UNIVERSITI MALAYSIA SARAWAK
UNIMAS

Faculty of Computer Science and Information Technology

IMAGE ENHANCEMENT FOR TEXT EXTRACTION ON MOBILE APPLICATION

Fiona Anak Kandou

Bachelor of Computer Science with Honours
(Multimedia Computing)
2015

IMAGE ENHAN[illegible] EXTRACTION ON [illegible] [illegible]ION

FIONA ANAK KANDOU

This project is submitted in partial fulfilment of the
requirements for the degree of
Bachelor of Computer Science with Honours
(Multimedia Computing)

Faculty of Computer Science and Information Technology
UNIVERSITI MALAYSIA SARAWAK

2015

UNIVERSITI MALAYSIA SARAWAK

THESIS STATUS ENDORSEMENT FORM

TITLE IMAGE ENHANCEMENT FOR TEXT EXTRACTION ON MOBILE APPLICATION

ACADEMIC SESSION: 2014/2015

FIONA ANAK KANDOU

(CAPITAL LETTERS)

hereby agree that this Thesis* shall be kept at the Centre for Academic Information Services, Universiti Malaysia Sarawak, subject to the following terms and conditions:

1. The Thesis is solely owned by Universiti Malaysia Sarawak
2. The Centre for Academic Information Services is given full rights to produce copies for educational purposes only
3. The Centre for Academic Information Services is given full rights to do digitization in order to develop local content database
4. The Centre for Academic Information Services is given full rights to produce copies of this Thesis as part of its exchange item program between Higher Learning Institutions [or for the purpose of interlibrary loan between HLI]
5. ** Please tick (√)

[ ]	CONFIDENTIAL	(Contains classified information bounded by the OFFICIAL SECRETS ACT 1972)
[ ]	RESTRICTED	(Contains restricted information as dictated by the body or organization where the research was conducted)
[√]	UNRESTRICTED	

Validated by

(AUTHOR'S SIGNATURE)

(SUPERVISOR'S SIGNATURE)

Permanent Address

No. 1F, Lorong RTM 6A/2,
Jalan Oya Lama,
96000 Sibu Sarawak

Date: 26/06/2015

Date: 26/06/2015

Note * Thesis refers to PhD, Master, and Bachelor Degree
** For Confidential or Restricted materials, please attach relevant documents from relevant organizations / authorities

DECLARATION

I hereby declare that this project is my original work. I have not copied from any other student's work or from any other sources except where due reference or acknowledgement is not made explicitly in the text, nor has any part had been written for me by another person.

.....

(FIONA ANAK KANDOU)

30227

JULY 2015

ACKNOWLEDGEMENT

First and foremost, I would like to express my greatest gratitude to my supervisor, Dr Jacey-Lynn Minoi, for her help and guidance in making it possible to complete my project to the best of my ability. I would also like to extend my appreciation to my examiner, Miss Emmy Dahliana Bt Hossain, for her criticism and suggestion that help me to improve my project. Finally, I would like to give my eternal gratitude to my family and friends who had help and support me in the duration of completing this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Images are captured using a mobile phone camera would create 2D image variants, such as uneven lighting, cluttered background and text-like object. Pre-processing of images aim to reduce the noise in the image. In this project, a mobile application on image enhancement is developed. The application uses a combination of Otsu's Threshold binarization and colour clustering method is proposed to separate the text from the background. Text in images often has a higher intensity or lower intensity than the background. Otsu's binarization is used to convert the image to binary image for a higher contrast between the text and background. However, Otsu's binarization cannot distinguish between different colour pixels that have the similar illumination. Colour clustering is used in conjunction to binarization method as it is able to differentiate the different colour pixels hence eliminating the problem of Otsu's binarization method.

