Human Shape Recognition using Fourier Descriptor

N. Md Tahir, A.Hussain, S. Abdul Samad, H. Husain and R. A. Rahman, *Members IEEE*

***Abstract*-The aim of this study is to investigate Fourier Descriptor (FD) as feature vectors for shape representation and recognition since FD is the best known boundary based shape descriptor and has proven to outperform most other boundary based methods in terms of accuracy. Furthermore, FD is also invariant to geometric transformations and has good noise tolerance. The main concern regarding FD is the number of terms that need to be maintained from the original Fourier transform for effective representation and description. A system that computed FDs of human and non human from their silhouettes; normalized the descriptors and further applied as feature vectors for recognition is developed. Initial results of experiment showed that using adequate number of both low and high frequency components could represent the shape based on high recognition rate achieved. The process of shape recognition using FDs looks promising.**

Keywords: Shape descriptor, Human recognition, Fourier Descriptor (FD)

I. INTRODUCTION

Low-level image features denote several of its perceived characteristics including shape, color, texture and spatial relationship. Among these features, objects shape of an image is the most vital feature since it represents the significant regions or relevant objects in the image. These make the shapes a very significant feature to human acuity. Generally, two categories for shape representations are available, one is based on the entire shape region (the pixels contained in the region), and the other is based on the boundary of objects (the pixels along the object boundary) [1] & [2]. A large amount of research work on shape representation has been done in past years. As a result, many techniques have been developed based on these features. However, there is no best technique so far based on specific application, all methods have their own strengths and weaknesses.

Manuscript received 25 February 2009.

Nooritawati Md Tahir and Ruhani Abdul Rahman are lecturers at [1]Faculty of Electrical Engineering Universiti Teknologi MARA (UiTM), 40450 Shah Alam, Selangor DE (norita_tahir@yahoo.com)

Aini Hussain, Salina Abdul Samad, and Hafizah Husain attached at [2]Dept of Electrical, Electronics and System, Faculty of Engineering, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor DE.
(norita@vlsi.eng.ukm.my)

FD approach is one of the techniques that are widely used in shape analysis. Chen et al. [3] studied the performances of Fourier descriptors and Hu's seven moment invariants for recognizing images with different spatial resolutions. Result showed the image recognition using FDs, achieve 100% recognition rate for spatial resolution not less than 64x64. Hazem et al. [4] presented a modified algorithm for Fourier descriptors, which are invariant under the image translation, rotation, and scaling in recognition of human faces. Simulation results showed that their algorithm is an efficient method for recognizing human faces even with rotation, scaling, changes in illumination as well as expression. Zhang et al. [5] & [6] have proposed a Generic Fourier Descriptor (GFD) desirable for image retrieval that has been tested on MPEG-7 region shape database. Comparisons have been made between GFD and MPEG-7 shape descriptor Zernike Moments Descriptor (ZMD) and results showed that the proposed GFD outperforms ZMD. The proposed GFD satisfies all the six requirements set by MPEG-7 for shape representation, that is, good retrieval accuracy, compact features, general application, low computation complexity, robust retrieval performance and hierarchical course to fine representation. Licsar et al. [7] have developed a static hand-gesture recognition system for the Human Computer Interaction based on Fourier descriptors for the classification of hand shapes. Further, Zahn et al. [8] showed in their work Fourier descriptors are very good features to use when dealing with particular types of shapes such as aircraft or alphanumeric characters.

Therefore, in this study, a method to represent human shape using Fourier Descriptor (FD) derived from shape signature is investigated. The aim is to use FD as object features to discriminate between human and non human. Further, a back propagation Artificial Neural Network (ANN) and Support Vector Machine (SVM) will be used to evaluate the recognition stage. Objects of non human include vehicles, trees, chair, racket, jacket, stand fan and some regular shapes namely, triangle, square and others positioned at various angle. Pre-processing and the conception of FD will be enlightened in Section 2. Section 3 described the classifiers chosen to perform the recognition task. Next, experimental and results achieved will be detailed in Section 4. Finally, in Section 5, we conclude our findings.

