HIERARCHICAL MARKOV RANDOM FIELDS FOR OBJECT-BASED IMAGE ANALYSIS

TRIYOGA WIDIASTOMO

March, 2016

SUPERVISORS:

Dr. V. A. Tolpekin
Prof. Dr. Ir. A. Stein

HIERARCHICAL MARKOV RANDOM FIELDS FOR OBJECT-BASED IMAGE ANALYSIS

TRIYOGA WIDIASTOMO

ENSCHEDE, THE NETHERLANDS, 2016

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics

SUPERVISORS:
Dr. V. A. Tolpekin
Prof. Dr. Ir. A. Stein

THESIS ASSESSMENT BOARD:
Prof. Dr. Ir. M. G. Vosselman (Chair)
Dr. Ir. B.G.H. Gorte (External Examiner, TU Delft)

ABSTRACT

Very high resolution satellite imagery have been increasingly used in many applications. Image analysis from a VHR satellite imagery requires information at object level to increase the accuracy. Classic pixel-based image classification methods only work based on the pixel information of the image and neglects the contextual elements (e.g., size, shape, texture, etc.). Therefore, this research is focused on developing a method that includes the contextual information as an input for image classification. Performing image classification at object level is a challenging task. Hierarchical Markov random field (MRF) approach is used in this research to incorporate contextual information of objects as prior information.

Hierarchical MRF accommodates the integration of pixel-level and object-level image classification. At the first level, pixel-based MRF allows the image classification based on the prior and likelihood energy function. The prior energy function is controlled by smoothness parameter considering the neighbourhood system. The likelihood energy function contains the maximum likelihood classifier. The energy is optimized by simulated annealing (SA) technique. At the second level, object-based MRF employ size and texture information in the prior energy function. This method is applied over an oil palm plantation area at two different spatial resolution, which is the 0.5 m and 2 m resolution.

The application of hierarchical MRF should improve the accuracy of the result by incorporating the contextual information. In this research, the hierarchical MRF is still difficult to employ the contextual information to each segment. This is proved by the appearance of improper size of segments in the results. The results obtained the highest quality in identifying oil palm tree crown compared to MLC and pixel-based MRF. The MLC method leads to a higher quality than pixel-based MRF. Even though, by employing the pixel-based MRF on the pan-sharpened image, the result achieved higher kappa κ values.

Key words: *Hierarchical Markov random fields, Object-based image analysis, Image segmentation, Class spectral separability, Simulated annealing, Grey level co-occurrence matrix, Oil palm tree crown.*

ACKNOWLEDGEMENTS

Al-hamdu lillahi rabbil 'alamin, all the praises belong to Allah SWT.

First of all I wish to express my utmost thanks and gratitude to my first supervisor, Dr. V. A. Tolpekin, who has always provided me a support in thinking and developing the ideas and implementations of this research. I would also wish to offer my thanks and gratitude to my second supervisor, Prof. Dr. Ir. A. Stein, for the fruitful discussion and input.

I want to express my deepest gratitude to my parents and my brother in Jakarta and also my brother in Delft. Thanks for the formidable support and help for my study here and throughout my life. It will be an honour for me to dedicate this work to my parents.

I would like to thank people that I have met here, especially to my Indonesian friends in Enschede. Thanks to my GFM classmates, for the time we shared together.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Motivation and problem statement

Remote sensing data provide a huge amount of information about earth surface that can be used in many applications related to urban scenes, monitoring of forests, disaster management, etc. Over the past decade, very high resolution (VHR) satellite imagery, an image at a spatial resolution below one meter, have been increasingly used in various applications. Various commercial VHR satellite images, such as IKONOS, QuickBird, GeoEye-1, WorldView-2, WorldView-3 and Pleiades, are sufficient to extract small features of some land cover types (Shackelford & Davis, 2003).

In current years, the development of image processing methods for the VHR satellite imagery become one of the most serious issues for the remote sensing researchers. Many studies focus on developing automatic data processing method for VHR satellite imagery. Automatic data processing method is essential for reducing the time of processing data and minimizing the inconsistency of the result of human interpretation. The methods are developed to reach a better analysis and interpretation result than a human operator.

Although VHR satellite imagery provides more detailed features than the moderate and low resolution, it may also lead to some problems. As the spatial resolution gets higher, the spectral variation within one class of land cover type rises (Salehi et al., 2011). High spectral variation may lead to a lower accuracy of classification result. This problem may occur in the pixel-based classification. This traditional classification method is not perfectly suitable for vegetation mapping that may have either high spectral variations within the same class or similar spectral value among different classes (Sha et al., 2008).

Geographic object-based image analysis (GEOBIA) has been widely suggested to be one of the solutions to handle the problems mentioned above. GEOBIA works more efficiently than pixel-based approach for higher resolution satellite imagery (Hay & Castilla, 2008). This approach works by grouping neighbouring pixels by a segmentation technique. In extracting feature from VHR satellite imagery using GEOBIA, some of the key features are size, shape, tone and texture.

In this research, Markov random field (MRF) model based algorithm is proposed. Using an MRF model, a spatial correlation between neighbouring pixels can be defined. It is assumed that the configuration of a pixel given the configurations of the entire image is equal to the configuration of a pixel given the configurations of its neighbouring pixels (Tso & Mather, 2009). In the MRF approach, the neighbours pixels have a higher probability to be classified in the same class than different classes (Kasetkasem et al., 2005). In this research, object-based MRF is applied in combination with pixel-based image classification. To do so, hierarchical MRF is proposed in this research. This method has been proved in object-based image analysis, such as image segmentation, texture modelling (Graffigne et al., 1995).

The current highest spatial resolution of VHR satellite imagery is a panchromatic band image at 30 cm and multispectral bands at 1.2 m resolution from WorldView-3. This is a typical spatial resolution specification for VHR images, which the panchromatic band has a higher resolution than other bands. Nowadays, many types of research require sub-meter spatial resolution imagery. To get a coloured image at finer spatial resolution, the image pan-sharpening technique can be applied. This technique fuses panchromatic image at higher resolution with multispectral images at lower resolution. This fusion method is suitable for visual interpretation but not for quantitative analysis because it introduces spectral or geometric distortions. Super-resolution mapping (SRM) is a technique to produce a map at a finer spatial resolution than the spatial resolution of the image source using information from soft classification (Tatem et al., 2002). This method is one of the possible methods to be incorporated in MRF model.

The aim of applying each of the techniques described above is to develop a hierarchical MRF for object-based image analysis (OBIA). In this research, the method will be applied for detecting and delineating oil palm trees. To the best of my knowledge, none of the previously developed methods combine the pixel-based and object-based analysis using hierarchical MRF and applied it for oil palm trees inventory application. Applying object-based image analysis for oil palm trees mapping may give some advantages since oil palm trees have some object features that different from another landcover type in a plantation.

Oil palm trees usually planted in a plantation with around 8 meter spacing from one tree to another. The diameter of the smallest oil palm tree crown is about 1 meter. Detecting and delineating oil palm trees using the traditional image classification method is difficult to get a high accuracy because the spectral variation within a class is larger than variation between the classes and some tree crown has a small size compared to the pixel size. Furthermore, this method may also be applied for another application with similar problems.

1.2. Research identification

The primary subjects focused in this research project can be defined through the following research objectives and research questions.

1.2.1. Research objectives

The main objective of this research is to apply and develop hierarchical Markov random fields for object-based image analysis in VHR satellite imagery and apply it to detect and delineate oil palm tree crown in oil palm plantation. The following are the sub-objectives:

- To design MRF probability model: prior, likelihood and posterior.
- To combine pixel level and object level image analysis using hierarchical MRF.
- To evaluate the performance of the applied method and compare the results with other methods.

1.2.2. Research questions

The research questions of this research are:

- How to design MRF probability models?
- How can pixel level and object level be combined using MRF?
- How combinations of smoothness parameter for pixel and segment affect the accuracy?
- How the most appropriate parameters for MRF should be determined?
- How the assessment of the results should be performed?
- Which size of oil palm tree crowns can be successfully identified?

1.3. Innovation

The novelty of this research are:

- Development of image analysis algorithm that combines pixel level and object level image analysis using hierarchical MRF.
- Application of hierarchical MRF object-based image analysis for detecting and delineating oil palm trees from very high resolution satellite imagery.

2. LITERATURE REVIEW

In this chapter, a review of the literature relevant to this thesis is given to obtain the required understanding for this research. The first section reviews the methods used and developed for image segmentation. The object-based image analysis is explained in the second section. The section 2.3, section 2.4 and section 2.5 presents the approaches related to this research. The last section presents the related works done for oil palm trees detection.

2.1. Image Segmentation

In image analysis context, image segmentation is a method to divide an image into relatively homogenous regions (Blaschke, 2010). A good image segmentation allows a uniform and homogeneous region and adjacent regions should have a contrast with respect to their characteristics and divided by a simple boundary (Haralick & Shapiro, 1985). Many techniques are available to implement the image segmentation, such as neural network architecture, MRF model, grey level histogram, region growing, and fuzzy set theoretic approaches.

Adams and Bischof (1994) developed a robust, rapid, and free of tuning parameters algorithm of seeded region growing (SRG) for image segmentation. The growing region is initialized by individual pixels or regions and then merging pixels into their adjacent seed region. Shih and Cheng (2005) applied an automatic SRG for a colour image segmentation. In their research, the seeds are chosen automatically. The result showed a satisfactory result compared to other algorithms.

Benz et al., (2004) employed multi-resolution fuzzy analysis of remote sensing data. They explained the method to combine the fuzzy methods and multi-resolution technique at the object level. The multi-resolution technique is able to form an extremely homogeneous objects. This technique requires the scale parameter to control the segmentation process Drăguţ, at al. (2014)

Senthilkumaran and Rajesh (2009) studied the edge detection technique for image segmentation by applying the soft computing approaches. They compared three soft computing approaches, which are fuzzy logic, genetic algorithm and neural network. The results indicated that the three approaches worked well in different particular condition. Bellon and Silva (2002) developed an algorithm to improve the image segmentation by edge detection. The algorithm is able to preserve the object topology and independent to noise.

2.2. Object-based image analysis

Object-based image analysis (OBIA) is a technique that considers the characteristics (spatial, spectral and temporal) of objects in segmenting an imagery (Hay & Castilla, 2006). OBIA has been introduced to overcome the limitations of the old-fashioned pixel-based image analysis that true geographical objects are missing and that contextual information is ignored. The main aim of using OBIA in remote sensing is to adapt the human interpretation of objects in providing accurate automatic / semi-automatic image analysis. Because of its advantages, OBIA has been widely used for many applications over the past decade.

Myint et al., (2011) compared the result of pixel-based and object-based urban landcover classification from VHR image. They used the nearest neighbour classifier to classify vegetation classes, which are grass and trees/shrubs while membership neighbour classifier was used in classifying other class using different parameter values used for every class. The result showed that object-based classification gives a significantly better accuracy than pixel-based classification.