ABSTRAK

Imej yang diambil menggunakan kamera telefon bimbit akan mewujudkan variasi imej 2D, seperti pencahayaan yang tidak sekata, latar belakang yang berselerak dan objek yang berasaskan teks. Pra-pemprosesan imej adalah bertujuan untuk mengurangkan ganguan dalam imej. Dalam projek ini, satu aplikasi mudah alih untuk peningkatan imej telah dicipta. Kaedah yang menggunakan gabungan Otsu's binarization *dan* Colour clustering *telah dicadangkan untuk mengasingkan teks dari latar belakang imej. Teks dalam imej sering mempunyai intensiti yang lebih tinggi atau lebih rendah daripada intensiti latar belakang.* Otsu's binarization *digunakan untuk menukarkan imej kepada imej binari untuk menghasilkan kontras yang lebih tinggi antara teks dan latar belakang imej. Walau bagaimanapun,* Otsu's binarization *tidak dapat membezakan antara piksel berwarna yang berbeza dan yang mempunyai pencahayaan yang sama.* Colour clustering *adalah digunakan bersama dengan kaedah* Otsu's binarization *kerana kaedah ini dapat membezakan piksel berwarna yang berbeza dan juga dapat mengatasi masalah yang dihadapi oleh kaedah* Otsu's binarization.

CHAPTER 1: INTRODUCTION

1.1 Introduction

Text in images plays an important role in describing the content of the image. The text provides semantic meanings to the object in the images as well as information about the environment. If this text can be automatically detected from a mobile phone camera, many useful applications can be provided for users. An example would be an application that translates the extracted text in a foreign language into a text that the user understands.

Text Extraction from an image is the process of extracting the relevant text information from a collection of images. It plays a major role in finding vital and meaningful information. Text extraction involves detection, localization, tracking, binarization, extraction, enhancement and recognition of the text from the captured image. Due to the rapid growth of available multimedia documents and growing requirement for information, identification, indexing and retrieval, text extraction from image using mobile phone cameras had gain many interest from researchers (Lee et. al., 2009). On the other hand, these text characters are difficult to be detected and recognized due to their differences in size, font type, orientation, as well as complex coloured and cluttered background.

1.2 Problem Statements

Mobile phones these days have built-in camera application pre-set to high-resolution. Additional mobile applications to shoot and edit captured photos to full resolution are also becoming easily available. Most of these applications are used to improve the quality of the captured images by using photo editing features such as correcting the brightness information of the image. However, editing features works on the whole images. Currently, there is only a

limited number of mobile application that is designed to remove either background or objects, such as text and people, which are usually done manually by incorporating Photoshop.

At present, there is not yet an application that allows removal of objects done automatically. Until today, it is still a challenge, in image processing, to extract text from the background on a captured image using mobile application. Images that are captured using a mobile phone camera would create image variants such as uneven lighting, distorted perspective, cluttered background and text-like object. Issues, such as separating the text from a cluttered with many objects and coloured background, and separating colourful text on plain background, and a bit of both of the prior mentioned, still needs to be addressed.

Hence, a mobile application that could enhance the captured image by creating a higher contrast between the text and the background and enhancing the text from the background for further process is proposed in this project. The developed image enhancement mobile application will be able to enhance colours on text and background in a dual tone images that either have darker colour text and brighter background or lighter colour text with dark background using a hybrid technique that combines Otsu's binarization and colour clustering method.

1.3 Objectives

The followings are the objectives of the project:

- To develop a mobile application that could capture an image, whereby the foreground objects are text and then extract the text from the captured image.
- To extract the text from the background by enhancing the image using a hybrid technique that combine binarization and colour clustering method.

1.4 Scope

The project is limited to the listed followings:

- It is to be developed for mobile phone that is running on Android platform.
- It must use a mobile phone that has a built-in camera.
- The images to be enhanced are limited to text-based images.
- The images are limited to dual colour tone with plain background.

1.5 Brief Methodology

The methodology used in this project is the Incremental Model. The Incremental Model is a method of software development where the model is designed, implemented and tested incrementally until the product is finished. The phases involved are requirement, design, implementation and unit testing, integration and system testing and operation and maintenance as shown in Figure 1.1.

(i) Phase 1: Requirement.

At this phase, all possible requirements of the system to be developed are gathered and analysed.

(ii) Phase 2: Design.