II. PRE-PROCESSING AND FEATURE EXTRACTION

A. Pre-processing

The shapes considered in this paper are outline shapes that can be described as single plane closed curves. The shapes are first converted to the form of gray level images and preprocessed. Pre-processing includes background subtraction, binarization, thresholding and median filtering for determining the shape silhouette [1] & [2]. Silhouettes do not have holes or internal markings, therefore the associated boundaries are conveniently represented by a single closed curve.

B. FD as Feature Extraction

Fourier transformation on shape signatures is widely used for shape analysis [1], [2], [5] & [6]. The Fourier transformed coefficients form the Fourier descriptors of the shape. These descriptors represent the shape of the object in a frequency domain. First, the N points forming the boundary of a region are required by taking all the pixels occupied by the boundary, or taking N samples from the boundary. This can be done by tracing the boundary counter clockwise. The region can be viewed as being in the complex plane with the ordinate being the imaginary axis and the abscissa being the real axis as depicted in Fig. 1.

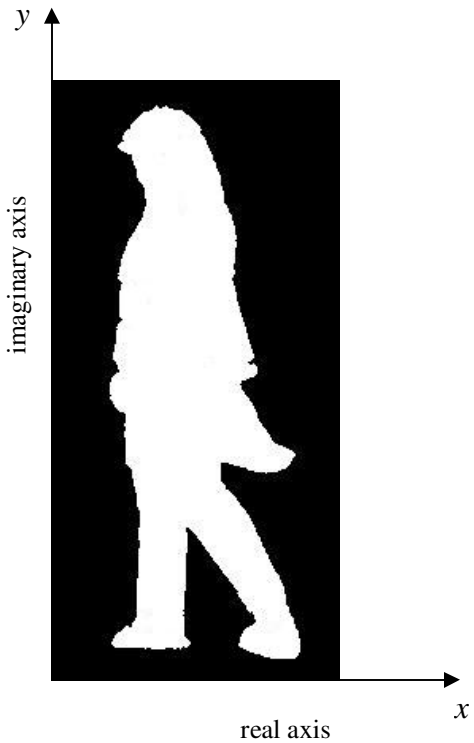

Fig 1 : Silhouette of a shape in the complex plane

The x-y coordinates of each point in the shape contour can be expressed in the form of $(x_k\ y_k)$ where $0 < k \leq N$-1. A complete set of coordinates describing the boundary is determined. With this notation, the contour can be expressed as coordinate series

$$s(k) = [x(k),\ y(k)], \text{ for } k = 0,1,2,\ldots\ldots N\text{-}1$$

and each term is a complex number

$$s(k) = x(k) + \mathrm{j}\ y(k) \tag{1}$$

The advantage of this representation is that it reduces a 2-D into a 1-D problem. The Discrete Fourier Transform (DFT) of this sequence is the FD of the contour. The DFT of $s(k)$ is:

$$A(u) = \sum_{k=0}^{N-1} s(k) e^{-j2\pi ku/N} \tag{2}$$

for $u = 0,1,2,3\ldots. N$-1, and $N = 40$ in this study. The complex coefficients $A(u)$ are called the FDs of the boundary. By applying inverse Fourier Transform to $A(u)$ restores the value of $s(k)$:

$$s(k) = \frac{1}{N} \sum_{u=0}^{N-1} A(u) e^{+j2\pi uk/N} \tag{3}$$

for $k = 0,1,2,\ldots\ldots N$-1

The normalized magnitude of the FD can eliminate dependency on size and used as a model for shape recognition. The high frequency descriptors contain information about finer details of the shape while the low frequency descriptors contain information about the general or global features of the shape. There are various drawbacks of using all the FD terms namely, the number of attributes could be too large and the model could be too specific for robust recognition. Thus, the optimal number of FD to be selected for shape recognition has to be determined first.

