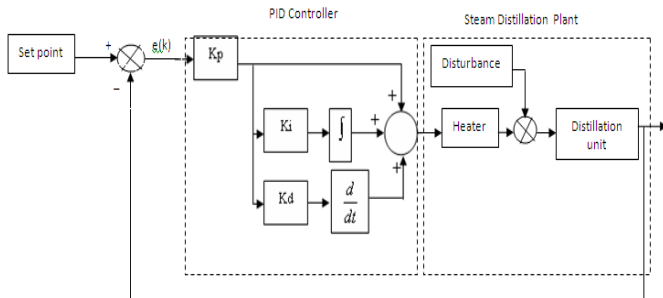

Figure 4 Block diagram of PID controller

The PID controller is give in equation (4) where Kp, Ki, and Kd is the coefficients for proportional (P), integral (I) and derivative (D) respectively whereas e (t) is the error.

$$Go = K_p e(t) + K_i \int e(t)\,dt + K_d \frac{de(t)}{dt} \tag{4}$$

There are three gain parameters consists in PID controller. In this study, GD and ZN tuning methods were applied to tune these parameters. The Ziegler-Nichols tuning method is well known and often forms the basis for tuning procedures used by control system vendors. In the ZN approach, simple formulas for PID controller settings are expressed in terms of ultimate gain K_u and ultimate period T_u of the process. The disadvantages of experimentally determining the critical parameters are that the system can be brought to a state of instability [15, 16]. To overcome these problems, gradient descent optimization is applied to tune these parameters and was done by simulation. Gradient descent is one of the optimization techniques to tune important parameter simultaneously. The method improves dynamic and steady state response and maintains the frequency at desired level [17, 18].Gradient descent tuning algorithm requires no knowledge of the plant to be controlled. This makes the algorithm robust to changes in the plant. It also makes the algorithm universally applicable to linear and nonlinear plants, with or without noise, with or without time delay. The algorithm achieves the tuning objective by minimizing an error function [19].

The PID parameters are tabulated in Table 1 for GD and Table II for ZN tuning method .The tuned parameters of PID controller that was done by simulation is tested on the real system.

TABLE 1

PID PARAMETERS USING GD

Parameter	
K_p	17.22
K_i	0.0000288
K_d	0.0001583

TABLE II

PID PARAMETERS USING ZN

Parameter	
K_p	26.8
K_i	0.0002057
K_d	16.575

V. RESULT AND DISCUSSION

i. Real Time Implementation

In this paper, PID controller is designed to control the steam temperature at the desired temperature.
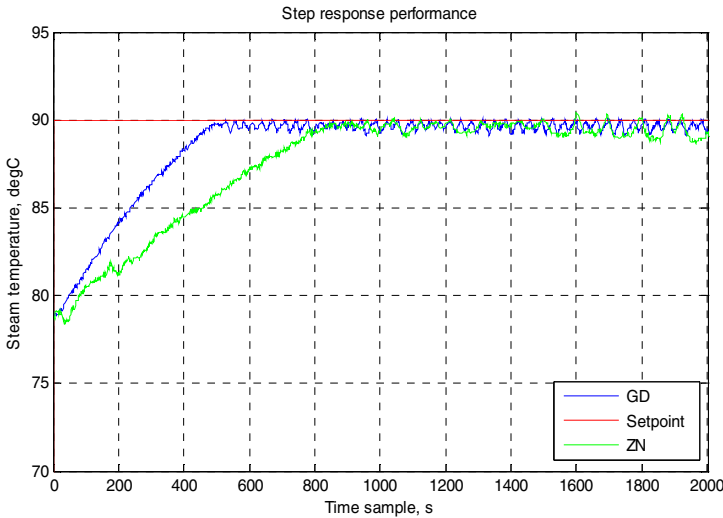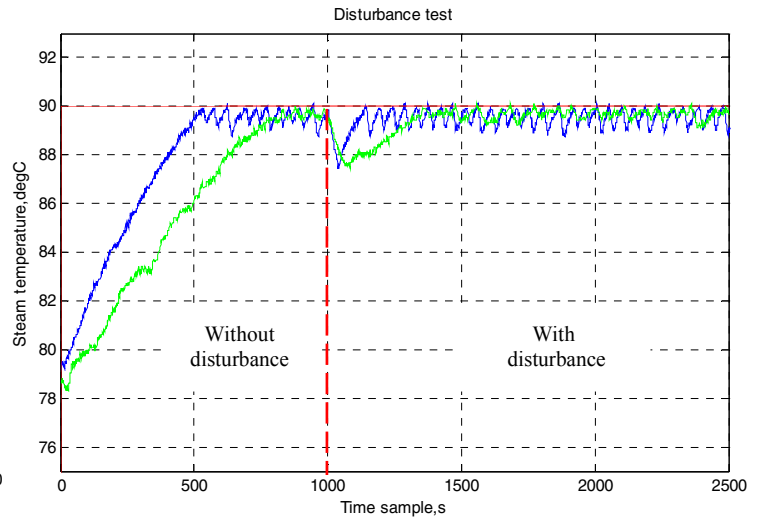
Figure 5 PID controllers step response

TABLE III

STEP RESPONSE PERFORMANCE

PID CONTROLLER	Rise time (s)	%OS	Settling time (s)	RMSE
GD	345	0.6700	1993	0.4239
ZN	658	0.9800	2667	0.6707

Figure 5 and Table III shows the step response performance for real time PID controller using GD and ZN tuning method. In term of the rise time, GD gave faster response with 345 s compared with the ZN with 658 s. There is different approximately 0.31% deviation for overshoot between GD=0.67% and ZN= 0.98%. For a settling time criteria, GD reach the faster settling time with 1993 s compared with ZN= 2667 s. Meanwhile, the roots mean squared error (RMSE) showed that PID controller using GD tuning method outperformed PID controller using ZN with 0.4239 and 0.6707 respectively.

Robustness test of the PID controller are based on: i) introduce disturbance and ii) set-point tracking. The objective of the introduce disturbance is to test ability of the PID controller in comprising an abrupt disturbance during running process. The disturbance is cut off the power supply from 1.5kW to 0W for 30 seconds and start at 1000 samples.

Figure 6 PID controllers response with disturbance

TABLE IV

DISTUREBANCE TEST PERFORMANCE

PID CONTROLLER	GD	ZN
T_{min}	87.62 $^{\circ}$C	87.42 $^{\circ}$C
Time taken to go back to the setpoint, (s)	151	386

From the result in Figure 6, after introduce the disturbance at sample 1000s, steam temperature decrease with T_{min}= 87.62°C and T_{min}= 87.42°C for GD and ZN controller respectively. However, time taken to go back to the set point after introduce disturbance for GD controller is two times faster which is only 151s compared with ZN controller = 386s. Significant improvement can be spotted to GD controller because provide fast and robust response in comprising an abrupt disturbance during running process.

The PID controller also assessed based on the set point tracking where the actual response is expected to follow the reference signal, r(k):

$$r(k) = \begin{cases} 80 \ ^{\circ}C; \ for \ 0 <k< 700 \\ 85 \ ^{\circ}C; \ for \ 701 <k< 1200 \\ 90 \ ^{\circ}C; \ for \ 1201 <k< 1700 \end{cases}$$
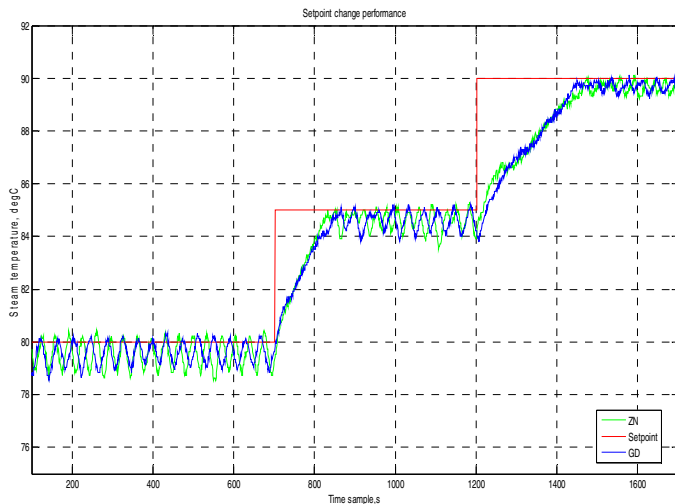
Figure 7 PID controllers with set-point tracking

Figure 7 illustrates the comparison between GD and ZN performance in controlling the system using step change input. As can be seen, at higher temperature level, both controllers performed well in tracking the stipulated set point. Regardless of lower temperature, GD still showed a consistent and better control performance compared to ZN.

VI. CONCLUSION

This work demonstrates the implementation of real time PID controller using GD and ZN tuning method for hydro-diffusion steam distillation essential oil extraction system. The performance of PID controller using GD tuning method reveals that this controller can be adapt for the system because being sensitive to parameters changes and robust as the response can compensate the load disturbance and set point change .

ACKNOWLEDGEMENT

REFERENCES

[1] H. Ebrahimzadeh, Y. Yamini, F. Sefidkon, M. Chaloosi, and S. M. Pourmortazavi, "Chemical composition of the essential oil and supercritical CO2 extracts of *Zataria multiflora* Boiss," *Food Chemistry,* vol. 83, pp. 357–361, 2003.

[2] J. M. Roldan-Gutierrez, J. Ruiz-Jimenez, and M. D. L. d. Castro, "Ultrasound-assisted dynamic extraction of valuable compounds from aromatic plants and flowers as compared with steam distillation and superheated liquid extraction," *Talanta,* vol. 75, pp. 1369–1375, 2008.

[3] F. Chemat, M. E. Lucchesi, J. Smadja, L. Favretto, G. Colnaghi, and F. Visinoni, "Microwave accelerated steam distillation of essential oil from lavender: A rapid, clean and environmentally friendly approach," *Analytica Chimica Acta,* vol. 555, pp. 157–160, 2006.

[4] E. Cassel, R. M. F. Vargas, N. Martinez, D. Lorenzo, and E. Dellacassa, "Steam distillation modeling for essential oil extraction process," *Industrial Crops and Products,* vol. 29, pp. 171–176, 2009.

[5] N. Kasuan, M. Yunus, M. H. F. Rahiman, S. R. S.Aris, and M. N. Taib, "Essential Oil Composition of Kaffir Lime: Comparative analysis between Controlled Steam Distillation and Hydrodistillation Extraction Process," *Proceedings of 2009 Student Conference on Research and Development (SCOReD 2009),* 2009.

[6] F. Peng, L. Sheng, B. Liu, H. Tong, and S. Liu, "Comparison of different extraction methods: steam distillation, simultaneous distillation and extraction and headspace co-distillation, used for the analysis of the volatile components in aged flue-cured tobacco leaves," *Journal of Chromatography A,* vol. 1040, pp. 1–17, 2004.

[7] P. Masango, "Cleaner production of essential oils by steam distillation," *Journal of Cleaner Production,* vol. 13, pp. 833-839, 2005.

[8] M. Rouatbi, A. Duquenoy, and P. Giampaoli, "Extraction of the essential oil of thyme and black pepper by superheated steam," *Journal of Food Engineering,* vol. 78, pp. 708–714, 2007.

[9] N. Kasuan, Z. Yusuf, M. N. Taib, M. H. F. Rahiman, N. Tajuddin, and M. A. A. Aziz, "Robust Steam Temperature Regulation for Distillation of Essential Oil Extraction Process using Hybrid Fuzzy-PD plus PID Controller," *World Academy of Science, Engineering and Technology* vol. 71, pp. 932-937, 2010.

[10] Z. Yusuf, Z. Janin, and M. N. Taib, "Application of Fuzzy Logic Controller for Glycerin Bleaching Process," *Proceedings of 2009 IEEE Student Conference on Research and Development (SCOReD 2009)*, pp. 438-441, 2009.

[11] N. Tajjudin, Z. Yusuf, M. H. F. Rahiman, and M. N. Taib, "Self tuning PID Controller for Steam Distillation Essential Oil Extraction System," *6th International Colloquium on Signal Processing & Its Applications (CSPA),* pp. 259-262, 2010.

[12] K. S. Tang, K. F. Man, G. Chen, and S. Kwang, "An Optimal Fuzzy PID Controller," IEEE Transactions On Industrial Electronics, vol. 48, pp. 757-765, 2001.

[13] Z. Yusuf, Z. Janin, and M. N. Taib, "Self Tuning Fuzzy Logic Controller with Smith Structure for Controlling Glycerin Bleaching Temperature," 2010 IEEE Control and System Graduate Research Colloquium, pp. 39-43, 2010.

[14] N. Ismail, N. Tajjudin, M. H. F. Rahiman, and M. N. Taib, "Modeling of Dynamic Response of

Essential Oil Extraction Process," *2009 5th International Colloquium on Signal Processing & Its Applications (CSPA),* pp. 298-301, 2009.

[15] M. Tajjudin, N. Ishak, H. Ismail, M. H. F. Rahiman, and R. Adnan, "Optimized PID Control using Nelder-Mead Method for Electro-hydraulic Actuator Systems," *2011 IEEE Control and System Graduate Research Colloquium*, pp. 90-93, 2011.

[16] R. Adnan, M. Tajjudin, N. Ishak, H. Ismail, and M. H. F. Rahiman, "Self-tuning Fuzzy PID Controller for Electro-Hydraulic Cylinder," *IEEE 7th International Colloquium on Signal Processing and its Applications*, pp. 395-398, 2011.

[17] X. Dong and D.-X. Zhou, "Learning gradients by a gradient descent algorithm," *Journal of Mathematical Analysis and Applications,* vol. 341, pp. 1018-1027, 2008.

[18] G. Mallesham, S. Mishra, and A. N. Jha, "Optimization of Control Parameters in AGC of Microgrid using Gradient Descent Method " *16th National Power Systems Conference,* pp. 37-42, 2010.

[19] F. Lin, R. D.Brandt, and G. Saikalis, "Self-Tuning of PID controller by adapitve interaction," *Proceedings of America Control Conference,* pp. 3676-3681, 2010.

.

Gait Recognition based on Lower Limb

Nurul Illiani Yaacob & Nooritawati Md Tahir
Faculty of Electrical Engineering
Universiti Teknologi MARA (UiTM)
40450, Shah Alam, Selangor DE, MALAYSIA
Corresponding author:norita_tahir@yahooo.com

Ramli Abdullah
Centre for Research and Innovation Management
Universiti Kebangsaan Malaysia (UKM)
Bangi, Selangor DE, MALAYSIA

Abstract - In this paper, a new approach is proposed for extracting feature from walking human based on the lower body. This approach utilized DCT to extract features followed by PCA as feature selection. Further ANN and PNN are utilised as classifiers. All developed method is tested with 10 subjects from CASIA gait database. Initial findings showed that the proposed method is viable based on 90% recognition rate for ANN and 80% for PNN.

Keywords – gait recognition, averaged silhouette, discrete cosine transform, principal component analysis, artificial neural network, probabilistic neural network

I.   INTRODUCTION

Currently, person verification based on 'the way they walk' is one of the promising biometric identification. It is well known that gait is an idiosyncratic feature of a person that is determined by, among other things, an individual's weight, limb length, footwear, and posture combined with characteristic motion [1]. Human gait sequences are based on spatio-temporal periodic signal [17]. Gait is one special biometric which can be detected and measured at a distance based on low resolution images. Conversely, several challenges in gait analysis and recognition study include robustness of foreground segmentation of the walking subject from the background scene, changes in clothing of the subject, variations in the camera viewing angle with respect to the walking subjects as well as changes in gait as a result of mood or speed change, or as a result of carrying objects [1] that contributed to recognition accuracy via gait. Earlier, at the beginning of the gait recognition existence, marked-based techniques are used that required expensive specialized hardware [2]. However, nowadays, most recent approaches based on computer vision approach have attracted most researchers' attention to utilised this approach for recognition of gait from videos as well as real time application. This strong interest is driven by a variety of potential applications such as visual surveillance, covert security and access control [3].

Current approach for gait recognition can be classified into two categories that is appearance based [4] [5] [6] [7] & [8] that deals directly with image silhouettes which is obtained as a results of background subtraction to extract features. For instance, Boulgouris et al. [4] applied Radon transform on silhouette to extract the feature from each frame and employed Linear Discriminant Analysis

(LDA) to reduce the dimension of accumulated feature vector of one period. Next, Hayder et al. [5] proposed a new combination between Radon Transform to extract feature along 0° and 180° and Principal Component Analysis (PCA) is used to reduce the dimension of feature followed by Euclidean distance for recognition purpose. Appearance based approach are lower computational cost and easier to be implemented [9].

Next, another category of gait recognition is model based [10] [11] & [12] which firstly model the image data followed by analysing the variation in its parameter for gait recognition. As in [13], Chew et al. has proposed a new technique using coupled oscillator for the thigh and lower leg with fewer parameters. Recognition rate that achieved for walking was better than running recorded using k-nearest neighbour. Further, another approach was also introduced by Xiaxi et al. [14] based on combination of three discriminative features specifically area, gravity centre and orientation of each body component that improved performance for gait recognition. Also, Imed Bouchrika et al. [15] proposed a new baseline analysis that deployed for gait recognition using elliptic Fourier descriptor to parameterise motion of the joint ankle, knee and hip joint and achieved better recognition rate. Amongst the two approaches discussed above, normally model based will contribute to higher accuracy since this method is capable to combat noise effectively whilst model free based or appearance based approach will contribute to lower computational cost [16] [30].

Figure 1 showed an example of gait cycle that is the time between successive foot contacts of the same limbs. One gait cycle took place once the reference foot is in contact with the ground and ends with subsequent floor contact of the same foot. As observed in Figure 1, variations between the feet and hand due to human walking could be extracted and utilised for recognition purpose. However in this study we deemed further to investigate only the variations due to lower limb as features for gait recognition and compared with entire body features. The CASIA gait database will be used as our inputs with Discrete Cosine Transform (DCT) as feature extraction followed by PCA as feature selection. Next, feed forward neural network (NN) and probabilistic neural network (PNN) will be evaluated as classifiers for gait recognition.
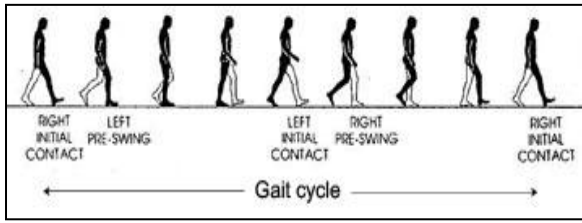
Figure 1 Gait Cycle

II.   PROPOSED METHODOLOGY

Figure 2 depicted the general block diagram of the proposed methodology in gait recognition. Firstly, the original gait sequences are pre processed to obtain sub body followed by DCT as feature extraction. To validate the effectiveness of sub body of gait silhouette as features, lower limb due to average gait silhouette will also be evaluated for comparison purpose.
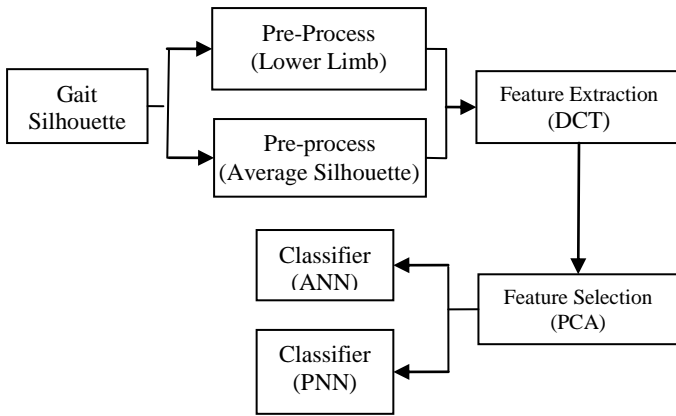

Figure 2: Overview of the Proposed Methodology for Gait Recognition

A.   Lower Limb

As stated earlier, the lower limb will be used by extracting the variations as features for recognition. Firstly, the centroid point is determined. Then, the lower limb is cropped based on the centroid. Fig. 3 showed the entire body and example of lower limb.

Figure 3: Entire body (Left) & Lower Limb (Right)

B.   Averaged silhouette

Averaged silhouette of subject in one complete cycle is normally utilised since this can contribute to memory reduction as well as optimize computational time [28]. Moreover averaging the silhouette generated robust features that can reduce noise [29]. As the width and height of the silhouette image vary due to walking, averaged of a sequence of silhouette is computed based on centroid too. The averaged silhouette is defined as follows:

$$A\,(x,y) = \frac{1}{n} \sum_{i=1}^{n} G\,(x,y) \tag{1}$$

where n is the number of frames in single cycle, x and y are values in 2D image coordinates and i is the frame number in sequences. Figure 4 depicted example of average silhouette in gait sequences.

(a)   (b)   (c)   (d)
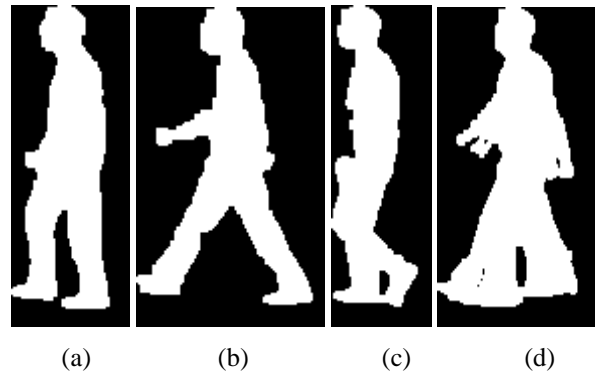
Figure 4: (a) Min. Width Silhouette (b) Max. Width Silhouette (c) Min Width Silhouette (d) Resultant of Averaged Silhouette

The two-dimensional DCT has the property that, for a typical image, most of the visually significant information about the image is concentrated in just a few coefficients of the DCT. For this reason, the DCT is often used in image compression applications. The equation involves are:

$$B_{pq} = \alpha_p \alpha_q \sum_{M=0}^{M-1} \sum_{N=0}^{N-1} A_{MN} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N},$$

where,
$$0 \le p \le M-1$$
$$0 \le q \le N-1$$
$$\alpha_p \begin{cases} 1/\sqrt{M} & ,p=0 \\ \sqrt{2/M} & ,1 \le p \le M-1 \end{cases}$$
$$\alpha_q = \begin{cases} 1/\sqrt{N} & ,q=0 \\ \sqrt{2/N} & ,1 \le q \le N-1 \end{cases}$$

These functions are called the basic functions of the DCT.

Then, the DCT coefficients, *Bpq*, can be regarded as the weights applied to each basis function. For this study, 8-by-8 matrices are applied then 64 basis functions is constructed as illustrated in Fig. 5.
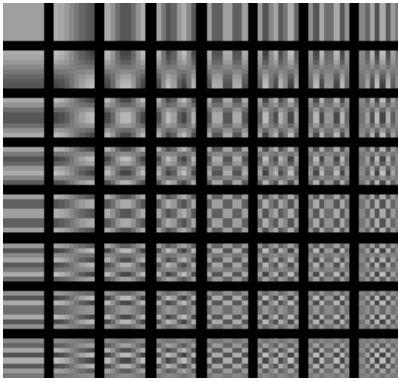
Figure 5: 64 Basis Functions from DCT

Fig. 5 illustrated that horizontal frequencies increased from left to right, and vertical frequencies increase from top to bottom. The constant-valued basis function at the upper left is known as the DC basis function, and the corresponding DCT coefficient B_{00} is called the *DC coefficient*. Because of the compressibility of the 2D-DCT, its low-frequency coefficients represent most information of the image. In this way, the characteristics of image can be efficiently extracted and the amount of data can be relatively reduced [21]. In this study, DCT is used as feature extraction.

D. PCA as feature selection

PCA is widely used in recognition field to reduce the dimensionality of the data. The goal of PCA is to reduce the dimensionality of the data while retaining as much as possible of the variations in the original dataset. PCA allows computation of linear transformation that maps data from a high dimensional space to a lower dimensional space [20] [32]. The strength of PCA as feature selection will be verified in this study.

E. ANN as Classifier

The pattern classifier (or recognizer) is one of the most important components in gait identification system [24]. Many approaches to analyze and recognize gait have been used especially k-nearest neighbours [26], neural networks [31] and SVM [27] [30]. T. Chau [25] stated that neural network facilitated gait analysis because of their highly flexible, inductive, non-linear modelling ability, unlike any other approaches. Hence we will continue to use ANN as classifier in this work.

F. PNN as Classifier

Besides ANN, PNN introduced by Specht [22] will also be investigated for classification task. PNN is a supervised NN that is widely used in the area of pattern recognition, nonlinear mapping, and estimation of the probability of class membership and likelihood ratios [23]. PNN is one of the classifiers that are relatively insensitive

to outliers and generate accurate predicted target probability scores.

III. EXPERIMENTAL ANALYSIS & DISCUSSION

To validate the viability of the proposed method, 10 subjects from the CASIA gait database is chosen as our database. Each person has 12 image sequences with 4 sequences for direction of 90 degrees to the image plane. The length of each sequence is not identical for the variation of the walker's speed, but it must ranges from 37 to 127.

A. Lower limb

Firstly, maximum stride of human walking is identified from the gait cycle to compute the lower limb. Dimension of the features extracted using DCT are 1650 x 10. Then the output features is reduced by PCA to 5 x 10. The classification is done based DCT followed by PCA as feature selection.

B. Average Silhouette

On the other hand, averaged silhouette is computed from the sequence of silhouette images based on one cycle. Further, the dimension features extracted based on DCT is 3900x10. Same as the lower limb, to verify the effectiveness of DCT, PCA acted as feature selection that reduced the dimension of features to 7 x 10.

C. Results and Discussion

Here, the result from the classification stage will be elaborated. The architecture of NN used in this classifier consisted of two layers with hyperbolic tangent sigmoid as activation function for hidden layer and linear transfer function for output layer. Besides that, training function used is Levenberg-Marquardt backpropagation which is the fastest backpropagation algorithm and reduces memory requirements. This network is further trained with 190 epochs. As for the PNN, the first layer has radial basis neurons with its weighted inputs using Euclidean distance function and its net input function using net product. The second layer has competitive neurons transfer function with its weighted input using dot product and its net inputs using sum net.

Experimental results obtained both lower limb and average silhouette with DCT solely as feature extraction followed by PCA as feature selection is tabulated in Table 1. For lower limb gait silhouette, the result obtained for ANN as a result of DCT as feature extraction and PCA as feature selection contributed to highest recognition rate specifically 90% for ANN whilst PNN attained 60% accuracy. As for average silhouette, once again ANN outperformed PNN based on recognition rate attained specifically 80% for ANN and 70% for PNN respectively. On the contrary, accuracy rate attained for both average silhouettes as well as lower limb based on DCT solely is poor as compared to implementation of PCA as feature selection that contributed to higher recognition rate for all cases. In addition, ANN as classifier surpassed PNN in all cases too.

Table 1: Accuracy Rate (%) for ANN and PNN based on the proposed method

Feature Vectors	ANN		PNN	
	Average Silhouette		Lower Limb	
DCT	60	60	60	50
DCT & PCA	80	**90**	70	60

IV. CONCLUSION

This paper presented a new gait recognition method based on DCT as feature extraction and PCA as feature selection. Based on the accuracy rate attained, it was proven that PCA as feature selection contributed to higher recognition rate as compared to DCT as feature extraction solely. Also ANN as classifier also obtained better result using lower limb as compared to PNN. Future works includes verifying the proposed method in a larger database and other condition of subjects such as wearing shoe, jacket, and others in variation of angle.

ACKNOWLEDGMENT

REFERENCES

[1] L. Lee W. E. L. Grimson, "Gait Analysis for Recognition and Classification", Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition (FGR.02), 2002.

[2] Bo Ye, Yumei Wen, "A New Gait Recognition Method Based on Body Contour", Control, Automation, Robotics and Vision, 2006. ICARCV '06, 2010, pp 1726 – 1729.

[3] Liang Wang, Huazhong Ning, Tieniu Tan, Weiming Hu," Fusion of Static and Dynamic Body Biometrics for Gait Recognition", Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set, 2003.

[4] N. V. Boulgouris and Z. X. Chi, "Gait recognition using radon transform and linear discriminant analysis," IEEE Trans. Image Process., vol. 16, no. 3, pp. 857–860, March 2007.

[5] Hayder Ali, Jamal Dargham, Chekima Ali, Ervin Gobin Moung, "Gait Recognition using Radon Transform with Principal Component Analysis", 2010 The 3rd International Conference on Machine Vision (ICMV 2010), 2011.

[6] Qi Yang , Kuisheng Qiu ,"Gait recognition based on active energy image and parameter-adaptive kernel PCA", Information Technology and Artificial Intelligence Conference (ITAIC), 2011 6th IEEE Joint International , 2011, pp: 156 – 159.

[7] Han Su, Feng-Gang Huang, "Human Gait Recognition Based On Motion Analysis", Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, August 2005.

[8] Sungjun Hong, Heesung Lee, Kar-Ann Toh, and Euntai Kim, "Gait Recognition Using Multi-Bipolarized Contour Vector", International Journal of Control, Automation, and Systems (2009) , pp: 799-808.

[9] Bo Ye, Yumei Wen, " A New Gait Recognition Method Based on Body Contour", Control, Automation, Robotics and Vision, 2006. ICARCV '06. 9th International Conference , 2006, pp: 1-6.

[10] Johnson, A.Y., Bobick, A.F., A Multi-View Method for Gait Recognition Using Static Body Parameters, LNCS, 2091, 301-11 (2001).

[11] Donghyeon Kim, Daehee Kim and Joonki Paik, "Model-Based Gait Recognition Using Multiple Feature Detection", Lecture Notes in Computer Science, Volume 5259/2008,pp: 1018-1029, 2008.

[12] Faezeh Tafazzoli, Reza Safabakhsh, "Model-based human gait recognition using leg and arm movements", Engineering Applications of Artificial Intelligence, Volume 23, Issue 8, pp: 1237–1246, December 2010.

[13] Chew-Yean Yam, Mark S. Nixon, John N. Carter, "Extended Model-Based Automatic Gait Recognition of Walking and Running", eprints.soton.ac.uk, 2001

[14] Xiaxi Huang and Nikolaos V. Boulgouris, "Model-Based Human Gait Recognition Using Fusion Of Features", Acoustics, Speech and Signal Processing, 2009. ICASSP, pp: 1469-1472, 2009.

[15] Imed Bouchrika and Mark S. Nixon, "Model-Based Feature Extraction for Gait Analysis and Recognition", MIRAGE 2007, LNCS 4418, pp. 150–160, 2007.

[16] Nurul Illiani Yaacob and Nooritawati Md Tahir, "Feature Selection For Gait Recognition",

[18] C. BenAbdelkader, R. Culter, and L. Davis, "Stride and Cadence as a Biometric in Automatic Person Identification and Verification," in Proc. 5th International Conference of Automatic Face and Gesture Recognition, Washington, DC, USA, 2002, pp. 372-377.

[19] L. Wang, T. Tan, H. Ning et al. "Silhouette Analysis-Based Gait Recognition for Human Identification," IEEE Trans. On Pattern Analysis and Machine Intelligence, vol. 25, pp.1505-1518, 2003.

[20] Ling-Feng Liu, Wei Jia and Yi-Hai Zhu, "Gait Recognition Using Hough Transform and Principal Component Analysis", ICIC 2009, LNCS 5754, pp. 363–370, 2009.

[21] Zhang Qian-jin and XU Su-li, "Gait-based Recognition of Human Using an Embedded Hidden Markov Models", Information Engineering and Computer Science, 2009. ICIECS, pp:1 – 4, 2009.

[22] D. F. Specht, "Probabilistic neural networks", Neural Networks, 1991, 1(3), pp. 109-118.

[23] D. F. Specht, H. Romsdahl, "Experience with adaptive probabilistic neural networks and adaptive general regression neural networks", IEEE International Conference on Neural Networks, 1994, 2, pp. 1203-1208.

[24] Jang-Hee Yoo and Doosung Hwang, "Automated Human Recognition by Gait using Neural Network", Image Processing Theory, Tools and Applications, 2008. IPTA, pp: i – iii, 2008.

[25] T. Chau, "A Review of Analytical Techniques for Gait Data. Part 2: Neural Network and Wavelet Methods", Gait and Posture, 13(2), pp.102-120, 2001.

[26] Lu Xiaoqi, Huang Hongmei, Zhang Baohua, "Recognition of human movement Based on Fourier Descriptors", Bioinformatics and Biomedical Engineering, 2008. ICBBE, pp: 1943-1946, 2008.

[27] Jiwen Lu, Erhu Zhang, "Gait recognition for human identification based on ICA and fuzzy SVM through multiple views fusion", Pattern Recognition Letters 28 pp: 2401–2411, (2007).

[28] Nini Liu, Jiwen Lu, Yap-Peng Tan and Zhenzhong Chen,"Enhanced Gait Recognition Based On Weighted Dynamic Feature", IEEE International Conference on Image Processing (ICIP), Nov. 2009.

[29] Sabesan Sivapalan, Rajib Kumar Rana*, Daniel Chen, Sridha Sridharan, Simon Denmon and Clinton Fookes, "Compressive Sensing for Gait Recognition", International Conference on Digital Image Computing: Techniques and Applications, 2011.

[30] Fourier descriptor for pedestrian shape recognition using support vector machine Tahir, N.Md., Hussain, A., Mustafa, Mohd.M., Samad, S.A., Husin, H. 2007ISSPIT 2007 - 2007 IEEE International Symposium on Signal Processing and Information Technology , pp. 636-641

[31] A speed limit sign recognition system using artificial neural network Ishak, K.A., Mohd Sani, M., Md Tahir, N., Samad, S.A., Hussain, A. 2006SCOReD 2006 - Proceedings of 2006 4th Student Conference on Research and Development "Towards Enhancing Research Excellence in the Region" , pp. 127-131

[32] Feature selection for classification using decision tree Tahir, N.M., Hussain, A., Samad, S.A., Ishak, K.A., Halim, R.A. 2006SCOReD 2006 - Proceedings of 2006 4th Student Conference on Research and Development "Towards Enhancing Research Excellence in the Region" , pp. 99-102

Human Shape Recognition using Radon Transform and Regularized Principal Component

A.R. Mahmud & Nooritawati Md Tahir[*]
Faculty of Electrical Engineering,
Universiti Teknologi Mara
40450, Shah Alam Selangor, Malaysia
[*]Corresponding Author: norita_tahir@yahoo.com

Abstract--**The aim of this study is to investigate the potential of Radon Transform and Regularized Principal Component Analysis as feature extraction for classification of human and non human. Several training algorithms are used for the neural network. The finding of the investigation shows that the best training algorithm is Lavenberg-Marquardt (LM). In addition, the execution time taken by LM is fastest among the training.The outcomes of the proposed method using LM are 0% False Rejection Rate (FRR) and 0% False Acceptance Rate (FAR)ona database of 100 images on each category.**

Keywords- Radon Transform, Principal Component Analysis, artificial neural network, training algorithm, classification rate, execution time.

I   INTRODUCTION

Computer vision is one of important research area today. The applications of computer vision are widely used in our life based on their functions. Wang et al. [3] used a codebook which is created based on segment of a pedestrian image. The segmentations are analyzed using improved Shape Context feature in order to build a hypothesis. The hypothesis is then verified using False Positive Pruning procedure. Shape-based recognition and classification for common objects is used by Ku et al. [4]. This approach divides objects into 4 different classes, which are human, four legs animal, vehicle and other object. The classification of the categories is conducted using Width to Height Ratio test followed by Base to Abdomen Ratio or Shape Boundary test. Sabzmeydani and Mori [5] implement Shapelet feature with Adaptive Boosting (AdaBoost) in categories pedestrian and non-pedestrian while Nguyen and Phung [6] used Convolutional Neural Network to do the same task. Object detection also can be analyzed using Principal Component Analysis (PCA). In [7, 8, 9], the combination of PCA and Support Vector Machine (SVM) are used to enhanced the feature extraction process. In order to conduct a feature extraction on nonlinear image, Kernel PCA is used. The employment of AdaBoost is to discover the most suitable features of different classes [10]. This approach produced detection rates above 95% with minimal false alarm rates. Vehicle detection and tracking can also be conducted using either PCA or SVM. Alvarez et al. [11] investigated vehicle and pedestrian detection using SVM classifier in extracting the features from the candidates. Objects detection in this approach used inward motion, textures, horizontal edges, vertical symmetry and entropy. Since the subject involve with movement, Kalman filter is used for the final task which is tracking process. Chen and Su [12] used PCA for moving object tracking in video sequence. The object is correctly detected and the path of the object pass through is able trace exactly. Fernando et al. [13] combined the used of multiresolution analysis, neural network and PCA in detection and tracking vehicle. It used an onboard camera which it also moving in detecting vehicle. Akoz and Karsligil [14] explored a new function of object detection by implementing partial vehicle trajectories in video-based accident analysis. Monitoring the normal flow of vehicle at intersection is recorded therefore it is used in differentiating normal and abnormal event.

Radon Transform (RT) was introduced by Johann Radon in 1917. It is an integral transform consisting of the integral of a function over straight line. Lin et al. [15] experimented RT, Multiwavelet and PCA in infrared face recognition. The first two methods are used to construct the low-frequency component while PCA functions as feature extraction. Motion blur will degrade the images used for analysis. Mohammadi et al. [16] and Sun et al. [17] propounded RT in dealing with this condition. In [16], Fourier Transform is used to extract the blur image and then implemented RT for extracting the motion direction in vehicle speed estimation while [17] estimated the parameter from image gradient in image restoration. Image recognition usually involve with respect to the shape of the object.

Mustaffa et al. [18] used Ridgelet Transform and RT in handling images of arbitrary size. Zhao and Liu [19] analyzed human activity by proposing adaptation of RT and edit distance matching scheme. Khan and Sohn [20] investigated human activity using PCA and RT. Training and recognition is done by Hidden Markov Model adaptation. RT also being used in gait recognition [21,22]. Zhang and Liu [21] presented Haar Wavelet and RT of binary silhouettes in gait representation and recognition by having feature database of gait. Guha and Ward [22] enhanced RT by producing Differential RT for gait recognition.

In this paper, image identification using RT and PCA are presented. Input of the RT will be silhouette of moving objects which are vehiclesas non human and human. The output from the RT is fed to PCA. Artificial Neural Network (ANN) is used to as classifier [26]. Several training algorithms in the toolbox is simulated and compared for their performance and accuracy. The training algorithms tested are BFGS, BR, GDX, LM, RP and SCG. The remaining of this paper is organized as follows: in Section II, RT, PCA and ANN are introduced. While in section III, results are presented. Finally, we concluded and discussed on future work in section IV.

II   METHODOLOGY

In this section the method proposed will be discussed. The overview of the proposed method is as shown in Figure1.
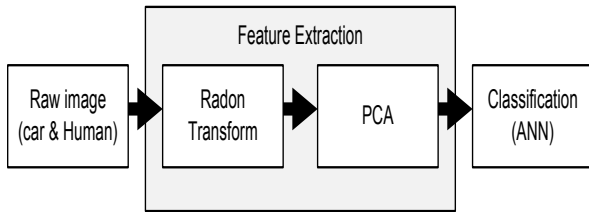

Figure 1: Proposed Method

A.   Radon Transform

Radon Transform of 2-D function $f(x, y)$ is illustrated in Figure 2. The projection of the 2-D function $f(x, y)$ on the axis s of θ direction produced a result of function $g(s, \theta)$ which is the process of integration along the line with normal vector in θ direction. Initial value of s equal to zero is used for the integration so that it passes the origin of (x, y)-

coordinate in order to obtain the projection result of $g(0, \theta)$ [23]. The points on this line satisfythe equation:

$$\frac{y}{x} = \tan(\theta + \frac{\pi}{2}) = \frac{-\cos\theta}{\sin\theta}$$

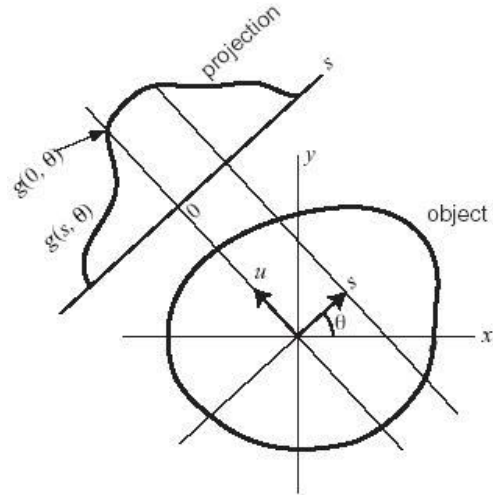$$\Longrightarrow x\cos\theta + y\sin\theta = 0$$

Figure 2: The Radon Transform computation.

The projection results in the integration of $f(x, y)$ only at the points satisfying the previous equation. With the help of the Dirac "function" δ, $g(0, \theta)$ is expressed as:

$$g(0, \theta) = \iint f(x, y) \cdot \delta(x\cos\theta + y\sin\theta) \, dxdy$$

Therefore, the following equation representing the line with normal vector in θ direction and distance s from the origin:

$$(x - s \cdot \cos\theta) \cdot \cos\theta + (y - s \cdot \sin\theta) \cdot \sin\theta = 0$$

$$\Longrightarrow x\cos\theta + y\sin\theta - s = 0$$

In general, the Radon transformation equation is as follows:

In this paper, the RT is tested at four different θ

$$g(s, \theta) = \iint f(x, y) \cdot \delta(x\cos\theta + y\sin\theta - s) \, dxdy$$

which are 0°, 45°, 90° and 135° in order to get different view of the images as shown in Figure 3. The choice of θ is taken from the first two quadrants because the opposite angle of θ will produced the mirror of the graph as shown in Figure 4. Each of the images produces different size of matrices which are difficult to evaluate. Therefore, PCA is implemented in order to overcome the problem.
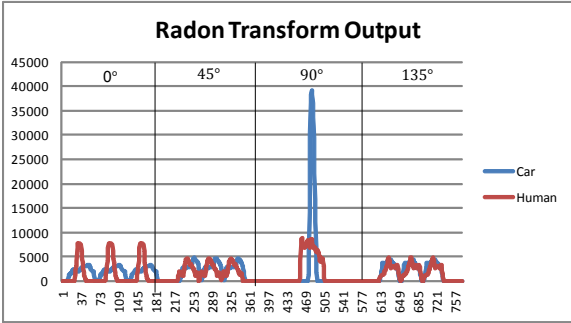
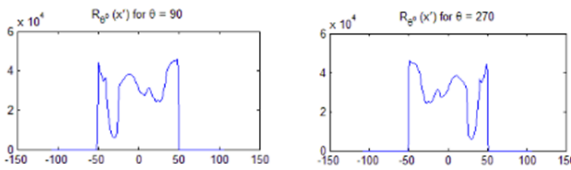Figure 3: Radon Transform output for non human and human with θ = 0°, 45°, 90° and 135°

Figure 4: Radon Transform output sample of image for opposite angle: 90° and 270° (mirror)

B. Principal Component Analysis

In this study, the input vector (matrix) dimension is large and different for each of the images. However, the elements of the vectors are interreled with each other. "This technique has three effects: it orthogonalizes the components of the input vectors (so that they are uncorrelated with each other); it orders the resulting orthogonal components (principal components) so that those with the largest variation come first, and it eliminates those components that contribute the least to the variation in the data set" [24].

Principal components are the variables of a dataset. Each variables generated by PCA are the same as the original data with a coefficient representing the whole data. The rank of the coefficients depending on the level of principal components with the first principal component having the maximum coefficient for the variance and the second component will have the second most variance. The variance for each of the principal components is not interreled and decreasing for the next principal component [27].

The new data generated by the PCA acted as input to the ANN classifier. The conventional PCA is tested. However it produced low accuracy for test and training of ANN. The data is than tested using Regularized PCA (RPCA) [25]. This technique is suitable for data that have many variables as this study has 772 variables for each image.It extracts leading eigenvectors based on number of principal components selected, for example is 6 in this experiment. It also keeps tracks ofextracted

eigenvalues and their sum of leading eigenvalues over the total sum of all eigenvalues.

C. Artificial Neural Network Architecture

Feedforward network is used in this study. It is set to have five hidden layers of sigmoid neurons (tansig) followed by an output layer of linear neurons of purelin. The maximum number of epoch used is 100. The training of the algorithm will stop if it reaches 100 even though the result is not satisfactory.

There are two groups of images used in this study comprised of 50 images of vehicles and 50 images of humans in standing or walking position. Then, each group is divided into two sets which are the training set and testing set. Based on the distribution of the images, each of the training algorithms is simulated where the performance and accuracy of the classification is analyzedand compared.

D. Performance Measures

Three statistical indices are introduced specifically accuracy, sensitivity and specificity using the computed results from True Acceptance Rate (TAR), False Rejection Rate (FRR), True Rejection Rate (TRR) and False Acceptance Rate (FAR).The detail of TAR, FRR, TRR, FAR, Accuracy, Sensitivity and Specificity are as below:

- *TAR*: Classifiers identify Human as Human
- *FRR*: Classifiers identify Human as Vehicle
- *TRR*: Classifiers identify Vehicle as Vehicle
- *FAR*: Classifiers identify Vehicle as Human
- *Sensitivity*: Classifiers ability to identify Human as Human
- *Specificity*: Classifiers ability to identify Vehicle as Vehicle

$$Accuracy = \frac{TAR + TRR}{TAR + FRR + TRR + FAR} \quad (1)$$

$$Sensitivity = \frac{TAR}{TAR + FRR} \quad (2)$$

$$Specificity = \frac{TRR}{TRR + FAR} \quad (3)$$

III   EXPERIMENTAL RESULTS AND ANALYSIS

In this section, the classification result will be discussed. In the first step of the experiments, each of the images is passed through a Radon Transform. It will create a matrix with four variables representing the degree angle of the RT at θ equal to 0°, 45°, 90° and 135°. For recognition purpose, the target is ranged between -1 and 1. The vehicle group is representing

by -1 and the human group representing by 1. Images that are correctly classified should fall between ±0.1 of the target values. Six training algorithms is evaluated as training algorithm namelyBroyden, Fletcher, Goldfarb, and Shanno (BFGS) quasi-Newton, Bayesian Regularization (BR), Gradient Descent with adaptive α and with momentum (GDX), Lavenberg-Marquardt (LM), Resilient Backpropagation (RP) and Scaled Conjugate Gradient (SCG).The result of the classification rate and execution time are shown in Figure 5 and Table 1.

The result for RPCA displayed that BFGS and BR produce the best result followed by LM and RP. However, when the execution time is taken into consideration, LM executed the process in the fastest time compared to other training algorithms. It is three times faster than BFGS and nine times faster than BR. The classification rate of LM is also similarto BFGS and BR except on the classification rate of training in human group. Therefore, LM is the best training algorithms in terms of execution time with promising result in classification rate.

Table 1: Result of ANN Performance Comparison for PCA and RPCA Execution Time

Method	conventional PCA (s)	Regularize PCA (s)
BFG	460.66	2.56
BR	489.22	3.12
GDX	0.81	1.97
LM	55.72	1.37
RP	0.81	1.87
SGC	1.00	1.73

Table 2: Classification results based on LM as training algorithm

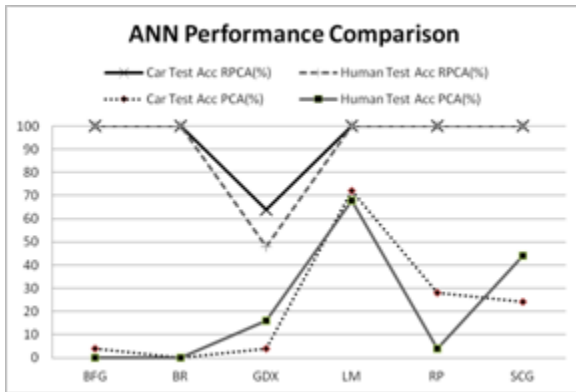| Method	Predicted			
	Conventional PCA (%)		Regularize PCA (%)	
Actual	Human	Vehicle	Human	Vehicle
Human	68	28	100	0
Vehicle	32	72	0	100

Figure 5: Result of ANN Performance Comparison for PCA and RPCA Test Accuracy

The result for conventional PCA proved that LM has the highest test accuracy among the algorithms. The execution time in conventional PCA also is longer compared to RPCA as shown in Table 1.

Table 2 illustrates the classification results based on the best training algorithm (LM) for conventional PCA and Regularize PCA. The relationship of columns and rows represented the value of TAR, FRR, TRR and FAR. TheLM classifications using Regularize PCA produce good results of 100% TAR, 0% FRR, 0% FAR and 100% FRR, while the conventional PCA results are 68% TAR, 32% FRR, 28% FAR and 72% FRR. These result shows that the Regularize PCA outperformed the conventional PCA.

Table 3 illustrates the Accuracy, Sensitivity and Specificityof LM for both PCA. Sensitivity and Specificity are correlated with each other. If the Sensitivity percentage is high and the Specificity percentage is low, this result in the system ability to classify human correctly but unable to do the same thing for non human and vice versa. In other words, the Sensitivity is the correctly classified rate for human and the Specificity is the correctly classified rate for vehicle. A good system should have high Sensitivity and Specificity values.

Table 3: Accuracy, Sensitivity, Specificity, TAR, FRR, TRR and FAR of LM

Performance Measure	Conventional PCA (%)	Regularize PCA(%)
TAR	68	100
FRR	32	0
FAR	72	100
TRR	28	0
Accuracy	70	100
Sensitivity	68	100
Specificity	72	100

IV CONCLUSION

In conclusion, this study has evaluated the effectiveness of RT and RPCA as feature extraction and selection. Initial findings using ANN as classification technique showed that LM is the fastest execution time with good result in classification accuracy. Further work includes using other classifiers namely Support Vector Machine and as classifier and additional database of human and non human.

Acknowledgment

REFERENCES

[1] N. Dalal, B. Triggs, "Histogram of Oriented Gradient for Human Detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.

[2] P.Dollar, C. Wojek, B. Schiele, P. Perona, "Pedestrian Detection: A Benchmark," IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 304-311.

[3] Liming Wang, Jianbo Shi, Gang Song, I-Fan Shen, "Object Detection Combining Recognition and Segmentation," Asian Conference on Computer Vision (ACCV), Vol. 4843 (2007), pp. 189-199.

[4] Z. K. Ku, C. F. Ng, S. W. Khor, "Shape-based Recognition and Classification for Common Objects - An Application in Video Scene Analysis," International Conference on Computer Engineering and Technology, 2010.

[5] P. Sabzmeydani, G. Mori, "Detecting Pedestrians by Learning Shapelet Features,"Computer Vision and Pattern Recognition, 2007.

[6] G.H. Nguyen, S. L. Phung, "Reduced Training of Convolutional Neural Networks for Pedestrian Detection," International Conference on Information Technology and Applications, 2009.

[7] N. M. Tahir, A. Hussain, S. A. Samad, H. Husain, "PCA-based Human Posture Classification," 35 JurnalTeknologi, UniversitiTeknologi Malaysia 46(D) Jun 2007, pp. 35-44.

[8] L.M. Borja, O. Fuentes "An Object Detection System using Image Reconstruction with PCA,"Proceedings of the Second Canadian Conference on Computer and Robot Vision, 2005

[9] S. Sun, "Ensembles of Feature Subspaces for Object Detection," The 6th International Symposium on Neural Networks, Lecture Notes in Computer Science, 2009, Vol. 5552, pp. 996-1004.

[10] S. Ali, M. Shah, "A Supervised Learning Framework for Generic Object Detection in Images,"Proceedings of the Tenth IEEE International Conference on Computer Vision, 2005.

[11] S. Alvarez, M. A. Sotelo, I. Parra, D. F. Llorca, M. Gavilan, "Vehicle and Pedestrian Detection in eSafety Application," Proceedings of the World Congress on Engineering and Computer Science, 2009.

[12] C.F. Chen, Y.T. Su "The Use of PCA for Moving Objects Tracking on the Image Sequence,"Canadian Conference on Computer and Robot Vision, 2005.

[13] S. Fernando, L. Udawatta, P. Pathirana "Visual Tracking of Vehicles using Multiresolution Analysis and Neural Network," Proceedings of the 4th International Conference on Information and Automation for Sustainability, pp. 355-360, 2010.

[14] O. Akoz, M. E. Karsligil, "Video-Based Traffic Accident Analysis at Intersections Using Partial Vehicle Trajectories," IEEE 17th International Conference on Image Processing, 2010.

[15] Z. Lin, F. Zhijun. W. Shengqian. Zhang Wenrui. Yang Fan. "Infrared face recognition based on Radon and multiwavelet transform,"IEEE International Conference on Communications Technology and Applications, 2009.

[16] J. Mohammadi, R. Akbari, M. K. Bahaghighat, "Vehicle Speed Estimation Based On the Image Motion Blur Using RADON Transform,"International Conference on Signal Processing Systems, 2010.

[17] H. Sun, M. Desvignes, Y. Yan, W. Liu, "Motion Blur Parameters Identification from Radon Transform Image Gradients," Industrial Electronics Conference, 2009.

[18] M. R. Mustaffa, F. Ahmad, R. Mahmod, S. Doraisamy" InvariantGeneralisedRidgelet-Fourier for Shape-based Image Retrieval," International Conference on Information Retrieval & Knowledge Management, 2010.

[19] H. Zhao, Z. Liu, "Shape-Based Human Activity Recognition Using Edit Distance," International Congress on Image and Signal Processing, 2009.

[20] Z. A. Khan, W. Sohn "Feature Extraction and Dimensions Reduction using R transform and Principal Component Analysis for Abnormal Human Activity Recognition," International Conference on Advanced Information Management and Service, 2010.

[21] H. Zhang, Z. Liu "Gait Representation and Recognition Using Haar Wavelet and Radon Transform," International Conference on Information Engineering, 2009.

[22] T. Guha, R. Ward, "Differential Radon Transform for Gait Recognition," IEEE International Conference on Acoustics Speech and Signal Processing, 2010.

[23] S. Venturas, I. Flaounas, "Study of Radon Transformation and Application of its Inverse to NMR," Paper for "Algorithms in Molecular Biology" Course, 4 July, 2005.

[24] http://matlabdatamining.blogspot.com/2010/02/putting-pca-to-work.html by Will Dwinnell.

[25] Yeh, Yi-Ren, KernelStat toolbox at http://dmlab1.csie.ntust.edu.tw/downloads.

[26] A speed limit sign recognition system using artificial neural network Ishak, K.A., Mohd Sani, M., Md Tahir, N., Samad, S.A., Hussain, A. 2006SCOReD 2006 - Proceedings of 2006 4th Student Conference on Research and Development "Towards Enhancing Research Excellence in the Region" , pp. 127-131

[27] Feature selection for classification using decision tree Tahir, N.M., Hussain, A., Samad, S.A., Ishak, K.A., Halim, R.A. 2006SCOReD 2006 - Proceedings of 2006 4th Student Conference on Research and Development "Towards Enhancing Research Excellence in the Region" , pp. 99-102

Video Stabilization Based on Point Feature Matching Technique

Labeeb Mohsin Abdullah, Nooritawati Md Tahir* & Mustaffa Samad

Faculty of Electrical Engineering
Universiti Teknologi MARA (UiTM)
40450, Shah Alam, SELANGOR
Corresponding author: norita_tahir@ yahoo.com

Abstract This study proposed an algorithm to stabilize jittery videos directly without the need to estimate camera motion. A stable output video will be attained without the effect of jittery that caused by shaking the handheld camera during video recording. Firstly, salient points from each frame from the input video is identified and processed followed by optimizing and stabilize the video. Optimization includes the quality of the video stabilization and less unallied area after the process of stabilization. The output of using such method showed good result in terms of stabilization and discarded distortion from the output videos recorded in different circumstances. Initial results showed that the proposed technique is suitable to be used and provide great deal of stabilization.

Keywords: image processing, video stabilization, point feature matching, salient points, image quality measurement

I.   INTRODUCTION

Recently, the market of handheld camera has growth rapidly. However, video capturing by non-professional user normally will lead to unanticipated effects. Hence, many researchers study such drawbacks to enhance the quality of casual videos. Currently, hardware stabilizers are attached to the cameras as effective solution. On one hand, pre-processing techniques such as nonlinear filters is applied to discard the unwanted noise. On the other hand, using multi-stages for pre and post processing could aggravate the existing problems according to errors accumulative. However, there are shortcomings related to process the videos with complicated motion such as multiple moving foreground objects [1].

Generally the process of stabilization have to go through three phases namely motion estimation, motion smoothing and image composition [2]. For the first phase the purpose is to estimate the motion between frames. After that, the parameters of estimated motion which is obtained from the first phase will be sent to motion compensation, where removing the high-frequency distortion and calculating the global transformation, which is very important to stabilize the current frame. Next, warping will be done by image composition for the frame under processing [9] .This three-step frameworks

are the essential steps in most of the video stabilization algorithms.

II. MATERIALS AND METHOD

This section will present the overview of the proposed methodology and implementation as depicted in Figure 1.
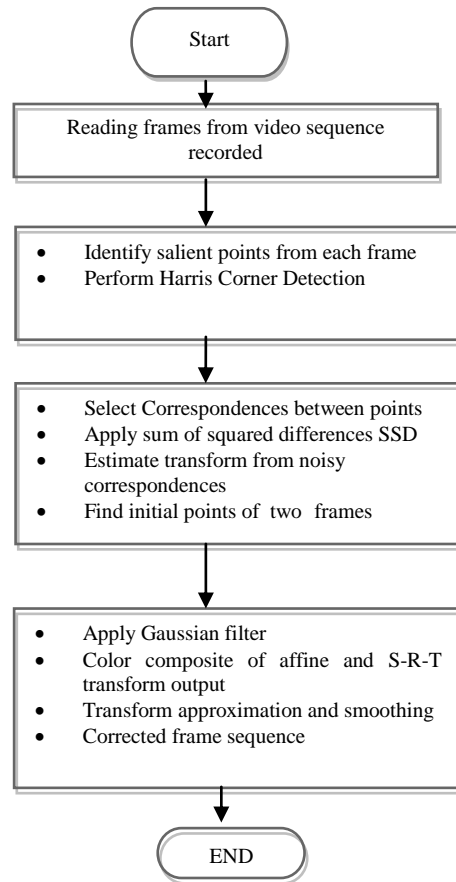

Figure 1: Overview of the proposed method of video stabilization.

A. Identification of salient points from each frame & Harris Corner Detection

The main goal of this step is to correct the distortion between the two frames by finding a transformation that will be done by applying an object system which returns affine transform [3]. The input for this stage should supply the object with a set of point correspondences between the two frames [4]. Firstly, the wanted points from the two chosen frames have to be identified followed by selecting the common correspondence between the frames. At this point, the candidate points for each frame are identified but to make sure that these points will have corresponding points in the second frame, it is necessary to find points around salient image features, like corners. Thus, Corner Detector System Object is used to find corner values using Harris Corner Detection which is one of the fastest algorithms to find corner values.

B. Select correspondences between points & SSD

After the salient points from each frame are obtained the correspondence between the points that are identified previously need to be picked [4]. For each point, the matching of lowest cost between the points that existed in frame A and B are also needed to be found for all points. Hence, it is necessary to divide the sequence of frames image into 9×9 block. The matching cost means the distance between frame A and B measured in pixel. To find this cost, the technique of Sum of Squared Differences (SSD) can be used between the consecutive frame images. Each point in frame A is compared with the points in frame B to find the lowest matching cost or in other words the shortest distance between them measured in pixels.

III. RESULTS AND DISCUSSION

In this section, the results attained based on the proposed methodology will be discussed. Table I showed the basic characteristics of each video utilized as database in this study. In addition, the values of the size and the number of bytes for the salient points existed in each video are also tabulated.

A. Strong corners Detection

Firstly, an algorithm is developed based on Harris and Stephens corner detection algorithm [2] to identify all salient points or strong corners from each frame. These points are considered as the anchor points as benchmark for points to be considered and vice versa. Sample of detected points obtained from two frames are as demonstrated in Figure 2. Furthermore, it is observed the total points covered are the same frame features for instance the salient points along the trees, corners of the sidewalk and the moving object.

Figure 2: The detected strong corners from both frames where they marked with green dots.

B. Corresponding points

Next, the initial correspondences between the points that is identified from the previous step will be invoked. Correspondences between the invoked points have to be picked for each point, for that purpose a matrix of 9 x 9 blocks will be extracted around each point from its consecutive image frames. The most important here is matching the cost between points by finding the Sum of Squared Differences (SSD) between the consecutive image regions of frames. Thus we have to find the lowest costs to consider them in the solution [8]. Figure 3 showed the same positions for the green color points of the initial corresponding points existed in both frames.
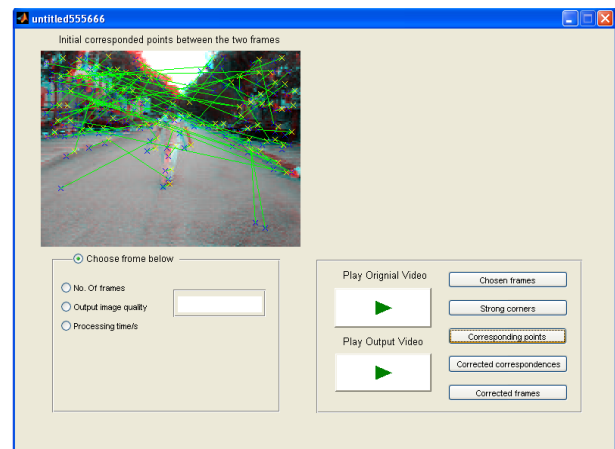

Figure 3: Corresponding points between frames.

However, not all these correspondence points are correct, which means many of them are redundancy points, but at the same time there is a significant number of outlier points as well. This lack will be considered in the next step. SSD will ensure to find the minimum cost matching

point in points B with the aid of features, which resulted a loop over points A that search for best matches in points B with features contribution.

C. Accurate correspondence

As mentioned above, there are several incorrect point correspondences but strong estimation of geometric transform between the two image frames can be determined using the random sample consensus algorithm (RANSAC) [5][6]. This algorithm searched through the given set of point correspondences specifically valid linear correspondence as in Figure 4.
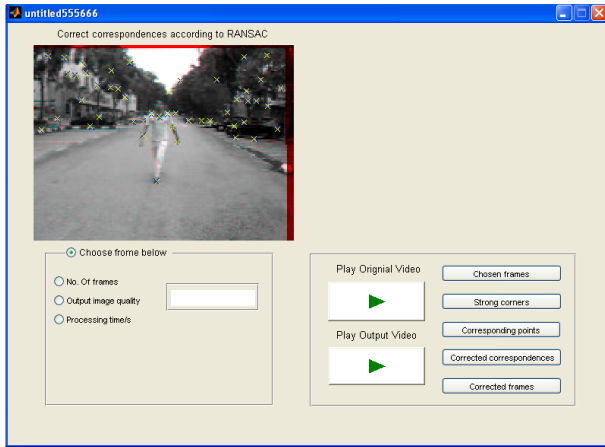

Figure 4: Correct correspondences according to RANSAC.

From Figure 4, the inliers correspondences consecrated in the image background, not in the foreground, which itself is not aligned; is observed. The reason stand behind this is the background features are far enough that act as if they were on an infinitely distant plane. We can assume that background plane is static and will not change dramatically between the first and second frame, instead, this transform is capturing the motion of the camera. Thus correcting process will stabilized the video. Furthermore, as long as the motion of the camera between frame A and frame B is minimize or the time of sampling the video is high enough, this condition is maintained. The RANSAC algorithm is repeated multiple times and at each run the cost of the result is calculated by projecting frame B onto frame A via Sum of Absolute Differences between the two image frames and the results attained is as in Table I.

Firstly, the effect of the number of corners is influenced by the output as resulted in Vid1 with highest matching point values but least SSD followed by Vid3 and Vid2 respectively. This indicated that Vid2 comprised of the maximum number of salient points to be handled since the SSD attained is the highest.

D. Frames Correction

Further, the raw mean video frames and the mean of corrected frame are computed as in Figure 5.

Figure 5: Corrected frames

The left image showed the mean of the raw input frames that resembled the distorted original video frame due to extreme jittery. On the right side is the mean of the corrected frames with less distortion. This proven that the stabilization algorithm worked well. Several more samples of corrected video frames are as depicted in Figure 6.

