REMOVING AL-QURAN ILLUMINATION

AMIRUL RAMZANI BIN RADZID

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS TESIS

JUDUL: REMOVING AL-QURAN ILLUMINATION

SESI PENGAJIAN: 2014/2015

Saya AMIRUL RAMZANI BIN RADZID
(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: N0 39, JLN BLI 3,
TAMAN BUKIT LARANG INDAH,
TELOK MAS,
75460 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.

REMOVING AL-QURAN ILLUMINATION

AMIRUL RAMZANI BIN RADZID

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
2015

DECLARATION

I hereby declare that this project report entitled

REMOVING AL-QURAN ILLUMINATION

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

STUDENT : _____ Date: _____

(AMIRUL RAMZANI BIN RADZID)

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 Radzid Bin Ideris and Rugayah Binti Sulaiman 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 implement the image processing technique which convert the original image into more meaningful image. The purpose of this project is to remove Al-Quran illumination using image processing technique. Al-Quran illumination content different pattern and texture. This is because Al-Quran came from many source of publish Makkah and Madinah. The expert analyst difficult to analyse and investigate the word of Al-Quran cause of its illumination. So, the first issues that being rise to analyse a word of the Al-Quran is the distraction came from illumination frame of the Al-Quran itself. The second issues is the variety pattern of illumination frame of the Al-Quran that need to be removed. The last issues is the perfect align for to differentiated illumination frame. Objective of this project are to investigate segmentation techniques for removing illuminations occur in Al-Quran and to develop segmentation technique for removing illumination for Al-Quran. The method use for implement this project is image segmentation which convert grey-scale image into binary set. This project also implement algorithm to eliminate frame from Al-Quran text word. The results from this project are Al-Quran text without illumination and calculate differenced text value of Al-Quran.

ABSTRAK

Projek ini melaksanakan teknik pemprosesan imej yang menukar imej asal ke dalam imej yang lebih bermakna. Tujuan projek ini adalah untuk membuang iluminasi Al-Quran menggunakan teknik pemprosesan imej. Iluminasi Al-Quran mengandungi pelbagai corak dan tekstur yang berbeza. Ini adalah kerana Al-Quran diperolehi daripada banyak sumber terbitan Makkah dan Madinah. Pakar penganalisis sukar untuk menganalisis dan mengkaji ayat Al-Quran keranahalngan daripada iluminasi tersebut. Oleh itu, isu pertama yang timbul adalah untuk menganalisis ayat Al-Quran dan apabila halangan yang timbul itu datang daripada bingkai iluminasi Al-Quran itu sendiri. Isu yang kedua ialah pelbagai jenis corak bingkai iluminasi Al-Quran yang perlu dibuang. Isu yang terakhir adalah menyelaraskan pembezaan yang sempurna untuk membezakan bingkai iluminasi Al-Quran. Objektif projek ini adalah untuk menyiasat teknik segmentasi untuk mengeluarkan iluminasi dalam Al-Quran dan untuk membangunkan teknik segmentasi untuk mengeluarkan iluminasi Al-Quran. Kaedah yang digunakan untuk melaksanakan projek ini adalah segmentasi imej yang menukar imej skala kelabu ke dalam set binari. Projek ini juga menggunakan algoritma untuk menbuang bingkai daripada huruf Al-Quran. Hasil daripada projek ini adalah teks Al-Quran tanpa iluminasi dan menghitung perbezaan nilai teks Al-Quran.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

UTeM	Universiti Teknikal Malaysia Melaka
IDE	Integrated Development Environment
JPEG	Joint Photographic Experts Group
PNG	Portable Network Graphics
TXT	Text

CHAPTER I

INTRODUCTION

1.1 Introduction

The term image segmentation refers to the partition of an image into a set of regions that cover it [1]. The purpose to used segmentation is to simplify and change the representation of the regions of an image to represent into something more meaningful and easier to analyse. Segmentation method will be used to decompose the page of Al-Quran into parts for further analysis. The process well enough controlled to extract only parts that need to be analysed further. An Al-Quran page was presented for segmenting its frame illumination using appropriate segmentation method. Extracting a complete word of Al-Quran in a page need a proper algorithm method. Segmentation also used to perform a change of representation of Al-Quran. Illumination of Al-Quran will be extracting out from a page that has contains a word and the result will be proceed to further process.

