

**Mobile Data Capture and Transmission
(Conversion Part)
Enhanced**

By

Nafizah Binti Abdul Talib

**Dissertation Report in partial fulfillment of
the requirement for the
Bachelor of Technology (Hons)
Information Communication Technology**

JANUARY 2006

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan

t

TA

1640

.N146

2006

1) Optical character recognition devices

2) IT/IS -- there's

CERTIFICATION OF APPROVAL

**Mobile Data Capture and Transmission
(Conversion Part)
Enhanced**

By

Nafizah Binti Abdul Talib

A project dissertation submitted to the
Information Communication Technology Programme
Universiti Teknologi PETRONAS
in partial fulfillment of the requirement for the
BACHELOR OF TECHNOLOGY (Hons)
(INFORMATION COMMUNICATION TECHNOLOGY)

Approved by,

(Mr. Low Tang Jung)

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
January 2006

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons



NAFIZAH BINTI ABDUL TALIB

ACKNOWLEDGEMENT

The author would like to express her heartiest gratitude, first and foremost to Allah, The Al-Mighty, for giving the strength and ability to carry out all the research and task required. Also the author wishes to convey the highest appreciation to the project supervisor, Mr. Low Tang Jung for his time in consulting and facilitating throughout the completion of the final year project. In addition, the author would like to thank the authors of previous research papers whom had greatly craft a mind opening thesis. Thank you for friends who provide the ideas and their piece of mind for the project. Special thanks to the previous developer for this project, Mohd Khalid Bin Ka'ab by giving the guidance and for all of his effort, hard work and contribution put together towards this project. Last but not least, the author wishes to dedicate her work to her beloved family and colleges. Without their support, she could never have completed this project.

Thank you.

ABSTRACT

The Mobile Data Capture and Transmission (Conversion Part) project is about turning a mobile phone into a portable document scanner. Focus more on converting images captured by camera phone to personal computer using OCR engine. OCR is an acronym for Optical Character Recognition that functions to convert picture text image into editable wording character. Users then would be able to edit the texts converted afterward. The main objective of the project is to enhance the previous OCR program in term of functionality and effectiveness. Therefore, the project should be able to produce a program that is more reliable and errorless. Besides, the program will hopefully be able to perform pre-image processing before being past to the OCR engine for conversion. Thus, users can use any type of mobile camera to scan the images that required and not limited to merely 1.3 mega pixels and above. This final year project involves researches on optical character recognition (OCR) concept as it is known as the best way to do the conversion process and the implementation of image processing techniques in term of pattern recognition. The enhancement of the conversion part of this project is also involved directly with transmission part where it concentrates more on enhancing captured images. The spiral methodology is being used in the completeness of this project. In the end, the final product of this project may perhaps perform a better function than the previous one. Therefore it will absolutely give more help and benefits to various groups of people to reduce time needed to complete their works.

CHAPTER 4:	RESULT AND DISCUSSION	18
	4.1 Conversion Process System	18
	4.2 Future Recommendation for Conversion Process System							23
	4.3 Project Findings	23
CHAPTER 5:	CONCLUSION	24
REFERENCES		25

LIST OF DIAGRAM

Diagram 1: The Spiral Model

Diagram 2: Data model of the Conversion part

Diagram 3: Context Level Diagram

LIST OF FIGURES

Figure 1: Conversion Process page

Figure 2: Processing Options page

Figure 3: Adjust Dot per Inch page

Figure 4: Filter Option page

Figure 5: Picture Region Function

Figure 6: Filters Functions

Figure 7: Result page

CHAPTER 1

INTRODUCTION

1.1 Background of Project

The project is to make use of scanners to scan hardcopy documents and convert them into editable text using optical character recognition (OCR) programs. Instead of retyping text, scanner users can quickly convert documents into computer editable form.

The idea is about turning a mobile phone into scanner where user could discreetly photograph contracts, documentation or other written information while on business trip such as working while traveling or in an important conference which later can convert them into a format for processing that allows user to edit before be printed out as a hard copy or stored in the computer for future references.

The project used optical character recognition (OCR) concept as it is known as the best way to do the conversion process whereby any captured image can be turned into an electronic character based file. The process involves photo-scanning of the text character-by-character, analysis of the scanned-in image, and then convert the character image into character codes, typically in ASCII. In OCR processing, the page image is scanned, then analyzed for light and dark areas which is known as the black and white area in the binary image in order to identify each alphabetic letter or numeric digit.

1.2 Problem Statement

1.2.1 Problem Identification

The process of converting a captured image into editable text accurately is not an easy as it seems. Many documents that can easily read cannot be accurately converted into editable text by an ordinary OCR program especially images that taken with a mobile camera. Usually, in the market today a mobile camera comes with standard VGA resolutions camera that captured images with less than 1 mega pixel which is only suitable for messaging or web use. Printouts taken from such images are extremely blurry. In the previous project, the OCR program can read and convert any images 1.3 mega pixels and above. For example, several of the characters may not be converted accurately despite the fact that the text is easy to read for the human eye. Current OCR pattern-recognition algorithms still require the input characters to be well-formed, smooth, and large enough so that the individual character elements are very distinct. OCR cannot recognize characters as human do because human have the ability to look subconsciously at different facets of what they are reading and use multilevel thinking to figure out and recognize words. The program also ignores any other images other than text to be recognized.