E. Quality

The output video quality is also measured based on the proposed methods. This is evaluated based on SVD based grayscale Image value and graphical measurement.

i. SVD Based Grayscale Image Quality

Singular value decomposition (SVD) is developed as a new measurement that can express the quality of distorted images either graphically that is in 2D measurement or numerically as a scalar measurement, both near and above the visual threshold. The experiments here utilized SVD based measurement that outperformed the normally used PSNR [10]. Equation 1 represented the computed value for this purpose:

$$M - SVD = \frac{\sum_{i=1}^{\left(\frac{k}{n}\right)x\left(\frac{k}{n}\right)}|D_i - D_{mid}|}{\left(\frac{k}{n}\right)x\left(\frac{k}{n}\right)} \qquad (1)$$

where:

D_{mid} represents the midpoint of the sorted D_{iS}
k is the image size
n is the block size
M –SVD is the measurement of Singular value decomposition

An example for the output quality for Vid1 based on Equation 1 with k=8, n=1, D_i and D_{mid} represented by 256×256 matrix attained M-SVD of ≈ 22.20. Hence, the numerical quality obtained from the three sample videos are tabulated in Table 1. As visualize in Figure 6, it can be seen that Vid3 obtained the best quality based on the calculated value that is 40.50% followed by Vid2 with 39.21% and Vid1 22.20%. This resembled that Vid1 has great distortion whilst Vid3 is least distorted.

Figure 6: Three input videos (Vid1, Vid2, Vid3) with different stabilization need to be done since recorded in different circumstances.

i. *Graphical measurement*

The criteria of measuring graphical quality in any image or frame can be done as shown in Figure 7. Graphical measurement will indicate the condition of video due to distortion. As seen in the graphical results, Vid1 has the most size of blocks among all frames blocks to indicate

that it is the worst video experienced distortion followed by Vid2 and Vid3. This result agreed with the SVD value calculated earlier.

Figure 7: Graphical measurement for each sample videos (Vid1, Vid2 & Vid3) as quality stabilization indication

IV. CONCLUSION

In conclusion, the video stabilization technique based on proposed method showed remarkable results in term of stabilizing high jittery videos suffered from distortion. Initial results also proven that due fusions of RANSAC algorithm, Gaussian filter, Harris, Stephens's and SAD efficiency stabilization process succeeded based on the output quality attained. Future work includes finding better feature detector and overcome the consequences of extreme shaking of handheld camera in feasible real time implementation for video stabilization.

Table I: Criteria of sample videos & Results (Type: RGB & Extension: AVI)

Sample 1st Frame	Sample Inputs	Size	Bytes No	Frames # & Length	SAD Value	Computational Time (s)	Quality Value (M -SVD)
	Vid1	2x139	1112	34 & 2	$8.85e^5$	8.81	22.20 %
	Vid2	2x66	528	73 & 3	0.0824	9.18	39.21%
	Vid3	2x128	1024	132 & 4	$3.553e^3$	9.28	40.50%

REFERENCES

[1] M. Gleicher and F. Liu., "Re-cinematography: Improving the camerawork of casual video," ACM Transactions on Multimedia Computing, Communications, and Applications, 5(1), pp 1- 28, 2008 .

[2] C. Harris and M.J. Stephens, "A combined corner and edge detector", Proc of Alvey Vision Conference, pp 147–152, 1988.

[3] Anu Suneja and Gaurav Kumar . "An Experimental Study of Edge Detection Methods in Digital Image", Global Journal of Computer Science and Technology, 10(2), 2010.

[4] http://www.mathworks.com/products/computer-vision/demos.html?file=/products/demos/shipping/vision/videostabilize_pm.html.

[5] Fischler, MA; Bolles, RC. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography." Comm. of the ACM 24, 1981.

[6] Tordoff, B; Murray, DW. "Guided sampling and consensus for motion estimation." 7th European Conference on Computer Vision, 2002.

[7] J. Jin, Z. Zhu, and G. Xu. "Digital video sequence stabilization based on 2.5D motion estimation and inertial motion filtering", Real-Time Imaging, 7(4):357–365, 2001.

[8] http://siddhantahuja.wordpress.com/tag/sum-of-squared-differences/

[9] M. Pilu. "Video stabilization as a variation problem and numerical solution with the Viterbi method". In Proceedings of Computer Vision and Pattern Recognition, pp 625–630, 2004.

[10] Aleksandra Shnayderman, Alexander Gusev, and Ahmet M. Eskicioglu "An SVD-Based Grayscale Image Quality Measure for Local and Global Assessment ",IEEE 15(2), 2006.

An Analytical Process of 2D Sonar Sensor Low Altitude Field Mapping for UAV Application

Mohamad Farid Misnan, Nor Hashim Mohd Arshad, Ruhizan Liza Ahmad Shauri, Noorfadzli Abd Razak, Norashikin Mohd Thamrin, Siti Fatimah Mahmud.

Faculty of Electrical Engineering
Universiti Teknologi Mara, UiTM
Shah Alam, Selangor.
mfmjawa@yahoo.com

Abstract— **Analytical processes is performed for calculating point distance using ultrasonic sensors, with the intention to resolve a low-altitude two-dimensional surface mapping using UAV. An experimental structure model was constructed to allow for dynamic one-dimensional movement of the sensor system in a horizontal plane. Sample experimental surfaces are chosen to be scanned by the sensor system and recreation of the two-dimensional profile was implemented using a 16-bit microcontroller unit. The recreated profiles are shown for comparison between the calculated and actual values, which proved to be nearly similar.**

Keywords-sonar; mathematical method;sonar mapping; height-measurement;low altitude;

I. INTRODUCTION

Sonar-based sensors have been well established for distance and range measurement applications [1][2]. Additionally they are also used for area mapping purposes. They could easily fused with different types of sensors to produce better results [4]. Capability of the sonar sensor to adapt with any other model in the market is also a factor it is been use by many researcher [4]. Many researches were done indoor, where height and distance measurement were performed between laboratory walls, floor and ceiling. Quite often sample measurements were taken on known object sizes, inside the known indoor environment parameters.

Construction array of the sensor are needed to get optimum mapped area [1]. Danilo et.al used sonar sensor for indoor mapping in long corridor and they used a rotary sensor which extracted feature in a 360° [2] environment. The rotation mechanism produces unwanted delays which could increase error in the mapped data. The structure used is meant for straight line detection and measurement. Similarly, Vassilis Varveropoulos used 180° sweep of sonar sensor which is updated at every 10° incremental angle. The resulting errors can be as high as 20 cm [3].

In high-altitude mapping application, Xiangming Xiao et.al performed mapping of paddy field by using Moderate Resolution Imaging Spectroradiometer (MODIS) sensor onboard the NASA EOS Terra satellite [5]. The area covered is relatively very large from very high altitude, without emphasizing the locality of the paddy field.

Similarly for Kithsiri Perera et.al, they performed land mapping by using MODIS where the information facility is provided by NASA, for underlining changes in vegetation in Southeastern Sri Lanka that is vital for control of greenhouse gas [6]. The intention is not for measurement and mapping of the covered area.

Some researchers used UAV as a means of transportation to perform mapping application because of its mobility [1]. Wang Feng et.al, performed research in UAV borne Real Time Road Mapping [7]. Wang Feng used an unmanned fixed-wing aircraft with double engine and double generator with flight altitude of 100 m to 4000 m for road mapping information. Wang Feng used camera as image sensor to measure the size of the road and it produced a large error of 1.28 m as compared to the actual measurement [7].

In sonar sensors application, the distance covered between the sensor and object is much smaller. It is important to get closer distance between transmitter and receiver to reduce error and increase accuracy [1]. The sensors described in this paper are similar as described in [1], namely the EZmax where transmitter and receiver are built-in together for greater accuracy [1]. Sonar sensor needs sufficient computational processes to produce accurate estimated distance of the measured height as described in this paper.

The sonar sensor construction and assembly described in this paper is practically different as compared to other researches [2][3]. Two sonar sensors are arranged projecting from a common point outward at some angle. An analytical process is used to get estimation of the height of the surface plane.

This paper focused on analytical process to measure actual height of a two-dimensional surface plane using ultrasonic sonar sensors that could be implemented in UAV for low altitude mapping purposes. Series of instantaneous points on the surface plane are measured dynamically with regards of their distance from the sonar sensor. The gathered points are used to create a proper two-dimensional map using an analytical process. The analytical process is meant to reduce the error and implemented in an embedded controller setup. A constructed structure model for the experimental setup is described.

II. PLATFORM SETUP

A. The structure for sonar sensor

An experimental structure setup is constructed as shown in Figure 1. Two vertical poles have a height of 155 cm and firmly hold a horizontal rail with a length of 160 cm. A set of two sonar sensors are placed on the horizontal rail.
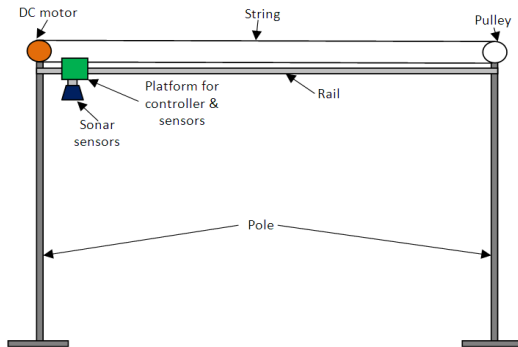

Figure 1. The complete model structure

The two sonar sensors are arranged inclined at about 30° with each other, projected towards the floor, as illustrated in Figure 2.
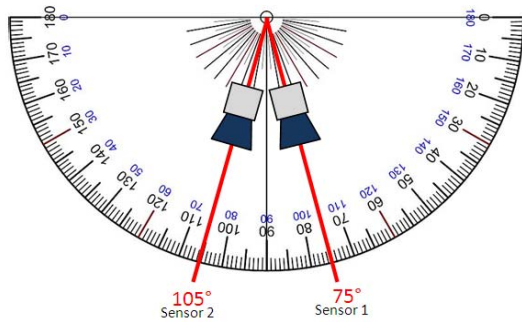

Figure 2. The sensors orientation

The sensors are mounted together with the controller on a platform. A detail actual picture of the sensor on the platform is shown in Figure 3. The platform could freely move along the horizontal rail between the two vertical poles. It is fixed to a closed loop string which is pulled by a DC motor around two pulleys at both ends of the rail. The speed of the DC motor is controlled to produce a linear velocity of about 0.01 m/s for the sensor platform. During experiment the platform moves from left to right to scan the surface profile underneath it.
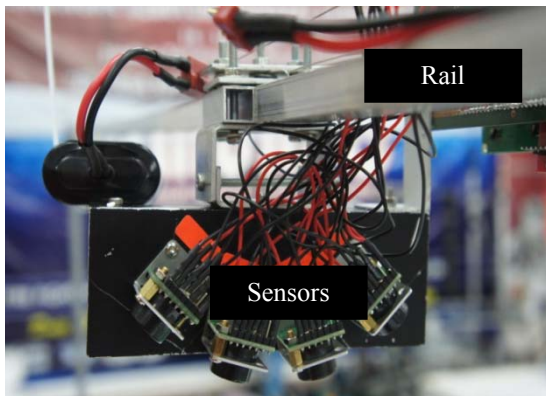

Figure 3. The sonar sensor mounted at the rail.

B. Sonar sensor and controller

The EZmax sonar sensors have the ability to measure a distance of up to 600 cm. In this paper we tested with a height measurement of up to 300 cm, as to enlarge the resolution of the digital controller.

The sonar waves generated converges and scattered and it has a half-beam width angle of 15 degrees [1]. In a research by M. F. Misnan et.al [1], it is known that the size of the resulting wave from sonar sensors are scattered and the area of detection is equal to the arc of radius equal to the measured distance and angle is ± 15 degrees as shown in Figure 4.
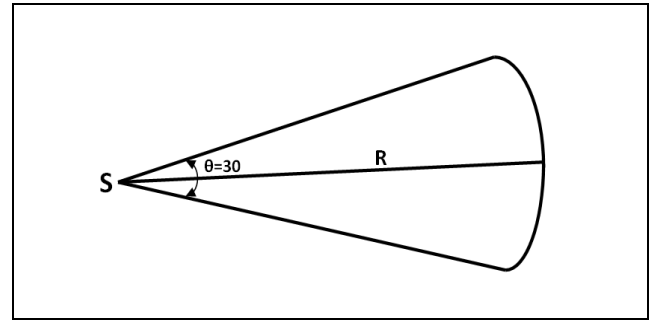

Figure 4. The arc angle of the sonar reflected.

A 16-bit PIC microcontroller (Figure 5) is used to control the DC motor and capturing data from the sensors. Outputs from the two sonar sensors are connected to the built-in ADC in the microcontroller. Equivalent digital values are recorded. Control software is developed using MPLAB IDE with MCC30 compiler.
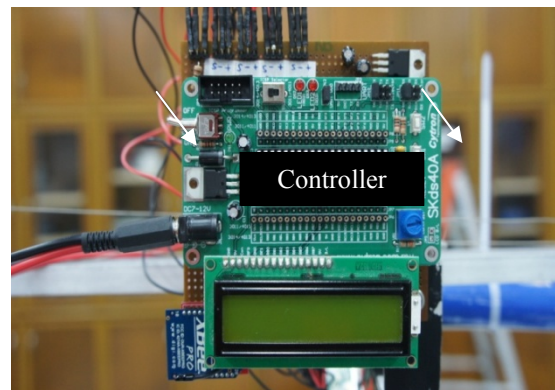
Figure 5. Controller.

III. IMPLEMENTATION AND RESULT

A. Surface profile

Several surface models with different heights are used to provide different height measurements for the mapping process. The test is done indoor. Data from the sonar sensors are recorded for processing and analysis.

B. Controller setting

The controller is programmed to run the sensor platform at a set rate from left to right. As the platform moves, the sonar sensors scan the surface profile underneath it at a fixed sampling rate. The built-in ADC produced equivalent digital values which are recorded and processed. Once the data are processed, they are wirelessly transmitted to a personal computer (PC).

Figure 6.   Block diagram of controller setup

C.   Mathematical model and process

A mathematical process model is needed to obtain actual measurements of height and surface mapping used for mapping applications.

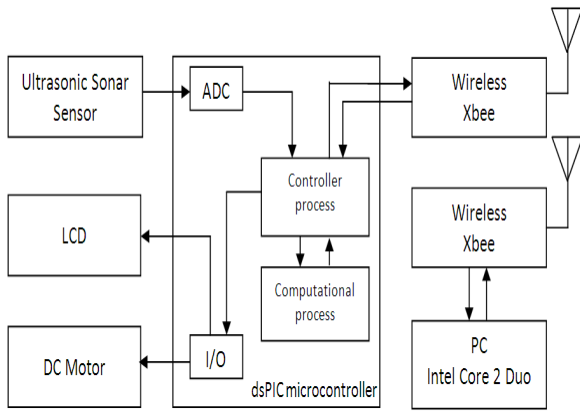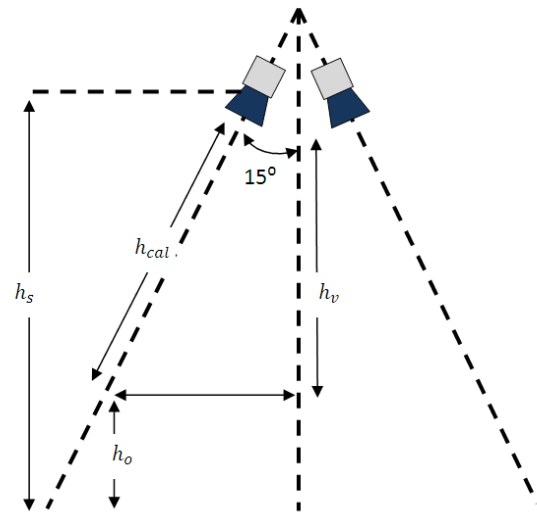Figure 7 shows the flow in processing the data from the sensors. They are converted to the size and actual height of the surface plane. The main process includes logging, digitizing, and converting the height and actual size of the surface plane.
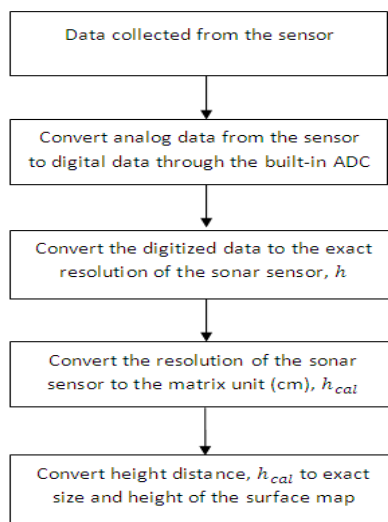

Figure 7.   The analytical and computational process

The direct distance measured from the sensor has been shown and calculated by M. F. Misnan et. al. [1], which is the distance measured is equal to the distance from the sensor to the surface plane. As shown in equation (1) and equation (2), the digitized data from the controller to the height measured in inches and it was converted to the unit matrix. Therefore, the distance between the sonar and the surface is known. The height measurements are shown in figure 8.

Figure 8.   Height measurement

Calculated digitized height in inches, h

$$h = \frac{20}{126}(n - bit)x\frac{20}{21} \qquad (1)$$

The calculated height, h_{cal} converted into unit cm, is shown as in equation (2);

$$h_{cal} = 2.54h \qquad (2)$$

To get the actual size of the heights and map the surface, the height of sonar sensors must be known. As mentioned earlier the sonar sensors are placed at 155 cm above the ground surface. After the direct distance between the sensor and surface plane h_{cal} and elevation of the sonar h_s are known, the vertical height h_v of the surface plane could be calculated.

By using Pythagorean Theorem, the actual height h_m of the sensor can be calculated and predicted based on equation (3).

$$\cos \theta = \frac{h_v}{h_{cal}} \qquad (3)$$

A mathematical process had been done to estimate the vertical height h_v as shown in equation (4).

$$h_v = h_{cal} \times \cos \theta \; ; \quad \theta = 15° \qquad (4)$$

Therefore, to get the surface plane height h_o, equation (5) is used and all data had been gathered to be analyzed and used for mapping.

$$h_o = h_s - h_v \qquad (5)$$

D.   Height and distance

The experimental structure is as set, with horizontal length (x-axis) of 155 cm and vertical height (y-axis) of 155 cm. The structure is as shown in Figure 9.
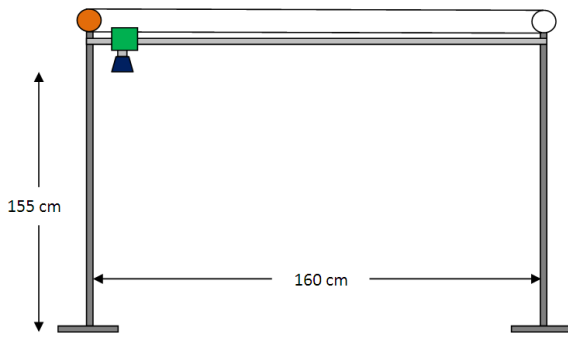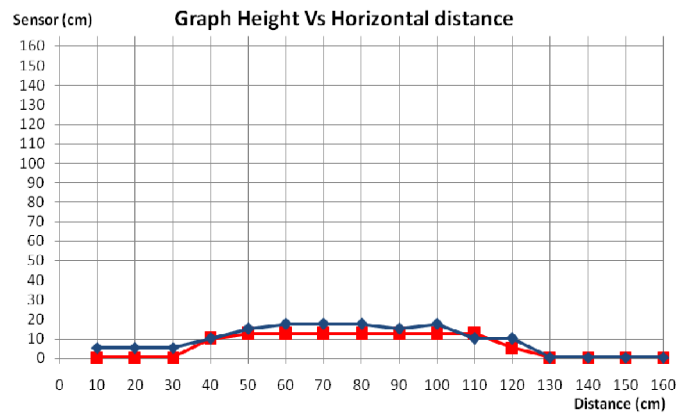
Figure 9.   The x-axis and y-axis structure setup

E.   Data from experiments.

Several different surface planes with different heights were tested using the experimental setup. They were placed under the structure for scanning by the sensors. The corresponding points along the horizontal plane axis were scanned by the sensors and the data gathered were analyzed.

The sensor platform is initially positioned at 0 cm on the horizontal axis (x-axis), which is at the leftmost of the structure. It moves to the right at a fixed speed, whereby points are sampled by the sensors at every 10 cm incremental position on the horizontal plane. The points are recorded appropriately for analytical process.

The first sample surface for the test is a rectangular box having height of 10 cm and length of 50 cm. It is placed 60 cm from the left pole. The detail arrangement is shown in Figure 10. The point values were scanned by the sensors and recorded by the controller, to be sent to the PC. Using the analytical method, the corresponding points were calculated and shown in a graph of height versus horizontal distance, which is shown in Figure 11.
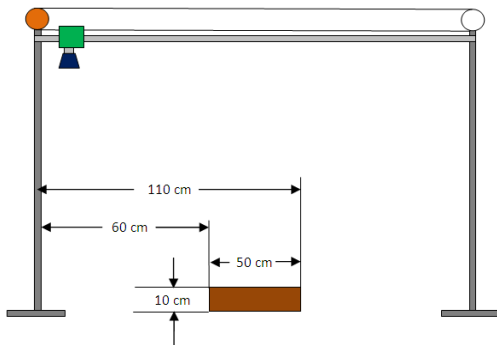
Figure 10. The surface of an object with 10 cm height and 50 cm length which is place between 60 cm and 110 cm at the horizontal axis.

Figure 11. Graph for sonar sensor data on the surface of an object with 10 cm height and 50 cm length.

Next test sample is a combination of three boxes arranged as shown in Figure 12(a) and 12(b). First box is the same box as in the first test with 10 by 50 cm dimension. The other two boxes have a height of 30 cm and length of 50 cm. The first edge of the box arrangement is placed 60 cm from the left pole. The recorded points are calculated using the analytical method and the result is shown in the graph in Figure 13.

(a) Side view

(b) Front view

Figure 12. Surface plane with 10 cm by 50 cm object 30 cm length and 50 cm objects placed left side after the center object.

290
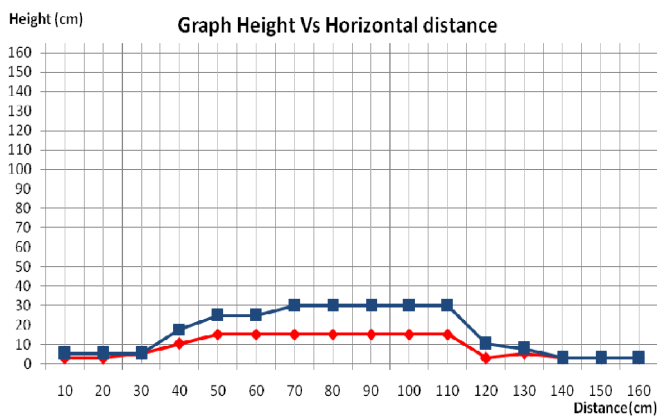
Figure 13. Graph for sonar sensor of the surface with object 10 cm height and 50 cm length is placed at the center of the rail and the object with 30 cm length and 50 cm length placed left side after the center object. Both objects placed between 60 cm and 110 cm at the horizontal axis.

From the two results, comparison can be made on the calculated measurements, especially the height, to the actual surface condition. The actual height of 10 cm is interpreted to an approximately about 15 cm by the calculation (graph 1 in Figure 11). For the second sample, the height of 30 cm is interpreted as almost 30 cm in the calculation (graph 2 in Figure 13).

IV. CONCLUSION

All data from the sensors had been gathered and processed by the analytical method which gave an estimation of the two-dimensional measurement of the sample surface planes. The results have displayed a certain degree of errors, especially on the lower height surface. Nevertheless, the collected and processed points could able to reconstruct the two-dimensional surface plane to a certain degree of accuracy. The 5 cm error projected from the results possibly caused by the limitation in the sonar sensor sensitivity itself. In real application, this accurateness could still be applicable in many different UAV applications which have practical ranges may be up to more than 5 meters. Another aspect need to be observed is the ability of the system to resolve height at points of sharp edges with abrupt changes in height. This naturally is caused by the wedged profile of ultrasonic wave be projected and reflected back.

With this in mind and further improvement on the system could be considered with the additional third-dimensional surface plane. Fusing sonar sensor with sensor types could possibly increase the sensitivity of height measurement.

REFERENCES

[1] Mohamad Farid Misnan, Norhashim Mohd Arshad, Noorfadzli Abd Razak, " Construction sonar sensor model of low altitude field mapping sensors for application on a UAV ", *2012 IEEE 8th International Colloquium on Signal Processing and its Applications (CSPA),* 23-25 March 2012, p.p 446 – 450.

[2] Danilo Navarro, Gines Benet & Milagros Martínez, "Line Based Robot Localization Using a Rotary Sonar", *IEEE Conference on Emerging Technologies and Factory Automation, 2007. ETFA,* 25- 28 Sept, p.p 896 – 899.

[3] Vassilis Varveropoulos,"Robot Localization and Map Construction Using Sonar Data"*, The Rossum Project,* p.p 10.

[4] Lindsay Kleeman and Roman Kuc, "Mobile Robot Sonar for Target Localization and Classification", *International Journal of Robotics Research*. Volume 14, Issue 4, August 1995, Pages 295-318.

[5] Xiangming Xiao, Stephen Boles, Jiyuan Liu, Dafang Zhuang, Steve Frolking, "Mapping paddy rice agriculture in southern China using multi-temporal MODIS images", *Remote Sensing of Environment 95 (2005)* , p.p 480–492

[6] Kithsiri Perera, Kiyoshi Tsuchiya, "Experiment for mapping land cover and it's change in southeastern Sri Lanka utilizing 250 m resolution MODIS imageries", *Advances in Space Research 43 (2009)* , p.p 1349–1355.

[7] Wang Feng, Wu Yundong, Zhang Qiang, "UAV Borne Real-time Road Mapping System", *Urban Remote Sensing Event, 2009,* p.p 1 – 7.

[8] Lindsay Kleeman and Roman Kuc, "Mobile Robot Sonar for Target Localization and Classification", *International Journal of Robotics Research*. Volume 14, Issue 4, August 1995, Pages 295-318.

[9] Hans Jacob S. Feder, John J. Leonard, Chris M. Smith, "Adaptive Concurrent Mapping and Localization Using Sonar", *Proceedings of the 1998 IEEE Inti. Conference on Intelligent Robots and System*, Victoria, B.C., Canada October 1998.

[10] Yu-Cheol Lee, Wonpil Yu, Jong-Hwan Lim, Wan-Kyun Chung and Dong-Woo Cho. (2008). "Sonar Grid Map Based Localization for Autonomous Mobile Robots", *2008 IEEE/ASME International conference on Mechatronics and Embedded Systems and Applications, MESA 2008*.

Analysis of Trusted Identity Based Encryption (IBE-Trust) Protocol for Wireless Sensor Networks

Yusnani Mohd Yussoff , Habibah Hashim, Mohd Dani Baba
Faculty of Electrical Engineering,
Universiti Teknologi MARA
Shah Alam, Selangor Malaysia
yusna233;habib350;mdani074@salam.uitm.edu.my

Abstract — **The peculiarity of Wireless Sensor Networks demands extra consideration during the design of the security protocol. One of the most challenging yet important security features in Wireless Sensor Network is in establishing trusted communication between sensor node and base station. While the term trusted has been widely used referring to valid nodes in the group, this paper discusses the term trusted based on the specifications of Trusted Computing Group (TCG) and presents an IBE-Trust security protocol utilizing well-known identity-based encryption scheme. The protocol incorporates ideas from Trusted Computing Group and Identity-based cryptosystem by Boneh Franklin in ensuring trusted and secured communications between sender and receiver. The proposed protocols were then modeled using the high-level formal language HLPSL and verified using the model checking tool AVISPA. Analysis on the proposed protocols is presented at the end of this paper**

Keywords- Formal Analysis, WSN, Trusted Computing, Security

I. INTRODUCTION

Based on the potential threats in WSNs and the growth of applications related to WSNs, demand for the security in WSNs should be considered at the very beginning of the development which is the design of the physical sensor. Furthermore, due to the eccentric characteristic of WSNs, the designer should also consider the following aspects in their security design; *Constraints faced by the sensor node such as limited power, computational capabilities and memory size; High accessibility of nodes by anybody thus exposing them to intruders and attackers; Information exchanges are easily accessible due to public communication channel and finally difficulties to monitor the condition of the deployed sensor nodes which may be due to its location, mobility and environment.*

While majority of the works done in WSN security have focused on the security of the network [1], our proposed works not only consider the physical design of the sensor node to protect important keys but also software aspects and the design aspects in establishing trusted communication between sensor node and base station with minimum communication overhead [2]. While the first induce extra efforts to thwart physical attacks the second will ensure authenticity, confidentiality and integrity of the data and networks. Besides, our proposed scheme guarantees only trusted sensor nodes can exist in the network.

At this stage, the authors believe that embedding the security parameters in the processor is the most suitable technique for securing wireless sensor node. This technique is believed to be capable of reducing the size of the sensor node, and processing time and preventing software and physical attacks as well as providing other benefits. Johann et al. in his paper [3] also concluded that hardware based security features need to be integrated into the processor to avoid vulnerabilities such as those which exist in today's personal computer. Besides secure implementation, the node also should communicate in a trusted environment. Tiago and Don [4] mentioned that the demand in trusted computing is driven by the potentially severe economic consequences due to unsecured embedded applications.

The following section is organized as follows. Section 2 discussed on physical design of the sensor node. Section 3 briefly discussed the identity based key agreement scheme. Section 4 introduces the proposed security framework followed by validation of the proposed protocol using AVISPA. Section 5 presents the significant benefits in memory utilization for the scheme and finally forming the conclusion in section 6.

II. PHYSICAL DESIGN OF THE SENSOR NODE

According to TCG documents [3], trust is defined as an entity that always behave in the expected manner for the intended purpose. In addition, from the TCG definitions, identity of the entity is equal to measurement. Therefore, for base station to "trust" the sensors that exist in its networks, the sensor nodes need to measure the entities in its platform, produce a measurement value and report the value to the base station. If the value reported is equal with the value measured before deployment, the base station will add the sensor node identity (ID) in the trust list thus enable the sensor node to participate in the networks.

In this process, the Root of Trust (ROT) which is an entity that must be trusted is located in the on-SoC ROM of the ARM11 processor. The integrity (I) of the image that is loaded into it which is the 1st Boot loader is assumed to be unmodifiable and therefore is always TRUE. The sensor node will only be able to complete the secure boot process only when the integrity in each level is true. The integrity of each level of operating system is evaluated by a set of equations. The final level secure boot equation takes into accounts the
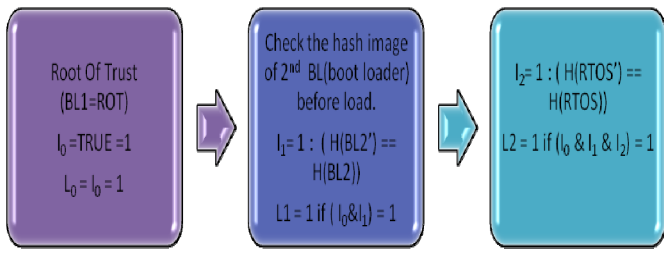
Fig. 1: Secure boot process

integrity of each layer below it, as shown in (1) and presented in Fig. 1. Finally the value generated from the secure boot process will then be used to establish trust relationship with the base station. Details on the secure boot process can be found in [5].

$$I = 1 \text{ if } (I_0 \ \& \ I_1 \ \& \ I_2) = 1 \qquad (1)$$

The main reason for having a secure boot process and secure memory devices is to protect the node from physical type of attacks such as false code injection. While the secure boot process confirms the integrity of the images running in the nodes, the following section discuss communication procedure in confirming secured communication between sensors and base station.

III.   ID-BASED KEY AGREEMENT SCHEME

ID based key agreement scheme is based on an Elliptic Curve Cryptography (ECC) type algorithm. IBE was proposed by Adi Shamir in 1984 and only in 2001, Boneh and Franklin [6] has successfully come out with a fully functional identity-based encryption scheme. IBE has simplified the certificate based public key encryption scheme by using publicly known unique identifiers to derive public keys and eliminate the needs of certificate authority.

In IBE, an arbitrary string is used as a public key. Public key can be calculated from any string such as email, project name or any other string. According to RFC 5408 [7], an IBE public key can be calculated by anyone who has the essential public while a cryptographic secret (master key) is needed to calculate an IBE private key, and the calculation can only be performed by a trusted server that has this secret. In WSN, the trusted authority or trusted entities is the base station which has to be placed in the most secure place and controlled directly by the network proprietor. Besides that, the existence of pre-deployment stage has offer better security and control environment for the key distribution phase. This criterion does not exist in other PKC infrastructure. Another characteristic that differentiates IBE from other server-based cryptography is that no communication is required with the server during encryption operation whereby the sender only needs to know the recipient's ID for it to encrypt the message. Moreover, IBE implementation also consumes less memory for storing public keys of the other nodes.

IV. PROPOSED IBE-TRUST SECURITY FRAMEWORK

The following section will briefly discuss proposed identity based authentication scheme to establish the security goals discussed earlier. The proposed scheme is motivated by trusted platform basic functions that measure and report the integrity of the platform which in this case is the base station [8]. With the nature of WSNs such as distributed, unsupervised and physically exposed, the need for a trusted platform is a must. By integrating trusted platform technique in the proposed security framework, base station can rely on the data it received from the sensor node.

Two basic trusted platform functions that is integrate into the security framework are: *Protected capabilities* that grant the user issuing the access command to protected locations such as in memory and register and *Integrity measurement and reporting*: Process of obtaining metrics of the platform characteristic (hash management value) and reporting to base station.

Simplified version of the protocol is best viewed in Fig. 2. Following section will discuss IBE_Trust protocol and verify the protocol using AVISPA tools. Noted that trusted is established through two different processes which are secure boot and reporting the generated value to base station. Following sections present two cases that we consider in the development of the IBE_Trust protocol followed by the analysis of the protocol using AVISPA tools.

Case 1: Sensor node reports its trust value to base station.

Throughout this analysis we will consider that before deployment, a node public key and common parameters have been saved in each sensor node. Base station also has a list of sensor node's ID and its Hm_A value (value generated after secure boot process before deployment).We also adopt the only intruder model in AVISPA which is Dolev-Yao that is able to perform overhear, intercept, alter or inject any message into the radio communication channel.

Fig. 2 depicts the IBE_Trust model for the Case 1. This process happens immediately upon node deployment at the intended location. The successfully boot up node will report its measurement value (Hm_A) to the trusted authority which in this case in the base station. Processes number 2 and 3 happened in the sensor node and base station will authenticate the sender with the newly Hm_A value.

The protocol using AVISPA syntax is as follows:

A → BS : Snd(A.BS{Na'.Hm_A'}_Kbs)
A: Request BS to authenticate on Hm value
BS : Rcv(A.BS.{Na'.Hm_A'}_Kbs) in (A.Hm_A, trustmap)

** BS compare the received Hm value with the value in trustmap list.

BS→A: Snd({Na'.Nb'}_Ka)

** BS Request A to authenticate on Nb Value

A: Rcv ({Na.Nb'}_Ka)

A → BS: Snd ({Nb'})_Ks
A: Request server to authenticate on Na value

Table 1 describes the notations used in the proposed scheme.

Table 1: Notations used in the proposed scheme

Symbol	Description
ID_A	Identifier of sensor node A
PvK_A	Private key of sensor node A
A.S	Sender ID.Receiver ID
Hm_A'	New trust value (DY) stage
Hm_A	Trust value at PDP stage
N_a', N_b'	Random Nonce
K_s , K_A	Public key BS and A
Snd	Send packet
_Ks	Encrypted packet with K_S public key
M	Message
d(Hm_A)	Decrypt(Hm_A)

As mentioned earlier, our protocol is based on IBE concept and therefore all the packets sent are encrypted with public key of the receiver which is the hash of the receiver ID. In this case, the sensor node, only needs to store the ID of the BS. To confirm confidentiality of the Hm_A value, the packet is encrypted with BS public key. Therefore only BS can decrypt the packet and compare the received value with the value in the list.

Unfortunately, the above processes produce attack on the authentication of the Hm_A value. Instead of BS send the start signal to A, the intruder has initiated the communication and asks A to send the Hm_A value to intruder. Although the secrecy of Hm_A value is not compromised, we have modified the protocol to prevent intruder from initiating the communication.

Fig. 2: IBE-Trust Model

Modified version:

A → BS : Snd(A.BS.{Na'.Hm_A'}_Kbs)
BS : Rcv(A.BS.{Na'.Hm_A'}_Kbs) in (A.Hm_A, trustmap)

**BS compare the receive Hm_A value with the value in trustmap list.

BS→A: Snd({Na'.Nb'}_Ka)

**BS: Request A to authenticate on Nb Value

A: Rcv ({Na.Nb'}_Ka)
A → BS: Snd ({Nb'})_Ks
A: Request server to authenticate on Hm value.

In the modified protocol, BS will only authenticate on the Hm_A value once the authentication on Na and Nb are successful. This protocol reported no intruder attacks and the goal of the protocol which are authentication on Hm_A and secrecy of Hm_A value is accomplished.

Upon successful authentication, BS will generate a new list contain trusted node's identity. This new list, which is smaller than the trust list can be distributed to sensors in its network or stored in base station for faster verification process. Nodes will only communicate with the nodes which its ID exist in the new list. Trusted nodes in the networks remain to be in trusted condition as long as it remains in the ON state. Once reboot or shutdown due to any reason, the nodes will need to re-authenticate with BS. Failure to authenticate will lead to node termination process where the node's ID will be removed from the trust list. All new nodes have to go through the same process before joining the group. Case 2 describes communication between sensor nodes.

Case 2: Communication with immediate neighboring node or Base Station

Two slightly different methods are discussed: The first method performs authentication by letting the sensor node keeps a list of all trusted node IDs in its memory while in the second method receivers checks the authenticity of the sender with the BS. The first leads to extra memory requirement in sensor node and add extra security entities (item need to be protected) and the latter contributes to extra communication overhead.

The packet sent by sender will include the source and destination address, its ID (ID_x, nonce value and encrypted data (encrypt with BS public key). Receiver then verifies the ID_x with the list of ID in his trustlist. If the ID is not valid, the packet will be discarded. In the later technique, instead of checking the trustlist in its memory, receiver will verify the validity of sender ID with the base station. BS will reply to receiver on the status and proceed to the next stage depending on the BS report. The nonce value is used to avoid replay attack in the network. Above protocol using AVISPA syntax is as follows:

(The HLPSL codes are simplified to reduce space).

1st technique:
A: Snd(A.B.IdA{Data.Na'}_Kb
B:Rcv(A.B.IdA{Data.Na'}_Kb/\ in (A.IdA,trustlist)
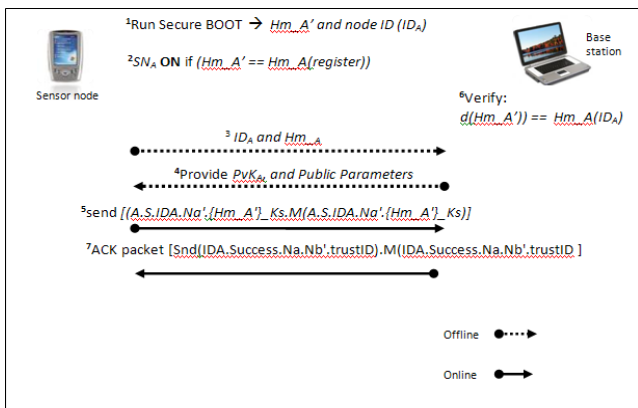B:Snd(B.A.Ok)
A: proceed with next packet

2nd technique:
A: Snd(A.B.IdA{Data.Na'}_Kb
B:Rcv(A.B.IdA{Data.Na'}_Kb
  Snd(B.S.A.IdA)
BS:Rcv(Snd(B.S.A.IdA) /\ in (A.IdA,trustlist) =|>
  Snd(S.B.Ok)
B:Rcv(S.B.Ok) =|> Snd(B.A.Ok)

A: proceed with next packet

It is clear that the 1st technique consumes less communication overhead compared to the 2nd techniques. Both techniques report no attacks on data secrecy and authentication. Further studies on energy and memory utilization are needed to confirm which technique performs better.

V. ANALYSIS ON MEMORY REQUIREMENTS

In this section, we evaluate the memory utilization for storing keys of our IBE-based key management scheme. The key pre-distribution scheme proposed by Eschenauer and Gligor[9], Du et.al[10, 11], and TinyIBE[12] is used for comparison. Although the WSN architecture presented above is between sensor node and base station only, our scheme is flexible to be implemented in hierarchal structure that may consist of base station, sensor node and cluster head.

Terminology used:

M = Number High end sensor (H) which is usually the cluster head

N = Number of Low end sensor (L)

In IBE-Trust scheme, each L and H sensors is preloaded with its private key, ID and equation to derive public key based on the ID. Thus H and L sensor is pre-loaded with a single key only. The total number of pre-loaded keys is:

$$M + N \qquad (2)$$

In the E-G scheme, each sensor is pre-loaded with m keys. The total number of pre-loaded keys in a network with M+N sensors is:

$$m*(M+N) \qquad (3)$$

In the Du et.al scheme (Asymmetric Pre-Distribution), total number of pre-loaded keys is: Where m = number of pre-loaded keys in sensors.

$$xM + yN \text{ where } xy=m^2 \qquad (4)$$

Latest, Du et al. presented a routing driven key-management scheme and the total number of pre-loaded keys is:

$$M(N+3) + 2N \qquad (5)$$

Finally, the TinyIBE total number of keys is fixed at:

$$3M \qquad (6)$$

From the above equations, the E-G and A-P scheme have "m" numbers of pre-loaded keys in each sensor node. Therefore, total numbers of keys in the networks increased linearly with the number of sensor nodes (H and L). This does not happen in IBE implementation where the numbers of keys mainly depends on the number of H and L sensors only. Fig. 3 present the comparison between the describe schemes. The parameters were set as: M = 30, m = 100, x = 500, y = 20. The resulting graph shows that our IBE-Trust scheme gracefully scales with the number of nodes in the network. However, TinyIBE scheme produces a constant number of keys because no private key is stored in the sensor node. Instead they use a session key for communications with primary cluster head and they assume the cluster head to have more battery power compared to L sensors for generating numbers of session key for each L sensor.

We can also see that, the routing driven scheme (ECC) proposed by [11] increase almost the same as the A-P scheme and at L=1500, total number of keys in ECC exceed A-P scheme (not shown in graph). This suggests the need for proper key management or key distribution scheme in IBE implementation to avoid linear increase in total number of keys in the networks.
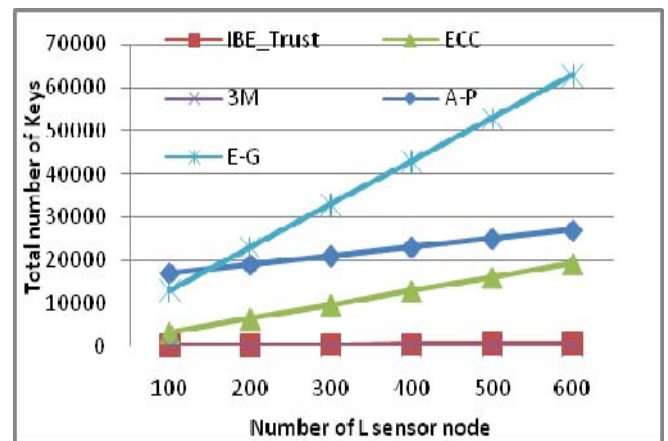
Fig.3: Pre-Loaded Keys in Networks

VI. CONCLUSION

This work proposed a security framework for WSN applications that that need high security features. This might include applications such as in body sensor networks, oil and gas, crucial financial information, un-critical military communications, medical data, and others. This paper focuses on: *confirming the integrity of images on the platform through secure boot process*, *establishing trusted platform status by reporting the hash management value to base station*, and *finally presenting an analysis of the IBE-Trust framework using AVISPA tool to confirm message confidentiality and secure authentication between sensor node and base station*

and between sensors. The trusted authentication mechanism proposed in the IBE_Trust framework will ensure the integrity of the data received in the data centric environment. Future work will be focused on real implementation of IBE_Trust framework and its performance analysis in real environment.

REFERENCES

[1] W. Hu, P. Corke, W. C. Shih *et al.*, "SecFleck: A public key technology platform for wireless sensor networks," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Verlag, 2009, pp. 296-311.

[2] Y. M. Yussoff, and H. Hashim, "Trusted Wireless Sensor Node Platform," in Proceedings of The World Congress on Engineering 2010 London, United Kingdom, 2010, pp. pp774-779.

[3] J. Grobschadl, T. Vejda, and D. Page, "Reassassing the TCG Specifications for Trusted Computing in Mobile Embedded Systems." pp. 84-90.

[4] T. Alves, D. Felton, and ARM, "TrustZone: Integrated Hardware and Software Security," *Technology in-Depth,* 3, 2004].

[5] H. Hashim, Y. M. Yussoff, and L. H. Adnan, "Secure Boot Process for Wireless Sensor Node."

[6] D. Boneh, and M. Franklin, "Idenetity-based encryption from weil pairing," *Advance in cryptology-crypto,* vol. 2139, pp. 29, 2001.

[7] L. Martin, G. Appenzeller, and M. Schertler, "RFC5408 - Identity-Based Encryption Architecture and Supporting," Network working Group, 2009.

[8] D. Challener, K. Yoder, R. Catherman *et al.*, *A practical Guide to Trusted Computing*: IBM Press, 2008.

[9] L. Eschenauer, and V. D. Gligor, "A key-management scheme for distributed sensor networks," in Proceedings of the 9th ACM conference on Computer and communications security, Washington, DC, USA, 2002.

[10] X. Du, Y. Xiao, M. Guizani *et al.*, "An effective key management scheme for heterogeneous sensor networks," *Ad Hoc Networks,* vol. 5, no. 1, pp. 24-34, 2007.

[11] D. Xiaojiang, X. Yang, C. Song *et al.*, "A Routing-Driven Key Management Scheme for Heterogeneous Sensor Networks." pp. 3407-3412.

[12] P. Szczechowiak, and M. Collier, "TinyIBE: Identity-based encryption for heterogeneous sensor networks." pp. 319-354.

Brainwave Sub-band Power Ratio Characteristics in Intelligence Assessment

A. H. Jahidin, M. N. Taib, N. Md Tahir, M. S. A. Megat Ali, S. Lias, N. Fuad, and W. R. W. Omar

Faculty of Electrical Engineering
Universiti Teknologi MARA
Shah Alam, Malaysia
aisyah23@gmail.com

Abstract—This paper discusses on the brainwave sub-band characteristics for different intelligence groups based on electroencephalogram (EEG) power ratio technique. The EEG datasets have been collected from 50 healthy subjects for two sessions; at relaxed, closed eye (CE) state as reference and at cognitively-stimulated state. In the stimulated state, subjects need to answer the intelligence quotient (IQ) test based on Raven's Standard Progressive Matrices (RPM). Sub-band power ratio from the two sessions were calculated and further analyzed to observe the pattern among different IQ groups. The results show that by implementing power ratio technique, the pattern of IQ groups, especially in the relaxed state can be clearly observed. It can be concluded that the value for alpha ratio is higher for high IQ group compared to low IQ group. In contrast to beta and theta ratio where high IQ groups have lower value compared to the low IQ group. This indicates that the ESD ratios can discriminate the characteristic of brainwaves for intelligence assessment.

Keywords-human intelligence; EEG; intelligence quotient; energy spectral density; power ratio

I. INTRODUCTION

EEG is a neurophysiological measurement of electrical activities of the brain [1]. EEG recorded in the absence of an external stimulus is called spontaneous EEG; EEG generated as a response to external or internal stimulus is called an event-related potential (ERP). The amplitude of a normal subject in the awake state recorded with the scalp electrodes is 10-100 µV [2]. In general, brain waves are commonly sinusoidal [3]. The distinguishable features that can be observed with specific activities or phenomenon manifesting itself in the form of varying amplitudes and frequencies range [4]. EEG frequencies can be categorized into four major bands, which include delta-δ (0.5–4 Hz), theta-θ (4–8 Hz), alpha-α (8–13 Hz) and beta-β (13–30 Hz) [5]. Recent studies have proven that at specific brain states, certain frequency bands may become more dominant [6].

Individual differences in intelligence are usually measured using psychometric tests such as intelligence quotient (IQ), mental rotation test and many others. These tests cover cognitive domains such as reasoning, processing speed, executive function, memory and spatial ability. Although cognitive domains are sometimes considered to be independent, differential psychology has firmly established that they are not: people who perform well in one domain also tend to perform well in the others. This is recognized in the term 'general intelligence', which is usually designated 'g' [7]. Some individual tests — such as Raven's Progressive Matrices (RPM) test, which is used to assess non-verbal reasoning [7], Binet-Simon scale and Spearman's g-factor [8-9]. The score from the test will be calculate and rank into several IQ level or groups in the specific participant distribution. However this approach has it drawback which is not culture-free [10]. Thus, there are ongoing studies to measure the human intelligence which relates on IQ by using and electroencephalogram (EEG) [11-15].

Higher alpha power which is less mental activity and more cooperation between brain areas are found in high intelligent persons (gifted and intelligent) compared to the average intelligent persons (creative and average) when solving closed problems [16]. Less intelligent subjects have a larger degree of anxiety while solving IQ tests that could affect the results. Positive emotions related to the activation of left hemisphere, while negative emotions related to the activation of right hemisphere. It indicates that larger activation of the brain of the less intelligent subjects cause a larger state of anxiety which source of high arousability [17].

Hence, this paper proposes to study the brainwave characteristics of different IQ groups using power ratio technique. The study focuses on the alpha, beta and theta band waves. The pattern distribution will be analyzed among and within the high, medium and low IQ groups.

II. METHODOLOGY

A. Subjects and Data Collection

The EEG data collection involved 50 volunteers and has been done at Faculty of Electrical Engineering, Universiti Teknologi MARA. The flow diagram of the methods that have been carried out for this research is shown in Fig 1. It consists of review on the related topic of EEG and human intelligence, EEG data collection, pre-processing, feature extraction analysis on brainwave sub-band characteristics.

EEG data were recorded in two sessions. The subjects need to relax with eyes closed for 3 min during the first session. Subsequently in the second session, the EEG will be recorded with the brain stimulated with IQ test for 10 min. In order to reduce body actions, logical thinking and non-verbal task performance based on RPM was employed. The subjects need to answer the IQ test through user-friendly graphical user interface (GUI). The range of the IQ test score is in between 0 to 150.
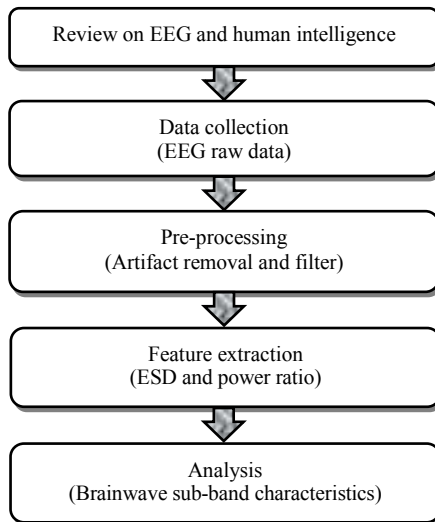

Figure 1. Flow diagram of methodology

EEG signal will be recorded from two paired bipolar recording at prefrontal area of brain region. Channel 1 is connected to the right hand side (RHS), Fp_2, while Channel 2 is connected to the left hand side (LHS) forehead, Fp_1. The reference is placed at the center of forehead, Fp_Z and both ear lobes in agreement with International 10/20 EEG electrode placement standard with sampling rate; 256 Hz. The recorded EEG signals will be processed to obtain the most suitable features as the inputs to the model structure later. The experimental setup for EEG recording is shown in Fig. 2.
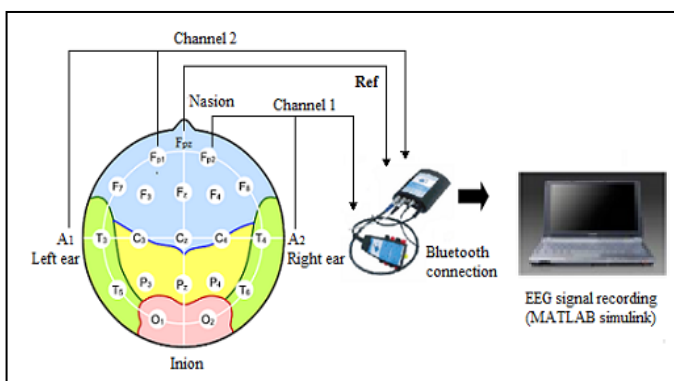

Figure 2. Experiment setup

B. EEG Analysis

In pre-processing stage, EEG raw data were analyzed offline to remove the artifact such as the baseline noise,

powerline interference and eye blink contamination. Initially, non-EEG signals having voltage characteristics of more than 100 µV were removed by deleting the sampled data in all channels [18]. Then it will be filtered into four basic sub-bands; δ, θ, α, and β band. Then, the power spectrum density (PSD) was calculated using Welch technique and Hamming windowing function. The length, NFFT was selected at 1024 with a window size of 256. The window overlaps at 50%.

PSD detects the highest power or energy at peak frequency. On the other hand, energy spectral density (ESD) covers entire energy distribution for each range of the frequency band. Hence ESD has been selected as feature instead of PSD. ESD can be obtained by computing the area of PSD curve, specific to the frequency range of the bands. The ESD versus frequency band can be plotted to validate the power spectrum of ESD.

Power ratio formula as shown by (1), (2) and (3) have been applied to the ESD data. The graphs for two sessions, left and right side of the brain had been plotted. Then, the normality were tested in Statistical Package for Social Science (SPSS) software using Shapiro-Wilk technique. Shapiro-Wilk has been selected due to the small size of samples. In Shapiro-Wilk test, if p-value is less than 0.05 ($p < 0.05$), the data is said significant but not normal distribution.

$$\text{Alpha ratio} = \frac{\alpha}{\alpha + \beta} \tag{1}$$

$$\text{Beta ratio} = \frac{\beta}{\alpha + \beta} \tag{2}$$

$$\text{Theta ratio} = \frac{\theta}{\alpha + \theta} \tag{3}$$

III. RESULTS AND DISCUSSION

A. IQ Groups

The IQ scores from the main data were analyzed offline. It has been grouped into three IQ groups based on the mean and standard deviation calculation. The IQ score's average is 96.9 with standard deviation (SD), 24.8.

Table 1 shows the respective of IQ groups and their range of IQ score. Results show that 78% of the sample were grouped under Index 2 (medium IQ), 12% were grouped under Index 3 (high IQ), and 10% is grouped under Index 1 (low IQ).

TABLE I.   IQ GROUP WITH RANGE OF IQ SCORE

| IQ group/index	Range of score	Subjects
Low - 1	IQ score < 72	5
Medium - 2	72.10<= IQ score <=121.70 (mean – SD and mean + SD)	39
High - 3	IQ score > 121.8	6

B. Distinct Characteristic of CE and during IQ Test

The mean of α, β and θ ratios for (a) closed eye and (b) during IQ test sessions both for left and right side of brain is shown in Fig. 3. In general, mean of α, β and θ ratios gives similar pattern for LHS and RHS side of the brain for the three IQ groups.

In Fig. 3 (a), it is clearly shows that the mean of α ratio decreased from high IQ followed by medium IQ to low IQ groups. In contrast, β and θ ratios gradually increases from high IQ to medium IQ and followed by low IQ groups.
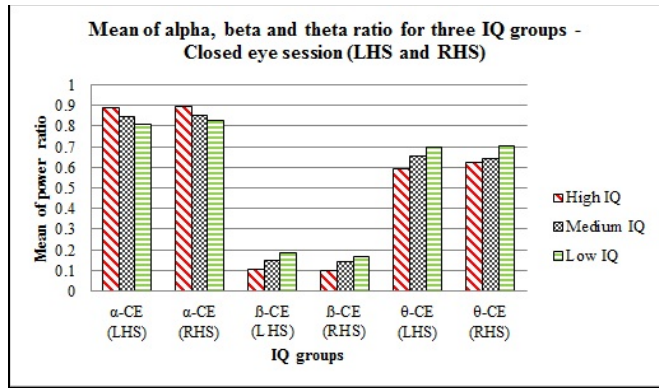
(a) Closed eye session
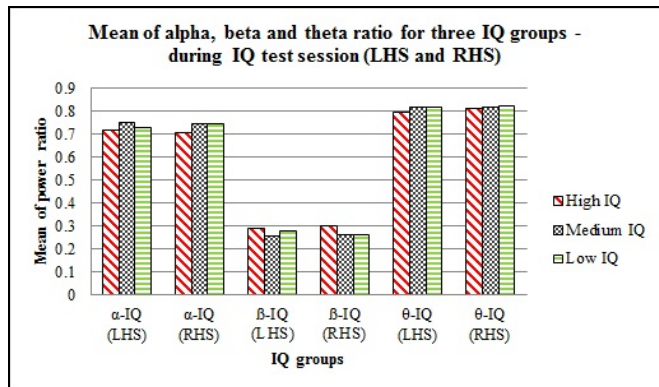
(b) During IQ test session

Figure 3. Mean of alpha, beta and theta ratio for two sessions (a) closed eye and (b) during IQ test sessions

During IQ test session, the pattern of ratio is not consistent for three power ratios as shown in Fig. 3 (b). For α ratio, medium group is the highest followed by low IQ group and lastly high IQ group. In contrast for β ratio, highest value is observed in high IQ group and lowest value in medium IQ group. For theta ratio, high IQ group obtained the lowest value, while the results for medium and low IQ groups look quite similar.

It can also be observed that in CE session, mean for α is the highest followed by θ and then β ratio for three IQ groups as shown in Fig. 3 (a). Fig. 3 (b) shows the mean for θ is highest and β is the lowest.

C. Comparison between CE and during IQ Test Sessions

Fig. 4 shows comparison for mean values of α, β and θ ratios for CE and during IQ test sessions. Comparison between the two sessions on the LHS of the brain indicates that α ratio decreases, while β and θ increase when the subject is doing the IQ test. Similar trend can be observed on the RHS of the brain.
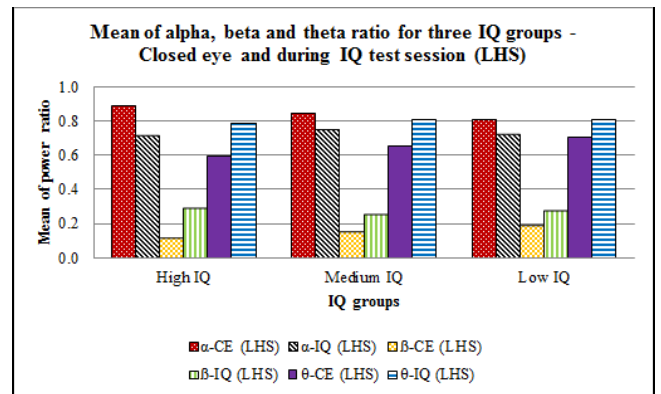

Figure 4. Mean of alpha, beta and theta ratio for two sessions

Table II summarizes the change in cognitive states for each sub-band power ratio. It shows that α ratio for CE is bigger compare to when subject answering IQ test. Yet, β and θ were smaller in CE session compare to during IQ test session.

TABLE II.   SUMMARY OF CHANGES IN COGNITIVE STATES FOR EACH POWER RATIO

Mean of power ratio (LHS & RHS)		
α-ratio	*β-ratio*	*θ-ratio*
Closed eye > During IQ test	Closed eye < During IQ test	Closed eye < During IQ test

D. Normality Test with Shapiro-Wilk Technique

The confidence interval (significant level, *p*) for mean is 95%. The null hypothesis of the test is that the population is normally distributed (if $p > 0.05$), and the alternative hypothesis is that it is not normally distributed (if $p < 0.05$).

Table III shows Shapiro-Wilk normality test of the dependent variables which is sub-band power ratios data for each IQ group, both LHS and RHS in CE session. It can be seen that $p < 0.05$ for sub-band α and β ratios in medium IQ group, therefore these variables are not normally distributed. These are true for both left and right side of the brain. The other dependent variables show $p > 0.05$ which means that the data is normally distributed.

The results for the Shapiro-Wilk normality test during IQ test session are shown in Table IV. The dependent variables are sub-band power ratios for LHS and RHS of each IQ group. Yet again, it indicates that the distribution is not normal for medium IQ group in all sub-band power ratios for LHS and RHS where $p < 0.05$. The other dependent variables are

normally distributed since $p > 0.05$. The pattern of distribution could be caused by the presence of outliers and the data needs to be further analyzed.

TABLE III.   SHAPIRO-WILK TEST OF POWER RATIO BAND FOR EACH IQ GROUP – CLOSED EYE SESSION

Closed eye session		Shapiro-Wilk	
Power ratio band	*IQ group*	*Statistic*	*Sig.*
Alpha ratio - left	low	.826	.130
	medium	.903	.003
	high	.874	.244
Alpha ratio - right	low	.848	.187
	medium	.809	.000
	high	.968	.878
Beta ratio - left	low	.826	.130
	medium	.903	.003
	high	.874	.244
Beta ratio - right	low	.848	.187
	medium	.809	.000
	high	.968	.878
Theta ratio - left	low	.813	.878
	medium	.957	.102
	high	.928	.147
Theta ratio - right	low	.859	.565
	medium	.979	.680
	high	.988	.983

TABLE IV.   SHAPIRO-WILK TEST OF POWER RATIO BAND FOR EACH IQ GROUP – DURING IQ TEST SESSION

During IQ test session		Shapiro-Wilk	
Power ratio band	*IQ group*	*Statistic*	*Sig.*
Alpha ratio - left	low	.907	.448
	medium	.846	.000
	high	.901	.382
Alpha ratio - right	low	.975	.448
	medium	.787	.000
	high	.948	.382
Beta ratio - left	low	.907	.448
	medium	.846	.000
	high	.901	.382
Beta ratio - right	low	.975	.904
	medium	.787	.000
	high	.948	.723
Theta ratio - left	low	.918	.514
	medium	.880	.001
	high	.852	.164
Theta ratio - right	low	.989	.976
	medium	.885	.001
	high	.976	.929

IV.   CONCLUSION

Experimental results revealed that power ratio can be applied as a form of feature to recognize the pattern of IQ groups. The value for α ratio is higher for high IQ group compared to low IQ group. In contrast for β and θ ratio which high IQ group has lower value compared to the low IQ group.

The distribution of data shows that some of the dependent variables are not normally distributed which caused by the outliers. Thus, the data would be further analyzed on the statistical part as future work. A few parameters should be check in normality test. This would include removal of outliers

to obtain better correlation. This part should be considered before can proceed to further analysis to get the best significant result. This is important in order to establish significant pattern unique to specific intelligence (IQ) bands.

REFERENCES

[1] J. W. Clark Jr., "The origin of biopotentials," in Medical Instrumentation: Application and Design, 4th ed., J. G. Webster Ed., New York: Wiley, 2010, pp. 126–188.

[2] K. Blinowska and P. Durka, "Electroencephalography (EEG)," in: Encyclopedia of Biomedical Engineering, John Wiley & Sons, 2006, pp.1–15.

[3] G. T. Henderson, E. C. Ifeachor, H. S. K. Wimalaratna, T. Allen, and R. Hudson, "Prospects for routine detection of dementia using the fractal dimension of the human electroencephalogram," *IEE Proc. Sci. Meas. Technol.*, vol. 147, no. 6, pp. 321–326, Nov. 2000.

[4] R. S. Khandpur, "Handbook of Biomedical Instrumentation," 2nd ed., New Delhi: Tata McGraw-Hill, 2007.

[5] L. Sornmo, and P. Laguna, "Bioelectrical Signal Processing" in Cardiac and Neurological Applications. Burlington, MA: Elsevier Academic Press, 2005.

[6] M. Teplan, "Fundamentals of EEG measurement," *Meas. Sci. Rev.*, vol. 2, no. 2, pp. 1–11, 2002.

[7] I. J. Deary, L. Penke, and W. Johnson, "The neuroscience of human intelligence differences," Nature Reviews Neuroscience, vol. 11, pp. 201-211.

[8] C. Spearman, ""General intelligence" objectively determined and measured," American Journal of Psychology, vol. 15, pp. 201-293, 1904.

[9] A. Binet, "New methods for the diagnosis of the intellectual level of subnormals," L'Annee Psychologique, vol. 12, 1905.

[10] J. L. Andreassi, Psychophysiology: human behavior and physiological response: Lawrence Erlbaum, 2006.

[11] W. Vogel and D. M. Broverman, "Relationship between EEG and test intelligence: A critical review," *Psychophysiology Bulletin,* vol. 62, pp. 132-144, 1964.

[12] W. W. Surwillo, "Human reaction time and period of the EEG in relation to development," *Psychophysiology,* vol. 8, pp. 468-482, 1971.

[13] R. W. Thatcher, D. North, and C. Biver, "EEG and intelligence: Relations between EEG coherence, EEG phase delay and power," *Clinical Neurophysiology,* vol. 116, pp. 2129-2141, 2005.

[14] A. C. Neubauer, R. H. Grabner, A. Fink, and C. Neuper, "Intelligence and neural efficiency: Further evidence of the influence of task content and sex on the brain-IQ relationship," *Cognitive Brain Research,* vol. 25, pp. 217-225, 2005.

[15] A. Fink and A. C. Neubauer, "EEG alpha oscillations during the performance of verbal creativity tasks: Differential effects of sex and verbal intelligence," International Journal of Psychology, vol. 62, pp. 46-53, 2006.