In OBIA approach, an image is subdivided into homogeneous regions based on both spectral reflectance values and its contextual information (shape, size, texture, etc.). This information hierarchically generated a network of image segments (Blaschke, 2005). Image segmentation process plays an important role in OBIA approach. Features are extracted with respect to their characteristics to distinguish one region from other different regions. Pixels that have the same characteristics will be grouped. Yu et al., (2006) quantified several features, including spectral features, textures, topographic features, and geometric features, for each object to classify a detailed vegetation map from an aerial imagery. In their research, a comprehensive information is highly required for the object features.

Two essential characteristics of pixels in satellite image are tone and texture (Zhang, 1999). The tone is simply the average grey level of a region while the texture is the spatial distribution pattern of grey levels in a region. In some cases, image segmentation using tone doesn't work properly because the average levels of different regions are similar to each other. Therefore, the texture segmentation algorithms are required to get a better result. Ryherd and Woodcock (1996) combined both spectral value and texture information from a medium spatial resolution satellite imagery to classify an area based on the trees density. The result shows that although incorporating texture gives a better result, the size information for each object is needed to improve the accuracy.

Mallinis et al. (2008) found the OBIA approach for delineating forest vegetation using VHR imagery. A multi-resolution segmentation was followed by a classification tree and compared the result with a nearest neighbour classifier. Peña-Barragán et al. (2011) incorporated several vegetation indices with textural features and crop phenology to achieve an object-based crop identification using medium resolution imagery. The textural feature can distinguish between permanent crops. However, the vegetation indices present around 90% in the model.

The texture is one of the important characteristics in OBIA. It gives information of tonal variation spatial distribution in an image (Haralick et al., 1973). The spatial relationship of pixels in a local neighbourhood is described to identify different objects. The texture may present different structures, which are periodic, semi-periodic and random.

One of the most common methods for texture analysis is grey level co-occurrence matrix (GLCM). GLCM measure the textural characteristic of an image by specifying the spatial relationship of the measured different combinations of grey levels. The number of grey levels occurs in the image determine the size of GLCM. The relationship between pixels may be described by considering the difference of direction and distance. Haralick et al., (1973) successfully used a statistical method to compute textural features in an image by retrieving the information of spatial relationship of pixel values in GLCM. In the context of trees analysis, GLCM has been proved to become one of the methods in classifying trees species using multispectral VHR aerial imagery (Ramezani, 2015).

2.3. Maximum likelihood classification

MLC is one of the most common supervised classification technique based on Bayesian probability (Richards, 2012; Tso & Mather, 2009). It assumes that each class in an image is normally distributed. In principal, maximum likelihood (ML) classifier assign pixels to a class that has the highest probability of membership (Maselli et al., 1994). Strahler (1980) included the prior probabilities to ML classifier to calculate the posterior probabilities of class membership which combine the pixel value and the weight of the class. In order to improve the performance of MLC, Ediriwickrema and Khorram (1997) developed the hierarchical MLC by estimating the prior probabilities to the pixels that probably did not properly classified. Despite its limitations, the performance of MLC is comparable to other methods, such as support vector machines (SVM) and decision tress (DT). Otukei and Blaschke (2010) compared the result of MLC, SVM

and DT in assessing landcover change. The result shows that all of the methods achieved more than 85% accuracy. Although, the DT obtained the highest result.

2.4. Hierarchical MRF

Markov random field (MRF) is a stochastic approach to model contextual information by includes both the prior and posterior distribution on the original imagery (Geman & Geman, 1984). In MRF model, every image is assumed to have Markov properties, which is based on spatial dependence. This assumption is suitable for pixels in an image of earth surface where two adjacent pixels are more likely belong to the same landcover class. Based on this assumption, the isolated pixels are likely to be disappeared to allow homogeneous region.

Maximum a posteriori (MAP) is one of the most used methods as a statistical criterion for MRF. The joint posterior probability of MRF labels is the objective of MAP. The MAP-MRF models derive the posterior distribution, determine the parameters and design optimization algorithms (Li, 2009). In MAP solution, when both prior distribution and likelihood function are known, the best that can be estimated is Bayes labelling. In Bayes labelling of MRF, the posterior distribution of an MRF become an important step. The posterior probability can be derived from prior and likelihood energy.

The MRF can explain the dependence of the grey level of a pixel in an imagery on the grey level of its neighbouring pixels (Cross & Jain, 1983). This dependency can be described, such as in first order neighbourhood systems where every pixel has four neighbours or in second order neighbourhood systems where every pixel has eight neighbours. Kato and Pong (2006) combined colour and textural feature using MRF to obtain an image segmentation model. The MRF defined in the nearest neighbourhood system.

Another application of MRF in image analysis is by applying hierarchical MRF. Yamazaki and Gingras (1995) modelled a hierarchical MRF that includes the observed intensity process and the hidden class label process. They used MAP criterion to estimate class label. Hu and Fahmy (1992) used hierarchical MRF to develop a new texture segmentation technique. They combined two particular MRF models. One is the multi-level logistic (MLL) model, which is an MRF model with conditional probability. The other is the binomial model for modelling textures inside the regions. The binomial model is for discrete random variables. This model is fit for satellite imagery segmentations because the grey levels in an image are discrete random variables. The hierarchical MRF is also suitable for synthetic aperture radar (SAR) image. Voisin et al., (2013) employed a quad-tree model to combine both amplitude SAR data and textural information in a hierarchical MRF model for urban area classification.

2.5. Super-resolution mapping

Aerial vehicle and space satellite are two most common platforms for acquiring remote sensing imagery. Both platforms have different specifications in acquiring imagery. One of the biggest differences is the spatial resolution of the acquired imagery. Nowadays the spatial resolution of imagery acquired from a satellite is lower than imagery acquired from an aerial vehicle. Spatial resolution plays an important role in landcover mapping using imagery data. A lower spatial resolution may contain mixed land cover class and a significant amount of mixed pixels may affect the result of image classification (Fisher, 1997). Therefore, sub-pixel classification may be one of the approaches to obtain higher accuracy of image classification result.

Super-resolution mapping (SRM) is a technique for producing thematic map at a finer spatial resolution than that of input data from remote sensing imagery (Atkinson, 2009). This technique works under the assumption of spatial context. The main focus in SRM is to estimate the location of sub-pixel level within a pixel. This may require a lead from the soft classification that used to estimate the class proportion of image pixels. This is only become a lead since the spatial distribution within a pixel is still unknown (Muad, 2011).

Soft classification methods are developed to classify an image at a sub-pixel level. Some of the soft classification techniques are linear mixture model, maximum likelihood classification (MLC), fuzzy c-means and support vector machines (SVM).

Some SRM algorithms that applied to the soft classification output require iterative optimization. A fine map generated from coarse output classification map. Then, it iteratively optimized by algorithms to get a satisfactory result. Mertens et al. (2003) employed genetic algorithms (GA) to obtain a super-resolution (SR) map from soft classification output. GA operated to find an optimal result by a set of solutions. Villa et al. (2011) proposed simulated annealing as a spatial regularization to achieve an SR map. HNN also applied to obtain SR map by maintaining the class proportion information of the soft classification (Muad & Foody, 2012).

SRM can be applied to both the input and the output as a post-processing of soft classification directly. Tatem et al. (2001) applied Hopfield neural network (HNN) as an energy minimization tool to obtain final classification map from a preliminarily classified map. Kasetkasem et al. (2005) applied the SRM technique directly to the input imagery data based on an MRF model. Ardila et al. (2011) considered a conditional probability function of both multispectral and panchromatic bands of raw VHR imagery in applying MRF-based SRM. Tolpekin and Stein (2009) used class proportions of a coarse input imagery to study the effects of class separability on the accuracy of MRF-based SRM. Su et al. (2012) combined pixel swapping and contouring techniques to obtain SR map. The class boundary was represented by smoothing the pixel swapping output to draw the boundary of classes from the contour of a similar class member.

2.6. Automatic Oil Palm Trees Detection

In the field of extracting information of oil palm trees from satellite imagery, works have developed to detect oil palm tree crown and estimating oil palm tree age by using satellite-based imagery. Oil palm trees usually grow in a large plantation area. Oil palm tree crown has some characteristics that can be differentiated from other objects in a plantation area, which are round shape, green colour and diameter generally ranges from sub-meter to 14 meters. Srestasathiern and Rakwatin (2014) conducted a research on oil palm tree detection with high-resolution multispectral satellite imagery. They hypothesized that palm trees are located at the local peak within the oil palm plantation. The rank transformation was performed to the index image to enhance the separability between tree crowns and other objects. The non-maximal suppression algorithm was used to detect the local peak. The appropriate window size was determined using semi-variogram analysis. Chemura et al. (2015) studied an approach for determination of the age of oil palm tree from high-resolution multispectral satellite imagery. They combined very high-resolution multispectral imagery data and regression techniques using a study case of Ejisu-Juaben district of Ghana. The relationship between age and crown projection area of oil palms was determined. The hierarchical classification using OBIA techniques on VHR satellite imagery was applied to determine the crown projection areas of oil palms. They combined crown projection areas obtained from the hierarchical classification and regression model developed from the field survey.

The studies explained above have its own potential to be applied in a particular condition in an image. This research is focuses on developing image segmentation technique by combining pixel level and object level image analysis will be plugged into a hierarchical MRF model. At the pixel level, the analysis consists of exploring the neighbourhood system and the ML classifier. At the object level, the area and texture information are included to allow object-based image classification. The SRM technique may be incorporated into the hierarchical MRF model to achieve a sub-pixel image classification.

3. STUDY AREA AND DATA

This chapter presents a brief description of the study area, remote sensing data and their pre-processing. Section 3.1 describes the study area, while Section 3.2 and 3.3 describes the satellite image data and reference data respectively. The section 3.4 explains briefly the software used in this research.

3.1. Study area

The study area is an oil palm plantation approximately located at 0° 58' 47.01" N and 118° 2' 6.57" E in Kutai Timur Regency, Kalimantan Timur Province, Indonesia. This regency is known as one of the leading areas in producing crude palm oil in Indonesia. The study area is a rectangular area of 5 km by 5 km which covers oil palm trees from various years of planted and diverse terrain. Figure 3.1 shows an overview of the study area.