The regions of illumination Al-Quran will be identify using segmentation method. The method of image segmentation that be used is the thresholding method. A colour of page of Al-Quran will convert to gray-scale image using Otsu's method. A grey-scale image will turn to binary image for further analysed. There is also a balanced historgram thresholding. The binary image will be analysed using the frequency of image produced by gray-scale image. The pattern and the texture of a page of an Al-Quran will be detected for further investigated. The patterns of the frame become the factor toward investigation to be analysed because the illumination

of Al-Quran came from variant sources and productions. The tracing techniques using the complexity of algorithm being implemented to detected the illumination that contains multiple variant of pattern.

Based on the project, the feature of this algorithm is to implement the extraction of the Al-Quran words toward more meaningful of Al-Quran analysis. The feature can detect words of Al-Quran to proper further analysed and investigated without distraction from its illuminations on the same page of Al-Quran

1.2 Problem statement(s)

The first issues that being rise to analyse a word of the Al-Quran is the distraction came from illumination frame of the Al-Quran itself. The second issues is the variety pattern of illumination frame of the Al-Quran that need to be removed. The last issues is the perfect align for to differentiated illumination frame.

A study of Qur'an manuscripts from around the world has revealed a number of particular styles of manuscript illumination associated with different country, state and regions. Different state has different own style will be found a wide variety of decoration illumination styles. When an Al-Quran manuscript is illuminated, it will be difficult to analyse and process the text of Al-Quran for further analysed the features. The structure, motifs and colours of the decorated double frames found are the key challenged to be removed.

1.3 Objective

Objective of this project are:

i. To investigate segmentation techniques for removing illuminations occur in Al-Quran.
ii. To develop segmentation technique for removing illumination for Al-Qurans.

1.4 Scope

Scope considered and applied in this study are:

i. The project to investigate segmentation techniques for detecting Al-Quran illuminations frame.
ii. The project to investigate segmentation techniques for convert image to binary and turn to RGB colour.
iii. The project to investigate segmentation techniques for removing Al-Quran illuminations frame.
iv. The project to investigate segmentation techniques for extracting Al-Quran text from a pages that contains illuminations frames.

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 follows:

i. This project will help analysis expert to extracting Al-Quran text from its illumination 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 help to extract a bundle of Al-Quran pages and take less time for extraction to produce the result. From the current Al-Quran, Al-Quran Al-Hakeem contains 608 pages, Al-Quran Al-Kareem (Saudi Style) contains 617 pages, Al-Quran Al-Majeed contains 855 pages and Mushaf Al-Madinah Quran Majeed contains 625 pages. By using this features, it will make easy for analysis expert to extract Al-Quran pages to Al-Quran words with efficiently and effectively.
iii. This project introduce suggested features for detecting illumination frame from any manuscripts. The features is detecting the continue point until it reaches the end point. By using this features, any unwanted line or region will be eliminated from the text. The method to implement the algorithm is precisely conduct to

make the output result as expected. The result of detecting illumination frames as Figure 1.1 below:

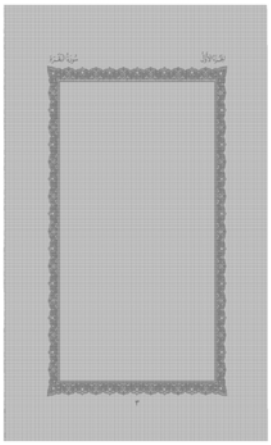
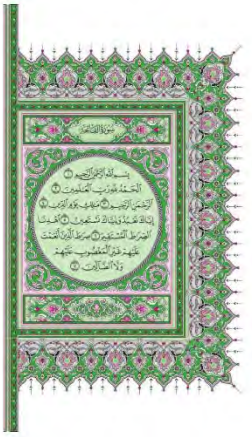
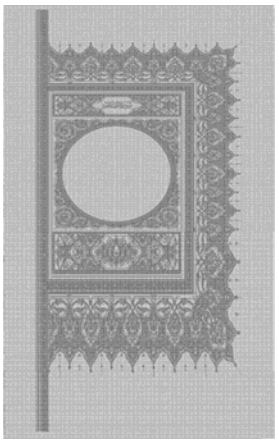
Data Test	Detacted Frame
 Image Al-Quran Al-Karim from Mushaf Al-Madinah Quran Majeed	 Image of Illumination from Mushaf Al-Madinah Quran Majeed
 Image Al-Quran Al-Karim from Mushaf Al-Madinah Quran Majeed	 Image of Illumination from Mushaf Al-Madinah Quran Majeed