1.2.2 Significance of the Project

On the whole, the completeness of this project would be able to cater needs for several groups of users such as researchers, reporter, journalists and academicians. Especially, when they are traveling or during business trips that may require them to copy any important information from journals, text speech, notes and so on for future references. Therefore, the Mobile Data Capture and Transmission project may offer new alternative

to the users as a simple application in which can be used and portable enough to be carried around such as mobile phones and PDA that act as a simple scanner. In addition Enhanced Mobile Data Capture and Transmission project may be able to perform better and provide more accurate OCR reading.

1.3 Objectives and Scope of Study

1.3.1 Objectives

1. To enhance the program in term of image availability and preservation as for now the application only be able to read text image only and neglect images other than text.
2. To make image lower than 1.3 mega pixels readable by the program such as images taken by VGA camera (0.3 mega pixels).
3. To apply image enhancement filter to the program so it can sharpened the image and perform better in term of readability.
4. To make the program be able to read format other than .bmp and .tiff.

1.3.2 Scope of Study

On the whole, the completeness of this project would be able to cater the problems that arise from existing OCR program especially in the pre-processing stage of the OCR reading. The research focus on the image processing techniques that can be applied to help increase the application performance and provide a more accurate reading without interfere the OCR engine. This means that the image enhancement is to be performed before the OCR reading.

CHAPTER 2

LITERATURE REVIEW

There are many concepts that can be used to develop an optical character recognition (OCR) application. OCR is most known as a technology that lets people to scan a printed page and convert it into text document that can be edited in any word processor.

Furthermore, OCR also known as a system that can provide a full alphanumeric recognition of printed or handwritten characters at electronic speed by simply scans the form. Basically, the OCR basic functions should allow forms that are containing characters images can be scanned through scanner and then recognition engine of the OCR system interpret the images and turn images of handwritten or printed characters into ASCII data (machine-readable characters). Therefore, OCR allows users to quickly automate data capture from forms, eliminate keystrokes to reduce data entry costs and still maintain the high level of accuracy required in forms processing applications. The OCR technology provides a complete form processing and documents capture solution. Usually, OCR uses a modular architecture that is open, scaleable and workflow controlled. It includes forms definition, scanning, *image pre-processing*, and recognition capabilities.

In the research, developer managed to find out on how to turn a mobile phone into a portable scanner. According to Christopher Dance (2004) [1], there are four (4) step processes to be followed in turning mobile phone into a scanner: (1) hold camera roughly one foot away from the document and photograph it, (2) fix blurriness and convert it to black and white, (3) adjust contrast and eliminate shadows or reflections caused by document surfaces, and (4) compress the captured image using a format

known as Fax Group 4 or G4. Those steps can be considered as important tips for making the project of Mobile Data Capture and Transmission to succeed with achieving all the intended objectives. Fundamentally, all the given steps have its' own strength and weaknesses. Thus, developer should extend the research from that point in order to find out what are the real steps could be applied for transforming camera phone into a scanner.

From that the developer had find out that the crucial part is in the *pre-processing* after images had been captured and scan before proceeds to the text recognition process. In this case, low resolution camera produces blurry images. This is where image processing appears to eradicate the problem arises. The image processing techniques can be used to perform the task. One of the issues in image processing is the detection of images edges. The latter usually refers to rapid changes in some physical properties, such as geometry, illumination, and reflectivity.

According to G. W. Wei and Y. Q. Jia [3]

“The edge detection process serves to simplify the analysis of images by dramatically reducing amount of data to be processed, meanwhile preserving useful structural information about the object.”

The standard edge detectors can be used to achieve the output. The Sobel and Prewitt detectors are finite-difference-based first-derivative operators, which pick up high-frequency responses at image edges. Moreover Canny has formulated edge detection as an optimization problem, and the Canny detector can be effectively approximated by the first derivative of the Gaussian function. Regardless of the great success of traditional image processing techniques, the problem of quality edge detection with many real-world images remains unsolved due to large amount of textures.

In return, the developer should deal first with the OCR accuracy in converting any captured image, text and so on. According to Douglas A. Bicknese (1998) [2], OCR accuracy varies from page-to-page depending on a number of variables. Many of the original book and journal pages have become brittle, faded, and foxed, while others have been written on or otherwise marred. Any of these factors can cause the program to mistake one letter for another. Therefore, the project must have one mechanism or method that can handle the situation.

CHAPTER 3

METHODOLOGY

3.1 Procedure

The project will be developed based on spiral methodology. The spiral methodology reflects the relationship of tasks with rapid prototyping, increased parallelism, and concurrency in design and builds activities.

