TasMuQ (Tashih Mushaf Al-Quran Authentication)

MOHD AFIF NAJMI BIN MOHD RADZIF

BORANG PENGESAHAN STATUS TESIS

JUDUL: TASHIH MUSHAF AL-QURAN AUTHENTICATION (TASMUQ)

SESI PENGAJIAN: 2016/2017

Saya MOHD AFIF NAJMI BIN MOHD RADZIF
(HURUF BESAR)

mengaku memben arkan tesis (PSM/~~Sarjana/Doktor Falsafah~~) ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Sila tandakan (/)

_____ SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
_____ TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
____ TIDAK TERHAD	

(TANDATANGAN PENULIS)

Alamat tetap: F480,
KAMPUNG SINGKIR GENTING,
BEDONG,
08110 MELAKA.

Tarikh: _____

(TANDATANGAN PENYELIA)

DR. MOHD SANUSI BIN AZMI
Nama Penyelia

Tarikh: _____

CATATAN: * Tesis dimaksudkan sebagai Laporan Akhir Projek Sarjana Muda (PSM)
** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa.

TASMUQ – TASHIH MUSHAF AL-QURAN AUTHENTICATION

MOHD AFIF NAJMI BIN MOHD RADZIF

This report is submitted in partial fulfillment of the requirements for the
Bachelor of Computer Science (Software Development)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA
2017

DECLARATION

I hereby declare that this project report entitled

TASMUQ – TASHIH MUSHAF AL-QURAN AUTHENTICATION

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT : _____ Date: _____

(MOHD AFIF NAJMI BIN MOHD RADZIF)

SUPERVISOR : _____ Date: _____

(DR. MOHD SANUSI BIN AZMI)

DEDICATION

To my beloved parents. teachers and friends.

ACKNOWLEDGEMENTS

First and foremost, I would like to praise and thank God Allah SWT , the almighty, who has granted countless blessing, knowledge, and opportunity to the writer, so that I have been finally able to accomplish the thesis.

I would like to show my greatest appreciation to my beloved parents Mohd Radzif Bin Ismail and Ruslina binti Said who have been giving me support and motivation throughout my project.

Apart from the efforts of me, the success of this thesis depends largely on the encouragement and guidelines from my supervisor. I would also like to thank Dr. Mohd Sanusi Bin Azmi for giving assistant to complete this project successfully.

Last but not least, for those who are not included, which help me formally or informally, I surely very grateful for all the help and assistance I could get. Thank you.

ABSTRACT

This project is about an image processing web application which convert the original image to more meaningful images and check the similarity on system which is provided on database. The purpose of this project is to authenticate the image by processing the image and get the information from the images, then the system will compare the information on image with database which are the Al – Quran valid resources. Al-Quran content have a different pattern and texture. This is because Al-Quran came from many source of publisher. The expert analyst difficult to analyze and investigate the Al-Quran word is from particular approve company or the wrong Al-Quran. So the first issue will be arising when Mushaf content is well protected and writing, printing and arrangement error will cause sensitive issue not only in Malaysia but in global level as well. Mushaf in Malaysia divide to printing in written book and in a digital format such as image format, pdf and software. Mushaf that are not approved are not allowed to sell to public. Propensity to occur the mistake is high. Therefore, this study developed Tashih Mushaf Al-Quran Application which then this application can be used by Mushaf Al-Quran publication company and Islamic Community not only in Malaysia or oversea.

ABSTRAK

Projek ini adalah satu aplikasi web pemprosesan imej yang menukar imej asal untuk imej yang lebih bermakna dan daftar persamaan pada sistem yang dibekalkan di dalam pangkalan data. Tujuan projek ini adalah untuk mengesahkan imej dengan pemprosesan imej dan mendapatkan maklumat daripada imej, maka sistem akan membandingkan maklumat pada imej dengan pangkalan data yang merupakan Al - Quran sumber sah. kandungan Al-Quran mempunyai corak yang berbeza dan tekstur. Ini adalah kerana Al-Quran datang dari banyak sumber daripada penerbit. Penganalisis pakar sukar untuk menganalisis dan menyiasat perkataan Al-Quran adalah dari tertentu syarikat atau yang salah Al-Quran meluluskan. Jadi isu pertama akan timbul apabila kandungan Mushaf dilindungi dengan baik dan penulisan, percetakan dan kesilapan perkiraan akan menyebabkan isu sensitif bukan sahaja di Malaysia tetapi di peringkat global juga. Mushaf di Malaysia membahagi untuk mencetak dalam buku yang ditulis dan dalam format digital seperti format imej, pdf dan perisian. Mushaf yang tidak diluluskan tidak dibenarkan untuk menjual kepada orang ramai. Kecenderungan untuk berlaku kesilapan adalah tinggi. Oleh itu, kajian ini dibangunkan Tashih Mushaf Permohonan Al-Quran yang kemudiannya permohonan ini boleh digunakan oleh Mushaf syarikat penerbitan Al-Quran dan Masyarakat Islam bukan sahaja di Malaysia atau luar negara.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