III. ANN AND SVM AS CLASSIFIERS

ANN and SVM have been chosen to evaluate the effectiveness of FD as feature vectors for recognition of human and non human. The ANN classifiers are well known for their abilities to express highly nonlinear decisions. This make them appropriate for recognition of complex patterns. They also possess the ability to maintain accuracy even when some input data are inappropriate and/or inadequate. In this study, the feed forward MLP neural network was used to evaluate the efficacy of the FD feature vectors. A three-layer NN with weights adjusted using the scaled conjugate gradient (SCG) algorithm [9] was employed. The MLP had an input layer consisting of number of neurons corresponding to the input features, one hidden layer and one output layer with two neurons to represent human and non human.

Next, the SVM, developed by Vapnik [10] as an implementation of structural risk minimization (SRM) was also considered. The idea behind SRM is that given a sequence of

hypothesis spaces of increasing complexity, one needs to choose the hypothesis space that minimizes the training error. Then from this same sequence of hypotheses, one must again choose the hypothesis that minimizes the upper bound of the generalization error. The SVM approximates the upper bound by performing these two tasks simultaneously and by controlling the size of the feature weights. This principle formulation of weight decay is used in neural networks to improve generalization [10] & [11]. In terms of geometrical interpretation, this is when an SVM chooses the optimal separating surface [10-12]. First, the SVM method was outlined for the linearly separable case. Kernel functions were then introduced to deal with non-linear decision surfaces. Finally, for noisy data, slack variables were introduced when complete separation of classes cannot be achieved [10-13]. In this study, the two classes SVM using RBF kernel is applied based on findings in [14], [15] and [16].

IV. EXPERIMENTAL RESULTS

To adopt FD, a boundary tracing is first required. The optimal number of FD required for human shape recognition is determined empirically. The data set used in this study consisted of 200 humans and 180 non human feature vectors. Cross validation method is chosen to estimate the classifier generalization error. A sample magnitude plots of human and non human is as shown in Fig. 2. From these plots, it is observed that only several first and last terms of the FD vectors contained the vital information of the shapes and these terms could be utilized to represent the shape contours. Therefore several silhouettes of human and non human are tested experimentally. We begin to classify the first 100 and last 100 terms of the FD vectors and evaluated the recognition rates using ANN and SVM as classifier. Both first and last term is then reduced in steps of 10. Table 1 illustrated the classification rate. We found out that the first 30 and last 30 terms in the FD contain most of the information. Thus, the 60 selected Fourier descriptors are opt for to characterize the shapes of human and non human. Fig. 3 demonstrated the 60 terms used for reconstructing the boundary of human and non human shapes to confirm that the FDs vectors selected is sufficient for representing the shape of the objects. Further, the reconstruction of shapes using only the first 30 terms or the last 30 terms of the FDs vectors are also performed. The purpose is to appraise the effect of these two terms have in the reconstruction process. Fig. 4 depicted the outcome of this evaluation. Using only the low term coefficient or the high term coefficient is inadequate for this purpose. This is because the high frequency descriptors contain information about finer details of the shape whilst the low frequency descriptors contain information about the general or global features of the shape.

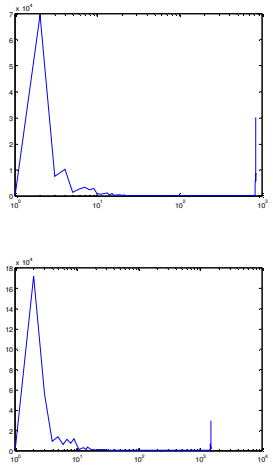

Fig. 2: Magnitude plots of FD for human (top); non human (bottom).