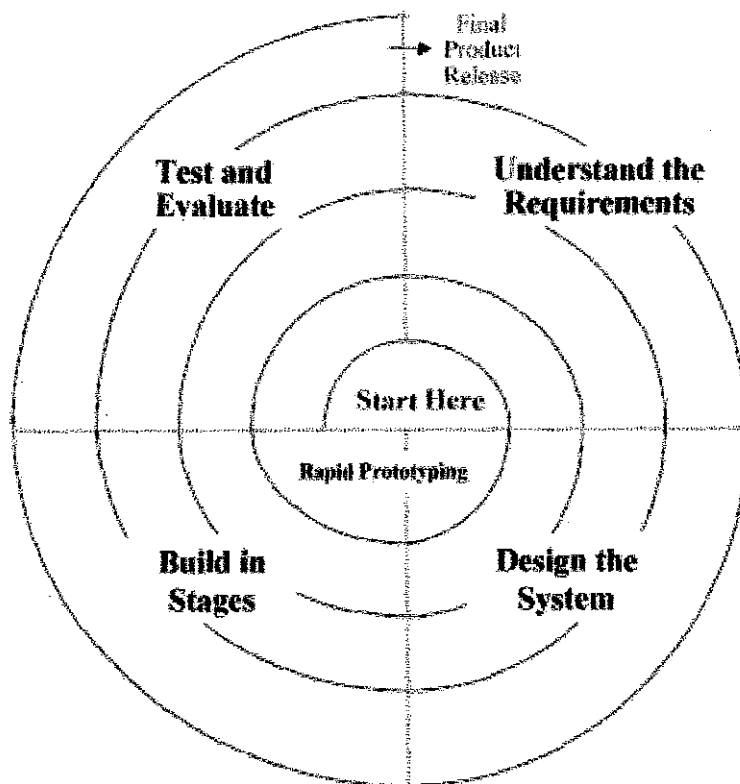


Diagram 1: The Spiral Model

3.2 The Spiral Methodology

3.2.1 Test and evaluate

Final product from previous project is treated as prototype. The prototype is tested to identify its strengths, weaknesses, and risks. Evaluate the result and recognize the improvement that should be made. The evaluation consists of three (3) main steps. The steps involved are: -

- *Test each features*: make a testing for each function or feature after the development for each of it has been completed.
- *Test each page*: integrate all the functions and features from the page and test the whole elements from start until finish.
- *Test the whole system*: integrate all the main pages and run a testing with using sample data.

3.2.2 Understand the requirement

From the test result objective of the project is derived. The requirements of the second prototype are defined in as much detail as possible.

The analyst approaches the users to find out what the needs of the new system. This is the most important phase of the methodology. Therefore, it is necessary to get accurate and correct information before starting a project. The tasks were: -

- Identified any relevant data regarding input and output that relevant to the project.
- Identified all the processes that involved in image processing.

- Identified process related to OCR reading.
- Identified possible project interfaces that needed to be upgraded.