UTeM	Universiti Teknikal Malaysia Melaka
TasMuQ	Tashih Mushaf Al-Quran
IDE	Integrated Development Environment
JPEG	Joint Photographic Experts Group
PNG	Portable Network Graphics

CHAPTER I

INTRODUCTION

1.1 Introduction

The term image segmentation is the process of partitioning a digital image into multiple segments. This term also refers to the partition of an image into a set of regions that cover it. The term image processing also give a broad definition is analyzing and manipulate digital images, in particular to improve quality. The goal is to use the segmentation is to simplify and change the representation of the image to represent into something more meaningful and easier to analyze when comparing the data from images with database that have valid authentication of Al-Quran words. The process is fairly well controlled to remove only the part that needs to be analyzed further. Al-Quran will be extract its frame using illumination of Al-Quran frame and using appropriate segmentation method. When extracting a complete word of Al-Quran in a page, this can analyze in a further method. The data on processing the image will be compared to the valid data in the database this will state the original information of pattern and the current valid information of Al-Quran.

The methods that used in image processing Mushaf Al-Quran is when extracting f triangular geometry features Mushaf image. The resulting of features from image extracting image from Mushaf will be used as model or test data. Valid Al-Quran copies

will be used as model while, copies of Mushaf Al-Quran in a market will be used as test data. Experiment will be execute using two approach which based on the distance of the closest data and supervision of supervised machine. Technical process of system review and correction of Mushaf Al-Quran.

Based on the project, the feature of this algorithm is to process an image from extraction of Al-Quran words analysis and process it to segmentation while extracting the data from image and compare the image with the valid Mushaf Al-Quran model on database.

1.2 Problem Statement(s)

The issue will be arising when Mushaf content is well protected and writing, printing and arrangement error will cause sensitive issue not only in Malaysia but in global level as well. Mushaf in Malaysia divide to printing in written book and in a digital format such as image format, pdf and software. Mushaf that are not approved are not allowed to sell to public. Propensity to occur the mistake is high.

The Al-Quran verse in internet or image cannot be verify immediately its required its own work and time to search from the correct information. This will be wasting work and time to search for correct information, may take hours or days to search for correct information.

A study of Qur'an manuscripts from around the worlds has revealed a number of publication of around the world with different style of manuscript associated with different country, state and regions. Different state has its own style and will be found a various type of decoration of frame. The challenged is when extracting word from picture that have various style of pattern and compare it with valid database model.

1.3 Objective

Objective of this project are:

i. To investigate the image processing techniques which required in order to process Al-Quran image.
ii. To develop an application for Al-Quran authentication from image processing

1.4 Scope

Scope considered and applied in this study are:

i. The project to investigate the current valid Al-Quran verses for segmentation techniques with illumination frame.
ii. The project to investigate the image processing Al-Quran data for comparing the data with valid data in database.
iii. The project to investigate the actual test data that can compare with valid data model.

1.5 Project Significance

This study is very important to research domain, knowledge in the field of extraction of features and applications to the real world. The project significance as follow:

i. This project will help analysis expert to extracting Al-Quran text from its binarization frame. Then, further investigate will be conducted to process the image. The data result from this project can be carried out to next level of image processing in segmentation.

ii. This project will be used by Al-Quran publication company as software for check and reviewing the validity of Mushaf Al-Quran such as writer, printing company, year, date, edition, Mushaf reviewer and current status of publication.

iii. This project also may be used by Islam Community, this software will be used to help the process of purchasing Mushaf Al-Quran, and report the error in Mushaf and to get an information about to be purchased Mushaf.

iv. Philosophically, this project will be covered several domain and knowledge. This features focus on image processing and segmentation and compare it with valid data of Mushaf Al-Quran. It involves technique that will give impact to image analysis field for further process.

1.6 Expected Output

The result from this features extraction will be produced Al-Quran text without its illumination frame. Pages of Al-Quran that contains text and illumination will be used as data test. The data test will be converted to binary set and then will be process to produce a result that be extracted from its illumination frame. The data then will be inverted to black and white color pixels' image because during the image analyses process, the image will convert to binary set. After process its will change to RGB (Red Green Blue) color by pixels, then the process will continue after the extraction. The segmentation of image will be executed, then it extracts the information from model and compare it with valid database resources.