Image Al-Quran Al-Karim from KSU - Electronic Mosshaf

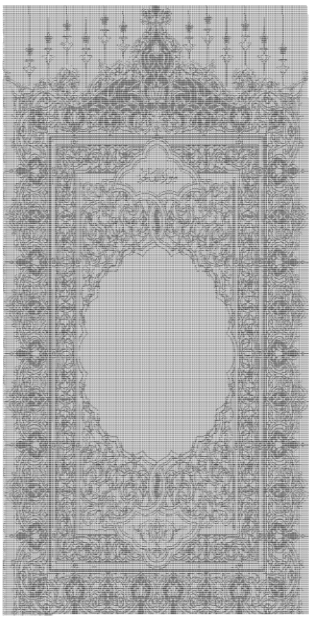


Image of Illumination from KSU - Electronic Mosshaf

Figure 1.1 Frame detection from Al-Quran pages.

iv. Philosophically, this project will be covered several domain and knowledge. This features focus on image processing and segmentation. It involve several techniques that will give impact to image analysis field for further process. By using technique segmentation and extraction, it using thresholding method and binarization method that already investigate.

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 result will be inverted to black and white colour pixels image because during the image analyses process, the image will covert to binary set. After process its will change to RGB (Red Green Blue) colour by pixels.

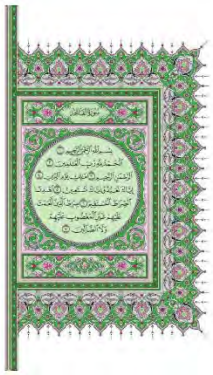
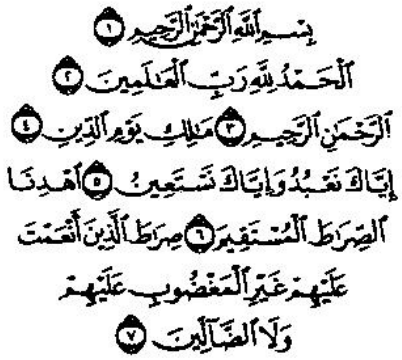
Input	Output
 Image Al-Quran Al-Karim from KSU - Electronic Mosshaf	 Image of text word from KSU - Electronic Mosshaf
 Image Al-Quran Al-Karim from Mushaf Al-Madinah Quran Majeed	 Image of text word from Mushaf Al-Madinah Quran Majeed
 Image Al-Quran Al-Karim from Mushaf Al-Madinah Quran Majeed	 Image of text word from Mushaf Al-Madinah Quran Majeed

Figure 1.2 Expected output removing Al-Quran illuminations frame.

1.7 Conclusion

This project will covered theory of image processing using image segmentation method to produced specific result that going to be used to fulfil project's objective. The concept of image segmentation is used to find the thresholding to turn a gray-scale image into a binary image. Meanwhile, the other concept is being apply to find the color similarity filter between 5x5 neighborhoods to find the point of marker. This is call the tracing techniques for detecting Al-Quran illumination frame. This techniques will help to extract Al-Quran page to only text word without its illuminations frame.

The purpose for this project is for produced more meaningful image for easier to analyses without distracting its' illumination. The result image can further investigate and process by analyses expert. This project can contributes huge impact for image processing domain specifically.

The representation of Al-Quran pages will be change due to decomposition process and will be produces expected result of Al-Quran text word. The intention from this project is going to remove the illumination frame from pages.

Chapter 2 wills covers 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 of image segmentation are 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 unique and the beauty of the Al-Quran is just not only on the holly word of the Al-Quran, but the written of the Al-Quran will make it more beautiful with the illumination that come from variety of pattern, texture and decoration as shown in Figure 1.1. However, this research project will focusing on how to separate Al-Quran text from its illumination. Then, it will become a part of research study to eliminate Al-Quran illumination.

Al-Quran was written in an Arabic word, thus this research can be refer to the research domain on an Arabic handwritten. Besides that, research on Latin also can be reference to eliminate the illumination on the Al-Quran.

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. Image processing will manipulate and process the image to produce as an image result with more meaningful image. The focusing domain project research is more toward removing Al-Quran illumination. Al-Quran have different illumination of its pattern and decoration. Thus, the process need to be well enough controlled to extract only parts that need to be analysed further.

2.2.2 Existing System

This project is focusing on removing Al-Quran illumination. Thus, it has not recently develop but the features of word segmentation already exist and research by Khairuddin Omar (2000), Mazani Manf (2002), Remon Redika (2008), Anton Heryantp (2008), Moalla et al. (2006) and Aiolli et al. (1999) 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. This segmentation technique used to find the thresholding to convert a gray-scale image into a binary image.

2.3 Project Methodology

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

Research framework of methodology are divided into two phases which is investigation phase and execution phase.