[16] N. Jaušovec, "Differences in Cognitive Processes Between Gifted, Intelligent, Creative, and Average Individuals While Solving Complex Problems: An EEG Study," Intelligence, vol. 28, pp. 213-237, 2000.

[17] U. Fidelman, "Neural transmission-errors, cerebral arousability and hemisphericity: Some relations with intelligence and personality," Kybernetes, Vol. 28 Iss: 6/7, pp.695 – 725, 1999.

[18] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller, "A fully automated correction method of EOG artifacts in EEG recordings," Clinical Neurophysiology, vol. 118, 2007, pp. 98-10.

Analysis of Thresholding Techniques for Asymmetrical Annular Flow Image

Muhamad Suhaidi Umar

Electrical and Electronic Department
UniversitiTeknologi PETRONAS
31750, Tronoh, Perak, Malaysia
suhaidiumar@gmail.com

Idris Ismail

Electrical and Electronic Department
UniversitiTeknologi PETRONAS
31750, Tronoh, Perak, Malaysia
idrisim@petronas.com.my

Abstract — **This paper present the result of a study the behaviour of asymmetrical annular two-phase gas-liquid flow at different gas and liquid flow rate by using several thresholding techniques. The behaviour of annular flow is analysed base on liquid film thickness which is calculated after thresholding process. Thresholding techniques are applied on the images which are captured by Electrical Capacitance Tomography (ECT) in vertical upward flow. The results at the end of this report show which thresholding techniques that able to yield good thresholding images and show the liquid film thickness of annular two-phase flow at different gas and liquid flow rate.**

Keywords - *thresholding; electrical capacitance tomography; annular flow; two-phase flow*

I. INTRODUCTION

Study of two-phase flow is applicable to wide range of applications such as in industrial plant, reactor and others. Two-phase flow in particular gas-liquid is usually discussed [1-2] because by controlling or monitoring gas-liquid flow in pipelines will improve the online measurement, increase quality control and for safety purpose. The flow patterns that can be occurred in vertical upward flow are bubble, slug, churn, wispy annular and annular [3].One of the most common gas-liquid flow patterns is annular flow [4-5]. It is characterized by a thin film of liquid form around the pipe's wall and gas core form at the centre of pipe.

One of the technologies that are capable of examining the liquid in annular flow is electrical capacitance tomography (ECT) [6-9]. ECT is based on non-invasive visualization technique which is capable of obtaining cross-sectional images from dynamic industrial process. Dyakowski *et al.* [8] and Ismail *et al.* [6] conclude that, ECT is preferred because it is flow regime independent, non-intrusive, low cost and minimal maintenance. ECT is developed to reconstruct the dielectric properties after the measurement of electrical capacitance taken between all possible pairs of electrode are done [8]. ECT will produce reconstructed images which contain ofliquid and gas components. The measurement of liquid such as liquid film thickness from those images is complex due to spatial resolution error and non-linearity in the image reconstruction techniques. One of the methods used is thresholding techniques.

In image reconstruction, thresholding technique is needed in order to separate between gas and liquid to two different phases. Thresholding algorithm will change the grey-scale of images which are taken by ECT to binary scale. In binary mode, two different components which are gas and liquid will be separated at the phase border. Previously, Ismail *et al.* used Otsu's threshold method for separating between gas and water [10].

Many thresholding techniques have been developed from last decade until now. In 1978, Ridler and Calvard [11] developed a thresholding algorithm by using iterative technique for choosing the threshold value. The iteration is continued until the difference of threshold's value between present and past is small. Besides that, Otsu [10] calculated the optimal threshold value based on the minimization of the weighted sum of the object and background pixels within-class variances. In 1986, Kittler and Illingworth [12] developed another thresholding technique named as minimum error thresholding. They assumed the images can be characterized by mixture distribution of objects and background pixels. Jawahar *et al.* [13] proposed a thresholding method based on fuzzy clustering and Kapur *et al.* [14] developed a method by applying the concept of entropy to thresholding.

In general, the threshold techniques are divided to two parts: global thresholding and local thresholding. In 1995, Trier and Jain [15] did a review among the thresholding techniques and conclude that the best global thresholding technique is Otsu's method [10] and the best local threshold technique is Niblack's method [16]. In 2004, Sezgin and Sankur [17] also realized that Otsu's method [10] is the best method for global thresholding.

Global thresholding like Otsu's method will show noise if the images have complex background. Then, the better way to solve the problems is to use the local thresholding techniques such as Niblack's method [16]. Niblack's method is a technique that based on the local mean and variance but produced some noise at the image's background. Then, Sauvola *et al.* [18] suggested a factor of R that shown in Eq. (9) to solve Niblack's problem [18]. Sauvola's method is capable of reducing the noise.

TABLE I. SUMMARY OF SEVERAL THRESHOLDING TECHNIQUES

Method	Otsu[10]	Niblack[16]	Sauvola[18]	Kuo (hybrid)[19]
Classification Techniques	histogram	histogram	histogram	histogram
Performance Quality	Global	Local	Local	Global + Local
Advantages	Faster algorithm	Improve Otsu's method	Improve Niblack's method	Improve Otsu's, Niblack's and Sauvola's methods
Drawbacks	Noise for complex background	Noise at the background	Lost some data	Slower algorithm
Complexity	Moderate	Moderate	Moderate	High

After that, Gorman [20] and Liu *et al.* [21] proposed a new technique by combining both global and local thresholding which is called as hybrid thresholding. The algorithm was done in order to improve and solve the problems that are occurred in global and local thresholding. In 2010, Kuo *et al.* [19] also proposed the hybrid thresholding technique for grey-scale images.

Table 1 is a summary of thresholding techniques that are used in this report. These thresholding techniques are applied to asymmetrical annular two-phase flow image. Then, thresholding techniques that are capable of showing good threshold image are used to calculate liquid film thickness in order to study of behaviour of liquid and gas at different gas and liquid flow rate

II. THRESHOLDING ALGORITHM

A. Otsu's Method

Otsu's method [10] is a global thresholding and histogram-based technique. Otsu chose the value of threshold by minimising the value of within-class variance, $\sigma_B{}^2$ of the whole image. Let, the level of image's grey-scale denoted by $[1,2,...,L]$ and the number of pixels at the level i is represented by n_i. Therefore, the total pixels in the images will be equal to $N = n_1 + n_2 + \cdots + n_L$.

The probabilities of histogram of the images will be:

$$p_i = {n_i}/{N} \qquad (1)$$

where $\qquad p_i \geq 0$ and $\qquad \sum_{i=1}^{L} p_i = 1$

Now, the image is divided to two classes; object, X_1 and background, X_2 at the level, y. Therefore, X_1 are the pixels that have value of level $[1,...,y]$ and X_2 is the pixels that have value of level $[y+1,...,L]$. The total probabilities for each class are represented in (2) and (3). Then, the class mean for both classes are represented in (4) and (5):

$$q_1 = \sum_{i=1}^{y} p_i \qquad (2)$$

$$q_2 = \sum_{i=y+1}^{L} p_i \qquad (3)$$

$$\mu_1 = \sum_{i=1}^{y} {ip_i}/{q_1} \qquad (4)$$

$$\mu_2 = \sum_{i=y+1}^{L} {ip_i}/{q_2} \qquad (5)$$

The classes' variances are given by:

$$\sigma_1{}^2 = \sum_{i=1}^{y} \frac{(i-\mu_1)^2 p_i}{q_1} \qquad (6)$$

$$\sigma_2{}^2 = \sum_{i=y+1}^{L} \frac{(i-\mu_2)^2 p_i}{q_2} \qquad (7)$$

The value of class probabilities, q, class mean, μ and class variance, σ^2 changed when the level y was changed. Therefore, Otsu find the level y that make the within-class variance, $\sigma_B{}^2$ is minimize by using (8):

$$\sigma_B{}^2(y) = q_1(y)\sigma_1{}^2(y) + q_2(y)\,\sigma_2{}^2(y) \qquad (8)$$

B. Sauvola's Method

Sauvola's method [18] improved and modified Niblack's method by adding a factor of R where R is maximum value of standard deviation, σ in the image. The equation is used in Sauvola's method shown in Eq. (9):

$$T(x,y) = \mu(x,y)\left[1 + k\left(\frac{\sigma(x,y)}{R} - 1\right)\right] \qquad (9)$$

where k is a bias factor between 0 and 1 and the value of k in this report is 0.8.

C. Hybrid Algorithm

Hybrid's thresholding was proposed by Gorman [20] and Liu *et al.* [21] by combining the advantages of global and local thresholding. In 2010, Kuo *et al.* [19] combined Otsu's method which is global thresholding with Niblack and Sauvola's methods which are local thresholding. The objective of this algorithm is to retain the picture information after binarization.

Firstly, hybrid's method will obtain the threshold value from global thresholding which is denoted by T_{global}. Then, the value of T_{global} will be used in local thresholding. The image is divided to sub-images or window, W by using size $m \times n$. Then number of pixels that have grey value less than T_{global} in W is represented as N. Then, the method choose another value from image's histogram and denoted by S where value of S is less than T_{global} and the selection of S is to make the number of pixels falling between S and T_{global} be a quarter of N.

Thus, the target pixel, p in window W will be threshold by using (10):

$$b(p) = \begin{cases} 1 & , I(p) > T_{global} \\ Local\ Thresholding & , T_{global} \geq I(p) > S \\ 0 & , I(p) \leq S \end{cases} \quad (10)$$

Local thresholding in (10) is used when the grey level of the pixels, $I(p)$ within W is between $, T_{global}$ and S. Local thresholding algorithm in (11) will calculate T_{local}:

$$T_{local} = \begin{cases} \mu_{local} - \dfrac{\sigma}{T_{global}} & , |I(p) - \mu_{local}| \geq \dfrac{T_{global}}{16} \quad (11.1) \\ \mu_{local} \times \left[1 + k\left(\dfrac{\sigma}{T_{global}} - 1\right)\right] & , |I(p) - \mu_{local}| < \dfrac{T_{global}}{16} \quad (11.2) \end{cases}$$

Where k is a preset constant, μ_{local} is the average grey value of all pixels in window, W and σ is local standard deviation of the grey value of pixels in window, W. Note that, equation (11.1) and (11.2) are Niblack's and Sauvola's method respectively. The local standard deviation equation is calculated by using (12):

$$\sigma = \sqrt{\dfrac{\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(I[W(i,j)] - \mu_{local})^2}{m \times n}} \quad (12)$$

Once the value of local threshold, T_{local} is obtained, thresholding is using (13):

$$b(p) = \begin{cases} 1 & , I(p) > T_{local} \\ 0 & , I(p) \leq T_{local} \end{cases} \quad (13)$$

III. EXPERIMENTAL PROCESS

All the algorithms were programmed in MATLAB and a Graphical User Interface (GUI) was built in order to run the experiment process. The experiment will be run at different locations in the pipeline, different thresholding method, different gas flow rate and different liquid flow rate. The annular two-phase images were captured at four different locations; inlet, outlet, throat1 and throat2 as shown in Figure 1 and Table II. Gas and liquid flow rate are used in this experiment are shown in Table III.
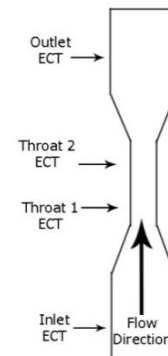

Figure 1. Four locations at pipe for image capturing

TABLE II. INNER DIAMETERS FOR EACH LOCATION IN PIPE

Location	Diameter (mm)
Inlet	32
Outlet	32
Throat1	16
Throat2	16

TABLE III. GAS AND LIQUID FLOW RATE USED IN EXPERIMENT

Gas Volume Flow Rate (L/m)	Liquid Volume Flow Rate (L/m)								
	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
550				X	X	X	X	X	X
600			X	X	X	X	X	X	X
650		X	X	X	X	X	X	X	X
700	X	X	X	X	X	X	X	X	X

Firstly, choose an image at certain conditions and examples of images is shown in Figure 2. After that, threshold method will be chosen for the next process. Then, the chosen image will be threshold by using chosen methods and produce threshold images like show in Figure 3 and 4. Then, liquid film thickness will be calculated. The thickness will be calculated based on the average thickness at 8 locations in the threshold images like shown in Figure 5.
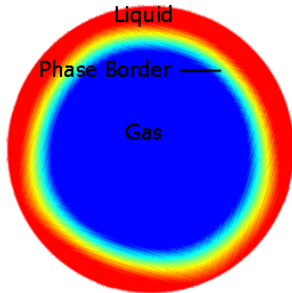

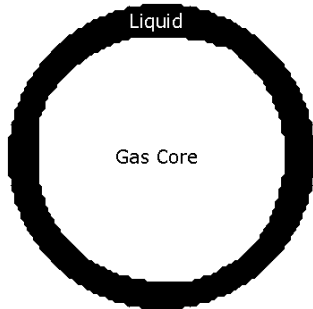Figure 2. Sample of annular two-phase image
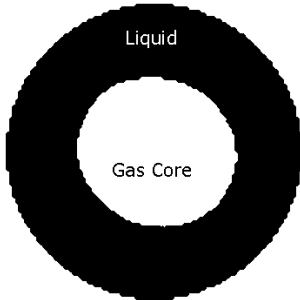

Figure 3. Sample of threshold image in inlet or outlet

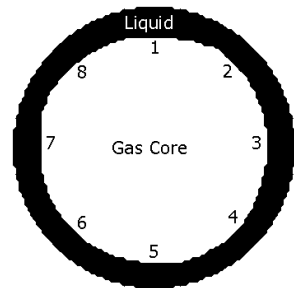Figure 4. Sample of threshold image in throat 1 and throat 2

Figure 5. Eight locations for averaging of liquid thickness
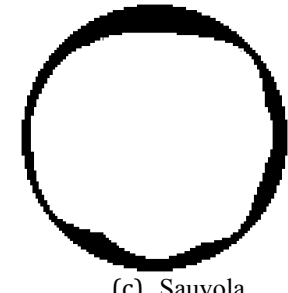
IV. RESULTS AND DISCUSSION

After all thresholding techniques which are Niblack's, Sauvola's, Otsu's and Kuo's methods are applied on the annular two-phase images, they produce different threshold images. Figure 6 show the threshold images from Niblack, Otsu, Sauvola and Kuo.
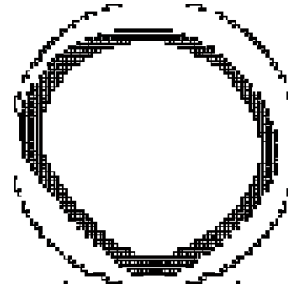
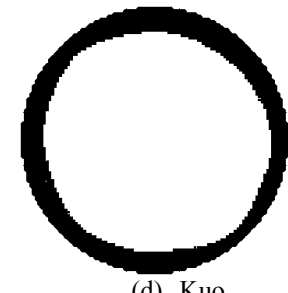(a) Niblack
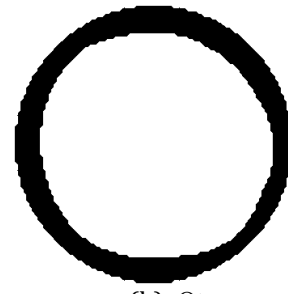
(b) Otsu

(c) Sauvola

(d) Kuo

Figure 6. Threshold image for four different threshold methods. The image was taken at 550L/m and 0.5 L/m

Based on Figure 6(a), Niblack method shows the noise in the image. Therefore, Niblack's method is not acceptable and Otsu's, Sauvola's and Kuo's methods will be used for thresholding of annular two-phase images and benchmarking evaluation.

Based on the experiment, liquid film thickness is increasing when liquid flow rate is increased at a constant gas flow rate. This behaviour gives same results either at inlet, outlet, throat 1 or throat 2 and similar results for all threshold methods. An example of distribution liquid film thickness is shown in Figure 7.
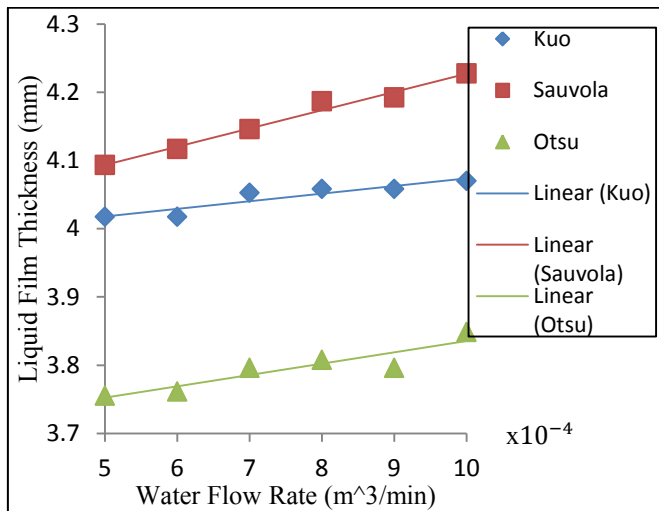

Figure 7. Liquid Film Thickness at 550 L/m of gas flow rate in Throat 1

For the case of constant liquid flow rate, the liquid film thickness is decreasing when gas flow rate increased. Similar behaviour can be observed at inlet, outlet, throat 1 or throat 2 and for all threshold methods. An example of distribution liquid film thickness is shown in Figure 8.
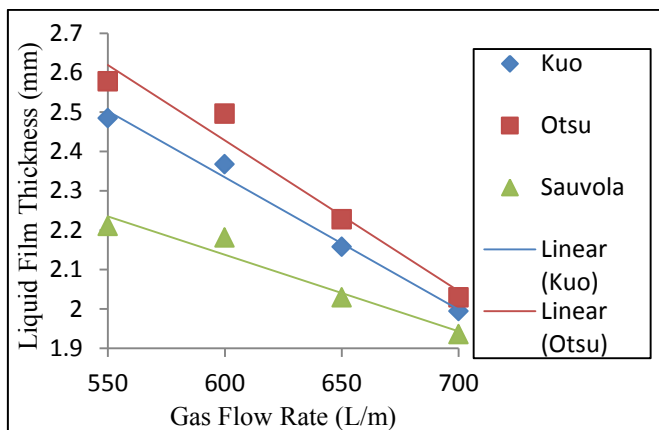

Figure 8. Liquid Film Thickness at 0.6 L/m of liquid flow rate in Inlet

From the same image 7 and 8, we also observed that, the liquid film thickness in throat are greater that liquid thickness in inlet or outlet. It is due to the liquid enters the converging section, there will be more resistance to the flow and it will cause liquid flooding at the inlet.
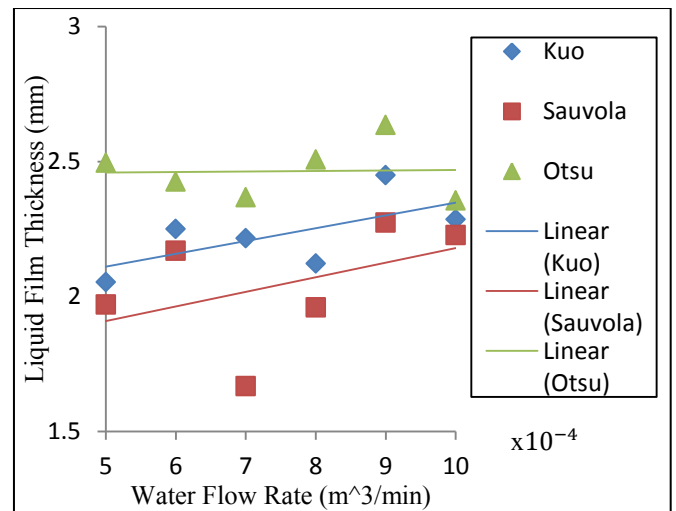

Figure 9. Liquid Film Thickness at 550 L/m of gas flow rate in Outlet

V. CONCLUSION

Asymmetrical annular flow images were captured by using Electrical Capacitance Tomography (ECT) and were analysed by using thresholding techniques which are Niblack's, Otsu's, Sauvola's and Kuo's method. Based on the chosen thresholding techniques, Niblack's method is unable to produce an acceptable post-threshold image. From Figure 6, the Niblack's method will produce noise and inaccurate image. Although the objective is to benchmark Otsu's performance with the more recent techniques, there are minimal differences between the three techniques. Based on the liquid film thickness trending, the film thickness will increase when as the liquid flow rate increased at constant gas flow rate but liquid film thickness will decrease when gas flow rate be increased. This is validated against the observed liquid film thickness during the experiment. Thus Otsu's with a simpler algorithm and implement would be a preferred technique for ECT image thresholding.

REFERENCES

[1] Y. Liu, J. Cui and W. Z. Li, "A two-phase, two-component model for vertical upward gas–liquid annular flow," *International Journal of Heat and Fluid Flow*, vol. 32, pp. 796–804, 2011.

[2] B. N. Kishore and S. Jayanti, "Amultidimensional model for annular gas–liquid flow," *Chemical Engineering Science*, vol. 59, pp. 3577 – 3589, 2004.

[3] J. R. Thome, "Two-phase flow patterns," in *Enginnering Data Book*.: Wolverine Tube Inc., 2010, ch. 12.

[4] N. S. Hall-Taylor, G. F. Hewitt and P. M. C. Lacey, "The motion and frequency of large disturbance waves in annular two-phase flow of air–water mixtures.," *Chemical Engineering Science*, vol. 18, pp. 537–552, 1963.

[5] M. Ishii and K. Mishima, "Droplet entrainment correlation in annular two phase flow," *International Journal Heat Mass Transfer*, vol. 32, pp. 1835–1846, 1989.

[6] I. Ismail, J. C. Gamio, S. F. A. Bukharia, W. Q. Yang, "Tomography for multi-phase flow measurement in the oil industry," *Flow Measurement and Instrumentation*, vol. 16, pp. 145–155, 2005.

[7] I. Ismail, "Measurement of wet gas flow and other two-phase processes in oil industry using Electrical Capacitance Tomogprahy," The University of Manchester, Manchester, 2009.

[8] T. Dyakowski, L. F.C. Jeanmeure and A. J. Jaworski, "Applications of electrical tomography for gas–solids and liquid–solids flow - A review," *Powder Technology*, vol. 112, pp. 174–192, 2000.

[9] Z. Huang, B. Wang, and H. Li, "Application of Electrical Capacitance Tomography to the void Fraction Measurement of Two Phase Flow," *IEEE Transaction on Instrument and Measurement*, vol. 52, no. 1, pp. 7-12, 2003.

[10] N. Otsu, "A threshold system method from gray-level histogram," *Systems, Man, Cybernectics*, vol. 19, pp. 62-66, 1979.

[11] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Transaction on System, Man, and Cybernetics*, vol. 8, no. 8, pp. 630-632, 1978.

[12] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognition*, vol. 19, no. 1, 1986.

[13] C. V. Jawahar, P. K. Biswas, and A. K. Ray, "Investigations on fuzzy thresholding based on fuzzy clustering," *Pattern Recognition*, vol. 30, no. 10, pp. 1605-1613, 1997.

[14] J. N. Kapur, P. K. Sahoo., and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Graph Model Image Processing*, vol. 29, pp. 273-285, 1985.

[15] O. D. Trier and A. K. Jain, "Goal-directed Evaluation of binarization methods," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 1191-1201, 1995.

[16] W. Niblack, "An introduction to digital image processing," *Prentice Hall*, pp. 115-116, 1986.

[17] M. Sezgin and B. Sankur, "Survey over image thresholding Technique and quantitive performance evaluation," *Journal of Electronic Imaging*, vol. 13, no. 1, pp. 146-168, 2004.

[18] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," *Pattern Recognition*, vol. 33, no. 2, pp. 225-236, 2000.

[19] T. Y. Kuo, Y. Y. Lai and Y. C. Lo, "A novel image binarization method using hybrid thresholding," *IEEE*, 2010.

[20] L. O. Gorman, "Binarization and multithresholding of document images using connectivity," *CVGIP, Graph, Models Image Processing*, vol. 56, no. 6, p. 494, 1994.

[21] Y. Liu and S. N. Srihari, "Document image binarization based on tecture features," *IEEE Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 540-544, 1997.

Investigation on Elman Neural Network for Detection of Cardiomyopathy

M. H. Ahmad Shukri, M. S. A. Megat Ali, M. Z. H. Noor, A. H. Jahidin, M. F. Saaid, and M. Zolkapli

Faculty of Electrical Engineering
Universiti Teknologi MARA
Shah Alam, Malaysia
hilmi.shukri@yahoo.com

Abstract—Deterioration of structure and function of heart muscle is indicative of a degenerative disease known as cardiomyopathy. As a result, the hypertrophic condition of the heart often revealed itself in the form of abnormal sinus rhythm that can be detected via an electrocardiogram (ECG). In order to reduce the risk of misinterpretation by cardiologists, a variety of computational methods have been suggested for automated classification of arrhythmias. This paper proposes to explore Elman neural network for detecting cardiomyopathy. A total of 600 ECG beat samples were acquired from an established online database. Initially, the signals were filtered to eliminate high-frequency interference and perform baseline rectification. Nine time-based descriptors from leads I, II and III were used for training, testing and validation of the network structures. A total of five hidden-node node structures were tested with four different learning algorithms. Results show that all the network structure managed to achieve more than 90% classification accuracy. The fastest convergence was achieved with the Levenberg-Marquardt algorithm with an average of 16 epochs.

Keywords-cardiomyopathy; Elman neural network; accuracy; sensitivity; convergence rate

I.   INTRODUCTION

Cardiomyopathy is a degenerative disease which leads to progressive deterioration in the structure and function of the cardiac muscle. The impairment caused by continued use of damaged and stretching of the hypertrophic heart muscle will lead to arrhythmic conditions, which can be observed through the electrocardiogram. The electrocardiogram is a specialized diagnostic tool that assists cardiologists in interpreting the heart condition. The electrical signal is representative of the myocardial activities throughout the cardiac cycle [1]. Hence, any abnormalities can then be benchmarked with the standard tracing. In the event of cardiac disease or abnormalities, the lead system allows cardiologists to locate the exact site of damage. Therefore enabling localized treatment and improving efficiency. However, visual observation by cardiologists imposes a risk of misinterpretation from the electrical tracings. Hence to overcome the problem, a more accurate and computerized approach would be highly desired [2].

Various techniques have been tested in designing an automated system for classifying cardiac diseases such as the premature ventricular contractions (PVC) [3] and bundle branch block arrhythmias [4]. Classification techniques ranging from the simple yet effective linear discriminant classifiers and k-nearest neighbour approach [5], to complex techniques such as the artificial neural network has been employed. Neural network have received continuous attention with the feedforward architecture being the most conventional form being used as performance benchmark [6-8]. Advancements in artificial intelligent techniques has garnered renewed interest in the field with the introduction of innovative concepts such as the radial basis function (RBF) neural network [6], self-organizing maps (SOM) [9], probability neural network (PNN) [6], recurrent neural network (RNN) [10], and support vector machines (SVM) [11]. In addition, combination with other forms of intelligent theory result in hybrid structures such as the neuro-fuzzy approach [12].

The study aims to investigate the performance of Elman neural network for detection of cardiomyopathy. Variations in the number of hidden nodes and learning algorithm implementations have also been observed. The performance of the proposed network is compared with the conventional feedforward neural network. Parameters such as classification accuracy, sensitivity, positive predictivity and training speed have been employed for benchmarking purposes.

The paper is organized in five distinct sections. Section II describes on the signal processing approach adopted for high-frequency noise removal and baseline correction, as well as feature extraction. Section III explains on the theoretical aspects of the Elman neural network topology. Section IV highlights on the learning algorithms. Then Section V explains on the performance indices. The results obtained for the different network structures and algorithm implementations are discussed in Section VI. Finally, Section VII concludes the finding of the study.

II.   SIGNAL PREPROCESSING

A.   Noise Removal

ECG signals for healthy, cardiomyopathy and a combination of other arrhythmic conditions were acquired from the online PTB Diagnostic ECG Database. The signals have been sampled at a frequency of 1 kHz. All signal processing work and intelligent classification tasks were carried out using MATLAB 2010b. At the initial stage of the study, the signals

which are prone to external interference has to be filtered from noises from electromyogram and motion artifacts, 50 Hz powerline interference, as well as baseline wander [13].

High-frequency noise removal was carried out in two stages. The first stage involves application of median filter to remove the high-frequency content. A filter order of 21 was selected since it provides the best noise reduction effect on the entire signal. The median filter is particularly effective in removing sharp overshoots which resulted from motion artifacts. The second stage uses a 10th-order moving average filter to enhance the smoothening effect. The median and moving average filters can each be represented by (1) and (2)

$$y_m = \text{median}\{y_{i-N},...,y_i,...,y_{i+N}\}. \tag{1}$$

$$y_a = \frac{1}{M}\sum_{m=0}^{M-1} y_m. \tag{2}$$

where $2N+1$ represents the order of the median filter, while M, represents the order of the moving average filter.
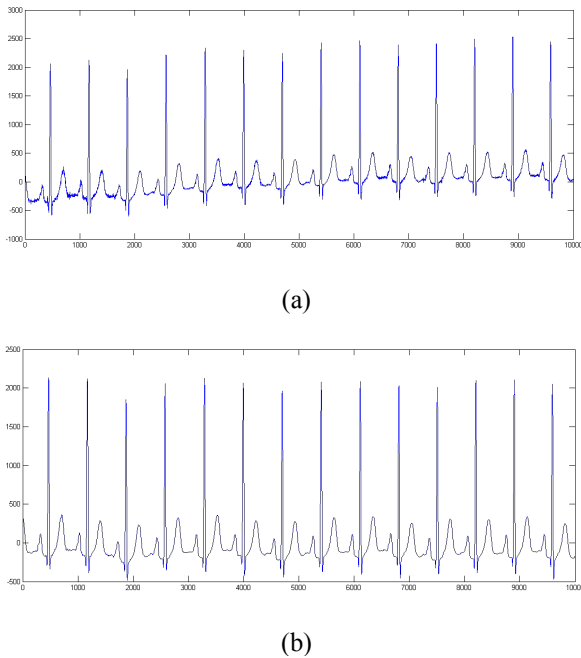

(a)

(b)

Figure 1. (a) Original 10 s data segment (b) Pre-processed waveform.

The last stage of the preprocessing involves rectification of the wandering ECG baseline using the two-stage polynomial fitting approach. Initially, a 4th-order polynomial was approximated and fitted to the entire signal, which reduces the baseline difference. Then the data were divided into segments of 10 s data. 2nd-stage approximation was performed for every smaller section of the data to obtain the 2nd-order polynomials. The solutions were then fitted to further eliminate the difference between the central tendencies of the signal to the zero-baseline. Fig. 1 shows the raw ECG and the preprocessed signal.

B. ECG Descriptors

200 ECG beat samples each have been obtained for healthy, cardiomyopathy and other arrhythmias. The study emphasizes on implementation of time-based descriptors as input to the network structure. Hence, the duration of P-wave, QRS-complex, and T-wave were measured from all bipolar limb leads using the median limit technique. The algorithm sets the output for any amplitude value exceeding the median threshold. Meanwhile, any amplitude value that is lower than the median is being reset. The resultant wave duration can then be obtained through a point-to-point inspection.

III. ELMAN NEURAL NETWORK

The study sets out to investigate Elman neural network for detecting cardiomyopathy. The network is built up of an input and output layer, as well as hidden layers. The number of hidden layers may vary according to different applications. However, one hidden layer should be adequate for function approximation up to an acceptable accuracy [14]. The defining feature of the Elman neural network is the existence of a context layer [10]. The additional layer allows the current state of the hidden vector variables to be returned as a delayed input element to the ensuing cycle of feedforward mapping process. The modified feedforward architecture in the Elman neural network is as shown in Fig. 2.
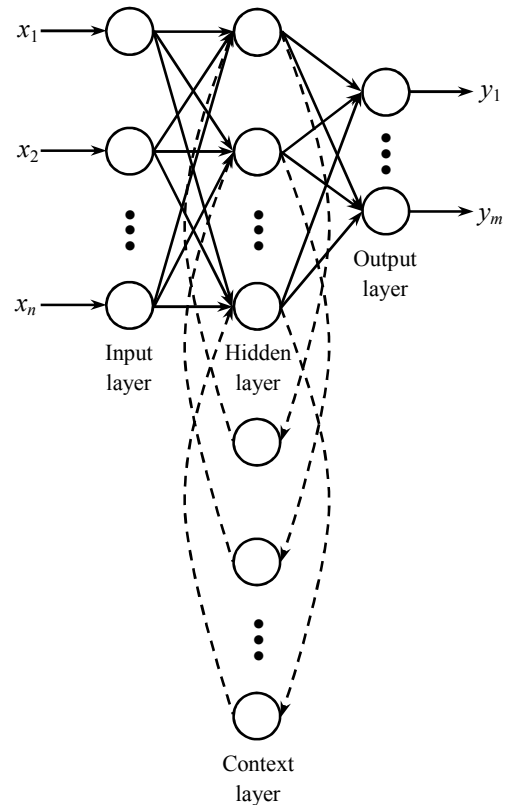

Figure 2. Elman neural network architecture.

The descriptors form an input vector, while the target vector functions as the learning outcomes. The descriptors will be mapped to a set of hidden variables u_j, by an activation

function, φ_1. By considering the presence of the context layer, the intermediate variable u_j of the jth hidden node can be represented by (3) where w_{ij} is the weighted connection between the ith input node to the jth hidden node, and u_j' is the delayed input from the previous forward propagation iteration. The biases in the hidden layer are represented by θ_j.

$$u_j = \varphi_1 \left[\sum_{i=1}^{M} w_{ij} x_i + \theta_j + u_j' \right]. \tag{3}$$

The hidden variables, u_j are then mapped to the output variables, y_k via an activation, φ_2 function through a similar transformation process. Forward propagation of information from the hidden layer to the output can be expressed by (4), where w_{jk} is the weighted connection between the jth hidden node to the kth output node. θ_k represents the bias in the output layer. For the purpose of the study, hyperbolic tangent has been selected as the activation functions φ_1 and φ_2.

$$y_k = \varphi_2 \left[\sum_{j=1}^{N} w_{jk} u_j + \theta_k \right]. \tag{4}$$

The error-connection learning rule is formed via feedback connections from the output to the hidden layer, and from the hidden to the input layer. In network learning, the weight of the connections is adjusted based on the error, e which is the difference obtained between the resultant output and the target vector. e can be defined by (5), where y_t is the target and y_k is the obtained output [15]. The error is then integrated with the learning algorithm to perform the backpropagation weight correction procedure.

$$e = y_t - y_k. \tag{5}$$

The study aims to classify ECG waveforms for three different health conditions. Hence, the target output is described by unique binary sequences. The ECG conditions and its corresponding target are as presented in Table I.

TABLE I.   ECG CONDITIONS AND TARGET OUTPUT

ECG Conditions	Target Output		
Healthy	1	0	0
Cardiomyopathy	0	1	0
Others	0	0	1

IV.   LEARNING TECHNIQUES

Learning algorithms play an important role in the training of any artificial neural network. Implementation of suitable algorithms will speed up convergence rate with better classification accuracy. For the purpose study, four algorithms with different computational approach will be implemented.

These include the gradient descent and the conjugate gradient methods, as well as the Newton and Hessian matrix approximation approach.

The most basic approach in the weight correction procedure is the gradient descent approach. The algorithm can be represented by (6) where w_n represents the weight and biases, while g_n and a_n each represents the gradient and learning rate at the current iteration. The discovery of an adaptive momentum training approach helps to increase the convergence rate by maintaining a steady balance between network stability with a fast learning rate [15].

$$w_{n+1} = w_n - a_n g_n. \tag{6}$$

In order to enhance network training and speed up the learning process is proposed via the conjugate gradient method. As shown by (7), the algorithm searches for the optimum step size with the performance function minimized.

$$w_{n+1} = w_n + a_n p_n. \tag{7}$$

The ensuing searching trend will be conjugate to the preceding iterations. The search procedure is as shown by (8), where several computational variants exist for β_n.

$$p_{n+1} = -g_n + \beta_n p_{n-1}. \tag{8}$$

The Fletcher-Reeves update defines the parameter, β_n as a ratio of the norm squared of the current gradient to the preceding ones. The expression is as presented by (9).

$$\beta_n = \frac{g_n^T g_n}{g_{n-1}^T g_{n-1}}. \tag{9}$$

An alternative to the conjugate gradient approach have been developed through the Newton's method. The algorithm can be expressed by (10) where A^{-1} is the Hessian matrix of the performance index at present weights and biases.

$$w_{n+1} = w_n - A_n^{-1} g_n. \tag{10}$$

The algorithm however is computationally expensive since it would require second-order derivatives of the performance indices to obtain the Hessian matrix. The proposed method has been most successful with the Broyden-Fletcher-Goldfarb-Shanno update procedure.

In order to overcome the computational complexity of the Newton's method, an approximation to the Hessian matrix has been simplified through the Levenberg-Marquardt approach. The algorithm is as shown by (11)

$$w_{n+1} = w_n - \left(J^T J + \mu I \right)^{-1} J^T e. \tag{11}$$

where e is the vector of network errors, J represents the Jacobian matrix, and μ is a constant.

Table II shows summarizes the abbreviation used to denote the learning algorithms used in the study, each representing the gradient descent and conjugate gradient techniques, Newton computational method and the Hessian matrix approximation approach.

TABLE II. LEARNING ALGORITHMS

Learning Technique	Acronym
Gradient descent with variable learning rate algorithm	GDX
Fletcher-Reeves conjugate gradient algorithm	CGF
Broyden-Fletcher-Goldfarb-Shanno algorithm	BFGS
Levenberg-Marquardt algorithm	LM

V. PERFORMANCE METRICS

In order to evaluate the classification performance of the Elman neural network, three performance indices will be used. These include sensitivity (Se), positive predictivity (Pp), and classification accuracy (Acc). The parameters can each be represented by (12), (13) and (14)

$$Se = \frac{T_P}{T_P + F_N} \times 100 . \tag{12}$$

$$Pp = \frac{T_P}{T_P + F_P} \times 100 . \tag{13}$$

$$Acc = \frac{N_T - N_E}{N_T} \times 100 . \tag{14}$$

where variable T_P denotes the number of correctly classified sample beats, while F_N represents the falsely classified ones. N_T represents the total number of beats, while N_E is the total number of errors.

VI. RESULTS AND DISCUSSION

70% from the 600 ECG beat samples have been allocated for training purpose, while another 30% for testing and validation tasks. The network performance for different learning algorithm implementations is as shown in Table III, IV, V and VI. The convergence rate is also indicated as a number of epochs.

Results show that the overall performance of the Elman neural network with different implementation in the number of hidden nodes and learning algorithms are acceptable, with all tests achieving more than 90% classification accuracy. Similar results can also be observed for sensitivity and positive predictivity. As for performance comparison, network training

using the Levenberg-Marquardt algorithm achieves the best results at an average of 99.2% classification accuracy, followed by Broyden-Fletcher-Goldfarb-Shanno algorithm at an average of 97.9%. Both the gradient descent and conjugate gradient algorithms managed to achieve an average classification accuracy of 96.9%.

TABLE III. NETWORK PERFORMANCE WITH GDX ALGORITHM

Node	Healthy		CM		Others		Acc	Epoch
	Se	Pp	Se	Pp	Se	Pp		
10	99.5	99.0	98.5	99.5	98.5	98.0	98.8	217
20	100	97.1	99.0	98.5	95.5	99.0	98.2	195
30	100	93.0	99.5	95.2	87.5	99.4	95.7	123
40	100	93.9	99.5	94.8	88.0	99.4	95.8	127
50	100	93.5	99.5	96.1	89.0	99.4	96.2	115

TABLE IV. NETWORK PERFORMANCE WITH CGF ALGORITHM

Node	Healthy		CM		Others		Acc	Epoch
	Se	Pp	Se	Pp	Se	Pp		
10	99.5	97.5	99.0	99.0	96.5	98.5	98.3	26
20	97.5	98.5	82.5	89.7	91.5	83.9	90.5	63
30	100	97.1	99.0	98.5	95.5	99.0	98.2	25
40	99.5	97.1	99.5	98.5	95.5	99.0	98.2	7
50	99.5	98.5	100	99.5	98.0	99.5	99.2	40

TABLE V. NETWORK PERFORMANCE WITH BFGS ALGORITHM

Node	Healthy		CM		Others		Acc	Epoch
	Se	Pp	Se	Pp	Se	Pp		
10	99.0	98.0	99.5	98.0	96.0	98.5	98.2	39
20	97.5	99.0	99.0	98.5	97.5	96.5	98.0	31
30	100	96.6	99.0	99.5	96.0	99.0	98.3	25
40	100	96.2	99.5	95.7	91.5	99.5	97.0	26
50	99.0	98.0	98.5	99.0	97.0	97.5	98.2	39

TABLE VI. NETWORK PERFORMANCE WITH LM ALGORITHM

Node	Healthy		CM		Others		Acc	Epoch
	Se	Pp	Se	Pp	Se	Pp		
10	98.5	98.5	99.0	98.0	97.0	98.0	98.2	24
20	100	97.6	100	99.5	97.0	100	99.0	13
30	100	100	100	99.5	99.5	100	99.8	15
40	99.5	99.5	100	99.5	99.0	99.5	99.5	19
50	100	99.0	100	99.5	98.5	100	99.5	11

A distinguishable difference however, can be observed in the rate of convergence. Results using the Levenberg-Marquardt algorithm converge the fastest an average of 16 epochs, followed by the Broyden-Fletcher-Goldfarb-Shanno

and the conjugate gradient algorithms, both at 32 and 77 epochs respectively. Trainings using the gradient descent method converge the slowest at an average of 155 epochs.

VII. CONCLUSION

Results show that the Elman neural network has the capability to be implemented for detecting cardiomyopathy. Five hidden-node and four learning algorithm implementations have been successfully tested with more than 90% classification accuracy. However, results also indicate that the Levenberg-Marquardt computational algorithm is the more efficient learning method through its fast convergence rate with excellent classification capabilities.

Future work involving the study on Elman neural network will also cover classification of other types of cardiac arrhythmia with a larger variety of features for network training. Performance benchmarking will also be conducted with other prospective neural network classifiers such as the conventional feedforward neural network and the hybrid multilayered perceptron network.

REFERENCES

[1] M. S. A. Megat Ali, C. Z. A. Che Zainal, A. Husman, M. F. Saaid, M. Z. H. Noor, and A. H. Jahidin, "Detection of cardiomyopathy using multilayered perceptron network," in *Proc. 2012 IEEE Int. Colloq. Signal Process. Its Appl.*, Malacca, 2012, pp. 436–440.

[2] M. S. A. Megat Ali, A. H. Jahidin, and A. N. Norali, "Hybrid multilayered perceptron network for classification of bundle branch blocks," in *Proc. 2012 Int. Conf. Biomed. Eng.*, Penang, 2012, pp. 531–535.

[3] A. Ebrahimzadeh, and A. Khazaee, "Detection of premature ventricular contractions using MLP neural networks: A comparative study," *Measurement*, vol. 43, no. 1, pp.103–112, Jan. 2010.

[4] M. S. A. Megat Ali, A. H. Jahidin, A. N. Norali, and M. H. Mat Som, "Classification of bundle branch blocks using multilayered perceptron network," in Proc. 2011 IEEE Int. Conf. Control Syst. Comput. Eng., Penang, 2011, pp. 531–535.

[5] I. Jekova, G. Bortolan, and I. Christov, "Assessment and comparison of different methods for heartbeat classification," *Med. Eng. Phys.*, vol. 30, no. 2, pp. 248–257, Mar. 2008.

[6] A. Ebrahimzadeh, A. Khazaee, and V. Ranaee, "Classification of the electrocardiogram signals using supervised classifiers and efficient features," *Comput. Meth. Prog. Bio.*, vol. 99, no. 2, pp. 179–194, Aug. 2010.

[7] M. Moavenian, and H. Khorrami, "A qualitative comparison of artificial neural networks and support vector machines in ECG arrhythmias classification," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 3088–3093, Apr. 2010.

[8] H. Khorrami, and M. Moavenian, " A comparative study of DWT, CWT and DCT transformations in ECG arrhythmias classification," *Expert Syst. Appl.*, vol 37, no. 8, pp. 5751–5757, Aug. 2010.

[9] A. Gacek, "Preprocessing and analysis of ECG signals - A self-organizing maps approach," *Expert Syst. Appl.*, vol. 38, no. 7, pp. 9008–9013, Jul. 2011.

[10] E. D. Ubeyli, "Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 1192–1199, Mar. 2010.

[11] M. R. Homaeinezhad, S. A. Atyabi, E. Tavakkoli, H. N. Toosi, A. Ghaffari, and R. Ebrahimpour, "ECG arrhythmia recognition via a neuro-SVM-KNN hybrid classifier with virtual QRS image-based geometrical features," *Expert Syst. Appl.*, vol. 39, no. 2, pp. 2047–2058, Feb. 2012.

[12] Y. Ozbay, R. Ceylan, and B. Karlik, "Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 1004–1010, Jan. 2011.

[13] L. Sornmo, and P. Laguna, Bioelectrical Signal Processing in Cardiac and Neurological Applications. Burlington, MA: Elsevier Academic Press, 2005.

[14] K. I. Funahashi, "On the approximate realization of continuous mappings by neural networks," *Neural Networks*, vol. 2, no. 3, pp. 183–192, 1989.

[15] M. Negnevitsky, Artificial Intelligence: A Guide to Intelligent Systems, 2nd ed., Harlow, Essex: Addison Wesley, 2005.

Mitigate the GPS Ionospheric Scintillation during Solar Flare

Norsuzila Ya'acob[1,2], Azwati Azmin[1], Azita Laily Yusof[1], Mohd Tarmizi Ali [1], Azlina Idris[1] and Darmawaty Mohd Ali[1]

[1]Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia.

[2]Institute of Space Science, Universiti Kebangsaan Malaysia,43600 Bangi, Selangor, Malaysia.

norsuzila@salam.uitm.edu.my, azwatiazmin@yahoo.com, azita968@salam.uitm.edu.my,mizi732002@salam.uitm.edu.my, azlina_idr@yahoo.com, darmawaty@salam.uitm.edu.my,

Abstract—**This paper to mitigate the Ionospheric Scintillation effect during solar flare happen on 7th June 2011. This scintillation data recorded by dual-frequency GPS Receiver, GPS ionospheric scintillation and TEC monitored (GISTM) at UKM, Malaysia (2.55°N, 101.46°E). Parameters on this study will focus on the amplitude scintillation in between 18:00 LT until 00:00 LT; 06:00 pm to 12:00 am. Observation showed that after the post sunset until midnight, the scintillation index in strong scintillation level, $S_4 \geq 0.4$.**

Keywords-component: Ionospheric Scintillation; GPS-Receiver

I. INTRODUCTION

There are many disturbance in Ionosphere such as scintillation, travel ionosphere disturbance (TID), Plasma bubble etc. Plasma density depletion region which appears in the equatorial ionosphere is generated in the bottom side of the F region after the sunset of the ionosphere through plasma instability [1]. Accompanying electron density irregularities are formed in the post-sunset period at low magnetic latitudes is inherently unstable due to the gravitational Rayleigh-Taylor (RT) instability processes operating on the steep upward gradient in the bottom side F-region [2].

The low latitude ionosphere suffers due to ionospheric scintillation [3]. This ionospheric scintillation happen from the irregularities arises in the F-layer ionosphere after the late evening and will last for several hours. Scintillation gives rapid amplitude and phase fluctuation for the single and dual frequency GPS receiver. Amplitude scintillation gives more significant effect to phase scintillation over equatorial region. Higher scintillation gives challenge to signal performance and effect degradation to the signal received by the GPS receiver [4].

For the GPS receiver, signal deterioration occurs because of several factors such as by natural, man-made like obstruction from building, Doppler shift, ionospheric scintillation multipath, jamming, evil waveforms and also antenna effects [5]. This degradation of signal performance will impact to the total loss of signal tracking errors. This partial of signal loss in tracking results in geometry degradation which effect to position accuracy [6].

Ionospheric scintillation gives serious threat to the performance of GPS receiver. Therefore, it is a significant mitigation to study the characteristic of the scintillation and consider the factor that effect to the ionospheric scintillation in order to overcome the problem such as signal loss and tracking errors. Several studies before also determine that ionospheric scintillation over the equatorial region varies with the season, local time, geomagnetic activity and the also the 11-year solar cycle. The maximum of solar activity estimate to be reached mid-2013. Adds of energetic particles, solar materials and gravity waves will disrupt the ionosphere region; where this additional energy striking to the magnetosphere and entering the upper atmosphere. Changes of these irregularities are rapid and significant.

The GNSS receivers that used as an ionospheric model no longer compute the reality position, while resulting in biased position estimated. Other solar cycle effects are damage to satellite electronics and increased drag on spacecraft, altering their orbits. This energetic wind also affects the Earth's magnetic field and can cause significant DC ground currents, potentially disrupting local power grids. During the solar maximum of solar cycle 22 (1989), the Hydro-Quebec electric power system failed, resulting in 6 million people in the US and Canada being without power for 9 hours [7].

In this research, the GPS data taken by using GPS Ionospheric and Scintillation Monitor in NovAtel binary format which being observed by Universiti Kebangsaan Malaysia (UKM) station in order mitigate the Ionospheric Scintillation effect during solar flare happen on 7th June 2011

II. IONOSPHERE

The ionosphere is made up with the ionized and uppermost part of the atmosphere and primarily depends on the Sun activity, as reflected by the 11-year solar cycle. Ionosphere ranges from the 50-1000km above the surface of the earth and generally classified into four regions, D, E, F1 and F2 as illustrated in the Fig.1 below. The D region typically extends

from 50-90 km and has a negligible effect on GPS frequencies. The E region typically extends from 90- 140 km and is produced by solar soft x rays and also has a negligible effect on GPS frequencies. The region F1 typically extends from 140-210 km and has a significant impact on GPS frequencies. The regions D, E, and F1 are associated with the daytime ultra violet (UV) ionization and not present at night. The F2 region typically extends from 210-1000 km. F2 region also is the most active region and its give influence to GPS frequencies in maximum level. During the evening hours the lower bound of the ionosphere rises to 200 km above the surface as the magnetosphere turns away from the Sun and fewer solar particles interact with the atmosphere. Radio signals from satellites pass through the ionosphere and experience a propagation delay or a travel time that is different than would occur in a vacuum, due to the presence of free electrons. Most of the electrons are concentrated in the F-layer, at a height of about 400 km.
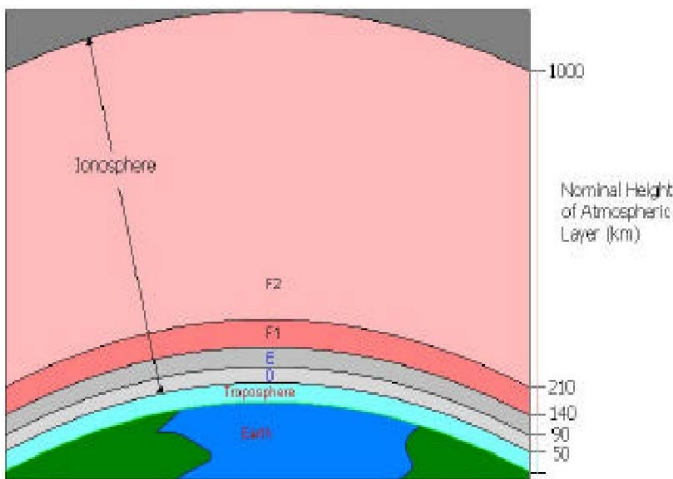
Figure 1. The Ionosphere

III.   DATA AND METHODOLOGY

This analysis motivated to mitigate the amplitude scintillation by classified the amplitude scintillation index in group level when the solar flare occurred.

A.   Data Source

In this research, the data recorded by the dual frequency GPS receiver (GISTM) at UKM, Malaysia during solar flare event on 7th June 2011, auroras on 8, 9th June 2011 and 2 days before solar flare, 5, 6th June 2011 in between 18:00 LT to 00:00 LT; 06:00 pm to 12:00 am. The amplitude scintillation computes by their S_4 value, which is defined as the standard deviation of the signal intensity (SI) that is received from satellites and is divided by its mean value [8]

Parameters on this study will focus on the amplitude scintillation value captured when the solar flare happen and figured out the scintillation index depend on the geographic satellite location extracted from the GPS Ionospheric Scintillation and TEC Monitoring (GISTM) Malaysia data. Amplitude scintillation, S_4 is extracted from GPS receivers and being plot in graph direct from simulation results.

B.   Procedure Process

In this analysis flowchart as shown in Fig. 2 below, the data collected by the GISTM system, then classified into required period. Data only focused in between 18:00 LT until 06:00 LT, whereby, most of the previous study determined that over the equatorial region, the ionospheric scintillation strongly happened after the sunset and last for the several hours until dawn.
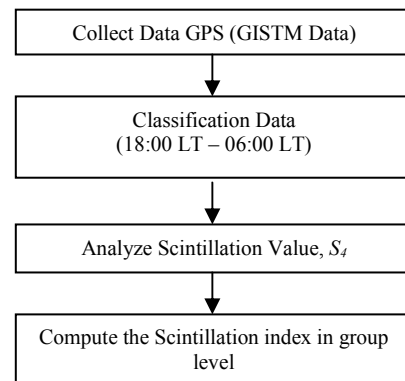

Figure 2. Flowchart in investigate Amplitude Scintillation

The amplitude scintillation can be monitored traditionally accomplished by computing the S_4 index. Sampling rate of amplitude scintillation is 50 Hz and is recorded over 60-second intervals. The GPS receiver also computes two parameters of S_4, which are the total S_4 (S_{4t}) and correction of S_4 (S_{4cor}). The total S_4 (S_{4t}) includes the effect of ambient noise. In the analysis, the corrected S_4 index excluding their influence of ambient noise is computed by using Eq. 1 as below:

$$S_4 = (S_{4t}^2 - S_{4cor}^2)^{1/2} \qquad (1)$$

The result value of S_4 can be categorized into four categories as Table 1 below.

Table 1: Scintillation Index

Scintillation Index	Types of Scintillation
$(0.2 \leq S_4 < 0.3)$	Weak scintillation
$0.3 \leq S_4 < 0.4$	Moderate scintillation
$S_4 \geq 0.4$	Strong scintillation
$S_4 < 0.2$	negligible

IV. RESULTS AND DISCUSSIONS

This study selected for time recorded by GPS Ionospheric Scintillation & TEC Monitoring (GISTM) data monitored at Institute of Space Science (ANGKASA), University Kebangsaan Malaysia (UKM) station, Malaysia. The scintillation of the GPS signals comparing the amplitude scintillation between the data separated into three conditions which is data during Solar flare event, on 7th June 2011 and Auroras on 8th June 2011.

The measured data are first recorded and then converted from binary to ASCII format by specially designed software called *ParseIsmr*. Then, analyze the data and computation of scintillation indices among amplitude scintillation level. All the data taken from GPS signal will be transferred using Matlab graph to analyze the higher value of amplitude scintillation. From the analyzed of amplitude scintillation value, it will be determine each of amplitude scintillation level either in strong, moderate, weak or negligible level. Below is the case study on the amplitude scintillation analysis at all the categories mentioned above.

A. Amplitude Scintillation on 7th June 2011

For the Solar Flare event, data was recorded on 7th June 2011 every minute for the whole day. This analyzed only focuses after the post-sunset which is from 18:00 LT until 00:00 LT; 06:00 pm to 12:00 am. Amplitude scintillation was analyzed for the different Satellite ID, PRN 20 and PRN 13 for time 18:00 LT till 00:00 LT.
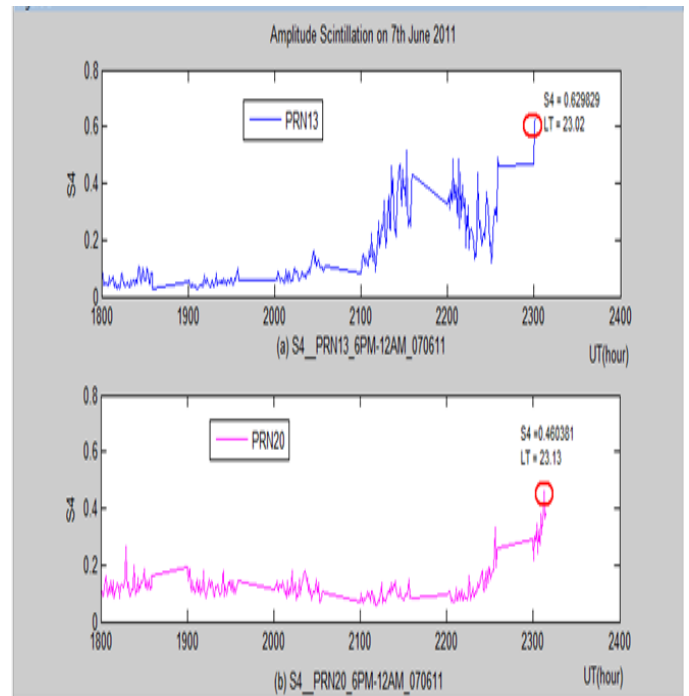

Figure 3. The S_4 at (a) PRN 13, (b) PRN 20 on 7 June 2011

Figure 3 shows graph for the amplitude scintillation during solar flare event recorded by Satellite ID PRN 13 and PRN 20. For the graph (a), we can see that in the beginning of 18:00 LT (6:00 pm) to 21:00 LT (09:00 pm), the scintillation value in range lower than 0.2, which is neglected, or consider no scintillation occurred. After time 21:45 LT (09:45 pm) to 23:02 LT (11:02 pm), the scintillation value increased up to 0.4, which is in strong level. From the graph (b), we can see that in the beginning of evening, 18:00 LT (06:00 pm), amplitude scintillation in range 0.1 to 0.2 which is in weak scintillation level. After late in the night, time 22:30 LT (10:30 pm), the scintillation value increase to range up to 0.4 until time 23:13 LT (11:13 pm). Both of these graph show that after the highest amplitude scintillation occurred, data is loss. Summary amplitude scintillation value shows in Table 2.

Table 2: The summary of amplitude scintillation value on 7th June 2011

Satellite number (PRN)	Amplitude Scintillation (S_4)	Time (LT)	Category
13	0.629829	23.02	Strong
20	0.460381	23.13	Strong

The highest amplitude scintillation recorded by PRN 13 which is S_4 value is 0.629829 at 23:02 LT, 7th June 2011, and followed by PRN20, scintillation value is 0.460381 at 23:13 LT.

This value shows that during solar flare event, amplitude scintillation in strong scintillation level, which is $S_4 \geq 0.4$. From this simulation also shows that the scintillation happens strongly in midnight and the next factor to consider is solar activity. Solar activities tend to suppress the occurrence of bubbles before local midnight and enhance their occurrence after local midnight. During solar maximum and during the peak season, bubbles may occur on 70-80% of the nights.

B. Amplitude Scintillation on 8th June 2011

For the amplitude scintillation on auroras, 8th June 2011, the data focus between time 00:00 LT to 06:00 Lt; 00:00am to 06:00 am. This data recorded to determine the effect of the amplitude scintillation last after solar flare happen recorded by Satellite ID PRN 9 and PRN 27.
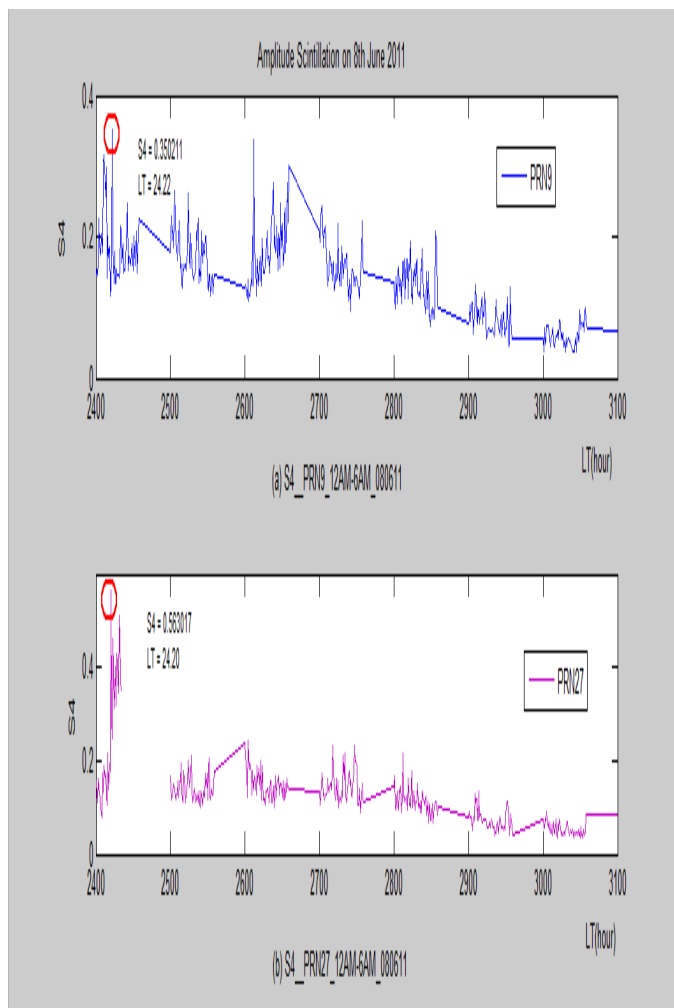
Figure 4. The S_4 at (a) PRN 9, (b) PRN 27 on 8 June 2011

Table 3: The summary of amplitude scintillation value on 7, 8 June 2011

Satellite number (PRN)	Amplitude Scintillation (S_4)	Time (LT)
27	0.563017	00.20
9	0.350211	00.22

According to the figure 4, graph show the amplitude scintillation recorded by PRN 9 and PRN 27. Both of the graphs show that, in the beginning of 00:00 LT (00:00 am), the amplitude scintillation in highest value, since the ionospheric scintillation strongly happen in mid-night and last after the several hours. For the graph (a), the amplitude scintillation last after 03:00 LT and for the graph (b), the signal lost in between time 00:20 LT until 01:00 LT and after that the amplitude scintillation value is neglected. This is because; system satellite for this PRN is still cannot capture the exact value. During this time also, it can assume that, the data has been provide by this PRN inaccurate. The summary of amplitude scintillation value on 7, 8 June 2011 shows in Table 3.

V. CONCLUSION

This study shows about the amplitude scintillation analysis. It shows that higher S_4 happen with value of 0.629829 at 23:02 LT recorded by PRN 13 during solar flare event, on 7th June 2011. Analysis of amplitude scintillation also show that after the sunset, post in midnight mostly all the data recorded with different Satellite ID in strong scintillation level, $S_4 \geq 0.4$. Higher S_4 remain for several hours in post midnight and data is loss when the higher amplitude scintillation occured. The objective of this project to investigate effect of solar activity to amplitude scintillation achieved since the value of S_4 gives worst case for the GPS signal.

REFERENCES

[1] MichiNishioka (nishioka@kugi.kyoto-u.ac.jp) Plasma bubble in the ionosphere, Department of Geophysics, Kyoto University,2005

[2] G.Li, BNin, L.Liu, W.Wan and J.Y.Liu," Effect of magnetic activity on plasma bubbles over equatorial and low –latitude regions in East Asia" 1Beijing National Observatory of Space Environment, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China, 2Institute of Space Science, National Central University, Chung-Li, Taiwan: 19 January 2009

[3] S. Datta-Barua, P. H. Doherty, S. H. Delay, T. Dehel, J. A. Klobuchar, "Ionospheric Scintillation Effects on Single and Dual Frequency GPS Positioning".

[4] Aramesh Seif, Mardina Abdullah, Alina Marie Hasbi, Yuhua Zou, "Observation of GPS Ionospheric Scintillation at UKM, Malaysia," IEEE International Conference on Space Science and Communication (IconSpace), 12-13 July 2011, Penang, Malaysia.

[5] Jiwon Seo, Todd Walter, Edward Marks, Tsung-Yu Chiou, and Per Enge, "Ionospheric Scintillation Effects on GPS Receivers during Solar Minimum and Maximum, International Beacon Satellite Symposium 2007, 11-15 June 2007, Boston, MA

[6] Glenn MacGougan, Gerard Lachapelle, Rakesh Nayak, Alexander Wang, "OVERVIEW OF GNSS SIGNAL DEGRADATION PHENOMENA", (87-100).

[7] Modul of Solar cycle 24 and its impact on GNSS positioning. GSV GPS Silicon Valley.

[8] "GSV4004B GPS Ionospheric Scintillation & TEC Monitor (GISTM) User's Manual", 200

Design and Development of 'Smart Basket' System for Resource Optimization

M. Z. H. Noor, M. H. A. Razak, M. F. Saaid, M. S. A. Megat Ali, and M. Zolkapli

Faculty of Electrical Engineering
Universiti Teknologi MARA
Shah Alam, Malaysia
adriexaffendi@gmail.com

Abstract—The paper proposes a 'smart basket' system as an alternative solution to optimize management of electrical and water consumption in domestic laundry activities. The device acts as a microcontroller-based digital weighing system, and weighs unwashed clothes prior to the laundry stage. In principle, the weights are measured through the use of force sensitive resistance (FSR) sensor, where the results are being displayed on liquid crystal display (LCD). The system is designed with a keypad to enable user selection of weights. Buzzer is used to alert users in case where the basket has reached the selected weight threshold. Results using the 'smart basket' system were compared with data from Tenaga Nasional Berhad (TNB) and Syarikat Bekalan Air Selangor (SYABAS), indicating that resource consumption can be reduced by as much as 33%.

