

COMMON PADDY DISEASE IDENTIFIER

TAN JOE WEE

A report submitted in partial fulfilment
of the requirements for the award
of the degree of
Bachelor of Computer Science (Software Engineering)

Faculty of Computer Systems & Software Engineering
University Malaysia Pahang

MAY 2011

| | |
|---|-------------------------------------|
| PERPUSTAKAAN 91 G UNIVERSITI MALAYSIA PAHANG | |
| No. Perolehan 068667 | No. Panggilan TA 1637 7365 |
| Tarikh 30 NOV 2012 | 2011 rs BC. |

ABSTRACT

Rice paddy is important to Malaysia in terms of economic, food supplement and even as a culture for some ethnic. Yet the paddy yield by Malaysia farmers has always been infected by the crops disease such as Blast and Brown Spots. LPP officers have to move a lot to help farmers in rural areas. Sometimes farmer that are rural are located far from help. The objective of the system (Common Paddy Disease Identifier) is to develop a mobile prototype system that can identify the disease found on the paddy plants. Extreme programming is use as an approach to develop the system. This is because it promotes rapid application development under constraint time. Image processing technique such as Otsu, Median Blur, Sharpen and Prosterise has been studied thoroughly to use it during required phases of image processing. The system develop could identified three different disease that are rice blast, rice sheath blight and brown spot disease. If the system would be able to develop it could help a lot of farmers where sometimes officers from LPP are located too far. Besides this it also can help promote understanding of plant disease for students who are interested.

ABSTRAK

Padi merupakan sesuatu bahan mentah yang penting bagi Negara Malaysia dari segi ekonomi, bahan makanan dan juga adat bagi sesetengah bangsa. Namun demikian, petani Malaysia yang menanam padi selalu dilanda oleh penyakit-penyakit padi tertentu seperti Blast dan Brown Spots. Tambahan lagi pengawai LPP terpaksa bergerak lang alik untuk menolong petani. Tambahan lagi sesetengah petani terletak jauh dari kawasan LPP. Objektif sistem (Common Paddy Disease Identifier) adalah untuk membangunkan sebuah system dimana system itu dapat menentukan jenis penyakit yang didapati dalam padi. “Extreme programming” digunakan sebagai salah satu kaedah metodologi dalam pembangunan system. Ini kerana iannya dapat membangunkan system dengan tangkas dalam masa yang dihadkan. Kaedah imej pemprosesan seperti Otsu, Median Blur, Sharpen and Prosterise telah pun dikajikan dengan teliti dan akan digunakan dalam fasa tertentu. Sistem yang dibangunkan akan dapat mengenali tiga jenis penyakit padi iaitu “ rice blast”, “rice sheath blight” and penyakit “brown spot. Jikalau system itu dapat direalisasikan iaanya dapat membawa banyak manfaat terutamanya kepada petani yang terletak jauh dari pengawai LPP. Selain itu, iannya dapat membawahkan manfaat kepada pelajar yang minat tentang penyakit padi.

TABLE OF CONTENTS

| CHAPTER | TITLE | PAGE |
|---------|--|------|
| | TITLE PAGE | |
| | SUPERVISOR DECLARATION | i |
| | STUDENT DECLARATION | ii |
| | DEDICATION | iii |
| | ACKNOWLEDGEMENT | iv |
| | ABSTRACT | v |
| | ABSTRAK | vi |
| | TABLE OF CONTENTS | vii |
| | LIST OF TABLES | viii |
| | LIST OF FIGURES | xii |
| | LIST OF ABBREVIATIONS | xiii |
| | LIST OF APPENDICES | xiv |
| 1 | INTRODUCTION | 1 |
| | 1.1 Introduction | 1 |
| | 1.2 Problem Statements | 3 |
| | 1.3 Objectives | 4 |
| | 1.4 Scope | 5 |
| | 1.5 Target User | 5 |
| | 1.6 Thesis Organization | 5 |
| 2 | LITERATURE REVIEW | 7 |
| | 2.1 Introduction | 7 |
| | 2.2 Image Processing Architecture and Design | 8 |
| | 2.3 Image Processing Techniques | 12 |
| | 2.3.1 Bilateral Median Filter | 12 |