3.2.2.1 Data Requirement

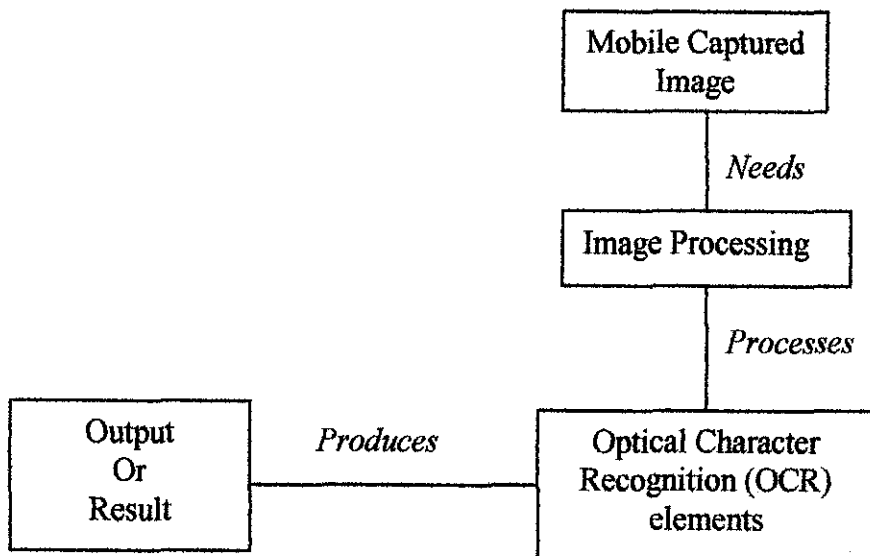


Diagram 2: Data model of the Conversion part

3.2.2.2 Process Requirement

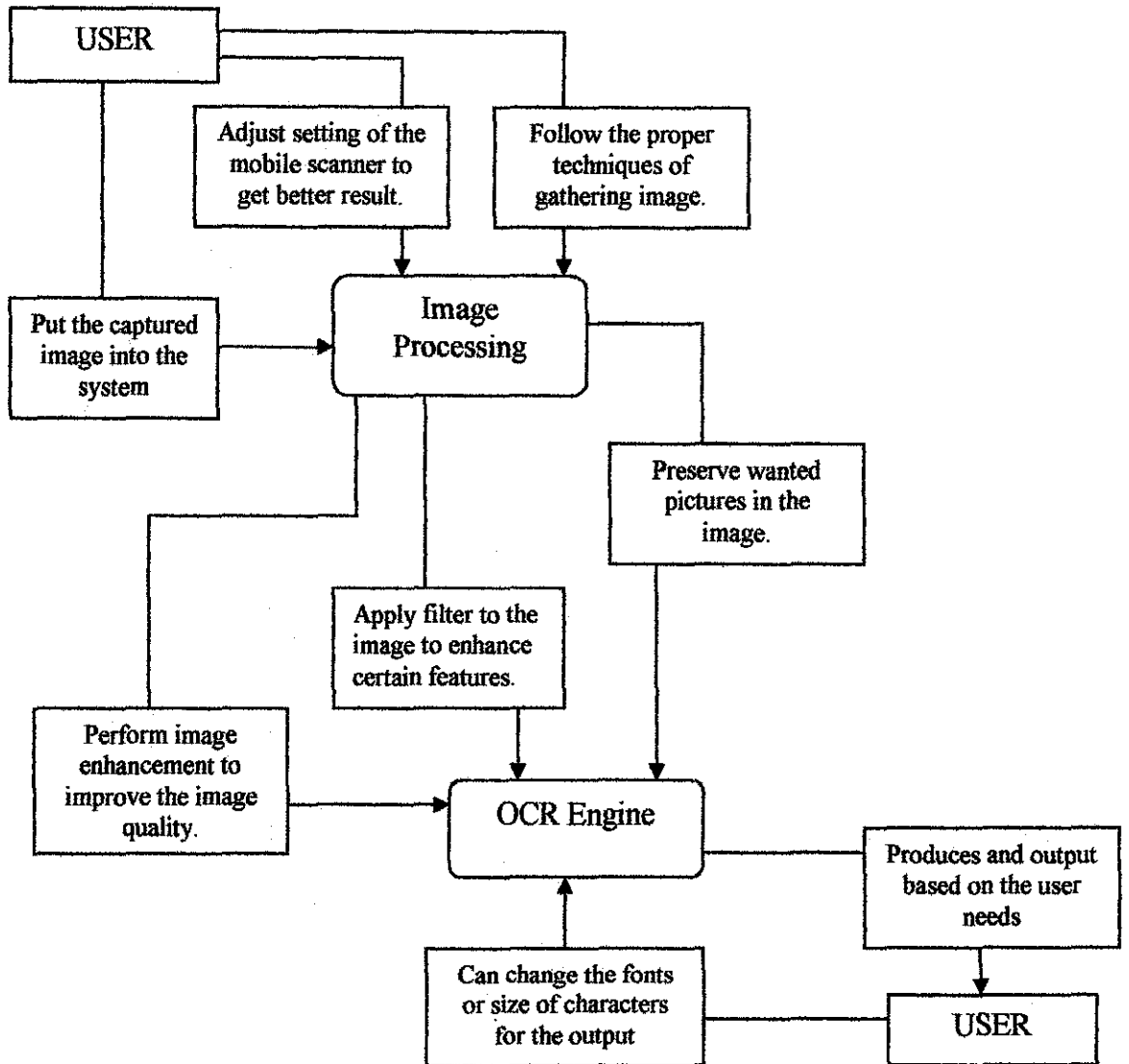


Diagram 3: Context Level Diagram

3.2.2.3 Interface Requirement

Based on both data and processes involved on this project, the interface requirement can be predicted to have at least four (4) interface designs. The interfaces can be listed below:-

- Conversion process page
- Processing options page
- Filter option page
- Adjust DPI (dots per inch) page

Each page will perform their function respectively in order to produce better image and provide better OCR reading for the Conversion part.

3.2.3 Design the system

Plan and design the second prototype based on the new requirement.

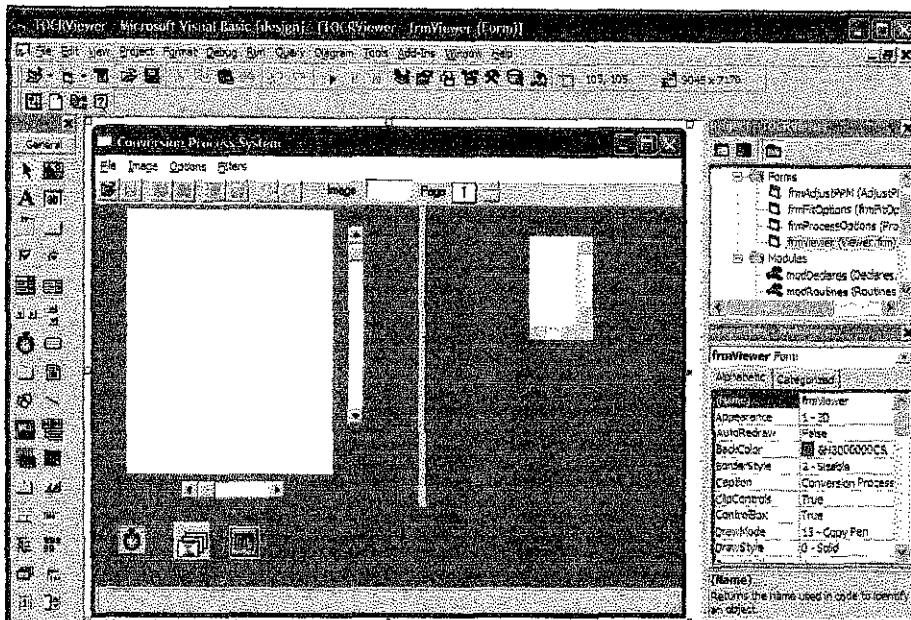


Figure 1: Conversion Process page

This is the main page in the system. Image that had been captured by the mobile device will be retrieve and process within this page. The output generated from the process can be saved to a word processing document to be edited later on. Some features have been upgraded in order to support the process of converting image. Such as a picture region button which the function is to preserve a region selected as picture in the output page.

