UNIVERSITI PUTRA MALAYSIA

# PLANT LEAF RECOGNITION ALGORITHM USING ANT COLONY-BASED FEATURE EXTRACTION TECHNIQUE

MOHAMMAD ALI JAN GHASAB

FK 2013 34

# PLANT LEAF RECOGNITION ALGORITHM USING ANT COLONY-BASED FEATURE EXTRACTION TECHNIQUE

By

MOHAMMAD ALI JAN GHASAB

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

December 2013

© COPYRIGHT UPM

## COPYRIGHT

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia



© COPYRIGHT UPM

**DEDICATIONS**

*Mum*

*Dad*

*and my Sister*



© COPYRIGHT UPM

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

**PLANT LEAF RECOGNITION ALGORITHM USING ANT COLONY-BASED FEATURE EXTRACTION TECHNIQUE**

By

**MOHAMMAD ALI JAN GHASAB**

**December 2013**

**Chair: Asnor Juraiza binti Ishak, PhD**

**Faculty: Engineering**

Plant recognition as a substantial subject of biology has occupied the minds of many botanists throughout the world to concentrate their efforts on the identification of unknown plant species with the aim of protection and other purposes. As a troublesome and gradual process, traditional methods of taxonomy of plants impede a high rate of performance for the taxonomist in this field. In the modern-day, improvements in the fields of artificial intelligence and soft computing have led to the field of automatic plant recognition being considered as a challenging topic due to the various uses of plants in medicine, food and industry. Although many studies have been undertaken to seek out a method that can be applied for the classification of numerous plants, there is still a lack of a highly-efficient system for the recognition of a wide range of different plants. The aim of this research is to contribute to the measurement of physiological dimensions of plant leaves by the proposed Auto-Measure algorithm to operate in an automatical manner which inherently requires an improvement in automatic feature extraction. Moreover, the ant colony optimisation technique will



© COPYRIGHT UPM

be applied as an expert algorithm to make a decision for the selection of optimal features in order to enhance the performance of a classifier for recognition of diverse species of plants. To do this, at first, based on the proposed algorithm, the physiological dimensions of leaves are automatically measured and with regard to these parameters, specified features such as shape, morph, texture and colour are extracted from the image of the plant leaf through image processing to create a reserved feature database to be used for different species of plants. Then, based on the characteristics of each species, decision making is done by means of ant colony optimisation as a search algorithm to return the optimal subset of features regarding the related species. Finally, the selected features are employed by a multi-class support vector machine to classify the species. The proposed method was applied to different kinds of plant and herb species for testing the system and it was found from the experimental results that the system, by eliminating redundant features, not only optimised the number of features in the subset, but also had a remarkably positive impact on the performance of the classifier in a way that implementation of the proposed method on almost 2830 leaves improved the average accuracy over all the plant databases to 96.66 %. Therefore, it can be concluded that the proposed method is capable of a high rate of classification of various plant species.

© COPYRIGHT UPM

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

**LOJI DAUN PENGIKTIRAFAN ALGORITMA MENGGUNAKAN SEMUT COLONY-BERDASARKAN CIRI-CIRI PENGEKSTRAKAN TEKNIK**

Oleh

**MOHAMMAD ALI JAN GHASAB**

**Disember 2013**

**Pengerusi: Asnor Juraiza binti Ishak, PhD**

**Fakulti: Kejuruteraan**

Pengecaman tumbuhan sebagai subjek yang agak penting dalam biologi telah menjadikan ramai botanis di seluruh dunia menumpukan usaha mereka dalam mengenalpasti spesis tumbuhan yang tidak dikenali bagi tujuan perlindungan dan lain-lain. Sebagai proses yang sukar dan berperingkat, keadah tradisional dalam taksonomi tumbuhan amat menjejaskan prestasi ahli taksonomi dalam bidang ini. Pada zaman moden ini, peningkatan dalam bidang pengajian perisian pintar dan pengkomputeran lembut, bidang pengecaman tumbuhan secara automatik menjadi topik yang mencabar disebabkan penggunaan tumbuhan secara meluas dalam bidang perubatan, makanan dan industri. Walaupun banyak kajian telah dijalankan bagi mencari kaedah yang boleh diaplikasi untuk mengklasifikasi pelbagai jenis tumbuhan, masih terdapat kekurangan dalam sistem yang efektif bagi pengecaman pelbagai jenis tumbuhan. Tujuan kajian ini adalah untuk menyumbang dalam pengiraan dimensi fisiologikal daun tumbuhan secara automatik berbanding manual, dimana ini menghasilkan satu penambahbaikan