For the design phase, the requirement specifications from first phase are studied and the system design is prepared. Design helps in specifying hardware and system requirements and also helps in defining overall system architecture of the project.

(iii) Phase 3: Implementation and Unit Testing.

With inputs from system design, this phase sees to it that the system is first developed in small programs called units, which are integrated in the next phase. Each unit is developed

and tested for its functionality which is referred to as Unit Testing. Iterations are done for the coming units.

(iv) Phase 4: Integration and System Testing.

In this phase, all the units developed in the implementation phase are integrated into a system after testing of each unit. After integration, the entire system is tested for any faults and failures.

(v) Phase 5: Maintenance and Operation.

This phase involves making changes in the system's software, hardware and documentation to support its operational effectiveness. It includes changes to improve the system's performance in the future.

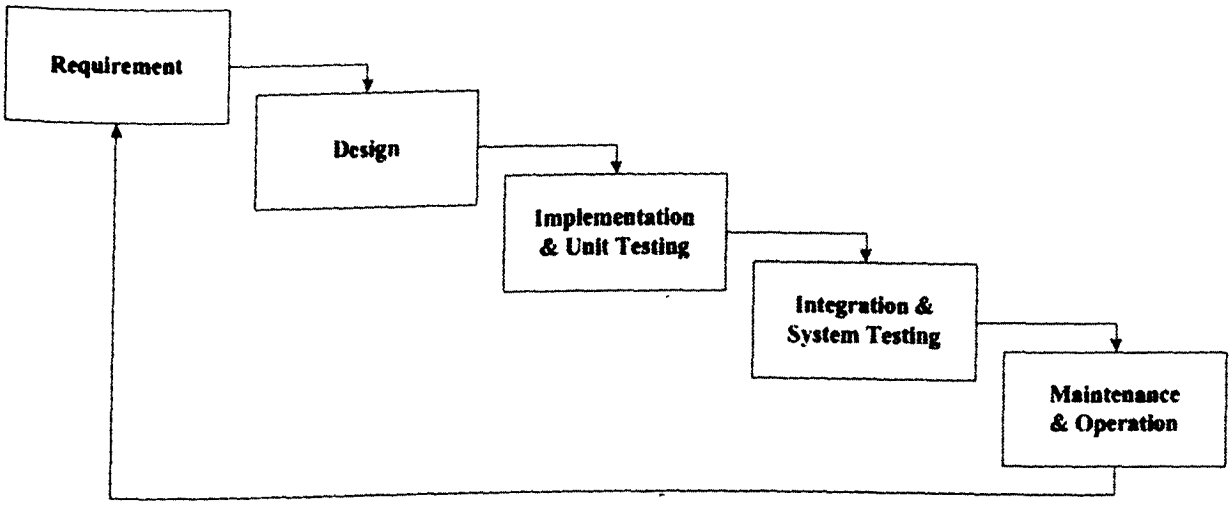

Figure 1.1: Incremental Model Diagram

1.6 Significance of Project

The significant of this application is that it can be used to enhance the text in the image from the background so that the text in the image is clearer. This project would open

huge opportunities for future projects whereby the text can be extracted and processed for further action such as text processing, text recognition, interpretation and/or text translation.

1.7 Expected Outcome

The outcome of the project is an image processing mobile application that runs on any Android smart devices to enhance the image by extracting the text from the background using image enhancement technique that combined Otsu's binarization and colour clustering method.

1.8 Project Schedule

It is a 2 semester project that runs for about 8 months. The detail of the schedule can be found in Appendix A.

1.9 Project Report Outline

This project is developed following the listed project outline.

Chapter 1: Introduction

This chapter describes the overview of the project. The sections that are included in this chapter are problem statement, objectives, scope or limitation, methodology used, significant of the project, expected outcome and project schedule.

Chapter 2: Literature Review

This chapter reviews on the current technique and method that is design and proposed by others on text extraction techniques. In this chapter, the details of techniques used to

extract text from the background of an image are identified and reviewed. The weakness and strength of each technique are analysed and listed out. This chapter also includes the proposed method to enhance the image for text extraction.