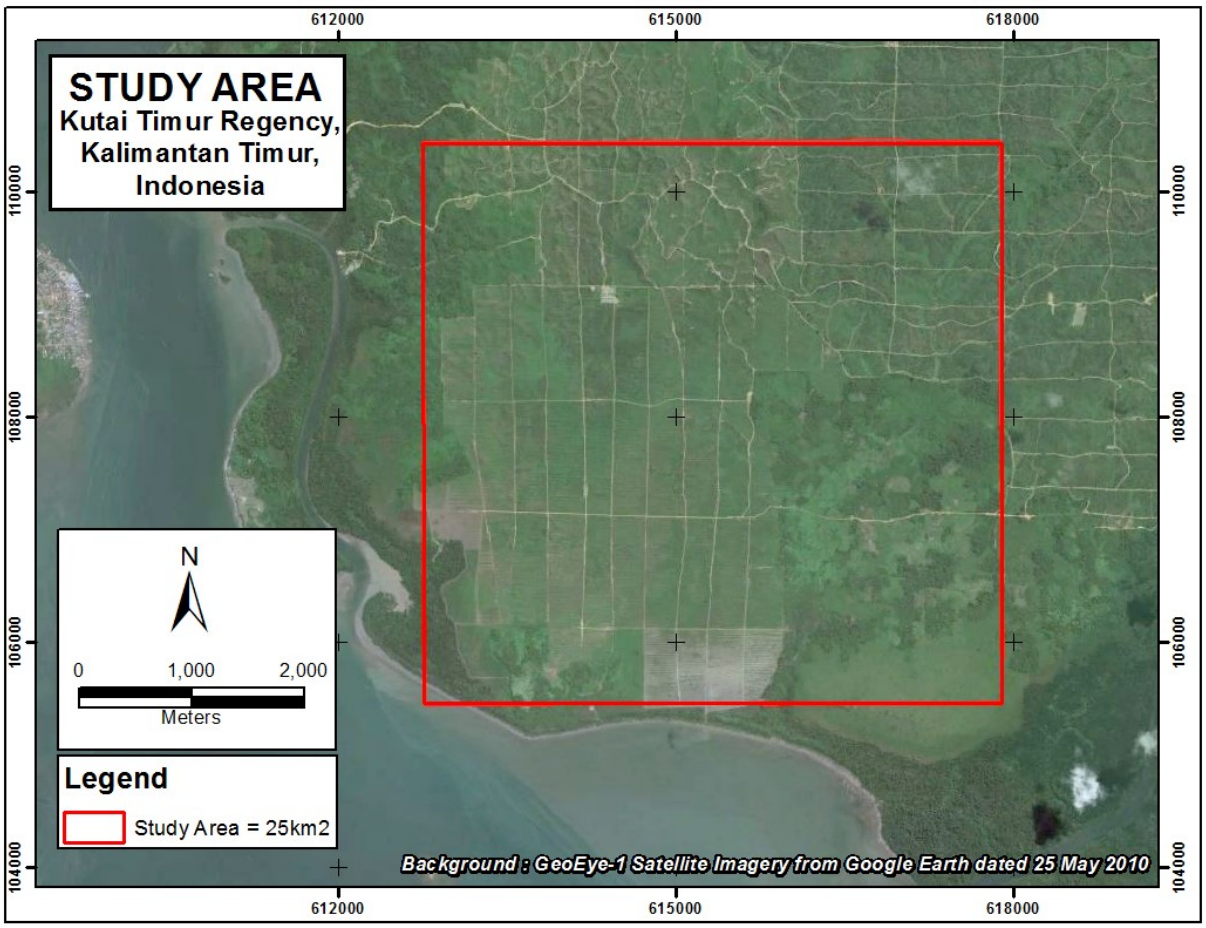

Figure 3.1: Overview of the study area

3.2. GeoEye-1 Satellite Imagery data

The GeoEye-1 satellite has multispectral and panchromatic bands with spatial resolutions of 2 m and 0.5 m respectively. The multispectral bands consist of blue (450 - 510 nm), green (510 - 580 nm), red (655 - 690 nm) and near-infrared (780 - 920 nm) spectral bands while the panchromatic band captures 450 - 900 nm part of electromagnetic spectrum. The data used in this research were acquired on 25 May 2010. It covers

25 km² area of oil palm plantation. In this research, the subset area used for acquiring the training and test dataset is 400 m². The subset of the image is shown in Figure 3.2. The image is considered as a good image for image classification since it has very low percentage of cloud cover at around 8%. It has also a moderate off nadir-angle at 16.68° and sun elevation angle at 59.85°. This can be seen in the image where some shadows occur in this image. Shadow may confuse the image classification process because it covers the true land cover lies on the ground. In this study, the shadow was classified as a class.

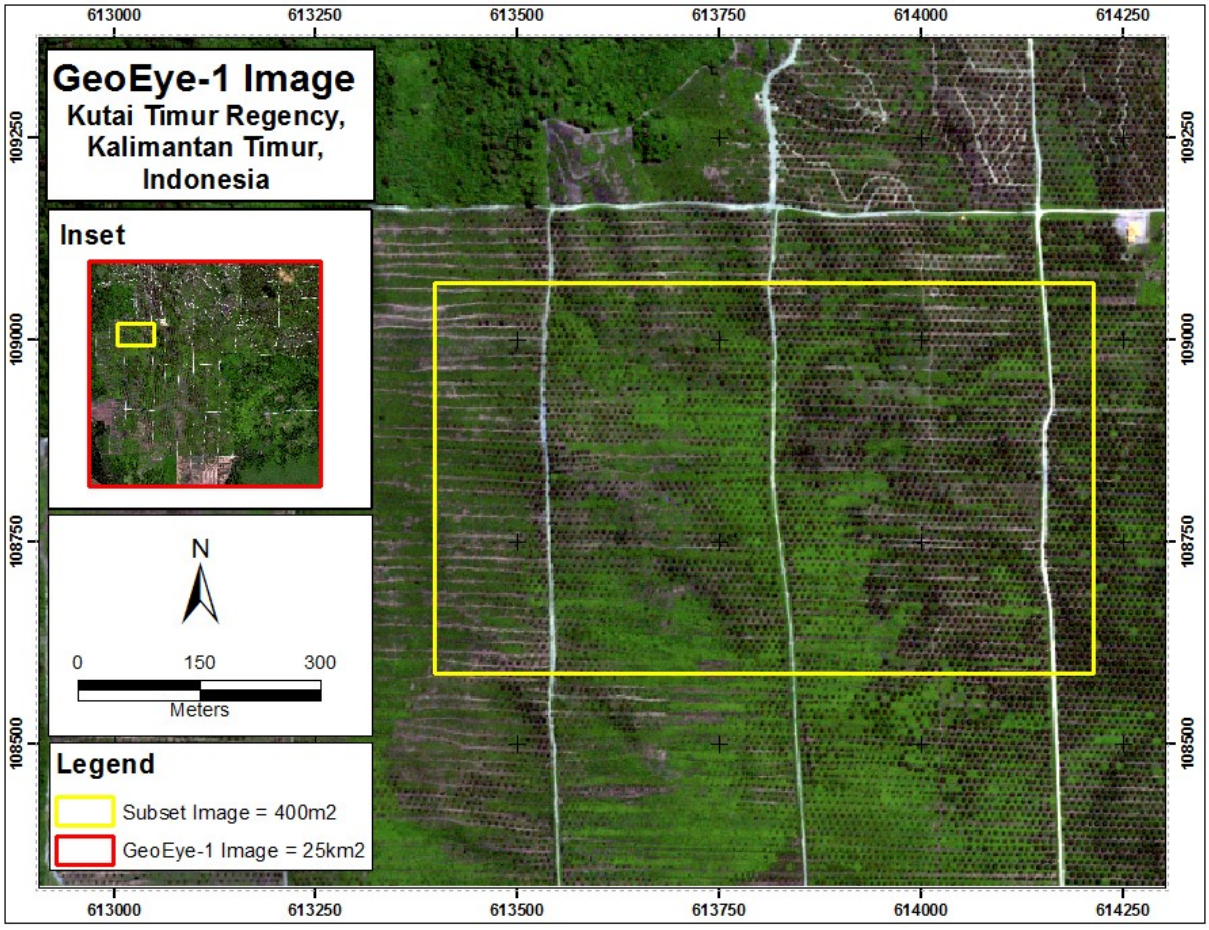

Figure 3.2: GeoEye-1 Satellite Imagery in 2 m natural colour mode

3.3. Reference data

Reference data is required in this research to implement accuracy assessment for the final result. The accuracy assessment measures how well the performance of the applied method compared to the reference data. The reference data is a vector layer of tree crowns that was derived by digitization of tree crowns from 0.5 m pan-sharpened GeoEye1 as shown in Figure 3.3.

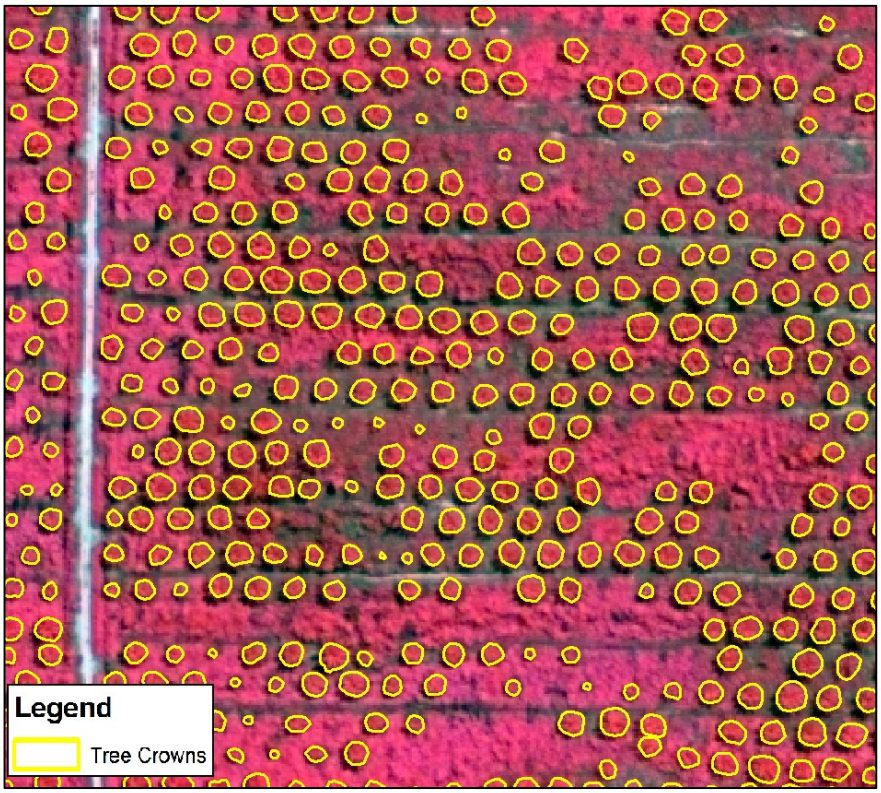

Figure 3.3: Example of reference vectors for tree crown with false infra-red colour background

3.4. Software

In this research, different software was used for applying the proposed methodology and data visualization. ArcGIS 10.3.1 was used to delineate the training sets and test sets from the VHR image, calculating area size of tree crowns and map visualization. Erdas Imagine 2015 was used to make the image subset and pan-sharpening the VHR image. Both ArcGIS 10.3.1 and ERDAS Imagine 2015 are well known commercial software for GIS and remote sensing. R version 3.2.2 software was used for implementing the method developed in this research and visualize the result. R software is an open-source software that able to compute and visualize statistical data developed by R Core Team (2015). More specifically, the packages used in this study is provided below:

- rgdal developed by Bivand et al., (2015)
- MASS developed by Venables and Ripley (2002)
- e1071 developed by Meyer et al., (2015)
- kernlab developed by Karatzoglou et al., (2004)
- rgl developed by Adler et al., (2016)
- Rcpp developed by Eddelbuettel and Francois (2011)

4. METHODOLOGY

4.1. Spectral classes definition

A supervised classification requires the definition of spectral classes and spectral statistics. In this context, training sets were used for estimation of class mean and covariance values of the imagery, while the test set was used for statistical analysis of the image classification accuracy. All the pixels in the imagery were labelled to represent the classes. To achieve a fair estimation of the elements of the class conditional mean vector and covariance matrix, a sufficient number of training pixels for each spectral class is required. In this research, the total number of training pixels is 800 pixels for the multispectral image and 12,598 pixels for the pan-sharpened image, which spread over the subset image as shown in Figure 4.1.