© COPYRIGHT UPM

dalam pengautomatan sarian ciri. Juga, untuk mengupah pakar algoritma dalam membuat keputusan untuk memilih ciri-ciri yang optimum bagi meningkatkan prestasi pengelas bagi mengenal spesis tumbuhan yang pelbagai. Untuk melaksanakannya, peringkat pertama adalah berdasarkan algoritma yang dicadangkan, dimensi fisiologi daun diukur secara automatik, dan berdasarkan parameter ini, ciri-ciri spesifik seperti bentuk, morph, tekstur dan warna disarikan dari imej daun tumbuhan melalui pemprosesan imej yang kemudiannya ciri-ciri tersebut dijadikan sebagai pangkalan data simpanan ciri-ciri yang akan digunakan kepada spesis tumbuhan yang berbeza. Kemudian, berdasarkan cirri-ciri setiap spesis, pemilihan ciri dibuat berdasarkan teknik ant colony optimisation sebagai algoritma carian untuk mengenalpasti cirri subset yang optimum bedasarkan spesis yang berkaitan. Akhirnya, ciri yang terpilih akan digunakan oleh mesin sokongan vektor pelbagai kelas untuk mengelaskan spesis tersebut. Kaedah yang dicadangkan digunakan kepada pelbagai jenis spesis tumbuhan dan herba yang berbeza sebagai ujikaji kepada sistem, dan didapati dari keputusan eksperimen bahawasanya dengan membuang ciri-ciri yang berulang di dalam sistem, bukan sahaja nombor ciri didalam subset dioptimumkan, bahkan ia juga menunjukkan impak positif yang bermakna kepada prestasi pengelas dengan cara yang pelaksanaan kaedah yang dicadangkan pada hampir 2830 daun improvrd ketepatan purata semua pangkalan data untuk 96.66 %. Oleh itu, dapatlah disimpulkan bahawa kaedah yang dicadangkan berkebolehan untuk klasifikasi pada kadar yang tinggi bagi pelbagai spesis tumbuhan disamping untuk generaslisasi yang pantas dari segi masa.

© COPYRIGHT UPM

## ACKNOWLEDGEMENTS

First of all,I would like to express my deepest appreciation to all those who provided me the possibility to complete this research.

A special gratitude I give to Dr. Asnor Juraiza, whose helped and encouraged me to coordinate my project especially in writing this report.Words fail me to express my appreciation to my parents for their dedication,love,inseparable support and prayers. To my sisters,thank you for being supportive and caring sibling.

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.

© COPYRIGHT UPM

I certify that a Thesis Examination Committee has met on 3 December 2013 to conduct the final examination of Mohammad Ali Jan Ghasab on his thesis entitled "PLANT LEAF RECOGNITION ALGORITHM USING ANT COLONY-BASED FEATURE EXTRACTION TECHNIQUE" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Master of Science.

Members of the Thesis Examination Committee were as follows:

**Name of Chairperson, Ph.D.**
Title (e.g. Professor/Associate Professor/Ir) – Omit if not relevant
Name of Faculty
Universiti Putra Malaysia
(Chairperson)

**Name of Examiner 1, Ph.D.**
Title (e.g. Professor/Associate Professor/Ir) – Omit if not relevant
Name of Faculty
Universiti Putra Malaysia
(Internal Examiner)

**Name of Examiner 2, Ph.D.**
Title (e.g. Professor/Associate Professor/Ir) – Omit if not relevant
Name of Faculty
Universiti Putra Malaysia
(Internal Examiner)

**Name of External Examiner, Ph.D.**
Title (e.g. Professor/Associate Professor/Ir) – Omit if not relevant
Name of Department and/or Faculty
Name of Organisation (University/Institute)
Country
(External Examiner)

**SEOW HENG FONG, PhD**
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

© COPYRIGHT UPM

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

**Asnor Juraiza binti Ishak, PhD**
Senior Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Chairperson)

**Azura binti Che Soh, PhD**
Senior Lecturer
Faculty of Engineering
Universiti Putra Malaysia
(Member)

**Mohammad Hamiruce Marhaban, PhD**
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

© COPYRIGHT UPM

---

**BUJANG BIN KIM HUAT, PhD**
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

## DECLARATION

### Declaration by Graduate Student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No.: _____

© COPYRIGHT UPM

**Declaration by Members of Supervisory Committee**

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____
Name of
Chairman of
Supervisory
Committee: _____

Signature: _____
Name of
Member of
Supervisory
Committee: _____

Signature: _____
Name of
Member of
Supervisory
Committee: _____

© COPYRIGHT UPM

## TABLE OF CONTENTS


| | Page |
|---|---|
| **DEDICATIONS** | ii |
| **ABSTRACT** | iii |
| **ABSTRAK** | v |
| **ACKNOWLEDGMENTS** | vii |
| **APPROVAL** | viii |
| **DECLARATION** | x |
| **LIST OF TABLES** | xiv |
| **LIST OF FIGURES** | xv |
| **LIST OF ABBREVIATIONS** | xviii |
| **CHAPTER** | |
| 1 **INTRODUCTION** | 1 |
| 1.1 Plant Taxonomy Biography | 1 |
| 1.2 Current Difficulties in Leaf Classification | 2 |
| 1.3 Problem Statement | 4 |
| 1.4 Objectives of the Research | 5 |
| 1.5 Contribution of knowledge | 5 |
| 1.6 Research Scope | 7 |
| 1.7 Thesis Layout | 8 |
| 1.8 Summary | 8 |
| 2 **LITERATURE REVIEW** | 9 |
| 2.1 Introduction | 9 |
| 2.2 Leaf Feature Extraction | 10 |
| 2.2.1 Shape-Based Descriptors | 10 |
| 2.2.2 Content-based Features | 14 |
| 2.2.3 Features Combination | 18 |
| 2.3 Feature Decision Making | 21 |
| 2.3.1 Search Starting Point | 22 |
| 2.3.2 Search Procedure | 22 |
| 2.3.3 Evaluation Function | 24 |
| 2.3.4 Search Stopping Criteria | 25 |
| 2.4 Ant Colony Optimization | 26 |
| 2.4.1 Theory of Ant Algorithm | 26 |
| 2.4.2 Applications of ACO in Feature Decision Making | 28 |
| 2.5 Support Vector Machine | 30 |
| 2.6 Summary | 31 |


© COPYRIGHT UPM

3 **METHODOLOGY** 33
3.1 Introduction 33
3.2 Research Framework 33
3.3 Image Source 34
3.3.1 Real Images 35
3.3.2 Controlling Image Databases 36
3.4 Data Acquisition & Image Preprocessing 38
3.5 Automated Feature Extraction Technique 39
3.5.1 Shape Feature Extraction Technique 41
3.5.2 Digital Morphological Feature Extraction Technique 47
3.5.3 Texture Features Extraction Technique 48
3.5.4 Color Feature Extraction Technique 50
3.6 Feature Decision Making with Ant Colony Algorithm 51
3.6.1 Structure of Feature Search Space 54
3.6.2 Probability Function 54
3.6.3 Selection Function 57
3.6.4 Evaluation Function 58
3.6.5 Pheromone Updating and Evaporation 59
3.6.6 Proposed ACOFSS Algorithm 60
3.7 Classification with Support Vector Machine 62
3.8 Summary 62

4 **RESULTS AND DISCUSSION** 64
4.1 Introduction 64
4.2 Automated Feature Extraction Results 64
4.2.1 Experimental Results of Automeasure Algorithm 64
4.2.2 Results of Automatic Construction of Feature Databases 66
4.2.3 Real Image Databases 66
4.3 Feature Decision Making Results 69
4.3.1 Initialize ACO Parameters 69
4.3.2 Experimental Results of Feature Decision Making 69
4.3.3 Analysis on Quality of Features Subsets 70
4.4 Classification Results 75
4.4.1 Real image Databases 76
4.4.2 Controlling Image Database 82
4.5 Comparison With Previous Approaches 84
4.5.1 Comparison on Classification 84
4.5.2 Comparison on Computation Time 85
4.6 Summary 86

5 **CONCLUSION** 87

**REFERENCES** 90
**APPENDICES** 98
**LIST OF PUBLICATIONS** 104

© COPYRIGHT UPM