Keywords-smart basket; laundry activities; optimized resources; microcontroller-based; digital weighing system

I. INTRODUCTION

In recent years, more sophisticated technologies have been designed to assist in daily living. Although having limitations, the advancements in terms of application have become more rapid [1]. A digital weighing scale is a device designed to measure weight of any living or non-living matter such as sand, water, flour, and many others. As a standard system, the digital weighing scale would comprise of input, mathematical model and an output. A sensing unit would feed the weight information as input, then being translated by the transfer function into a measureable unit that is finally being displayed as an output. Various types of sensors can be implemented for weight measurements which include load cell, pressure sensor, strain gauge, etc.

The research sets out with the objective to develop digital weighing system using force sensitive resistance (FSR) sensor. FSR sensor works on the principle where its resistance value changes with the applied pressure or force onto a surface area. The sensor is capable of measuring weight within a range of 10 g to 10 kg. The sensor is integrated with a microcontroller unit to translate the weight information into outputs on liquid crystal display (LCD) and alert buzzer. The analogue information from the FSR sensor is interfaced with the microcontroller unit through the analogue-to-digital converter (ADC). Since the research aims to conserve the necessary resources, the system is designed as a 'smart basket'.

The paper is organized in four distinct sections. Section II details out the methodology on material selection and design approach used in the research. Section III discusses on the results pertaining the performance of the FSR sensor and system output. A comparative brief comparative study is also conducted on the effect of the device usage on resource consumption. Lastly, the findings are concluded in Section IV.

II. METHODOLOGY

A. Design Overview

The design of the 'smart basket' system is based on the concept of digital weighing scale. The scope of research covers both the software and hardware development aspects. The flow of research methodology on prototype development is as shown in Fig. 1.
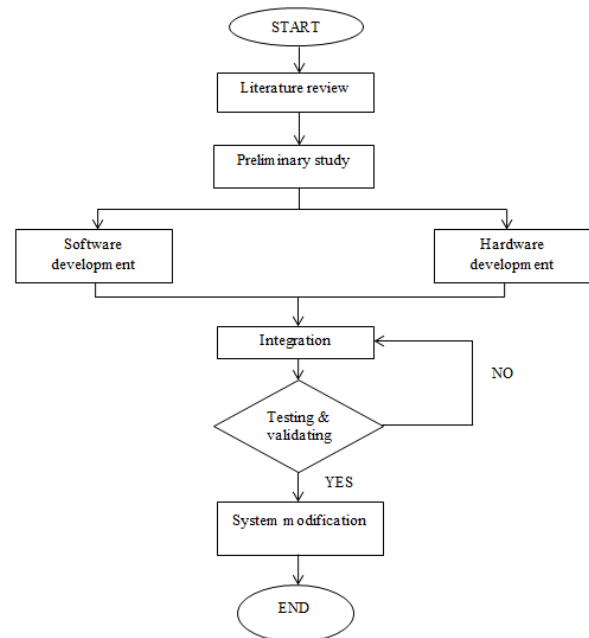

Figure 1. Flow of research methodology on prototype development.

Initially, a preliminary study was conducted on the currently existing digital weighing scales. In this stage,

approaches on concepts of weight measurement using microcontrollers have been reviewed. Then the development stage is split into the programming and hardware development. Selection of suitable sensors and microcontroller units are vital to fulfill the design requirements. Prototype testing follows soon after the completion of both programming and hardware elements. Continual improvement is then performed through functional tests and troubleshooting tasks.
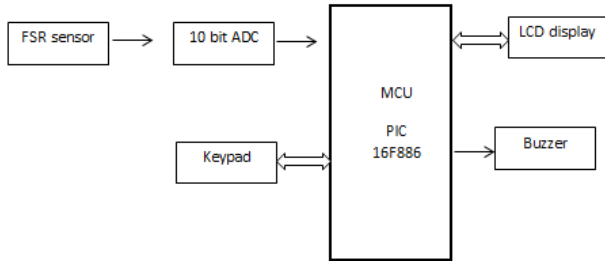

Figure 2.   Block diagram of proposed system.

Fig. 2 shows the general block diagram of the proposed system. All the related components have been considered based on the interaction between the input and output elements. Components involved in the system set-up include the FSR sensor, keypad unit, buzzer and LCD output. For the purpose of the study, PIC16F886 microcontroller has been selected as the information processing unit.
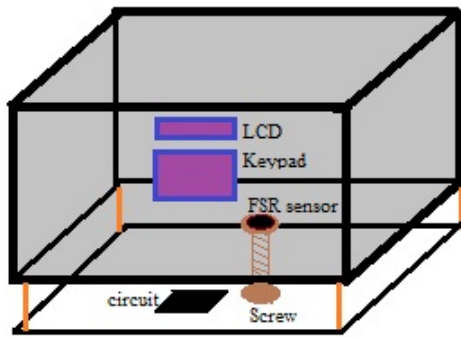
Figure 3.   Mechanical design of 'smart basket' system.

Fig. 3 illustrates the mechanical design aspects of the device. The FSR sensor is placed in the middle of the basket unit with a screw supporting it from collapsing when load is being applied.

B.   Force Sensitive Resistance Sensor

In essence, the FSR sensor is basically a resistor which changes its resistance value, in accordance with pressure that is being applied [2]. The resistance decreases with increasing pressure. The sensor is comprised of two polymer films, one is a conducting surface while the other one is an electrode with a printed surface. Both of the surfaces are separated by a spacer. These surfaces come into contact with each other when force is being applied, resulting in short circuit to the electrode on the

printed surface. Hence the resistance of the FSR sensor is reduced. Fig. 4 shows the three distinct layers of active area, plastic spacer, and conductive film that build up the FSR sensor.
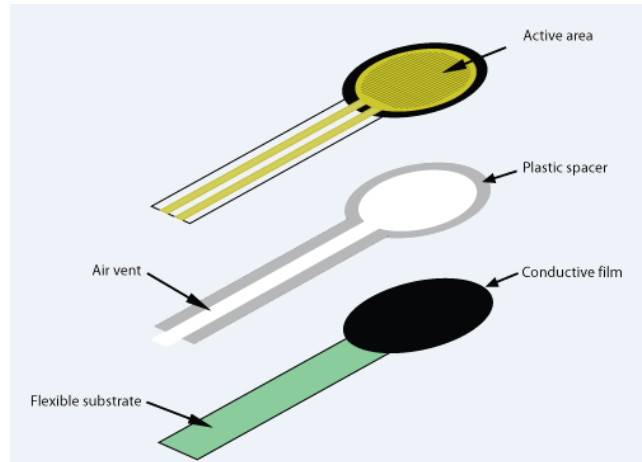

Figure 4.   FSR sensor layout.

C.   Microcontroller Unit

The PIC16F886 microcontroller is a device that has internal memory. In order to control the 'smart basket' system, the microcontroller acts as the central processing unit. The model was selected due to special features which include the capability to control the designed system. There are 28 I/O pins, and an internal oscillator with frequency range of 8MHz to 31 KHz. In addition, PIC16F886 has 368 bytes of RAM, 256 bytes of EEPROM, 10-bit ADC, 16-bit enhanced capture, comparator, 10-bit PWM (10-bit), and an In-Circuit Serial Programming (ICSP) capability [4]. The pin diagram is as shown in Fig. 5.
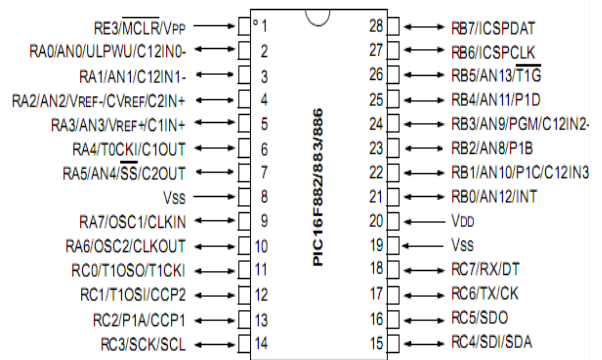

Figure 5.   Pin diagram for PIC16F886 microcontroller.

D.   Firmware and Interfacing

The designed system measures the weight via the FSR sensor and displays the value on LCD. Initially, users can select the weight limit. As load is being increased, the FSR sensor will feed the analogue information into the microcontroller unit

via the ADC. When the measured weight exceeds the limit setting, the buzzer will be triggered to alert the user. The firmware was written in MPLAB IDE. Fig. 6 shows the flowchart for the control firmware.
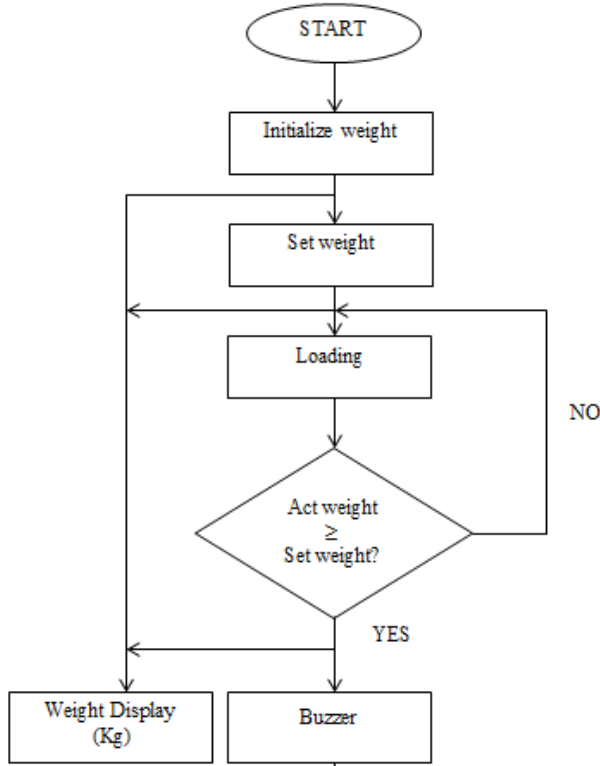

Figure 6.   Flowchart of the control firmware.

In order to provide stable supply voltage to the system, the voltage regulator were used to regulate the voltages between 9 V to 12 V directly from the source. The LM7805 were selected to maintain the voltage output at 5 V. The overall circuit connection is as shown in Fig. 7.
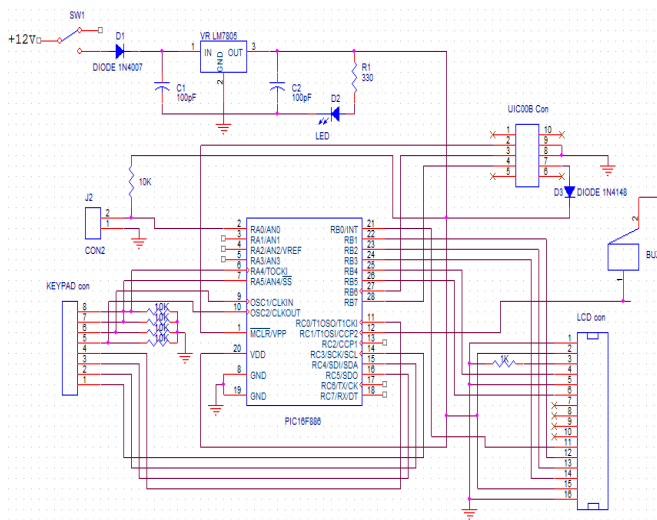
Figure 7.   Circuit connections.

III.   RESULTS AND DISCUSSION

A.   Voltage Drop for Different Weights

Voltage drops were compared for different weights that were measured between 1 kg to 8 kg using multimeter and theoretical calculation as expressed by (1). Data and analysis of the voltage drops are as shown in Table I and Fig. 8.

$$V_{Out} = \frac{R_{FSR}V_{In}}{R_{FSR} + R_M} . \qquad (1)$$

Table 1 shows the relation between weight and voltage drop. The values between the weight and voltages are inversely proportional. An increase in the weight will result in decrease of voltage. The difference between calculated and measured voltage varies by as much as 11.8%. The variation may be induced by external factors imposed by crude fabrication method on the device assembly.

TABLE I.   THEORETICAL AND MEASURED VOLTAGE DROP

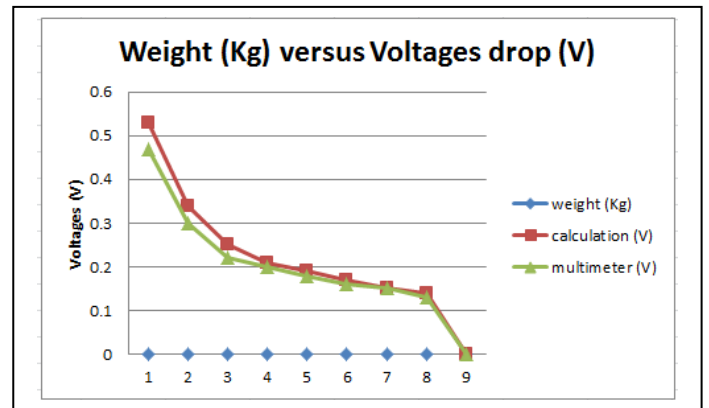| Weight (kg)	Voltage (V)		Difference (%)
	Theoretical	Measured	
1	0.53	0.47	11.3
2	0.34	0.30	11.8
3	0.25	0.22	12
4	0.21	0.20	4.7
5	0.19	0.18	5.3
6	0.17	0.16	5.9
7	0.15	0.15	0
8	0.14	0.13	7.1

Figure 8.   Comparison between theoretical and measured voltage.

B.   Reliability and Functionality Test

In order to test the reliability of the 'smart basket', the fixed weight is measured for five times. From this experiment, the functionality of the buzzer, voltage detection and the LCD was also tested. The test results are as outlined in Table II

TABLE II.   . RELIABILITY AND FUNCTIONALITY TESTS

Weight (kg)		Voltage (V)	Buzzer	LCD (kg)
1	1st	0.48	On	1
	2nd	0.48	On	1
	3rd	0.48	On	1
	4th	0.54	Off	0
	5th	0.49	On	1
2	1st	0.31	On	2
	2nd	0.31	On	2
	3rd	0.31	On	2
	4th	0.31	On	2
	5th	0.48	Off	1
3	1st	0.23	On	3
	2nd	0.27	Off	2
	3rd	0.23	On	3
	4th	0.23	On	3
	5th	0.23	On	3
4	1st	0.19	On	4
	2nd	0.19	On	4
	3rd	0.19	On	4
	4th	0.19	On	4
	5th	0.19	On	4
5	1st	0.17	On	5
	2nd	0.17	On	5
	3rd	0.17	On	5
	4th	0.17	On	5
	5th	0.17	On	5
6	1st	0.16	On	6
	2nd	0.16	On	6
	3rd	0.16	On	6
	4th	0.16	On	6
	5th	0.15	On	7
7	1st	0.15	On	7
	2nd	0.15	On	7
	3rd	0.15	On	7
	4th	0.15	On	7
	5th	0.15	On	7
8	1st	0.13	On	8
	2nd	0.13	On	8
	3rd	0.13	On	8
	4th	0.13	On	8
	5th	0.13	On	8

The aim of the reliability and functionality test is to observe the system whether it is functions according to intended design specifications. This can be observed for fixed weight measurement of 1 kg, 2 kg, 3 kg, and 6 kg loads. The disparity may most likely be caused by the crude mechanical fabrication of the device that affects load distribution on the surface area, excessive surface bending and the sensor performance itself.

C. Comparative Study on Resource Consumption

A comparative study was conducted on the effectiveness of 'smart basket' system on frequency of laundry activity, as well as reducing water and electricity consumption. A study conducted by Qingdao Smad Electric Appliances [6] discovered that for a 7 kg washing machine, the consumption of water and electric usage is 59.7 l and 1.031 kW per cycle. In general, a person would spend three times per week for laundry activities [7]. Based on estimation for one standard family size of four, the weight would increase to 14 kg per week. Implementation of the 'smart basket' system shows that the user can efficiently reduce the need for laundry activities to twice per week. Hence, the electrical and water consumption would be reduced. Table III shows the effect of the system on the frequency of laundry activities and resource usage.

TABLE III.   EFFECTIVENESS OF 'SMART BASKET' SYSTEM

Parameters	Conventional method	'Smart basket' system
Frequency per week (times)	3	2
Water consumption (l)	179.1	119.4
Electric consumption (kW)	3.093	2.062

With reference to Syarikat Bekalan Air Selangor (SYABAS), the water tariff for $1m^3$ is RM 0.57 [8]. Meanwhile, the tariff set by Tenaga Nasional Berhad (TNB) for domestic electric consumption is RM 0.218 per kW [9]. The effectiveness of the device on optimizing the cost of resource consumption is shown in Table IV. Calculation shows that the use of 'smart basket' system is capable of reducing the monthly cost spent on laundry activities by as much as 33%.

TABLE IV.   COST OF MONTHLY RESOURCE CONSUMPTION

Parameters	Conventional method	'Smart basket' system
Weekly water consumption	RM 0.102	RM 0.068
Weekly electric consumption	RM 0.674	RM 0.450
Monthly cost	RM 3.32	RM 2.22

IV.   CONCLUSION

The research has successfully designed and developed the 'smart basket' system. Studies have demonstrated the practicality of the device as a control mechanism in reducing electrical and water consumption, as well as the monthly cost incurred through the laundry activities. Reliability and functionality tests indicate the further improvement is still needed on sensor selection, and mechanical design aspects.

REFERENCES

[1] Yang Chenghui2, College of Electrical   Engineering , Northwest University for Nationalities Lanhou, China 730030 "Circuit Design of Electronic Weighing system", 2010.

[2] "FSR" http://www.ladyada.net/learn/sensors/fsr.html

[3] G. E. Kim, "Force Sensor Resistor," unpublished.

[4] Microchip Technology inc, " PIC16F882/883/884/886 datasheet," 2009.

[5] Interlink Electronics Inc., "FSR Force Sensing resistor Integration Guide and Evaluation Parts Catalog," 2002.

[6] Qingdao Smad Electric Appliances Co., Ltd. Shandong, China (Mainland).

[7] "How to estimate how much water use each day", http://portal.unesco.org

[8] Syarikat Bekalan Air Selangor (SYABAS), "Water Tariff," http://www.syabas.com.my/consumer/water-bill-water-tariff

[9] Tenaga Nasional Berhad (TNB), "Pricing and Tariff," http://www.tnb.com.my/residential/pricing-and-tariff.htm.

[10] U. Hashim, and M. N. Haron, "Design of digital display system for ISFET pH Sensor by Using PIC Microcontroller Unit (MCU)", in Proc. 2009 1st Asia Symp. Quality Elect. Design, Kuala Lumpur, 2009, pp. 148–152.

The Design and Development Of Automatic Fish Feeder System Using PIC Microcontroller

M. Z. H. Noor, A. K. Hussian, M. F. Saaid, M. S. A. M. Ali, M. Zolkapli
Faculty of Electrical Engineering
Universiti Teknologi MARA Malaysia
40450 Shah Alam, Selangor, Malaysia
mzhakim@salam.uitm.edu.my

Abstract— this research devoted to reduce the labor cost as well as develop better pellet dispense system. Subsequently, this research was proposed to design an automatic fish feeder system using PIC microcontroller application. The device developed combines mechanical and electrical system in controlling fish feeding activity. This device, basically consists of pellet storage, former, stand, DC motor and microcontroller. The pellets controlled by DC motor which located under the pellet storage. A control system was then attached to this device allowing the fish to be fed at the right cycle time as required or predefined by user. Timer was employed in this device to control the motor rotation attached to sphere former, which dispense the pellets into the water. The pellets dispensed into the marking area of the pond based solely on the rotation speed of the motor itself. The controller came with a keypad giving user more option in determining the suitable speed for the motor depends on their cattle. In short, the pellets in the automatic fish feeder system will be controlled by the rotation speed of DC motor.

Keyword-PIC Microcontroller; Animal feeder; fish feeder

I. INTRODUCTION

An automatic fish feeder is an electronic gadget or device that has been designed to dispense the right amount of pellets at a particular time. In addition, such system also demonstrated the capability in repeating the task daily and accurately, hence promising efficiency and productivity in fish farming field in long run. In general, two basic concepts which are fixed and mobile conceived the automatic fish feeder [1]. This device fed fish following the right schedule and amount pre-defined by user, therefore avoiding the issue of overfeeding. To date, many of the fish farmers still stick with the manual feeding system. By utilizing the traditional manual feeding system, it means that more work forces would be needed by owner of fish farm in handling certain jobs, particularly in cleaning the feeder, refilling the pellet and even repair or maintenance procedure. All these process consumed considerably more time and energy compare to the automatic fish feeding system. Moreover, the benefits automatic fish feeding system emerged as the areas of the farm increased. For larger area, the manual feeding system users will certainly face difficulty in managing the entire feeding schedule.

Based on previous research, there were some efforts taken in order to replace the inefficient and unproductive manual feeding system. For instance, the Simple Automatic Fish Feeder which employed the timer in dispensing the pellets. However, this system was limited by the ability of dispensing pellets at a constant speed. Next, it was the Automatic feeder for marine mammals [3]. This feeder was found to be highly reliable with a very low rate of breakdowns. Unfortunately, these feeders were not specifically to handle [3]. In addition, this system caused unbalance growth of fish as the dispensed pellets focused at one part only [3].

Discovering the shortcoming of current automatic feeding systems had motivated the research to develop a system which overcome previous systems' flaw, hence giving more advantages and benefits to the owner and workers. As a remedy of dispensing pellets solely at one part, the new system was able to dispense pellets into the desire area based on the rotation speed of the motor, combined with suitable cycle time. The system resulted in more systematic feeding schedule which certainly, will directly decrease the labor cost. This new automatic system was also designed in such a way that it can replace of human activity. This "Automatic Fish Feeder Using PIC Microcontroller Application" offered the user control feeding time up to 24 daily feeding cycles depending on the timer employed in the system and the optional reset time on the feeder. Also, the amount of pellets dispensed depended on the length of the feeding cycle adjusted on the feeder itself [2].

The rotation speed in new developed system will determine the travel distance of pellets before they landed at the desire area. Resistance was strung in the rotor circuit or adjusts the voltages of electrical machinery circuit is one of the traditional method of control speed [4] due to the fact that this method was easily set up, however some shortcomings come with such easy method. Consequently, a new kind of speed control method known as PWM (pulse width modulation) was considered in the new feeding system. Today, such speed regulating system has been widely used in the motor control speed. The power energy can be fully utilized and the circuit efficient is very high with the broad usage of PWM technological [5].

II. METHODOLOGY

There are several ideas and methods required in order to target the research's main objectives. The Figure 1 below aimed to aid researcher in explaining the methodologies of

designing and implementing automatic fish feeder using PIC microcontroller application which encompassed system design, hardware and software development as well as the circuit design.
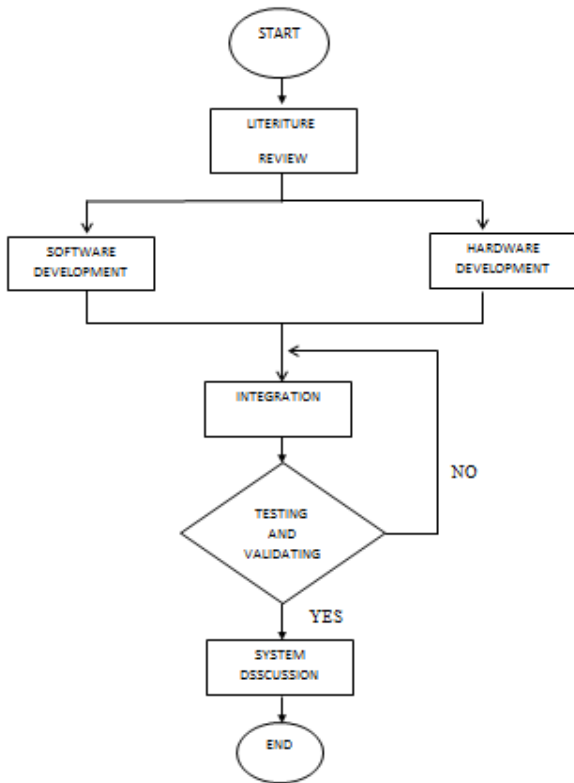

Figure 1: Methodologies of designing Automatic Fish Feeder System.

Figure 1 depicted the flowchart of the procedure conducted. Firstly, the literature review involved studies and collecting information, particularly from previous research. Researcher then reviewed the current issue or development state of automatic feeding system, and made comparison among the researches to identify the flaw of existing systems. Problems or shortcomings were identified and remedies were proposed in in hardware and software development stage. In integration stage, all circuit developed will be combined and integrated with the programming code to initiate the system. The data was initialized and processed by microcontroller on the program with the user interface. Later, the data processed by microcontroller was then sent to device as output of this system akin to user demanded or ordered. The output for this system would be the speed and the torque, generated by the DC motor and control by the microcontroller. Once the system started or failure occurred, researcher proceeds to testing and validating stage. As a result, data collected form system, especially output voltage drop and the distance of the pellet travelled will be discussed in system discussion stage.

A. System Design

The design of this Automatic fish feeder comprised of four main parts, namely main controller, pellet storage, stand and spreader. The controller of this system, a 4x4 Keypad functioned as input device which provided the user abilities to set timer and motor speed to spread the pellet into the water. Apart from that, LCD display played an important role in illustrating the data entered by the user before DC motor start to operate. PIC 16F886 controller was employed as main controller output of DC motor. In order to control the speed rotation of DC motor, L293D motor driver which work by PWM (Pulse Width Modulation) technique was utilized. As for the software design, programming of PIC16F886 was done with the aid of C language. The conceptual diagram as shown in figure 2 where the major components are included:

- PIC16F886 Microcontroller
- 4x4 Keypad
- 2x16 LCD (Liquid Cristal Display)
- L293D (Motor Driver)
- DC Motor
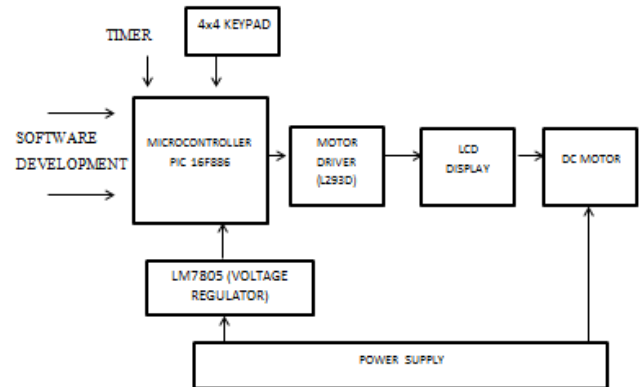- LM7805 (Voltage Regulator)
- Power Supply

Figure 2: Conceptual diagram of the system.

B. Mechanical Design

All appropriate components used in hardware design stage were assembled to set up the feeding device. This system was simple in construction and operation, also relatively inexpensive, as well as reliable in the period of system operation. The storage that attached at the top stored the pellets and the amount decreased across the time as the motor start running. Pellets in the sphere former will be dispensed at target area based on the motor speed. The shape of the former was designed corresponding to sphere shape ensuring the spreading process more efficient compare to the other shape, especially on the amount of pellets stored. It was proven by math that given the same surface area, sphere outperformed other shape in terms of their volume. Meanwhile, during the operation of dc motor, there were several holes that allowed the pellet to be dispensed into the water. Figure 3 below shows the mechanical design of Automatic Fish feeder.
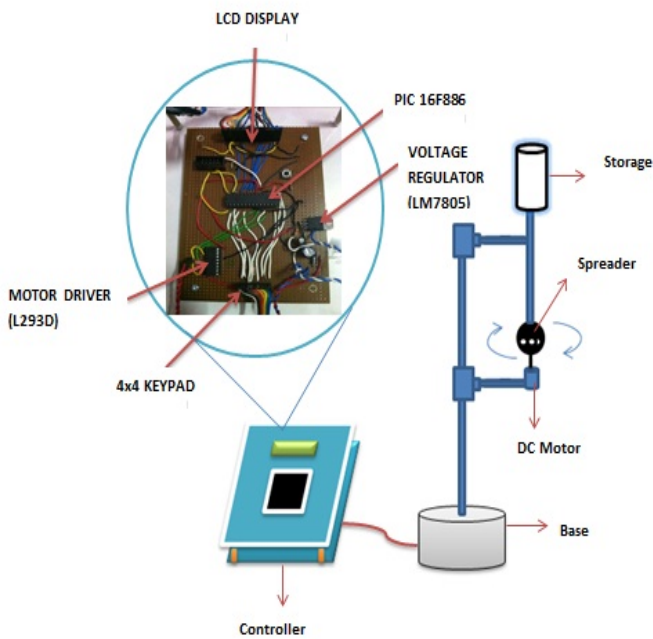
Figure 3: Mechanical Design of the system.

C. L293D motor driver for controlling motor

L293D is designed to provide bidirectional drive current of up to 600mA at voltage from 4.5V to 36V. It is to drive inductive load such as relays, solenoid, DC and bipolar stepping motor as well as other high current or high voltage loads in positive supply application. Each pair of channels is equipped with an enable input to simplify use as two bridges. This device is suitable for the use in switching applications at frequencies up to 5 kHz. For this research, L293D was applied to control the motor speed rotation. Without this component, the motor will not operate as programmed.
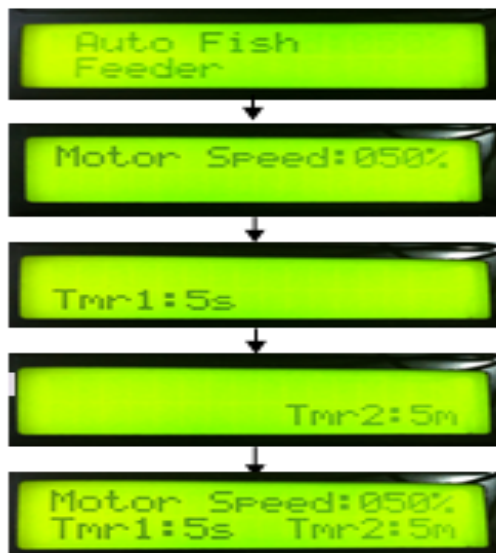
Figure 4: Flow of the system interface.

During the operation, the system would display information as in figure above and ready for the next operation. The speed option screen requested user to set the speed that suitable corresponding to the size or area of the pond itself. There are different speeds of choices depend on the PWM (Pulse Width Modulation) that had been programmed in the microcontroller at the development stage. The motor speed directly proportional to the distance travelled by pellet, hence increasing the motor speed will result in longer travel distance. In the timer case, two types of timers employed in this research. Timer 1 responds to the time entered by user and controls for the feeding time.

This timer will activate the motor at specific time requested by user. Timer 1 came with 2 options that allowed the user to set the time second (s) or minute (m). In the other hand, Timer 2 received the desired time delay after each feeding process entered by user. Timer 1 will be activated and repeated each time Timer 2 finished its calculation, and this process will continue until the system shut down. Once input parameters had been completely set up by the user, the motor will start to operate and activate the feeding mechanism, in which pellets dispensed into the pond according to the predefined rotation speed of the motor. Finally, the pellets will be shot into the marked area of the water surface. Figure 4 below shows the flows of the system interface.
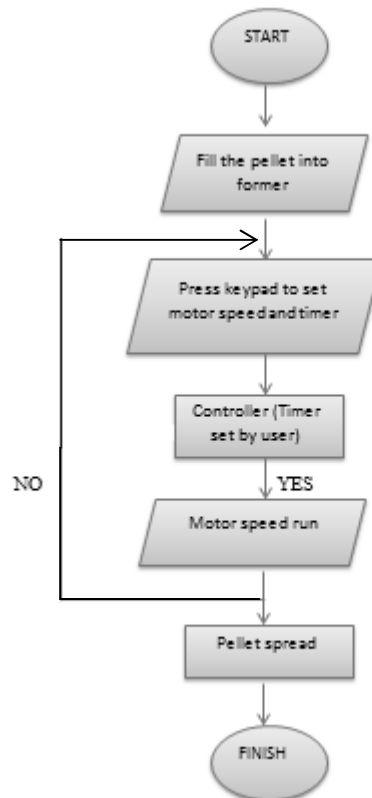

Figure 5: Operation of the system.

III. RESULT AND DISCUSSION

A. *Voltage drop at DC motor output*

The study of designing Automatic Fish Feeder using PIC microcontroller is to find out the way of rotation speed of motor base on the percentage of PWM value set.

TABLE I.   TABLE OF AVERAGE VOLTAGE WITH PWM VALUE SET TYPE STYLES

PWM (%)	Motor Condition	Voltage Drop
10	OFF	4.06
20	OFF	5.41
30	ON	6.53
40	ON	7.31
50	ON	8.38
60	ON	8.71
70	ON	9.86
80	ON	10.27
90	ON	10.91

Table 1 above shows the voltage drop at the output DC motor depending on the percentage of the PWM use. Base on the result, the DC motor do not function or operate during the PWM are set to 10% and 20%. The voltage drop are too low and effect the rotation speed of the DC motor. This system is reliable since the other PWM percentage will run the DC motor smoothly. This result also shows that the DC motor will be run properly when the voltages are in between 6V until 11V. Figure 6 below shows the graph of PWM value set versus the voltage drop at DC motor output.
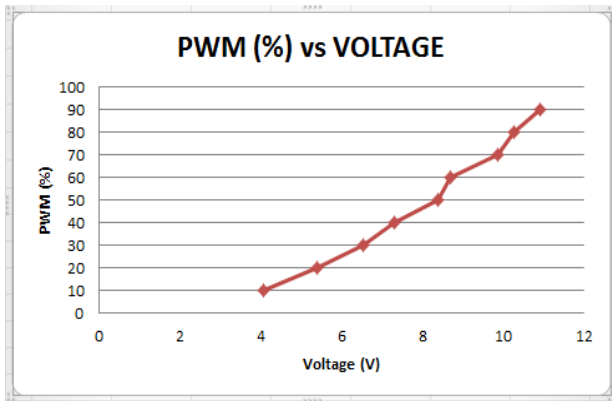
Figure 6:   Graph of PWM value set versus the voltage drop at DC motor output

B. *Analysis of the pellets distance*

TABLE II.   TABLE OF MAXIMUM AND MINIMUM PELLET DISTANCE
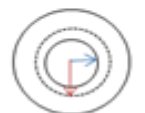
PWM (%)	Area of the pellet distance — MIN, MAX	Description
30		MIN = 39 cm MAX= 64 cm
40		MIN = 45 cm MAX= 70 cm
50		MIN = 81 cm MAX= 99 cm
60		MIN = 112 cm MAX= 118 cm
70		MIN = 132 cm MAX= 145 cm
80		MIN = 150 cm MAX= 160 cm
90		MIN = 164 cm MAX= 176 cm

Table 2 above shows the data of maximum and minimum distance of the pellet spread and the marking area of the pellet distance. From the observation, the pellet that has been spread into the water has different distance depending on PWM value set. Therefore, the user has their own choice to feed their fish based on the size of the pond itself. Basically, the small of PWM value set is suitable for small fish feeding because the small fish did not have to compete each other in order to get the pellet compare with the big fish fed. The distance of the pellet spread play an important role to ensure that the good and efficient of the fish growth.

C. *The pellet usage deal with timer*

TABLE III. TABLE OF MAXIMUM AND MINIMUM PELLET DISTANCE

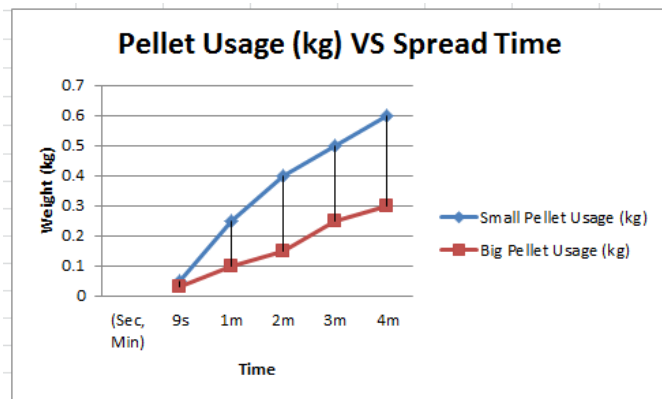| Timer (Sec, Min)	Small Pellet Usage (kg)	Large Pellet Usage (kg)
9s	0.05	0.03
1m	0.25	0.10
2m	0.40	0.15
3m	0.50	0.25
4m	0.60	0.30

Figure 7: Graph type of pellet usage versus spreading time

Figure 7 above shows the analysis of the type of pellet usage with the spreading time. The maximum speed (90% PWM value set) has been use for this analysis. The size of the pellet will be effect to the pellet usage in this system. From the observation, the small pellet will be dispensing faster than the big pellet. The dispensing rate for small pellet is around 250g/m and this is 100g/m for the large pellets.

IV. CONCLUSION

This research is one of the alternative and easy ways to feed the fish at the right cycle time. Nowadays, the entire fish farmer has to hire more labor cost to handle any kind of work dealing with the feeding fish. So this research can reduce the owner to hire more workers and also reduce the time needed. The objectives of this research were successfully achieved through the 3 analysis that have been done. The speed rotation of the motor have been controlled by PWM technique and the different voltage drop at the DC motor output will determine the suitable speed that should be use.

Next, the analysis of the pellet distance which allows the user to choose the suitable marking area based on their cattle. The analysis of the pellet usage will make the pellet use to fed the fish more reliable and avoid the pellet usage waste. The main purpose of this research is to create and develop an innovation which is useful to overcome some of the problems that occurred nowadays. PIC microcontroller has been used for monitoring the present control system. The aim of this research is to be able to control the DC motor by using microcontroller application. Therefore, this system allows the owner to adjust the cycle time and dispensing time as required.

This research can be modified for more advance in it operation. One of the modifications is to control the pellet flow from the pellet storage into the sphere former. By doing so, the research can be more efficient in form of pellet flow control. Another modification is tried to make some modification of sphere former by having a gate to open or close when the system is running.

REFERENCES

[1] M.E.I Zulkefly, Development of PLC controled Ariel Fish Feeding System, Bachelor Degree Thesis in Mechanical Engineering, Universiti Tun Hussien Onn Malaysia 2010.

[2] Martin George Ljungqvist, Michael Engelbrecht Nielson, Bjarne Kjaer Ersboll, Stina Frosch."image Analysis of Pellet Size For Control System in Industrial Feed Production", Technical University of Denmark, October 201.

[3] Allen Goldblatt, "Automatic feeder for marine mammals", Laboratory for marinne mammal Research University of Tel Aviv, 1992.

[4] Ding Jie, Wang kecheng. A new Type Speed Regulator for a PWM Speed Regulating System.Anshan Institute of science and technolgy, 2003.

[5] Zhijjun Liu, Lianzhi Jiang, "PWM Speed Control System of DC motor Based on AT89S51", School of Electrical and Information Engineering Liaoning Institute of Science and Technology, 2011

[6] Varadi, Mechanized Feeding in Aquaculture, *Inland Aquaculture Engineering*, ADCP/REP/84/21, Food and Agriculture Organization (FAO), 1984.

[7] L. Wong, Redesign and Detail of Analysis of a Tiger Prawn Food Feeder, *Bachelor Degree Thesis in Mechanical Engineering*, Universiti Tun Hussein Onn Malaysia, 2005.

[8] Yeoh, F.S Taip, J. Endan, R.A. Talib and M.K Siti Mazlina. "*Development of automatic Feeding Machine for Aquaculture Industry*". Department of Process and Food Engineering, Faculty of Engineering, Universiti Purta Malaysia, 2009.

[9] Song jian, Jiang Junsheng, Zhao wenliang. "*The DC-motor PWM Speed regular system Base on Single Chip microcomputer*". Study on Agricultural Mechanization,2006,1(1)

[10] C.M.Chang, W.Fang, R.C.Jao, C.Z.Shyu, J.C.Liao, Development of intelligent feeding controller for indoor intensive culturing of ccl. Aquacultural Engineering, 2005.

The Development of Self-service Restaurant Ordering System (SROS)

M. Z. H. Noor, A. A. A. Rahman, M. F. Saaid, M. S. A. M. Ali, M. Zolkapli

Faculty of Electrical Engineering
Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia
mzhakim@salam.uitm.edu.my

Abstract— **This paper is about The Development of Self-service Restaurant Ordering System (SROS). The manual restaurant ordering system is relies on a lot of manpower to handle all the process from taking order from customers, placing order, tidy up the table and cleaning the dishes. Therefore this system is developed to reduce the number of manpower in ordering task and at the same time reduce the monthly cost for the restaurant. Customer can make their order through the system and directly stored to the database. The system is designed using Microsoft Visual Studio 2008 and Microsoft Office Access 2007 to give a better solution for the manual system. A preliminary testing suggested that the SROS has a potential to replace the manual restaurant ordering system.**

Keywords-component; food ordering system, interactive user interface

I. INTRODUCTION

Nowadays, most of the restaurant around the world using the manual ordering system. This system is using a waiters and waitress to take an order from the customers. This system relies on large numbers of manpower to handle customer reservation, inquiry, ordering food, placing order, reminding dishes [1].

This typical method is kind of wasting of time and energy when there are a lot of costumers at that time. Moreover it may be cause a misunderstanding between the customer and waiter. However, if there are too many waiters are hired, it may be waste of resource during non-peak hour [2]. It also will give an extra-work to the cashier to record all the transaction.

There are some early efforts have been made to replace this manual ordering process, for example is a PDA-based wireless food ordering system. However this system is only replacing paper and pen used by the waiter to take an order [3]. A miscommunication still can occur since customers have to order through the waiter. Another early system is A Customizable Wireless Food Ordering System with Real time Customer Feedback (CWOS-RTF) [4]. This system requires the customer to make an order through their smart phone using web based application. However, not all the customer has a smart phone and its gives a problem if they do not have one. Another system is Restaurant Waiter Paging System [4]. This system allows customers to call for a waiter but it is still using manual ordering process.

Therefore, the research has been done to develop a system which will give a lot more benefit to both restaurant owner and customers. The system will improve all the lack from the previous systems. Customer can directly make an order from the system and misunderstanding between customers and waiters can be reduced to minimal [4]. Moreover, it also will improve the data collection since order make by the customer is directly sent to the database. It will reduce time waiting by the customer and restaurant owner can reduce the expenses on manpower.

The Self-Service Restaurant Ordering System (SROS) developed a Graphical User Interface (GUI) using a Microsoft Visual Studio 2008 and Microsoft Office Access 2007 for the database. It is requiring customer to order via touch screen device that placed on each table in the restaurant. Customers are able to view the menu, price and make an order directly using the system. Then, their order will be sent to the database in restaurant and will be view on the computer screen at the kitchen for food preparation.
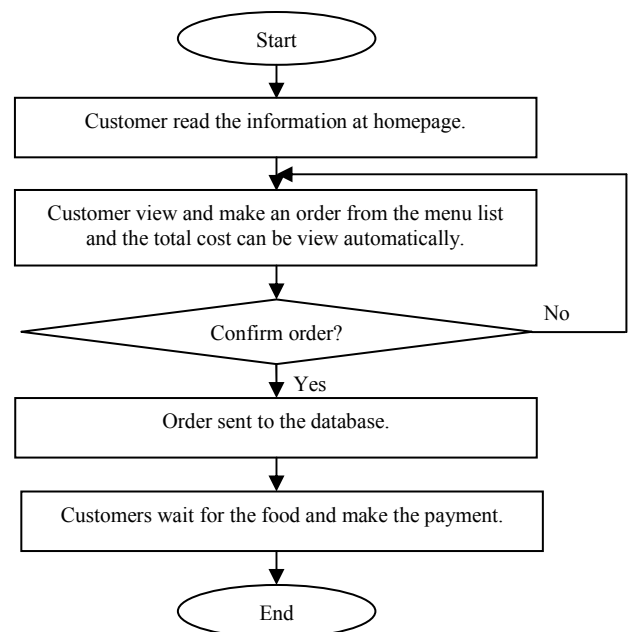
II. METHODOLOGY

Figure 1.   Work flow for the SROS

The work flow for developed the SROS as shown in Figure 1. When customers enter the system, the home page of the system will be appearing. This page contains the information of the restaurant and the promotion price when available. Then, customers will navigate to the menu page to view the entire menu and then can be proceed to make their order. They can simply add and remove their desired menu and the total cost is automatically counted.
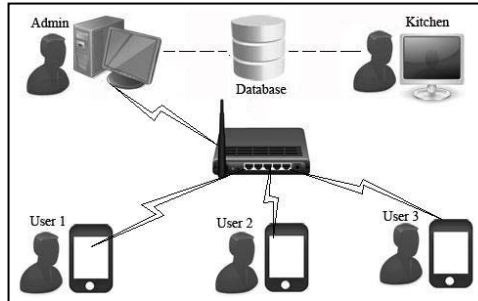

Figure 2. SROS functional correlation

The SROS functional correlation as shown in Figure 2. The connection between users and server is connected by the router through the LAN cable for a better connection [6]. The order will be saved into the database in the admin computer. This order can be also viewed by the kitchen computer for the food preparation [7].

A. Login Form

This login form is requiring the password for user to enter the system. The admin can change the password anytime for the security purpose.

Figure 3. Login GUI form

Figure 4. Password changer option

B. Home page form

It is a main page which is the user see after they access from the login form. This page is showing a restaurant banner, picture and information for promotion purposes.

Figure 5. Home page for SROS

C. Menu ordering form

This is a most important page of the system. This page is used by the customers to make their orders as shown in Figure 6. It is showing a menu pictures, names and prices. Customers has to choose their desired menu by clicking the on the button which is contain the menu picture. Customer can see their selected menu on the table beside this tab. Meanwhile the total price is automatically calculated every time customer chooses their order. However customer can remove their unwanted menu by clicking remove button below the table. If they satisfied with their order they must clicked the confirm button below the table. This order then sent to the database for data collection and food preparation. The admin are not going to print the receipts in order to limit the used of the paper.

Figure 6. Menu ordering form

D. Table number editing form

This form is used by the admin to update the table number. It is including add table and remove table. This table information is very useful to differentiate the order from the customers.
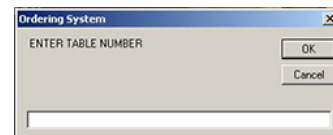

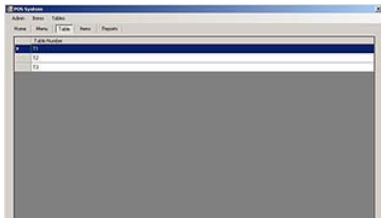Figure 7. Table number editing form

Figure 8.   List of restaurant table

E.   Menu updating form

This is an items manager used by the admin to updates the menu. It is including add menu, remove menu and edit menu. This updated menu then saved in the database and can be view in items tab. The menu information is including an items barcode, menu name, buy price, selling price and picture.

Figure 9.   Menu updating form

Figure 10. Example of menu list

F.   Restaurant profit data form

It is a fast tool used by the admin to view the profit of the restaurant in a certain period. It is containing overall profit, total profit between two dates, total profit and item for all time and total profit and item between two dates. It is one of the best features offered by this system compare to others.
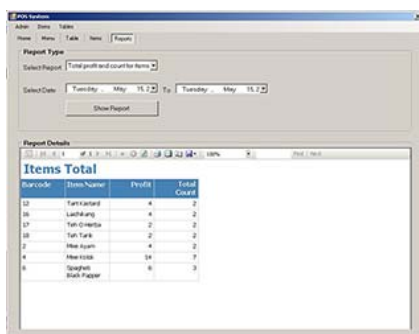

Figure 11. Profit data form

G.   SROS database

This is a very important part of the system. It is used to restore all the items information that save by the admin including its barcode, name, buy price, selling price and picture. It is also used to save the password of the system and the table number information. Moreover it is very important because all the transaction of the restaurant will saved in this database and will be used for future revision.
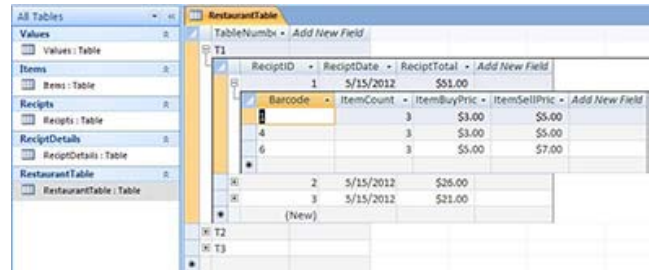

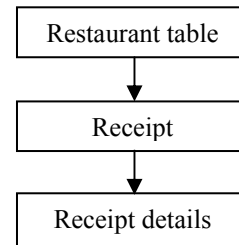
Figure 12. Example of SROS database

Figure 13. SROS ordering database relationship

In Figure 13, restaurant table database show a list of a table in the restaurant. Inside this table is containing a receipt table. It is a place where the order saved in form of receipt ID, receipt date and total cost. The last one is a receipt details table. It is a place where order saved in form of barcode, quantity of the menu, item buy price which is mean that the cost of the menu and the last one is an item sell price. The kitchen attendant can close this last table each time the order is done and mark the last column.

III.   RESULT AND DISCUSSION

There are several testing and evaluations have been done to check the performance of the system. The first test is to check the functionality of the system whether it is work as expected result. The result of the test can be seeing in Table 1 below.

TABLE I.   SROS ANALYSIS RESULTS

Test Module	Expected Result	Actual Result
Initialize the system by adding all items to the database.	All items can be viewed in database and ready to use in GUI.	Success
Update the password.	Admin login with updated password.	Success
Add and remove desired items to the table. (before	List of items and quantity selected can be viewed beside the order	Success

confirmation)	form.	
Run calculation on input read from interface.	Total cost is automatically calculated once users add an item.	Success
Send confirmed order to the database.	Database record all the transaction that have been made.	Success
Report generation.	The profit in certain period can be easily viewed from the system.	Success
Performance.	All interfaces are fully displayed within less than 2 seconds.	Success

Initialization the system by add all the menus and tables is a procedure taken by the admin to fill in the selling the offered menu in the menu adder form as shown in Figure 14. The example of menu offered database as shown in Figure 15.

Figure 14. Menu adder form

Figure 15. Example of menu list database

The initial password given is a 'password'. However, admin can change the password anytime to secure the system.

Figure 16. Password changer form

Menu page is a GUI where the customers have to deal with. They can simply adding their desired menu to the list by clicking the button. The quantities of the menu selected are counted by how many times the button clicked. Meanwhile customers can remove the items before confirmation.

The total cost is calculated automatically each time customer add the items. This customer will give a benefit to the customer to control their expenses. The purposes of the list of items selected and total cost is to prevent the receipt printing as a way to control the used of paper.

Figure 17. Menu selection including total cost

Once the confirmation button is clicked, this order is send to the database for the record. This record is easy to review by the admin as it is differentiate by sections.

Figure 18. Menu ordered database

SROS profit database as shown in Figure 19 is a fast tool used by the admin to monitor the restaurant profit. From this section, admin can review which menu is most ordered and which one is not. So that, admin can easily budget the expenses on the materials.

Figure 19. SROS profit database

There are some analysis have been done to compare the SROS with the manual system. The Table II and calculation below show the different in manpower cost and time for both methods.

TABLE II.   COMPARISON BETWEEN SROS AND MANUAL SYSTEM COST AND TIME

Analysis	SROS	Manual	Reduce
Manpower cost	RM6400	RM8800	27.3%
Servicing time	20 minutes	27 minutes	29.5%

A. Cost for manpower

One of the objectives of the system is to reduce the number of manpower since ordering process is replace by the system. From the observation, the average salary and number of workers is taken from the restaurants around Section 7, Shah Alam, Selangor.

1) *Manual/Existing system*

```
Cost    =   Salary x No. of workers
Waiter  =   RM1200 x 4
        =   RM4800

Chef    =   RM2000 x 2
        =   RM4000

Total   =   RM4800 + RM4000
        =   RM8800
```

2) *Self-service Restaurant Ordering System (SROS)*

```
Cost    =   Salary x No. of workers
Waiter  =   RM1200 x 2
        =   RM2400

Chef    =   RM2000 x 2
        =   RM4000

Total   =   RM2400 + RM4000
        =   RM6400
```

The calculation below show the percentage of cost can be saved by the restaurant. Large numbers reduction can give more profit to the restaurant. So that owner of the restaurant can use this cost to upgrade the restaurant to make the customer more comfortable or expense on promotion.

```
Percentage of man power cost (%)
= [(Manual – SROS) / Manual] x 100%
= [(RM8800 – RM6400) / RM8800] x 100%
= [RM2400 / RM8800] x 100%
= 27.3%
```

B. Servicing Time

Since the system are replacing a lot of task that done by the human before, so it is reduced a lot of times. The calculation below show the different of using time between existing system and SROS for each order made. The data is the average time during peak time in restaurants around Section 7, Shah Alam, Selangor.

1) *Manual/Existing ordering system*
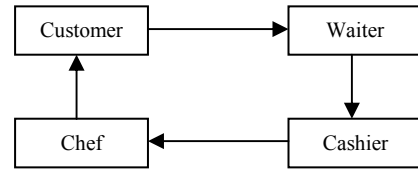

Figure 20. Manual ordering system flow

```
Waiting            =   5 min
Order time         =   5 min
Record menu        =   2 min
Food preparation   =   15 min

Total service time =   5 + 5 + 2 + 15
                   =   27 minutes
```

2) *Self-service Restaurant Ordering System (SROS)*

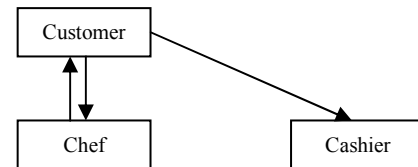Figure 21. SROS ordering system flow

```
Waiting            =   0 min
Order time         =   5 min
Record             =   0 min
Food preparation   =   15 min

Total service time =   5 + 0 + 0 + 15
                   =   20 minutes
```

Figure 20 and 21 show the manual and SROS ordering system flow. From the calculation for total service time, it is show a reduction of ordering time for SROS compares to the manual ordering system. The different total service time between manual and SROS ordering system is 7 minutes which is about 25.9%.

```
Percentage of total service time (%)
=   [(Manual – SROS) / Manual] x 100%
=   [(27 – 20) / 27] x 100%
=   [7 / 27] x 100%
=   25.9%
```

C. Customer satisfaction

The automatic system is giving a lot of benefit to both restaurant owner and customers. There is a study have been done the see the customers satisfaction between automatic ordering system and manual ordering system on a few criteria to 43 respondents [4].

- User friendly

- Successful in sending order

- Save times

- Easy to understand/communicate

From the analysis, about 91% of the respondent preferred automatic ordering system than manual system. This result show that the acceptance of the system among the customers.

The system is really easy to use by the customers to make an order. It saved a lot of times and the misunderstanding between customer and waiter can be avoided since the order is directly send to the database at the counter.

The design of the system is also user friendly and easy to understand by the customer to make their first move.

IV. CONCLUSION

In this paper, the Self-Service Restaurant Ordering System (SROS) has been designed to replace the manual system. The system served a lot of benefit to both customer and restaurant owner.

Microsoft Visual Studio 2008 is chosen to develop this system since it is easily used to edit the graphical interface. Meanwhile MS Access is sufficient to accommodate storage and queries of data.

Since a lot of task is replace by the system, the owner of the restaurant can reduce the number of manpower and reduce the cost of monthly expenses. For customer, the system will reduce the time waiting and misunderstanding can be reduced to minimal. This is really important thing during peak hour to make sure the customer satisfy with the service.

Some element can be added to the system in the future to make it more interactive and user friendly. One of the elements is to add the entertainment page such as games, movies or music. The customer can used this application during food preparation and prevent from bored of waiting during peak hour. Another addition is the system can be connected with web based application. Is it used by the customer to make a reservation or pre-order before enter the restaurant. In conclusion, the system is very suitable in a real-time to give more benefit to the business.

REFERENCES

[1] Y. Z. Yuan, Z. Weibing. "The Research and Realization of Wireless Ordering System Based on Embedded Technology" International Conference on Computer Application and System Modeling, 2010.

[2] Y. C. Tan, et al., "Automated Food Ordering System with Interactive User Interface Approach" Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Malaysia, 2010.

[3] S. N. Cheong, et al., "Design and Development of Multi-Touchable E-Restaurant Management System" Faculty of Engineering Multimedia University, Malaysia, 2010.

[4] N. A. Samsudin, et al., "A Customizable Wireless Food Ordering System with Real time Customer Feedback" IEEE Symposium on Wireless Technology and Application (ISWTA), 2011.

[5] D. Green, J. Curtis "Restaurant Waiter Paging System" US Patent 6,366,196, Tennessee, 2000.

[6] C. S. Chang, et al., "Development and Implementation of an E-Restaurant for Costumer-Centric Service Using WLAN and RFID Technologies" Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan, 2008.

[7] M. H. A. Wahab, et al., "Implementation of Network-based Smart Order System" Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn, Malaysia, 2008.

[8] K. J. Patel, et al., "PDA-based Wireless Food Ordering System for Hospitality Industry – A Case Study of Box Hill Institute" Centre for Information and Communications Technology and Super Cisco Academy Training Centre (Asia Pacific) Box Hill Institute Box Hill, Victoria, Australia, 2007.

[9] S. N. Cheong, et al., "Enriching Dining Experience with the Multi-Touchable Entertainment Applications" Faculty of Engineering Multimedia University, Malaysia, 2010.

[10] XU Hongzhen, et al., "Wireless Food Ordering System Based on Web Services" Department of Computer Science and Technology, East China Institute of Technology Fuzhou, China, 2009.

Analysis of EEG Signal from Right and Left Hand Writing Movements

C.W.N.F. Che Wan Fadzal, W. Mansor, L. Y. Khuan

Faculty of Electrical Engineering,
Universiti Teknologi MARA
40450 Shah Alam,
Selangor, Malaysia
wahidah231@salam.uitm.edu.my

Abstract—**This paper studies on the characteristics of electroencephalogram (EEG) which generated from writing using right and left hand. The EEG signals were recorded from 4 channels, C3, C4, P3 and P4 and processed using band pass filter (8-30Hz). Two method of analysis were performed; Fast Fourier transform and power spectral density. The results showed that Power Spectral Density can be used to distinguish right and left hand writing movements from EEG signals.**

Keywords-component; Electroencephalogram; writing, Fast Fourier Transform, Power Spectral Density

I. INTRODUCTION

Electroencephalogram (EEG) is used in Brain Computer Interface (BCI) to convert mere reflections of central nervous system (CNS) activity into messages that act on the world [1]. BCI has been used to control instruments using signal from human or animal brain which is recorded either invasively or noninvasively [2, 3]. It is a communication system that allows interaction between human and external devices.

There are 4 main components of brain; cerebrum, cerebellum, the brain stem and the thalamus [4]. The largest part of the brain is cerebrum. Its function is to give us information on what happen around us and how to respond with it. It is divided into two modules that is right and left hemisphere as shown in Figure 1. Muscles of our right body are controlled by the left hemisphere and the left body muscles are controlled by the right hemisphere.
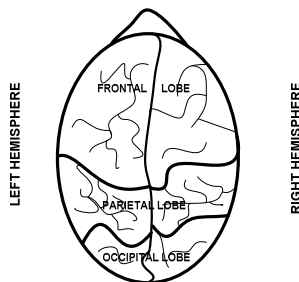

Figure 1: Location of right and left hemisphere in cerebrum of brain

Numerous investigations have studied EEG signals from right and left hand movements. Zhang et al [5] extracted features of EEG signal obtained from imagined right and hand movements using power spectral entropy. Cincotti et al [6] studied EEG signals from imagined right and left finger movements. However, none of the researchers investigates the EEG signal characteristics obtained from right and left hand writing movements.

This paper discusses the analysis of EEG signals generated from right and left handed writing activities. Fast Fourier Transform and Power Spectral Density were used to distinguish between the left and right hand writing movements.

II. BACKGROUND THEORY

EEG signal has unique pattern which can be revealed from Fourier Transform analysis. This signal can be recorded by placing multiple electrodes on the cortex, under the skull or certain locations on the scalp. Applying Fourier transform on the signal allows the frequency content to be extracted. There are 5 major brain wave types can be identified; delta (0.5-4Hz), theta (4 – 7.5Hz), Alpha (8 -13 Hz), Beta (14 – 26Hz) and Gamma (above 30Hz, mainly up to 45Hz) [7].

Fast Fourier Transforms (FFTs) is an algorithm that will compute discrete Fourier transform (DFT) and it's inverse. Equation (1) above is the equation defined the DFT of an N-length. To examine the frequency of the signal generated, frequency analysis done due to the types of EEG signal in time domain [8].

$$X(k) = \sum_{n=0}^{N-1} x[n] \exp\left(-\frac{j2\pi kn}{N}\right) = \sum_{n=0}^{N-1} x[n] W_N^{kn} \text{--(1)}$$

X[k] is evaluated for $0 \leq k \leq (N-1)$

W_N^{kn} is a periodic functions with a limited number of distinct values.

A. Power Spectral Density(PSD)

Power spectral density has been used to investigate the EEG signal power for various cases. These includes in the study of human brain changes after consuming alcohol [9] and validation of neural mass model [10]. The power spectral density (PSD) indicates how the sequence's power or energy is distributed in the frequency domain, and is widely-used to measure the random signals and noise.

The general PSD equation is given by Equation (2),

$$Pxx(X) = \sum_{m=-\infty}^{\infty} \emptyset xx[m]\exp(-jXm) \quad -----(2)$$

$Pxx(X)$ is related to the spectral power distribution of a digital sequence [6].

Power spectral density function (PSD) shows the strength of the variations energy as a function of frequency. PSD shows at which frequency variations are strong and at which frequency variations are weak.

III. METHODS

A. EEG Recording

The EEG datasets used in this study was collected from right and left-handed subjects. Four channels have been used to record EEG data that is C3, C4, P3 and P4. However, in this study, EEG obtained from C3 and C4 channels was used since these channels associate with activities produced from right and left hands.
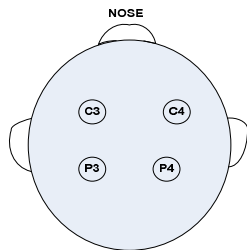

Figure 2: Location of electrode placements.

In EEG recording stage, subjects were asked to write a passage for 3 trials. In each trial, EEG was recorded in 120 seconds as shown in Table 1. Before a new trial was conducted, the subjects were asked to relax in 20s and then focus on the task.

B. Data pre-processing

The recorded EEG signals consist of artifacts which interfere with the embedded information and need to be removed. Mehrdad Fatourechi et al [11] has reported that this artifacts are eye movements, known as electrooculogram (EOG) and muscle movements, known as electromyogram (EMG) in the context of BCI systems. In order to remove the artifacts, a band-pass filter with frequency of 8-30Hz was used. EOG and EMG has a wide frequency range with EOG has frequency below 4Hz meanwhile EMG with frequency higher than 20Hz – 300Hz [11, 12].

C. Frequency Analysis

The EEG signals from the left and right handed subjects were first analysed using Fast Fourier Transform. After frequency analysis, power spectral density (PSD) was then computed. Here, Welch's method was used to calculate the power spectral density. Welch's method has been used by Meng Hu et al to classify normal and hypoxia EEG since it enables the basic information of how power distributes as a functions of frequency to be revealed [13].

IV. RESULTS AND DISCUSSIONS

Frequency spectrum and power spectral density of EEG signals obtained from the right handed and left handed writers are shown in Figure 3 to Figure 6. It is obvious that there are different pattern of frequency spectrum produced from right and left handed writers. This shows that electrode C3 and C4 are suitable electrode placement to clearly differentiate EEG signals obtained from right and left- handed writers.

Figure 3 and Figure 4 show the frequency spectrum of EEG signal from male and female subjects that perform writing using right hand. From the frequency spectrum, it can be seen that writing activities are more apparent from EEG signals recorded at electrode C3 compared to C4. The PSD results also give the same indication where the maximum PSD at electrode C3 for right handed subject is more (3.917 db/Hz) than that obtained at electrode C4 (2.647 db/Hz).