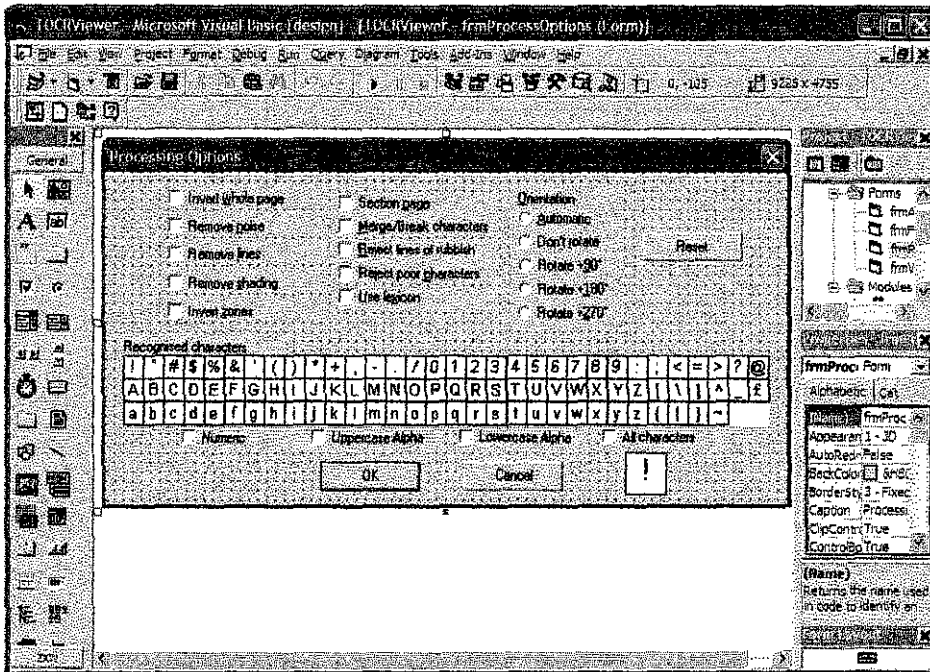


Figure 2: Processing Options page

This is the processing options page. It will isolate any category of recognized characters that users do not want to see when the system produces the output. Users can choose their selection of converted text by ticking in the appropriate boxes. As a result, the converted document will display no numeric as required by the users. In addition, users can also set aside any character from the 94 characters that can be seen on Figure 2. Users can simply right click on the character and it will not be appeared on the output. In other words, this page offers flexibility to the users in order to customize on whatever kind of output that they want it to be as a result.

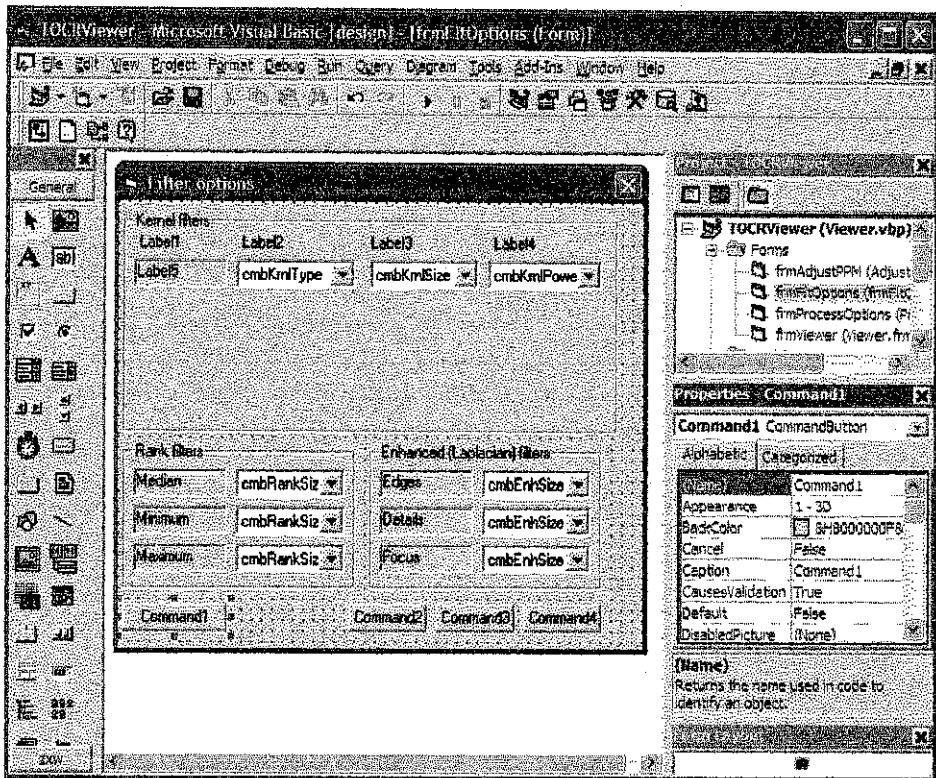


Figure 4: Filter Option page

In this page, user can set the option of filter that they want to apply to the image. There are three type of filter that can be adjusted to the user’s need which is Kernel filters, Rank filters and Enhanced filters.

3.2.4 Build in stage

It is a development stage whereby a new system is developed. This is usually a scaled-down system, and represents an approximation of the characteristics of the final product.

In this stage all the findings from the previous phases will be integrated and referred in order to start designing the source code for each feature and function for the system and as well as to accommodate with all project requirements. Several things have been done in order to ensure everything would be ready as expected. The procedures are: -

- Upgraded the three existing pages in the system to meet with the new project requirements which is to apply the filter, captured picture region and test the accuracy.
- Established connection between interfaces and the modules.
- Developed each function and feature in every page to meet with the requirements.
- Responded with all changes about the requirements for this project when some unnecessary situations occur.
- Repeat the whole procedure again to make sure all the system is working.

3.3 Tools and Equipment

The following are the tools suggested to be used during the development of the project. Project development is done using Microsoft® Visual Basic 6.0, image testing is in MATLAB 6.1 using the Image Processing Toolbox and Easy Graphic Converter is used during the testing stage whereby format conversion need to be done.

3.3.1 Software

- Microsoft® Visual Basic 6.0
- MATLAB 6.1 (Image Processing Toolbox)
- Easy Graphic Converter

3.3.2 Other Tools

- Hand phone or PDA with digital camera that supports MMS and any wireless facilities such as Infra Red and Bluetooth.

CHAPTER 4

RESULT AND FINDINGS

The result of the project is the usage of new technology or application, which will be known as smart mobile scanner version 2.0 after it has been successfully produced from Mobile Data Capture and Transmission project. Basic idea for this project is to turn digital camera phone into a portable scanner.

From researches and studies, there are ways to make the project become reality. It depends on developer to use and manipulate the concept as a guideline to start develop and achieve all the objectives stated in the project.

As a result, Optical Character Recognition (OCR) concept has been developed and being implemented on the project. The project is still on progress. For the project, optical character recognition (OCR) concept would play a major role on this project, particularly on conversion part. It is interesting to learn and know the OCR. People around the world are using the OCR concept or applications in order to make their products can perform multitasking jobs and saving their quality time.

4.1 Conversion Process System

As stated earlier, developer had to focus on the converting part for Mobile Data Capture and Transmission project. Therefore, the conversion process system should satisfy the requirements of the conversion part to the system's objectives. In other words, the system will process any image captured by camera phone. The process starts when the image is transferred to server either through MMS via email or using any wireless technology such as Infra Red or Blue Tooth.

Then, users may pick up the image and use this application to convert it into another format that can be edited and printed out.

From developer's point of view, the objectives help the system runs smoother and more accurate. The main objectives for the system can be listed below:-

1. To convert captured image into a new format which allow users to do some editing process and later print it out.
2. To allow users to customize the output to be produced by the system with using the Processing Options page.
3. To make desirable picture from the image be preserve in the output.
4. To apply image filter to the image to enhanced it for more accurate reading.
5. To automate the process of defining resolution on any captured image that may use the system for conversion purposes.
6. To allow users to change the fonts type or size in order to make the result looks clearer and tidy enough with perfect arrangement of words.

Previously, only text image will be converted. This is the main function of the OCR engine. In the enhanced project, despite concentrating on the accuracy of the OCR reading there are also an additional function added which is to make the pictures in the image be preserved in the output page to be saved later.

The procedure is click at the *picture region* button and selects the picture that wanted to be preserve in the output. Then perform the OCR reading. So, after the page is being read, the picture will be available in the output.

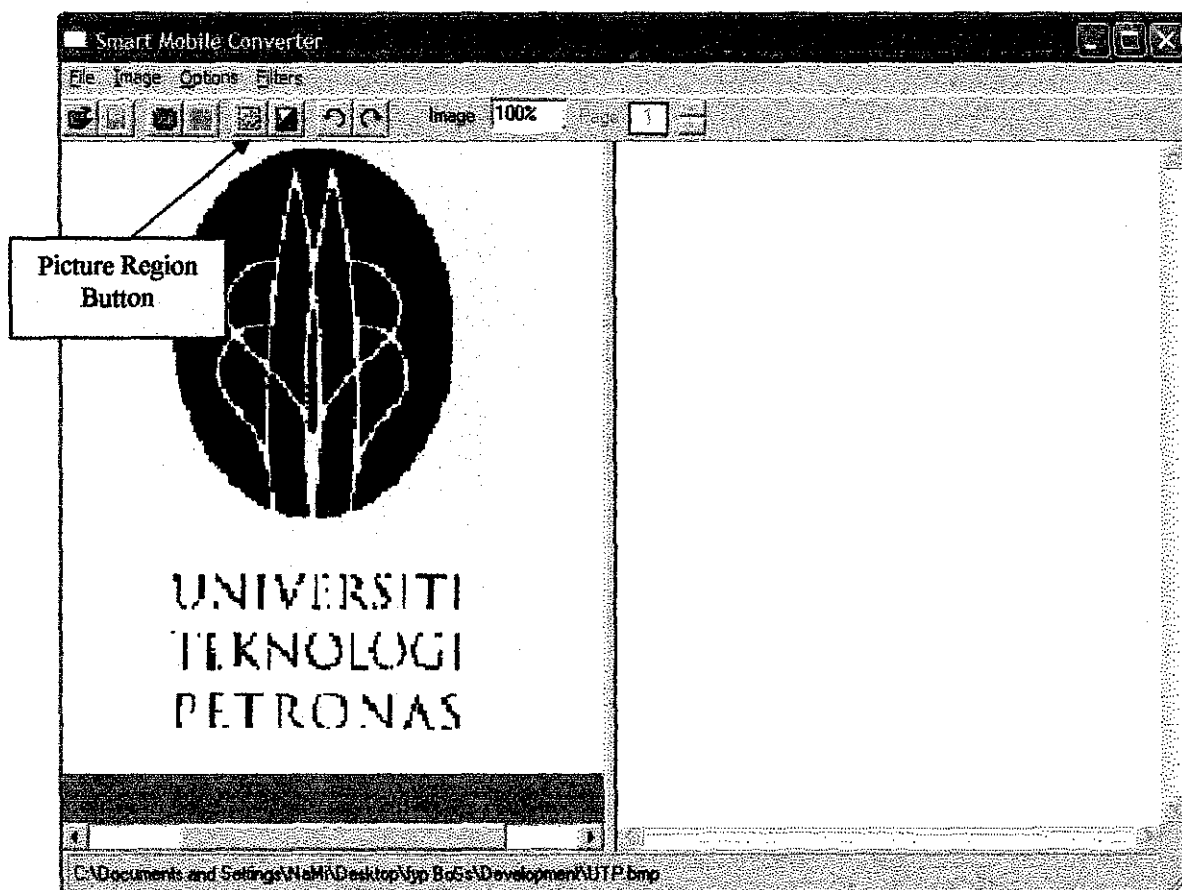


Figure 5: Picture Region Function

To make the OCR reading more accurate some filter might need to be applied to the image. This can be done by selecting the appropriate filter from the filters menu list. Click at the enhancement technique from the list and the result will be applied to the input image. Continue with the OCR reading after that. For more perfect conversion, user can play around with the filters functions before the page is being passed to the OCR engine for conversion process.

