

**REMOTE SENSING TECHNIQUE FOR OIL PALM AGE
CLASSIFICATION USING LANDSAT-5 TM SATELLITE**

SHAMALA VADIVELU

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS TESIS

JUDUL: REMOTE SENSING TECHNIQUE FOR OIL PALM AGE CLASSIFICATION USING LANDSAT-5 TM SATELLITE

SESI PENGAJIAN: 2013/2014

Saya _____ SHAMALA VADIVELU
(HURUF BESAR)

Mengaku membenarkan tesis (PSM/~~Sarjana/Doktor~~Falsafah) ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hak milik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Silatandakan (/)

_____ SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
_____ TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
_____ TIDAK TERHAD	

(TANDATANGAN PENULIS)

(TANDATANGAN PENYELIA)

Alamat tetap: Lot 3248, Batu4
JalanSekolah , RantauPanjang, 42100
Klang.

DR. CHOO YUN HUOY
NamaPenyelia

Tarikh : _____

Tarikh : _____

CATATAN: *Tesis dimaksudkan sebagai Laporan Projek Sarjana Muda (PSM)
**Jika Tesis ini SULIT atau TERHAD, sila Lampirkan surat daripada pihakberkuasa

**REMOTE SENSING TECHNIQUE FOR OIL PALM AGE
CLASSIFICATION USING LANDSAT-5 TM SATELLITE**

SHAMALA VADIVELU

This report is submitted in partial fulfilment of the requirements for the Bachelor of
Computer Science (Artificial Intelligence)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2014

DECLARATION

I hereby declare that this project report entitled

**REMOTE SENSING TECHNIQUE FOR OIL PALM AGE
CLASSIFICATION USING LANDSAT-5 TM SATELLITE**

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT : _____ DATE: _____

(SHAMALA VADIVELU)

SUPERVISOR: _____ DATE: _____

(DR CHOO YUN HUOY)

DEDICATION

I dedicate my final year project report to my family and friends. To my supervisor, Dr. Choo Yun Huoy and Dr. Asmala Ahmad, for guiding and helping me to finish up this project. I would like to express deep gratitude to my beloved parents for a life-long love and affection. They have been very supportive and encouraging in completion of my thesis and throughout the years of my studies. On top that, I also would like to dedicate this report to my close friends and family who have been very supportive throughout the project development.

ACKNOWLEDGEMENT

I would like to greatly thank the contribution of the following individuals in helping me throughout the project. Firstly is my supervisor, Dr. Choo Yun Huoy and evaluator, Dr. Asmala Ahmad who have given me a lot of encouragement, guidance and support from the initial to the final level of my project which successfully developed an understanding of the project in me. Besides that, I also would like to thank my parents for being so supportive throughout the projects and helped me a lot in term of moral support. They were so helpful where my parents accompanied me to the study area. And not forgetting the Eng Soon Plantation coordinator, Mr. Perumal who has helped in providing the information required for this study. Lastly, I would like to thank my class peer for their cooperation and camaraderie and to all those who supported me in any aspect throughout the project.

ABSTRACT

Age of oil palm is an important variable used in carbon and yield forecasting studies. Conventionally, the age classification of oil palm was made manually by mapping the plantation area. This technique is time consuming and difficult to classify a large area of hectare which causes difficulty for organisation like MPOB to analyse on the yield production. With remote sensing, the nature of acquiring the oil palm age through spectral response is more convenient. Limited studies were concerning the performance of current technique in classification of oil palm age. Mostly were using traditional parametric statistical approaches. Moreover, previous studies for vegetation age prediction were carried using different remote sensing approaches consisting of different resolution and measurements of data. This project demonstrates the procedure/algorithm to classify age of oil palm trees using LANDSAT-5 TM remote sensing data. The study were conducted in two phases; where phase I is the land cover classification whereas phase II is the oil palm age classification. Firstly, region of interest (ROI) was identified and drawn in order to supply training and testing pixels for the supervised classification. Maximum likelihood (ML) classifier was used for land cover classification with overall accuracy of 85.69%. Whereas, three classifiers were studied, such as: ML, Neural Network (NN) and Support Vector Machine (SVM) for oil palm age classification. Two sets of training set, smaller training set and larger training set were compared to obtain a good result. The accuracy of the classifications was assessed by using confusion matrix and decision boundary analysis. In conclusion, SVM could be used to classify oil palm age as it performs the best with highest overall accuracy of 91.89%. It is stable of limited amount and quality of training data. Further study can focus on hybrid techniques for age classification with accuracy assessment using ground truth image instead of ROI.