Meanwhile, the frequency spectrum of EEG signals from left hand writing movements obtained at electrode C3 in figure 5 and Figure 6 is similar to that obtained at electrode C4. However, the maximum PSD of EEG recorded at electrode C3 is more than that obtained at electrode C4, which clearly shows that C4 gives more indication of left hand writing movements.
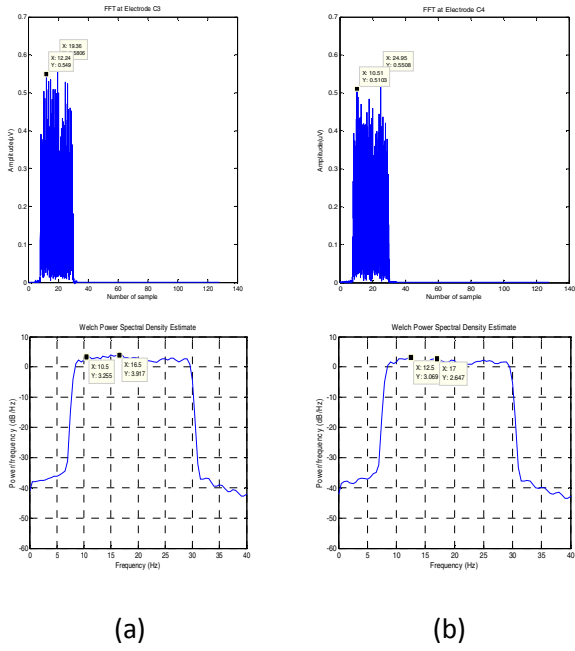
TABLE I. TIME SEQUENCE OF ONE TRIAL OF WRITING ACTIVITY

20s	10s	90s
Relax	Start writing	Writing the passage

(a)   (b)

Figure 3:Frequency spectrum and power specral density of EEG signal obtained from right handed male writer at (a) C3 and (b) C4

(a)   (b)

Figure 5: Frequency spectrum and power spectral density of EEG signal obtained from left handed male writer at electrode (a) C3 and (b) C4

(a)   (b)

Figure 4 : Frequency spectrum and power spectral density of EEG signal obtained from right handed female writer at electrode (a) C3 and (b) C4
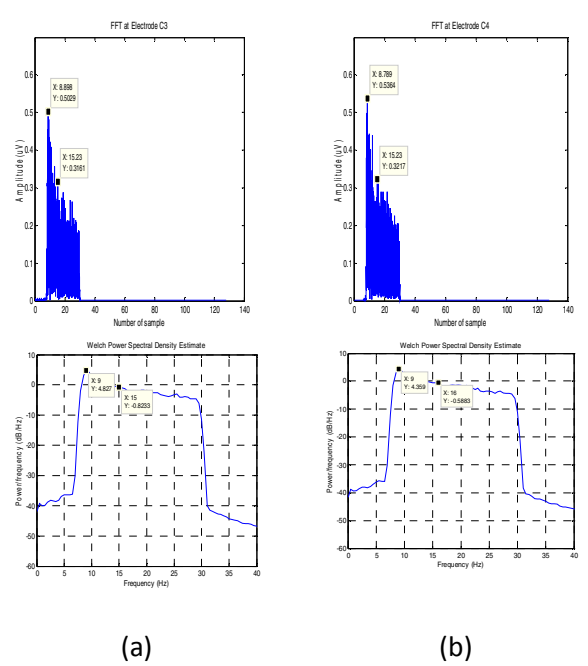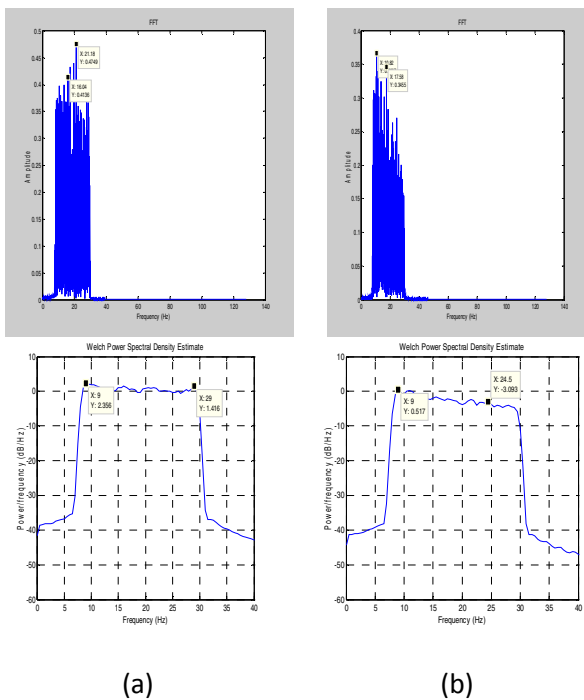

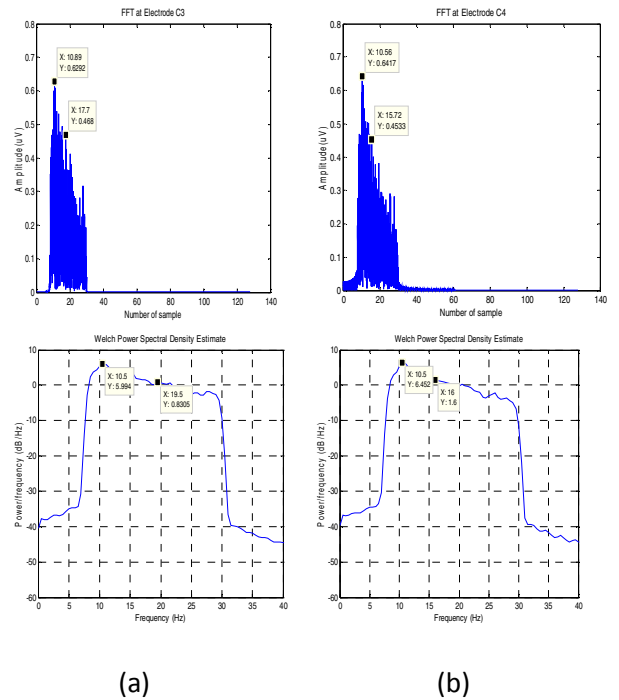(a)   (b)

Figure 6: Frequency spectrum and power spectral density of EEG signal obtained from leftt handed female writer at electrode (a) C3 and (b) C4.

335

The detail maximum power (db/Hz) of EEG signal at each electrode for right and left handed writers is shown in Table II and Table III.

TABLE II. THE MAXIMUM POWER (DB/HZ) OF EEG FROM RIGHT-HANDED WRITING ACTIVITY

Subject	Max.Power at Electrode C3(db/Hz)	Max.Power at Electrode C4(db/Hz)
1	3.917	2.647
2	1.416	-3.093
Average	2.6665	-0.223

TABLE III. THE MAXIMUM POWER (DB/HZ) OF EEG FROM RIGHT-HANDED WRITING ACTIVITY

Subject	Max.Power at Electrode C3 (db/Hz)	Max.Power at Electrode C4 (db/Hz)
3	-0.8233	-0.5883
4	0.8305	1.6
Average	0.0036	0.50585

Figure 7 and Figure 8 show the dominant electrode placement for left and right handed writer based on the power spectral density plot. It is obvious that for right-handed writer, the power of EEG signal from electrode C3 is more than that produced from electrode C4. Meanwhile for left-handed writer, the EEG signal power spectral density is larger than that obtained from electrode C3.This is because for right-handed person, the EEG signal is activated at left-hemisphere in the brain meanwhile for left-handed person, the signal is activated at right-hemisphere in the brain.

The results obtained from this study are in agreement with that reported by Zhang et al [5]. They have found that power spectral entropy of EEG signal from imagined right hand is dominant at channel C3 [5]. Thus, channel C3 and C4 can be used to record EEG signals for imagined right and left hand movements. Since this method is suitable for imagined movement, it is also good for actual hand movement either for right and left-handed writer.

In another study, Yang has discovered that there are some differences between the EEG waveforms from the two sides of brain, left and right [6]. They found that EEG waveform from C4 location was more suppressed than that from C3 when

movements from left hand were imagined [6]. This means that the left and right hand movements may produce different characteristics of EEG signals.
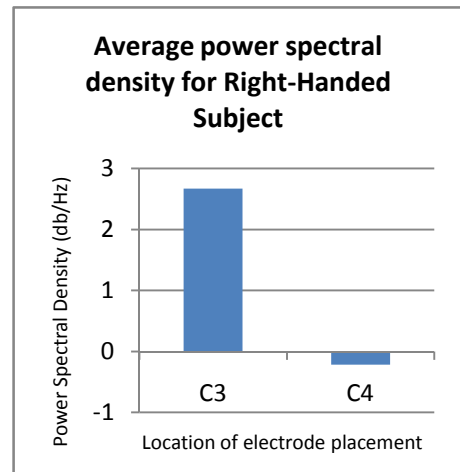

Figure 7: Average power spectral density for right-handed subject.
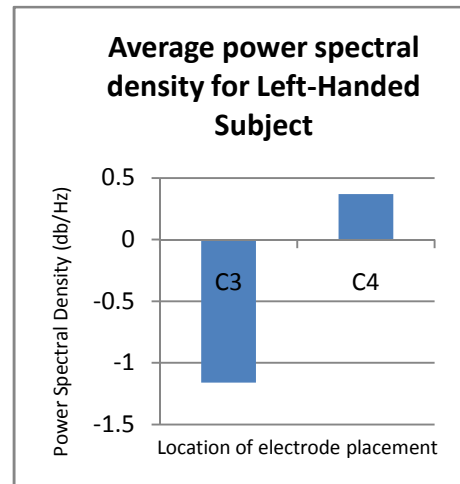
Figure 8: Average power spectral density for left-handed subject.

V. CONCLUSION

EEG signal from right-handed and left-handed writing have been analysed using frequency analysis and power spectral density are discussed in this paper. Four electrodes were used to acquire the EEG signals, however only EEG signals from electrode C3 and C4 were analysed. It was found that the power spectral density of EEG signal from right-handed and left-handed writing is different. EEG signal obtained from right handed writing produced higher power spectral density (PSD) in the signal is recorded at electrode C3. For left-handed writing, the suitable electrode placement is C4.

REFERENCES

[1] Joseph N. Mak, Jonathan R. Wolpaw," Clinical Applications of Brain-Computer Interfaces: Current State and Future Prospects", *IEEE Rev Biomed Eng*, vol (2),pp.187–199, 2009.

[2] L.Kauhanen ,T.Nykopp, M.Sams,"Classifiction of single MEG trials related to left and right index finger movements",Clinical Neurophysiology,vol 117,pp. 430-439, 2006.

[3] Siuly,Y.Li,P.Wen,"Classification of EEG Signals using Sampling Techniques and Least Square Support Vector Machines", RSKT '09 Proceedings of the 4th International Conference on Rough Sets and Knowledge Technology,pp. 375 – 382,2009.

[4] C.W.N.F. Che Wan Fadzal, W. Mansor, L. Y. Khuan, "An Analysis of EEG Signal Generated From Grasping and Writing", In Proceedings of 2011 IEEE Conference on Computer Application & Industrial Electronics , pp. 535-537, 2011.

[5] A.Zhang,B.Yang,L.Huang,"Feature Extraction of EEG signals Using Power Spectral Entropy", Biomedical Engineering and Informatics, vol 2, pp.435-439, 2008.

[6] F. Cincotti, L. Bianch, J. R. Millhn , J. Mourifio, S. Salinari4, M. G. Marciani, F. Babilooi, "Brain Computer Interface: the use of Low Resolution Surface Laplacian and Linear Classifiers for the Recognition of Imagined Hand Movements", Proceeding of IEEE EMBS, pp 655-658, 2001.

[7] S.Sanei,J.A.Chambers,"EEG Signal Processing",John Wiley & Sons Ltd,England,2007.

[8] P.A.Lynn,W.Fuerst,"Introductory Digital Signal Processing with Computer Applications:Second Edition",John Wiley & Sons Ltd,The Atrium,Southern Gate,England, 2007.

[9] D. Wu., Z. Chen, R. Feng, G. Li, T. Luan, "Study on Human Brain After Consuming Alcohol Based on EEG signal", Proc. of IEEE on Computer Science and Information Technology, pp 406-409, 2010.

[10] Melissa Zavaglia, Laura Astolfi, Fabio Babiloni, and Mauro Ursino, "The Effect of Connectivity on EEG Rhythms, Power Spectral Density and Coherence Among Coupled Neural Populations: Analysis With a Neural Mass Model", IEEE Transactions On Biomedical Engineering, Vol. 55, No. 1, pp 69-76, 2008.

[11] M.Fatourechi,A.Bashashati,R.K.Ward,G.E.Birch," EMG and EOG artifacts in brain computer interface systems" , Clinical Neurophysiology, Vol 118(3), pp. 480-494, 2007.

[12] M. R. Ahsan, M. Ibrahimy, O. O. Khalifa, "Electromygraphy (EMG) Signal based Hand Gesture Recognition using Artificial Neural Network (ANN)", Int. Conference on Mechatronics, pp 1-6, 2011.

[13] Meng Hu, Jiaojie Li, Guang Li,*, Xiaowei Tang and Qiuping Ding," Classification of Normal and Hypoxia EEG Based on Approximate Entropy and Welch Power-Spectral-Density",International Joint Conference on Neural Networks Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada,July, 2006.

Adaptive Feedforward Zero Phase Error Tracking Control for Minimum Phase and Non-minimum Phase Systems – XY Table Real-Time Application

Ramli Adnan[#1], Hashimah Ismail[*2], Norlela Ishak[#3], Mazidah Tajjudin[#4] and Mohd Hezri Fazalul Rahiman[#5]

[#]Faculty of Electrical Engineering Universiti Teknologi MARA, Shah Alam, Selangor
[*]Faculty of Engineering, Universiti Selangor (UNISEL), Bestari Jaya, Selangor
[1]ramli324@salam.uitm.edu.my
[2]hashimah@unisel.edu.my

Abstract— **This paper presents the realization of Adaptive feed-forward Zero Phase Error Tracking Control (ZPETC) for real-time XY table application. A minimum and non-minimum phase mathematical models were obtained from input-output experimental data using different sampling times to represent the true plant of XY table via Matlab system identification toolbox. These two models were used to test the effectiveness of designed controller. In this paper, ZPETC without factorization of zero polynomials is utilized and the gain compensation filter is replaced by Finite Impulse Response (FIR) filter. The gain adaptation mechanism is done using Recursive Least Square (RLS) algorithm. The contour error was used as indicator for effective tracking performance. Real-time application results show that the designed controller can effectively enhance the tracking performance of the system and provide superior two dimensional contour plots for the XY table.**

Keywords— Adaptive ZPETC; Real-Time Control; System Identification; Motion Control

I. INTRODUCTION

The XY table has been widely used in electronic manufacturing, laser cutting, wire cutting and CNC machining. For these types of application, precision is very important. In order to achieve high precision positioning tracking system, the tracking error for each axis must be minimal. This would lead to great challenge to the researchers. The tracking controls for the individual axis can indirectly reduce the contour error [1]. According to Wang et al. [2] the x-y plane contour accuracy is dominated by position and tracking error of each axis. Satisfying the accuracy requirement is very important particularly for direct input used as reference for a servo motor. This kind of system usually represented by a non-minimum phase (NMP) system due to fast sampling-time used during modeling to avoid the lost of plant dynamic characteristics. A non-minimum phase system will have internal stability problem when the inverse closed-loop transfer function is used directly as feed-forward controller. A discrete-time non-minimum phase model posses one or more zeros outside the unit circle. However, using slow sampling-time during data acquisition process, the minimum phase plant model can be acquired.

The feed-forward Zero Phase Error Tracking Control introduced by Tomizuka [3] is effective and has been used widely in solving unstable system as in [4]–[7]. This technique is suitable only for low frequency. Many researchers have modified this technique to achieve better control performance [8] – [11].

The main objective in feedforward ZPETC is to find the optimum gain filter so that the overall gain is close to unity. To avoid the unwanted phase error, Yeh and Hsu [12], Mustafa [13] and Adnan *et al.* [14] used ZPETC without factorization of zeros polynomial. In this method, gain filter is proposed as FIR filter. The filter coefficient can be solved using comparing coefficient method as presented in [15] and Laurent series expansion method [16]. To obtain the optimum coefficients values, adaptive control technique proposed by Adnan *et al.* [14] is considered. Further improvement is presented in this paper by adding model reference as the excitation signal to find the optimum value of the coefficients.

The plant mathematical models were identified from ac servo XY table. To avoid the mechanical complexity and application of physic laws, system identification technique is used. This technique used input/output data performed in time domain to derive the model transfer function. In this part, sampling-time is varied from 45ms to 60ms to change the zero location from outside to inside the unit circle. For the purpose of results analysis, contour error is measured for each testing. Contour error is described as the minimum distance from point to point to a prescribed contour. For a constant radius circular contour, the error calculation is simple [17].

The rest of the paper is organized as follows: Section II discussed on XY table plant setup and mathematical modeling; Section III discussed the design of Adaptive feed-forward ZPETC; Section IV is on results and discussion; and finally, Section V is the concluding remark.

II. HARDWARE SETUP AND MODEL IDENTIFICATION

A. Hardware Setup

The XY table used in this paper consists of a servomotor connected to a roller shaft for each axis, two servo drivers, two displacement precision sensors, Advantech PCI 1240 motion card, Advantech PCI 1716 data acquisition (DAQ) card and a host PC set. The displacement sensor resolution is 1 μm. The minimum movement of the motor is 1μm. The hardware setup for the whole system is given in Figure 1. Microsoft Visual C++ 6.0 programming language was used for system control and interfacing. For model identification, Matlab system identification toolbox was used. For data and results analysis, Matlab software was also utilized.

The control algorithm and system interfacing is embedded in a developed C++ program. The reference input is given in the form of a mathematical equation in the program. The program is linked to DAQ and motion card for the input/output interfacing. The controlled signal will generate output voltage from DAQ card to the motor driver to control the servomotor. To have a precise measurement, the output from a precision displacement sensor is acquired instead of the displacement output from the servomotor encoder. Motion card retrieves data and sends them to the program. The constraint of the system is that the controlled signal voltage is saturated at ±10V. This will limit the speed of the motor.
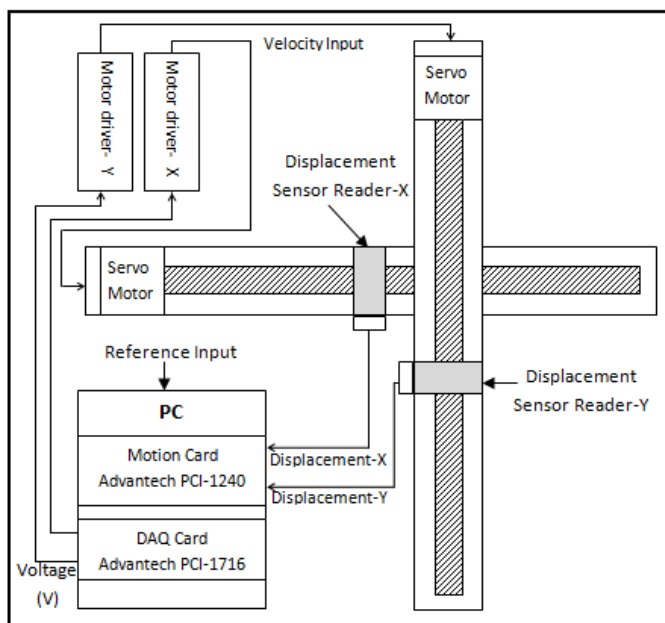

Figure 1. Hardware Setup

B. Model Identification

To acquire the system mathematical model transfer functions, system identification technique was applied. This technique models a system based on the input-output data. Data collection for input-output open-loop test of the plant was done using developed Visual C++ console programming

and data acquisition were done through Advantech PCI-1716 interface card and Advantech motion card. The input signal for both axes was generated using three different frequencies that based on Equation (1).

$$u(k) = \sum_{i=1}^{p} a_i \cos \omega_i t_s k \tag{1}$$

a_i : amplitude

ω_i : frequency (rad)

t_s : sampling time (sec)

k: integer

For model identification, a third-order ARX331 model is selected to represent the nearest model of true plant. This selection will produce satisfactory third-order transfer function. To test the performance of the controller, a minimum and non-minimum phase models will be used. The models were obtained from Open-loop input/output data response collection using sampling-time of 45ms and 60ms. Reducing the sampling time will make the zero be placed outside the unit circle. Open-loop input/output data response collection for the x and y axis is done simultaneously to achieve a perfect XY table system and to avoid dynamic miss-matching between X axis and Y axis problem [18].

The obtained non-minimum phase model using 45ms sampling-time is given by Equation (2) for x axis and Equation (3) for y axis.

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{1.132 z^{-1} + 1.1614 z^{-2} - 0.4516 z^{-3}}{1 - 0.7498 z^{-1} - 0.2873 z^{-2} + 0.03625 z^{-3}} \tag{2}$$

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{0.996 z^{-1} + 1.17 z^{-2} - 0.3076 z^{-3}}{1 - 0.7298 z^{-1} - 0.3098 z^{-2} + 0.03974 z^{-3}} \tag{3}$$

The pole and zero plots are shown in Figure 2 below.
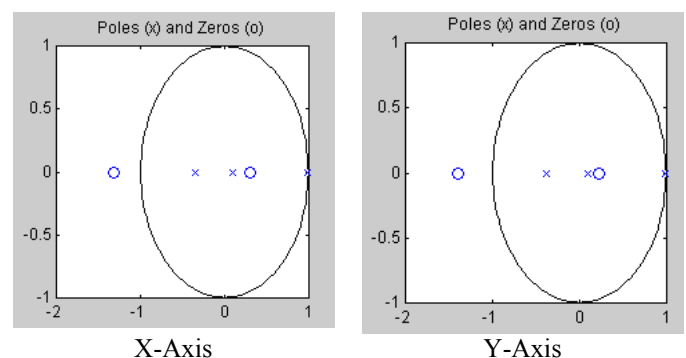

Figure 2. Zero-pole plot for x and y axis for 45ms sampling time

For the minimum phase model using 60ms sampling time, the best fits for x-axis is 93.62% and for y-axis is 96.28%. The plant x-axis mathematical model in the form of discrete-time open-loop transfer function is given by Equation (4) and for y axis is by Equation (5). The pole and zero plots are shown in Fig. 3.

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{1.729\,z^{-1} + 0.7567\,z^{-2} - 0.3407\,z^{-3}}{1 - 0.7869\,z^{-1} - 0.2983\,z^{-2} + 0.08543\,z^{-3}} \qquad (4)$$

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{1.584\,z^{-1} + 0.7458\,z^{-2} - 0.1583\,z^{-3}}{1 - 0.8051\,z^{-1} - 0.2438\,z^{-2} + 0.04917\,z^{-3}} \qquad (5)$$
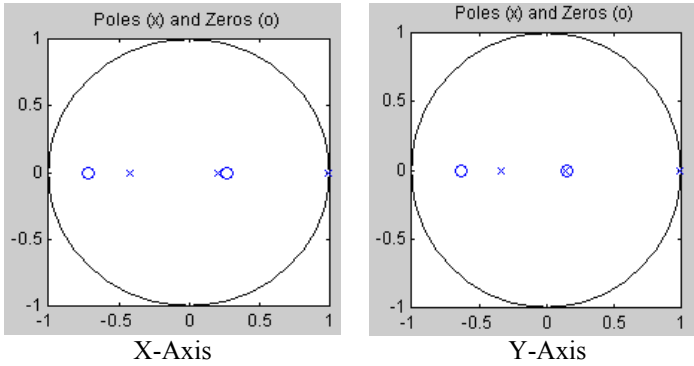
Figure 3. Zero-pole plot for x and y axis for 60ms sampling time

III. CONTROLLER DESIGN

This section describes the development and implementation of the controller design. Feedback and Adaptive feed-forward ZPETC controller are applied to the system. Feedback control has always come with the feed-forward control for 2DOF control. The control system block diagram is given in Figure 4.
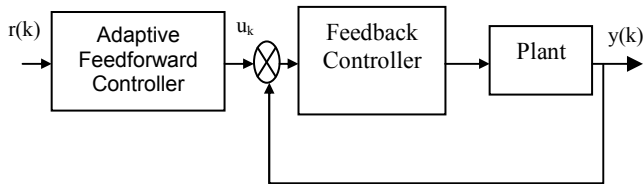
Figure 4. Control System Block Diagram

Figure 5 describes the feedback control system using pole-placement method. The pole-placement controller enables all closed-loop poles of the system to be placed at desired location. This will produce stable output performance.
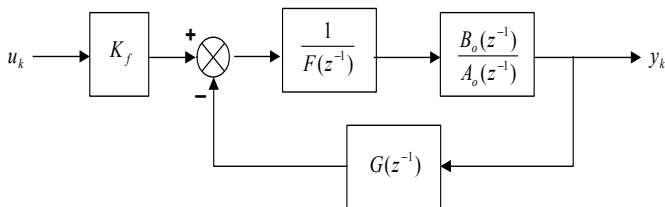
Figure 5. Pole placement feedback control system block diagram

Generally, for the feedback control, we need to know the poles location. A pole position which is inside the unity circle is considered for the controller. Other poles cancelled each others. The range of considered pole-placement controller is between 0 to 1. For slow response, pole position is set large and for fast response it is set small. Details for pole placement feedback control and how to find equation $F(z^{-1})$ and $G(z^{-1})$ can be obtained from [16].

For the Adaptive feedforward ZPETC, let the closed-loop transfer function of the feedback system be presented by:

$$G_{cl}(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} = \frac{z^{-d} B_c(z^{-1})}{A_c(z^{-1})} \qquad (6)$$

$$A_c = 1 + a_1 z^{-1} + a_2 z^{-2} + \ldots\ldots + a_{n_a} z^{-n_a}$$

$$B_c(z^{-1}) = b_o + b_1 z^{-1} + b_2 z^{-2} + \ldots + b_{n_a} z^{-n_b}$$

$$d = time\ delay$$

The ZPETC structure without factorization of zeros polynomial based on [13] is used in this paper. The control structure is shown in Figure 6.
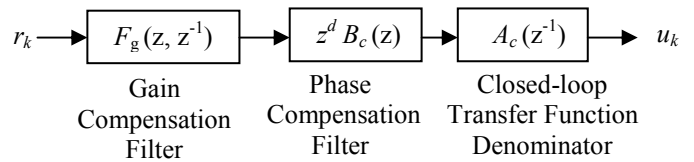
Figure 6. ZPETC without factorization of zeros polynomial

The same approach as applied in [12], [13] and [14] is followed to avoid phase error. Finite Impulse Response (FIR) symmetry filter is used. The filter can be represented by Equation (7).

$$F_g(z\ z^{-1}) = \sum_{k=0}^{n\alpha} \alpha_k \left(z^k + z^{-k} \right) \qquad (7)$$

The cost function to represent the error between desired and actual frequency response is given by Equation (8).

$$J(\alpha_i) = \left\| 1 - B_c(z^{-1}) B_c(z) \sum_{k=0}^{n_\alpha} \alpha_k \left(z^k + z^{-k} \right) \right\|_{l_2} \qquad (8)$$

The objective here is to find a set of α_k so that can minimized the cost function. By minimizing the cost function of Equation (8) close to zero, the equation can be transformed to be:

$$B_c(z^{-1}) B_c(z) \sum_{k=0}^{n_\alpha} \alpha_k \left(z^k + z^{-k} \right) = 1 \qquad (9)$$

The optimal set of α_k is estimated using the recursive least square (RLS) parameter estimation algorithm. This system can be represented by the block diagram given in Figure 7.

Figure 7. ZPETC without factorization of zeros

TABLE I. CONTOUR RMSE SUMMARY FOR ADAPTIVE FEEDFORWARD ZPETC

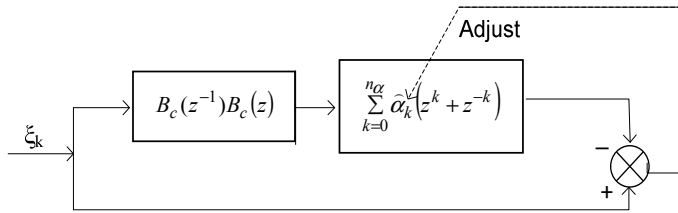| Adaptive ZPETC Filter Order (N)	Root Mean Squared Error (mm)	
	Non-Minimum Phase Model	Minimum Phase Model
2	0.1399	0.1136
5	0.1389	0.1125
10	0.1379	0.1115
15	0.1377	0.1111

Based on [14], in ensuring the generated signal ξ will always persistently exciting, a low level noise signal which its frequency spectrums close to the reference signal frequency spectrums is superimposed to the reference signal. The noise addition will not affect the system as it will not pass through the tracking system. It is used only for parameter estimation. To improve this parameter estimation part, this paper proposes model reference signal to be used instead of reference signal. The complete block diagram of the system utilizing this method is given in Figure 8.
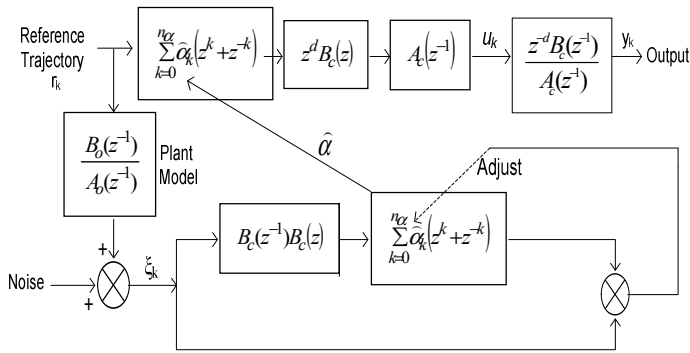

Figure 8. ZPETC without factorization of zeros

The implementation of the proposed structure was done on real-time experiment using Visual C++ console programming. The forgetting factor used was 0.95.

IV. RESULT AND DISCUSSION

The results for real-time experiments are analyzed to show the effectiveness of the proposed controller for both minimum and non-minimum phase systems. The reference input, r_k for x-axis is a sine function and for y-axis, a cosine function is used to produce circular contour in x-y plane. The equations of the input signal are given by Equation (11) for x-axis and Equation (12) for y-axis.

$$r_k = 40\sin(0.015t) \quad (11)$$
$$r_k = 40\cos(0.015t) \quad (12)$$

Experiments were conducted using same inputs to the two types of model namely, minimum phase and non-minimum phase as stated in Section II (B). This is to test the effectiveness of the controller for different types of system model. The gain compensation filter order, n_α (N) used in the experiments were 2,5,10 and 15. The root mean squared of the contour errors for all testing are summarized in Table I.

From the summarized contour root mean squared error given in Table I, the error is very small for both types of systems. The minimum phase tracking performance shows better results due to zero location away from unity circle. This imply that the proposed ZPETC controller can be applied to both types of system and much better as compared to Tomizuka [3] ZPETC that can only be use for non-minimum phase system only. The contour error difference is about 25 μm. By increasing the filter order, the contour error is reduced about 1 μm. These results show that the proposed controller can significantly control the XY table irrespective of system types. Since the contour error differences between the filter order is so small, only one plot for each model are presented. Both plots are for 15th order filter and are given in Figure 9 and Figure 10.
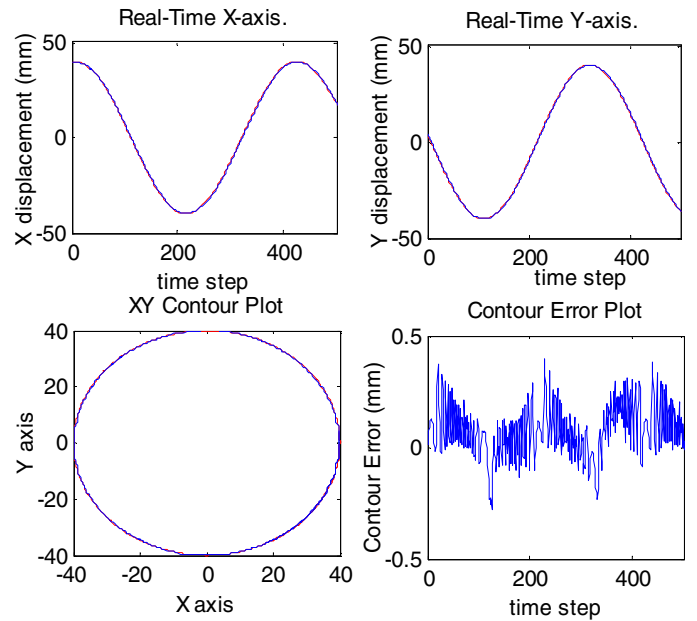
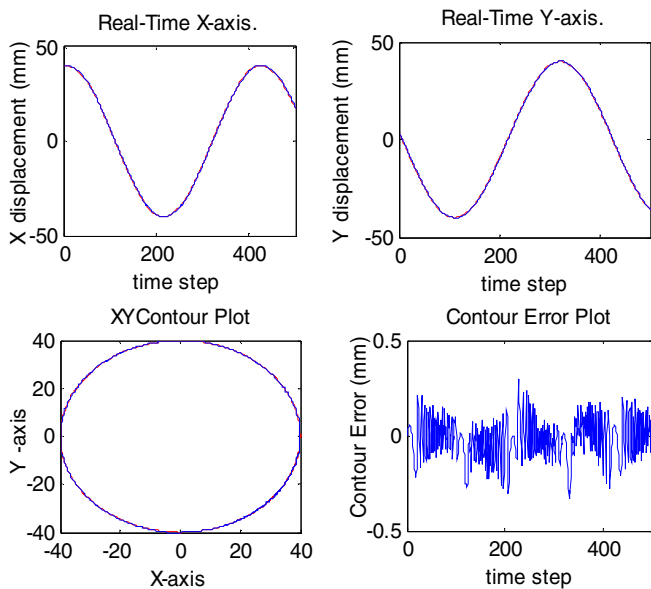Figure 9. Experimental results using 15th order filter for non-minimum phase model

Figure 10. Experimental result using 15th order filter for minimum phase model

The real-time experimental results for both models show impressive tracking performance achieved using the proposed controller. Good tracking for each axis will indirectly improve the contour performance. Better tracking performance can also be achieved by raising the filter order so that the controller will have more degree of freedom to approximate the overall transfer function to unity. Simulation results on the effect of filter order can be referred to [19]. However, for real-time application, the voltage saturation has limited the tracking performance.

V. CONCLUSION

Trajectory adaptive ZPETC was successfully developed and implemented to the XY table. The proposed controller shows that it can provide impressive tracking performance for both minimum phase and non-minimum phase systems. Reducing the tracking error will indirectly lead to better contour. The contour error for both cases are very small and minimum phase system gives slightly better performance compared to non-minimum phase system due to zero distance from the unity circle. As a conclusion, the paper verifies that by real-time experiments, the proposed controller has showed impressive tracking performance for both types of discrete-time systems.

REFERENCES

[1] R. Ramesh, M.A. Mannan and A.N. Poo, " Tracking and Contour Error Control in CNC Servo systems", International Journal of Machine Tools & Manufacture , Vol. 45, pp. 301-326, 2005.

[2] L. Wang, F. Jin and Y. Sun, "Contour Control for Direct Drive XY Table", *IEEE International Conference on Mechatronics and Automation,* August 2009, China, pp. 4919-4923.

[3] Tomizuka, M., "Zero Phase Error Trackig Algorithm for Digital Control",*ASME J.,of Dyn. Sys. Meas., and Control*, vol.109,pp.65-68, March 1987.

[4] L. J. Huang and M. Tomizuka, "A Self Paced Fuzzy Tracking Controller for Two-Dimensional Motion Control",*IEEE Transaction on Systems, Man and Cybernetics*, Vol. 20, No.5, pp. 1115-1123. 1990.

[5] H. S. Lee and M. Tomizuka, "Robust Motion Controller Design for High-Accuracy Positioning System", *IEEE Transaction on Industrial Electronics,* Vol. 43, No.1. 1996.

[6] Y. Wang, Z. H. Xiong and H. Ding, " Robust Internal Model Control with Feedforward Controller for a High-speed Motion Platform", *IEEE International Conference on Intelligent Robots and Systems*, pp. 187-192, 2005.

[7] Z. Jamaludin, H. V. Brussel, G.Pipeleers and J.Swevers, " Accurate Motion Control of XY High-speed Linear Drives Friction Model Feedforward and Cutting Forces Estimation",*Elsevier Journal on CIRP Annal -Manufacturing Technology 57*, pp. 401-406, 2008.

[8] Haack, B. and Tomizuka, M., "The Effect of Adding Zeroes to Feedforward Controllers", *ASME J.,of Dyn. Sys. Meas., and Control*, vol.113,pp.6-10, March 1991.

[9] J. Z. Xia and C. H. Menq, "Precision Tracking Control of Non-minimum Phase Systems with Phase Error", *ASME Int. J. Control*, Vol. 61, No. 4, pp. 791-807, 1995.

[10] S. Cheng and P. C. Chou, "Adaptive ZPETC Method for Dual-Stage Actuator Controller Design In Miniature Optical Disc Drive", *IEEE Transaction*, 2008.

[11] C.Y. Lin and P. Y. Chen, "Precision Motion Control of a Nano Stage using Repetitive Control and Double Feedforward Compensation", *SICE Annual Conference*, Taiwan, 2010.

[12] S.S. Yeh, & P.L. Hsu, "An Optimal and Adaptive Design of the Feedforward Motion Controller". *IEEE/ASME Transactions on Mechatronic*, 4(4), pp.428–439,1999.

[13] Mustafa, M.M., "Trajectory-Adaptive Digital Tracking Controllers for Non-minimum Phase System Without Factorization of Zeroes", *IEEE Proc.Control Theory Appl.*,vol.149, no. 2,pp. 157-162,2002.

[14] R. Adnan, A.M. Samad, and M. Mustafa, "Real-Time Control of Non-Minimum Phase System Using Trajectory Adaptive ZPETC", 7th *International Colloquium on Signal Processing & Its Application*, pp. 78-82,Malaysia, 2011.

[15] R. Adnan, A.M. Samad, N. Md Tahir, M.H. Fazalul Rahiman and M. Mustafa, "Trajectory Zero Phase Error Tracking Control using Comparing Coefficients Method", 5th *International Colloquium on Signal Processing & Its Application*, pp. 385-390,Malaysia, 2009.

[16] R. Adnan, M.H. Fazalul Rahiman and A.M. Samad, "Model Identification and Controller Design for Real-Time Control of Hydraulic Cylinder",6th *International Colloquium on Signal Processing & Its Application*, Malaysia, 2010.

[17] J. Yang, Z. Li, H. Wang, Y. Lou and Z. Long, " Direct Contour Error Compensation for Biaxial Contouring Control Systems Based on a Global Fixed Coordinate Frame", 8th World Congress on Intelligent Control and Automation, Taiwan, 2011.

[18] L.M. Wang, Q. Yang, Y.B. Sun and C. F. Liu, "Iterative Learning Cross-coupled Control for XY Table Based on Real-time Contour Error Estimation", *Advanced Materials Research*, Vol. 383-390, pp. 7054-7059, Switzerland, 2012.

[19] Ismail, H., Ishak, N., Tajjudin, M., Rahiman, M. H. F., & Adnan. R , "Robust Trajectory Tracking Controller for XY Table ", *2012 IEEE 8TH International Colloquium on Signal Processing & Its Applications (CSPA 2012)*, 23-25 March 2012

A Study on TNB Transmission Line Route Sustainability and Suitability Using GIS-AHP

Faizah Husain[*1], Nur Aishah Sulaiman[*1], Khairil Afendy Hashim[#2] and Abd. Manan Samad[*1]

Pixelgrammetry & Al-Idrisi Research Group (Pi_ALiRG)

Centre of Studies Surveying Science and Geomatics

Faculty of Architecture, Planning and Surveying

[*1]Universiti Teknologi MARA Malaysia, Shah Alam, SELANGOR

[#2]Universiti Teknologi MARA, Arau, PERLIS

Email: faizah_husain@yahoo.com ; dr_abdmanansamad@ieee.org

Abstract—**This study defines the sustainable and suitable route planning for transmission line by using Analytic Hierarchy Process (AHP) based on Geographical Information System (GIS) approached. The criteria needed to be taken into account to define the route of the transmission line were evaluated with expert who doing this business in the available system for this study. For determine of the suitable route, many factors which affect the route should be considered all together. Some of these factors are (i) technical factors (ii) environmental factors (iii) social-economic factors. Each of these factors basically corresponds to a spatial dataset. Routes defined with AHP (Analytical Hierarchy Process) to minimize economic, environmental and time cost and depending on the quality of data used.**

Keywords-Transmission line route; AHP (Analytical Hierarchy Process); GIS (Geographical Information System).

I. INTRODUCTION

Tenaga Nasional Berhad (TNB) is a private body that has tasked by the government to deliver electricity to consumers especially in Peninsular Malaysia with a low rate of disturbance possible on the system used. Transmission line is a medium to transmit total power electricity from other stations and is regarded as a project that is important in the energy business.[1] If the transmission line is not completed the electrical energy cannot be transmitted to the user. Disruption transmission line system can also affect the growth economy where investors would dispute the ability of the energy system this power if investors want to invest.

The rapid development of the country in the industrial, housing, telecommunications, information technology and infrastructure to give effect to the high electricity demand. [2] Factor increase in the construction technology, high population density, land prices are rising, modern technology and saving space, cost and energy led to a new era in the planning of electricity transmission in the present and future. [3] Such a case, a lot of power transmission lines route have been proposed to follow up the consumers demand or to connect a new power station. But due to the rapid development of rural delivery routes electrical difficulty for difficult for selection site new transmission line route to works carried out. There are several others such as the awareness of landowners, increasing of construction costs and optimum path finding [4].

Transmission line is a complex process which involves local, state and federal agencies. The route of the line has to approve by the agencies and then another agencies. The transmission line routing process a highly complex, as transmission lines are not suitable, and the public are concerned about health issues like electromagnetic field.

The optimal goal in proposed new transmission line is effectively to avoid the negative impacts on community (e.g., neighborhood and historic site), environment (e.g., wetland and floodplains), and engineering (e.g., slope and access) while ensuring reliability, safety, and cost saving for utility. [5]. In general, the process of the planning and design of transmission line consists of the following 5 phases such as planning, survey basic route, detailed route and route for implementation. [4]

Existing standards are subject to agreement with Tenaga Nasional Berhad to consider the following factors: [1]

1) Capabilities and requirements of the electricity substation for certain planning zones.
2) How to electricity distribution and type of channel.
3) The cost of supply and maintenance.

A GIS system for analysis was favored because it can perform optimal route predictions based selection by incorporating multiple influence factors into its analysis. The critical factors to address are the perspectives of society, the environment and engineering capabilities in determining the most suitable transmission line routes. GIS can be effectively used for transmission line route managing electricity distribution facilities. It has the ability to improve upon traditional practices and hence today it has become an important aspect in power delivery [6].

The type of data requires for GIS application on transmission line such as: [5]

1) Landuse map and landcover map.
2) Locations of transformers.
3) Dynamic data associated with invidual transformers.
4) Distribution of high tension wires line map.
5) Utility and network analysis based on terrain features.
6) To investigate every site per the requipment.

In environment GIS applications, generating surfaces is a frequently imposed requirement in the early stage of analyses. The importances of generating the surfaces are used as the basic information to perform further spatial analyses in various applications including site suitability of transmission line.

GIS is very suitable for use in the transmission line route as a technical tool. The planning the construction of a route involving public, this may be make a delayed decision in the approval line. Therefore, GIS is used to increase public involvement in the process of selection route and decrease in involved in the approval line. This technique can avoid public resistance to the planning and design. [7]

The AHP can be used in determining the suitability GIS-based landuse. [8] AHP decision making planning a variety of methods can be combined in each case, when a decision must be made by selecting between alternative and effective solution, this approach is used in solving a variety of different situations in which the goal is to make decision in various fields such as government, private, environmental, economic, health and education.

II.   AIM AND OBJECTIVE

The aim of this study is to find the suitable route for transmission line by using AHP in GIS approach. The objectives for this research study are:

1) To explore the application of GIS in transmission line route under TNB operation.
2) To determine and establish susceptible routes is applied to minimize the cost function to be taken into consideration for locating the transmission line tower. .
3) To determine the criteria to be taken into consideration for locating the transmission line tower.

III.   PROBLEM STATEMENT

In recent years, the problems faced by the government and private bodies in the driving force and require a lot of the time to plan and suitability of any infrastructure on site, especially if the project involved is large. Traditional way, not only requires time even long hike up the cost of project. The process of construction of transmission line towers can be divided in three main phases of the planning phase, tender phase and construction phase. There is a communication gap or do not understand each other, inaccurate data compilation, storage and collection of information is between the identified factors contribution to the delay transmission line construction project. In this sector of information and data storage in a database and graphics are the most suitable for realizing the project and obtain accurate results.

IV.   TRANSMISSION LINE ROUTING

A typical best-practice process for route selection of a transmission line is decided from the following main considerations. [9]

1) Shortest length, hence least capital cost.
2) Ease during construction and ease in maintenance of the line.
3) Requipment of future loads (substations) near the proposed route so that the line can be easily connected.
4) Required separation distance from parallel communication.
5) Avoiding of forest areas as well as wild life sanctuaries.
6) Cost of securing and clearing right of way (ROW).
7) Maintaining statutory distance from airports.

The areas to be avoided as far as possible while selecting the route of the line are as follows. [18]

1) Tough inaccessible areas where approach is difficult.
2) Towns and villages, leaving sufficient margin for their growth.
3) Areas subject to floods gushing nalas during rainy seasons, tanks, ponds, lakes and natural hazards.
4) Wooded areas with high trees or fruits bearing trees involving payment of heavy compensations for cutting of the trees.
5) Swamps and shallow land subject to flood, marshy areas, low lying lands, rivers beds, and earth slip zones, involving risk to stability to foundations.
6) High hillocks/ hilly areas/ sand dunes and areas involving abrupt changes in levels and requiring too many long spans.
7) Series of irrigation wells.
8) Rifle shooting areas and other protected areas such as army/ defence installation and ammunition depots.
9) Areas which involve risk to human life, damage to public & private properties, religious places, cremation grounds, quarry sites and underground mines, and gardens, orchards and plantations.
10) Areas which will create problems of right of way and way leaves.
11) Building/storages areas for explosives or inflammable materials, bulk oil storage tanks, oil or gas pipelines.

The evaluation process has been classified into three categories such as natural, social and technical. Each of those categories includes the following layers. [5]

- Natural Environment
  - Habitats
  - National park.
- Social Environment
  - Cultural assets
  - Temples
  - Agricultural
  - Forest area
  - District for urban planning and Airports.

- Technical Environment
  - Dangerous district
  - Snowfall, Salty breeze, Thunderstorm and Wind pressure.

Table 1: Evaluation Criteria for Appropriate point [8]

Criteria	Factors	Indicators
Technical	Executive	The observance of Authorized and legal privacy Absence of noise interferencing Points with lower angle
	maintenance	Accessibility Resistance in difficult weather condition Security
	Economic	Establishing time Minimum
Natural environment	Absence of environment danger	Earthquake Flood Hurricane
	No damaging to the natural environment	
Social environment	Observing the general rules	Avoid crossing the residential areas and public facilities Avoid crossing agricultural land and Antiquities
	Coordination with regional construction plans	

A transmission line is a complex system. It relates closely to geographic location, environment, geological condition and many other factors. It can defined in many difference data, diagrams, figures, files and other information, which are difficult to use and update them together and effective on each step of transmission line route planning, construction, operations and management by manual method. So, there is a need to raster datasets for GIS based route determining. In the process, the most important step is database design. After this, optimum route are generated via intermediate processes and produced data. The conceptual models is composed (Fig 1) [9]
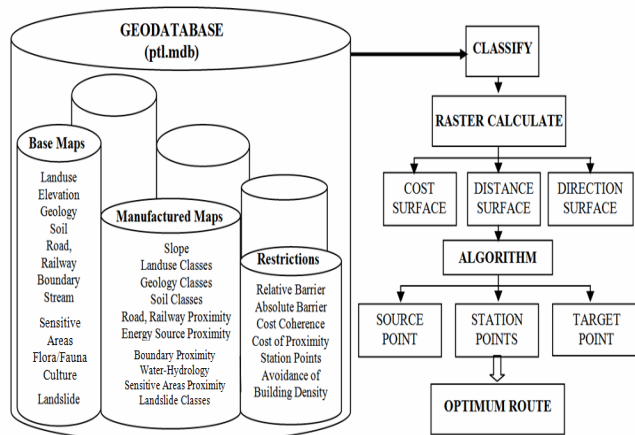
GIS data can derived based on the funneled approach of the analysis phase (Fig 2). These data can give physical barriers or assumed start and end location of the route. The dataset included landuse, building density, landslide, soil, slope, linear infrastructures, flora/fauna, protected area and lakes are composed. [1]
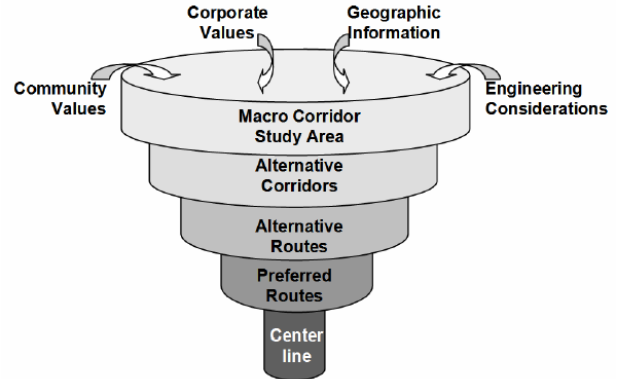

Fig 2: Funneled Approach [9].

The steps for find to optimum transmission line route are as follows: [Table 2]

1) Factor selection and raster conversion: to define the factors and factor weights that effect the transmission line route.
2) Pixel-based calculation and weight surface: to classified raster data layers and transmission line route cost can be show.
3) Absolute/relative barriers: the areas absolutely impossible to be passed thought and the area with high passage costs can be modeled as "barriers" on the surface.
4) Source/destination/stops: source and targets points
5) Optimal routing: the final result for route is obtained.

Fig 1: A conceptual model for optimum transmission line route [9].

Table 2: Factors and Sub-factors Weight [9]

Land Use	20%	Linear Infrastructure	10%
Open Land, Pasture	1	Rebuild Existing Transmission Line	1
Rocky	3	Parallel Existing Transmission Line	1.4
Forest	4	Parallel Road ROW	3.6
Agriculture	7	Parallel Gas Pipelines	4.5
Residential/Settlement	9	Parallel Railroad ROW	5
Building Density	**15%**	Background	5.5
0-0.5 Buildings/Acre	1	Future Road Plans	7.5
0.5-0.2 Buildings/Acre	3	Parallel Expressways	8.1
0.2-1 Buildings/Acre	5	Road /Rail ROW	8.4
1-4 Buildings/Acre	7	Scenic Highway ROW	9
4-25 Buildings/Acre	9	**Stream**	**5%**
Landslide	**15%**	Branch	1
Potential	7	Brook	3
Old	8	Wide-Brook	5
Active	∞	Stream	7
Soil	**10%**	River	9
Non - Agriculture	1	**Flora - Fauna**	**5%**
Agriculture	3	Flora	∞
Agriculture (Best)	7	Fauna	∞
Agriculture (The Best)	9	**Protected Area**	**5%**
Slope	**10%**	Level III	5
0 – 10	1	Level II	7
10 – 20	5	Level I	∞
20 – 30	7	**Lakes and Pond**	**5%**
> 30	9	Background	1
		Lakes and Ponds	9

V. ANALYTIC HIERARCHY PROCESS (AHP)

A. An overview of Analytic Hierarchy Process (AHP)

Analytic Hierarchy Process (AHP) is a multi criteria decision and structured techniques for organizing and analyzing complex decision. This method has been developed by Saaty (1980) and has been widely studied and refined them. More applications in AHP and it can use around the world for a wide variety of results in areas such as government, business, industry, health care, and education [10] [11]. The steps in the AHP method are as follows [8]:

1) Defines the problem and determines the desired solution.
2) Create a hierarchical structure, which begins with a general purpose, followed by a sub-criteria, criteria and possible alternatives at the lowest level criteria.
3) Make paired comparison matrix that reflects the influence of the relative contribution each element of each single purpose criteria above. Comparison based on "judgment" of the decision to assess their level of an element than other elements.
4) Doing so obtained judgment paired comparison of all nx [(n-1) / 4]
5) Calculate the eigenvalues and test the consistency, if not consistent then the data acquisition repeated.
6) Repeat steps 3, 4 and 5 for the entire hierarchy level.
7) Calculating the eigenvectors of each matrix of paired comparisons. The eigenvectors is the weight of each element. This measure is to synthesize the judgment in the determination priority of the elements at the lowest hierarchical level until the achievement of goals.
8) Check the consistency of the hierarchy. If a value greater than 10% of the data evaluation judgment should improved.

Decision-making process is basically the best to choose an alternative. As do the structuring of the problem, determination of alternatives, determination of possible values for the setting the value, preference statement to time, and the specification of preferences over risk. However the widening alternative that can be specified as well as the possibility assessment, the limitations that still surrounds the base of comparison in the form of a single criterion.

The main equipment Analytical Hierarchy Process (AHP) is to choose a functional hierarchy of the main input of human perception. The hierarchy, a complex and unstructured problems solved in groups. Then the groups are arranged into a hierarchical form. AHP compared with the other advantages are: [12]

1) Hierarchy structure, as a consequence of the selected criteria, to the deepest sub-criteria.
2) Into account the tolerance limit of validity up to the inconsistency of the various criteria and the alternative chosen by decision makers
3) Into account the endurance or resistance output sensitivity analysis of decision making. In addition, the AHP has the ability to solve problems that multi-objective and multi-criteria based on a comparison of the preferences of each element in the hierarchy.

Thus, this model is a comprehensive model of decision making. Rationale principle of AHP in solving problems with logical analysis explicit, there are three principles that underlie the AHP thinking, namely: the principle of preparing the hierarchy, the principle of setting priorities, and the principle of logical consistency.

The principle of hierarchy is set to illustrate and describe the hierarchy, by breaking the problem into elements separated. How to break down knowledge, our complex thoughts into the basic elements, then this section is broken down into its parts, and so on in a hierarchical manner. Elaboration of a lower hierarchical goal basically is aimed to obtain a measurable criterion. In certain cases, may be more advantageous to use a goal at a higher hierarchy in the analysis process. The lower in describing a goal, the easier it will determine the size of the objective criteria-criteria. However, there are times when the decision analysis process does not require a translation that is too detailed. [13]

B. The basic principles of the AHP

How does the role of the matrix in determining priorities and how to establish consistency. Set the priority of the elements by making paired comparisons, with the scale set by Saaty appeal.

Table 3: The Saaty Rating Scale [10].

Intensity of importance	Definition	Explanation
1	Equal importance	Two factors contribute equally to the objective.
3	Somewhat more importance	Experience and judgment strongly favour one over the other.
5	Much more importance	Experience and judgment strongly favour one over the other.
7	Very much more importance	Experience and judgment strongly favour one over the other. Its importance is demonstrated in practice.
9	Absolutely more importance	The evince favouring one over the other is of the highest possible validity.
2,4,6,8	Intermediate values	When compromise is needed.

Weighting matrix obtained from the comparison in pairs, must have cardinal and ordinal relations, as follows: [14]

Cardinal Relationships: a_{ij}. $a_{jk} = a_{ik}$
Ordinal Relations: $A_i > A_j$, $A_j > A_k$, then $A_i > A_k$

Relationship can be express from above in two terms as follows: [10]

1) By looking at the multiplicative preferences, such as when apples better than oranges 4 times and 2 times more delicious juice from the melon, apples better than 8 times that of the melon.

2) By looking at transitive preferences, such as apples better than oranges, grapefruit and melon better than, the more delicious apples from the melon.

In the real situation will be some deviation from the relationship, so that the matrix is not perfectly consistent. This happens because of inconsistencies in a person's preferences. To model the AHP, comparison matrix is acceptable if the value of consistency ratio ≤ 0.1. CR value ≤ 0.1 represents the value of a good level of consistency and accountability. Thus the value of CR is a measure of the consistency of a comparative matrix in pairs in the opinion. If the index of consistency is quite high, it can be revised judgment, that the RMS deviation of the line sought (aij, and Wi / wj) and revise judgment on the line that has the greatest priority. It is difficult to obtain perfectly consistent, in life, such as in a variety of specific life circumstances often affect the preferences that are subject to change. If apples are preferable to orange juice and bananas are preferable, but the same people can be more like bananas than apples, depending on time, season, and others. But to some degree of consistency in setting priorities for each element is necessary to obtain valid results in the real world. The inconsistency ratio of maximum tolerable 10%.

VI. GEOGRAPHIC INFORMATION SYSTEM (GIS)

GIS is a computer-based tool for the storage, manipulation and analysis of geographical referenced information. A geographic information system (GIS) integrates hardware, software, and data for capturing, managing, analyzing, and displaying all forms of geographically referenced information. GIS allows us to view, understand, question, interpret, and visualize data in many ways that reveal relationships, patterns, and trends in the form of maps, globes, reports, and charts. [15]

The main components of GIS are computer systems, and geospatial data users. System for GIS consists of computer hardware and software and design procedures for procurement support and data entry, processing, analysis and display of spatial data. GIS data sources are from digit maps, aerial photographs, satellite images, statistical tables and other related documents. [16]. Geospatial data can be classified into graphic data (geometric data) and data attributes (thematic data). Graphic data has three elements, namely the point (node), line (sculpture) and area (polygon) either in the form of a vector or raster topographic geometry, size, shape, position and orientation.

Fig 3: The components of GIS [16].

GIS provides the facility to address data needs for the functions discussed above. The GIS capability to handle both digital cartographic data and attribute information database related to a plan or map. GIS can store coordinate map for the location of point, linear features and areas.

These features are characteristics that must be stored in a database. When all these data are stored, both digital map information and database attributes can be processed simultaneously.

This ability is a key function in most applications that require space planning and environmental .GIS variety of physical data is expected to overcome the problem of spatial data processing. When it is done of advantages in terms of: [17]

1) Storage of spatial data efficiently and in a specific format that is easy to use.
2) Facilitate the updating and review information.
3) Geospatial data and information easily reproduced, analyzed and displayed.
4) The end result of value-added.
5) Geospatial data can be shared and exchanged freely.
6) Increase staff productivity.
7) Saving time and costs.
8) The increase in development decisions.

In generally use GIS to overcome the problems of treatment and management of geospatial data. In addition, analysis using this system can be used to improve and achieve development results and management more efficient and effective.

VII. CONCLUSION

The analytical hierarchy process (AHP), which was developed by Thomas Saaty in the 1970s, has been successfully used for organizing and analyzing complex decision. It also applied in many other problems solving such as in project planning, route selection and construction works. AHP has also been integrated into a decision-support system with geographic information system (GIS) for route selection. GIS approach allows viewing many of the data used to report summary. Other than that it also allows the decision about the identity or the information which is disclosed to clear through graphical displays in producing intermediate and final results. Besides able to display

information on transmission line tower project management, system is also capable and will assist if there are any problems in determining the exact location. GIS technology used in utility sectors is emerging as an efficient planning and decision making tools. The ability of GIS to integrate common database operations such as query and statistical analysis make it different from other traditional systems. Therefore it is valuable in planning strategies for transmission line route suitability and sustainability.

ACKNOWLEDGMENT

Pixelgrammetry and Al-Idrisi Research Group *(Pi_ALiRG)*; Research and Management Institute (RMI-UiTM); Centre of Studies Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying, UiTM Shah Alam and UiTM Perlis; Malaysia Tenaga Nasional Berhad (TNB) are greatly acknowledged.

REFERENCES

[1] T. N. Berhad, "Tenaga Nasional Berhad Corporate Report 2003," www.tnbonline.com/about/AnnualReports/AnnualReport2003.pdf
[2] D. A. P. P. E. Robert Beazer, "Intergrating Social Altitudes and Spatial Data in Sitting Power Lines using the DotSpatial Framework," depart. of Geosciences, Idaho State University, Pocatella ID.
[3] M. H. H. Zainoren Shukri, A.A.M.zin, "Application of Distance Function as Back-up Protection for Grid Transformer in TNB 500/275/132kv System," 14-16 Sep 2005 2005.
[4] J. Glasgow, French, s., Zwick, P., Kramer, L., Richardson, S., Berry and K., J., "A Consensus Method Finds Preferred Routing " 2008.
[5] M. Murata, "A GIS Application for Power Transmission Line Siting," *ESRI User Conference Proceedings, USA,* 1995.
[6] N. V. v. S. S. Kulkarni, "improving efficiency of power sectors by using GIS," *National Conference on Geo-informatics,* 2003.
[7] A. Schmidt, J, "Implementation a GIS Methodology for Sitting High Voltage Electric Transmission Line " *Papers in Resourses Analysis,* vol. 11, p. 17, 2009.
[8] M. H. Homa Faml Bahmani, Fatemeh Shabannejad, "An AHP based way to evaluate appropriate points for installing power towers and finding the best way for power transmission lines by GA Algorithm," Department of Artificial, iran, 2011.
[9] J. K. Berry, "Optimal Path Analysis and Corridor Routing, Infusing Stakeholder's Perspective in Calibration and Weighting of Model Criteria," *Berry & associates spatial information system,* 2000.
[10] T.L.Saaty, "the analytic hierarachy process (AHP)," 1991.
[11] T. L.Saaty, "Decision making with AHP," vol. 1, 2008.
[12] Y. Shu and L. Chunhua, "An enhanced routing method with Dijkstra algorithm and AHP analysis in GIS-based emergency plan," in *Geoinformatics, 2010 18th International Conference on*, 2010, pp. 1-6.
[13] H. Fang Cheng and C. Jiah-Shing, "A study on multi criteria decision making model: interactive genetic algorithms approach," in *Systems, Man, and Cybernetics, 1999. IEEE SMC '99 Conference Proceedings. 1999 IEEE International Conference on*, 1999, pp. 634-639 vol.3.
[14] S. H. M. Evangelos Triantapaphyllou, "Using the Analytic Hierarchy Process for Decision Making in Engeneering Application:Some Challenges," vol. 2, pp. 35-44, 1995.
[15] V. yildirm, "a GIS route determine in linear engineering structures information " presented at the xxiii international FIG cagress, 2006.
[16] Aronoff, "Geographic Information Systems, A Management Perspective, Ottawa Ontario, WDL publication Canada," 1993.
[17] S. V. A. K. G. Alekhya Datta, "GIS Application In Power TRansmission Line Sitting."
[18] Transmission Line Route Selection and Acquisition, (2011) http://www.transmission-line.net/2011/03/transmission-line-route-selection-and.html

The Comparison of Human Body Electromagnetic Radiation between Down Syndrome and Non Down Syndrome Person for Brain, Chakra and Energy Field Stability Score Analysis

Mastura Rosdi , Ros Shilawani Sheikh Abd Kadir , Zunairah Hj Murat and Nadiah Kamaruzaman
Faculty of Electrical Engineering
Universiti Teknologi MARA Malaysia
40450 Shah Alam, Selangor, Malaysia
ros885@salam.uitm.edu.my

Abstract - **This paper present a study on the comparison of human body Electromagnetic radiation of Down Syndrome and Non Down Syndrome persons. Frequency data were recorded using human body radiation detector from 15 persons from National Association of Down Syndrome Malaysia and 15 persons from Universiti Teknologi MARA, Shah Alam with age range between 6-28 years old. The body Electromagnetic radiation was analyzed based on three analyses which are brain analysis, chakra analysis and energy field stability score analysis. The data was analyzed using software by converting the frequency into color and categorized the human aura. From these findings, Brain Analysis indicated that Non Down Syndrome persons were slightly relax and creative by 93.33% as compared to Down Syndrome persons by 80%. For Chakra Analysis, Non Down Syndrome persons are better in communication, handling stress and their mind rather than person with Down Syndrome by 80%. For the Energy Field Stability Score Analysis, results demonstrated that 26.67% of Non Down Syndrome persons having a stable personality while 20% persons for Down Syndrome persons. Additionally, 80% of Down Syndrome persons have dynamic personalities while 73.33% for Non Down Syndrome persons. By completing this research, it can be concluded that the human body radiation detector can identify the human body Electromagnetic radiation or aura between Down Syndrome and Non Down Syndrome person.**

Keyword: Electromagnetic, Human Body Radiation Detector, Down Syndrome, Chakra

I. INTRODUCTION

Down Syndrome is most common referring to chromosomal condition associated with mental retardation [1]. This condition is caused by the existence of all or part of an extra 21st chromosome. By through chorionic villus sampling or amniocentesis during pregnancy or in a baby at birth, a down syndrome can be identified in a fetus. Chorionic villus sampling (CVS) is prenatal diagnosis to verify chromosomal or genetic disorder in the fetus [2] .The chromosomal disorder occurs in every 1 in 650 live births [1]. Regularly Down Syndrome is related with a few impairment of cognitive ability and physical growth and a particular set of facial characteristics [3,4]. Individuals with Down Syndrome usually have a low intelligence, ranging from mild to moderate disabilities.

Aura is the electromagnetic energy field that exists around the human body, plants and animals [5]. Every living organism has their own radiation whereby everyone has different frequencies among their body. The radiation field encloses the human as a sphere of radiation [6,7] and depicted as endogenous energy fields generated by and contained within the body. By the biological system, this phenomenon has been described to be electromagnetic (EM) field generated [8,9].

