

SHAPE-BASED RECOGNITION USING COMBINED JACCARD AND MAHALANOBIS MEASUREMENT



**RESEARCH MANAGEMENT INSTITUTE (RMI)
UNIVERSITI TEKNOLOGI MARA
40450 SHAH ALAM, SELANGOR
MALAYSIA**

**BY:
DR. SITI SALWA SALLEH
DR. FAKARIAH HANI MOHD ALI
PROF DR. DAUD MOHAMAD
ASSOC PROF DR. MEGAWATI OMAR**

OKTOBER 2012

Contents

- 1. Letter of Report Submission iii
- 2. Letter of Offer (Research Grant)..... iv
- 3. Acknowledgements **Error! Bookmark not defined.**
- 4. Enhanced Research Title and Objectives vi
- 5. Report 1
 - 5.1 Proposed Executive Summary 1
 - 5.2 Enhanced Executive Summary..... 1
 - 5.3 Introduction 2
 - 5.4 Brief Literature Review 7
 - 5.5 Methodology..... 10
 - 5.6 Results and Discussion 13
 - 5.7 Conclusion and Recommendation..... 17
 - 5.8 References/Bibliography 19
- 6. Research Outcomes..... 24
- 7. Appendix 25

3. ACKNOWLEDGEMENT

PENGHARGAAN

Setinggi-tinggi penghargaan dan ribuan terima kasih diucapkan kepada semua pihak yang terlibat secara langsung dan tidak langsung bagi membolehkan penyelidikan ini disiapkan dengan sempurna.

Diantaranya :

Prof. Dr. Azlinah Mohamed
(Dekan Fakulti Sains Komputer dan Matematik)

Prof Dr. Abdul Bakar Majeed
(Penolong Naib Canselor (Penyelidikan))

dan

Semua individu yang telah terlibat di dalam penyelidikan dan yang telah memberikan kerjasama dan sokongan di dalam menjayakan penyelidikan ini

5. Report

5.1 Proposed Executive Summary

A shape-based classifier can be built to recognize shapes using a distance measure approach. Therefore, this research is going to produce a new classifier by integrating the Mathematical equations of Jaccard and Mahalanobis. This new classifier will recognize 2-dimensional objects faster and more accurately, which is within seconds instead of minutes. The present Mahalanobis distance measure suffers from delayed processes as its computation time formula reaches $O(n^2)$ for n-dimensional feature's vector. This means the larger the data the longer it takes. Thus by combining the Mahalanobis math equation with Jaccard's, the expected resulted classifier will not be affected by data size. Furthermore, human drawing and sketches cannot be in similar size, similar pen pressure, clarity and orientation at all time, thus the combination of the 2 Mathematical equations above will address these issues. Secondly, the combination of 2 Mathematical equations is expected to enable the classifier detect sketches on real time with guiding assistance, to see if one's drawing is correct, as such, immediately guiding a person to re-draw. The methodology will include an analysis of Jaccard and Mahalanobis equations, studies its measurement characteristics, strength and drawbacks. This will be followed by combining the math equations, and implementation. At the implementation stage, a data collection and analysis will be done. Data analysis consists of data filtering, feature extractions, shape normalization and segmentation tasks. Finally, the new classifier will be tested to recognize the the shapes collected earlier. The expected outcomes are i) a combined Jaccard and Mahalanobis distance measures, ii) a shape recognition classifier that can recognize shapes in seconds iv) real time checker that guide help users to draw. Continuity prospect: The classifier can be used for teaching purposes at various levels; for example for students in engineering, medicine, math and computer science mechanical drawing lesson; and also for medical therapy such as psychomotor dyslexia therapy and exercises.

5.2 Enhanced Executive Summary

Shape recognition is an extensive research ares in recent years as a result of the growing technologies of pen-based interface and increasingly in volume of digital images. Numerous shape recognition approaches have been developed in handling variance of image transformations and strokes for free-hand sketching however promising result has yet been proven. Complexity in the computation measure, expensive in computation time, requires large training data set and extensive training are not practical, especially in sketching recognition. Similarity measurement such as Jaccard distance, Mahalanobis distance and others are commonly used in recognition tasks which offer a simple computation and do not require large training data set. In conventional Jaccard distance, the way to measure the similarity based on overlapped

1. Letter of Report Submission