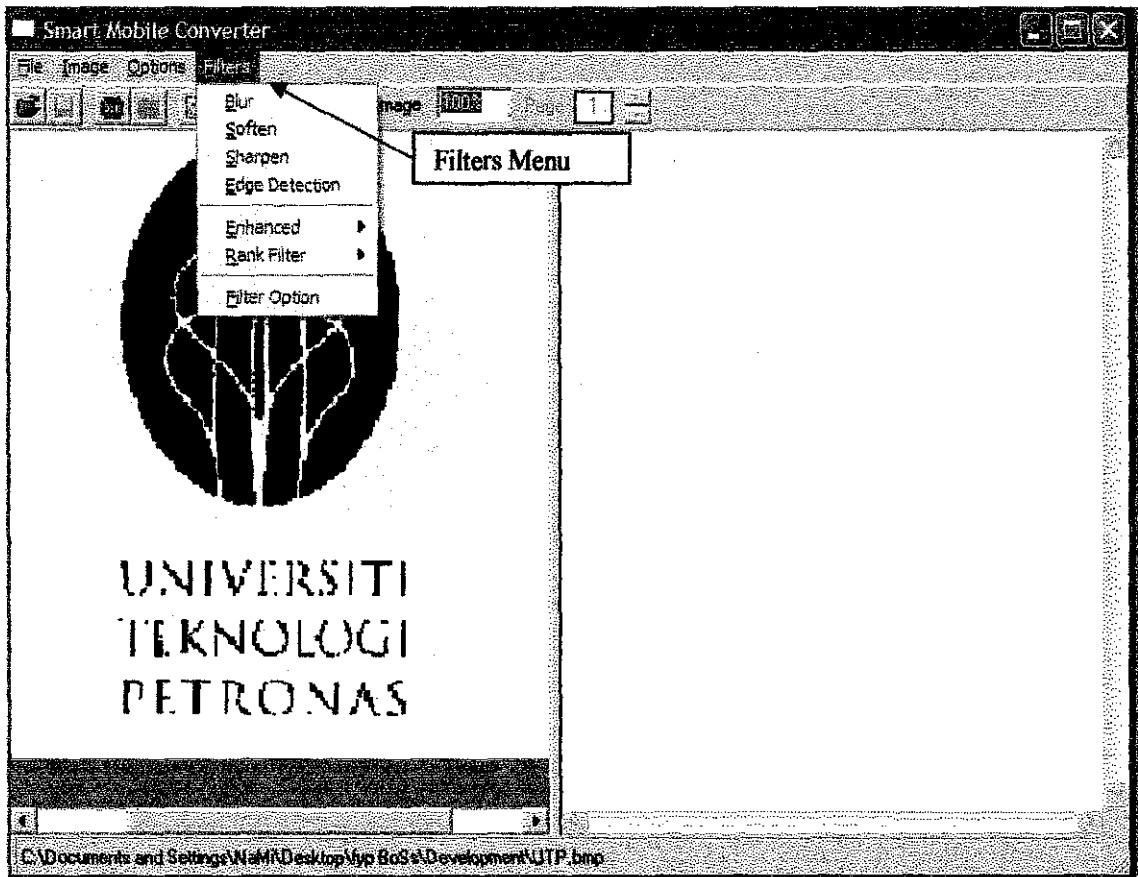


Figure 6: Filters Functions

Click on the OCR button to proceed with the OCR reading. The result can be seen as in *Figure 7*. User then can save the output result in the suitable word document to be used later.