Human body radiation detector is a technology which gives information and objective interpretations for all auras and identifies the type and function present in specific regions of human body. Human body radiation detector accurately identifies and interprets 15 colors of bioenergy which measure the frequency at human body at various points and distances and it is basically measures in megahertz (MHz) [5].

In this research, the analysis interpretations of human body Electromagnetic radiation were divided into three parts which are Brain Analysis (point 40), Chakra Analysis (point 35-40) and the Energy Field Stability Score Analysis (point 1-40). Figure 1 shows the standard of the human body levels that was analyzed
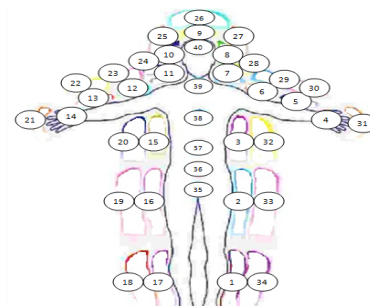

Figure 1. Standard human body level

Table I shows the colors and frequencies identified by human body radiation detector in order to analysis the human body Electromagnetic radiation.

TABLE I. THE COLORS AND FREQUENCIES IDENTIFIED BY HUMAN BODY RADIATION DETECTOR [10]

Colour	MHz	Colour	MHz	Colour	MHz	Colour	MHz
Gold	4.0-4.1	Gold	15.7-16.5	Gold	62.5-66.0	Gold	249.8-264.2
Yellow	4.2-4.3	Yellow	16.6-17.4	Yellow	66.1-69.9	Yellow	264.3-279.7
Green	4.4-4.6	Green	17.5-18.5	Green	70.0-74.1	Green	279.8-296.7
Cyan	4.7	Gray/Black	18.6-18.7	Cyan	74.2-74.9	Cyan	296.8-299.7
Gray/Black	4.8	Gray/Black	18.8-19.3	Gray/Black	75.0-77.5	Gray/Black	299.8-310.2
Cyan	4.9	Cyan	19.4-19.6	Cyan	77.6-78.5	Cyan	310.3-314.2
Blue	5.0-5.1	Blue	19.7-20.7	Blue	78.6-83.1	Blue	314.3-332.7
Navy	5.2-5.5	Navy	20.8-22.0	Navy	83.2-88.1	Navy	332.8-352.7
Purple	5.6-5.8	Purple	22.1-23.3	Purple	88.2-93.4	Purple	352.8-373.7
Orchid	5.9-6.1	Orchid	23.4-24.6	Orchid	93.5-98.6	Orchid	373.8-394.7
Silver	6.2	Silver	24.7-25.0	Silver	98.7-100.0	Silver	394.8-400.2
White	6.3	White	25.1-25.3	White	100.1-101.1	White	400.3-404.9
Burgundy	6.4-6.5	Burgundy	25.4-26.2	Burgundy	101.2-104.8	Burgundy	405.0-419.4
Red	6.6-6.9	Red	26.3-27.7	Red	104.9-111.1	Red	419.5-444.5
Rose	7.0-7.3	Rose	27.8-29.4	Rose	111.2-117.8	Rose	444.6-471.4
Orange	7.4-7.8	Orange	29.5-31.2	Orange	117.9-124.8	Orange	471.5-499.4
Gold	7.9-8.2	Gold	31.3-33.0	Gold	124.9-132.1	Gold	499.5-528.5
Yellow	8.3-8.7	Yellow	33.1-34.9	Yellow	132.2-139.8	Yellow	528.6-559.4
Green	8.8-9.2	Green	35.0-37.0	Green	139.9-148.3	Green	559.5-593.4
Cyan	9.3	Cyan	37.1-37.4	Cyan	148.4-149.8	Cyan	593.5-599.4
Gray/Black	9.4-9.6	Gray/Black	37.5-38.7	Gray/Black	149.9-155.1	Gray/Black	599.5-620.5
Cyan	9.7-9.8	Cyan	38.8-39.2	Cyan	155.2-157.1	Cyan	620.6-628.5
Blue	9.9-10.3	Blue	39.3-41.5	Blue	157.2-166.3	Blue	628.6-665.4
Navy	10.4-11.0	Navy	41.6-44.0	Navy	166.4-176.3	Navy	665.5-705.4
Purple	11.1-11.6	Purple	44.1-46.7	Purple	176.4-186.8	Purple	705.5-747.4
Orchid	11.7-12.3	Orchid	46.8-49.3	Orchid	186.9-197.3	Orchid	747.5-789.4
Silver	12.4-12.5	Silver	49.4-50.0	Silver	197.4-200.1	Silver	789.5-800.5
White	12.6	White	50.1-50.6	White	200.2-202.4	White	800.6-809.9
Burgundy	12.7-13.1	Burgundy	50.7-52.4	Burgundy	202.5-209.7	Burgundy	810.0-838.9
Red	13.2-13.8	Red	52.5-55.5	Red	209.8-222.2	Red	839.0-889.1
Rose	13.9-14.7	Rose	55.6-58.9	Rose	222.3-235.7	Rose	889.1-942.0
Orange	14.8-15.6	Orange	59.0-62.4	Orange	235.8-249.7	Orange	942.0-1000

II. LITERATURE REVIEW

A research done by Husna Abdul Rahman et.al [5] studied the analysis of correlation between body mass index and human physical condition using Resonant Field Imaging system (RFI). The objective is to investigate the relationship of different health condition with different BMI category such as underweight, normal, overweight and obese. This research have shows that aura is highly characterized and affected by the emotional and physical condition of the person. Besides that, aura is the electromagnetic energy field that exists around the human body, plants and animals. From the findings, it can be concluded that the samples with normal BMI are generally healthier. This shows that there is a clear and direct psychosomatic connection to the area of bioenergy and implies that pure and high frequency electro field energy is projected from parts of the body.

According to a research done by Gregorij Kurillo, Barbara Bajd [11] which is studied about grip force control in healthy children and children with Down Syndrome. The purpose of this study was to assess the difference in the grip force control between 10-year old healthy children and children with Down Syndrome. The research indicated that finger movements are controlled by the central nervous system which regulates the movement of the hand and arm muscles to act in synergy. In addition that, the central nervous system receives dynamic feedback information from visual sensors as well as from other body sensors placed in the skin, muscles and joints while regulating the motor output. The participants performed three diverse tracking tasks with the aim to obtain more information on their sensory-motor control associated to grasping. This research indicated that the children with Down Syndrome can develop their performance throughout tedious practice. It also shows that children with Down Syndrome lack many of the fundamental skills related to grasping such as precision of movement, coordination of force, speed and reduced efficiency in performance as compared to the healthy children.

A research done by American Academy of Pediatrics [12] studied the Health supervision for children with Down Syndrome. The purpose of this research is to study about health for children with Down Syndrome. This finding indicates that most of the Down Syndrome persons have intellectual disability in the mild which IQ 50-70 to moderate IQ 35-50 range. It also the hearing and vision impairment that would slow the performance of the person with Down Syndrome.

A study on a communication skill in children with Down Syndrome is done by L.Kumin [14]. The objective of this study is to analyze speech, language and communication for person with Down Syndrome. The findings indicated that children with Down Syndrome are usually good communicators. They are able to interact socially right from immaturity but they have specific difficulty with learning grammar as well as with developing clear speech. In the other words, those Down Syndrome children having a speech and language delay.

Another research by Roberts, JE et.al [15] studied on the language and communication development in Down Syndrome. This research indicated that persons with Down Syndrome having a different language and communication skills. The research has focused on the effectiveness of communications intervention strategies due to ear problems and hearing loss, low gain hearing aids or other amplification devices can be useful for language learning. The aid communication that often used is pointing, body language, object as well as graphics.

III. SIGNIFICANCE OF RESEARCH

Nowadays, all resources had discussed about the meaning of Down Syndrome, the contributing factors to this disease, their symptoms and characteristics. However, the research on human body Electromagnetic radiation of Down Syndrome person has not been done yet.

A theoretical study had shown that each human has a different vibration of Electromagnetic radiation. The Electromagnetic radiation gives information about human physical health, emotional and psychological condition. By having this study, it is very useful in determining the health condition of Down Syndrome and Non Down Syndrome persons. They will also lead to know the relationship between Down Syndrome and Non Down Syndrome characteristics and their aura category. Hence, the vibration of the human Electromagnetic radiation can be measured using very sensitive frequency detector and the human health condition can be identified based on the aura colors [10].

IV. METHODOLOGY

This research mainly concern on analyzing the results for the human body Electromagnetic radiation for Down Syndrome and Non Down Syndrome persons. Frequency data were recorded using human body radiation detector for 15 persons from National Association of Down Syndrome Malaysia and 15 persons from Universiti Teknologi MARA, Shah Alam with age range between 6-28 years old. After collecting all the data, it was then transferred into a program which to get the interpretation of the colors of human body aura and data were analyzed. Lastly, a conclusion had been made.

Three types of analysis were done which were brain analysis, chakra analysis and the energy field stability score analysis. While taking the readings, a control environment needs to be constructed in order to avoid any influenced by environment. Hence, it must be taken at the same location, same detector used and also stand in the open place. Figure 2 shows during the frequency measurement.

Figure 2. During frequency measurement

V. RESULT AND DISCUSSION

Below show some of the results that were obtained from the measurement. All the results were interpreted based on [10]. It was divided into three types of analysis.

A. Brain Analysis
B. Chakra Analysis
C. Energy Field Stability Score Analysis

A. Results From Brain Analysis

Brain Analysis was done to determine the predominant brain state of the persons. All the interpretations involved were referred to [10] as shown in table II. The following results were analyzed based on the frequency that is present at the Third Eye Chakra (Point 40) and their respective color was referred on table I.

TABLE II. THE MEANING FOR EACH BRAIN STATE [10]

Brain State	Meaning
Beta	Normal waking consciousness
Alpha	Moderate and relax condition
Theta	Deep sleep
Delta	Very deep sleep

Figure 3 shows the results of the Brain Analysis according to the type of brain states located. From the findings, it was obviously shows that both of the Down Syndrome and Non Down Syndrome persons fall in the Alpha brain state. According to [10], those persons who were in the Alpha brain state are meant that their brain state was usually related with the moderate and relax condition. Based on the result, the Non Down Syndrome persons were slightly higher compared to person with Down Syndrome due to Alpha Brain State. This research has been shown that although one brain state may predominant at any given time, it actually depending on the activity level of the person. The remaining brain states were present in the mix of brain state at all times. In the simple situation, while somebody is an alpha state, there were also exists in that persons brain a component of beta, theta and delta.
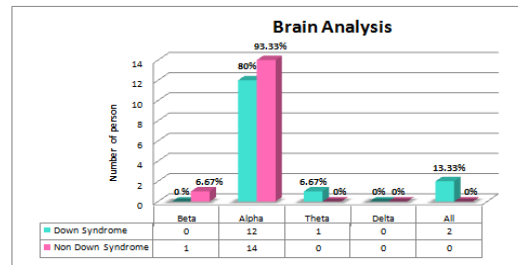

Figure 3. Brain analysis

B. Results From Chakra Analysis

There are six points of Chakra categories involved in this research which are Base Chakra (point 35), Sacral Chakra (point 36), Solar Plexus Chakra (point 37), Heart Chakra (point 38), Throat Chakra (point 39) and Third Eye Chakra (point 40). For Chakra analysis, there are detailed interpretations of color for each chakra points as provided in [10].

Table III and IV demonstrated the number of distribution persons due to each categories and colors for Down Syndrome and Non Down Syndrome persons respectively.

TABLE III. DOWN SYNDROME PERSONS DUE TO EACH CATEGORIES AND COLORS

Categories	Colors	Point 35 (Base)	Point 36 (Sacral)	Point 37 (Solar Plexus)	Point 38 (Heart)	Point 39 (Throat)	Point 40 (Third Eye)
Excellent	Gold	3	1	1		1	2
	Silver				1	2	
Good	White		1				
	Orchid		1		1	2	1
	Cyan				1	1	
Moderate	Rose	1	1				3
	Purple	2	1	1		2	1
	Blue	3	2	3	2		1
	Navy	1	2	2	4	2	4
	Green	1		1	1		
	Yellow	1	3	2	1	2	1
	Orange			1	1		1
Poor	Black		1			1	1
	Burgundy	2	1	1	1	1	
	Red	1	1	3	2	1	

TABLE IV.   NON DOWN SYNDROME PERSONS DUE TO EACH CATEGORIES AND COLORS

Categories	Colors	Point 35 (Base)	Point36 (Sacral)	Point 37 (Solar Plexus)	Point 38 (Heart)	Point 39 (Throat)	Point 40 (Third Eye)
Excellent	Gold	2	2		1		
	Silver						
Good	White				1		
	Orchid	1	1	1		2	2
	Cyan						1
Moderate	Rose			2	1		2
	Purple		4	2	1	2	1
	Blue	2	3	3	5	2	3
	Navy		1	2	2	1	1
	Green	2	1	1	1	2	1
	Yellow	1	1	1	1	3	2
	Orange		2	1	1	2	
Poor	Black	4	1	1	1	1	1
	Burgundy	2					1
	Red	1		1			

Figure 4 demonstrated the comparison of the Down Syndrome and Non Down Syndrome persons at the Base Chakra. It shows that most of the Down Syndrome persons were falls into the moderate category with 60% and the rest goes into excellent and poor categories with both 20% respectively. According to [10], base chakra was concerning the sexuality, self-identity and predominant physical or material focus in the person's life.
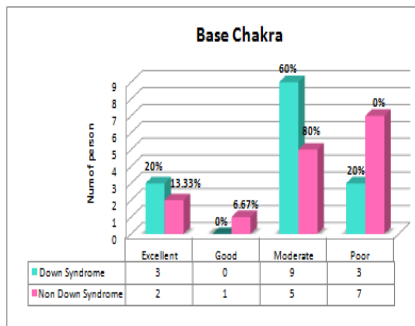
Figure 4. Comparison of Down Syndrome and Non Down Syndrome at Base Chakra

Referring to figure 5, most of the Down Syndrome and Non Down Syndrome persons were goes to the moderate category at Sacral Chakra analysis. According to [10], this category demonstrated that the persons has more related to jubilance or physical vitality and healing. It means that if they are happy, nervousness or else, it can seem and detected based on their emotional and personality and it was supported by [11]. However, the Non Down Syndrome persons were higher rather than person with Down Syndrome where 80% delegate for Non Down Syndrome whereas 60% for Down Syndrome. It means that the Non Down Syndrome persons having a better of physical vitality and healing compared to person with Down Syndrome. These seem agreed with [11-12]. It said that those peoples were always being rational and intellectual thinking in their life as well as in their relationship also as compared to Down Syndrome persons.
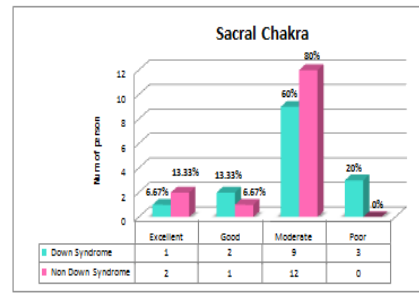
Figure 5. Comparison of Down Syndrome and Non Down Syndrome at Sacral Chakra

Figure 6 shows the comparison between the Down Syndrome and Non Down Syndrome persons at Solar Plexus Chakra. It shows that the majority of Down Syndrome and Non Down Syndrome persons were falls into the moderate category. According to [10], Solar Plexus Chakra is referring to calmness or nervousness of person. From this finding, it was shows that 80% of Non Down Syndrome persons were goes to the moderate category while 66.67% for Down Syndrome persons. It seems to agree that the Non Down Syndrome persons have a better in handling their calmness or nervousness as compared to Down Syndrome persons as supported based on [11].
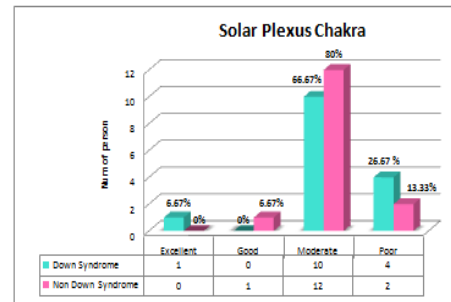

Figure 6. Comparison of Down Syndrome and Non Down Syndrome at Solar Plexus Chakra

For Heart Chakra analysis as shown in figure 7, most of the Down Syndrome and Non Down Syndrome persons were goes into the moderate category with 60% and 80% respectively. According to [10], moderate category demonstrated that the person seeks to be seriousness in their communication in their everyday life. They were also able to be open in their expression and communication of emotion.
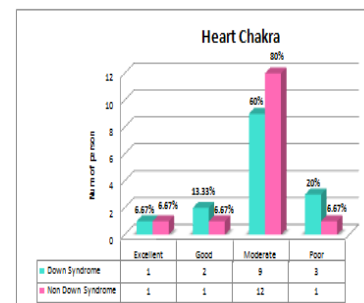

Figure 7. Comparison of Down Syndrome and Non Down Syndrome at Heart Chakra

Referring to figure 8, it was obviously shown that 80% of Non Down Syndrome persons were falls into moderate category at Throat Chakra. According to [10], the moderate category shown that those person that referring to this type of category means that they were jubilance in their communication in a joyful ways, despite of their sentiment as well as their feeling also at the time and it was supported by [13-15].
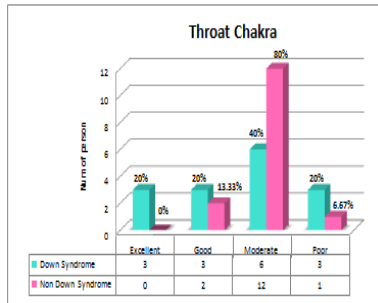

Figure 8. Comparison of Down Syndrome and Non Down Syndrome at Throat Chakra

Figure 9 demonstrated the comparison of Down Syndrome and Non Down Syndrome persons at Third Eye Chakra. From the findings, it was indicated that most of the Down Syndrome and Non Down Syndrome persons were falls into moderate category with 73.33% and 66.67% respectively. From [10], it was described that the third eye chakra was more about the clairvoyance, sensitivity, intuition and intellectual activity also the predominant mental activity as agreed in [11-12]. However, it was shown that there were none of the Non Down Syndrome persons that fall into excellent category at the third eye chakra.
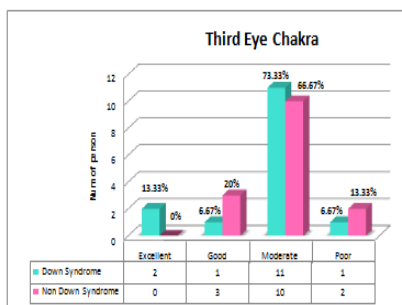
Figure 9: Comparison of Down syndrome and Non Down Syndrome at Third Eye Chakra

Based on the figure 10 and figure 11, most of the Down Syndrome and Non Down Syndrome persons were falls into moderate category. Regarding to the moderate category at Throat Chakra (point 39), it shows that Non Down Syndrome persons were higher than persons from Down Syndrome. This supported by [10,14-16] that Non Down Syndrome persons have a better in communication, handling stress and their mind as compared to Down Syndrome persons.
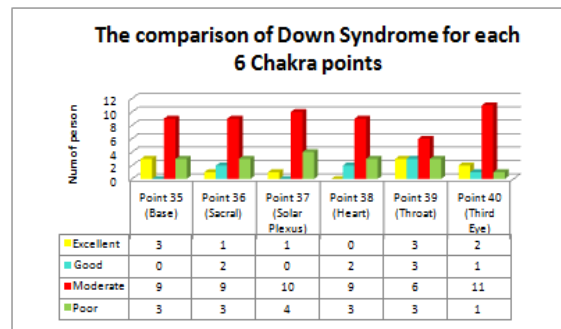
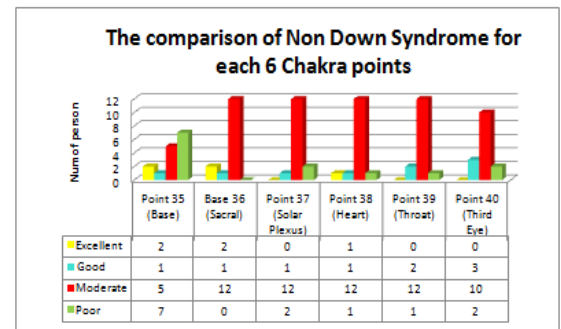Figure 10. Comparison of Down Syndrome persons for each 6 chakra points

Figure 11. Comparison of Non Down Syndrome persons for each 6 chakra points

C. *Results From Energy Field Stability Score Analysis*

For the Energy Field Stability Score Analysis, all the 40 points in the standard human body point have been measured to produce score of sub-ranges as shown in table V. It was analyzed in order to know how stable or how dynamic a person energy field is.

TABLE V. THE DESCRIPTIONS OF THE SUB-RANGE SCORE

Score of sub-ranges	Descriptions
0-50	A person has a dynamic (changing) personality
60-100	A person has a stable personality

Figure 12 determined the comparison of the Energy Field Stability Score between Down Syndrome and Non Down Syndrome persons. These findings indicated that most of the Down Syndrome persons fall in sub-ranges 0-50. According to [10], for those persons, their energy field exhibits a dynamic personality at this time. Based on the results, it shows that as much 80% of Down Syndrome persons were having a dynamic personality while 73.33% for Non Down Syndrome persons. However, a number of Non Down Syndrome person were falls in sub-ranges 60-100 were higher rather than persons from Down Syndrome with 26.67% for Non Down Syndrome and 20% for Down Syndrome person. It supported by [10] that the personality of Non Down Syndrome persons were generally more stable as compared to Down Syndrome persons. Person with Down Syndrome shows more difficulties

in the processing of facial expressions of emotions as compared to person from Non Down Syndrome. It seem to agree with the finding from [9,16-17], for this sub-ranges, the personally of persons generally shown in stable.
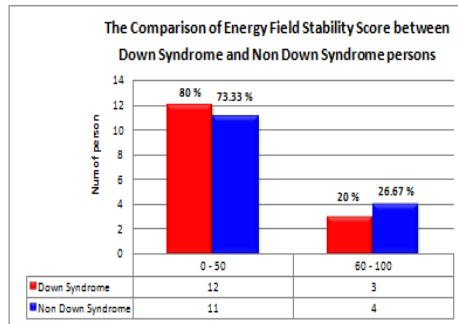

Figure 12. The comparison of the Energy Field Stability Score between Down Syndrome and Non Down Syndrome persons

VI. CONCLUSION

From these findings, Brain Analysis indicated that Non Down Syndrome persons were slightly relax and creative by 93.33% as compared to Down Syndrome persons by 80%. For Chakra Analysis, Non Down Syndrome persons having a better in communication, handling stress and their mind rather than person with Down Syndrome by 80%. For the Energy Field Stability Score Analysis, results demonstrated that 26.67% of Non Down Syndrome persons having a stable personality while 20% persons for Down Syndrome persons. Where as much 80% of Down Syndrome persons were having a dynamic personality while 73.33% for Non Down Syndrome persons. By completing this research, it can be concluded that the human body radiation detector can identify the human body Electromagnetic radiation or aura comparison between Down Syndrome and Non Down Syndrome person.

ACKNOWLEDGMENT

REFERENCES

[1] Jeremy Svendsen and Alexandra Branzan Albu, Naznin Virji-Babul, "Analysis of Patterns of Motor Behavior in Gamers with Down Syndrome," University of Victoria, BC, Canada, Down Syndrome Research Foundation Vancouver, BC, Canada.

[2] A PubMed search yields 168 papers using chorionic villous as of June 15, 2011.

[3] Gordon Grant; Peter Goward; Paul Ramcharan; Malcolm Richardson (1 May 2010). *Learning Disability: A Life Cycle Approach to Valuing People*. McGraw-Hill International. p. 43-44. ISBN 978-0-335-21439-6.

[4] This discussion by Myron Belfer, M.D., book by Gottfried Lemperie, M.D., and Dorin Radu, M.D. (1980). "Facial Plastic Surgery in Children with Down's Syndrome (preview page, with link to full content on plasreconsurg.com)". p. page 343. Retrieved 2009-07-22.

[5] Husna Abdul Rahman, Siti Nurdiana Mohamad Rameli, Ros Shilawani S. Abd Kadir, Zunairah Haji Murat and Mohd Nasir Taib, "Analysis of correlation between Body Mass Index and human physical condition using Resonant Field Imaging system (RFI)," RF and Microwave Conference 2008. December 2008

[6] M. Greenwood, "Acupuncture and The Chakras," in *Medical Acupuncture*. vol. 17, 2006, pp. 27-32.

[7] A. R. Liboff, "Toward an Electromagnetic Paradigm for Biology and Medicine," *Journal of Alternative and Complementary Medicine,* vol. 10, pp. 41-47, 2004.

[8] Siti Zura A.Jalil, Mohd Nasir Taib, Hasnain Abdullah, Megawati Mohd Yunus, "Frequency Radiation Characteristic Around the Human Body," Razak School of Engineering Universiti Teknologi Malaysia International Campus, Faculty of Electrical Engineering Universiti Teknologi MARA, IJSSST,Vol.12, No. 1, ISSN: 1473-804x online, 1473-8031 print

[9] B. Rubik, "The Biofield Hypothesis: Its Biophysical Basis and Role in Medicine," Journal of Alternative and Complementary Medicine, vol. 8, pp. 703-717, 2002

[10] "Resonant Field Imaging: Aura Brain and Imaging System", Technician's Manual form Scientific and Clinical Applications. Innovation Technologies and Energy Medicine (ITEM). 2011

[11] Gregorij Kurillo, Barbara Bajd, "Grip Force Control in Healthy Children and Children with Down Syndrome," Serbia & Montenegro, Belgrade, November 22-24, 2005.

[12] American Academy of Pediatrics Committee on Genetics (February 2001). "American Academy of Pediatrics: Health supervision for children with Down syndrome". *Pediatrics*107 (2): 442–49. DOI:10.1542/peds.107.2.442.PMID 11158488.

[13] Richards, M.C. Toward wholeness: Rudolf Steiner education in America.1980.University Press of New England, N.H

[14] L.Kumin. "Communication skills in children with Down Syndrome: A guide for parents", 1994, Rockwille, MD,Woodwill House.

[15] Roberts, JE; Price, J; Malkin, C (2007). "Language and communication development in Down syndrome". *Ment Retard Dev Disabil Res Rev* 13 (1): 26–35.DOI:10.1002/mrdd.20136

[16] R. Sudirman, T. M. Kuan, C.Y. Yong, E. Supriyanto, "Effective Support System for Language Assessment and Training of Special Children," Faculty of Electrical Engineering, Faculty of Biomedical Engineering & Health Science Universiti Teknologi Malaysia, 2010 IEEE EMBS Conference on Biomedical Engineering & Sciences (IECBES 2010), Kuala Lumpur, Malaysia, 30th November - 2nd December 2010.

[17] S. Bagnator, 1T. Neisworth, and S. Munson, (1997) LinKing assessment and early intervention: An authentic curriculum-based approach. Toronto: Paul H. Brookes Publishing.

Trajectory ZPETC Design Using Comparing Coefficients Method: The Studies on the Effects of Zero Locations to Tracking Performance

Ramli Adnan[*1], Norlela Ishak[*2], Mazidah Tajjudin[*3], Hashimah Ismail[#1], Muhammad Nursalam Mashuti[*4], and Yahaya Md Sam[**1]

*Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Malaysia
#Faculty of Engineering, UNISEL, Bestari Jaya, Selangor, Malaysia
**Faculty of Electrical Engineering, UTM, Sekudai, Johor, Malaysia
[1]ramli324@salam.uitm.edu.my
[2]norlelaishak@salam.uitm.edu.my

Abstract—**Hydraulic actuator has been widely used in industrial applications because of it properties exhibiting linear movements, fast response, smooth reversal and accurate positioning of heavy load. The integration of electronic and hydraulic technology has becoming increasingly common for precision positioning. Reducing the position tracking error of the hydraulic actuator is a challenging task but advances in control strategies have resulted in cylinders that transmit high forces with a high degree of positioning accuracy. This paper presented the design of ZPETC parameters using comparing coefficients method and studies the effects of zero locations to tracking performance. The plant model used was approximated from input-output experimental data using Matlab system identification toolbox. In order to have 2 plant models with different zero locations, 2 different sampling times were used. The studies show that tracking performance is good when the zero is far from the unity circle.**

Keywords - Feed-forward Control; ZPETC; Real-Time Control; Pole-Placement; Digital Tracking Control

I INTRODUCTION

In tracking control system, achieving a perfect tracking is what it is all about. Zero tracking error is the objective of a perfect tracking control (PTC) system and one of the methods in achieving this is by using a feed-forward controller. PTC can only be implemented to minimum phase system but not to the non-minimum phase system due to unstable zero conditions. The objective of tracking control design for non-minimum phase system is such that the plant will follow a desired path as closely as possible. This requires unity gain and zero-phase for the overall system in the relevant frequency band.

The natural nonlinear property of hydraulic cylinder had challenge researchers in designing suitable controller for motion control or tracking control. In the past few years, researchers have investigated the use of digital feed-forward controller to improve the performance of servo system[1]. The feed-forward controller is capable of cancelling all the poles and zeros, hence, creating a unity overall transfer function but in most real-world process, it is difficult to get this type of perfect system. A plant zeros located on or outside the unity circle will cause the system to be unstable and to overcome this problem, many methods have been introduced. One of the method is called Stable Phase-Zero Cancelling (SPZC) which was proposed by Tomizuka [2] but this method was unable to eliminate the phase error and the gain error left by the zero outside the unity circle. Realizing this, Tomizuka [2] proposed another method that has the ability to cancel not only all the poles and zeros but it is also capable of eliminating the phase error left by the SPZC hence increasing the controller's performance and this new method is called the Zero Phase Error Tracking Control (ZPETC). On such designs, the zero phase error tracking controller (ZPETC) cancels the closed-loop poles and cancellable zeros, at the same time, eliminates phase error induced by non-cancellable zeros.

The ZPETC can provide the overall system with frequency characteristics such that phase is zero for all frequencies and the gain is unity at only zero frequency. To avoid the unwanted phase error, Yeh and Hsu [3], Mustafa [4] and Adnan et al.[5] used ZPETC without factorization of zeros.

This paper discussed the design of ZPETC parameters using comparing coefficients method and studies the tracking performance due to the effect of non-minimum phase zero locations. The plant model used was approximated from input-output experimental data using Matlab system identification toolbox. In order to have 2 plant models with different zero locations, 2 different sampling times were used.

The rest of the sections were organized as follows: Section II discussed the Zero Phase Error Tracking Control; Section III presenting the plant models; Section IV discussed the controller design using comparing coefficients method; Section V discussed the real-time control and simulation studies; Section VI is on results and discussion; and finally, Section VII is the conclusion.

II ZERO PHASE ERROR TRACKING CONTROL

Tomizuka [6] proposed a tracking control system as shown in Figure 1 with two-degrees-of-freedom (2-DOF) controller.
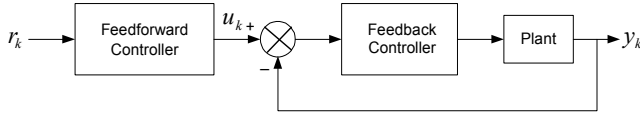

Figure 1. Tomizuka 2-DOF Controller

Without the feed-forward controller, the reference signal continuously varying and mixed with the closed-loop system dynamics, which keeps the tracking error inside the system without trying to eliminte it. The feed-forward controller is needed so that the reference signal can be pre-shaped by the feed-forward controller, so that more emphasis to the frequency components that were not properly taken care off by the feedback system can be provided [7].

The closed-loop system transfer function, Gcl without feedforward controller can be given in terms of discrete-time model:

$$G_{cl}(z^{-1}) = \frac{z^{-d}B_c(z^{-1})}{A_c(z^{-1})} \tag{1}$$

$$A_c(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a}$$

$$B_c(z^{-1}) = b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b}$$

where $n_a \geq n_b$ and d is a time delay. The factor $B(z^{-1})$ can be factorised into minimum and non-minimum phase factors.

$$B_c(z^{-1}) = B_c^+(z^{-1})B_c^-(z^{-1}) \tag{2}$$

where $B_c^+(z^{-1})$ represents the minimum phase factor and $B_c^-(z^{-1})$ represents the non-minimum phase factor. The ZPETC that proposed by Tomizuka [6] has three blocks as shown in Figure 2.
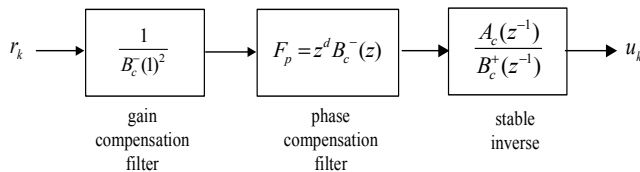
Figure 2. Conventional ZPETC Block Diagram

The improved version of ZPETC that used and implemented as feed-forward controller in this paper is shown in Figure 3.
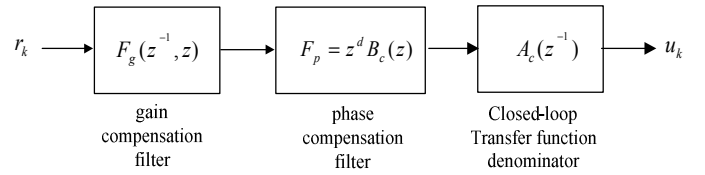

Figure 3. ZPETC without factorization of zeros

The same with any others ZPETC, the controller design mainly focused on the selection of appropriate gains compensation filter to ensure that the overall gain is unity.

The same approached that taken by Yeh and Hsu [3], Mustafa [4] and Adnan et al. [5] was used to ensure that the gain compensation filter, F_g of Eq. (3) does not introduce any phase error that might jeopardize the whole objective of this study.

$$F_g(z, z^{-1}) = \sum_{k=0}^{n_\alpha} \alpha_k (z^k + z^{-k}) \tag{3}$$

The cost function to represent the error between the desired and actual frequency response is given by:

$$J(\alpha_k) = \left\| 1 - B_c(z^{-1})B_c(z) \sum_{k=0}^{n_\alpha} \alpha_k (z^k + z^{-k}) \right\|_{l_2} \tag{4}$$

The design objective here is to find a set of α_k so that the cost function of Eq. (4) is minimized. For finite α_k, the cost function cannot be made zero for all frequencies. Thus, minimizing Eq. (4) will yield the following equation:

$$B_c(z^{-1})B_c(z) \sum_{k=0}^{n_\alpha} \alpha_k (z^k + z^{-k}) = 1 \tag{5}$$

The optimal set of α_k can be obtained by expanding Eq. (5) to polynomial of positive and negative power of z, and then compare the coefficients of the same power.

III PLANT MODEL

The plant used in this studies is an electro-hydraulic actuator that is used for quater-car system. The plant is shown in Figure 4. Using Matlab system identification toolbox, the ARX331 model can be obtained from input-output experimental data of open-loop test. In this studies, two model were approximated from two set of data with different sampling-time. The first model was approximated from experimental data of 40ms sampling time, whereas the second model was from data of 50ms sampling-time.
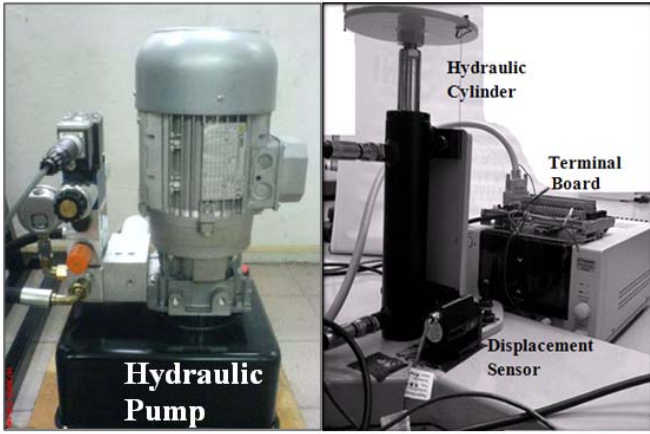
Figure 4.   Electro-Hydraulic Actuator

The obtained plant models are as follows:

Model 1:

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{0.0039z^{-1} + 0.0029z^{-2} - 0.0055z^{-3}}{1 - 1.907z^{-1} + 0.9414z^{-2} - 0.0341z^{-3}} \tag{6}$$

Eq. (6) can be simplified as follows:

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{0.0039 \, z^{-1}(1 + 0.7576 \, z^{-1} - 1.4253 \, z^{-2})}{1 - 1.907 \, z^{-1} + 0.9414 \, z^{-2} - 0.0341 \, z^{-3}} \tag{7}$$

The pole-zero plot of Eq. (7) is given in Figure 5.
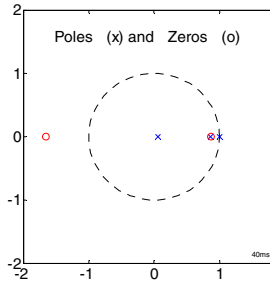
Figure 5. Pole-zero plot of Model 1

From the plot, one zero is located outside and far away from the unity circle, but, another zero which is inside the unit circle is quite near to the circle. Thus, the model obtained is a non-minimum phase model.

Model 2:

$$\frac{B_o(z^{-1})}{A_o(z^{-1})} = \frac{0.008704z^{-1} + 0.003684z^{-2} - 0.008833z^{-3}}{1 - 1.58z^{-1} + 0.3938z^{-2} - 0.1861z^{-3}} \tag{8}$$

The poles-zeros plot of Model 2 is given in Figure 6.
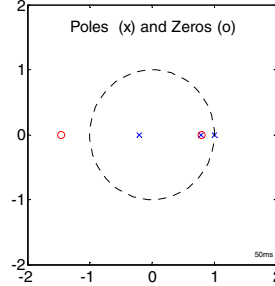
Figure 6. Pole-zero plot of Model 2

From the plot, one zero is located outside but not that far away from the unity circle as Model 1. Another zero which is inside the unit circle also is quite distance from the circle. Thus, the model obtained is also a non-minimum phase model.

Various zero position for plant model can be obtained by varying the sampling-time. A minimum phase model can be obtained using bigger sampling time whereas the non-minimum phase model can be obtained using smaller sampling time [8]. The reason why these two plant models were used in this study is to check whether the position of zero outside the unity circle affects the performance of obtaining the desired output.

IV CONTROLLER DESIGN USING COMPARING COEFFICIENTS METHOD

From Eq. (7), the optimal set of α_k for the 10th order gain compensation filter is obtained as follows:

$$B_c(z^{-1}) = 1 + 0.7576z^{-1} - 1.4253z^{-2}$$
$$B_c(z) = 1 + 0.7576z - 1.4253z^2 \tag{9}$$

From Eq. (5) and Eq. (9), the following equation is derived:

$$(1 + 0.7576z^{-1} - 1.4253z^{-2})(1 + 0.7576z - 1.4253z^2)\sum_{k=0}^{n_\alpha} \alpha_k(z^k + z^{-k}) = 1, \text{ and}$$

$$\left[3.6055 - 0.3222(z + z^{-1}) - 1.4253(z + z^{-1})\right] \cdot \left[2\alpha_0 + \alpha_1(z + z^{-1}) + \alpha_2(z^2 + z^{-2}) + \alpha_3(z^3 + z^{-3}) + \cdots + \alpha_{10}(z^{10} + z^{-10})\right] = 1 \tag{10}$$

By expanding Eq. (10) to polynomial of positive and negative power of z and compare the coefficients of the same power, the following equation is obtained:

$$
\begin{bmatrix}
7.2109 & -0.6444 & -2.8506 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-0.6444 & 2.1801 & -0.3222 & -1.4253 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-2.8506 & -0.3222 & 3.6054 & -0.3222 & -1.4253 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1.4253 & -0.3222 & 3.6054 & -0.3222 & -1.4253 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1.4253 & -0.3222 & 3.6054 & -0.3222 & -1.4253 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1.4253 & -0.3222 & 3.6054 & -0.3222 & -1.4253 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1.4253 & -0.3222 & 3.6054 & -0.3222 & -1.4253 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1.4253 & -0.3222 & 3.6054 & -0.3222 & -1.4253 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1.4253 & -0.3222 & 3.6054 & -0.3222 & -1.4253 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1.4253 & -0.3222 & 3.6054 & -0.3222 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1.4253 & -0.3222 & 3.6054
\end{bmatrix}
\begin{bmatrix}
\alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_7 \\ \alpha_8 \\ \alpha_9 \\ \alpha_{10}
\end{bmatrix}
=
\begin{bmatrix}
1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{bmatrix}
\tag{11}
$$

Solving Eq. (11) will give the values of optimum α. The values obtained are listed in Table I.

TABLE I. Optimal set of α_k for 10th order gain compensation filter for Model 1 and Model 2

k	Model 1 α_k	Model 2 α_k
0	0.3579	0.5666
1	0.3776	0.0362
2	0.4692	0.7405
3	0.3097	0.0310
4	0.3157	0.4782
5	0.2284	0.0238
6	0.2078	0.3001
7	0.1497	0.0160
8	0.1243	0.1749
9	0.0753	0.0081
10	0.0559	0.0803

The same procedure was used to obtain the optimun values of α from Eq. (8) for Model 2. From Table I, it can be observed that the values of α are converging to zero as the filter order, k increased.

V REAL-TIME CONTROL AND SIMULATION STUDIES

A. Real-Time Control

The tracking control structure that used for real-time studies is shown in Figure 7. The structure is divided into two parts: feed-forward control; and feedback control. The feed-forward control block is using the trajectory ZPETC where the controller parameters are determine using the proposed comparing coefficients method. The controller parameters as given in Table I are used in the trajectory ZPETC structure. The feedback control block is using pole-placement method to determine its controller parameters. The controllers parameters used are as follows:

Model 1:
$t_1 = 1 - 0.87z^{-1}$
$K_f = 96.59$
$F(z^{-1}) = 1 + 0.9366z^{-1} + 0.4823z^{-2}$
$G(z^{-1}) = 24.5043 + 74.5365z^{-1} - 2.9675z^{-2}$

Model 2:
$t_1 = 1 - 0.82z^{-1}$
$K_f = 50.63$
$F(z^{-1}) = 1 + 0.21896z^{-1} - 0.2165z^{-2}$
$G(z^{-1}) = 62.1614 - 6.9397z^{-1} - 4.5606z^{-}$

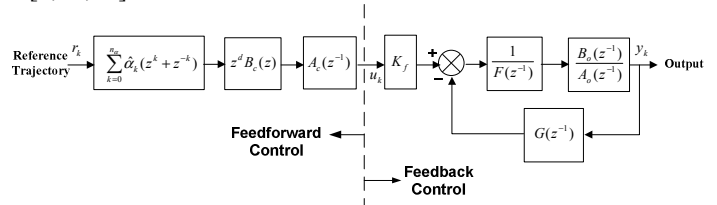The literature on calculating the given parameters is available in [9,10,11].

Figure 7. Tracking structure for real-time control

A. Simulation Studies

The trajectory zero phase error tracking control without factorisation of zeros is shown in Figure 8. Due to the effect of poles-zeros cancellation, the control structure was simplified as given in Figure 9. It can be seen that only the zero polynomial equation of the plant model is needed.
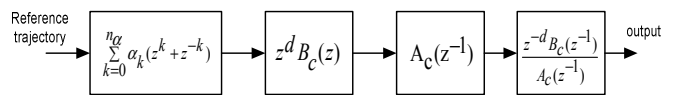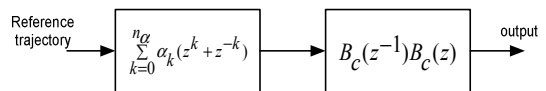

Figure 8. Trajectory ZPETC Structure

Figure 9. Tracking control structure for simulation studies

VI RESULTS AND DISCUSSION

In order to determine the correct filter order to be used in the trajectory ZPETC, the frequency response of the ZPETC given in Figure 8 are plotted using the transfer functions of Model 1 and Model 2. The resulting frequency responses are given in Figure 10 and Figure 11. From Figure 10, it can be observed that an approximate overall unity gain can be achieved when using filter order, N≥30.
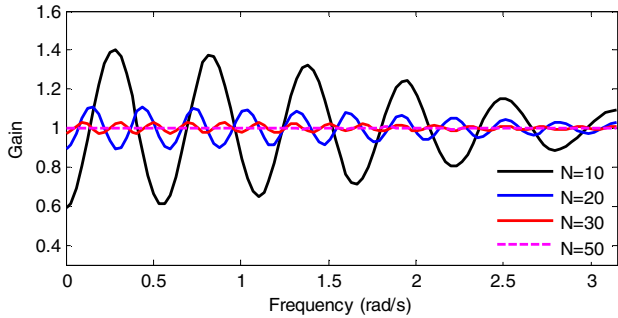

Figure 10. Frequency response of 10th,20th,30th and 50th filter order ZPETC using to Model 1
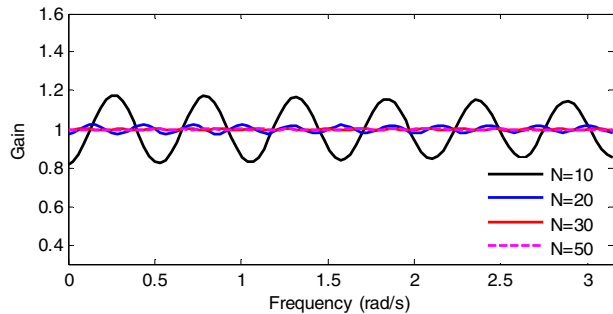
Figure 11. Frequency response of 10th,20th,30th and 50th filter order ZPETC using Model 2

From Figure 11, it can be observed that an approximate overall unity gain can be achieved when using filter order, N≥20. Thus, the degree of difficulty is harder for Model 1.

The real-time experimental results for Model 1 are shown in Figure 12 and Figure 13. Figure 12 is using filter order N=20 which yield tracking rmse of 0.1954in. Figure 13 is using filter order N=50 which yield tracking rmse of 0.1414in. The improve tracking performance can be expected from the frequency response of Figure 10.
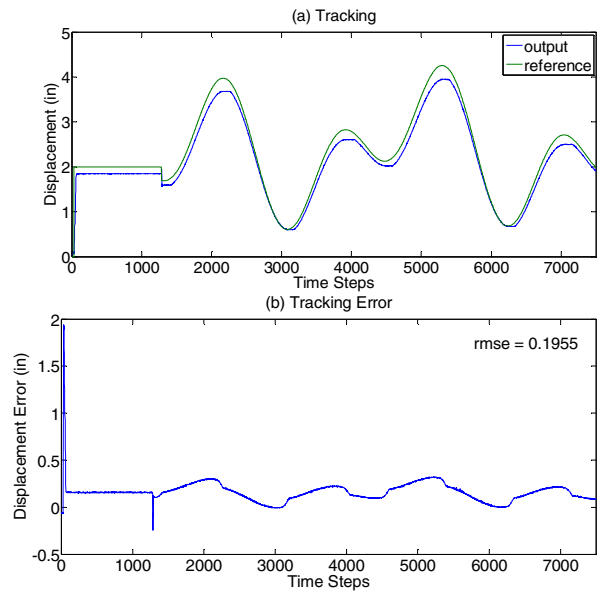

Figure 12. Real-time tracking control for Model 1 using filter order N=20 with rmse 0.1954 in
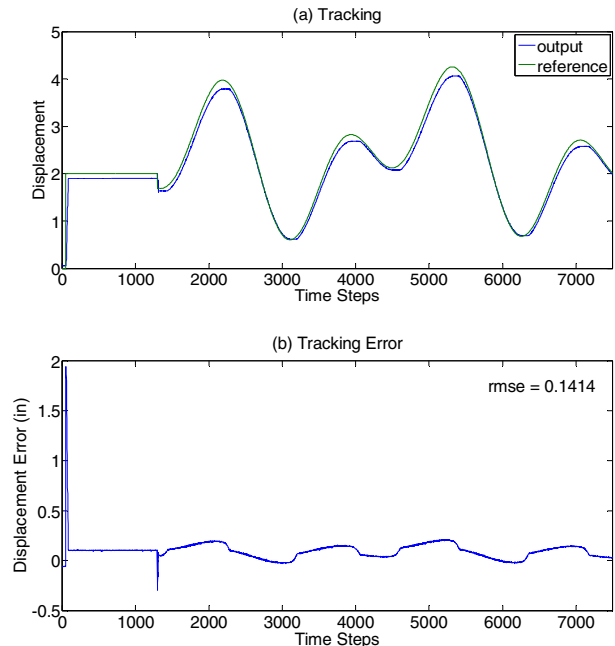

Figure 13. Real-time tracking control for Model 1 using filter order N=50 with rmse 0.1414 in

The real-time experimental results for Model 2 are shown in Figure 14 and Figure 15. Figure 14 is using filter order N=20 which yield tracking rmse of 0.1149in. Figure 15 is using filter order N=50 which yield tracking rmse of 0.0953in. The improve tracking performance can be expected from the frequency response of Figure 11.
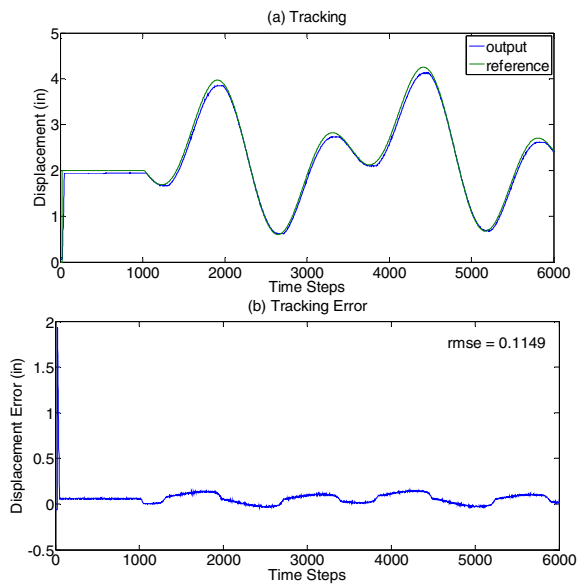
Figure 14. Real-time tracking control for Model 2 using
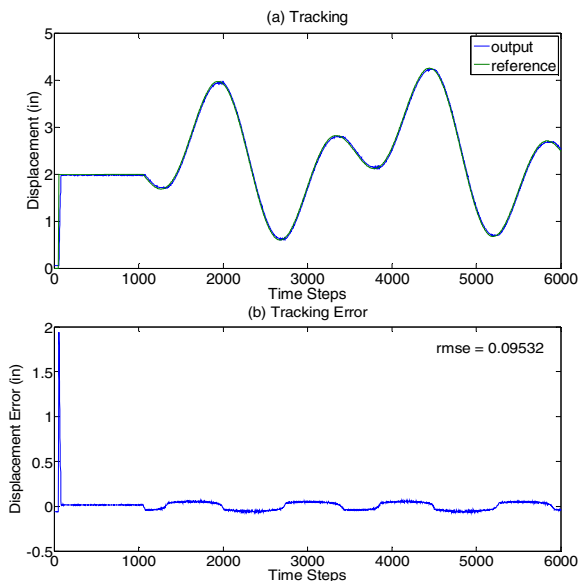filter order N=20 with rmse 0.1149 in

Figure 15. Real-time tracking control for Model 2 using
filter order N=50 with rmse 0.0953 in

VII CONCLUSION

The studies on the effect of zero locations to tracking performance of trajectory ZPETC that was designed using comparing coefficients method has been presented. Results show that when there is zero, whether minimum phase or non-minimum phase, when it is near to unit circle, the design controller could not achieved good tracking performance when low order filter is used. Low order filter could not approximate overall unity gain for the tracking system due to small degree of freedom. Higher order filter provides more degree of freedom in approximating overall unity gain for the tracking system

REFERENCES

[1] X. Liu, Y. Wu, and Y. Liu, "Digital Servo Control Based on Adaptive Zero Phase Error Tracking Controller & Zero Phase Filter", Chinese Control and Decision Conference, China, 2010.

[2] M. Tomizuka, Zero Phase Error Tracking Algorithm for Digital Control ASME Journal of Dynamic Systems, Measurement, and Control, vol. 109, pp. 65-68, 1987

[3] S.S. Yeh, & P.L. Hsu, "An Optimal and Adaptive Design of the Feedforward Motion Controller". IEEE/ASME Transactions on Mechatronic, 4(4), pp.428–439,1999.

[4] Mustafa, M.M., "Trajectory-Adaptive Digital Tracking Controllers for Non-minimum Phase System Without Factorization of Zeroes", IEEE Proc.Control Theory Appl.,vol.149, no. 2,pp. 157-162,2002.

[5] R. Adnan, A.M. Samad, and M. Mustafa, "Real-Time Control of Non-Minimum Phase System Using Trajectory Adaptive ZPETC", 7th International Colloquium on Signal Processing & Its Application, pp. 78-82,Malaysia, 2011.

[6] Tomizuka, M., "On the Design of Digital Tracking Controllers", ASME J. of Dyn. Syst., Meas., and Control, vol. 115, pp.412–418, 1993

[7] Ko. R.C., Good, M.C. and Halgamuge, S.K., "Adaptive Calibration of Feedforward Controller for Laser Profiling Machines", Proc. *Information, Decision and Control Conference (IDC '99)*, Adelaide, Australia, pp. 157–162, 1999.

[8] Astrom, K.J. and Wittenmark, B.,"*Computer-Controlled Systems:Theory and Design*", 3rd ed., Prentice Hall, Englewood Cliffs, N.J.,1997.

[9] H. Ismail, N. Ishak, M. Tajjudin, M. H. F. Rahiman & R. Adnan, "Robust Trajectory Tracking Controller for XY Table", 2012 IEEE 8th International Colloquium on Signal Processing & Its Application, pp. 153-158, Malaysia, 2012.

[10] H. Ismail, N. Ishak, M. Tajjudin, M. H. F. Rahiman & R. Adnan, "Positioning and Tracking Control of an XY Table", 2011 IEEE International Conference on Control System, Computing and Engineering, pp. 343-347, Malaysia, 2011.

N. Ishak, M. Tajjudin, H. Ismail, M. H. F. Rahiman, Y. M. Sam & R. Adnan, "Tracking Control for Electro-Hydraulic Actuator Using ZPETC", 2011 IEEE Control and System Graduate Research Colloquium, pp. 94-97, Malaysia, 2011.

A Statistical Approach for Orchid Disease Identification using RGB Color

Nurul Hidayah Tuhid, Noor Ezan Abdullah, N.M Khairi[*], M.F.Saaid, Shahrizam M.S.B, & Hadzli Hashim

Faculty of Electrical Engineering
Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia
[*]Corresponding Author: khairimhd@yahoo.com.sg

Abstract - **This paper presented a statistical approach for recognition of orchid diseases using RGB color analysis. As for features, the scale infection and black leaf spot disease of the orchid have been chosen in this study. Orchid plant with these two category disease samples were taken from a local home orchid collector and captured using digital camera in a controlled environment. The RGB components are extracted as features and statistical analysis specifically error plot and T-Test are utilized for differentiation between orchid either with scale or black leaf spot disease. Initial findings showed that the proposed method is capable to differentiate these two category diseases.**

Keywords- RGB color, statistical approach, orchid diseases, image, signature.

I. INTRODUCTION

Plant diseases are among the factors that can affect the healthy growth of the plant and further reduce the farm output. In addition, plant diseases will also contribute to additional budget in order to combat these diseases. However, identification of plant diseases is done manually using the naked eye of farmers. Farmers with vast knowledge will be the expert based on their experience whilst farmer with lesser exposure or familiarity in plant disease will prone to make mistake in identifying the type of plant diseases. Hence this method of identification is not optimal and will also cause an incorrect use of pesticides or fungicides during disease treatment [1].

As for our study, we have chosen orchid plant diseases to be experimented and analysed. This is due to the orchid nature as the second largest plant population in the world also having numerous diseases that need to be managed by the orchid farmers for optimum farm output. The orchid diseases include virus infection, bacteria and fungus [2, 3] that may infect the leaf, stem or root of orchids plant. Every disease infection can be differentiated through the sign occurred at the infected area. Hence, early accurate identification of disease signature is critical to prevent the major damage in farm by giving accurate treatment.

Amongst the type of diseases in orchid, there are possibilities of two different diseases with similar sign symptom or same signature. This situation may result in mistaken identification especially by new orchid planters. Moreover, there are also the possibility occurred between the signature of the disease infection with the signature of nutrients deficiency or cultural mistake [3]. Also, image processing technique is one of the popular methods for disease analysis and classification to identify the plant diseases [4, 6].

Hence, in this study development of technique for orchid disease identification is implemented. We used the statistical approach using RGB color to differentiate between black leaf spot disease and scale infection on Vanda orchid leaf. Even though these two diseases can be differentiated by naked eye, this study will contribute as an initial effort in development of vision system for disease identification in orchid.

II. METHODOLOGY

Figure 1 shows the research work flow. There are five phases that include sample preparation, equipment calibration, image acquisition in controlled environment, data extraction and experimental and analysis.
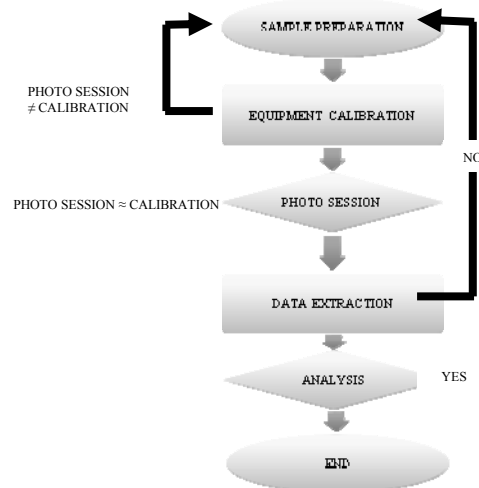

Figure 1: Research flowchart

A. Sample preparation

As mentioned earlier, a total of 119 infected leaves are collected from home gardener. The infected leaves are separated based on the two diseases to be tested which are leaf spots and scales infection. The samples are divided into two sections and resize to 10cm each.

B. Equipment calibration

The camera height is set at 41" and perpendicular to the sample being captured. The optimum distance from lens of camera to the sample is 12". If the distance is more than 12", the image of sample will reduce and inappropriate since the distance is not within tripod limitation. Spotlight is placed at 36" from the sample with the height of 64" and facing 45^0 to the sample. This position is able to offer the optimal coverage of uniform lighting and no shadows as well. The amount of light is calibration at 2363±41.877 lux.

C. Image Acquisition

This session is done in the Advanced Signal Processing (ASP) lab, Faculty of Electrical Engineering, UiTM, Shah Alam. The session is done in the morning. Images are captured under surrounding light monitored by the data logging light meter, Extech HD450. This equipment acted as the medium to store data of light value measured during photo session in lux unit. Further, the plant image is captured by Canon EOS 600D with 100 ISO speed, and 34mm focal length. The aperture is set at F4.5 and shutter speed used is 1/50 sec fixed at 5184x3456 pixels. The amount of light for image capturing session is 2371±29.789lux.

To ensure that calibration and image capturing session are viable, error plot is used. Results obtained showed that image capturing session and calibration almost overlap which mean only small differences between each other. Thus, it is concluded that the work can proceed to the next phase since the quality of images in this session is maintained as calibrated.

D. Data extraction

Next, the RGB color features are extracted after completion of the photo session. The sum of each color component (Red, Green, and Blue) was computed for every pixel lesion area to get the mean value. MATLAB program is used to extract images. Figure 2 shows the step for data extraction using MATLAB program. Firstly, the image resized from original 5184x3456 pixels to 500x500 pixels. This step is to assured that all the images' size is same. Then, RGB component is filtered using median filter to remove noise. Median filter can overcome edge blurred and protecting edge of image [6].
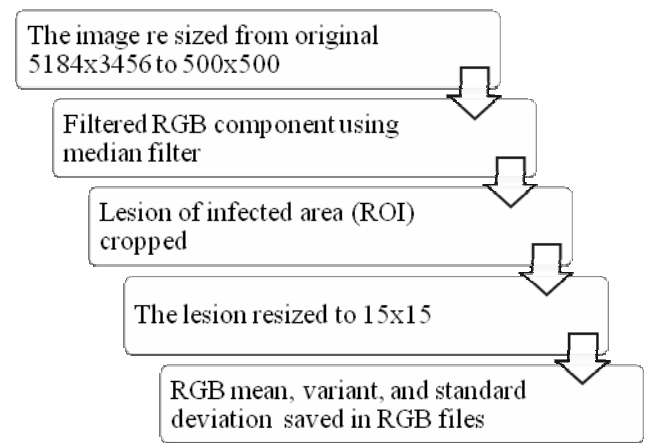

Figure 2: Procedure for feature extraction using MATLAB program.

Next, the lesion of infected area is cropped as much as the disease occurs in image. The lesion of infected area is called region of interest (ROI) which re sized to 15x15 since the lesion is too small and need precisely select. The size can assure that the selected region is ROI. Lastly, RGB mean, variant, and standard deviation saved in RGB files that have directed in the program. The data converted to Microsoft Excel program which is used to analyze in SPSS program. RGB mean is the interest data used to conduct analyses which are error bar plot and t test.

E. Statistical Analysis

Error plot is one of the available methods for demonstrating variability of data graphically. This plot can be used to indicate errors or uncertainty within the extracted features. Moreover, error plot can point out level of measurement accuracy might be [8] & [9]. Also, error plots can be used to visually compare between two quantities whether differences are statistically significant or otherwise. With error plot, it could also reveal how accurate a statistical fit the data involved in the experiment to a given function.

On the other hand, a T-test is a standard technique for measuring the statistical significance of two set of independent variables. It takes a single feature and the class associated with the data samples and measures the significance of the class variables in predicting the means of the feature [7]. The measure that T-test produces is the p-value for the feature set. With T-test, a conclusion that there is evidence that the two group means or otherwise as a basis to accept or reject the null hypothesis [10].

In this study, the SPSS software is used to perform the statistical analysis based on the features extracted from the R, G, B color component for categorization of the two disease type.

III. RESULTS AND DISCUSSIONS

In this section, the experimental results attained based on the extracted features as well as verification of identification is viable based on statistical approach will be discussed.
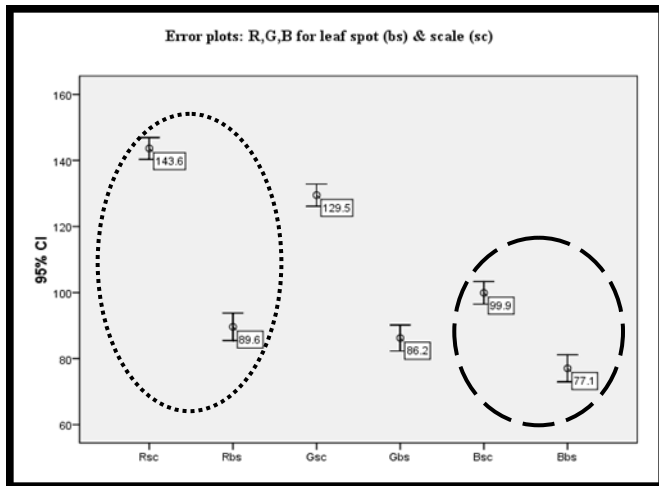

Figure 3: Error plot for RGB color of leaf spot (bs) and scale (sc)

A. Error Plot

Based on graph in Figure 3, it can be seen that the two diseases are clearly discriminated between each other with significant difference of 54, 43.3 and 22.8 for R, G and B color model respectively as indicated by the rounded dash lines. The most left is in the round dash lines is to demonstrate the significant difference between 'Red' Component of the two disease whilst the extreme right rounded dash lines is to indicate that there is a significant difference of the two disease under investigation based on 'Blue' Component as features.

Additionally, the two plots excluded from those two rounded shape represent the 'Green' component and it is clearly seen the 'green' feature vectors too contributed to significant differences. To further support this finding, inference which is t-test reinforces is implemented.

B. T- Test

Next, results based on implementation of T-test to confirm that there is a significant difference between the two category diseases will be elaborated as in Figure 4. Firstly, the mean difference is significant at level 0.05 and let:

• H0= There are no significant differences between leaf spot s disease and scale disease's mean;
• H1= There are significant differences between leaf spots disease and scale disease's mean.

Figure 4: Inferential statistic for t-test output

Figure 4 shows an inferential statistic for t test output in SPSS.The last column labeled "Sig. (2-tailed)" gives the two-tailed p value associated with the test which is the interest value to accept or reject the null hyphotesis. Firstly, at p value of 0.000, if $p \leq \alpha$, where p is equal to Sig. (2-tailed) and α, value is set to 0.05,then H_0 can be rejected.

In this analysis, the p value is less than 0.05 and hence null hyphotesis H_o is rejected and alternative hyphotesis is accpeted to confirm that there is a significant difference between these two diseases under experiment. Hence, the paired samples t test successfully revealed that statistically approach is capable to differentiate between mean of:

• pair 1: $t(172) = 20.429$, $p = 0.000$, $\alpha = 0.05$;
• pair 2: $t(172) = 16.281$, $p = 0.000$, $\alpha = 0.05$;
• pair 3: $t(172) = 8.623$, $p = 0.000$, $\alpha = 0.05$.