Chapter 3: Methodology

This chapter described the analysis of the requirement needed in the system. The method used in developing the system is defined. The hardware and software used are defined and described in detail.

Chapter 4: Implementation

This chapter describes the development of the application. Design phase and implementation phase is defined in detail. Each function will be tested to make sure it works properly. The application is then taking to the evaluate phase.

Chapter 5: Testing

This chapter covers the application testing. It is a high level software testing process in which testers verify that all related functions can operate in the same environment. The testing process ensures that all subcomponents are integrated successfully to provide expected results.

Chapter 6: Conclusion and Future Works

This chapter summarizes the finding from the development of the system and outlines the potential future work to enhance the system. All the constraints faced when implementing this project is listed. The project achievements and weaknesses are also stated.

1.10 Conclusion

The overview of this project is explained in this chapter. The project is scheduled according to the methodology used, which will be detailed later in Chapter 3. The objectives are stated clearly, which are based on the problem statements given.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter focuses on reviewing the current methods and techniques of image enhancement for text extraction from an image that have been proposed. It also listed the strengths and weaknesses of each approach. These methods are compared with each other to come up with a proposed new method to enhance image in order to extract text.

2.2 Text Extraction

Text extraction of an image is a complicated task because every image has a complex background, uneven lighting and different style of fonts. Images captured using mobile phone camera includes many unwanted environmental effects such as uneven illumination and perspective distortion. The image is also affected by image noise. Image noise is a random variation of brightness or colour information in images that is not present with the object itself, and is usually an aspect of electronic noise. It is an undesirable by-product of image capture that adds unnecessary information in the image.

Figure 2.1 shows the process of text extraction. Images with texts are the input whereby the text of the images will be extracted. The image will first undergo pre-processing. After the pre-processing process, it is possible to recognise the characters of the text. The text is detected in the image before it is recognised by the character of the text. This will ensure that it is possible to extract the information data of the text for the purpose of identification, indexing and information retrieval. The result of text extraction is the information of the image.

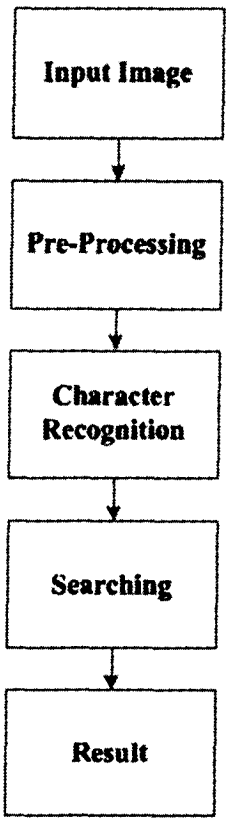

Figure 2.1: Text Extraction Process

2.3 Image Enhancement

Image enhancement is concerned with pre-processing the captured image so that it is optimized for text extraction. Pre-processing of an image aims to reduce the noise in the image. Pre-processing of an image is necessary as it will lead to successful binarization. For the purpose of extracting the text from the image, the input image must be converted to a binary image. By binarizing the image, it is possible to recognize the text in the image.

2.4 Review of Text Enhancement Techniques

This section reviews the text enhancement techniques that others had proposed. It also describes the strength and weakness of each technique.

2.4.1 Image Thresholding

Thresholding is the simplest of the image processing method. Threshold can be used to create a binary image. Binary image are made of pixel with only two possible values, a 0 or 1. In text extraction, the 0 value represents the background pixel of the image and the 1 value represents the foreground pixel of the image.

Most binarization techniques followed the architecture in Figure 2.2. Pre-processing of the image is necessary to reduce noise as it will lead to successful binarization. Binarization techniques can be divided into Global Thresholding and Local Thresholding (Bawa & Sethi, 2012).