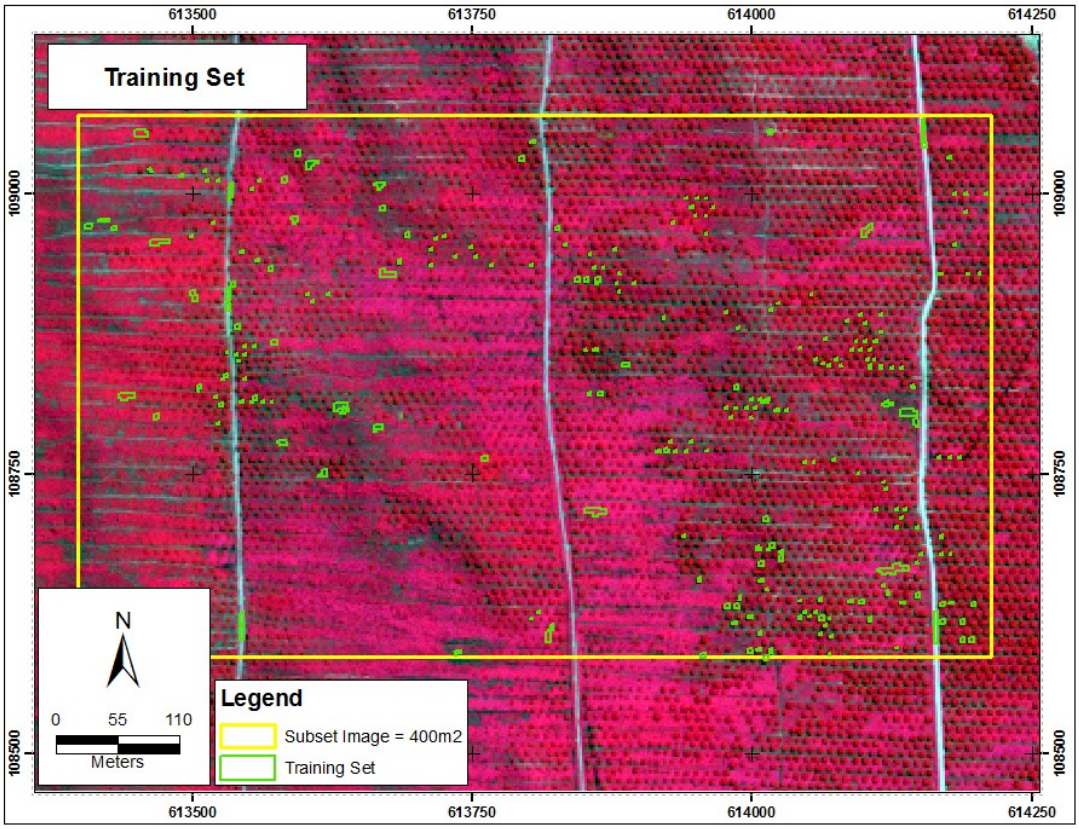

Figure 4.1: Training set inside subset image in false infra-red colour mode

Based on the visual interpretation of the image, five land cover classes were defined. The main focus of this study is the tree crown, but those classes should be defined with respect to the spectral value of all pixels in the imagery. The landcover classes consist of *trees*, grass, *unpaved road*, *bare soil*, and *shadow*. Considering the spectral value based on visual interpretation, there are some slight variations within tree crown class and there are some confusions between tree crown and *grass*. So, the training sets should be defined carefully to obtain a reasonable result. The training set was made in both multispectral and pan-sharpened image. The later was used as the reference for delineating the boundary since it has a finer spatial resolution.

The values of one band against another were compared using a feature space. It shows a scatterplot with a dot for every pixel in the pixel. The pixel position in the feature space image is defined by the spectral values for the two chosen bands in the 2D feature space. Each dot in the feature space has a colour associated with

its class type. 3D feature space compares values of three bands in a three dimensional graph. The feature space of each class represented in an ellipse with a colour associated with its class type. A well separable class is one of the conditions that should be meet in order to achieve a high accuracy of classification. A significant overlap within classes should be prevented in the class feature space. A classification process would be less likely to produce an error in distinguishing class label if the overlap is small (Richards, 2012). In classifying remote sensing imagery, the spectral values within one class should be close while the spectral values of different classes should be well separated.

Divergence, transformed divergence (TD), Bhattacharyya distance and Jeffries-Matusita (JM) distance are widely used to quantify class separability. Divergence is measured using the definition of the likelihood ratio. In Divergence, a small increase will lead to a much better classification accuracy. TD is used to minimize the effect of well-separated classes that may increase the average divergence value. JM distance is used to transform Bhattacharya distance. The JM distance has a tendency to suppress high separability values while overemphasizing the low separability ones (Gunal & Edizkan, 2008). It provides a trustworthy result since it performs more like the probability of correct classification (Swain & King, 1973). Tolpekin and Stein (2009) studied the effects of class separability on the accuracy of MRF-based SRM on a synthetic image. The values of class separability and scale factor are highly related to the optimal value of smoothness parameter, which resulted from the highest accuracy of classification.

The class spectral separability was evaluated using the Jeffries–Matusita distance (JM) in both multispectral and pan-sharpened mode. Equation 4.1 shows the Bhattacharyya distance. According to the equation 4.2, the JM distance transforms the Bhattacharyya distance to a value ranging from 0 to 2. A distance of two spectral values with JM distance close 2 indicates a high degree of accuracy in classifying a pixel into one of those classes. This condition is preferred in class separability.

$$B_{\alpha\beta} = \frac{1}{8}\left(\mu_\alpha - \mu_\beta\right)^T \left(\frac{C_\alpha + C_\beta}{2}\right)^{-1} \left(\mu_\alpha - \mu_\beta\right) + \frac{1}{2}\ln\left(\frac{\left|\frac{C_\alpha + C_\beta}{2}\right|}{\sqrt{|C_\alpha|.|C_\beta|}}\right) \tag{4.1}$$

$$JM_{\alpha\beta} = 2\left(1 - e^{-B_{\alpha\beta}}\right) \tag{4.2}$$

4.2. Markov random field

The contextual dependencies of image pixels can be modelled by MRF (Li, 2009). Let x be a random field on Y with random variables $\{x_1, x_2, \ldots, x_m\}$ with regards to a neighbourhood system N. A random field that considers its neighbours is a MRF if the following three MRF properties are satisfied by its probability density function (Tso & Mather, 2009):

1) Positivity: $P(x) > 0$ for all possible configurations of x. In most cases, this condition is satisfied. The joint probability $P(x)$ is driven by local conditional properties (Besag, 1974).
2) Markovianity: $(x_i|x_{Y-i}) = P(x_i|x_{Ni})$. This property indicates only neighbouring pixels have a direct interactions with each others.
3) Homogeneity: $P(x_i|x_{Ni})$ is the same for all sites i, regardless of the relative position of the site.

A neighbourhood system plays an important role in MRF. The first-order neighbourhood system has four pixels around the central pixel as shown in Figure 4.2 (a). Second-order neighbourhood system has four corner boundaries of the central pixel as shown in Figure 4.2 (b). A clique is a part of a neighbourhood system. It can be a single site, a pair of sites or triple of neighbouring pixels as shown in Figure 4.2 (c). In this research, a pair of sites clique is taken into consideration. Horizontal, vertical and diagonal neighbours are included in prior energy function in MRF model.

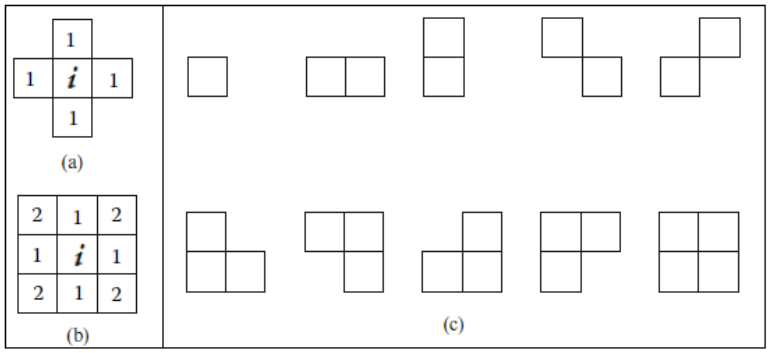

Figure 4.2: (a) First-order neighbourhood system; (b) Second-order neighbourhood system; (c) Cliques corresponds to the neighbourhood system (source: Li, 2009)

The hierarchical MRF is used to apply two levels image classification technique. The first level is at the pixel level and the second level is at the object level. Both levels work in parallel to produce an object-based classified image. At the pixel level, MRF is used to incorporate the contribution of a local neighbourhood to a pixel. MRF at the object level includes the contribution of area and texture to allow object-based image classification.

4.3. MRF at pixel level

The MRF model at the pixel level in this research follows the research by (Ardila et al., 2011). In this setup, they allow SRM technique in the MRF model by including scale factor S. An image classification process using SRM considers multispectral image y that consists of K bands at spatial resolution R, this corresponds to a square area in the ground of size R^2. The pixel locations $b_i \in B$, where B is the $M_1 \times M_2$ pixel matrix. The super-resolution map (SR map) c is a classified map at finer spatial resolution $r < R$ resulted from SRM method. The scale factor $S = R/r$ is considered as an integer. Each b_i covers the same area on the ground as S^2 finer resolution pixel $a_{i|j}$. A is a pixel matrix with size $(SM_1) \times (SM_2)$. The SR map c is defined on the set of pixel locations A and cover the same area as multispectral image y. In order to obtain the nominal scale for the result, $S = 1$ is applied in this research.

Multispectral image x is assumed to have K bands at spatial resolution r on the set of pixels A. Image x is not observed by satellite, while image y is observed by satellite and is considered as spatially and spectrally degraded observations of image x. Furthermore, it is assumed that each pixel in image x can be assigned to a unique class $c(a_{j|i}) = \alpha, a_{j|i} \in A$. The relationship between image y at pixel b_i and x can be established by the degradation model:

$$y(b_i) = \frac{1}{S^2}\sum_{j=1}^{S^2} x(a_{j|i}) \tag{4.3}$$

A classified image is modelled by using MRF that specifies a prior probability model in the following:

$$P(c) = \frac{1}{z_1}\exp\left[-\frac{U(c)}{T}\right] \tag{4.4}$$

$$P(y|c) = \frac{1}{z_2}\exp\left[-\frac{U(y|c)}{T}\right] \tag{4.5}$$

$$P(c|y) = \frac{1}{z_3}\exp\left[-\frac{U(c|y)}{T}\right] \tag{4.6}$$

Where,

$U(c)$ = prior energy function of map c
$U(y|c)$ = likelihood energy function to multispectral image y given the true SR map c
$U(c|y)$ = posterior energy function of SR map c given the multispectral image y
T = constant termed temperature
z = partition function = $\sum_{all\ possible\ combinations\ of\ c} exp\left[-\frac{U(c)}{T}\right]$

According to Bayes theorem

$$P(c|y) \propto P(c)\,P(y|c) \tag{4.7}$$

The energy functions can be expressed by

$$U(c|y) = U(c) + U(y|c) \tag{4.8}$$

4.3.1. Prior energy function

Considering the pair of sites clique for the neighbouring system, the prior energy $U(c)$ is expressed by:

$$U(c) = \sum_{ij} U\left(c(a_{j|i})\right)$$
$$= \sum_{ij} \sum_{l \in N(a_{j|i})} w(a_l) I\left(c(a_{j|i}), c(a_l)\right) \tag{4.9}$$

Where,

$U\left(c(a_{j|i})\right)$ = the local contribution to the prior energy from pixel $c(a_{j|i})$
$N(a_{j|i})$ = the neighbourhood system of pixels $a_{j|i}$
$w(a_l)$ = the weight of the contribution from neighbour pixel $a_l \in N(a_{j|i})$
$I(\alpha, \beta)$ takes the value 0 if $\alpha = \beta$ and 1 otherwise

The weight of the contribution can be modelled as:

$$w(a_l) = q \cdot \varphi(a_l) \tag{4.10}$$

The value of parameter q ranges from 0 to ∞ with higher values leading to a smoother solution. It controls the magnitude of the weights. As this model gives a smoot result, it reduces different class label between neighbourhood systems. $\varphi(a_l)$ is employed as an isotropic expression that depends only on the distance $d(a_{j|i}, a_l)$.

4.3.2. Likelihood energy function

The likelihood model considers a normal distribution of a single pixel b_i with value $y(b_i)$ with mean vector μ_i and covariance matrix C_i. Both can be defined as

$$\mu_i = \sum_{\alpha=1}^{L} \theta_{\alpha i} \mu_\alpha \tag{4.11}$$

$$\mathrm{C}_i = \sum_{\alpha=1}^{L} \theta_{\alpha i} C_\alpha \tag{4.12}$$

Where,

$\theta_{\alpha i}$ = the proportion of the class α in the pixel b_i

A probability model is designed for the class conditional density function $P(y|c)$. Gaussian distribution is used because it assumes that the classes of pixels are normally distributed. This assumption also used to produce maximum likelihood classification. For a K dimensional space, the likelihood of $y(b_i)$, given $c(a_{j|i})$, is formed as

$$P(y|c) = \prod_{i,j} U\left(y(b_i)\middle|c(a_{i|j})\right)$$

$$= \prod_{i,j} (2\pi)^{-K/2} |C_i|^{-1/2} \, exp\left(-\frac{1}{2}(y(b_i) - \mu_i)' C_i^{-1} (y(b_i) - \mu_i)\right) \tag{4.13}$$

$$U(y|c) = \sum_{i,j} U\left(y(b_i)\middle|c(a_{i|j})\right)$$

$$= \sum_{i,j} \left[\frac{1}{2}(y(b_i) - \mu_i)' C_i^{-1} (y(b_i) - \mu_i) + \frac{1}{2}\ln|C_i|\right] \tag{4.14}$$

4.3.3. Posterior energy function

Posterior energy function is needed for pixel class labelling based on MRF. $P(c|y)$ is the posterior distribution of an MRF. This distribution is derived by combining the prior and likelihood models. So the posterior energy can be defined by:

$$U(c|y) = q \sum_{ij} \sum_{l \in N(a_{j|i})} \varphi(a_l) I\left(c(a_{j|i}), c(a_l)\right) + \; U(y|c) \tag{4.15}$$

The equation above is divided by $1 + q$ and define the smoother parameter λ:

$$\lambda = \frac{q}{1+q}, 0 \le \lambda \le 1 \tag{4.16}$$

This can be expressed by:

$$U(c|y) \propto \lambda \sum_{ij} \sum_{l \in N(a_{j|i})} \varphi(a_l) I\left(c(a_{j|i}), c(a_l)\right) + \; (1-\lambda)\, U(y|c) \tag{4.17}$$

The smoothness parameter λ controls the contributions from prior and likelihood models. The value ranges from 0 to 1. It will neglect the likelihood model and assign all pixels to the same value if the smoothness parameter λ is 1.

4.4. MRF at object level

The MRF at the pixel level will create groups of pixels in the same class. At the object level MRF, this groups of pixels are treated as segments $\{S_1, S_s, \dots S_n\}$. By incorporating the contribution of area and texture, each segment may change its size and shape. Some of them may get bigger and some get smaller. Some segments may also disappear.

4.4.1. Area contribution

The segments consist of pixels belong to class *trees* is the main interest to be classified in this research. The size of tree crown is needed as one of the prior information of object-based image analysis. The shape of tree crown is roughly circular, so it was calculated by:

$$A = \pi r^2 \tag{4.18}$$

The size of tree crown was calculated in ArcGIS 10.3.1 software by digitizing 600 tree crowns in a circular shape. The tree crown statistics were done in R software for the probability density functions. Figure 4.3 (a) show the example of tree crowns samples using the panchromatic image as the background while Figure 4.3 (b) show the logarithmic normal distribution of tree crown area in square meter unit. This area unit was converted into pixel unit because this unit is used in this image analysis.

The size of other classes also estimated for object level prior information. The class *shadow* was estimated with the same method used for class *trees*. The size of *shadow* sample was calculated by delineating *shadow* objects in the image. The number of sample for class *shadow* are much smaller than the sample for class *trees*, but enough for a proper estimation. Class *unpaved road*, *bare soil* and *grass* were estimated by manually estimate the density function of the area size of each class. The probability of segment with area A_k given the true class C can be modelled by following the logarithmic normal distribution:

$$P(A|C) = \frac{1}{A_k \sigma_c \sqrt{2\pi}} . e^{-\frac{(\log A_k - \mu_c)^2}{2\sigma_c^2}} \tag{4.19}$$

The energy function can be expressed by

$$U(A|C) = \log\left(A_k \sigma_c \sqrt{2\pi}\right) + \frac{(\log A_k - \mu_c)^2}{2\sigma_c^2} \tag{4.20}$$

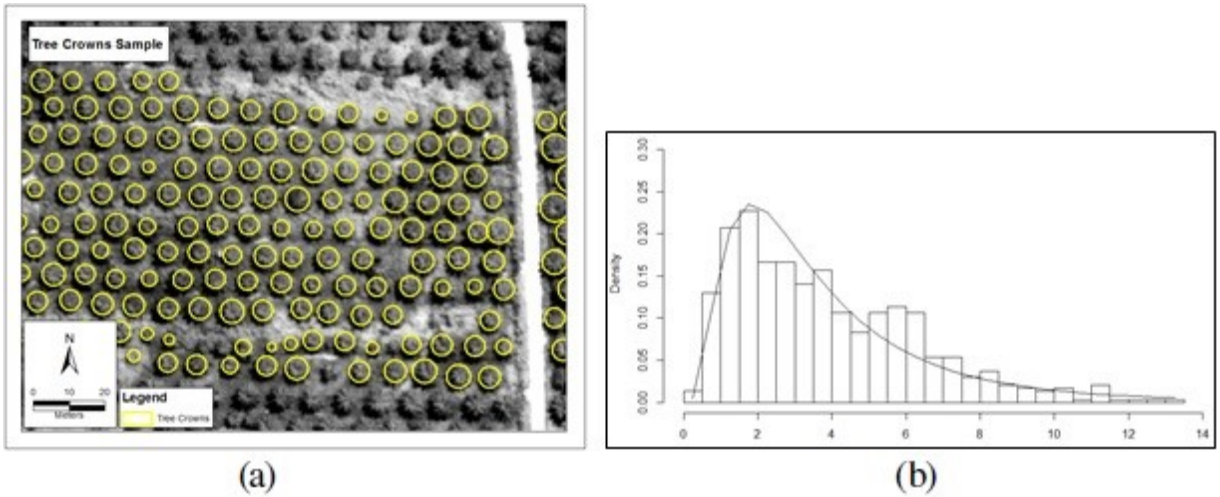

Figure 4.3: (a) Tree crowns digitization sample, (b) logarithmic normal distribution of tree crowns

4.4.2. Texture contribution

The texture contribution used as one of input for classification because texture becomes an advantageous information in object-based classification. It cannot be used properly for the multispectral image, due to the insufficient spatial resolution. In this research, GCLM is used to compute textural characteristic in the image. Lag=1 pixel is used to compute because the small size of the tree crown. A larger lag would not effective in observing the textural feature of the tree crown. In texture analysis, pixel i and its neighbouring pixel j are a neighbour in term of pixel intensities, not pixel coordinates. Several statistics can be derived from the GLCM that provides information of textural features in an image, which are:

- GLCM mean

$$\mu_i = \sum_{j=0}^{N-1} i(P_{i,j}), \qquad \mu_j = \sum_{i=0}^{N-1} j(P_{i,j}) \tag{4.21}$$

- GLCM variance

$$\delta_i^2 = \sum_{j=0}^{N-1} P_{i,j}(i-\mu_i)^2, \qquad \delta_j^2 = \sum_{j=0}^{N-1} P_{i,j}\left(j-\mu_j\right)^2 \tag{4.22}$$

- GLCM energy

$$\sum_{i,j=0}^{N-1} P_{ij}^2 \tag{4.23}$$

- GLCM entropy

$$\sum_{i,j=0}^{N-1} P_{i,j}(logP_{i,j}) \tag{4.24}$$

- GLCM contrast

$$\sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2 \tag{4.25}$$

- GLCM dissimilarity

$$\sum_{i,j=0}^{N-1} P_{i,j}|i-j|^2 \tag{4.26}$$

- GLCM correlation

$$\sum_{i,j=0}^{N-1} P_{i,j}\frac{\left(i-\mu_i\right)\left(j-\mu_j\right)}{\sigma_i\sigma_j} \tag{4.27}$$

- GLCM homogeneity

$$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \tag{4.28}$$

- GLCM maximum probability

$$\max\{P_{i,j}\} \tag{4.29}$$

- GLCM cluster shade

$$\sum_{i,j=0}^{N-1} P_{i,j}\left(i-\mu_i+j-\mu_j\right)^3 \tag{4.30}$$

- GLCM cluster prominence

$$\sum_{i,j=0}^{N-1} P_{i,j}\left(i-\mu_i+j-\mu_j\right)^4 \tag{4.31}$$

In this research, the texture contribution is modelled by using a normal distribution. The probability of a segment with a GLCM measure G_k given the true value of class C can be described by

$$P(G|C) = \frac{1}{\sigma_c\sqrt{2\pi}} . e^{-\frac{(G_k-\mu_c)^2}{2\sigma_c^2}} \tag{4.32}$$

The energy function can be expressed by

$$U(G|C) = \log\left(\sigma_c\sqrt{2\pi}\right) + \frac{(G_k-\mu_c)^2}{2\sigma_c^2}$$
$$= \frac{(\log G_k-\mu_c)^2}{2\sigma_c^2} \tag{4.33}$$

4.4.3. Combining contextual contributions into a posterior energy function

In order to accomplish the object-based image analysis, the contextual contributions for segment level MRF should be included into the posterior energy model $U(c|y)$. This will combine the prior and likelihood energy at energy level U_{pix} and the energy at segment analysis for area and shape contribution U_{seg}.

$$U_{pix} = \sum_{i=1}^{N}\left(U\left(y_i|c_i\right) + \sum_{j\in N_i} U\left(c_i|c_j\right)\right) \tag{4.34}$$

$$U_{seg} = \sum_{k=1}^{N_{seg}} U_k(A_k, c_k) \tag{4.35}$$

$$U_{total} = U_{pix} + U_{seg}$$
$$= \sum_{i=1}^{N}\left(\lambda\, U(y_i|c_i) + (1-\lambda)\sum_{j\in N_i} U\left(c_i|c_j\right)\right) + \lambda_{seg}\sum_{k=1}^{N_{seg}} U_k(A_k, c_k) \tag{4.36}$$

Simulated annealing with Metropolis-Hastings sampler is used to optimize the change from energy U_1 to energy U_2.

$$\Delta U = U_2 - U_1 \tag{4.37}$$

A system may share energy with its neighbours, so it is not energetically isolated. This share is defined by a temperature T. The Boltzmann energy factor, $e^{-\frac{E}{T}}$ is proportional to probability.

$$\rho = e^{-\frac{\Delta U}{T}} \tag{4.38}$$

4.5. Image segmentation

Image segmentation is a method in image analysis where the image partitioned into homogenous regions based on similarities in attribute values. Region growing is one of the approaches in image segmentation. This approach is initialized by merging regions from a seed pixel and growing iteratively to extend the region (Bins et al., 1996). In this research, the seed is chosen automatically because the position of the seed will not influence the segmentation result. Seeded region growing (SRG) is robust, swift and free of tuning parameters (Fan et al., 2005). The aim of SRG is to segment an image into regions with respect to a set of

n seeds. The initial seeds are replaced central pixels of the homogeneous region R. All the labelled pixels are called allocated pixels and the others are called the unallocated pixels. Let H be the set of all unallocated pixels which are adjacent to at least one of the labelled regions:

$$H = \{(a_{j|i}) \notin \cup_{i=1}^{n} R_i \,|N(a_{j|i}) \cap \cup_{i=1}^{n} R_i \neq 0\} \quad (4.39)$$

Where $N(a_{j|i})$ is the second-order neighbourhood system of the pixel $(a_{j|i})$ as shown in Figure 4.2 (b).

The neighbourhood system controls how many neighbours taken into the consideration. A higher order neighbourhood system require higher computational time. On the other hand, a lower order neighbourhood system has a limited area in describing the relationships between pixels. In this research, the second-order neighbourhood system is used considering the tree crown size in both 2 m and 0.5 m images. Neighbouring segments should have different homogeneous characteristics.

At object level, MRF works by forming a homogeneous segments. After the initial pixel level MRF, a number of segments or in the SRG called as regions. Each region consists of one or more pixel belong to the same class. A region shares a common boundary with its adjacent regions. Given the pixel $a_{j|i}$ is the boundary pixel of region R_i, all the pixels of its adjacent regions in the second-order neighbourhood system (pixels $a_{j\pm1|i\pm1}$) is explored. When the pixel at pixel location $a_{j\pm1|i\pm1}$ is changed to the class of region R_i in the iterative process of object level MRF, it will merge into the region R_i and become the boundary pixel of region R_i.

The application of image segmentation in a classified image would change the initial pixel labelling. Figure 4.4 shows the example of local changes possibilities in image segmentation due to the changes in pixels. Changes in one pixel may change the number of segments in an image. Local changes may also change the size of two or more segments by changes in the boundary pixel. Isolated pixels could be eliminated after image segmentation process. The local changes are calculated at segment level. These changes would make a difference in the energy before and after the image segmentation. This energy difference is optimized by using simulated annealing with Metropolis-Hastings sampler.

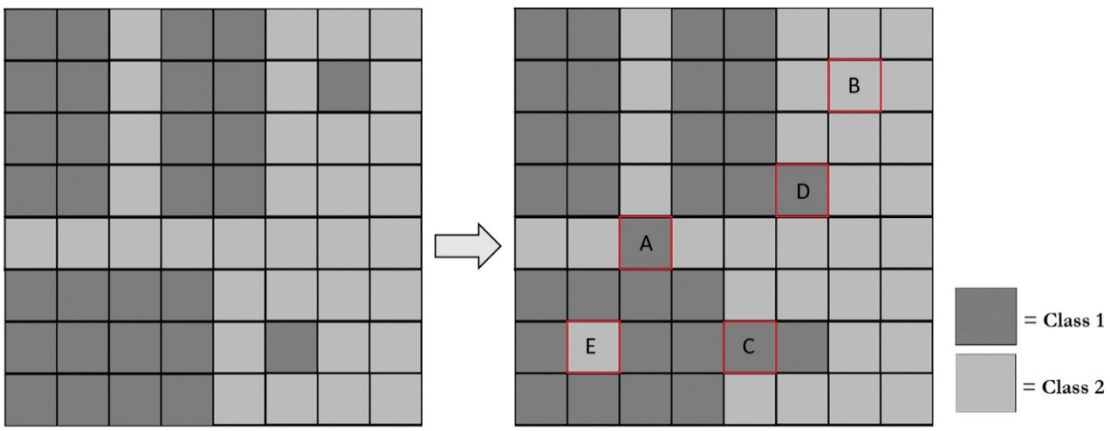

Figure 4.4: Examples of local changes possibilities due to image segmentation; A: Changes in one pixel that merged several segments of Class 1 into one segment and divided Class 2 into several segments; B: Elimination of isolated pixel; C: Changes in one pixel that merged several segments of Class 1; D: Changes in boundary pixel that change the size of two adjacent segments; E: Isolated pixel emerges

Figure 4.5 illustrates the changes of one pixel in four iteration process. The pixel that change in this scenario is pixel $a_{j|i}$ in the segment S_m that belongs to the class C_3. At the first iteration, this pixel changes into class C_1. Consequently, this pixel belongs to the same class with segment S_k that has pixel $a_{j+1|i-1}$ as its pixel boundary. Since it is located in the neighbourhood system $N(a_{j|i})$, the segment S_m merges into segment S_k. At the second iteration, the pixel $a_{j|i}$ change into class C_3. Therefore, it forms a new segment S_m. At the third iteration, the pixel $a_{j|i}$ change into class C_4. Hence, it belongs to the same class as segment S_l and segment S_p that located in pixel $a_{j|i-1}$ and pixel $a_{j|i+1}$. Because they belong to the same class and located in the same neighbourhood system, the merge into segment S_l since it has the largest area before the third iteration. At the fourth iteration, the pixel $a_{j|i-1}$ change into class C_5. Accordingly, it split the segment S_l into two segments, which are segment S_l and segment S_p. Because it has the same class as segment S_o and segment S_n that located in pixel $a_{j+1|i}$ and pixel $a_{j-1|i}$, they merge into segment S_o.

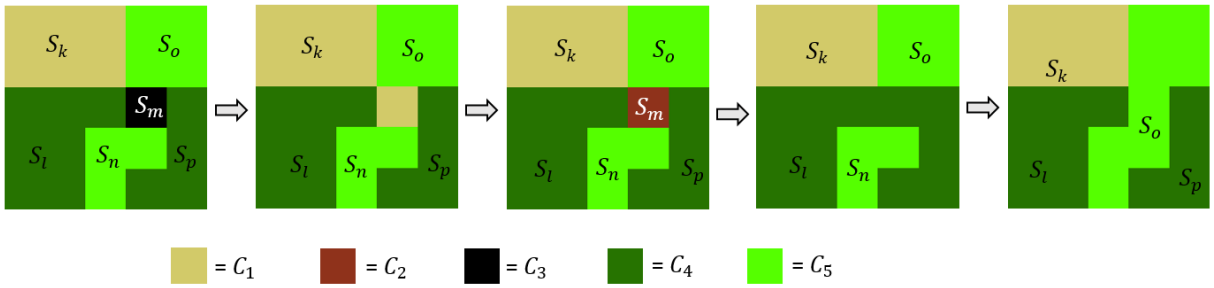

Figure 4.5: The segments changes because the changes of a pixel

The energy changes of segments involved in the local changes of pixel $a_{j|i}$ can be described by:

$$U_{seg_1} = U_k(S_k|\,C_1) + U_l(S_l|\,C_4) + U_m(S_m|\,C_3) + U_n(S_n|\,C_5) + U_o(S_o|\,C_5) + U_p(S_p|\,C_4) \tag{4.40}$$

$$U_{seg_2} = U_k(S_k + 1|\,C_1) + U_l(S_l|\,C_4) + 0 + U_n(S_n|\,C_5) + U_o(S_o|\,C_5) + U_p(S_p|\,C_4) \tag{4.41}$$

$$U_{seg_3} = U_k(S_k - 1|\,C_1) + U_l(S_l|\,C_4) + U_m(S_m|\,C_2) + U_n(S_n|\,C_5) + U_o(S_o|\,C_5) + U_p(S_p|\,C_4) \tag{4.42}$$

$$U_{seg_4} = U_k(S_k|\,C_1) + U_l(S_l + 5|\,C_4) + 0 + U_n(S_n|\,C_5) + U_o(S_o|\,C_5) + 0 \tag{4.43}$$

$$U_{seg_5} = U_k(S_k|\,C_1) + U_l(S_l - 5|\,C_4) + 0 + 0 + U_o(S_o + 4|\,C_5) + U_p(S_p|\,C_4) \tag{4.44}$$

The image in the above scenario consists of six segments before the iteration process starts or at the Seg_1. The table 4.1 presents the size changes of each segments at every iteration process. The segment S_k change at the second iteration when the segment S_m merges into their segment. They increase by one pixel at the Seg_2. The segment S_l change at Seg_4 with increase by five pixels after merged with the segment S_m and segment S_p. The segment S_m consists of one pixel at Seg_1 and Seg_3, when it belongs to class C_3 and class C_2. It disappear at Seg_2, Seg_4 and Seg_5 when it merged into another segments and forms a large segment. The segment S_n change at Seg_5 when it merged into segment S_o. The segment S_o change at Seg_5 when pixel $a_{j|i}$ and segment S_n merged into their segment. The segment S_p merged into segment S_l at the third iteration.

	Seg_1 (pixel)	Seg_2 (pixel)	Seg_3 (pixel)	Seg_4 (pixel)	Seg_5 (pixel)
S_k	6	7	6	6	6
S_l	7	7	7	12	7
S_m	1	0	1	0	0
S_n	3	3	3	3	0
S_o	4	4	4	4	8
S_p	4	4	4	0	4

Table 4.1: The size changes caused of each segments

4.6. Parameter estimation

4.6.1. Smoothness parameter at pixel level

The smoothing parameter at the pixel level (λ) is an essential parameter in MRF-based image analysis. The optimal value of λ should be defined beforehand. It balance the conditional energy and the prior energy in the optimization process and determine the smoothness of the classification result. The estimation of λ can be done by several experiments or by estimation of training data. Trial and error experiment may be time-consuming while estimation from training data is computationally expensive (Tolpekin & Stein, 2009). Trial and error experiments were used in this study to estimate λ. In this study, a subset image of 2 m multispectral GeoEye-1 satellite imagery was used to save the time. The optimal value of λ estimated by assessing the kappa κ value resulted from smoothness parameter λ values of 0.5, 0.7, 0.9 and 0.95 in this subset area at fixed other parameters (T_0=4, T_{upd} =0.95). The mean and standard deviation σ of kappa κ and overall accuracy value were calculated from 10 experiments of each λ value.

4.6.2. Simulated annealing parameters

Simulated annealing (SA) is widely used and works well as an optimization algorithm. It has been applied in wide range of areas since its introduction in 1983 (Aarts et al., 2005). The term of SA originally came from the way that metals and some liquids, cool and crystallize. Annealing is a process when the temperature cooled slowly and thermal equilibrium achieved at each stage of the gradual temperature lowering (Brooks & Morgan, 1995). Kirkpatrick (1984) studied the computational efficiency of simulated annealing when applied to graph partitioning and the traveling salesman problems. The results showed that simulated annealing is efficient for sufficiently large samples.

SA is able to avoid local optima and obtain the global optima. In this research, the energy optimization controlled by two main parameters, initial temperature (T_0) and updating temperature (T_{upd}). In this study, the cooling schedule is explained by $T = T_0 \times T_{upd}$. Initially, the SA process starts at high T_0 value and then cools down based on defined T_{upd}. Under SA, it is assumed that the T_0 will decrease iteratively until reaching the minimum energy solution. The T_{upd} should be defined carefully considering the duration of the computation time and the classification result. The T_{upd} controls the rate of temperature decrease. The process will stop if there is no more pixels are updated. The final result is expected to be achieved by applying the smoothness parameter λ. The smoothness parameter should be defined carefully since it is influential to classification result. The energy minimization is used for maximization of pixels class labelling.

The values of parameters for simulated annealing process are estimated in the multispectral image. Initial Temperature (T_0) and updating temperature (T_{upd}) are two parameters of simulated annealing. T is the temperature resulted from the iteration process. A lower T cause a tight coupling between neighbouring pixels and the image appear more regular while a higher T cause a loose coupling leading to chaotic

appearance (Geman & Geman, 1984). In this study, these parameters are estimated in the same subset image used in smoothness parameter estimation. 10 experiments of each value of T_0 and T_{upd} are used to obtain optimal value.

4.6.3. Smoothness parameter at object level

The optimum value of smoothing parameter at the object level (λ_{seg}) is estimated in a small subset. The subset area is smaller than the subset used for estimating smoothness parameter at pixel level λ. The result of each λ_{seg} value is measured by its kappa κ value and its quality in identifying the tree crown. The values measured in this parameter estimation are 0.1, 0.3, 0.5, 0.7, 1, 5 and 10. The value of λ_{seg} ranges from 0 to ∞. The λ_{seg}=10 is chosen as the maximum value because it has already give the example of the high λ_{seg} value result.

4.7. Accuracy assessment

Accuracy assessment will be calculated from the error matrix of the tree crown class and background class in the final classification result. An error matrix is a powerful way to illustrate the accuracy of classification result. The major diagonal of error matrix shows the proper classification result of each class and the rest of other values of the matrix represent the misclassified pixels.

User and producer accuracy and kappa κ value were calculated from the error matrix. User accuracy was corresponding to the error of commission (inclusion). It was calculated by dividing the number of correct pixels by the total number of pixels. Producer accuracy was corresponding to the error of omission (exclusion). It measures how well the reference pixels for each class have been classified. Kappa κ value is a discrete multivariate technique to measure the accuracy assessment of classification result (Congalton, 1991). It is necessary to also include the user and producer accuracy and not to only rely on kappa κ because kappa κ is a ratio, which problems may occur (Pontius & Millones, 2011).

The classification results of object-based MRF was also be assessed by compared with four classified images resulted from two other existing classification methods. First alternative method to be compared with the object-based MRF results is MLC. It is one of the most common methods for land cover classification. Second, the MRF-based classification that was implemented in both 2 m multispectral and 0.5 m pan-sharpened GeoEye-1 image.

The main concern in this research is an object level image analysis. To get a fair assessment for object-based super resolution map, an object level accuracy assessment should be performed. Ardila et al., (2011) developed a method to apply object level accuracy assessment. They defined two type of errors will be identified based on a reference map. The first type of error is the area inside tree crown that identified as a non-tree crown. The second type of error is the area outside tree crown that identified as a tree crown. Then, a geometric error (GE) were calculated based on this errors.

In this research, the object level accuracy assessment will be performed by identifying three status exclusively on detected and undetected tree crown class, which are:

- True positive: Tree crown correctly detected as tree crown
- False positive: Background class incorrectly detected as tree crown
- False negative: Tree crown incorrectly detected as background class

Figure 4.6 illustrates the implementation of the accuracy assessment. The accuracy assessment initialized by splitting the area inside and outside the tree crown area of the reference data. The segments inside the tree crown area are split into tree crown segment and non-tree crown segment. The segments outside the tree

crown also split with the same method. The tree crown segment inside the tree crown area are identified as the true positive. The non-tree segments inside the tree crown area and the tree segment outside the tree crown area is identified as the false negative and false positive respectively. The quality of the image classification performance was described by

$$Quality = \frac{True\ positive}{True\ positive + False\ positive + False\ negative}.100\% \quad (4.45)$$

The quality close to 100% represents a good quality result of the method in identifying tree crown.

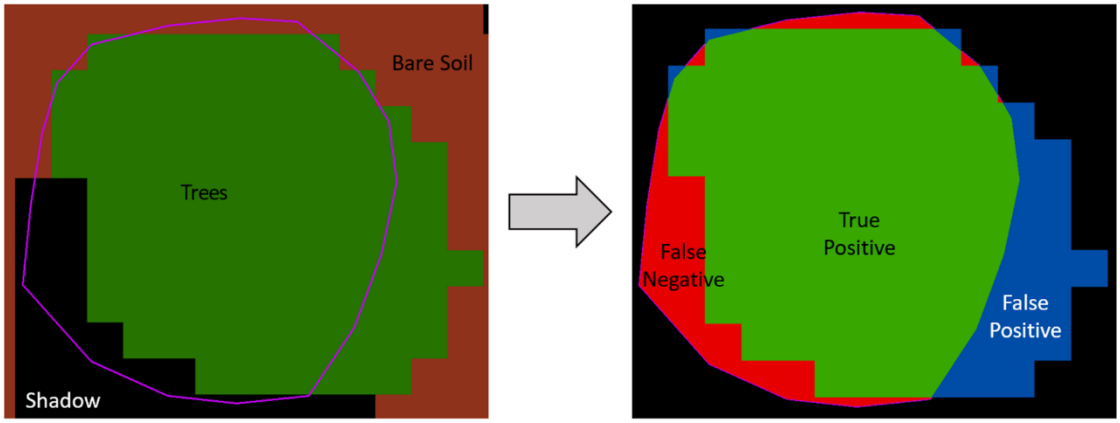

Figure 4.6: Accuracy assessment to measure the quality of the method in identifying tree crown

5. RESULTS

This chapter presents the results of all methods implemented in this research. It is divided into four main sections which cover the results of the spectral classes' definition based on the training set in the first section. In section 5.2, the results of all experimental results using the existing methods are given. The third section presents the result of the developed method. The last section covers the accuracy assessment performed for the developed method.

5.1. Spectral classes definition

In this image classification, five different colours represent five different classes, which are gold for the *unpaved road*, brown for *bare soil*, black for *shadow*, dark green for *trees* and light green for *grass*. The results presented in this section are highly related to the number of training samples used as the training area for a supervised classifier. Enough number of training samples is required in order to provide a proper training for the ML classifier used in this research.

The creation of the training samples in this research was not a simple effort to reach an enough number of training samples. The class *unpaved road* has a narrow shape. In the multispectral mode, the width of the road is only three pixels at its widest point. Class *shadow* and *trees* are very small. In the multispectral mode, their size is vary from 1-4 pixels. The use of higher spatial resolution as a training sample makes the training area might be contaminated by their neighbour class in the multispectral image. Therefore, the 0.5 m pan-sharpened image was used in identifying objects to help the creation of training samples. The number of training samples is not equal for all classes because some of the drawbacks mentioned above. The digitization of training dataset obtained 491 pixels in the multispectral image and 7,591 pixels in the pan-sharpened image.

5.1.1. Contingency Matrix

Contingency matrix contains the number of pixels which are properly classified. The contingency matrix of training sets in the multispectral image is presented in Table 4.1. It shows that most of the pixels are correctly classified. Every class has a different number of pixels, which depend on the number of pixels for each class in the image. The class *unpaved road* has the highest accuracy with 100% for both producer's and user's accuracy. The class *shadow* has the highest error of commission with 2.4%. The highest error of omission is class *trees* with 1.39% error. Contingency analysis for training sets significantly depends on a choice of classification method and number of training pixels. The number of training pixels for the multispectral image is not too large in respect to the number of bands. This may influence the high accuracy of the contingency analysis. The method used in calculating the contingency analysis also influencing the result. In this research, the contingency analysis is measured for maximum likelihood classifier.