This implies that there are significant differences between leaf spots disease and scale disease's mean.

IV. CONCLUSION

In conclusion, the proposed feature extraction can be used to classify the diseases by RGB color analysis. Moreover, identification method based on statistical approach namely error plots and T Test analysis showed that the two diseases can be distinguished significantly with color components as feature vectors. Future work includes testing more orchid leaf diseases as sample to confirm the robustness of the proposed method in this research. Also, there are other orchid diseases that show similarities physically such as blackened leaf, yellow leaf and many more but due to different disease as well as dissimilar root cause. Finally, the proposed method will need to be evaluated in field testing or uncontrolled environment.

REFERENCES

[1] R. Pydipati, T.F. Burks, W.S. Lee,' Identification of citrus disease using color texture features and discriminant analysis', Computers and electronics in agriculture 52, Elsevier, pp 49-59, 2006

[2] Kuo-Yi Huang ,' Application of artificial neural network for detecting Phalaenopsis seedling diseases using color and texture features', Computers and electronics in agriculture 52, Elsevier, pp 3-11, 2007

[3] Orchid Society of South East Asia – OSSEA,'Book - Orchid Growing In the Tropics, Time Edition, pp 135-157, 2003

[4] A.Camargo, J.S. Smith, 'An image-processing based algorithm to automatically identify plant disease visual symptoms', Biosystem Engineering 102, Elsevier, pp 9-21, 2009.

[5] A.Meunkaewjinda, P. Kumsawat, K.Attakitmongcol and A.Srikaew. "Grape Leaf Disease Detection from Color Imagery using Hybrid Intelligent System". In the proceedings of ECTI ,2008.

[6] The Math Work, Inc. "Video and Image Processing techniques".2010 pp. 1-7.

[7] Nooritawati Md. Tahir and Hany Hazfiza Manap, 'Parkinson Disease Gait Classification based on Machine Learning Approach', Journal of Applied Sciences, 12: 180-185, 2012.

[8] Manap, H.H.; Tahir, N.M.; Yassin, A.I.M.; , "Statistical analysis of parkinson disease gait classification using Artificial Neural Network," Signal Processing and Information Technology (ISSPIT), 2011 IEEE International Symposium on , vol., no., pp.060-065, 14-17 Dec. 2011.

[9] N.Md.Tahir, Hussain, A., Mustafa, Mohd.M., Samad, S.A., Husin, H. 'Fourier descriptor for pedestrian shape recognition using support vector machine 2007 ISSPIT 2007 - 2007 IEEE International Symposium on Signal Processing and Information Technology , pp. 636-64

[10] N.M Tahir., Hussain, A., Samad, S.A., Ishak, K.A., Halim, R.A 'Feature selection for classification using decision tree', 2006 SCOReD 2006 - Proceedings of 2006 4th Student Conference on Research and Development "Towards Enhancing Research Excellence in the Region" , pp. 99-102

A Study on Flood Risk Assessment for Bandar Segamat Sustainability Using Remote Sensing and GIS Approach

Nur Aishah Sulaiman[*1], Faizah Husain[*1], Khairil Afendy Hashim[#2] and Abd. Manan Samad[*1]

Pixelgrammetry & Al-Idrisi Research Group (Pi_ALiRG)
Centre of Studies Surveying Science and Geomatics
Faculty of Architecture, Planning and Surveying

[*1]Universiti Teknologi MARA Malaysia, Shah Alam, SELANGOR
[#2]Universiti Teknologi MARA, Arau, PERLIS
Email: chicah.sulaiman@gmail.com ; dr_abdmanansamad@ieee.org

Abstract— **Flood is one of the natural disaster usually take place that in the southern states of Malaysia such as Johor, Pahang and Malacca during rainy season. The number of area affected from this event has increase over the past years. This paper describes a study made to explore the application of GIS and Remote Sensing in flood risk assessment for Bandar Segamat sustainability. The objective is to explore and understand the flood risk assessment characteristics, Remote Sensing (RS) and Geographic Information System (GIS). Open source software such as QuantimGIS and Infoworks RS is used as tools in processing the data. The result show a significant movement of water in Segamat city which causing such loss in terms of property damages and lives. A simulation of flood risk and sustainability at Bandar Segamat was developed which indicate the most, moderate and less affected area in Segamat. The main contribution of the study was to prove that anyone that involved directly or indirectly in GIS or RS field are able to perform the similar study without the need of high cost, great experience and with basic knowledge of the field study.**

Key Words: ASTER, LiDAR, DEM, Remote Sensing, Geographic Information System, Flood simulation model, Infoworks RS, QuantumGIS.

I. INTRODUCTION

Flooding is one of the natural disasters that usually occur in Malaysia. It usually occurs due to monsoon season or the heavy rainfall. Every year government has to spend huge amount money when dealing with flood disaster. During the flooding events in the past few years, Bandar Segamat has being reported as one of the tremendously affected areas in Johor when it comes to heavy rainfall. The flooding event on December 2006 has given negatives impact and great loss to the residents of Bandar Segamat district [3]. These flooding events resulted in significant losses not only to the residents but also to the government in terms of wealth, money and risk their lives. The residents have to prepare themselves physically and mentally in order to deal with these flood events every time monsoon season.

Segamat River is 23km in length with average of 40m width and 14m above mean sea level. There are 13 river networks along Bandar Segamat River include Rangla Rver, Medoi River, Kedondong River, Temangar River, Anak Kemelah River, Juasseh River, Panting River, Temalek River, Tugal River, Peta River, Tenang River, Kenawar River and Jenalin River which act as water supplier when drought or rainy season [3].

In 2008, a 3D flood modelling on "A Real Time Simulation and Modeling of Flood Hazard" in Kuala Lumpur has been developed by Jasrul Nizam Ghazali and Amiruddin Kamsin [4] from Universiti Malaya. As a result, a prototype of a flood system is developed. It has verified the usability of 3D computer graphics and fluid simulation particle. This study however is focusing on the essential and the ability of computer graphic in modeling the flood.

In this research study, the main objective is to explore the application of GIS and Remote Sensing in flood risk assessment for Bandar Segamat sustainability which indicates the most, moderate and low affected area due to flooding events. The suitability and capability of the open source techniques such as QuantumGIS in terms of conducting the environmental issue also highlighted in this study. By using both GIS and RS techniques and approaches, a flood risk simulation is generated. The PCI Geomatica, Infoworks River Simulation and QuantumGIS are used as a tool in generating the model. The data used are topography map, Remote Sensing (RS) imagery, Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and Geographic Information System (GIS) data such as Light Detection and Ranging (LiDAR) and Digital Elevation Model (DEM).

II. AIM AND OBJECTIVES

The aim of the research study is to explore the application of GIS and Remote Sensing in flood risk assessment for Bandar Segamat sustainability. The objectives for this research study are:
1) To explore and understand the flood risk assessment characteristics, Remote Sensing (RS) and Geographic Information System (GIS).
2) To develop a simulation of flood risk and sustainability at Bandar Segamat.
3) To assess, evaluate and analyze the flood risk and sustainability of Bandar Segamat.

III. PROBLEM STATEMENT

Every year, the southern states of Malaysia such as Johor, Pahang and Malacca would be preparing themselves of having possible major floods. On December 2006, it is reported that a total of nearly 60 000 residents had being moved from their residence to a safer location. In Johor, the affected areas are Segamat, Batu Pahat, Kluang and Kota Tinggi. Some of the residents claim that the flood events on December 2006 and January 2007 are one of the worst flooding event occurred at their residential area. Each year the damage estimated due to the flood disaster was around RM 1 billion. Nowadays, RS and GIS technologies usually used as a tool when dealing with environmental studies. By using the technologies, a simulation of flood risk and sustainability at Bandar Segamat can be developed. In the study area will highlighted the area which will be fully, partially or not covered with flood at a certain level of elevation. It will provide better visualization which area is affected the most by the flood disaster.

IV. FLOOD AS A NATURAL DISASTER

Flood is a natural disaster that takes place in many countries all around the world. It is a disaster phenomenon that occur when the water level arise more than it is usual state, mainly during rainy season. Flood can be defined as an area which is suddenly covered by water at a temporary time [14]. It usually occurs on a dry place where it is not regularly drown by water on its surface. According to Federal Emergency Management Agency [18], flood can be described as a temporary condition of a dry land area to be partially or completely covered by water. The water may come from overflow of rivers, mountain torrents, due to rainy season and flood from the sea in the coastal area.

Flood is a disaster that cannot be control. The time taken for an area to be flooded maybe differ depends on what are the reasons and how this flood happened. Some flood events that occur may take hours or even more than a day to develop. The residents may have ample time in preparing themselves to deal with the flood and evacuate to a safer place. However, some flood events may develop a lot faster than the other [17]. These flash flood events are very dangerous where it may cause the lost of lives. The wave may have the force in sweeping any huge object that is in its pathway such as trees and cars.

There are increasing numbers of area affected by flood disaster in these recent years. In order to cope with flood disaster, many governments agencies from all around the world and researchers have make use of technology to come out safety framework. Geographic Information System (GIS) and Remote Sensing (RS) are some of the technology that has being used for the past decades when dealing with natural disaster. According to Hassan [5] when it comes to environmental studies, Geographical Information System (GIS) technique and Remote Sensing (RS) imagery is very useful. Hassan [5] also stated that by using GIS and RS approach, the 3D view technology can be generated thus providing much better and clearer observation and realistic visualization of the flooding event.

A. Flood Hazard in Malaysia

Malaysia is located in South-East of Asia and situated near the equator line. Malaysia's climate is categorized as equatorial, being hot and humid throughout the year. Every time monsoon season arrived, the East Coast country of Malaysia such as Kelantan and Terengganu would be preparing themselves of having possible major floods that may occur in some parts of area of the states. The Southern countries of Malaysia will be facing the flooding event due to the heavy rainfall that occurs every year.

Even though Malaysia is free from natural disasters such as volcanic eruption, typhoon and earth quake, the citizens of Malaysia however need to be prepare in the possibility of facing the major flooding. On December 2006, it is reported that a total of nearly 60 000 residents had being moved to the temporary relief centers which have been provided by the government to accommodate all the flood victims. However the number increasing to almost 100 000 residents have to evacuate their house on January 2007 due to the second wave of flood occurred. Figure 3 shows flooding in Bandar Segamat on December 2007.

Figure 3 Flood in Bandar Segamat on December 2007 [21]

In Johor, the affected areas are Bandar Segamat, Batu Pahat, Kluang and Kota Tinggi. Some of the residents claim that the flood events on December 2006 and January 2007 are one of the worst flooding event occurred at their residential area. It has affected their economic contributor activities such as industrial and agriculture sectors. These flooding events that often occur during monsoon season resulted in significant losses not only to the residents but also to the government in terms of wealth, money and risk their lives. The residents have to prepare themselves physically and mentally in order to deal with these flood events every time during monsoon season. Two common floods that usually occur are monsoon flood and flash flood Monsoon flood usually occur due to the monsoon season mainly from November to March Monsoon season also affected east coast of the Peninsular of Malaysia, Northern part of Sabah and the Southern part of Sarawak which contributed to heavy rainfall [5].

According to Gasim [3], there are many factors that contribute to arising of flooding events in the country which can be divided into, nature and human activities. Nature factors include the monsoon season, the increasing of heavy rainfall due to climate change, the shallow river and urban area which is defined with saturated sediments which not allowing water to enter the sediments. The human activities factor includes the growth of unorganized squatter settlements with poor drainage [3]. The flood events may occur due to the poor drainage and the unorganized development [7]. Ghazali [4] stated that few studies shown that rapid urbanization in major cities as the main cause. Typhoon Utor was blamed for contributing the heavy rain in four states in southern of

Malaysia which include Johor, Malacca, Pahang and Negeri Sembilan [15]. Every year, Malaysia government has to spend RM 1 billion due to flooding event [8].

According to National Security Council [13], the flooding event on 2006 and 2007 was one of the costliest flooding events that take place in the country which have cost almost RM1.1 billion. It was reported by BBC NEWS [16] that on January 2007 more than 100000 residents were evacuated from their home due to heavy flooding. There are more than 300 relief centers were set up in order to accommodate the flood victims. The flood disaster has result in damaging the economic growth of the affected area and causing the residents loss their property and belongings [4]. The disaster leads to massive destructions to the buildings, transportations, houses, landscapes and plants. Figure 4 shows the distribution of flood prone areas in Peninsular Malaysia**.**


Figure 4 Distribution of flood prone areas in Peninsular Malaysia [5]

V. REMOTE SENSING (RS) AND GEOGRAPHIC INFORMATION SYSTEM (GIS) APPROACH IN FLOODING EVENTS APPLICATION

Most of the environmental studies data came from Geographic Information System (GIS) and Remote Sensing (RS) [4]. It is widely used nowadays when dealing with natural disaster such as flooding and land slide. Geographical data such as Light Detection and Ranging (LiDAR) and Digital Elevation Model (DEM) and remote sensing imagery are some of the substantial data that can be used when it comes to environmental studies.

Light Detection and Ranging (LiDAR) can be used in many applications which include in flood plane mapping. It has become one of the major sources of digital terrain information [9]. By using GPS methods, LiDAR data is collected with reference to the ground data control points [1]. Raber [9] added that LiDAR has being widely used in the application of determine and acquiring the elevation data for trees, shrubs, buildings and Earth's surface. It acts as one of the Geographic Information System (GIS) data that can be used to generate Digital Elevation Model (DEM). In order to implement a hydrodynamic model, the most important information is roughness or surface data, topographic and terrain data. According to Tarekegn [12], in representing the river terrain geometry and floodplain topography, the topographic data is usually posed by Digital Elevation Model (DEM). Figure 5 shows the digital elevation model DEM.
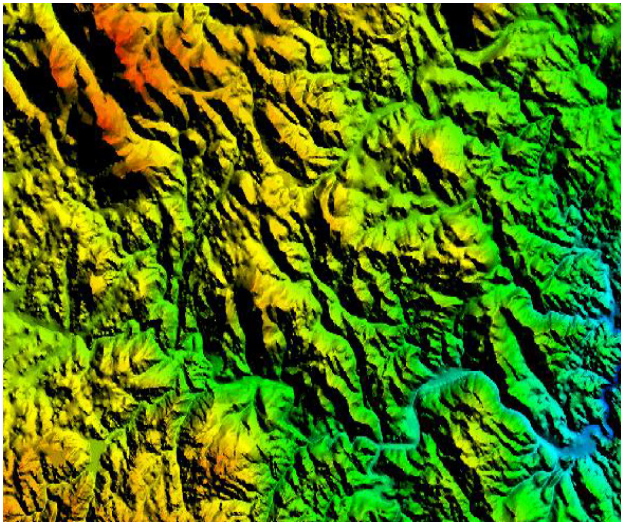
Figure 5 Digital Elevation Model (DEM) [19]

Chang [2] defined Digital Elevation Model (DEM) as the representation of elevation points in a regular array. Satelite imagery, radar data and LiDAR data are some of the alternatives sources of DEM that can be found [2]. Two stages are involve in generating the Digital Elevation Model (DEM) which include the production of 3-Dimensional Triangular Irregular Network (TIN) and digital terrain model in a rectangular block of cells, GRID [10]. Ghazali [4] also added that there was a research done by University of Texas proves that thirty two out of fifty state highway agencies that are related to hydrology and highway work responded on implementing GIS for mapping and data management. Ghazali [4] stated that environmental experts have agreed that GIS technology played such an important role in generating real-time simulation.

Remote sensing has become one of the substantial tools in dealing with flooding events as GIS. Most researchers found that remote sensing data provide a real time data which allowing the preparation of maps on limited time [6]. Aerial photograph is likely the most reliable source of remotely sensed flood area data. Due to the highly cost associated with airborne acquisitions the researchers are always tried to find the alternative ways for flood monitoring and management [11]. According to Sanyal [10], Landsat Multi Spectral Scanner (MSS) with 80 m resolution was the earliest data available for satellite imagery. The main investigation on the field was concentrated on flood prone area of United States of America (USA). In these past years, it can be seen that the researches started to try new alternatives of satellite imagery data for flooding event. One of the data that is frequently used nowadays in studying the flooding events is Spaceborne Thermal Emission and Reflection Radiometer (ASTER) which is developed by National Aeronautics and Space Administration (NASA) and Japan government. According to Terekegn [12], DEMs of ASTER are well known as having high accuracy in near-flat regions and smoothly sloped areas.

VI. RESEARCH METHODOLOGY

The research methodology for this research study is organized into five main stages which includes stage 1: project planning, stage 2: literature review, stage 3: data acquisition, stage 4: data processing, and stage 5: data analysis and representation. In stage 1 (project planning), it is started by stating the aim, objectives and the problem statement of the research study. At this stage, the selection of the study area is made. Bandar Segamat is selected as the study area due to the flooding events that gradually occurred for the past few years. Stage 2 (literature review) is where all the relevant journals,

articles and any references is collected act as reviews to the previous papers that has been studied. Stage 3 is where the data acquisition process being highlighted. The data used are LiDAR, ASTER satellite image and topography map of Bandar Segamat. Stage 4 explained on the results of the sata processing. PCI Geomatica and QuantumGIS software is used as tools in processing the data. The results of the research study will be analyze and being represented in the form of map and simulation of flood risk model. Figure 6 shows the flowchart data processing steps used in this research study.
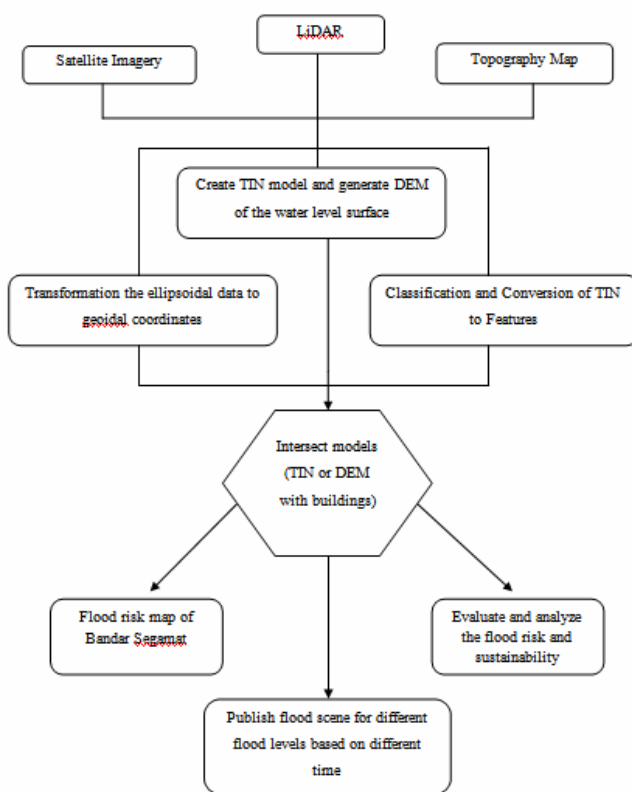
Figure 6 Data Processing Steps

VII. CONCLUSION

Flooding is the most common natural disaster occurs in Malaysia. Every year the government need to spend almost a billion Ringgit due to flood event all over the country. There are studies done by researches on how to decrease the effect of the flooding event through flood risk management, mapping and flood modelling. The usage of GIS and RS technologies in natural disaster are increasing for the past few years. This can be proof by the

growth in number of studies in flood hazard modeling in Malaysia that are using the GIS and RS application. In this research, it is focusing on the suitability and effectiveness of GIS and RS tools in order to be implemented in environmental studies. Due to the development of technologies nowadays, it is not possible for the integration of GIS and RS technologies to be extended to more users that may come from different field of study. With basic knowledge on the field of study, with appropriate tools being used and low cost spend when conducting the application, every one who came from different background of study can exploring them selves in this field. The main contribution of the study was to prove that anyone that involved directly or indirectly in GIS or RS field are able to perform the similar study without the need of high cost, basic knowledge and great experience.

ACKNOWLEDGEMENT

REFERENCES

[1] Adda, P., Mioc, D., Anton, F., Mcgillivray, E., Morton, A., Fraser, D., Eb, C., et al. (2010). 3D FLOOD-RISK MODELS OF GOVERNMENT INFRASTRUCTURE. *WebMGS 2010:1st International Workshop on Pervasive Web Mapping, Geoprocessing and Services, XXXVIII-4,* (pp. 6-11). Como, Italy.

[2] Chang, K.-tsung. (2010). *Introduction to Geographic Information System* (Fifth.). New York: McGraw-Hill.

[3] Gasim, M. B., Surif, S., Mokhtar, M., Ekhwan, M., Toriman, H., Rahim, S. A., & Bee, C. H. (2010). Flood Analysis of December 2006: Focus at Segamat Town, Johor. *Sains Malaysiana, 39*(December 2006), pp. 353-361.

[4] Ghazali, J. N., & Kamsin, A. (2008). A Real Time Simulation and Modeling of Flood Hazard. *12th WSEAS International Conference on SYSTEMS, , July 22-24, 2008* (pp. 438-443). Heraklion, Greece.

[5] Hassan, A. J., Ab Ghani, A., and Abdullah, R. 2006. Development of flood risk map using GIS for Sg. Selangor Basin. *In: Proceeding of the Sixth International Conference on ASIA GIS, 9-10 Mar 2006, UTM Skudai, Johor ,Malaysia*

[6] Kale, V. (2009). Research on Floods in India. *Available on www. iypeinsa. org/updates-09/art-7. pdf,* 2003-2005. Retrieved from http://www.iypeinsa.org/updates-09/art-7.pdf

[7] Ngai Weng Chan, (1995),"Flood Disaster Management in Malaysia: An Evaluation of the effectiveness of government Resettlement Schemes", Disaster Prevention and Management, Vol. 4 Iss: 4 pp. 22–29

[8] Norrudin, N. H. R. bin. (2007). *Kajian kejadian banjir di sungai skudai pada 2006 dan 2007*. Universiti Teknologi Malaysia.

[9] Raber, G. T., Jensen, J. R., Hodgson, M. E., Tullis, J. A., Davis, B. A., & Berglund, J. (2007). Impact of Lidar Nominal Post-spacing on DEM Accuracy and Flood Zone Delineation. *THE AMERICAN SOCIETY FOR PHOTOGRAMMETRY AND REMOTE SENSING*, *73*(7), pp. 793-804.

[10] Sanyal, J., & Lu, X. X. (2004). Application of Remote Sensing in Flood Management with Special Reference to Monsoon Asia: A Review. *Natural Hazards*, *33*(2), 283-301. doi:10.1023/B:NHAZ.0000037035.65105.95

[11] Schumann, G., Bates, P., & Horritt, M. (2009). Progress in integration of remote sensing-derived flood extent and stage data and hydraulic models. *Reviews of GeoPhysics*, *47*(2009), pp. 1-20. doi:10.1029/2008RG000274.1.INTRODUCTION

[12] Tarekegn, T. H., Tamiru Haile, A., Rientjes, T., Reggiani, P., & Alkema, D. (2009). Assessment of an ASTER-generated DEM for 2D hydrodynamic flood modeling. *International Journal of Applied Earth Observation and Geoinformation*, *12*(6), pp. 457-465.

[13} Majlis Keselamatan Negara, (2009),
"*Pengurusan Bencana Tanggungjawab Bersama*".
(On-line, Retrieved 6 May 2012)
http://www.mkn.gov.my/mkn/mampucms/upload/fokus/PENG URUSAN BENCANA TANGGUNGJAWAB BERSAMA.pdf

[14] Official Journal of the European Union, (2007),
"*Diretive 2007/60/EC of the European Parliament and of the Council*".
(On-line, Retrieved 6 May 2012)
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:288:00 27:0034:EN:PDF

[15] The Star (2006), "Typhoon Utor to Blame,"
(On-line, Retrieved 6 May 2012)
http://thestar.com.my/news/story.asp?file=/2006/12/21/nation/16 380186&sec=nation

[16] BBC NEWS (2007), "Malaysia flood test aid efforts;"
(On-line, Retrieved 5 May 2012)
from http://news.bbc.co.uk/2/hi/asia-pacific/6262125.stm

[17] "Floods." (On-line, Retrieved 5 May 2012)
http://environment.nationalgeographic.com/environment/natural-disasters/floods-profile

[18] "Definition on Flood."
(On-line, Retrieved 5 May 2012)
http://www.fema.gov/business/nfip/19def2.shtm

[19 "Digital Elevation Model (DEM)."
(On-line, Retrieved 5 May 2012)
http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/ DEMs

[20] "Banjir Menyerang Kembali di Segamat Johor, (2007),
(On-line, Retrieved 25 June 2012)
http://norzaidi.blogspot.com/2007/12/banjir-menyerang-kembali-di-segamat.html

GUI System for Enhancing Blood Vessels Segmentation in Digital Fundus Images

Ahmad Zikri Rozlan, Nor Syazwani Mohd Ali, Hadzli Hashim.
Faculty of Electrical Engineering
Universiti Teknologi MARA
Malaysia

Abstract: **An automated image processing system has the potential to aid ophthalmologist in diagnosing eye retinopathy diseases by presenting changes of retina features. One of the algorithms involved in the process is blood vessels segmentation. This paper proposes development of a Graphical User Interface (GUI) system that produces enhancement of blood vessels segmentation in digital fundus images. The system can help ophthalmologist in improving morphological procedures by observing significant features in the processed images for early diagnosis. The process of GUI creation is by using Qt Creator software while blood vessels segmentation methods utilize Canny and other morphology-based techniques. The outcome is a better processed image that can be the basis for further image enhancement so that eventually a diagnostic tool for diabetic retinopathy can be produced.**

Keywords: digital fundus image, blood vessels, image processing, Graphical User Interface.

I. Introduction

Diabetes is one of the chronic diseases that have no cure. According to Malaysian Diabetes Association, in 2007 nearly 1.2 million people in Malaysia have diabetes. Regrettably, more than half of them did not aware they have the disease [1]. Due to diabetes, patients are likely to have other complications such as skin disorder, nerve damage, or even eye damage known as diabetic retinopathy (DR).

In Malaysia, DR is the commonest cause of vision lost in adult of working age [2]. It is a disease which occurs due to excess glucose in the bloodstream. The excess glucose weakens the tiny blood vessel in the retina, resulted the blood vessel to swell and bleed. This can affect eyesight, which can ultimately lead to blindness. Therefore, early detection of DR is very important as it can aid in controlling severity of diabetes thus preventing blindness. Current practice in Malaysia for DR screening is using direct ophthalmoscopy [2]. Such method is labor intensive and time consuming. However in ophthalmology, digital fundus camera provides fundus image data which can be used for computerized detection of DR. By combining with an advanced image processing technique, a quick and effective mass screening for DR would be highly possible.

In order to aid ophthalmologist to diagnose eye retinopathy diseases better, graphical user interfaces (GUI) software system is proposed in this work. This research delineates a step along the path to providing a user friendly diagnostic support tools for the early detection of DR. We focus on the blood vessel in the retinal fundus image that is affected by diabetes. The proposed GUI is created using the Qt Software Development Kit.

II. Background

There are several features of DR such as exudates, blood vessels, haemorrhages and microaneurysms. Among other structures, retinal blood vessels are the most significant structures in fundus images [3]. They contain information which can offer lot of useful parameters for the diagnosis systemic diseases. However, fundus images generally needed to be pre-processed in order to correct the problems occur from nonuniform illumination and the existence of noise. Lili Xu *et. al.* [3] and Youssef *et. al.* [4] proved that in RGB image, the green component has the highest contrast between the blood vessels and the background compared to the blue and red components image. Furthermore, green component is less noisy compared to other two thus, it contains significant information that can be extracted from the fundus image. Due to this advantage, each sample was later transformed into greyscale image for further segmentation of blood vessels. As a result, the blood vessels appear darker in the grayscale image. This fact is strongly supported when the authors in [5] used green channel of RGB components in their wavelet analysis because it displayed the best vessels or background contrast.

There were many techniques being proposed in image pre-processing process. The authors in [4] pre-processed the images using median filtering followed by contrast enhancement using top-hat morphological operations to eliminate noise. Cornforth, *et. al.* [5] used mean filter of sized 5x5 pixels for pre-processing input images in order to reduce noise effects.

After the images had undergone pre-processing phase, blood vessels extraction was done through application of segmentation technique. There are several segmnentation methods used in the detection of retinal blood vessels. The most common are Canny and Morphology edge detection. Both methods were proposed by many authors in this related work. In [4], a simple morphological closing operations was described to extract the blood vessel tree. By subtracting two morphological closing with different sizes structering elements has resulted in brighter areas of blood vessels and

darker background with contrast higher than that of the original image. Meanwhile Kumari *et. al.* [6] have extracted blood vessel of various thicknesses using morphological opening and closing operations. The authors proposed that subtraction of closed images across two different sizes disk shaped structuring element has resulted in the blood vessel segmentation of image.

Canny edge detector is the most widely used compared to morphological matched filter as applied by Xiaolin *et. al.* [7]. However, the authors suggested to use bilateral filter rather than Gaussian filter to enhance the Canny edge detector's performance. Bilateral filter does not only compute weight of distances between pixel locations, but also considered the differences between gray values of the center pixel and its neighborhood pixels. Thus the Gaussian noises and the vessels background noises can be eliminated using the bilateral filter.

Nowadays, there are many types of equipment used to aid ophthalmologists in their medical work. Therefore, lot of methods was also being proposed by researcher in order to detect diseases such as DR. However, GUI vision system is still scarce; particularly a system that can compiles all methods together and embedded in such that it can be more users friendly. Vision system that is based on Web-telescreening platform and called as DrishtiCare was proposed by Joshi GD *et. al.* [8] This platform integrated a variety of value-added fundus image analysis components and the service offered as a Web-based platform so that it can be more cost-effective, easy to use, and scalable. The authors designed a practical interface which enabled with new visualization features that helped in examination cases by experts.

In addition, VAMPIRE system was proposed by Perez-Rovira *et. al.* [9]. This software allows user to load a set of images of various random sizes. Through its interface, the user can modified the results efficiently, corrected any wrong measurements or neglect any images. All measurements and results about the images can be saved in Excel files for further numerical analysis. With VAMPIRE, user could detect features of the vasculature such as vessel width, branching angles, and vessel tortuosity.

III. Materials & Methods

A. Image Acquisition

For blood vessels segmentation, digital fundus images were taken from the DRIVE database [10]. DRIVE is a database from the internet that is valid to be used by any researchers globally for their research work in this area. It consists of 40 samples of fundus images, together with manual segmentations of the vessels. The 40 images have been divided into two set, 20 for training and 20 for test. The entire images have been manually segmented by an

ophthalmologist. All the images in this DRIVE are captured using Canon CR5 non-mydriatic 3CCD camera at 45◦ field of view (FOV). The images are in standardize size of 768×584 pixels, 8 bits channel and have a FOV of approximately 540 pixels in diameter. The images are compressed in JPEG-format.

B. Blood Vessels Segmentation

1) Canny's Algorithm

Canny algorithm was applied in the detection of blood vessels in this work. The smoothing filter using Gaussian filter is the first step needed in canny edge detection as described in Fig. 1. After the image has been convolved with a symmetric Gaussian, the edge direction is estimated from the gradient of the smoothed image intensity surface [11]. Then the gradient magnitude is non-maximum suppressed in that direction. Thresholds were set based on local estimates of image noise using that algorithm. It makes use of two thresholds which were upper threshold and lower threshold [12]. When a pixel has a value above the upper threshold, then it is set as an edge pixel. Meanwhile a pixel that has a value more than the lower threshold and also was the neighbor of an edge pixel, it is set as an edge pixel as well. A pixel is not set as an edge pixel when the pixel's value is below the lower threshold or above the lower threshold but not the neighbor of an edge pixel.
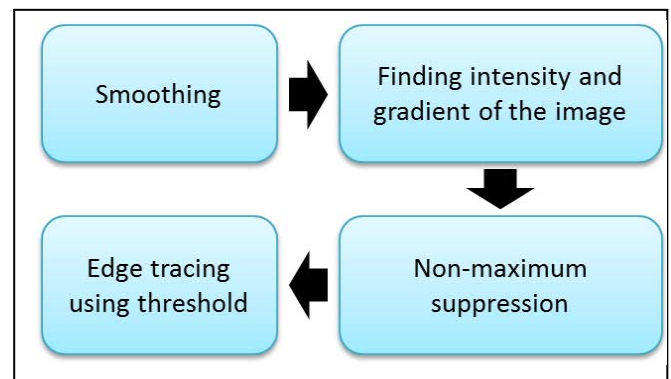

Fig. 1: Steps taken in Canny algorithm

2) Morphology's Method

To detect blood vessels we used morphology opening and closing operations because they can extract the blood vessel of various thicknesses. Since green component was recommended to be used [4], higher contrast of blood vessels that appears as dark regions in brighter background were narrowed further by morphological dilation. Fig 2 shows the fundus image being transformed into the respective RGB color components. The very small dark area should be eliminated from the image if the dilated image is then eroded using the same structuring element while the larger region returns to their initial size [4].
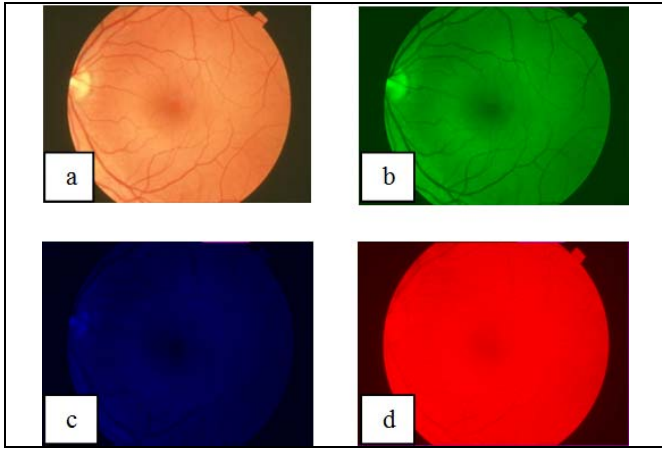
Fig. 2: (a) original image, (b) green component, (c) blue component, (d) red component

Morphological closing is defined as dilation followed by erosion. A disk shaped structuring element is applied in this method. The dilation and erosion of A by B is defined as [4]:

$$A \oplus B = A_1(x, y) = \max\left(A(x-i, y-j) + B(i, j)\right) \qquad (1)$$

$$A \ominus B = A_2(x, y) = \min\left(A(x-i, y-j) + B(i, j)\right) \qquad (2)$$

$$where \quad i, j \in B$$

The vessels have disappeared when closing by larger element. Thus, the subtraction of closed images against two diverse sizes has resulted in brighter areas of blood vessels and darker background with contrast higher than that of the original image. The result is the segmented blood vessel of the image [6]. The subtraction operation is as follows:

$$C' = \left(A \oplus B_2\right) \ominus B_2 - \left(A \oplus B_1\right) \ominus B_1 \qquad (3)$$

To segment the entire blood vessel, the difference between the two structuring elements is set at 4 pixels. The image C' was later thresholded and by using Laplacian and Gaussian (log) operator for significant edge detection. After that, morphological thinning (erode) was performed to obtain the skeleton of blood vessel.

C. *Additional function in GUI system*

A user friendly GUI application using Qt software was used to provide a friendlier platform. This user friendly may help the medicine's people in their work to do the research on the DR. Moreover, through this software, each channels of the input images for red, green and blue component conversion is provided in a button or menu form. Besides noise filtering buttons, additional feature such as histogram is also provided. From the histogram, values of mean, median, mode, variance, and standard deviation can be displayed for numerical investigation in further future work.
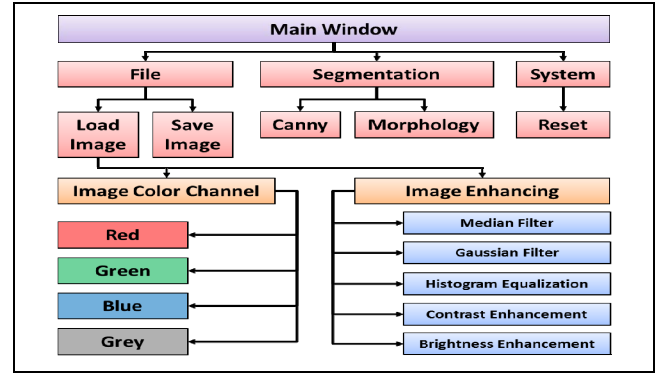
IV.   Result

Fig 3: Vision system in this project

Fig. 3 shows the concept of the vision system developed in this project so that to make sure that system is arranged consistently and can be run smoothly without any problem.
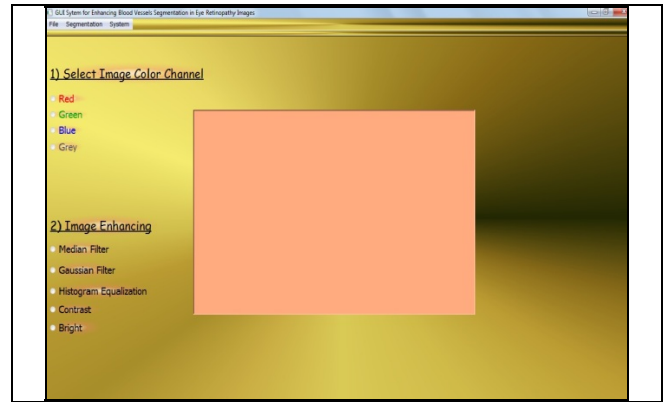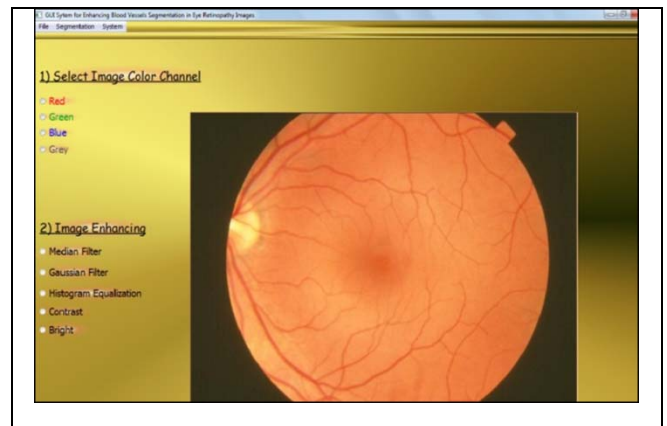

Fig 4: The "Main Window'

Fig. 5: Input image appeared after clicked the "OPEN IMAGE" button.

After running the application, a window appears as shown in Figure 4. Before user can upload an image, all other buttons except "open image" button are disabled. These buttons are enabled only after the selected image is uploaded as in Fig. 5. The next stage involves options given

for users to choose the various type of image enhancement. They can either choose median filter, Gaussian filter or contrast.

In this project, median filter was set to a 3x3 neighbourhood sized. This is the default parameter setting. As for the Gaussian, the neighbourhood size was set to 5x5. For contrast enhancement, the default multiplication value setting was set to 2. Basically contrast enhancement works by multiplying each pixel in the image with the set parameter.

Histogram equalization was proposed as an alternate method to increase the image contrast. The images displayed are as in Fig. 6. In addition, there are buttons which allow user to select the color channel of image as their choice. There are 4 buttons available which one of it set for red, green, blue and grey color. One button is created to reset the image.

(a)   (b)
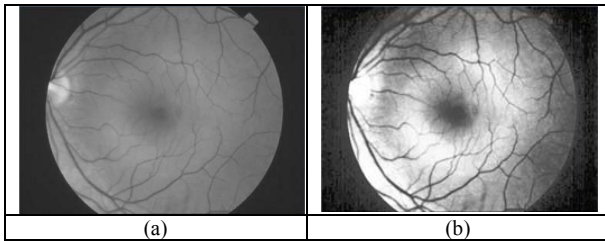
Fig 6: (a) Grayscale image, (b) grayscale image after histogram equalization

Besides, user also able to choose the histogram displayed as their extra information. There are 3 buttons available which one of its present red color channel histogram, blue color channel histogram and green color channel histogram. The images displayed are as in Fig. 7 below.
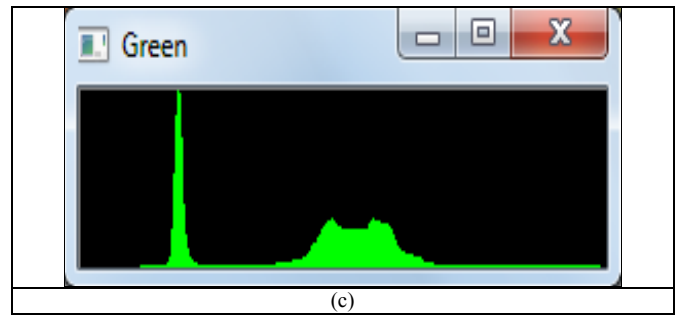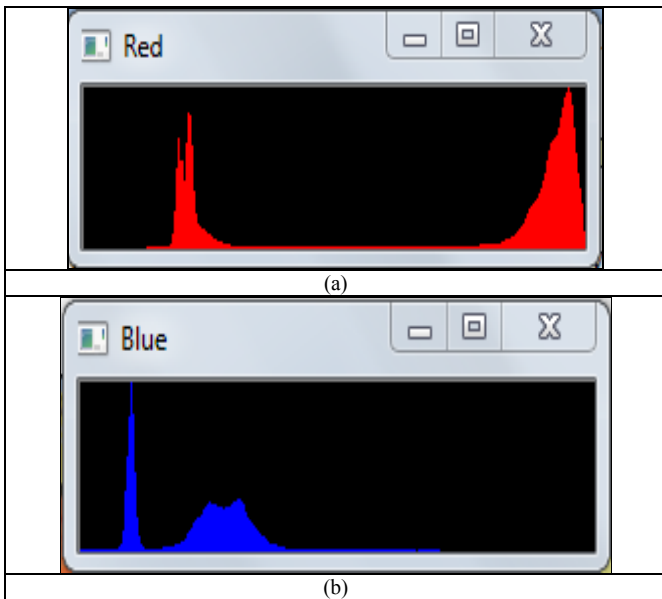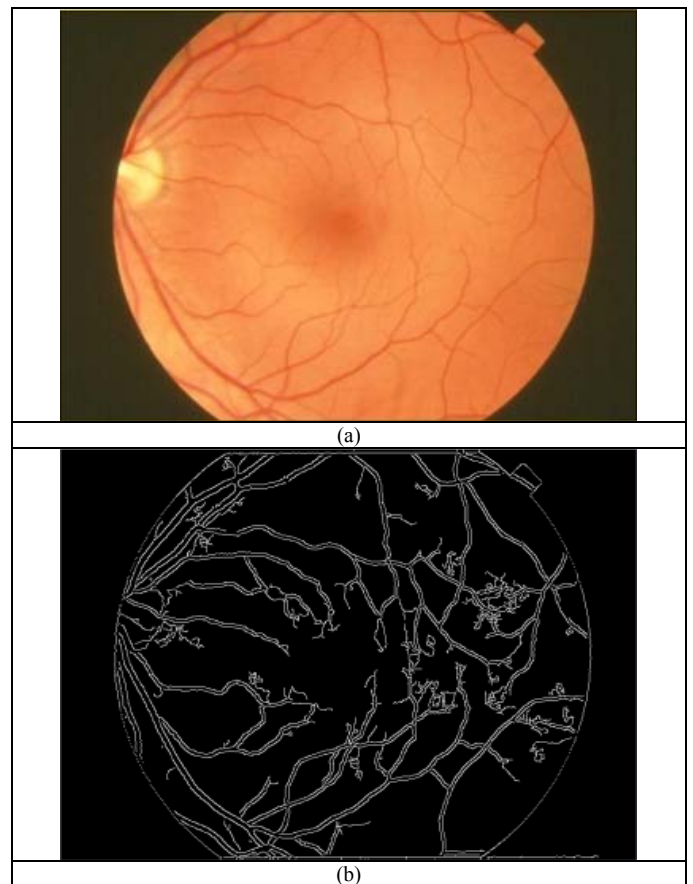

(a)

(b)

(c)

Fig 7. Histogram displayed (a) red channel histogram, (b) blue channel histogram, and (c) green channel histogram

User can now observed the processed image by superimpose it with the original fundus image. The resultant image is depicted in Fig. 8, where all the significant segmented blood vessels are clearly visible outcome from the applied proposed method.
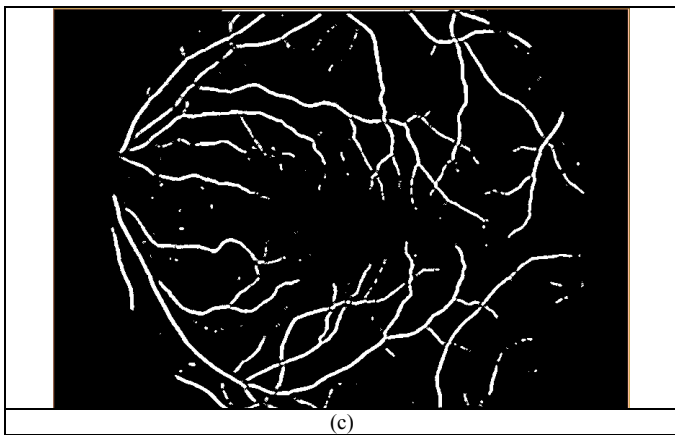

(a)

(b)

(c)

Fig 8. (a) Original image, (b) Blood vessel segmentation using Canny's method, (c) Blood vessel segmentation using Morphology method

V.   Conclusion

This research work has successfully developed a simple and friendly graphical user interfaces (GUI) that provides enhancement of blood vessels segmentation in retinal fundus images. The GUI can aid users to perform blood vessels segmentation easily in the retinopathy images. Canny and morphology edge detection algorithm were proposed in order to extract any blood vessel features. Several image enhancements' method such as median filtering, Gaussian filtering, histogram equalization, contrast and bright was added in this system to improve the displayed image before it was being segmented for significant features information. The system would help ophthalmologist to perform the appropriate morphological procedures by observing the required processed images for early diagnosis.

As far as the objective of this work, the system's application is able only to detect the blood vessels in the retinal images. Detection of other features of DR (exudates, haemorrhages and microaneurysms) and classification whether the patient has eye diseases such as DR or not will be proposed as extension of this initiative work. Besides, the ability for user to dynamically set the various parameter values (iteration, threshold, etc) also set as future development.

REFERENCES

[1]   *Diabetes Facts & Figures*. Available: http://www.diabetes.org.my

[2]   "Screening for diabetic retinopathy," in *Health Technology Assement Report*, M. D. D. Health Technology Assessment Unit, Ed., ed: Ministry of Health Malaysia, 1999.

[3]   L. Xu and S. Luo, "A novel method for blood vessel detection from retinal images," *BioMedical Engineering OnLine,* vol. 9, p. 14, 2010.

[4]   D. Youssef, N. Solouma, A. El-dib, M. Mabrouk, and A. B. Youssef, "New feature-based detection of blood vessels and exudates in color fundus images," in *Image Processing Theory Tools and Applications (IPTA), 2010 2nd International Conference on*, 2010, pp. 294-299.

[5]   H. J. J. D. J. Cornforth., et al. (2005). , *Development of retinal blood vessel segmentation methodology using wavelet transforms for assessment of diabetic retinopathy,* vol. Vol.11.

[6]   V. V. Kumari, N. Suriyaharayananm, and C. T. Saranya, "Feature Extraction for Early Detection of Diabetic Retinopathy," in *Recent Trends in Information, Telecommunication and Computing (ITC), 2010 International Conference on*, 2010, pp. 359-361.

[7]   S. Xiaolin, C. Zhenhua, M. Chuang, J. Yonghang, Z. Y. Duan, L. G. Wang, and S. H. Chang, "Retinal vessel tracking using bilateral filter based on Canny method," in *Audio Language and Image Processing (ICALIP), 2010 International Conference on*, 2010, pp. 1678-1682.

[8]   J. S. Joshi GD, "DrishtiCare: A telescreening platform for diabetic retinopathy powered with fundus image analysis," *Diabetes Science and Technology,* vol. 5, pp. 23-31, 2011.

[9]   A. Perez-Rovira, T. MacGillivray, E. Trucco, K. S. Chin, K. Zutis, C. Lupascu, D. Tegolo, A. Giachetti, P. J. Wilson, A. Doney, and B. Dhillon, "VAMPIRE: Vessel assessment and measurement platform for images of the REtina," in *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 3391-3394.

[10]   J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *Medical Imaging, IEEE Transactions on,* vol. 23, pp. 501-509, 2004.

[11]   M. Shah, *Fundamentals of Computer Vision*, 1997.

[12]   E. Nadernejad, S. Sharifzadeh, and H. Hassanpour, "Edge Detection Techniques: Evaluations and Comparison," *Applied Mathematical Sciences,* vol. 2, pp. 1507-1520, 2008.

Transient Response Characteristics and Deflection Analysis of Microactuator.

Sarah Addyani binti Shamsuddin, Nurul Akmal bin Zakaria, Nina Korlina Binti Madzhi,
Faculty of Electrical Engineering,
Universiti Teknologi MARA, Shah Alam,Selangor

Abstract—This research have been done in order to know the effect of mechanic characteristics to the microactuator such as the microcantilever beam deflection. From there, we also observed the effect of voltage and initial condition to the microcantilever displacement and the transient response characteristics. In this research, we used analytical method and simulation method with Matlab Simulink and Matlab. We observed that the best output response which have fastest response, low overshoot and low steady state error is when d=3.30 and k=3. We also found that the initial condition will only affect the starting point of the response. The increment of voltage will increase the value of displacement and the increment of the voltage did not give any effect to the transient response characteristics.

Keywords—Damping Factors, Spring Constant, Electrostatic Microactuator, Transient Response, and Transient Response Characteristics.

I. INTRODUCTION

Microsensors represent a large section of Microsystems market. Microsensors offer the advantage of replacing conventional sensors in a one to one fashion while saving weight, energy and cost [1]. With the miniaturisation of the system, the possibility to have an actuator is classically done via the electrostatic forces [2].But the modelisation of such a system is a real problem. In static mode, analytic results can be developed in order to underline the non linear characteristics of electrostatic actuation and estimate displacement [3]. As application, in biomedical specifically, microcantilevers are capable of label-free detection of marker proteins related to diseases, even at a low concentration in solution [4-13]. Microcantilevers, operated in a viscous fluid, have also enabled the real-time monitoring of protein-protein interactions [4, 8-11]. Furthermore, microcantilevers are able to recognize the specific protein conformations [14] and/or reversible conformation changes of proteins/polymers based on the principle of analyte adsorption onto the microcantilever surface which cause it to deflect [15]. It shows that microcantilevers have to build with high sensitivity and the miniature system had broadened the range.

The effect of mechanic characteristics (mass, damping factor, and spring values) to the microcantilever deflection had been obtained. Since all the spring constant,k and damping ratio,d had been identified, based on the physical model, the mathematical model and analysis will be extracted. The mathematical analysis or differential equation will be transfer

to find the transient response and all their characteristics such as time delay, t_d, rise time, t_r, settling time, t_s, steady-state error, e_{ss}, and percent overshoot, %OS. The output transient response had been analysed to see the effect of input voltage to the displacement and also to the transient response characteristics.

II. MODEL OF MICROACTUATOR SYSTEM

The microactuator such as microcantilever can be image as a cantilever beam which consists of a metal beam. The cantilever beam is fixed at one end while the other side is free and easy to deflect which the deflection can be measured as the displacement. When the beam deflected, tension occurred at one side and compression occurred at another different side. This condition shows why microcantilever or microactuator can be modeled as mass-spring-damper system. In our system, we consider that there is an input voltage which applied through two parallel plates. The system can be model as in **Fig. 1**.
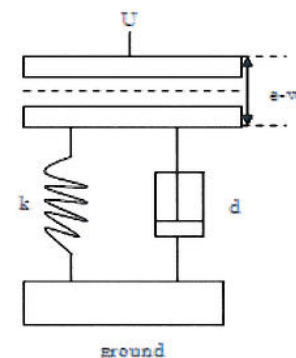

Fig. 1 Model of electrostatic microactuator.

From the model of the electrostatic microactuator, the equation had been extracted by Francais [16] based on the mechanical law and electrostatic actuator which can be expressed as equation (1)

$$m\frac{d^2w}{dt^2} + d\frac{dw}{dt} + kW = \frac{\varepsilon U^2}{2(e-w)^2}S \dots\dots (1)$$

Where;

    d = damping coefficient
    k = spring constant

The other parameters values as stated in **Table 1** had been used based on Francais [16] and Mita,Toshiyoshi [17].

Table 1 List of Parameters

Parameter	Symbol	Value used
Mass	m	1.0kg
Dielectric Constant	ε	8.85x10^{-12} F/m
Gap	e	5μm
Insulator thickness	w	0.5μm
Plate area	S	(200μm)^2

III. MATLAB SIMULINK AND LABVIEW IMPLEMENTATION

A. Parameters of Critically Damped

Using Matlab Simulink, the equation of the mass-spring-damper system is implemented to obtain the critically damped response parameters of damping ratio, d and spring constant, k with the fastest response, low overshoot and lowest steady state error, e_{ss}. **Fig. 2** shows the block diagram that had been used in the Matlab Simulink and LabV. The value of percent overshoot is obtained by using equation (2) and (3), Nise [18].
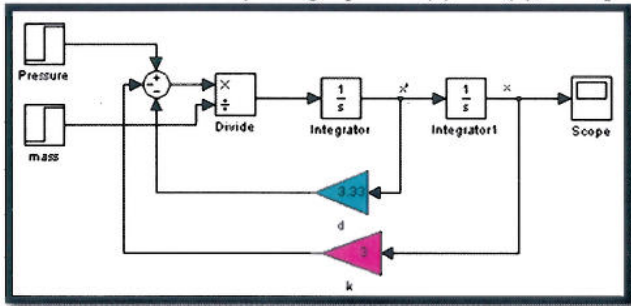

Fig. 2 Block diagram of the mass-spring-damping system.

To get the value of percent overshoot is calculated based on the damping ratio ,equation (2) and equation (3) as stated by Nise [18]:

$$\zeta = \frac{d}{2\sqrt{mk}} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2)$$

Where:

ζ = damping ratio
m = mass
k = spring constant

$$\%OS = e^{-\left(\zeta\pi/\sqrt{1-\zeta^2}\right)} x\ 100\% \ldots \ldots (3)$$

IV. ANALYTICAL METHOD

A. Steady-state Equation

With the values of spring constant, k and damping factor, d that had been obtained previously, which are d=3.33 and k=3.0, are substitute into equation (1). Using initial condition as stated by Neziric, et al. [19], $x(0) = 0.1\ and\ x'(0) = 0$

the steady state equation can be obtained. The initial condition in our system can be implemented as $W_c(0) = 0.1\ and\ W_c'(0) = 0$.

Based on calculation of a second order differential equation, and substitute the initial condition, the steady state equation for the system had been obtained as below:

$$W = W_c(t) + W_P(U)$$
$$= 0.1e^{-1.65t}cos0.5268t + 0.3132e^{-1.65t}sin0.5268t$$
$$+ 2.9136x10^{-9}U^2 - 6.4099x10^{-9}U$$
$$+ 5.10849x10^{-9} \ldots \ldots \ldots \ldots ..(4)$$

V. MATLAB IMPLEMENTATION

A. Displacement versus Voltage

The spring constant, k=3 and damping coefficient, d = 3.33 is substitute into equation (1). The equation is converted into s domain equation and the transfer function is obtained. The transfer function is as below:

$$W(s) = \frac{8.7407x10^{-9}U^2}{s^2 + 3.33s + 3} \ldots \ldots \ldots ..(5)$$

To get the transient response characteristics, we compared the equation (5) with general second-order transfer function as below and find the values of natural frequency and damping ratio. Then we substitute the values into equation (7), (8) and (3) [18].

$$G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \ldots \ldots (6)$$

Where;

ω_n = natural frequency
ζ = damping ratio

$$T_s = \frac{4}{\zeta\omega} \ldots \ldots .. \ldots \ldots \ldots .. (7)$$

$$T_p = \frac{\pi}{\omega\sqrt{1-\zeta}} \ldots \ldots \ldots \ldots (8)$$

With Matlab implementation, the characteristics of the transient response is obtained and the output responses versus time have been plotted. Based on the information from these graphs, the graph for displacement versus voltage can be plotted.

B. Pull-in voltage

From damping factor, d and spring constant, k parameters, we can calculate the pull-in voltage. An electrostatic MEMS switch is comprised of two parallel plate electrodes, one of which is fixed and the other is suspended by a compliant structure that allows it to move. When a voltage is applied across the two plates, opposite signed charges build up on the

two plates which exert an attractive force that pulls the plates together. Because of the nonlinearity of the force resulting from the electrostatic attraction, the plates experience an equilibrium bifurcation that results in a "pull-in" effect. This pull-in effect occurs at a voltage level defined by the design of the structure and is referred to as the pull-in voltage.

Equation (4) shows a lumped parameter model of a standard parallel plate electrostatic MEMS switch. The equation of motion for the movable plate as stated by Gregory N. Nielson [20]

$$ mx'' + bx' + kx = \frac{\varepsilon_0 u^2 s}{2(e-w)^2} \dots \dots (9) $$

Where m is the mass of the plate, k is the spring constant of the suspending structure, b is the damping coefficient of the system, ε_0 is the permittivity of free space, s is the overlap area of the two electrodes, e is the initial gap between the plates, and w is the insulator thickness of the movable plate from its equilibrium position. According to standard operation, quasistatic conditions are assumed which allows the velocity and acceleration terms to be neglected. Solving then for the voltage gives [20]

$$ V_{PI} = \sqrt{\frac{2kx(e-w)^2}{\varepsilon_0 s}} \dots \dots \dots (10) $$

Equation (4) has a local maximum at $w = {^e/_3}$. This point is the position at which the pull-in phenomenon occurs. The voltage associated with pull-in position is [20]

$$ V_{PI} = \sqrt{\frac{8ke^3}{27\varepsilon_0 s}} \dots \dots \dots \dots (11) $$

From the equation (6), we can calculate the pull-in voltage. The voltage is increased until it exceeds the pull-in voltage which causes the movable plate to snap into the pulled-in position. The pulled-in position is defined by the dielectric material or mechanical stops to keeps the two plates from touching and electrically shorting. [20].

VI. RESULT AND DISCUSSION

Using Matlab Simulink and LabView, the output response of the block diagram is obtained, and the response is observed in order to get the best parameters of damping coefficient, d and spring constant, k. The output responses for each fixed spring constant and also for each damping constant have been finalized to get the critically damped responses which have fastest response with low overshoot and steady-state error.

A. Response of Fixed Damping Coefficient, d and Varies Spring Constant.

The first section had been done by the damping coefficient, d fixed and varying the spring constant. The fixed value had been tested with six constant values which are 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 as shown in **Fig 3** (a-f). For each damping

coefficient value, we vary spring, k value from 0.1, 0.2, 0.3, 0.4, 0.5, 1.0, 5.0, 10.0, 50.0, 100.0, 150.0, 200.0, and 250.0.
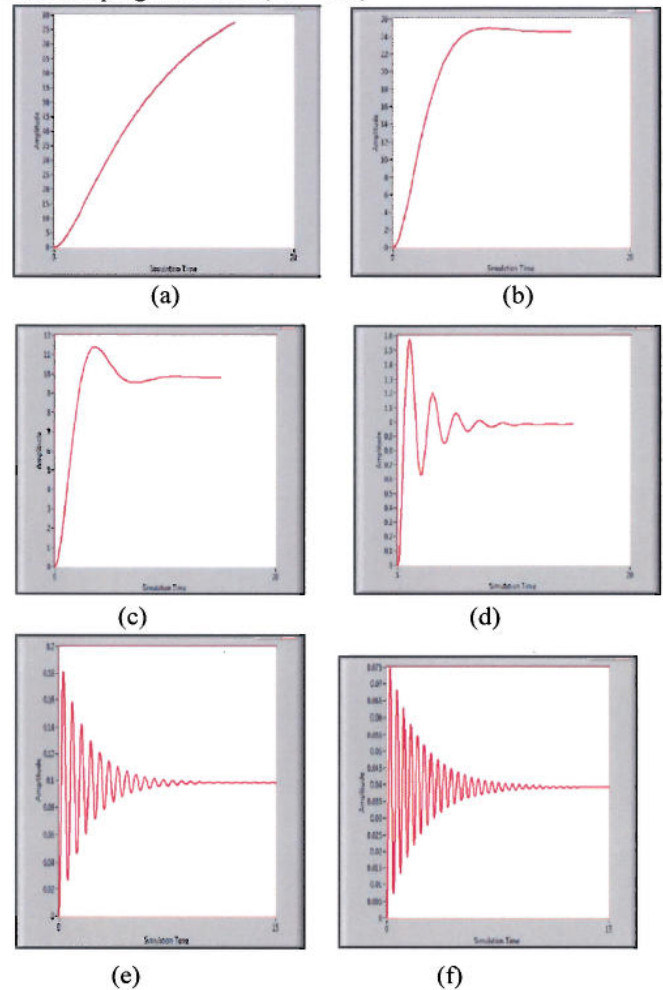
For Damping coefficient, d = 1.0;

(a)   (b)

(c)   (d)

(e)   (f)

Fig. 3 The output transient response of Displacement versus Time when damping coefficient, d =1.0 and spring constant, (a) k = 0.1, (b) k = 0.4, (c) k = 1.0, (d) k = 10, (e) k = 100 and (f) k = 250.

Based on **Fig. 3** above by varying spring constant, we can see that the output responses are in different types. (a) Overdamped, (b) Critically damped, (c) Underdamped and (d,e,f) Undamped.

### B. Response of Fixed Spring Constant, k and Varies Damping Coefficient, d.

The second section had been done with the first section where the spring constant, k is fixed and varying the damping coefficient, d. The fixed value also had been tested with six constant values which are 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0. For each spring constant, k value we vary damping coefficient value, with 0.1, 0.2, 0.3, 0.4, 0.5, 1.0, 5.0, 10.0, 50.0, 100.0, 150.0, 200.0, and 250.0.
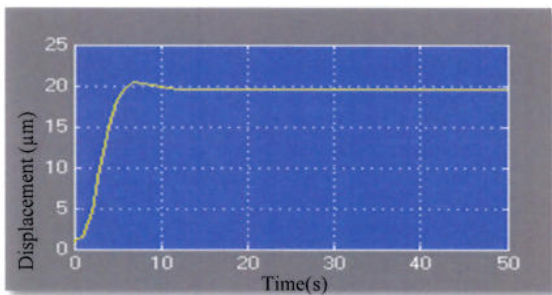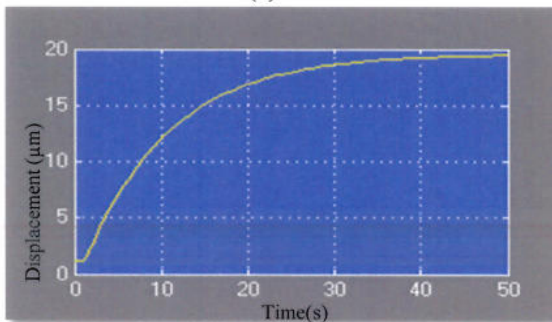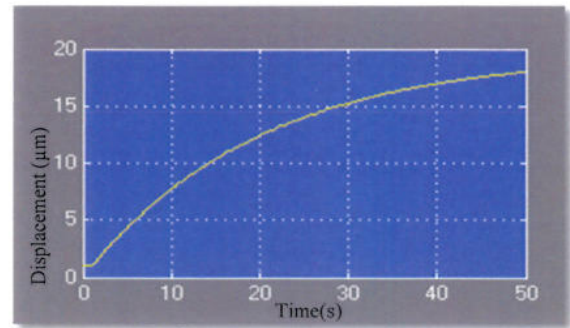
4

For spring constant, k = 0.5:
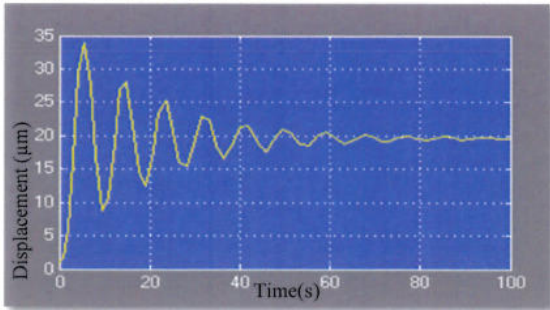
(a)
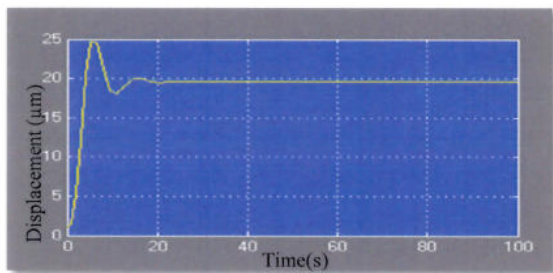
(b)
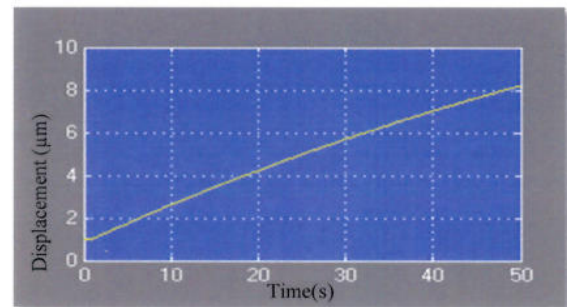

(c)

(d)

(e)

(f)

Fig. 4 Output transient response of Displacement versus Time when spring constant, k=0.5and damping factors, (a) d = 0.1, (b) d = 0.5, (c) d = 1.0, (d) d = 5.0, (e) d = 10.0 and (f) d = 50.0.