TABLE 1
CLASSIFICATION RESULTS USING DIFFERENT NUMBER OF FD FEATURE VECTORS

# of FD selected for first and last term	Classification Rate (%)	
	ANN	SVM
100	96	97
90	97	97
80	95	98
70	96	98
60	97	97
50	97	98
40	97	98
30	**97**	**98**
20	83	87
10	74	80

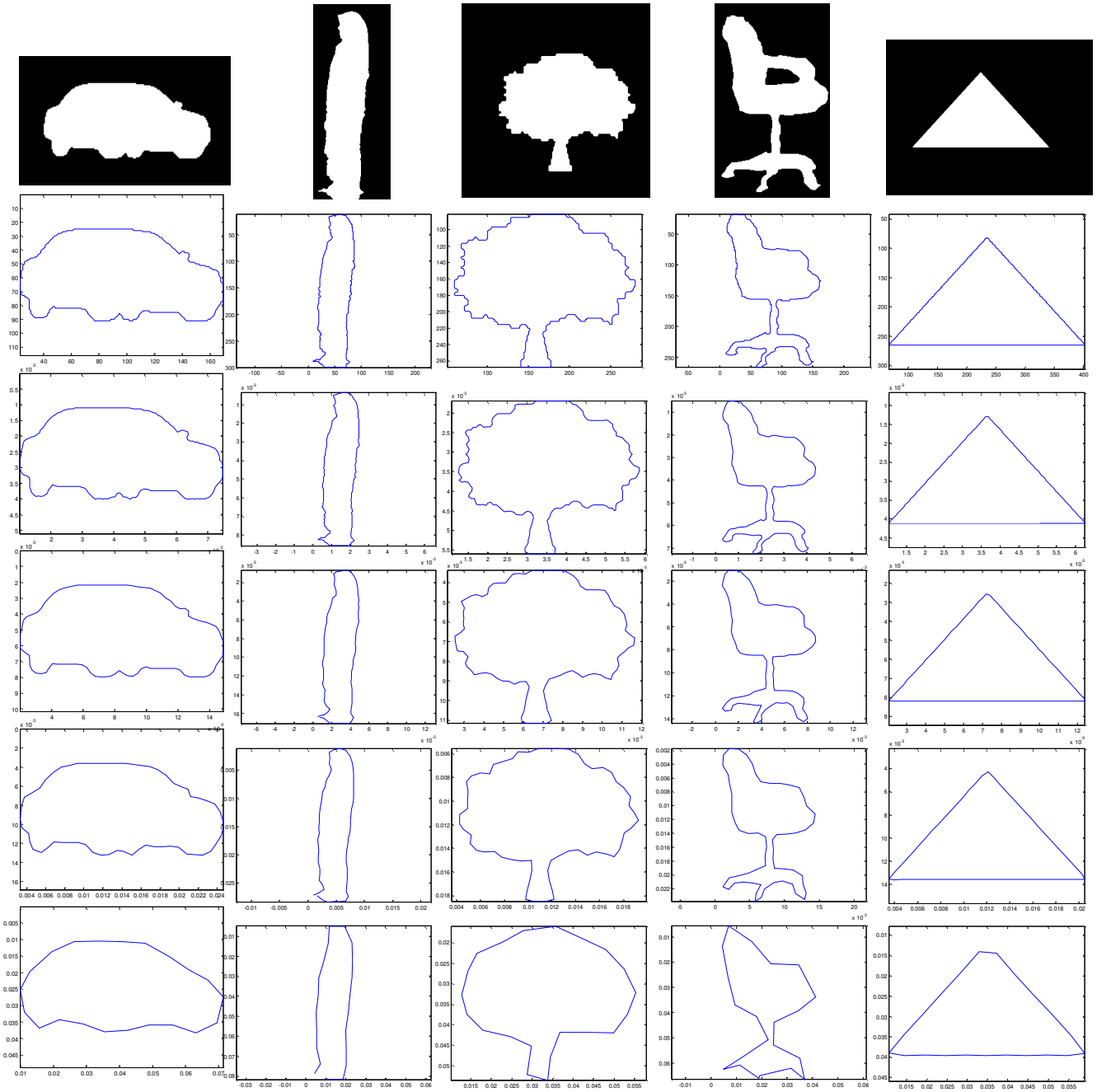

Fig.3: Some sample of reconstruction images of human and non human. From top to bottom: Original image: reconstructed image using all coefficients; first 100 and last 100 terms; first 50 and last 50 terms; first 30 and last 30 terms and only first 10 and last 10 Fourier coefficients.