ABSTRAK

Usia kelapa sawit adalah pembolehubah penting yang telah digunakan dalam banyak kajian, seperti dalam kajian karbon dan ramalan hasil. Sebelum ini, pengelasan umur atau pengelasan kelapa sawit telah dibuat secara manual dengan pemetaan kawasan ladang. Teknik ini memakan masa dan sukar untuk organisasi seperti MPOB untuk mengelaskan kebesaran hektar. Dengan „remote sensing“, sifat memperoleh umur kelapa sawit melalui tindak balas spektrum adalah lebih mudah. Kurang kajian mengenai prestasi teknik semasa untuk klasifikasi umur kelapa sawit. Kebanyakannya adalah menggunakan pendekatan tradisional statistik parametrik. Selain itu, kajian sebelum ini untuk ramalan umur tumbuhan telah dijalankan menggunakan pendekatan dengan sensor satelit berlainan yang terdiri daripada resolusi dan pengukuran data yang berbeza. Projek ini menunjukkan prosedur / algoritma untuk mengklasifikasikan umur pokok kelapa sawit menggunakan LANDSAT-5 TM data. LANDSAT-5 TM adalah sensor yang merekodkan dalam 7 band spektrum. Kajian ini telah dijalankan dalam dua fasa; di mana fasa I adalah klasifikasi kawasan manakala fasa II adalah klasifikasi umur kelapa sawit. Pertama, kawasan telah dikenalpasti untuk membekalkan “training” dan “testing” piksel. ML merekod hasil yang cukup baik dengan ketepatan keseluruhan 85.69% . Manakala, tiga kaedah klasifikasi, seperti: ML, NN dan SVM digunakan untuk pengelasan umur kelapa sawit. Dua set set latihan, set latihan yang lebih kecil dan set latihan lebih besar dibandingkan untuk mendapatkan hasil yang baik. Ketepatan pengelasan dinilai dengan menggunakan matriks kekeliruan dan analisis sempadan keputusan. Kesimpulannya, SVM boleh digunakan untuk mengelaskan umur kelapa sawit dimana ketepatannya adalah yang tertinggi iaitu 91.89%. Ia boleh klasifikasi dengan jumlah dan kualiti data latihan yang terhad. Kajian lanjut boleh memberi tumpuan kepada teknik-teknik hibrid untuk pengelasan umur dengan penilaian ketepatan menggunakan imej kebenaran tanah dan bukan ROI.

TABLE OF CONTENTS

CHAPTER	SUBJECT	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGEMENTS	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	xii
	LIST OF FIGURES	xv
	LIST OF ABBREVIATIONS	xviii
	LIST OF ATTACHMENTS	xix
CHAPTER 1	INTRODUCTION	
	1.1 Project Background	1
	1.2 Problem Statement	3
	1.3 Objective	4
	1.4 Scopes	4

1.4.1	Software Scope	4
1.4.2	Data Scope	4
1.4.3	Area of Study Scope	5
1.4.4	User Scope	7
1.5	Project significance	7
1.6	Expected Output	8
1.7	Conclusion	8
CHAPTER 2	LITERATURE REVIEW AND ANALYSIS	
2.1	Introduction	9
2.2	Life of Oil Palm	10
2.3	Remote Sensing Interest	14
2.3.1	Review of remote sensing	14
2.3.2	Age Classification and Remote Sensing Interest	18
2.4	Supervised Classification	21
2.4.1	Supervised Maximum Likelihood	22
2.4.2	Support Vector Machine	23
2.4.3	Neural Network	25
2.5	Analysis	26
2.5.1	Problem analysis	27
2.5.2	Data requirement	27
2.5.3	Software and Hardware Requirement	28
2.5.3	Ancillary Data Requirement	28
2.6	Conclusion	29
CHAPTER 3	METHODOLOGY AND DESIGN	
3.1	Introduction	30
3.2	Phases	31