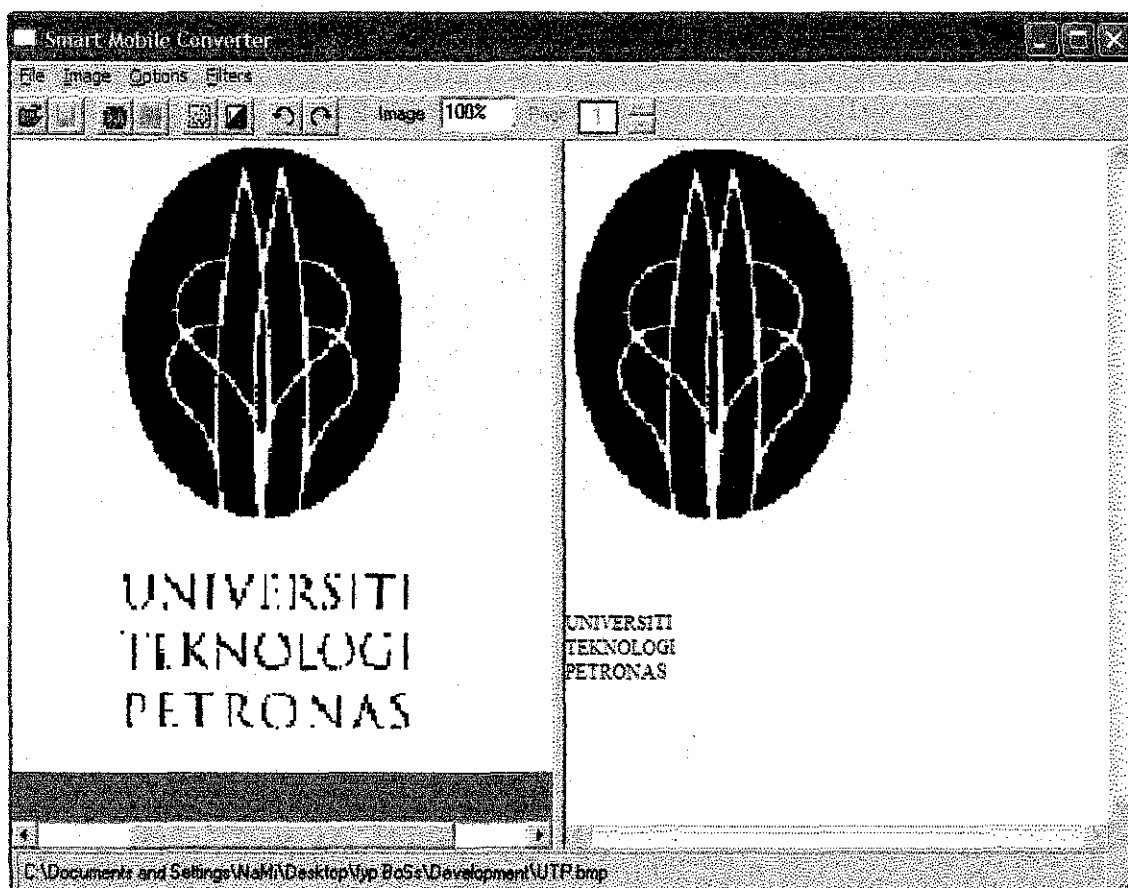


Figure 7: Result page

4.2 Future Recommendations for Conversion Process System

There are still additional features that can be added as future enhancement to the conversion process system. The enhancement can be:-

1. The formats that can be read on the system should be extended such as .png, and .ico.
2. On the result section, the system may apply dictionary concept to make the editing process easier for example including spelling check and other language characters check.
3. Provide reading for other characters as well as Arabic words, Chinese words, and Indian words.

4.3 Project Findings

As a whole, optical character recognition (OCR) concept would be the main implementation on this project, mostly on conversion part. Most of scanners are using the OCR concept or applications in order to make their products can perform multitasking jobs. For that reason, the Mobile Data Capture and Transmission project implement the OCR concept in order to convert the captured image into a format that can be edited and printed by the user. In addition to the previous project, image processing has been upgraded so that the OCR reading can produce a more accurate result. Hopefully, the outcome of this project can provide an alternative view to mobile camera.

CHAPTER 5

CONCLUSION

This project was developed to help groups of people that working like they are traveling where everything have to be portable and mobility is the issue. Information is the key aspect of the jobs and everything needs to be recorded for later used. Thus, with mobile camera phone all hassle can be avoided. Image of text document can be captured and transfer it to the laptop which contain the Conversion Process system, which is able to reduce time needed to complete the work.

From technical point of view, this project should be able provide a lot of benefits to various groups of users particularly in lessening their daily workloads. It is expected that the outcomes produces from the research and project could become a point of reference for other developers to make the application perform better in the future.

For the conversion part system alone, it has a lot to offer to the users with enhanced features and functions that will make the project more usable and reliable. However there are still rooms for improvement for future use that can be done. As for now, the project manages to keep up with the new project requirements and cater all the objectives.

REFERENCES

1. Christopher Dance, *Patented Xerox Technology Could Turn Camera Phone into Portable Scanner*, Retrieved November 17, 2004 from <http://www.phonecontent.com/bm/news/gnews/584.shtml>
2. Douglas A. Bicknese, 1998, *Measuring the Accuracy of the Optical Character Recognition*.
3. Jeffrey J. Tsay. 2000, *Visual Basic 6 Programming: Business Applications with a Design Perspective*, Prentice Hall, Upper Saddle River, New Jersey.
4. G. W. Wei and Y. Q. Jia , Synchronization-based image edge detection <http://www.math.msu.edu/~wei/PAPER/p81.pdf>
5. I.T. Young, J.J. Gerbrands, L.J. van Vliet, *Image Processing Fundamentals* <http://www.ph.tn.tudelft.nl/Courses/FIP/noframes/fip-Contents.html>
6. Rudolf K. Bock, 7 April 1998, *Image Enhancement*. <http://rkb.home.cern.ch/rkb/AN16pp/node127.html>
7. Clifford Watson, 1994, *An Image Processing Tutorial for Beginning Undergraduate Students*. <http://www.cs.washington.edu/research/metip/tutor/tutor.html>
8. Rafael C. Gonzalez, Richard E. Woods, 2002, *Digital Image Processing*. <http://ee.sharif.edu/~dip/Files/GonzalezSampleChapter.pdf>

9. Steve McMahon, 1998-1999, *VB Image Processing Sample*.
<http://www.vbaccelerator.com/codelib/gfx/imgproc.htm#sectop>

10. icameraphones.com, 2005, *Complete Info On Camera Phones*.
<http://www.icameraphones.com/buying-guide/camera-mobile-zoom.htm>