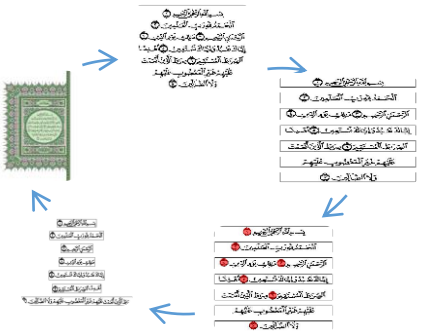

Figure 1-1.1 : Extracting Images

1.7 Conclusion

This project will have covered theory of image processing using image segmentation method to produced specific result that going to be extract the data from image and comparing the data test with actual model of valid Al-Quran verses. The concept of this project is when image segmentation is used to find data and extract the data test from image and divide the image into several data test such as frame, row, verses. This technique then will be compare with actual valid data and give the valid information for the images.

The purpose for this project is for comparing the data test with valid resources of Al-Quran in a database for authentication. The result of processing image can further the investigation of comparing the result and process by analyses expert. This project can contribute huge impact for image processing domain specifically. The intention from this project is going to comparing the data test of Al-Quran and actual data from database with a valid Al-Quran authentication.

Chapter 2 will cover about literature review and project methodology.

CHAPTER II

LITERATURE REVIEW AND PROJECT METHODOLOGY

2.1 Introduction

The research about implementation of segmentation was already propagating through several of project and features. For example, image segmentation of the medical imaging, locate tumors and other pathologies, object detection, pedestrian detection and also face detection.

This topic will explain about facts and finding related with domain, existing system and technique. Besides that, this topic will discuss about project methodology and project requirement related with software requirement, hardware requirement and other requirements. Lastly, this topic will briefly describe about project schedule and milestones.

2.2 Facts and Findings

The uniqueness and beauty of the Quran is not only to holy words of Al Qur'an, but the writing of the Qur'an will make it even more beautiful with the lighting that comes from a variety of patterns, textures and decorations. However, this research project will focusing on comparing Al-Quran text from the image processing with the actual data in database. Then, it will become part of research study to comparing the Al-Quran text with the actual data.

Al-Quran is literally meaning “the recitation”, it is the central religious text of Islam, which Muslim believe to be a relation from Allah. Al-Quran was written in an Arabic word; thus this research can be referring to the research domain on an Arabic handwritten.

Nowadays, research on Arabic handwritten by using Jawi and Parsi already introduce by Anton Heryantp (2008) on Arabic word, Khairuddin Omar (2000), Mazani Manf (2002), Remon Redika (2008) on Jawi word. On Latin word are being introduce by Moalla et al. (2006), Aiolli et al. (1999).

Segmentation techniques on this time only focus on separation of word into sub-word and sub-word into characters. On Arabic domain, research of separation word into sub-word introduced by Mazani Manf (2002) and Remon Redika (2008). However, on Jawi are being introduced by Khairuddin Omar (2000).

Segmentation technique of word onto sub-word introduced by Mohammad Faidzul (2010) using features of trace transform and feature comparison. Remon Redika (2008) using same technique with Mazani Manf (2002) where it will separate sub-word based on distribution of black pixels. On Jawi domain, Khairuddin Omar (2000) using statistic features of non-moment change and the distribution of black pixels.

All of this technique only focusing on segmentation of word into sub-word and also sub-word into character, but not focus on eliminate the illumination.

2.2.1 Domain

The domain of this project is focusing on image processing and image data comparing with actual data model. Image processing will manipulate and process to produce as an image result with more meaningful image, then the result of image will be extract to it and compare the result with actual database of Al-Quran. The focusing domain project research is more toward on comparing the data test with validation of Al-Quran. This, the process need to be well enough controlled to compare the only parts that need to be analyzed further.

2.2.2 Existing System

This project is focusing on validation of Al-Quran. Thus, it has not recently develop but some features of Arabic Segmentation already exist and research by Mohd Sanusi Azmi (2015), Amirul Ramzani (2015) as already described on sub-topic 2.2.1 Domain.

2.2.3 Technique

Technique to be used to develop this features is image processing using technique image segmentation and data comparing with data model. This technique can be used to extract and image and comparing the data with actual data model.

2.3 Project Methodology

This topic will be discussed about a research methodology that going to be used to fulfil project's objective. It will be focusing on conceptual role during the execution of the project. Execution of the research framework will be also discussed in this topic.

Research framework of methodology are divided into three phases which is investigation phase, execution phase and development phase.

Research framework methodology	Investigation	1. Analyses the problem.	
		2. Eliciting the problem requirement.	
		3. Analyses Al-Quran pages of design and decoration.	
		4. Investigate the technique and method for image processing	
	Execution	5. Collecting test image.	
		6. Model of features.	6.1Segmentation
			6.2Detection (removing)
		7. Result data test.	
	Development	8. Collecting data of analysis	
		9. Comparing test data with valid Al-Quran resources	
		10. Result of analysis	

Figure 2.1 : Table of Research Framework

2.4 Project Requirement

There are some tools to be used before, during and after developed this features. These tools have their own features and characteristic to build this project. It helps developer to execute the project with easy.