From the result in **Fig. 4**, it shows that by varying damping factors, it gives a large effect to the output transient response.The characteristics of the transient response such as settling time, T_s, time delay, T_d, rise time, T_r, steady-state error, e_{ss} and total displacement, W that had been obtained from the output response are summarized as in **Table 2** below. Since the value of the percent overshoot at the responses are too small and cannot be seen, we use equation (2) and (3) to get the values.

Table 2 Value of Transient Response Characteristics after varying spring constant, k.

d	k	$T_s(s)$	$T_d(s)$	$T_r(s)$	$W(\mu m)$	%OS
0.5	0.11	13.5	0.5	7.5	91.5	2.72
1.0	0.33	10.0	0.3	5.0	29.7	0.39
1.5	0.70	7.25	0.2	3.5	14.0	0.17
2.0	1.24	5.5	0.2	2.5	7.90	0.16
2.5	1.79	4.8	0.2	2.0	5.47	0.027
3.0	2.62	4.0	0.2	1.9	3.74	0.043

Table 3 Value of Transient Response Characteristics after varying damping factor, d.

k	d	$T_s(s)$	$T_d(s)$	$T_r(s)$	$e_{ss}(\%)$	$W(\mu m)$	%OS
0.5	1.28	8.7	0.5	4.3	0	19.6	0.1247
1.0	2.00	7.1	0.3	3.5	1.28	9.80	∞
1.5	2.20	5.0	0.2	2.9	0	6.53	0.1636
2.0	2.53	4.2	0.2	2.0	0.89	4.90	0.1863
2.5	2.87	4.0	0.2	1.8	0	3.98	0.1123
3.0	3.30	3.9	0.2	1.7	1.22	3.27	0.0053

From **Table 2** and **Table 3** can be observed that, the fastest output response occur when damping coefficient, d and spring factor, k are set at higher value. For example, k=3 and d=3.30 where the rise time, T_r is 1.7s and settling time, T_s is 3.9s. It also has low steady-state error, e_{ss} and percent overshoot, %OS which are 1.22% and 0.0053%. The displacement is 3.27μm. However, in **Table 3**, it can be observed that at low parameter which is k=0.5 and d=1.28 has largest displacement, W=19.6μm but the response are slowest at the rise time, T_r and settling time, T_s of 4.3s and 8.7s. For biomedical application, it needs to be in fast response condition. Which means, from these six values of k, the best condition is when k=3.0, and d=3.30.

A. *Effect of Voltage to Microactuator Displacement Value*

To find the effect of voltage to the displacement, we have analysed by using two methods. One is by using analytical method and another one is using transfer function simulated with Matlab. From equation (4), several time, t condition have been tested to observed the how time varies affect the system performance. For results with initial condition, the time condition of t=5s have been used as example to observed the effect of voltage to the displacement in two different condition; with initial condition and without initial condition.

1) With Initial Condition (*At t = 5s*)

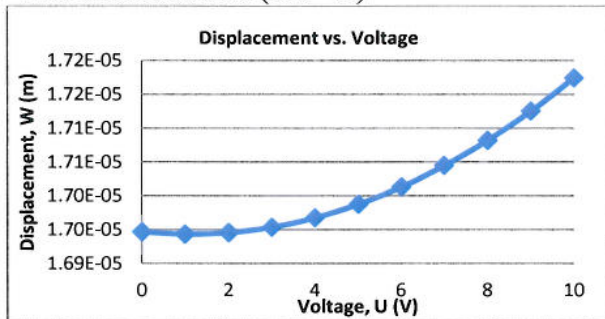Fig. 5 Output response of Displacement versus Voltage at time,t = 5s.
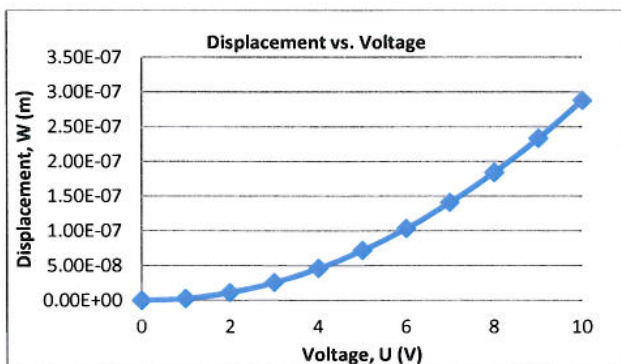
2) Without Initial Condition.

Fig. 6 Output response of Displacement versus Voltage without initial condition.

From **Fig. 5** and **Fig. 6**, we can see that by using the different methods it shows that the displacements are affected by voltage variation. The value of displacement increases when the voltage is increase.

B. *Initial Condition Effects to Microactuator Displacement.*

This result is based on the comparison of the microactuator displacement of two different conditions which are; with initial condition and without initial condition. The data of time, t is at 5s has been taken as an example with initial condition to compare the similarities and the differences of the two conditions. Based on **Fig. 6**, it is observed that for equation (4) with initial condition and equation (5) without initial condition, they do not affect the shape of the displacement versus voltage curve. We observed the initial condition will only affect the value starting point.

C. *Pull-In Voltage*

From equation (11), we calculate the pull-in voltage with analytical method. Then we get the value pull-in voltage.

V_{PI} = 16.56 volt

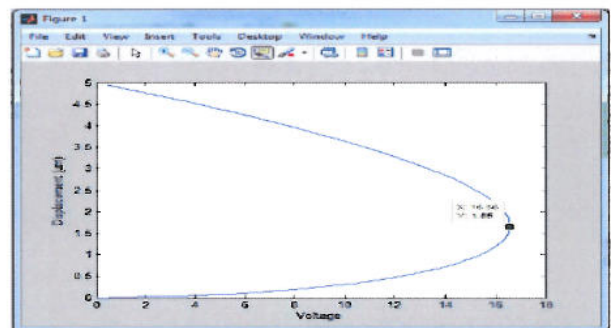Using Matlab application, the pull-in voltage is represented graphically as below:

Fig. 7 Pull-in voltage.

From **Fig. 7**, we can see that the pull-in voltage value is 16.56 volt. The pull-in voltage depends on the spring constant, k value. By increasing the spring constant, k value will increase the pull-in voltage. The supply voltage is increased until it exceeds the pull-in voltage. If the supply voltage value more than pull-in voltage, the two plates will stick.

C. *Voltage Effect to the Transient Response Characteristics.*

By using Matlab, based on equation (5), we observed that the voltage variation only affects the displacement and does not haveany effect to the value of transient response characteristics such as settling time, peak time and percent overshoot. These happened because from the equation (3), (7) and (8) the parameters that effects the transient response characteristics are only damping factor and normal frequency

values. It also can be observed by equation (5) which had been simulated by using Matlab, the voltage,U values varied at the numerators of the transfer function, different with damping factor and neutral frequency which placed at the denominator. This caused the input voltage values were not affected the transient response characteristics.

VII. CONCLUSION

To perform as a microactuator which used mostly in biomedical application, it should have fast response, low percent overshoot, and low steady-state error. In order to get the fast response, we need to have high values of mechanic characteristics such as spring constant, k and high damping coefficient, d. Both which are also the gain of the system, gave extreme effects to the transient response characteristics. When the voltage applied to the system, it will affect the displacement. When we increased the voltage, the displacement will also increase. Using the different method, with initial condition and without initial condition, the difference is only at the starting point. To obtain the output which have high displacement and can start at zero, we can use without initial condition with high voltage input applied. If the supply voltage value is more than pull-in voltage, the two plates will stick. Finally, by using Matlab Implementation, we found that the transient response characteristics are not affected by varying the input voltage but it is only affected by the spring constant, k and damping coefficient, d.

ACKNOWLEDGEMENT

REFERENCES

[1] G. El-Hak. The MEMS Handbook. *CRC Press.*
[2] S. K. R. Zengerle, M. Ritcher, A. Ritcher, "A Bidirectional Micropump," 1995.
[3] P. O. S. Senturia, H.Yie, X. Cai, J. White, "Selfconsistent Simulation and Modeling of Electrostatically Deformed Diaphragms," 1994.
[4] Y. Arntz, Seelig, J.D., lang, H. P., Zhang, J., Hunziker, P., Ramseyer, J. P., Meyer, E., Hegner, M., Gerber, C., "Nanotechnology," 2003.
[5] W. Wu, Datar, R.H., Hansen, K.M, Thundat, T., Cote, R.J., Majumdar, A., "Nat. Biotechnol," pp. 856-860, 2001.
[6] J. H. Lee, Yoon, K.H. Hwang, K.S., Park, J., Ahn, S., Kim, T.S, "Biosens. Bioelectron," 2004.
[7] J. H. , Hwang, K.S., Park, J., Yoon, K.H., Yoon, D.S., & Kim, T.S., "Biosens. Bioelectron," 2005.
[8] T. Braun, Barwich, V., Ghatkesar, M.K., Bredekamp, A.H., Gerber, C., Hegner, M., Lang, H.P., "Phys. Rev. E.," 2005.
[9] R. McKendry, Zhang, J., Arntz, Y., Strunz, T., Hegner, M., Lang, H.P., Baller, M.K., Certa, U., Meyer, E., Guntherodt, H.J., Geber, C. , "Proc. Natl. Acad. Sci.," 2002.
[10] K. S. Hwang, Lee, J.H., Park, J., Yoon, D.S., Park, J.H., Kim, T.S, "Lab Chip," 2004.
[11] N. Backmann, Zahnd, C., Huber, F., Bietsch, A., Pluckthun, A., Lang, H.-P., Guntherodt, H.-J., Hegner, M., & Gerber, C., "Proc. Natl. Acad. Sci. ," 2005.
[12] G. Wu, Ji, H., Hansen, K., Thundat, T., Datar, R., Cote, R., Hagan, M.F., Chakraborty, A.K., & Majumdar, A. , "Proc. Natl. Acad. Sci. ," 2001.
[13] C. A. Savran, Knudsen, S.M., Ellington, A.D., & Manalis, S.R. , "Anal. Chem.," 2004.
[14] R. Mukhopadhyay, Sumbayev, V.V., Lorentzen, M., Kjems, J., Andreasen, P.A., & Besenbacher, F. , "Nano. Lett.," 2005.
[15] W. Shu, Liu, D., Watari, M., Riener, C.K., Strunz, T., Welland, M.E., Balasubramanian, S., & McKendry, R. , "J. Am. Chem. Soc. ," 2005.
[16] D. O. Francais, "Analysis of an Electrostatic Microactuator with help of Matlab/simulink: transient and frequency characteristics.."
[17] M. M. a. H. Toshiyoshi, "An Equivalent-Circuit Model for MEMS Electrostatic Actuator using Open-Source Software Qucs," 2009.
[18] N. S. Nise, *Control Systems Engineering,* 2008.
[19] S. I. Emir Neziric, Vlatko Dolecek, Avdo Voloder, "An Analysis of Damping type Influence to Vibration of Elastic Systems," presented at the 14th International Research/Expert Conference, Mediterranean Cruise, September 2010.
[20] M. J. S. Gregory N. Nielson, Olga B. Spahn, Gregory R. Bogart, Michael R. Watts, Roy H. Olsson III, Paul Resnick, David Luck, Steven Brewer, Chris Tiggers, Grant Grossetete, "High-Speed, Sub-Pull-In Voltage MEMS Switching," 2008.

Multi-Criteria Selection For Power Plant Using GIS

Rosniza Idris and Zulkiflee Abd Latif

Centre of Studies Surveying Science & Geomatics,
Faculty of Architecture, Planning & Surveying
Universiti Teknologi MARA (UiTM),
40450 Shah Alam, Selangor, Malaysia
zulki721@salam.uitm.edu.my

Abstract - **Malaysia has been well enough qualified in order to operate own nuclear power plant for the long term energy needs preparation in electricity demands. The government has taking a step forward in order to reduce the reliance on gas. In this study, the role of AHP and GIS in determining the optimum location of nuclear power plant is explained, criteria for the site selection are outlined and the results for finding the optimum location of nuclear power plant are included. Raub District in Pahang was selected as a study site due to its topography meeting most of the investigated criteria. Seven criteria are considered in this analysis such as low population density, residential area, the access of the river, topography, land use, electricity infrastructure and land ownership. Moreover, the following steps are performed such as represent the criteria in GIS layers, undergo the GIS analysis in order to get the several propose sites and classify the criteria according to the weight value. The site suitability map is effective for assessing and reveals the rates of suitability for the nuclear power plant. The results of the site suitability assessment showed that Site E located at Ulu Dong was the most suitable area for locating the nuclear power plant, with covers an area of 10063.86 acres or 38 % of the total area Furthermore, the map provides the foundation for decision makers in develop the nuclear power plant location.**

Keywords: Analytical Hierarchy Process, Nuclear Power Plants, Multi-Criteria

I.   INTRODUCTION

Global energy consumption is significantly increase as increase in the population. An estimation being made by International Energy Agency, a 53% increase in global energy consumption is foreseen by 2030 [1]. Currently, Malaysia started to explore the benefit of the nuclear power plant for current and future generation. It is important for every Malaysian to comprehend that nuclear is not just an element to create explosive but also create power to generate electricity. As a fast industrializing country, the demand for electricity will continue to rise and keep to the same trend as GDP [2]. Malaysia source of energy is highly dependent on crude oil, natural gas, hydropower, coal and coke (Table 1).

TABLE 1. PRIMARY ENERGY SUPPLY SHARE IN MALAYSIA [3]

Primary energy supply	Amount (ktoe)		Share (%)	
	1990	2008	1990	2008
Crude oil	12,434	24,494	61.1	38.2
Natural gas	5,690	27,800	27.9	43.4
Coal and coke	1,326	9,782	6.5	15.3
Hydropower	915	1,964	4.5	3.1
Total	20,365	64,040	100	100

According to Malaysian Atomic Energy Licensing Board (AELB), Malaysia is well enough qualified in order to operate own nuclear power plant for long term energy needs in electricity demands. It is because of the full rely on coal and gas for getting their electricity power is no longer can adapt with the future demands. Subsequently, the proposal of finding the suitable site for a nuclear power plant has been held in order to overcome these issues. Analytical Hierarchy Process (AHP) is a multi-criteria decision making method proposed by T.L. Saaty [4]. AHP has been widely applied for site selection, evaluation of projects, selection of power plants, demand site management and electric load forecasting [5][6][7][8]. AHP method is a structured approach for dealing with multi-criteria decision problems where the decision process are defined hierarchy [4]. AHP technique can divided into three-step procedure. First, create judgment matrix by pair wisely comparing all the factors at one level of the hierarchy with respect to each factor in the immediate level. Then, compute the eigenvector of judgment matrix corresponding to largest eigenvalue. Finally, calculate the composite priority vector from the local priorities associated with each judgment matrix.

The main objective of this study is to determine the optimum site for nuclear power plant according to criteria of sustainable development including safety and environmental issues. AHP has been proposed as a systematic approach to set priorities and to determine the weight of each criteria. The ranking analysis to select the most suitable site with

regard to the potential criteria in constructing a nuclear power plant is also been investigated.

II. STUDY SITE AND BUILDING AHP MODEL

The study site was Raub, Pahang which covers an area of 2,271 km², situated between Titiwangsa Range and Benom Mountain Range at west of Pahang (3º 48' 00"N, 101º 52' 00" E). Raub district consist of seven(7) sub-district, namely Batu Talam, Sega, Semantan Ulu, Dong, Ulu Dong, Gali and Tras.

AHP has been an effective tool in structuring and modeling muti-objective problems which can assist decision makers to evaluate a problem through a series of pairwise comparison of relative criteria [7]. The decision regarding the selection of optimum area for nuclear power plant in Raub was evaluated according to the safety and environmental issues. For the purpose of this study, it is recommended to follow the criteria according to the Regulatory Guide of General Site Suitability Criteria for Nuclear Power Plant by Office of Nuclear Regulatory Research from U.S Nuclear Regulatory Commission (1998). Besides, the criteria discuss are used only in the initial stage of site selection because it does not provide the detailed information on the relevant factors. Besides, there are many criteria of importance to point in this issue such as criteria from Malaysian Nuclear Agency (2008) and Siting Nuclear Power Plant in Australia (2007). The criteria involved are

  i.   Residential area → the population area that 8km radius away from the proposed site.
  ii.   Topography → flat land or land with a slope of less than 20º.
  iii.   Land ownership → government or private land.
  iv.   Land use → forest area or open land.
  v.   Near to water sources → site near a river or along coastline.
  vi.   Population density → low population density.
  vii.   Electricity infrastructure → appropriate existing electricity infrastructure

III. GIS DATA PREPARATION

Followings are the criteria needed for GIS data preparation and priorities of each of the criteria;

  i.   Residential Area

   The preferred sites are most influence at the area that away from the residential area. The probable area must be located around 8 kilometers radial distance. The population area of 8 km radius away from the proposed site will be given the higher weight value.

  ii.   Topography

   Contour data are needed to acquire a land that has a slope of less than 20º. In order to get display of surface, Triangular Irregular Network (TIN) must be done first using ArcMap 3D Analyst. The areas that have been categorized as a flat land and have a slope less than 20º are to be considered as a suitable site for the purpose of installing the nuclear power plant.

  iii.   Land ownership

   Land ownership that belongs to the government would be more preferred. But, if there are the private owned lands, it will be considered. Reserve land will be excluded. Government land of the proposed sites will be given the higher weight value.

  iv.   Land use

   Land use involved in selection of nuclear power plant are in area of Raub are forest, residential, open land and reserve land. Forest area of the proposed sites will be given the higher weight value.

  v.   Proximity to water sources

   To evaluate the suitability of sites, there should be the availability of water in the quantities needed for nuclear power plant for the purpose of cooling system. Luckily, the area of Raub has several flow of river on it. Site near a river or along coastline of the proposed sites will be given the higher weight value.

  vi.   Population density

   According to the Nuclear Regulatory Research from U.S Nuclear Regulatory

Commission (1998), the nuclear power plant should be far away from the densely populated area. The areas that have a low population density are more preferred. This is for the purpose of reducing the potential of property damage in any type of accident and plan for the emergency planning. Low population density of the proposed sites will be given the higher weight value.

vii. Sub district

The sub district data which is in the line features is use to separate the suitable area for the use of classification process.

viii. Electricity Infrastructure

Electricity system loses energy as electricity is transported through the transmission and distribution of network. To reduce the transmission losses, the nuclear power plant would prefer to be sited near the main centers of demand. Appropriate existing electricity infrastructure of proposed sites given the higher weight value
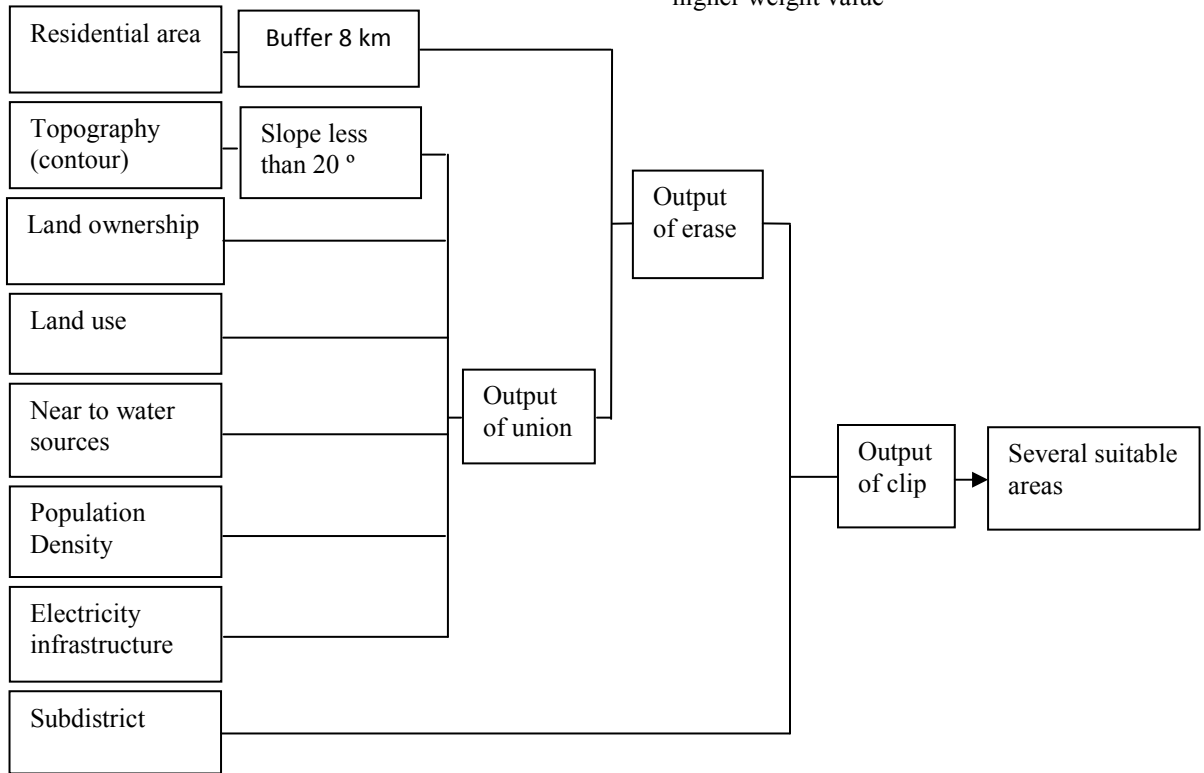
Figure 1. Methodology of GIS Analysis

A.   Site Verification

In order to obtain the optimum site for the nuclear power plant in Raub, Pahang, the intergrating of Analytical Hierarchy Process (AHP) with the Geographical Information System (GIS) has been done. Before proceed to the AHP approach, all the results obtained by using the Geographic Information System GIS) will be reviewed by visiting the proposed site. The aim of this process is to ensure that the results obtained are valid and can be used without any obstacles, cost and time.

B.   Classification according to weight value

Determination of the criteria's weight was depending on the importance of each criteria in comparison with the others in the same factor group. The weight value obtained will be use in calculating the final analysis in order to determine the optimum site for nuclear power plant. The weights value may be applicable if the consistency ratio is less than 0.1. If the consistency ratio is exceeding 0.1, the Analytical Hierarchy Approach (AHP) process needs to be done again. The purposes of this application are to facilitate users in computation and obtain results more quickly.

TABLE 2. RESULTS FROM SITE VERIFICATION

CRITERIA	SITE A	SITE B	SITE C	SITE D	SITE E
Low population density	Low	Low	Low	Moderate	Low
Residential Area (radius 8 km)	Yes	Yes	Yes	Yes	Yes
Near to water sources	Yes	Yes	Yes	Yes	Yes
Topography (Slope less than 20º)	Flat land	Flat and hilly area	Flat and hilly area	Flat area	Flat area
Land use of forest or open land area	Forest (watershed area)	Forest (watershed area)	Forest (watershed area)	Forest and open land (watershed area)	Forest (watershed area)
Appropriate existing electricity infrastructure	No	No	No	No	No
Land ownership	Government	Government	Government	Private	Government

IV. RESULTS AND DISCUSSION

There are five(5) suitable sites were determined using AHP. The sites are Site A, Site B, Site C, Site D, Site E (Fig. 3). The results of the site suitability assessment showed that Site E located at Ulu Dong was the most suitable area for locating the nuclear power plant, with covers an area of 10063.86 acres or 38 % of the total area. Land use of this area is mainly forest that has very low population density and situated away from the residential area. There is high water sources contribute by the Sungai Dong as a primary river for the purpose of cooling process for the plant. Besides, it is an advantage for the Site E because this land is owned by the government. Thus, it is possible to proceed with the nuclear power plant due to the government owned land.

The less suitable area for installing the nuclear power plant is located at Site D (Gali 2 sub-district). This site covers an area of 9270.54 acres. This region has low water sources for the cooling purpose and situated on the open land that currently developed. Therefore, population density of this area is a bit higher than others and was given low weight value. Besides, Site D is owned by private owner which needs a lot of time and procedure in land acquisition. This make a reason why the government owned land is more preferred.
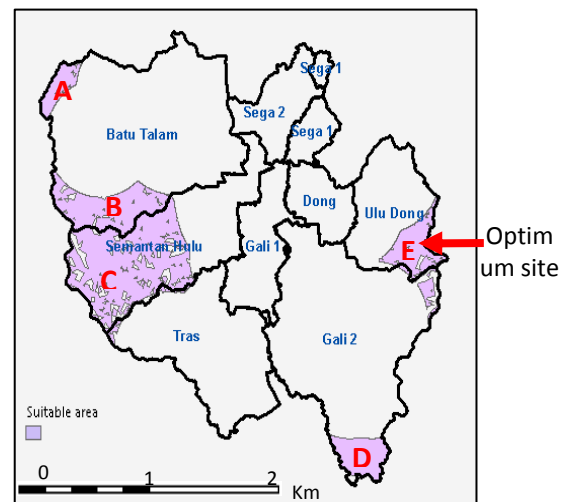

Figure. 2. The flow of weight calculation.

TABLE 3. THE CLASSIFICATION OF PROPOSED SITE ACCORDING TO WEIGHT VALUE

Proposed Site	Weight Value	Percentage (%)	Rank
Site A	0.197	20	2
Site B	0.166	17	3
Site C	0.135	14	4
Site D	0.124	12	5
Site E	0.377	38	1

Figure. 3. Optimum site for nuclear power plants

V. CONCLUDING REMARKS

A GIS-based assessment of the site suitability for nuclear power plant can help decision makers to identify the optimum areas while promoting the security and preservation of environmental areas. The suitability map obtained from this study can provide the decision makers with useful information to determine the trends and spatial distribution of suitability classes for the whole area of Raub, Pahang.

The AHP approach was successfully approve as a systematic tool for site suitability assessment which integrates with GIS.

VI. ACKNOWLEDGEMENT

REFERENCES

[1] Akash, B.A, Mamlook, R. & Mohsen, M.S., *Multi-criteria selection of electric power plants using analytical hierachy process*. Electric Power Systems Research.Vol 52: pp. 29-35, 1999.

[2] Ong, H.C., Mahlia, T.M.I. & Masjuki, H.H., *A review on energy scenario and sustainable energy in Malaysia*. Renewable and Sustainable Energy Reviews. Vol. 15: pp. 639-647, 2011.

[3] NEB, National Energy Balance 2008. Selangor, Malaysia: Malaysia Energy Centre; 2009.

[4] Saaty, T.. *The Analytic Hierarchy Process*, McGraw-Hill, New York, 1980.

[5] Wang, Y., Gu J. & Hu, B. *Practical comprehensive model for medium and long-term load forecasting based on AHP*. IEEE Trans. East China Electric Power, Vol.33: pp. 28–31, 2005.

[6] Sanjay, V. & Ramachandran, M. *Multicriteria evaluation of demand side management (DSM) implementation strategies in the Indian power sector*. IEEE Trans. Energy. Vol.31: pp. 2210–2225, 2003.

[7] Bilal, A. A., Rustom, M. & Mousa, S. M., *Multicriteria selection of electric power plants using analytical hierarchy process*. IEEE Trans. Electric Power Systems Research, Vol.52: pp. 29–35, 1999.

[8] Jang, W. & Xu. Y. *Reasonable structure of power network and optimum configuration of power source*. IEEE Trans. Power System Technology, Vol.22: pp. 42–45, 1998.

Processing Capability Improvement through Parameter Optimization of A Sludge Drying Plant

M. Z. Mohamed & M. F. M. Ghazali

Engineering Department

PETRONAS Penapisan Melaka Sdn Bhd

Melaka, Malaysia

mohdzurix@petronas.com.my

S. M. Idrus & N. A. Wahab

Faculty of Electrical Engineering

Universiti Teknologi Malaysia, UTM

Johor, Malaysia

Abstract— **The Sludge Drying Plant (SDP) is the final processing facility of Effluent Treatment System (ETS) that produces bio-sludge cake before it is sent out for final disposal. Due to the process disturbances the amount dry solid is reduced tremendously. The principal objective of this study is to perform a parameter optimization for the SDP through identification on the current performance through actual process data analysis and identification of the subsystems for optimization. On disposal, Kualiti Alam will charge the dried cake sludge based on the weight, not on the content; therefore it is only sensible to export the dried cake at maximum dryness. Optimized controls for the SDP ensures a maximum weight percent solution (wt%) is squeezed out through complex computing from available process parameters .**

Keywords-Sludge Drying Plant; process design; parameter optimization; effluent treatment system

I. INTRODUCTION

The end product of a typical oil and gas processing plant is normally sent to the Effluent Treatment System (ETS), or some similar industries may refer it as Water Treatment Facility, to be treated before they are being disposed in already made available methods, be it dispersing the treated liquid into the sea, unused sludge as landfill or through a government controlled body for disposal [1]. In most cases of an oil refinery, the output of the ETS, also referred as bio-sludge is sent to the Sludge Drying Plant (SDP) in which it will perform the final processing batch of a wet sludge to produce a cake like product, is sent for disposal through the government controlled body with a cost on the industry. The main purpose of the SDP is to treat the bio-sludge generated from the existing ETS or Wastewater Treatment Plant to a suitable dryness for ease of disposal. However, the problem faced by some specific Oil & Gas Processing Plant is the inability to sustain the low processing cost for dry solid disposal and in turn reflect the un-optimized processing equipment and high financial implication. The dryness of the processed bio-sludge is measured in weight percent solution (wt%), in which the efficiency is considered high when the wt% is high. In an ideal case with all the series of processes are at optimum level of operating, the wt% may reach as high as 90 wt%. Lacking of control due to process disturbances will reduced the amount of dry solid tremendously. A number of industries have already applied parameter optimization and have significantly showed

a high impact on finance [3]. For parameter optimization, data collection is based on normal actual operating parameters such as flow rate, percentage of dry solid and solid weight, each taken on random intervals to increase the chances of identifying the different bio-sludge properties per process batch. Picture 1 is a typical setup of a Sludge Drying Plant.

Picture 1. A typical installation of a Sludge Drying Plant on an Oil and Gas Processing Plant.

II. PARAMETER OPTIMIZATION

System optimization requires understanding on the type of process and medium and the sensors that the system is managing for each of the main components and also the equipment capability to deliver at optimum level. For a successful parameter optimization on a complex process as the SDP, as much data is required to assess the current performance. Data collection for parameter optimization is summarized in Table I. Optimization on the sensors may include defining the best measurement range. Each of the sensors and equipment is analyzed and compared with the targeted operating parameter to reach 90wt% optimization.

TABLE I.   DATA COLLECTION FOR PARAMETER OPTIMIZATION

Item	Optimization description
Flow rate and pressure	Data analysis to determine the process bahaviour.
Chemical consumption	Data collection on the volume usage per batch process
Pump operation and lubricant oil level	Capture pump's efficiency through abnormal noise and vibration during operation.
Mixer operation	Observation on abnormal noise and vibration during operation.
Polymer feeder operation	Leak test at injection facility.
Filter press system	Inspection for abnormal noise and vibration. Leak inspection during feeding mode.
Diaphragm pump operation	Pressure test, leak test
Programmable logic controller	Data collection on PLC cycle time

Figure 1.   A representation of a simple 3 layer artificial neural network with one output.

III.   DATA ANALYSIS

A pilot study on data collection to estimate the current deviation from target is shown in Table II. These data are taken on random intervals to ensure variation on the sample properties during normal operation. Efficiency deviation of the sensors and the equipment from the targeted optimized level are treated as disturbance. The observed dry solid, solid weight and flow rate are benchmarked against the desired output. Based on the pilot study results, the calculated deviation from the target ranges from 45.7% to 52.7%. On this normal process the percentage of dry solid only reaches up to 35% dry solid compared to a target of 90% dry solid.

TABLE II.   DATA COLLETION ON PILOT STUDY

Dry Solid (%)	Solid weight (Ton)	Flow rate (m3/hr)	Deviation from target (%)
25	0.47	2.4	52.3
27	0.45	2.7	48.7
30	0.44	2.1	52.3
29	0.46	2.2	50.7
35	0.45	2.6	45.7

a. Only showing a few number of data for paper reference.

IV.   RESULT AND DISCUSSION

Pilot study showed there is correlation between amount of dry solid and flow rate. Analysis also showed that the current efficiency of the overall system is only 50% efficient. Results from the pilot study showed that this process is indeed nonlinear. To close the gap of 40% optimization, a model based control strategy can be implemented as this nonlinear process can be optimized through Artificial Neural Network (ANN) [2]. To perform such optimization on SDP's complex process, the processes executed has been distinguished and treated as separate subsystem during performance gap identification but treated as a whole during ANN implementation. The identified subsystems on the SDP are the main programmable logic controller (PLC) that controls the overall process sequence, the chemical dosing system as the coagulate agent, and the non-fuel dryer system. On parameter optimization for all these subsystems, the SDP will produce up to 90wt% solid content. Due to process nonlinearity and inconsistent repeatability, the process optimization can be visualized through ANN shown in Figure 1 [4].

The parameters on the input layer of the neural network must be a crisp representative of the processes in order to achieve the desired output. To implement ANN, the optimization on the identified three major subsystems is treated separately. To prevent too many hidden layer on the ANN, the development of ANN is implemented separately on each subsystems. The outputs of these subsystems will then be an input to the overall ANN, as a complete system. Figure 2 is visualization on the ANN to be implemented.
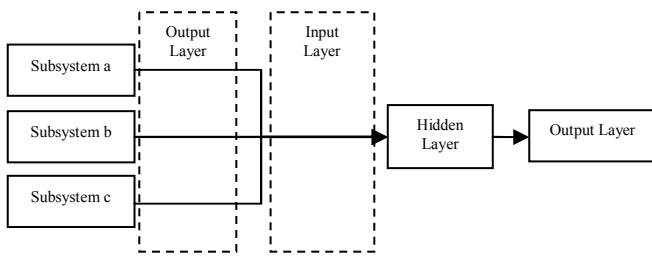
Figure 2.   A brief visualization of the ANN implementation on there subsystems' output as an input on another hidden layer.

In producing the algorithm, mathematically, ANN function (i.e. $f(x)$) is defined as a composition of other functions (i.e. $g_i(x)$). In which this can further be defined as a composition of other different functions. Putting this together, this can be represented as a network structure, with arrows depicting the dependencies between variables, as shown in Figure 2.

In ANN, a widely used type of composition is the nonlinear weighted sum, where:

$$f(x) = K\left(\sum_i w_i g_i(x)\right),$$

where K, also commonly referred to as the activation function, is some predefined function.

Training a neural network model simply means selecting one model from the set of allowed models that minimizes the cost criterion, similarly as stated in a Bayesian framework, whereby, determining a distribution over the set of allowed models. A few numbers of algorithms are available for training neural network models, in which they can be viewed as a straightforward application for optimization and as statistical estimation.

For the ANN on the SDP to be implemented with the ability to learn, a different set of task must be implemented. For example, given a specific task for an ANN to solve, and a class of functions, (i.e. F), learning means using a set of observations to find $f^* \in F$ which solves the task in the most logical and optimal sense.

The process is further elaborated on the defining a cost function:

$$C : F \rightarrow \mathbb{R},$$

in such that, for the most logical and optimal solution, that is defined by f^*, the logic must be within:

$$C(f^*) \leq C(f)\ \forall f \in F$$

In a logical sense, no solution has a cost less than the cost of the optimal solution.

The cost function (i.e. C) is an important concept in ANN learning. The measurement of how far away a particular solution from the desired solution is determined through the cost function. This learning algorithms search on the solution space to find the most suitable function that has the smallest possible cost.

For Sludge Drying Plant application, where the solution is highly dependent on some data attained from sensors, the cost must necessarily be a function of the observed value (i.e. Process Value (PV)), otherwise the ANN will not be modeled in relation to the data. This can be defined as a statistical calculation to which approximations can be made. To further elaborate this application on the SDP, consider the problem of a subsystem optimization via the model f, which minimizes

$$C = E\left[(f(x) - y)^2\right],$$

for data pairs (i.e. (x, y)) extracted from a distribution (i.e. $\mathcal{D}$).

In a practical situation, we would only have N samples from distribution $\mathcal{D}$ and thus, we would only minimize

$$\hat{C} = \frac{1}{N}\sum_{i=1}^{N}(f(x_i) - y_i)^2.$$

This would then cause the cost to be minimized over a sample of the data rather than the entire data set.

As for real time learning, where $N \rightarrow \infty$ some other form of machine learning must be used, where the cost is minimized as each new sample is seen.

V. CONCLUSION

Pilot study showed a nonlinear process and utilizing conventional controls is not enough for optimization. With current performance, the SDP is only capable of performing up to an average of 50% efficiency. Converting the SDP's current performance into money value, the industry is losing up to RM1.3 million annually on normal operation due to un-optimized system. To close the 40% gap from the targeted performance, this study has identified three major subsystems namely the main PLC, the chemical dosing system and the non-fuel dryer system, that each requires fine tuning. For future study to ensure reliability, repeatability and efficiency of the SDP, an ANN must be implemented. This would increase the output cake up to 90wt% solid content. To further increase the SDP capability and efficiency, more analog sensors, equipped with Line Monitoring system must be installed, with diagnostics features. There are many different methods today to implement ANN with learning capability and merging the hardware redundancy controller would increase the uptime of the system and comparative block would ensure the calculation is done two ways and compared for deviation.

REFERENCES

[1] A.Bhardwaj, S. Hartland, "Study of demulsification of water-in-crude oil emulsion," J. Dispersion Sci. Technol. 14, 1993.

[2] D. C. Psichogios, and L. H. Ungar, "Direct and indirect model-based control using artificial neural network," Ind. Eng. Chem. Process Des. Dev., Vol. 30, pp. 2564-2573, 1991.

[3] P. A. Miderman, and T. J. McAvoy, "Neural net modeling and control of municipal waste water process," Proceedings of American Control Conference, pp. 1480-1484, 1993.

[4] S. A. Rounds, "Development of a neural network model for dissolved oxygen in the Tualatin River Oregon," Proceedings of the Second Federal Interagency Hydrologic Modeling Conference, Las Vegas, 2002.

A Study on THD Reduction by Active Power Filter Applied Using Closed-loop Current Controlled AC-AC SPMC Topology

Mazratul Firdaus Mohd Zin, Noraliza Hamzah, Mohammad Nawawi Seroji
Faculty of Electrical Engineering, Universiti Teknologi MARA,
40450 Shah Alam, Selangor, Malaysia

Abstract— **An active power filter with single phase matrix converter topology is presented to reduce the THD of the input current, output current and output voltage of the system which is the main analysis in this paper. AC-AC converter topology switching strategy is used in the SPMC which is controlled by the current control closed-loop circuit or also known as active current wave-shaping technique. In order the AC-AC SPMC to operate and synthesize as an active power filter, the input supply current is being compensated by current control closed-loop circuit to produce the SPWM signal to be injected to the switches used which are the IGBTs. The work has been done using Matlab/Simulink simulation and the results are being compared to the same simulation done in open-loop method without using the active current wave-shaping technique.**

Index Terms— **Insulated Gate Bipolar Transistor (IGBT), Single Phase Matrix Converter (SPMC),Sinusoidal Pulse Width Modulation (SPWM), Active Power Filter.**

I. INTRODUCTION

The matrix converter has received considerable attention in recent years because of some appealing operational characters. Namely, sinusoidal input and output waveforms, bidirectional power flow, controllable input power factor, absence of energy storage reactive elements, and compact size[1].

Direct ac–ac converters have a number of advantages compared to dc link converters. In such converters, power is converted from ac fixed-frequency fixed-voltage to ac variable-frequency variable-voltage without any intermediate dc link[2]. These converters are becoming popular due to the availability of better switching devices. The matrix converter introduced by Venturini and Alesina in 1979 is the most general converter-type in the family of ac–ac direct converters. Its topology was first proposed by Gyugyi in 1976. The converter offers an all-silicon solution for ac–ac power conversion[3]. The single phase version named single phase matrix converter was first introduced by Zuckerberger[4]. Other SPMC topology was studied by S.H. Hoseini[5] and Abdullah Khoei[6].

However, the application of matrix converter may contribute high harmonic distortion that could affect the utility system and the electricity distribution. It does not only disturb the utility system, many other problems might occur due to the harmonic distortion. Thus, filtering circuit is required in order to attenuate the total harmonic distortion produced. Filtering is a method to reduce harmonics in an industrial plant when the harmonic distortion has been gradually increased or as a total solution in a new plant. There are two basic ways to achieve harmonics cancellation or reactive power compensation and improve the power quality of a power system. They are passive filtering and active filtering [5].

This paper represents the implementation of current control closed-loop circuit to perform as harmonic filtering method in AC-AC single phase matrix converter by using Matlab/ Simulink in order to improve the total harmonic distortion of the input current, the output voltage and output current. A comparison procedure is carried out between a simulation of SPMC with open loop operation and SPMC with current control closed-loop circuit.

II. SINGLE PHASE MATRIX CONVERTER

The Single-Phase Matrix Converter consists of a matrix of input and output lines with four bidirectional switches connecting the single-phase input to the single-phase output at the intersections[4]. It comprises of four ideal switches S1, S2, S3 and S4 capable of switching between states without any delays, conducting current in both directions, blocking forward and reverse voltages (symmetrical devices)[2]. Figure **1** shows a single-phase matrix converter.
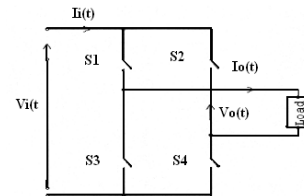

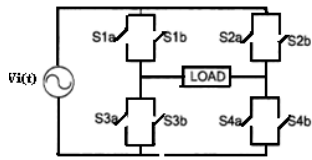Figure 1: Representation of a Single-Phase Matrix Converter

Figure 2: SPMC circuit configuration

Nowadays, there is no bidirectional switch semiconductor that has the capability of blocking voltage and conducting current in both directions available in the market that is required for the application of the matrix converter[7]. However, as in Figure 3, in order to fulfill the requirement, two IGBTs are used back to back that functions as bidirectional switch for the matrix converter. It is common emitter anti-parallel IGBTs with diode pair. Diodes are used to provide reverse blocking capability of the switch itself[3]. The IGBT is popularly used since its high speed switching capabilities and high current carrying capacity desirable amongst researchers for high power application. In SPMC, there would be four bidirectional switches required.
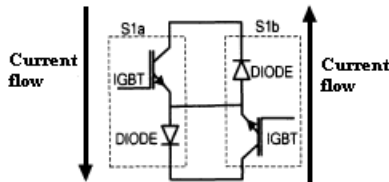
Figure 3: IGBT module

In this paper, the analysis on the SPMC would be for ac-ac converter. The frequency been used is 50Hz.
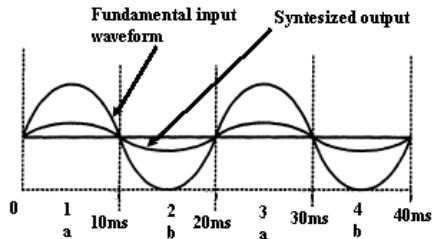
Figure 4: Sinusoidal input and synthesized output

The input and output voltage is given by (1) and (2) respectively while loads represented in (3).

$$v_i(t) = \sqrt{2}V_i \sin\omega_i(t) \qquad (1)$$

$$v_o(t) = \sqrt{2}V_o \sin\omega_o(t) \qquad (2)$$

$$v_o(t) = Ri_o(t) + L[i_o(t)/dt] \qquad (3)$$

Subscript i denote input, whilst o denotes output.

III. SWITCHING STRATEGY

The switching strategy used in these studies is based on implementation of SPWM as proposed in[3]. Power switches comprising IGBTs in the SPMC circuit are controlled where the switching angles, of the 4 bi-directional switches uses the form S (i = 1,2,3,4 and j = a,b), where 'a' and 'b' represent drivers one and two respectively. The switching pattern is based on TABLE 1[8];

• At any time 't', only two switches Sij (i = 1,4 and j =a) will be ON state and conduct the current flow during the positive cycle of input source (state 1), with S2b turn ON for commutation purpose.

• At any time 't' only two switches Sij (i = 1,4 and j = b) will be in 'ON' state and conduct the current flow during negative cycle of input source (state 2), with S2a ON for commutation purpose.

TABLE 1: SEQUENCE OF SWITCHING CONTROL

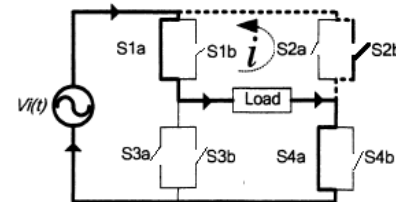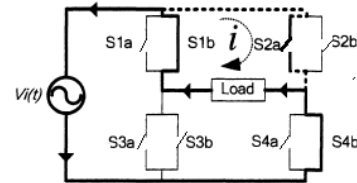| Switches	Operation	
	Positive cycle (State 1)	Negative cycle (State 2)
S1a	ON	OFF
S1b	OFF	ON
S2a	OFF	ON
S2b	ON	OFF
S3a	OFF	OFF
S3b	OFF	OFF
S4a	OFF	OFF
S4b	SPWM	SPWM

Figure 5: State 1 (positive cycle)

Figure 6: State 2 (negative cycle)

For the SPMC switching algorithm in this paper, the input

frequency is 50Hz and the output frequency will also be 50Hz. So that, state 1(for positive cycle) and state 2 (for negative cycle) are used for the switching strategy in this paper.

During positive cycle of the input source, S1a and S2b is maintained in ON state. Here, S4a is the controlling switch to synthesize the SPWM pattern. S1a acts to complete the loop for the SPWM return and acts in conjunction with S2b to provide freewheeling operation whenever S4a is turned OFF. Due to the turn-off IGBT characteristic, where the tailing-off of the collector current will create a short circuit with the next switch turned on, this is impossible for the SPMC switching sequence to be instantaneous and simultaneous. As a solution, the nature of operation of the commutation period has to be extended over the dead time period to allow for the energy to dissipate and hence the current reversals due the inductive loads are eliminated[2].

During negative cycle of the input source, S1b and S2a are maintained at continuously ON state while S1b to complete the loop for the SPWM return and act in conjunction with S2a to provide freewheel operation whenever S4b in turned OFF. Here, S4b is the controlling switch to synthesize the SPWM pattern.

IV. HARMONIC ATTENUATION

Conventionally, passive filter is used widely in the industry application in solving the harmonics pollution problem due to its low cost and high efficiency[9]. However, there are also disadvantages of using a passive filter. The resonance may occur between power impedances and passive filter. Moreover, the effect of harmonics suppression is highly influenced by impedance and its own parameters[10]. Then, the solution of the harmonics suppression has been discovered by developing an active filter[10]. Active filter is another effective way of providing harmonic attenuation and reactive power compensation. High power active components are becoming cheaper and offering improved performance as the power electronics technology has been developing and expanding power electronics industry. These filters can be applied as a controller for harmonic current or voltage elimination, terminal voltage regulation, neutral current compensation, reactive power compensation, voltage balance improvement and voltage flicker suppression. On the other hand, it still has the drawbacks. As in series connected active filter, it only compensate voltage harmonics and also for the short circuit at the load end due the short circuit current pass through the series transformer winding thus may overload the transformer[11].

As in previous studies,[9] shunt hybrid active power filter was used to attenuate different order harmonic components applied in three phase system that is connected to non-linear loads. Same goes with [10], three phase system with shunt hybrid active filter was studied to prove that it could improve the performance of the passive filter. Differently with [12], the analysis of classification and characteristics of active filter has been done. The review has contributes loads of aid in analyzing

and evaluating the various technique in a subjective fashion of the active filters. A study on single switch parallel active filter in a single phase hybrid active filter has been done in [13] that focusing on the application of shunt active power filter in conjunction with a passive filter.

In this paper, passive filter and series active power filter are used in the simulation in order to compare the results obtained from each type of filter application. The active power filter operation that is applied to the project is based on the studies in [14-16]. However, the application of those previous papers was done to a rectifier single phase matrix converter topology in order to compensate the supply current. Meanwhile, in this paper AC-AC single phase matrix converter topology is utilized for the current control closed-loop circuit operation to reduce the ripple current on the other hand improving the THD and the output voltage of the system.

V. SIMULATION AND PROPOSED CURRENT CONTROL CLOSED-LOOP CIRCUIT

As in previous study, a controlled rectifier using SPMC topology incorporating active power filter was done to ensure that the supply current is continuous, sinusoidal and in phase with the voltage[14-16]. In classical rectifier with DC capacitor filter, a discontinuous supply current is drawn which contained high harmonic distortion level affecting the quality of the power supply system.

However, by applying direct AC-AC converter using SPMC topology, different switching strategy is used as mentioned above in chapter 3. The supply current produced contains ripple current that contributes to high total harmonic distortion level. Other than that, the output voltage of the AC-AC SPMC is a bit lower than the supply voltage 150V (peak).

In order to solve the problems occurred; a supply current control closed-loop circuit is implemented to the AC-AC SPMC to produce the sinusoidal pulse width modulation (SPWM) to control the operation of the SPMC switches. In the proposed technique, low pass RL filter (R=10ohm, L= 0.1mH) is used in conjunction with the current control closed-loop circuit in order to mitigate the ripple of the supply current. High frequency harmonic content is being rejected and passes through the lower frequency. As for the current control closed loop circuit, the supply current is being tapped from input supply before the low pass filter using a current sensor and being compared to the reference signal. The compared signal then is corrected using peripheral integral controller. The corrected signal is compared with the triangular repeating sequence to produce the SPWM. The SPWM signal is then injected to the IGBTs' gate to control the switching sequence. A switching frequency of 6 kHz is proposed in this paper in order to generate instantaneous switching that is required for the SPMC.

VI. RESULTS AND DISCUSSION

The single-phase matrix converter has been simulated using SimPower system block in Matlab/Simulink software program. The operation of ac-ac converter is applied according to the switching strategy as mentioned above. The SPMC is modeled with 50Hz frequency AC single phase supplies at 150V (rms) and was loaded by passive R-L load. At the first stage, the simulation has been done using open loop method with a constant of 0.7 modulation index without using the PI controller to produce the SPWM signal to be injected into the switches involved. With the passive filter attached in between the input supply and the switches, the input current ripples is being reduced with 13.35% THD. The input current amplitude is 18.15A (peak) or 12.83A (rms). The output current measured is 2.847A (peak) or 2.013A (rms) with 3.21% THD. The output voltage recorded as 142.4V (peak) or 100.7V (rms). All the waveforms obtained also have been captured and recorded.

Then, the next similar simulation using the same SPMC topology with a passive RL filter insertion in conjunction with current control closed-loop circuit is done as shown in figure 8. PI controller is used to do the correction for the error compared between the input current tapped from the input supply with the reference current signal. With that, the input current is corrected by reducing the ripples. The input current measured is 24.41A (peak), 17.26A (rms) with 3.86% THD while the output current measured is 5.088A (peak), 3.598A (rms) with 3.24% THD and the output voltage is 179.9V (rms) or 254.4 V (peak).

From the simulation result recorded, it is shown that the output voltage has been stepped up from 142.4V (peak) to 254.4V (peak) which contribute to 78.65% of percentage increase. Meanwhile, the input current is 34.49% increased and the output current has 78.71% of percentage increase. On the other hand, the input current THD has been reduced from 13.35% to 3.86% makes it 9.49% of reduction.
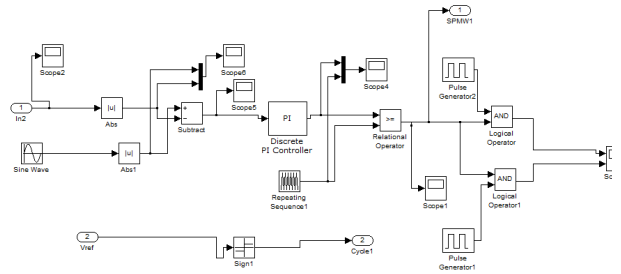

Figure 7: Single Phase Matrix Converter with Current Control Closed-loop circuit

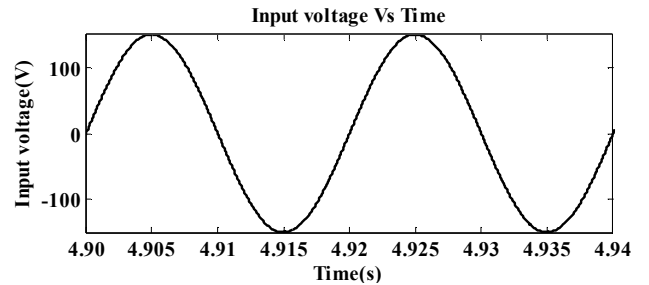Figure 8: Current Control Closed-loop Circuit

Figure 9: Supply voltage

a)

b)
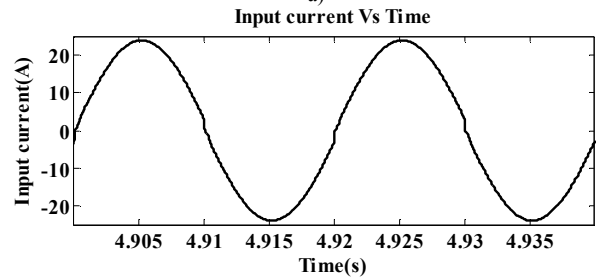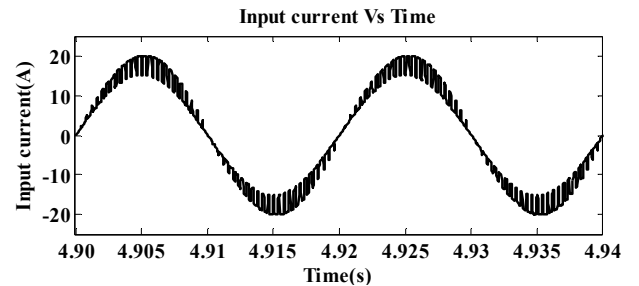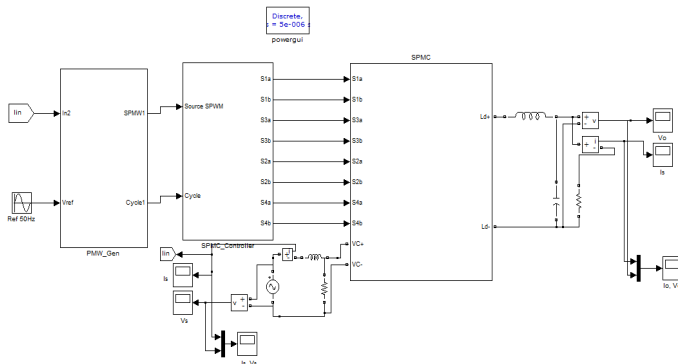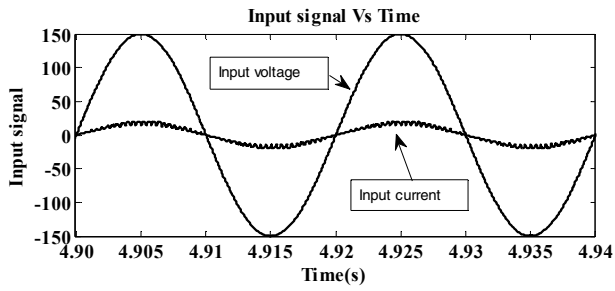
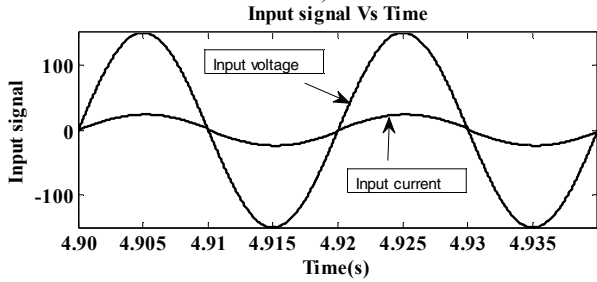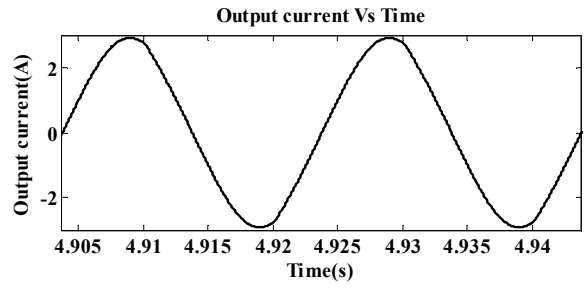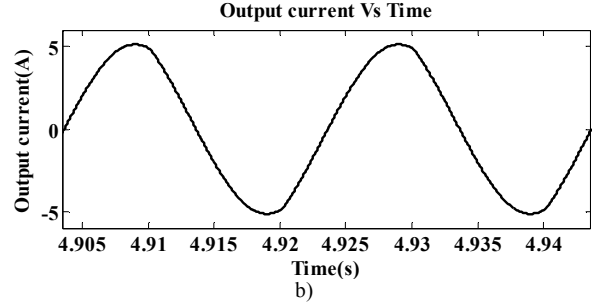Figure 10: Supply current; a)Before compensation b) After compensation

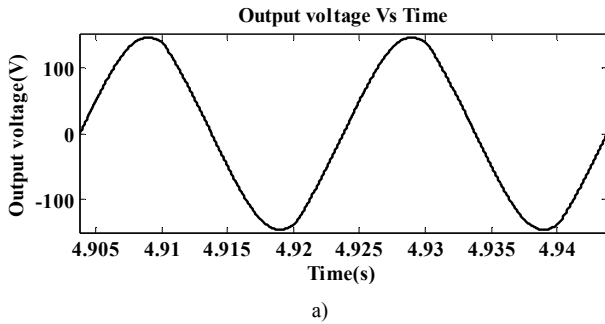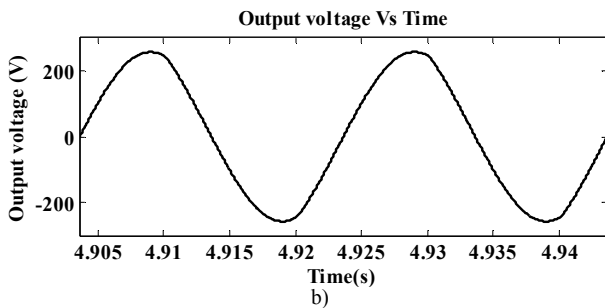Figure 11: Input signal; a) before compensation b) after compensation

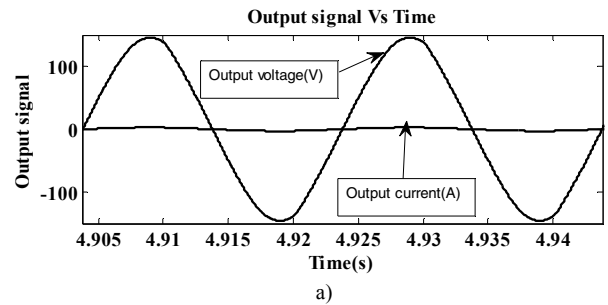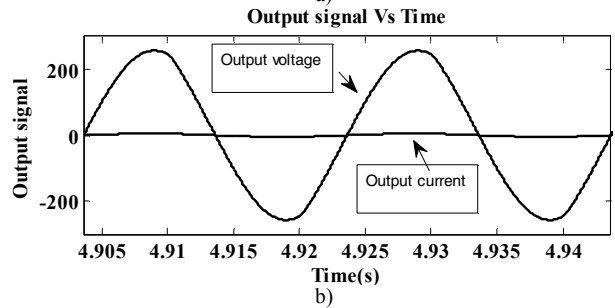Figure 13: Output current; a) Before compensation b) After compensation

Figure 12: Output voltage; a) Before compensation b) After compensation

Figure 14: Output signal; a) Before compensation b) After compensation

TABLE 2: COMPILATION OF RESULT

Signal	Analysis	Before compensation	After compensation
Input	*Fund. component*	18.15(peak)	24.41(peak)

current	*THD*	13.35%	3.86%
Output current	*Fund. component*	2.847(peak)	5.088(peak)
	THD	3.21%	3.24%

VII. CONCLUSION

In this paper, a study on the simulation of a direct AC-AC converter using single phase matrix converter topology in conjunction with the application of the active power filter operation has been implemented. It is proven from the result obtained that by using the current control closed-loop circuit, the AC-AC SPMC operation itself could be improved in terms of the THD reduction of supply current and also the output voltage of the system.

REFERENCE

[1] A. Alesina and M. G. B. Venturini, "Analysis and design of optimum-amplitude nine-switch direct AC-AC converters," *Power Electronics, IEEE Transactions on,* vol. 4, pp. 101-112, 1989.

[2] Z. Idris, M. K. Hamzah, and M. F. Saidon, "Implementation of Single-Phase Matrix Converter as a Direct AC-AC Converter with Commutation Strategies," in *Power Electronics Specialists Conference, 2006. PESC '06. 37th IEEE*, 2006, pp. 1-7.

[3] Z. Idris, S. Z. Mohammad Noor, and M. K. Hamzah, "Safe Commutation Strategy in Single Phase Matrix Converter," in *Power Electronics and Drives Systems, 2005. PEDS 2005. International Conference on*, 2005, pp. 886-891.

[4] A. Zuckerberger, D. Weinstock, and A. Alexandrovitz, "Single-phase matrix converter," *Electric Power Applications, IEE Proceedings -,* vol. 144, pp. 235-240, 1997.

[5] S. H. Hosseini and E. Babaei, "A new generalized direct matrix converter," in *Industrial Electronics, 2001. Proceedings. ISIE 2001. IEEE International Symposium on*, 2001, pp. 1071-1076 vol.2.

[6] A. Khoei and S. Yuvarajan, "Single-phase AC-AC converters using power MOSFETs," *Industrial Electronics, IEEE Transactions on,* vol. 35, pp. 442-443, 1988.

[7] Z. Idris, M. K. Hamzah, and A. M. Omar, "Implementation of Single-Phase Matrix Converter as a Direct AC-AC Converter Synthesized Using Sinusoidal Pulse Width Modulation with Passive Load Condition," in *Power Electronics and Drives Systems, 2005. PEDS 2005. International Conference on*, 2005, pp. 1536-1541.

[8] S. Firdaus and M. K. Hamzah, "Modelling and simulation of a single-phase AC-AC matrix converter using SPWM," in *Research and Development, 2002. SCOReD 2002. Student Conference on*, 2002, pp. 286-289.

[9] H. Guichun, "The Algorithm of a Reference Compensation Current for Shunt Active Power Filter Control," in *Business Intelligence and Financial Engineering (BIFE), 2010 Third International Conference on*, 2010, pp. 57-61.

[10] Y. Ma and L. Zhu, "An analysis and simulation of shunt hybrid active power filter," in *Electrical Machines and Systems, 2007. ICEMS. International Conference on*, 2007, pp. 90-93.

[11] N. A. Rahim, S. Mekhilef, and I. Zahrul, "A single-phase active power filter for harmonic compensation," in *Industrial Technology, 2005. ICIT 2005. IEEE International Conference on*, 2005, pp. 1075-1079.

[12] M. El-Habrouk, M. K. Darwish, and P. Mehta, "Active power filters: a review," *Electric Power Applications, IEE Proceedings -,* vol. 147, pp. 403-413, 2000.

[13] N. F. A. Rahman, M. K. Hamzah, S. Z. M. Noor, and A. S. A. Hasim, "Single-phase hybrid active power filter using single switch parallel active filter and simple passive filter," in *Power Electronics and Drive Systems, 2009. PEDS 2009. International Conference on*, 2009, pp. 40-45.

[14] R. Baharom and M. K. Hamzah, "A New Single-Phase Controlled Rectifier Using Single-Phase Matrix Converter Topology Incorporating Active Power Filter," in *Electric Machines & Drives Conference, 2007. IEMDC '07. IEEE International*, 2007, pp. 874-879.

[15] R. Baharom, M. K. Hamzah, A. Saparon, and I. R. Ibrahim, "Studies on control electronics implementation of Single-phase Matrix converter operating as AC-DC Converter with active power filter," in *Power and Energy Conference, 2008. PECon 2008. IEEE 2nd International*, 2008, pp. 758-763.

[16] K. S. Muhammad, R. Baharom, and M. K. Hamzah, "Advance battery charger topology using SPMC incorporating active power filter," in *Science and Social Research (CSSR), 2010 International Conference on*, 2010, pp. 548-553.