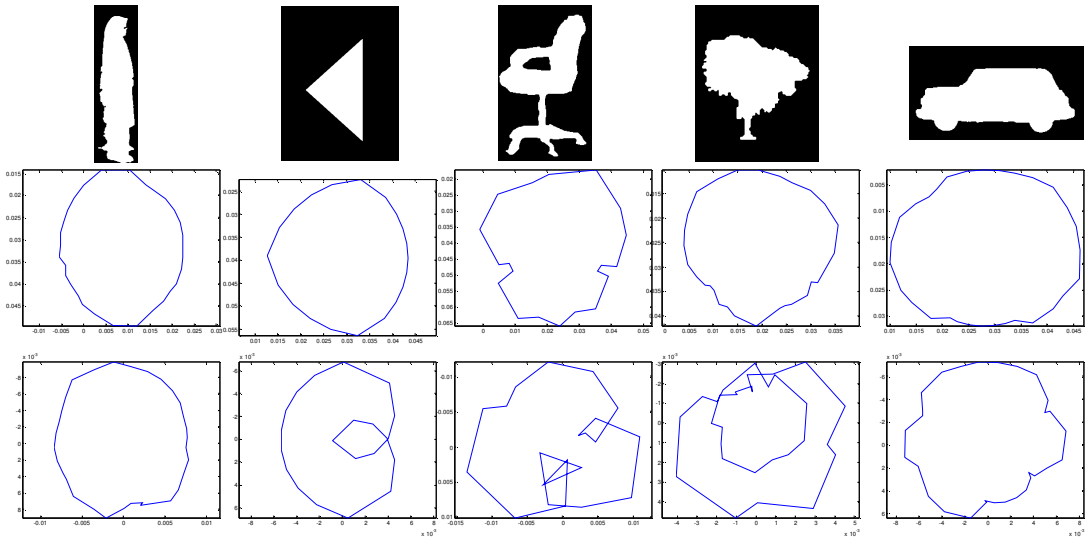

Fig 4: Some sample of reconstruction using only low and high terms.
From top to down: Original image; reconstructed image using only first 30 Fourier coefficients and only last 30 terms Fourier coefficients.

V. CONCLUSIONS AND FURTHER WORKS

In conclusion, an approach for human shape recognition using FD as the shape signatures was developed. It was found that using a reduced number of FD components specifically 60 normalized descriptors is sufficient for shape representation and classification of human and non human with above 96% classification rate with ANN and SVM as classifiers. This work could be applied for pedestrian recognition task in traffic scene analysis by eliminating irrelevant objects that are unavailable in traffic scene. Objects of focused for this purposes will be human, vehicles and trees. The number of FDs required to represent these shapes will also be evaluated.

ACKNOWLEDGMENT

This work was supported by MOSTI under the IRPA Grant No: 03-02-02-0017-SR0003/07-03. The authors also acknowledge Prof Burhanuddin Yeop Majlis as the Programme Head and UiTM for the UiTM-JPA SLAB scholarship awards.

REFERENCES

[1] Gonzalez R, Woods R, "Digital Image Processing", 2nd Edition, Prentice Hall, New Jersey, USA, 2000.

[2] Mitra, S. K. " Digital Signal Processing: A computer based approach", 2nd Edition. *McGraw-Hill,* New York, 2002.

[3] Q Chen, EM Petriu and X Yang ,"A Comparative Study of Fourier Descriptors and Hu's Seven Moment Invariants for Image Recognition", *Proceeding of IEEE Canadian Conference on Electrical and Computer Engineering 2004,* pp103-109, Canada, May 2004.

[4] M Hazem, El- Bakry Mohy and A. Abo-Eboud Kamel, "Integrating Fourier Descriptors and PCA with Neural Networks for Face Recognition ", *Proceeding of Seventeenth N.ational Radio Science Conference*, Egypt, Feb. 2000.

