



**UNIVERSITI PUTRA MALAYSIA**

**DESIGN AND DEVELOPMENT OF FICUS SPECIES DATABASE AND  
2D LEAF IMAGE IDENTIFICATION SYSTEM**

**TEE CHIOU HAU**

**FBSB 2010 14**



**DESIGN AND DEVELOPMENT OF *FICUS* SPECIES DATABASE AND 2D  
LEAF IMAGE IDENTIFICATION SYSTEM**

**By**

**TEE CHIOU HAU**

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in  
Fulfilment of the Requirements for the Degree of Master of Science**

**October 2010**



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfillment of the requirements for the degree of Masters of Science

**DESIGN AND DEVELOPMENT OF *FICUS* SPECIES DATABASE AND 2D LEAF IMAGE IDENTIFICATION SYSTEM**

By

**TEE CHIOU HAU**

**October 2010**

**Chairman: Associate Professor Norihan bt. Mohd. Saleh, PhD**

**Faculty: Biotechnology and Biomolecular Sciences**

Plants are important part of the ecosystem in the world. Numerous studies had been done on the richness of plants diversity. There are many plant databases currently available online. A search for useful information regarding a particular plant can be performed easily through the databases using text such as the name of a plant. However, this task will be difficult if only image of the plant is available. Thus, to facilitate the search using image of a leaf, a simple two-dimensional (2-D) *Ficus* Identification Database System (FicIDS) was developed. This system can perform search using both text information and image of the plant. Basically, FicIDS system was focused on *Ficus* species. *Ficus* species were chosen due to its variable leaf shapes and its significance in our local herbal industries. Currently, there is a high demand for the natural products derived from this plant, particularly from *Ficus deltoidea*. But, it is very difficult to identify a *Ficus* plant correctly since there are more than 100 different *Ficus* species and more than 20 different varieties of *F.*



*deltoidea* available in Malaysia. Furthermore, there is no proper documentation of this plant. The FicIDS system was designed and developed to identify *Ficus* plants based on the leaf image and to store the data about these species. Herbarium specimens of *Ficus* plant were prepared as evidence for the plants used in this study. Images of herbarium specimen and live plant materials were captured and stored in the database. Additional text information on *Ficus* plants were also collected from various sources to build up the database. Microsoft Office Access database management system was used to develop the plant database on Windows XP platform. 2-D leaf images identification system was constructed using the MATLAB R2006a program. The shape and size of plant leaves were used as the main features to identify a particular *Ficus* species. The process of image identification system comprised of four steps, namely image acquisition, image preprocessing and features extraction, computing of descriptors values and normalization, and *k*-nearest neighbor classification and decision making based on Euclidean distance. Thirteen descriptors values were used in identification of the image, which include aspect ratio, circularity, area convexity, rectangularity, sphericity, eccentricity, and 7 invariant moments. The accuracy of leaf image identification system was evaluated by using 130 leaf images corresponding to 6 *Ficus* species and 4 varieties of *Ficus deltoidea*. The evaluation of overall performance of FicIDS system showed that 120 (92.31%) of the tested leaf images were successfully identified by the system. However, the 2-D FicIDS system has some limitations with respect to image identification. The system requires single intact leaf with white background for identification. These limitations may be overcome by using three-dimensional (3-D) leaf image where more features of leaf such as leaf texture or venation can be included to improve the image recognition performance.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Sarjana Sains

**REKABENTUK DAN PEMBINAAN PANGKALAN DATA SPESIES *FICUS*  
DAN SISTEM IDENTIFIKASI IMEJ DAUN 2D**

Oleh

**TEE CHIOU HAU**

**Oktober 2010**

**Pengerusi: Profesor Madya Norihan bt. Mohd. Saleh, PhD**

**Fakulti: Bioteknologi dan Sains Biomolekul**

Tumbuhan merupakan sebahagian daripada ekosistem yang penting dalam dunia ini. Pelbagai penyelidikan telah dijalankan ke atas kepelbagaian diversiti tumbuhan. Terdapat banyak pangkalan data tumbuhan yang boleh didapati dalam talian pada masa ini. Pencarian maklumat yang berguna mengenai sesuatu tumbuhan boleh dilaksanakan dengan mudah melalui pangkalan-pangkalan data ini dengan menggunakan teks seperti nama tumbuhan. Namun begitu, tugas ini akan menjadi payah jika hanya imej tumbuhan tersedia ada sahaja. Oleh yang demikian, untuk membantu pencarian maklumat mengenai tumbuhan dengan menggunakan imej daun, satu sistem identifikasi dan pangkalan data *Ficus* (FicIDS) yang dua dimensi (2-D) telah dibinakan. Sistem ini boleh melakukan pencarian dengan menggunakan kedua-dua maklumat teks dan imej mengenai tumbuhan. Pada dasarnya, sistem FicIDS fokus pada spesies *Ficus*. Spesies *Ficus* telah dipilih sebab bentuk daunnya yang kepelbagaian dan kepentingannya dalam industri herbal tempatan kita. Pada masa ini,