3.2.1	Phase One- Analysis Preliminary Study	32
3.2.2	Phase Two – Design	32
3.2.3	Phase Three - Experimental and Results	32
3.2.4	Phase Four- Testing/Evaluation	33
3.2.5	Phase Five- Conclusion	33
3.3	Project Schedule and Milestone	34
3.4	Designs on Oil Palm Age Classification	35
3.4.1	Input	37
3.4.2	Selecting ROI for Land cover	37
3.4.3	Generate Random Sample for Land Cover Classification	38
3.4.4	Maximum Likelihood Classification	38
3.4.5	Non-oil Palm Masking	38
3.4.6	Subset/Resize	39
3.4.7	Selecting ROI for Oil Palm Age	39
3.4.8	Generate Random Sample for Oil Palm Age Classification	40
3.4.9	Supervised Classification using ML,SVM and NN	40
3.4.10	Output	41
3.4.11	Accuracy Assessment/Testing, Compare and Analysis	41
3.5	Conclusion	41
 CHAPTER 4 IMPLEMENTATION AND RESULTS		
4.1	Introduction	42
4.2	Data pre-processing	42
4.2.1	Input data process	43
4.2.2	Region of Interest (ROI)	45

4.2.2.1	ROI Selection for Land Cover	45
4.2.2.2	ROI Selection for Age Class	46
4.2.3	Generate Random Sample	47
4.2.3.1	Random sample for Land Cover	47
4.2.3.2	Random sample for Age Class	48
4.2.4	Non-oil Palm Masking	49
4.2.5	Subset/Resize	49
4.3	Classification	50
4.3.1	Land Cover Classification using Maximum Likelihood	51
4.3.2	Oil Palm Age Classification using SVM, NN and ML	52
4.4	Output	54
4.5	Accuracy Assessment/Testing	55
4.6	Conclusion	56
 CHAPTER 5 ACCURACY ASSESMENT/TESTING		
5.1	Introduction	57
5.2	Confusion Matrix Accuracy Analysis	57
5.2.1	Comparison between ROI and ML Classification on Land Cover	59
5.2.2	Comparison between ROI and SVM Classification on Oil Palm Age	64
5.2.3	Comparison between ROI and NN Classification on Oil Palm Age	67

5.2.4	Comparison between ROI and NN Classification on Oil Palm Age	71
5.3	Decision Boundary Analysis	74
5.4	Conclusion	77
CHAPTER 6	CONCLUSION	
6.1	Introduction	78
6.2	Strengths	78
6.3	Weakness	79
6.4	Proposition for Improvement	80
6.5	Conclusion	80
	REFERENCES	81
	APPENDICES	85

LIST OF TABLES

TABLE	TITLE	PAGE
2.1	Varieties of Oil Palm	11
2.2	Description of LANDSAT TM Bands	16
2.3	LANDSAT TM Band Combinations	17
2.4	Software and Hardware requirement	28
3.1	Final Year Project Schedule and Milestone	34
3.2	Selection of ROI for Land Cover	37
3.3	Selection of ROI for Oil Palm Age	39
4.1	Sets of Random Samples	48
4.2	Total Covered Area of Each Class Using SVM NN and ML Classifier for Set A	54
4.3	Total Covered Area of Each Class Using SVM, NN and ML Classifier for Set B	55
4.4	Overall Accuracy of Each Training Set for Age Classification	56
5.1	Confusion Matrix in Pixels for ROI vs. ML	60
5.2	Confusion Matrix in Percentage for ROI vs. ML	61
5.3	Confusion Matrix in Commission and Omission; ROI vs. ML	62
5.4	Confusion Matrix in Producer Accuracy and User Accuracy for ROI vs. ML	62
5.5	Confusion Matrix in Pixels for ROI vs. SVM	65
5.6	Confusion Matrix in Percentage for ROI vs. SVM	65

5.7	Confusion Matrix in Commission and Omission for ROI vs. SVM	65
5.8	Confusion Matrix in Producer Accuracy and User Accuracy for ROI vs. SVM	66
5.9	Confusion Matrix in Pixels for ROI vs. NN	68
5.10	Confusion Matrix in Percentage for ROI vs. NN	68
5.11	Confusion Matrix in Commission and Omission for ROI vs. NN	68
5.12	Confusion Matrix in Producer Accuracy and User Accuracy for ROI vs. NN	68
5.13	Confusion Matrix in Pixels for ROI vs. ML	69
5.14	Confusion Matrix in Percentage for ROI vs. ML	69
5.15	Confusion Matrix in Commission and Omission for ROI vs. ML	70
5.16	Confusion Matrix in Producer Accuracy and User Accuracy for ROI vs. ML	70
5.17	Confusion Matrix on Ground Truth Image (Pixels) Between SVM vs. NN	72
5.18	Confusion Matrix on Ground Truth Image (Percent) Between SVM vs. NN	73
5.19	Confusion Matrix on Ground Truth Image (Commission and Omission) for SVM vs. NN	73
5.20	Confusion Matrix on Ground Truth Image between SVM vs. NN	73
5.21	Overall Comparison of All Matrixes	77
6.1	Results of Classification Commission Error of SVM, NN and ML	79

LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	Image of Landsat 5	5
1.2	Bukit Kerayong Oil Palm Estate	6
2.1	Statistic on Producer and Exporter of Oil Palm	10
2.2	Average of Productive Oil Crop	11
2.3	Growth of Oil Palm a) Young Seedling b) Seedlings with Bifid Leaf	12
2.4	Nursery at Bukit Kerayong	12
2.5	Design for Mature Oil Palms	13
2.6	Average Yields versus Age of Oil Palm	13
2.7	Electromagnetic Spectrum	14
2.8	Concept of classification of remotely sensed data	22
2.9	Linear support vector machine example	24
2.10	Generic Three-Layer Neural Network Structure	25
2.11	Land Cover Map	28
2.12	Oil Palm Plantation Map	29
3.1	Phases of Experiment	31
3.2	Oil Palm Age Classification Process	36

3.3	Flow of classification	40
4.1	Image in combination of VNIR (Visible near Infra-Red) (4) , red (3) , green (2)	43
4.2	Basic stats	44
4.3	Min, Mean, Mean and Standard Deviation of Input Image	44
4.4	Histogram of Frequency versus Data Value	45
4.5	10 ROI for Land Cover	46
4.6	Selection of ROI for four classes	46
4.7	Selection of ROI (a) Image with 40% training pixels, (b) Image with 60 % testing pixel	47
4.8	Sample (a) Image with 40% training pixels, (b) Image with 60 % testing pixel of set D	48
4.9	Non-oil palm masking	49
4.10	Subset of Oil Palm Region	50
4.11	Min, Mean, Mean and Standard Deviation of Input Image	50
4.12	Basic stats for Subset Image	50
4.13	ML classification	51
4.14	Basic statistic of ML Classified Image	51
4.15	Histogram of Number of Pixels versus Data Value	52
4.16	Classification of Oil Palm Age Using Set A	52
4.17	Classification of Oil Palm Age Using Set B	53
5.1	Image of (a) 60% ROI and (b) ML classification	59
5.2	Classified Image using (a) SVM, (b) NN and (c) ML Classifier	64
5.3	Comparison of Age Classification Using (a) SVM and (b) NN	71
5.4	Decision Boundary for SVM Classification	75
5.5	Decision Boundary for NN Classification	75
5.6	Decision Boundary for ML Classification	75

LIST OF ABBREVIATIONS

DMC	-Disaster Management Constellation
ENVI	-Environment for Visualizing Images
LAI	-Leaf Area Index
MCI	-Multi-Coefficient Image
ML	-Maximum Likelihood
AVHRR	-Advanced Very High Resolution Radiometer
MODIS	-Moderate-resolution Imaging Spectroradiometer
NN	-Neural Network
OBIA	-Object-Based Image Analysis
SVM	-Support Vector Machine
TM	-Thematic Mapper
MPOB	-Malaysian Palm Oil Board

LIST OF ATTACHMENTS

ATTACHMENT	TITLE	PAGE
A.1	Decision Boundary Coding for SVM	86
A.2	Decision Boundary Coding for NN	87
A.3	Decision Boundary Coding for ML	89

CHAPTER I

INTRODUCTION

1.1 Project Background

Satellite remote sensing data constitute a significant potential source of information on our environment, provided they can be adequately interpreted. Remote sensing is the observations and measurement of objects from a distance where objects or the recorders are not in contact under investigation. Remote sensing is one of the technologies used in image processing. Certain physical properties of objects are determined by remote sensing through the measurement of some kind of energy that is emitted; transmitted, or reflected from an object. Remote sensing has been used in many fields such as vegetation, meteorological analysis, inspection and prevention of geological disaster and military purposes. According to study by Bacour, vegetation has a major influence on the exchange of energy between the atmosphere and the earth's surface as it is fundamental element of earth surface (Bacour et al.2002). Remote sensing has been recognized as a reliable method for prediction of various biophysical and biochemical vegetation variables (Cohen et al. 2003). First report of remote sensing being implemented in agricultural management, although indirectly used, begun with the mapping of soil resources from aerial photography in 1929 by Kellogg. Remote sensing was possible for oil palm plantation with the improving in the quality and availability of remote sensing as oil palm has become the most important commodity crop in Malaysia .Based on an

article by UNEP Global Environmental Alert Service (GEAS), in Asian region, Malaysia and Indonesia, records the highest production where 85 per cent of global production of oil palm takes place there. 45% of the world's oil palm production is by Malaysia and it exports 80% of its total production. Malaysia is being under large scale plantation system operating as a nucleus of many smallholder producers (Butler, et al., 2009). Conventionally, the age detection or classification was made manually by mapping the plantation area. From my ground survey on 28th March, according to the plantation officer, they keep track of the plantation age by mapping up the plantation area manually. He did mentioned that this manually method sometimes cause inaccurate due to carelessness in data interpretation by the workers. This file system wasn't systematic and may cause loss of the data. Besides that, acquisition of information on a large area of oil palm plantation are also difficult and time consuming if the data are kept manually as the whole area might be owned by various plantation owners. With remote sensing technology, the nature of acquiring the oil palm age through spectral response is more convenient.

The economic vitality of oil palm crop requires accurate and timely of its agronomy for best management strategy. Fruit bunches which are harvested from oil palm trees being used to produce palm oil. One of the important factor of fruit bunches production is the age. Oil palm starts its production at the age of two years old. Its optimum production is at the age of six to ten years after planting. Age of oil palm is an important variable which has been used in many studies, like in carbon study and yield forecasting. In terms of carbon study, age is a variable in algometric equation for estimating oil palm biomass and carbon stocks. Oil palms also produce twice more oil than rape seed and almost four times more than soy beans, groundnut and sunflower per hectare per year (Tan, et al., 2009). Therefore, age prediction using remote sensing is seen to be vital. Usage of remote sensing could actually be useful for some of the parties such as Malaysian Palm Oil Board (MPOB) in term of prediction of annual yield as age of the oil palm highly correlate with the yield production. The data used in this study is satellite data. The satellite data come from several bands (multispectral) of Landsat-5 TM (Thematic Mapper). Images from these high resolution satellites can have 1m or sub meter pixel resolutions.

1.2 Problem Statement

The study of age classification using remote sensing technology offers a new alternative way over the conventional methods practised elsewhere. Vegetation age has commonly used in study of yield production, carbon stocks and monitoring of yields. The mapping and monitoring of vegetation biochemical and biophysical variables is important for spatially distributed modelling of vegetation productivity, evapotranspiration, and surface energy balance. Measurements of this feature of vegetation manually would be labor-intensive and costly, so it is practical on experimental plots using remote sensing. Conventionally, the age classification of oil palm was made manually by mapping the plantation area. This technique is time consuming and difficult to classify a large area of hectare which causes difficulty for organisation like MPOB to analyse on the yield production. Nature of acquiring the oil palm age through spectral response is more convenient. The study on the relationship of spectral measurements from LANDSAT 5 TM data with oil palm age is a little shallow. LANDSAT 5 TM is being used due to its capability of capturing spectral wavelengths sensitive to vegetation such as band 2, band 3 and band 4. Images from this high resolution satellite can have 1m or sub meter pixels resolutions. Limited studies were concerning the performance of current technique in classification of oil palm age. Mostly were using traditional parametric statistical approaches. Furthermore, the existing techniques are also being studied on different satellite sensors consisting of different resolution and measurements of data. Thus, the aim of this study is to compare the relationship of spectral measurements from Landsat-5 TM satellite data with oil palm age and develop procedure for vegetation classification. The study is also to propose a technique for oil palm age classification. This study will be helpful for some of the organizational boards to decide the actions to remedy low yield regions as one could predict the yield that supposed to obtain for a particular age.

1.3 Objective

In order to achieve aim of the study, the study embarks on the following objectives:

1. To compare the spectral measurements from satellite for oil palm age classification.
2. To propose a procedure for vegetation extraction and classification.
3. To propose a techniques for oil palm age classification using a LANDSAT-5 TM satellite data.
4. To verify the results, analyse and make accuracy assessment of the produced result.

1.4 Scopes

The scope of the study is divided into four parts: software scope, data scope, area of study scope and user scope.

1.4.1 Software Scope

Mainly, there are two software tools being used in this study. Firstly is the ENVI 4.5 toolbox which is used to process and analyse the image of LANDSAT-5 TM satellite data while the second would be MATHLAB R2010a for representing and analysing decision boundary.

1.4.2 Data Scope

The data that is used for this study is a **Landsat 5 Thematic Mapper (TM)** satellite data dated on 22nd August 2005. Landsat 5 TM was a low Earth orbit