| | |
|---|-----------|
| 2.3.2 Pattern Matching | 14 |
| 2.3.3 Otsu Method | 15 |
| 2.3.4 K-Means | 17 |
| 2.3.5 Sobel Operator | 19 |
| 2.3.6 Laplace | 20 |
| 2.3.7 Canny Edge Detection Algorithm | 21 |
| 2.4 Software Approach | 22 |
| 2.4.1 Operating System | 23 |
| 2.4.2 Platform Development Environment | 25 |
| 2.5 Software Methodologies | 27 |
| 2.5.1 Extreme programming | 27 |
| 2.5.2 Rapid Application Development | 28 |
| 2.5.3 Spiral Model | 30 |
| 2.6 Paddy Disease | 31 |
| 2.6.1 Brown Spot(Bintik Perang) | 32 |
| 2.6.2 Sheath Blight (Hawar Seludang) | 33 |
| 2.6.3 Rice Blast(Karah) | 35 |
| 2.7 Conclusion | 37 |
| 3 METHODOLOGY | 38 |
| 3.1 Introduction | 39 |
| 3.2 Extreme programming on EasySnap Development | 39 |
| 3.2.1 Architectural Spike | 41 |
| 3.2.2 Release Planning | 42 |
| 3.2.3 Iteration | 43 |
| 3.2.4 Acceptance Test | 43 |
| 3.2.5 Small Release | 44 |
| 3.3 Design and Architecture of EasySnap | 44 |
| 3.3.1 Image Acquisition | 45 |
| 3.3.2 Image Enhancement and Restoration | 48 |
| 3.3.3 Segmentation | 50 |
| 3.3.4 Representation and Description | 52 |
| 3.4 Software and Hardware | 53 |
| 3.5 EasySnap Program Architecture | 55 |
| 3.6 Conclusion | 56 |

| | | |
|----------|--|-----------|
| 4 | IMPLEMENTATION | 57 |
| | 4.1 Introduction | 58 |
| | 4.2 Implementation of the system base on | |
| | Extreme Programming | 58 |
| | 4.2.1 Architectural Spike | 58 |
| | 4.2.2 Release planning | 59 |
| | 4.2.3 Iteration | 60 |
| | 4.2.4 Acceptance test | 60 |
| | 4.2.5 Small release | 60 |
| | 4.3 Algorithm and code use to implement image | |
| | processing into the system | 61 |
| | 4.3.1 Blur | 61 |
| | 4.3.2 Brightness and Contrast | 63 |
| | 4.3.3 Posterise | 65 |
| | 4.3.4 Count Number of Pixel (Identify Blight) | 67 |
| | 4.3.5 Count Number of Pixel (BrownSpot) | 69 |
| | 4.4 Conclusion | 71 |
| 5 | RESULT AND DISCUSSION | 72 |
| | 5.1 Introduction | 72 |
| | 5.2 Result of the System | 73 |
| | 5.2.1 Interface of the System | 73 |
| | 5.2.3 Accuracy of the system | 77 |
| | 5.2.3 Result of Image Processing | 74 |
| | 5.3 Factor Affecting Accuracy of the System | 78 |
| | 5.3.1 Ideal Case versus Real Case | 78 |
| | 5.3.2 Correct Object Orientation | 79 |
| | 5.3.3 Different User Preference | 79 |
| | 5.3.4 Disease is based on specific color | 79 |
| | 5.4 Advantages and Disadvantages of the System | 80 |

| | | |
|----------|-----------------------------------|-----------|
| | 5.4.1 Usage of JavaScript | 80 |
| | 5.4.2 Usage of HTML5 | 80 |
| | 5.4.3 Save Cost for Data Transfer | 80 |
| | 5.4.4 Security Issues | 81 |
| | 5.4.5 Speed of the system | 81 |
| | 5.5 Conclusion | 81 |
| 6 | CONCLUSION | 82 |
| | 6.1 Conclusion | 82 |
| | 6.2 FutureWorks | 83 |
| | REFERENCES | 84 |
| | APPENDICES | 88 |

LIST OF TABLES

| TABLE NO. | TITLE | PAGE |
|--------------|---|------|
| 2.1 | Comparison of Operating System | 23 |
| 2.2 | Minimum requirement for Symbian application | 24 |
| 2.3 | Comparison table between selected smart phones | 25 |
| 2.4 | Summaries paddy disease and signs and symptoms | 36 |
| 3.1 | Comparison between Ideal Case and Real Case | 42 |
| 3.2 | Example of paddy applies using Gaussian | 49 |
| 3.3 | Example of paddy applies during segmentation stage | 50 |
| 3.4 | Example of paddy applies during representation and segmenetation stage | 52 |
| 3.5 | Software use on development stage | 53 |
| 3.6 | Hardware use on development stage | 54 |
| 3.7 | Hardware use during execution stage | 54 |
| 5.1 | Result of each image being process | 75 |
| 5.2 | Result of EasySnap | 78 |

LIST OF FIGURES

| FIGURE NO. | TITLE | PAGE |
|------------|---|------|
| 2.1 | Key Stages in Digital Image Processing | 8 |
| 2.2 | Schematic Overview of Paddy Disease Diagnosis System | 10 |
| 2.3 | Modularized architecture of the overall system | 11 |
| 2.4 | Comparison between Bilateral Median Filter and Adobe Photoshop Filter | 13 |
| 2.5 | A diagram on pattern matching between testing model and learned model | 14 |
| 2.6 | Image of paddy disease using Otsu method | 16 |
| 2.7 | Series of image Lena+ Gaussian noise and Lena+ Salt and Peper noise using Otsu and K-Means | 18 |
| 2.8 | Series of image Man + Gaussian noise and Man + Salt and Pepper noise using Otsu and K-Means | 18 |
| 2.9 | Image of Lena using Sobel Operator | 19 |
| 2.10 | Image of Lena using Laplace Operator | 20 |
| 2.11 | Image of Lena using Canny Edge Detection Algorithm | 22 |
| 2.12 | Example of Extreme Programming Model | 27 |
| 2.13 | Example of Rapid Application Development | 29 |
| 2.14 | Example of Spiral Model | 31 |
| 2.15 | Life cycle of the Brown spot disease | 33 |
| 2.16 | Life cycle of the Sheath Blast disease | 34 |
| 2.17 | Life cycle of the Sheath Blight disease | 35 |
| 3.1 | Example of Extreme Programming Model | 40 |
| 3.2 | Design and Architecture of EasySnap | 44 |
| 3.3 | Example of Performing Image Acquisition | 46 |
| 3.4 | Histogram of Original Image | 51 |
| 3.5 | Histogram of Posterise Image | 51 |

| | | |
|-----|--|----|
| 3.6 | Architecture of EasySnap program | 55 |
| 4.1 | Snippet Code for Blur | 62 |
| 4.2 | Snippet Code for Brightness and Contrast | 64 |
| 4.3 | Snippet Code for Posterise | 66 |
| 4.4 | Snippet Code for Identify Blight | 68 |
| 4.5 | Snippet Code for BrownSpot | 70 |
| 5.1 | Interface of the System | 75 |
| 5.2 | Output of the System(Brownspot) | 76 |
| 5.3 | Output of the System (Blast) | 76 |
| 5.4 | Output of the System(Unkwown) | 77 |

LIST OF ABBREVIATIONS

| | |
|------|---|
| UMP | Universiti Malaysia Pahang |
| GUI | Graphic User Interface |
| JS | Javascript |
| HTTP | Hypertext Transfer Protocol |
| LKPP | Lembaga Kemajuan Perusahaan Pertanian Negeri Pahang |
| RGB | Red Green Blue |
| MB | Megabyte |
| KB | Kilobyte |
| ES | EasySnap |
| SAD | System Analysis and Design |
| SDLC | Software Development Life Cycle |
| XP | Extreme Programming |
| TIFF | Tagged Image File Format |
| JPEG | Joint Photographic Expert Group |

LIST OF APPENDICES

| APPENDIX | TITLE | PAGE |
|-----------------|--------------|-------------|
| A | Gantt Chart | 89 |
| B | Test Cases | 90 |
| C | User Manual | 106 |

CHAPTER 1

INTRODUCTION

1.1 Introduction

Paddy or Padi which are known in Malaysia plays a very important role in our country agricultural economy and also our diet intake. Paddy that are harvest by machine is then being use to polish the rice until it is white in colour. Rice are then being grade by an automated machine. Rice is graded base on the percentage of broken rice. The higher percentage of broken rice in a distributed packet, the lower the grade of rice is. Unlike westerners where potatoes are their main diets, rice is a more common diet in our country and most of the Asian region. It is also our main source of carbohydrate intake.

Rice does not only play an economic role but also through religion and culture. For example colour rice is use by Hindu to decorate their kolam during the Tamil month of “Markazhi” which falls on

December to January. According to IRRI(International Rice Research Institute) website, Agriculture provide 12% of GDP in Malaysia. This proves that Malaysia has a potential market in rice paddy agricultural. The foreign media reported that the global price for rice had jumped to US\$700 (about RM2,240) per ton, three times more than the price five years ago.

As Malaysia is also importing rice, any hike in the commodity's global price would create impacts in the local market, even though the heat is yet to be felt.[1] The drastic increase of rice price prove that Malaysia is incapable to provide the rice market demand. This is partially the drawback of paddy rice plantation is that it took too long to gain harvest. Using modern technology rice paddy can only be harvested two to three times per year. Paddy such as this are the” Padi MR 232” and Paddy MR220 (Padi Mas) which were introduce by MARDI [2].

During the period of plantation to growth many factors will cause the rice paddy quality and quantity .For example the quality of the soil, weather and even rain. In favorable irrigated environments, for instance, sheath and stem diseases (especially sheath blight, sheath rot, and stem rot) are particularly important, together with weeds. In contrast, the array of important rice pests is very wide in the unfavorable rain fed-lowland environments, including blast (especially panicle blast), sheath rot, sheath blight, brown spot, several insects (stem borers in particular), and weeds.[3]. At the end it will affect the quality and yield result during harvesting period. Although paddy farmers are being trained by Malaysia Agricultural Research and Development

Institute (MARDI) experts to recognize paddy diseases in order to ensure that early prevention or treatment is taken, but errors might occur in order to determine the type of diseases and control. However, this method is

labor intensive and time consuming, and it is impossible to accurately estimate the infected areas and severity in large-scale farming.[4]

Training a person to be expert in recognizing paddy disease will require time, training and cost. At the end, an expert will still make mistakes. This is because it is still gauge by a human. To reduce this problem, software that uses image processing is develop. Computer vision as compare to human is more accurate. It do not make mistakes and can be improve by algorithm and method.

This software is called EasySnap. Sample of photos are taken from the rice paddy are then being uploaded into the software. The software would then scan the photo. At first image will be process to remove the noise produce by the camera or the sun rays. Image will then be process and identify the problem cause by the disease. It will then provide instruction on how to solve the problem. EasySnap would also provide a database for user to know more detail about the pest and disease. This will provide a better alternative method from the traditional method.

1.2 Problem Statement

The problem statements are listed out as below:

- i. If were to follow the traditional method. A paddy farmer is usually train at MARDI to be able to analyze the type of disease. If the paddy farmer is unable or fail in analyzing, the paddy farmer would contact a specialist or an expert in that field to help identify the type of disease. Once the disease been identify the expert will then advice the paddy farmer on methods of treatment or even prescribe toxic to kill the disease. This shows that it rely on human judgment and evaluation. It might sometimes be wrong. A good expert might be rare and take many years of experience.

- ii. An expert usually is very well knowledgeable in that field. Sometimes even if a disease is identify correctly but the prescription of medicine or treatment might be wrong. This is might be again due to human error. But sometimes an expert might forget certain steps or precaution. Or an expert might diagnose wrongly. Thus the whole paddy crop might be destroyed.
- iii. There has been very limited software that uses to help paddy farmer in identifying the disease. The software available in the market usually runs on desktop/laptop which proves time consuming and inconvenience. Photo that are taken are require to be upload and then its photo are process. If a mobile application is develop. Paddy farmer could easily identify the disease conveniently.

1.3 Objectives

The objectives are listed out as below,

- i. To identify Malaysia rice paddy common base on external paddy skin layer.
- ii. To diagnosis and suggest solution for the disease.
- iii. To develop a mobile prototype system,

1.4 Scope

The scopes are listed out as below,

- i. Only be able to identify rice blast, rice sheath blight and brown spot disease
- ii. To develop a mobile prototype system that supports Symbian 5th Edition Series only

1.5 Target User

The target are listed out as below,

- i. Farmers with low education and knowledge on rice paddy disease.
- ii. For students who wanted to know about rice paddy disease.

1.6 Thesis Organization

The thesis contains 6 chapters. Chapter 1 will discuss about the introduction of the thesis, problem statement, scope and target user. This is to give reader a brief idea about the background of the thesis and its recommended solution to solve it. In chapter 2 it will discuss about the literature review. It explains the research being done base on this thesis. Article, journals, books and even video documentary are being read and watch to gain a fundamental knowledge on the subject. It will also discuss about the previous researcher method or technique use. Overall it will discuss anything that is relevant to the thesis topic.

Chapter 3 will discuss about the methodology being use. It will discuss and give justification on the techniques, hardware and software that are being use. Compare and comparison are also discussed in here. Chapter 4 document all the process involves in developing the project.

Chapter 5 will discuss about the data and result. Table is being tabulate and calculation is being done to give comparison on the result. Lastly conclusion is being made in chapter 6. It will make conclusion on the overall project summary this includes every summary from each chapter. For future reference new design, new approach and ideas are being written in the last section.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In this section of literature review is separate into five sections. The first section is to study the architecture of other people system. This can give a guideline to determine the architecture of Common Paddy Disease Identifier system (Easy Snap). The second section would review on other people's technique use in their project that is related to Common Paddy Disease Identifier system. This could give an idea on the current trend of technique use in the market. Common problems that face by previous software developers could also be known and could be avoided.

The third and fourth section discuss about the method, tools and technology that are use in the previous years. Method is discussed and reviews to gives reader a better understanding on software approach. Tools also play a very important role for a software developer. This is why usage of tools is also including in the literature review. If a software developer is not

knowledgeable on the tools he knows and chooses the wrong tools. He or she would then have problems during software development stages.

The last section would discuss on the common paddy disease that are face in today world. Reason of cause and symptoms would be discussed. This is to help the system in finding the object of interest and to improve the accuracy of the symptom.

2.2 Image Processing Architecture and Design

Architecture and design in image processing discuss about the flow of image being process until an output had reach. Through studying the designs, the advantages and disadvantages could be determined. For most digital image processing, it would follow similarly as figure 1.

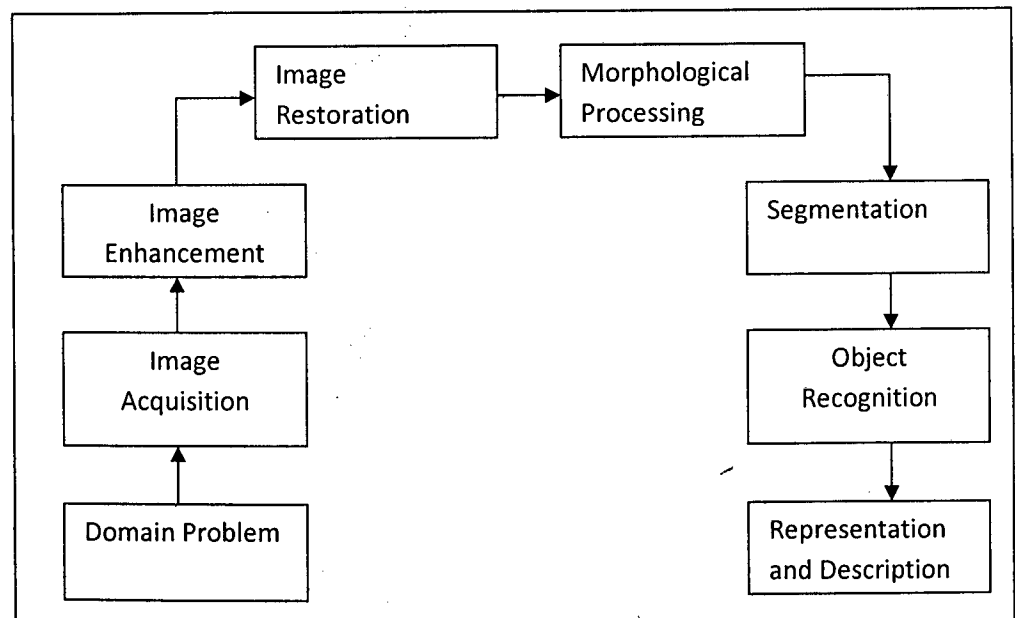


Figure 2.1: Key Stages in Digital Image Processing- resources adopted from [19]

Raw image that transfer into digital form begin at image acquisition stage. Later image are enhance and restore using blur, smoothen, sharpen, soften and gamma correction. This is to improve the condition of the digital image. At later stage morphological processing and segmentation is perform unto the digital image. Object of interest are divided from the background.

There are various techniques such as SUSAN, Sobel operator and Otsu method is use. Object is would then be recognize. Description and justification are given at the end of the stage. Below are the summary of each stages in image processing

- i. Domain Problem: This section is where case study is been studied and image processing can be implement as an alternate method or not.
- ii. Image Acquisition: This refers to image that been capture or taken are converted to binary or digitalized.
- iii. Image Enhancement: Condition of image is improved by using technique such as controlling contrast and brightness and sharpen.
- iv. Image Restoration: This process is to remove any form of degradationsuch as motion blur and noise.
- v. Morphological Processing: Object of interest is being shape to improve the final result by adding or removing pixel neighbor pixel.
- vi. Segmentation: Segmentation is a process to divide background and object of interest.
- vii. Object Recognition: Object of interest is identified base on object characteristic.

- viii. Representation and Description: Reason and justification of object being chosen are listed out.

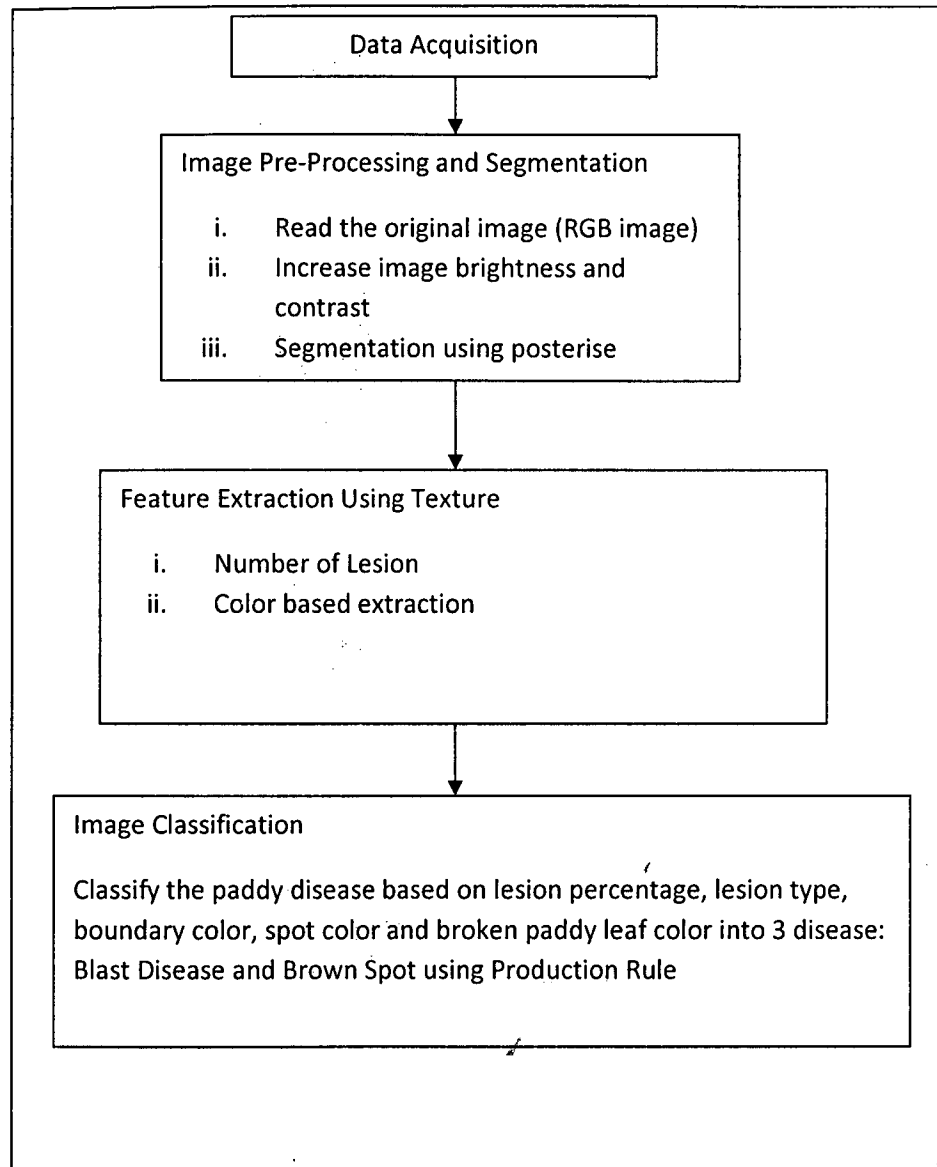


Figure 2.2: Schematic overview of Paddy Disease Diagnosis System- resource adopted from [4].

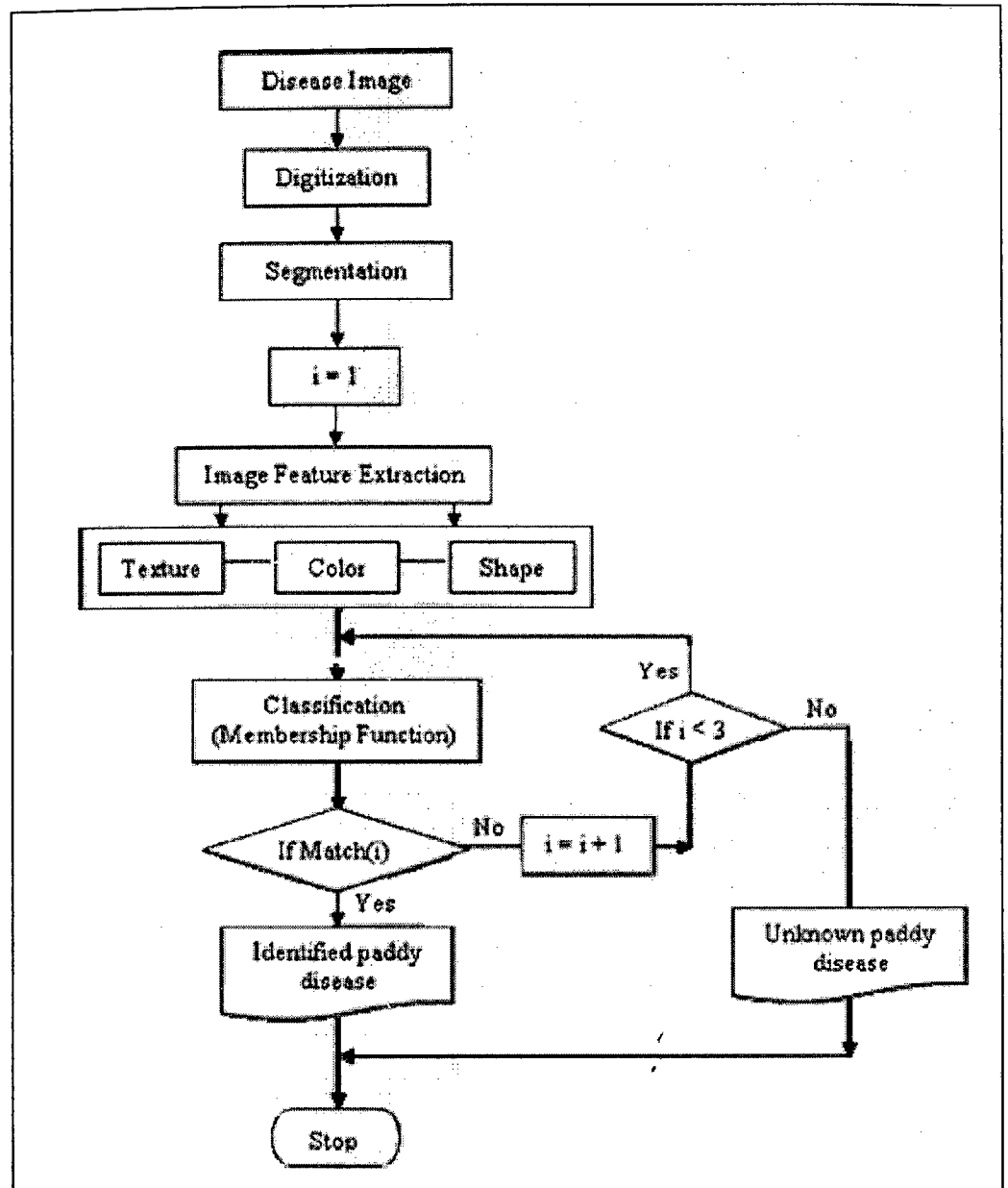


Figure 2.3: Modularized architecture of the overall system-adopted resource from [20]

Figure 2 and figure 3 are example of architecture on paddy disease image processing. Both architecture uses different approach but has similarities in

architecture as shown in figure 1. Image are digitalize, enhance, segmentation and then object recognize. There are no absolute correct when comes to design.

2.3Image Processing Technique

In this section we would discuss the technique use by other people. Technique is very important in image processing because technique could affect the result of the experiment. A computer vision does not function like a human eye. Computer has to use certain algorithm or techniques to conduct their function. This is because what computer sees are only pixel and bytes.

2.3.1 Bilateral Median Filter (Noise Remover)

Noise would refer to the speckle or dust when an image is shooting. It would blur or disturb the image. Hence, when an image processing is done the noise would affect the result or the outcome of the process. Median filter is usually used to remove the speckle or dust from an image.

Median filter usually works by correcting the pixel value. When there is a noise in between two pixels. The noise pixel would be usually higher than the other two pixels. Median filter would correct the noise pixel by comparing the neighbor pixel. If the neighbor pixel were both 1 hence the noise pixel would be corrected to 1 also.

The only con here would be the amount of data can be handling. Median filter uses iteration looping and its very exhaustive to the CPU. Therefore