[5] D. S. Zhang and G. Lu, "Content-Based Shape Retrieval Using Different Shape Descriptors: A Comparative Study", *Proceeding of IEEE International Conference on Multimedia and Expo (ICME2001),* Tokyo, Japan, pp.317-320, August 22-25, 2001.

[6] D. S. Zhang and G. Lu. "Shape Based Image Retrieval Using Generic Fourier Descriptors", *International Journal of Signal Processing: Image Communication*, August 2001.

[7] Attila Licsár, "Hand-Gesture Based Film Restoration", *Proceeding of 4th International Conference on Enterprise Information System",* pp95-103*,* Spain, April 2002.

[8] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Trans. Computers*, vol. 21, 269- 281 1972.

[9] C. Bishop, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995.

[10] Vapnik VN, "The nature of Statistical Learning Theory", Springer, New York, 1995.

[11] N. Cristianini, J. Shawe-Taylor, Support Vector Machines, Cambridge University Press, Cambridge, 2000.

[12] Burges CJC, " A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery 2, 121-167, 1998.

[13] Gunn S, "Support Vector machines for classification and regression", *ISIS Technical Report*, University of Southampton, 1998.

[14] J. Qin, S.H. Zhong, "SVM Face Recognition Based on Gabor Featured Key Points," *Proc of IEEE Conference on Machine Learning and Cybernetics*, Aug. 2005, pp. 5144-5149.

[15] B. Moghaddam B, M.H. Yang "Gender Classification with Support Vector Machine," *Proc of IEEE Conference on Automatic Face and Gesture Recognition*, 2000.

[16] Hua S, Sun Z, "Support Vector Machine approach for protein subcellular localization prediction", *Bioinformatics 17*, 721-728, 2001.

Nooritawati Md Tahir is an associate professor in Faculty of Electrical Engineering, Universiti Teknologi Mara, Shah Alam, Selangor, Malaysia. She obtained her BEng in Electrical Engineering from the Universiti Teknologi Mara (UiTM) in 1988, MSc (Electronics & Communication) in 1991 from the University of Liverpool, UK and PhD in Pattern Recognition form Universiti Kebangsaan Malaysia in 2008. Her research field includes image processing, pattern recognition and computational intelligence.

Aini Hussain is a professor of intelligent signal processing and pattern recognition at the Universiti Kebangsaan Malaysia since February 2007. She received her bachelor degree in Electrical Engineering from Louisiana State University, USA; MSc in Control Engineering from University of Manchester Institute of Science and Technology (UMIST), England and PhD in Neural Network with concentration in automatic speech recognition from the Universiti Kebangsaan Malaysia in 1985, 1989 and 1997, respectively. She is currently the head of the Electrical, Electronic and Systems Engineering Department of the National University of Malaysia (UKM). Her current research interests include signal processing, intelligent pattern analysis and recognition, computational intelligence, bio-signal processing and biologically inspired networks.

Salina Abdul Samad obtained her BSc in Electrical Engineering from the University of Tennessee, USA and a PhD from the University of Nottingham, England. Her research interests are in the field of digital signal processing, from algorithm design to software and hardware implementation. She is now employed by Universiti Kebangsaan, Malaysia as a professor.

Hafizah Husain obtained her BSc in Electrical Engineering from University of Hartford, USA in 1986, MSc in Mechatronics from De Montfort University, UK in 1995 and PhD in 2007 from Universiti Teknologi Malaysia. She is currently a lecturer at the EESE Department of Universiti Kebangsaan Malaysia. Her current research interests are in the area of soft computing, fuzzy logic and artificial neural network.

Ruhani Ab Rahman is an associate professor in Faculty of Electrical Engineering, Universiti Teknologi Mara, Shah Alam, Selangor, Malaysia. She obtained her BEng in Electrical & Electronic Engineering from the Portsmouth Polytechnic in 1988, UK and MSc (Data Telecommunication & Networks) in 1998 from the University of Salford, UK. Her research field includes computer networks, mobile networks and wireless sensor.