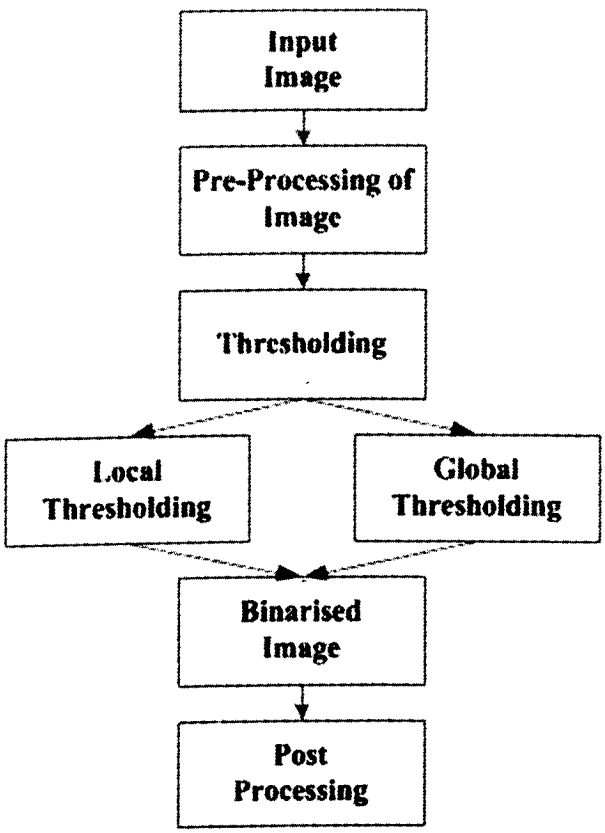

Figure 2.2: Framework for Binarizing Image

Global thresholding techniques process the whole of the image. These techniques work well when the input of the image has a bimodal histogram in which the grey level of the histogram has two equivalent peaks. In Global thresholding, a single value of threshold for the whole image is found. The output for the Global thresholding is a binary image. As these techniques produced strong peaks in the grey-scale histogram, the techniques are suitable for images with even illumination (Bawa & Sethi, 2012).

On the other hand, Local thresholding techniques are composed of local or adaptive. Accordingly, the image input is divided into a number of fixed-sized sub-images by choosing an appropriate window size. For each sub-image, the threshold is computed by using the information in the individual sub-image. Each of the sub-images is binarized independently. Local thresholding techniques are more suitable for images with uneven lighting as each of the sub-images producing their own simpler histogram (Bawa & Sethi, 2012).

Strengths and Weaknesses

The strength of thresholding method is that this method is able to convert grey-scale image to binary image for successful text extraction from the background. Regardless of whether the image have coloured background or not, the threshold method is able to convert it into binary image. Global thresholding is capable of handling images that have uniform illumination while Local thresholding is able to manage images that have uneven illumination.

The weakness of thresholding is that Local thresholding produces more noise in the non-text area of the image. Local thresholding also has a weakness in which it is difficult to choose proper window size. Thus, the thresholding will take a long time to process the image.

2.4.2 Interactive Differential Evolution

Interactive Differential Evolution (IDE) is a proposed automatic image enhancement tool for smartphone. This concept utilises the vector difference. It enhanced the colour of the images that subjectively give the output image a better looks compared to the original. This is done by changing the intensity of the image pixel (Lee & Cho, 2012).

An image usually consists of vectors. It contains information such as brightness, contrast, gamma, and colour. IDE represents all individual as NP dimensional real-valued parameter vectors:

$$x_{i,G}, i = 1,2,3, \dots, NP$$

where i is individuals and G is generations.

Trial vector $m_{i,G}$ is generated by mutation and crossover operations. Three individual members are randomly chosen and the difference of two vectors is multiplied by mutation control constant. The output is added to the other individual to generate the mutant vector. The calculation for a mutant operation is indicated as follows:

$$m_{i,G} = x_{r1,G} + F \cdot (x_{r2,G} - x_{r3,G})$$

where $x_{r1,G}$, $x_{r2,G}$, $x_{r3,G}$ are the randomly generated individuals and mutually different. $m_{i,G}$ is the mutant vector in generation G. F is the mutation control constant to suppress the amplification.

After the mutant filter is generated, crossover operation is conducted. The equation is as follows: