STUDY ON LEAVE IMAGE PROCESSING WITH APPLICATION IN HERBAL CLASSIFICATION AND EARLY DETECTION OF CHILI PLANT DISEASE

by

ZULKIFLI BIN HUSIN
(1040210539)

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy in Computer Engineering

School of Computer and Communication Engineering
UNIVERSITI MALAYSIA PERLIS

2016

UNIVERSITI MALAYSIA PERLIS

DECLARATION OF THESIS

Author's full name : ZULKIFLI BIN HUSIN

Date of birth : 29 / 03 / 1975

Title : STUDY ON LEAVE IMAGE PROCESSING WITH APPLICATION IN HERBAL CLASSIFICATION AND EARLY DETECTION OF PLANT CHILI DISEASE.

Academic Session : 2011 / 2012

I hereby declare that the thesis becomes the property of Universiti Malaysia Perlis (UniMAP) and to be placed at the library of UniMAP. This thesis is classified as :

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*

☐ **RESTRICTED** (Contains restricted information as specified by the organization where research was done)*

☑ **OPEN ACCESS** I agree that my thesis is to be made immediately available as hard copy or on-line open access (full text)

I, the author, give permission to the UniMAP to reproduce this thesis in whole or in part for the purpose of research or academic exchange only (except during a period of _____ years, if so requested above).

Certified by:

_____
SIGNATURE

750329-02-6525
_____
(NEW IC NO. / PASSPORT NO.)

Date :_____

_____
SIGNATURE OF SUPERVISOR

PROF. DR. ALI YEON MD SHAKAFF
_____
NAME OF SUPERVISOR

Date : _____

NOTES : * If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentially or restriction.

GRADUATE SCHOOL
UNIVERSITI MALAYSIA PERLIS

PERMISSION TO USE

In presenting this thesis in fulfilment of a post graduate degree from the Universiti Malaysia Perlis, I agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by my supervisor(s) or, in their absence, by the Dean of the Graduate School. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without any written permission. It is also understood that due recognition shall be given to me and to Universiti Malaysia Perlis for any scholarly use which may be made of any material from my thesis.

Request for permission to copy or to make other use of material in this thesis whole or in part should be addressed to

Dean of Graduate School
Universiti Malaysia Perlis (UniMAP)
No. 112 & 114, Tingkat 1, Blok A, Taman Pertiwi Indah,
Jalan Kangar-Alor Setar,
Seriab, 01000 Kangar, Perlis.

DEDICATION

I would like to dedicate this thesis firstly to my wife, Siti Nur Adh Dhuha Binti Derani, for her patience and guidance through this process. Secondly to my parents, Allahyarham Hj. Husin Bin Awang and Allahyarhamah Hjh. Siti Maryam Binti Md Saman, whom always believed in my potential, even in the hardest time.

Finally to all my family. They provided me with hope at a time in my life when I needed it the most. I am most certainly say that this would not have been possible for me if these people had not been in my life.

ACKNOWLEDGEMENTS

All praises and gratitude to Allah s.w.t, who has guided and blessed me in finishing my project successfully.

In the first place I would like to record my gratitude to Prof. Dr. Ali Yeon Bin Md Shakaff for his supervision, advice and guidance from the very early to the final stage of this project which enabled me to develop an understanding of this project. His involvement has triggered and nourished my intellectual maturity that I will benefit from, for my future.

Words fail me to express my appreciation to my wife, children, parents and all family for their dedication, love, inseparable support and prayers.

My special thanks goes to my friends, Abdul Hallis Abdul Aziz, Rohani S. Mohamed Farook, Azizi Harun, Mohamad Syahir Bin Mohd Kamar and all members in Sensor Labs and Computer staff for their warm support and encouragement, each from unique perspective. To others who have assisted in any way, I express my sincere gratitude.

Finally, I would like to thank everybody who was important to the completion of the project, as well as expressing my apology that I could not mention personally one by one.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
BPNN	Back-Propagation Neural Network
CCD	Centroid Contour Distance
CMY	Cyan-Magenta-Yellow
CMYK	Cyan-Magenta-Yellow-Black
CNN	Cellular Neural Network
HSI	Hue-Saturation-Intensity
HSV	Hue-Saturation-Value
MMC	Move Median Centre
MPP	Maximum Power Point
RGB	Red-Green-Blue
SVD	Singular Value Decomposition
VOC	Volatile Organic Compound
PNN	Probabilistic Neural Network
GSP	Generalized SoftMAX Perception
RBFNN	Radial Basis Function Neural Network
RBPNN	Radial Basis Probabilistic Neural Network
PCA	Principle Component Analysis
PHOG	Pyramid Of Histograms Of Orientation Gradient
SVM	Support Vector Machine
ACH	Angle Code Histogram
PDA	Procreates Discriminant Analysis
DPLS	Discriminant Partial Least Squares
QDA	Quadratic Discriminant Analysis
DFA	Discriminant Function Analysis
VOCs	Volatile Organic Compounds
LWF	Leaf Width Factor
MLP	Multi-Layer Perceptron

LVQ-ANN	Learning Vector Quantization Neural Network
DAN2	Dynamic Architecture For Artificial Neural Network
RBF-SVM	Radial Basis Function Support Vector Machine
SVD	Singular Value Decomposition
GUI	Graphic User Interface
HSV	Hue, Saturation, Value
HSI	Hue, Saturation, Intensity
ROI	Rectangle Of Interest

Kajian Pemprosesan Imej Daun dengan Applikasi Pengkelasan Herba dan Pengesanan Awal Penyakit Tumbuhan Cili

ABSTRAK

Herba telah digunakan secara meluas dalam penyediaan makanan, perubatan dan industri kosmetik. Pengetahuan tentang jenis-jenis herba yang perlu digunakan bagi tujuan yang disebut diatas adalah sangat penting. Pengenalpastian dan penentuan jenis-jenis herba ini masih dijalankan secara manual dan menjurus kepada kesilapan manusia. Merekabentuk sistem pengenalan spesis herba secara automatik dan bersesuaian adalah sangat penting bagi meningkatkan kecekapan pengkelasan spesis herba. Cili *(Capsicum Annum and Capsicum Frutescen)* adalah sayuran berbuah yang penting yang digunakan dalam hidangan makanan oleh majoriti orang Asia. Penanaman cili telah menjadi bidang yang sukar dan teliti kerana kelemahan kepada pelbagai serangan dari mikro-organisma, penyakit bakteria dan serangga yang meninggalkan tanda-tanda perbezaan pada daun, batang atau buah. Kaedah manual menggunakan racun perosak dan bahan kimia yang sewenang-wenangnya digunakan pada keseluruhan ladang. Bagi meningkatkan proses tersebut, merekabentuk dan membangunkan sistem pengesanan penyakit yang mudah dan automatik adalah penting. Beberapa kajian telah dijalankan di dalam bidang pengkelasan spesis tumbuhan menggunakan beberapa factor-faktor tertentu (bentuk dan saiz daun). Pengkelasan dilaksanakan melalui pengguanaan beberapa teknik pemprosessan imej. Walau bagaimanapun, kajian literatur menunjukan bahawa masih terdapat kekurangan dalam bidang pengkelasan spesis tumbuhan herba. Oleh itu, penyelidikan ini memfokuskan pendekatan pengkelasan terhadap bentuk ciri-ciri tekstur dan warna daun-daun herba. Kombinasi teknik-teknik ini digunakan dalam morfologi pemprosessan imej seperti SVD dan skeleton mampu melaksanakan pengkelasan spesis herba tanpa mengira bentuk dan saiznya. Tambahan pula, teknik tersebut menunjukkan kemampuan untuk mengesan penyakit awal tumbuhan cili melalui pemeriksaan ciri-ciri daun menggunakan teknik model warna HSV. Sistem pengecaman spesis herba yang dicadangkan menggunakan algorithma rangkain neural dan teknik pemprosessan imej bagi melaksanakan pengkelasan keatas dua puluh spesis herba. Seratus sampel bagi setiap spesis diumpukkan kedalam sistem dan didapati mencapai ketepatan pengkelasan pada kadar 98.9%. Perkara yang menjadi keutamaan bagi sistem tersebut mampu mengenalpasti spesis daun herba walaupun ia dalam keadaan kering, basah, koyak atau cacat. Selain itu, kaedah pengesanan awal penyakit tumbuhan cili yang baru secara automatik berdasarkan warna dan tekstur ciri dengan menggunakan model warna HSV dan teknik BPNN melalui sistem sokongan keputusan pintar dibentangkan dalam kajian ini. Sistem yang dicadangkan menggunakan teknik pemprosesan imej pada satu ribu sampel tumbuhan cili dan ketepatan pengesanan tersebut pada 97.7%. Kecekapan dan keberkesanan kaedah yang dicadangkan dalam mengecam tumbuhan herba dan mengesan penyakit awal tumbuhan cili yang telah ditunjukkan oleh eksperimen.

Image Processing Studies on Herbs Leave Classification and Early Detection of Chili Plant Disease as an Application

ABSTRACT

Herbs have been widely used in food preparation, medicine and cosmetic industry. Knowing which herbs to be used would be very important in these applications. The current way of identification and determination of the types of herbs however, is still being done manually and prone to human error. Designing a convenient and automatic recognition system of herbs species is essential since this will improve herb species classification efficiency. Chili *(Capsicum Annum and Capsicum Frutescen)* is an important fruiting vegetable used in majority of Asian dishes. Chili cultivation has been a very difficult and meticulous task due to its vulnerability to various attacks from micro-organisms, bacterial disease and pests which leave distinguished marks on leaves, stems or fruits. Current manual method applies pesticides and chemicals indiscriminately throughout the farm. To improve the process, development of an automated disease detection is essential. There are a few research that have been done in classification of the plant species using certain factors (leaf shape and size). The classification are accomplished through several image processing techniques. However, the literature shows that there are still a gap in classifying the herb plants species. Therefore, this research focuses on classification approach to the shape, texture features and colors of the herbs leaves. The combination of techniques used in morphology image processing i.e. SVD and skeleton would be able to classify the species of herb regardless of the shape and size. In addition, the techniques demonstrate the capability to detect early plant chili disease through leaf features inspection using HSV colour model technique. The proposed herbs species recognition system employs neural networks algorithm and image processing techniques to perform classification on twenty herbs species. One hundred samples for each species went through the system and the recognition accuracy was at 98.9%. Most importantly the system is capable of identifying the herbs leaves species even though they are dried, wet, torn or deformed. Additionally, a novel method of early automatic recognition for plant chili disease based on color and texture features using a HSV color model and BPNN technique via intelligent decision support system is presented in this research. The proposed system employs image processing technique on one thousand chili plant samples and the recognition accuracy was at 97.7%. The efficiency and effectiveness of the proposed methods in recognizing herbs plant and detecting early plant chili disease are demonstrated by the experiments.

CHAPTER 1

INTRODUCTION

1.1 Introduction to Herbs Species Classification

Malaysia is an agricultural country with a lot of people who work in the field based on the herbal industry, especially food, drugs and cosmetics. Classification of herb is useful only if herb can be classified among of herbs. Botanists classified herbs based on their knowledge and expertise, but for laymen, herb classification is still a complicated task. Classification of herbs can be made simpler by using computer aided system because there are more than one-half million of herbs inhabiting on the Earth. Without automation, classification becomes a complex task.

As the use of herbs as remedies becomes popular nowadays, a more effective way to find and identify various types of herbs plants which can be obtained from the natural environment, needs to be developed. With many species becoming extinct due to human development, botanists are urged to research as many species as possible before many of them become extinct. The research is mainly performed by recording species distribution and the time dependent evolution of the species. The process requires botanists identifying herb species on the field and noting down relevant data such as mutations and local growth

density. The process are done manually mostly, with minimum assistant from computers or automated methods. With automation, the herb species recognition can be made faster, thus enable more efficient sampling for any vegetation.

Nevertheless, the advantages are not only limited to the recording of endangered species but also the fast rate and affordability of the world transportation has led to the introduction of intrusive herb species which could endangered the local environmental balance. Having updated information of these species is very critical for local agriculture and following their evolution is important for a healthy local ecosystem. Until recently, there were insufficient digital applications to support the accomplishment of the aforementioned surveys. However, the advent of mobile computing (laptops, smartphones, tablets, etc.) has opened up new possibilities in computer-aided applications. For instance, herbs identification has always been a very time consuming task, as very unclear features may differentiate members of the same families.

Although herbs have many features (i.e. branch shape, dimension, and area of development), their leaf is one of the most defining features. In identifying leaves, even experienced botanists normally rely on feature trees to correctly determine the herb species. It will take time to process and determining the species. The leaves of plants have different characteristics in terms of shape, colour and pattern of veins but the shape and the edges are the features most commonly used. The criticality of leaf shape as a defining species feature is acknowledged by the scientific community and has been focused in many publications explaining automated recognition techniques. (P. N. Belhumeur, et al., 2008)

The purpose of this research is to develop the herb recognition system through image processing technique to guide botanists to identify the herb species on the field. The idea of the system is as simple as taking a photo of a leaf and comparing it automatically to the list in the original set of herb species database. The list would contain about 20 species, all displayed at once. The botanist would then decide the correct species visually. This would require less effort than navigating through a feature tree, since humans process visual information faster than text.

1.2 Introduction to Early Detection of Plant Chili Disease

Chili is a major horticultural commodity and one of the fruiting vegetables which are popular in an Asian countries, especially in Malaysia. Apart from being used in daily dishes, chili is also known as medicinal food since long time ago. Traditionally, chili has been known to help treat certain illness such as blood circulatory system problem, reduce and remove mucus from the lungs to cure bronchitis, influenza and asthma. Chilies contain flavour and antioxidants that help protect the body from cancer. Additionally, chili as a food contributes significant source of vitamins A, B, C and E to human body.

Chili planting requires suitable range of temperature, rainfalls and soil acidity. Temperature range of (15-32) ºC and rainfall between (1500-2000) mm per year would be most suitable for chili cultivation. Temperature lower than the specified range would render the growth stunted, while higher temperature would cause the flower to fall during flowering season thus reducing yield significantly. Chili could be grown in most soil types including

peat and sand. However pH range required is 4.5-7 and if too low crop may die from soil poisoning and if pH is too high crop could be unproductive (N. Kamisan, 2011). With this specific requirement and need for productive chili cultivation, the use of smart technology seems imperative to ensure chili cultivation industry becomes more productive and competitive.

A systematic approach to ensure fertility and quality of chili plants needs to be developed so that the chili production will increase. Most of farmers refuse to cultivate chili in the rainy season due to potential increase of chili disease. Therefore, implementing a novel design and very cost effective techniques for protecting the chili plants and its fruits is vital. Traditionally, the diagnoses techniques of plants chili diseases such as transmission electron microscopy and immunological techniques are still used for detection and identification of viral particles and fungi (B. Zechmann & G. Zellniga, 2009). With the suitability of current technology, a system to detect early plants chili diseases automatically is very important to be developed to reduce the maintenance costs, save on fertilizer use and excessive use of pesticides in order to provide a good ventilation to the environment. However, the main challenge is to enable control in a short period of time to detect the disease in population of plants before they spread throughout the farm.

Consequently, several techniques are needed to provide automatic on-site detection to investigate visual appearance of the appropriate part of the plants using image processing techniques and examine the Volatile Organic Compounds (VOC) of plants. There are two options which might produce an attractive way for non-destructive technique for plant testing. These techniques need to feed the knowledge of decision support system to advice

farmers the options of intervention and controlling their farms from plants diseases. (L. Jullada, et al., 2008).

1.3 Problem descriptions

1.3.1 Problem statement in herbs species classification

Since most of the herbs plants that grow in the forest depletion and hard to find then, an important aspect is needed to protect the herb from extinct either through monitoring or recording of any herb species. A typical way to gather herbs is through the experience of previous botany experts. This conventional technique is less effective and less efficient compared to the computerized system. Furthermore, this approach is more difficult and too slow if used in a large area. With the increasing in computer usage, classification of herb species that were previously done from traditional techniques become easier and faster. Therefore, this research has introduced a more effective system for the classification of herbs species through image processing techniques by using a leaf as sample for species classification process.

There are two main challenges related to using leaf as sample for species classification. First, the form of herb leaf which may differ with different herb species. For example, if different herb species have a different herb leaf shape characteristic. The differences in the characteristics of the margin or vein structures may be quite significant and the differences are often more apparent shape, although not visible to the botanists. In many

cases, the size of leaf is largely determined by the environment while the shape is inherited naturally.

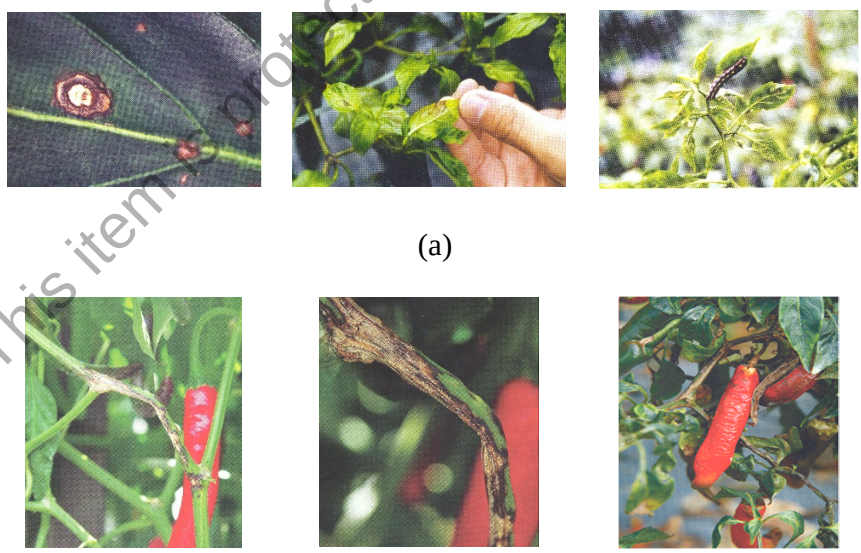
Second, the shape of leaf is the easiest ways to be removed automatically. If the leaves are imaging on a black background or a plain white then, a threshold technique can be used to separate the leave from the background using the pixels bordering leaf background. Furthermore, the gross structure of the leaf can be maintained even though the leaf specimen is damaged probably through the ages. For example, there are many dry leaves change the colour of leaf from green to brown colour. Therefore, colour is not commonly informative feature to classify the species of herb plants. Until now, no commercial device is available that is able to perform the classification of herb species through form, colour and odor. The literature also indicates that, there is a research gap for image processing technique in herb plant species classification. The information about the techniques that were previously developed by researchers will be described in more details in the literature review section.

Previous researches concentrated mainly on good quality and minimally deformed or non-deformed leafs as samples for processing. Some other researchers carefully selected their leaf samples, so they are of almost the same sizes for good detection. This approach however allows them to only used simple edge detection technique as feature extraction for image processing. Most researchers also mainly use color transformation technique for image processing. Nevertheless, most dry leafs change colour from green to brown over time. As such, colour is not a distinguishable feature to be used for herb species classification. The combination of techniques used in morphology image processing in this thesis (i.e. SVD and skeleton) would be able to classify the herb species regardless of the shape and size.

Additionally, the techniques would also be able to classify normal shape and also deformed leafs (i.e. damaged, dried or colour changed).

1.3.2 Problem statement in early detection of plant chili disease

Generally, there are two factors or symptoms that could bring death and destruction to plants such as chili life (insects, bacteria, fungi and viruses) and non-living (temperature, excessive humidity, poor lighting, inappropriate soil pH, pollution air causing insufficient nutrient). Common symptoms of the disease include abnormal growth of chili leaves, colour distortion, stunted growth, withered and broken chili. Figure 1.1 shows some samples of plant chili diseases.



(a)

(b)

Figure 1.1: Samples of Plant Chili Disease
(a) Spots and Bacterial Disease (b) Fitoptora Disease