terdapat permintaan yang tinggi terhadap produk semulajadi daripada tumbuhan ini, terutamanya daripada *Ficus deltoidea*. Walaubagaimanapun, untuk mengecam satu tumbuhan *Ficus* dengan tepat adalah payah, memandangkan terdapat lebih daripada seratus spesies *Ficus* yang berlainan dan lebih daripada 20 varieti *F. deltoidea* di Malaysia. Tambahan pula, dokumentasi mengenai tumbuhan ini adalah terhad dan tidak lengkap. Sistem FicIDS ini telah dirancang dan dilaksanakan untuk mengecam tumbuhan *Ficus* berdasarkan imej daun dan untuk menyimpan data mengenai spesies ini. Spesimen herbarium tumbuhan *Ficus* telah disediakan sebagai bukti bagi tumbuhan yang digunakan dalam kajian ini. Imej spesimen herbarium dan tumbuhan yang hidup telah ditangkap dan disimpan dalam pangkalan data. Maklumat tambahan teks mengenai tumbuhan *Ficus* juga telah dikumpulkan daripada pelbagai jenis sumber untuk membina pangkalan data. Sistem pengurusan pangkalan data Microsoft Office Access telah digunakan untuk membina pangkalan data ini dalam pelantar Windows XP. Sistem identifikasi imej daun 2-D telah dibina dengan menggunakan program MATLAB R2006a. Bentuk dan saiz daun tumbuhan telah digunakan sebagai fetur utama untuk mengecam sesuatu spesies *Ficus* yang tertentu. Proses dalam sistem identifikasi imej terdiri daripada 4 langkah, iaitu pemerolehan imej, pemprosesan awal imej dan ekstraksi fetur, pengiraan nilai diskriptor dan normalasi, pengkelasan 'k-nearest neighbor' dan penghasilan keputusan berdasarkan jarak terdekat Euclidean. Tiga belas nilai diskriptor yang digunakan dalam identifikasi imej merangkumi nisbah aspek, 'circularity', 'area convexity', 'rectangularity', 'sphericity', 'eccentricity' dan tujuh 'invariant moments'. Sebanyak 130 imej daun daripada 6 spesies *Ficus* dan 4 varieti *Ficus deltoidea* telah digunakan untuk menilai ketepatan sistem identifikasi imej ini. Penilaian mengenai prestasi keseluruhan sistem FicIDS menunjukkan bahawa 120 (92.31%) daripada imej daun

yang diuji telah berjaya dikenal pasti dengan menggunakan sistem itu. Namun begitu, sistem FicIDS 2-D ini mempunyai batasan berhubung dengan identifikasi imej. Sistem ini terbatas kepada penggunaan daun tunggal yang lengkap dengan latar belakang putih sahaja untuk pengecaman imej. Pembatasan ini boleh diatasi dengan menggunakan imej tiga dimensi (3-D) daun, di mana lebih banyak fitur daun seperti tekstur atau pemveinan daun boleh digunakan untuk memperbaiki prestasi pengecaman imej.

## ACKNOWLEDGEMENTS

This thesis could not have been accomplished without the help of many individuals. First and foremost I would like to express my sincere gratitude to my thesis supervisor, Associate Prof. Dr. Norihan Mohd Saleh, for providing me the opportunity to conduct this research under her supervision and guidance. Her guidance and support were invaluable. Also, I would like to thank Associate Prof. Dr. Mohd Puad Abdullah and Associate Prof. Dr. Rahmita Wirza O. K. Rahmat, my supervisory committee members, who provide much insight and experience to this project. Their assistance and comment for improvements made this thesis a success.

Next, I would like to express my appreciation to Dr. Mohd Nazre Saleh and the staffs of Herbarium of Faculty of Forestry, UPM, for their willingness in helping me during the preparation of herbarium specimens. Special thanks also to Mr. Adam bin Ali and the staffs from Department of Agriculture, Serdang, for providing the plant materials used in this study. Also, I would like to thank Mr. Shamsul Khamis and Mr. Roslan Mamat for the opportunity to work together with them during the identification of *Ficus* species and the varieties of *Ficus deltoidea*.

Thanks to my family, for their love, support, encouragement and understanding. Also, I would like to extend the hand of appreciation to my sisters, who assisted me in overall thesis preparation. Finally, I would like to express my special thanks to my friends and colleagues for their constant support.



I certify that an Examination Committee has met on 12<sup>th</sup> October 2010 to conduct the final examination of Tee Chiou Hau on her Master thesis entitled "Design and development of *Ficus* species database and 2D leaf image identification system" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the student be awarded the degree of Masters of Science.

Members of the Examination Committee were as follows:

**Nor ‘Aini Abdul Rahman, PhD**

Department of Biochemistry  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Chairman)

**Suhaimi bin Napis, PhD**

Associate Professor  
Department of Cell and Molecular Biology  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Internal Examiner)

**Adam Leow Thean Chor, PhD**

Department of Cell and Molecular Biology  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Internal Examiner)

**Amir Feisal Merican Aljunid Merican, PhD**

Associate Professor  
Institute of Biological Science  
Faculty of Science  
Universiti Malaya  
(External Examiner)

---

**SHAMSUDDIN SULAIMAN, PhD**

Professor and Deputy Dean  
School of Graduate Studies  
Universiti Putra Malaysia

Date:

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Masters of Science. The members of the Supervisory Committee were as follows:

**Norihan bt. Mohd. Saleh, PhD**

Associate Professor  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Chairman)

**Mohd. Puad Abdullah, PhD**

Associate Professor  
Faculty of Biotechnology and Biomolecular Sciences  
Universiti Putra Malaysia  
(Member)

**Rahmita Wirza O.K. Rahmat, PhD**

Associate Professor  
Faculty of Computer Science and Information Technology  
Universiti Putra Malaysia  
(Member)

---

**HASANAH MOHD GHAZALI, PhD**

Professor and Dean  
School of Graduate Studies  
Universiti Putra Malaysia

Date:

## **DECLARATION**

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or at any other institution.

---

**TEE CHIOU HAU**

Date: 12 October 2010

## TABLE OF CONTENTS

	<b>Page</b>
<b>ABSTRACT</b>	ii
<b>ABSTRAK</b>	iv
<b>ACKNOWLEDGEMENTS</b>	vii
<b>APPROVAL</b>	viii
<b>DECLARATION</b>	x
<b>LIST OF TABLES</b>	xiii
<b>LIST OF FIGURES</b>	xiv
<b>LIST OF APPENDICES</b>	xvi
<b>LIST OF ABBREVIATIONS</b>	xvii
<b>CHAPTER</b>	
1 <b>INTRODUCTION</b>	1
2 <b>LITERATURE REVIEW</b>	5
2.1 Taxonomy and Uses of <i>Ficus</i> Species	5
2.2 Digitization of Herbarium	11
2.3 Plant Data Management System	15
2.4 Leaf Image Identification System	20
2.5 Features Selection and Shape Descriptors	24
3 <b>MATERIALS AND METHODS</b>	28
3.1 Plant Materials and Research Framework	28
3.2 Preparation of Herbarium Specimens	30
3.2.1 Collection of Plant Specimens	33
3.2.2 Recording of Field Data	33
3.2.3 Cleaning of Plant Specimens	33
3.2.4 Capturing of Plant Images	34
3.2.5 Preservation of Plant Specimens	34
3.2.6 Pressing and Drying of Plant Specimens	34
3.2.7 Mounting of Plant Specimens	35
3.2.8 Classification and Labeling of Plant Specimens	36
3.2.9 Storage of Herbarium Specimens	37
3.3 Development of <i>Ficus</i> Species Database	37
3.3.1 Database Planning	38
3.3.2 Collection and Analysis of Database Requirements	39
3.3.3 Database Design	41
3.3.4 Application Design	48
3.3.5 Implementation	50
3.4 Development of Identification System	55
3.4.1 System Planning	55
3.4.2 System Design	56
3.4.3 System Implementation	62
3.5 System Testing and Evaluation	73
3.5.1 Testing of Database System	73
3.5.2 Evaluation of Leaf Image Identification System	73
3.5.3 Overall System Evaluation	74



4	<b>RESULTS AND DISCUSSION</b>	75
	4.1 Determination of Herbarium Specimen	75
	4.2 Database System	80
	4.3 FicIDS Identification System	88
	4.4 Overall System Performance	97
	4.4.1 Database System	97
	4.4.2 Leaf Image Identification System	99
	4.4.3 Overall System	101
5	<b>SUMMARY, CONCLUSION AND RECOMMENDATIONS FOR FUTURE RESEARCH</b>	104
	5.1 Summary and General Conclusion	104
	5.2 Future Enhancements	106
	<b>REFERENCES</b>	108
	<b>APPENDICES</b>	117
	<b>BIODATA OF STUDENT</b>	129
	<b>LIST OF PUBLICATIONS</b>	130