	Unpaved Road	**Bare Soil**	**Shadow**	**Trees**	**Grass**
Unpaved Road	106	0	0	0	0
Bare Soil	0	152	0	0	0
Shadow	0	2	121	1	0
Trees	0	0	0	213	0
Grass	0	0	0	2	203

Table 5.1: Contingency matrix of training sets in multispectral image

The contingency matrix of training sets in the pan-sharpened image is also measured for maximum likelihood classifier. Based on Table 4.2, the overall accuracy shows a satisfying result. Similar to the result of contingency analysis in the multispectral image, class *shadow* and *trees* have the highest error of commission and error of omission respectively. The number of training pixels in the pan-sharpened image is large enough to give a more fair contingency result. The training area for both images is the same, but in the pan-sharpened image, the same training area covers roughly 16 times more pixels compare to the multispectral image.

	Unpaved Road	Bare Soil	Shadow	Trees	Grass
Unpaved Road	1581	1	2	8	0
Bare Soil	10	2353	48	0	0
Shadow	0	88	1818	181	9
Trees	0	3	50	3180	1
Grass	0	0	0	25	3240

Table 5.2: Contingency matrix of training sets in pan-sharpened image

5.1.2. Feature space

Figure 5.1 shows the feature spaces of the 2 m multispectral image at significance level 0.9. The feature space in Figure 5.1 (A) shows that there is some overlap between *trees* and *grass*. A little overlap also appears between the *unpaved road* and *bare soil* and between *bare soil* and *shadow*. The feature space in Figure 5.1 (B) shows that there is a large overlap between *trees* and *shadow*. The feature space in Figure 5.1 (C) shows that it has similarities with feature space of Band 4 and Band 3, but the overlap between *trees* and *grass* is smaller and the overlap between the *unpaved road* and *bare soil* is slightly larger. In 3D feature space shown in Figure 5.1 (D), it is shown that one class can be separated from other classes at significance level 0.9. Overall, in all feature spaces the class *unpaved road* has a significantly larger area than all other classes. It implies that the class *unpaved road* has the highest variance in this image. The overlap between classes may lead to a lower accuracy result of the image classification.

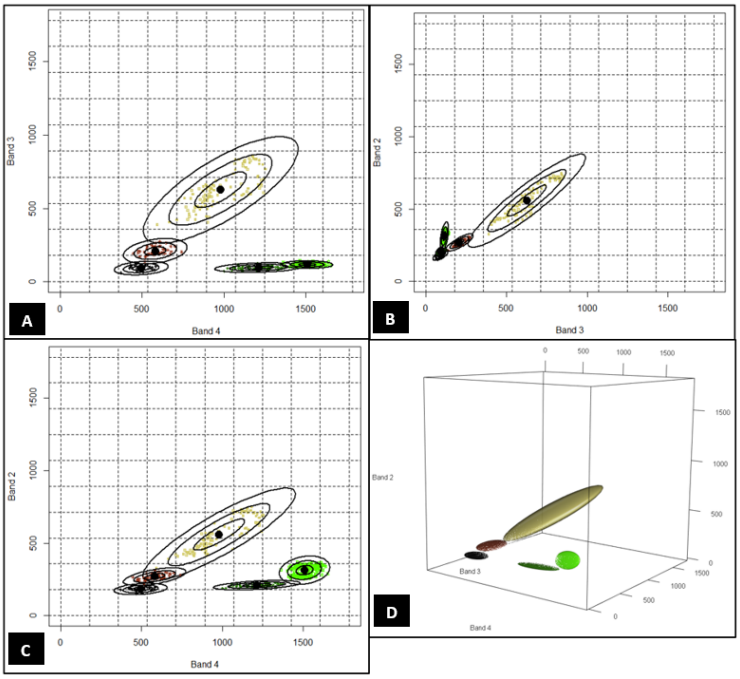

Figure 5.1: Feature spaces of multispectral image at significance level 0.9

Figure 5.2 shows the feature spaces of 0.5 m pan-sharpened image at significance level 0.9. In this feature spaces, the overall overlap between classes is larger than the feature space of multispectral image as expected. The dots are a lot denser than those in 2 m multispectral image. This is caused by the training sets in the panchromatic image has more pixel and higher variance. The feature space in Figure 5.2 (A) and in Figure 5.2 (B), some of the grass pixels is sparsely overlapped with *trees* and *shadow*. The feature space in Figure 5.2 (C) has a larger overlap between each class although it has similar shape compared to feature space in Figure 5.2 (A). The overlap is smaller in 3D feature space compare to all 2D feature spaces. The small overlap is between the *unpaved road* and *bare soil* and between *trees* and *grass*. The class *unpaved road* covers the largest area in all of the feature spaces.

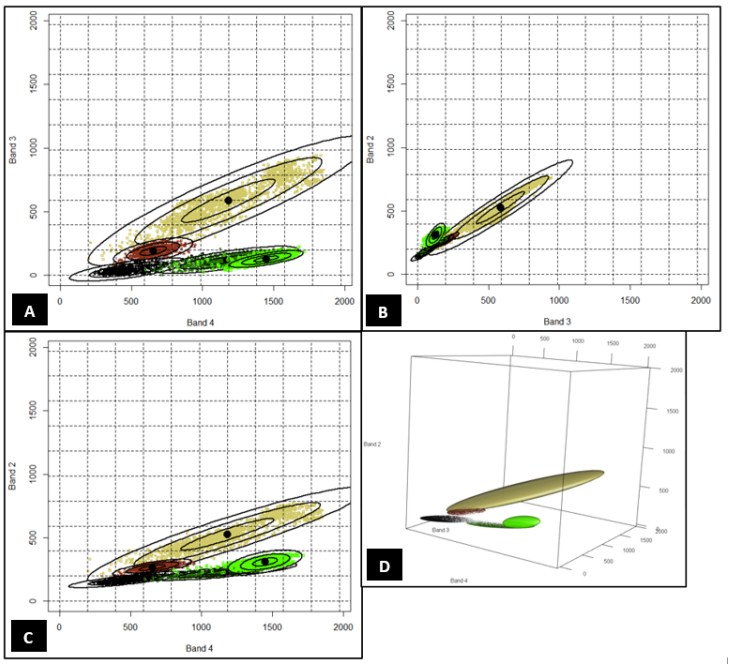

Figure 5.2: Feature spaces of pan-sharpened image at significance level 0.9

5.1.3. Class separability

The separability measured is an indirect estimation of correct classification. The class spectral separability values in multispectral mode show that every class has a high degree of separability from the other classes as presented in Table 5.3. Some of them have the highest possible values while the lowest value is 1.937. The lowest value of class spectral separability is between *grass* and *trees*. Both classes are vegetation, so they are close in term of spectral value compare to other classes. The class spectral separability of *bare soil* from *shadow* is 1.951. This might be caused by the *shadow* that occurs on the *bare soil* and by the reflectance value of *bare soil* that closest to shadow than the other classes.

	Unpaved Road	Bare Soil	Shadow	Trees	Grass
Unpaved Road	0	1.989	1.999	2.000	2.000
Bare Soil	1.989	0	1.951	2.000	2.000
Shadow	1.999	1.951	0	1.997	2.000
Trees	2.000	2.000	1.997	0	1.937
Grass	2.000	2.000	2.000	1.937	0

Table 5.3: Class separability of multispectral image

The class spectral separability values in pan-sharpened mode show that the spectral separability of *trees* from *shadow* is 1.777, which is the lowest value compare to other values. By using visual interpretation, in higher spatial resolution, there are exist some shadows within the tree crown area. The second lowest spectral separability is the spectral separability of *bare soil* from *shadow* with a value of 1.832. Overall, Table 5.4 presents a high degree of class spectral separability between classes and the spectral separability in pan-sharpened mode is slightly lower with the one in multispectral mode.

	Unpaved Road	Bare Soil	Shadow	Trees	Grass
Unpaved Road	0	1.965	1.988	1.999	2.000
Bare Soil	1.965	0	1.832	1.990	2.000
Shadow	1.988	1.832	0	1.777	2.000
Trees	1.999	1.990	1.777	0	1.938
Grass	2.000	2.000	2.000	1.938	0

Table 5.4: Class separability of pan-sharpened image

5.2. Parameters estimation

5.2.1. Smoothness parameter at pixel level

Table 5.5 illustrates the smoothness parameter λ estimation of 2 m multispectral image. The mean kappa κ value is decreasing following the increasing of λ. The λ value from 0.5 to 0.9 is not significantly different. While the λ value of 0.95 is leading to much lowest result accuracy amongst all the λ value. The standard deviation σ value of kappa κ and overall accuracy resulted from λ 0.95 is the highest. The computation time of the MRF classification using λ 0.95 is also the most expensive.

λ	Mean κ	σ κ	Mean Overall Accuracy	σ Overall Accuracy
0.5	0.959	< 0.001	96.777	< 0.001
0.7	0.958	0.001	96.732	0.068
0.9	0.929	0.002	94.418	0.196
0.95	0.853	0.005	88.597	0.408

Table 5.5: Smoothness parameter at pixel level estimation of 2 m multispectral image

Table 5.6 presents the smoothness parameter λ estimation of the 0.5 m multispectral image. The highest mean kappa κ is reached by employing λ value of 0.7 and 0.9. The λ value of 0.7 also leads to the highest overall accuracy mean score. Meanwhile, λ value of 0.9 is the second highest in overall accuracy mean at slightly lower score than λ value 0.7. The standard deviation σ of kappa κ and overall accuracy resulted from all parameter value are very small.

λ	Mean κ	σ κ	Mean Overall Accuracy	σ Overall Accuracy
0.5	0.938	< 0.001	95.149	0.022
0.7	0.942	0.001	95.443	0.049
0.9	0.942	0.001	95.438	0.065
0.95	0.938	< 0.001	95.139	0.02

Table 5.6: Smoothness parameter at pixel level estimation of the 0.5 pan-sharpened image

Both tables above indicate that a smoothness parameter estimations is not sufficient if it only refers to the kappa κ value and overall accuracy. The accuracy difference resulted from each of the smoothness parameter

λ is not significant. Furthermore, to make a better decision of choosing the best parameter, a visual interpretation of the parameter result is required. It gives a better understanding of how each of the smoothness parameter λ value affected the classification result. Based on this result, λ value 0.7 is chosen for both multispectral and pan-sharpened images.

5.2.2. Simulated annealing parameters

The simulated annealing parameters estimation result for the multispectral image shown in Table 5.7 for parameter T_0. The T_0 value 0 leads to the highest mean kappa κ value while T_0 value 5 leads to the lowest mean kappa κ value. Nevertheless, the difference is only as small as 0.005. Similar results also appear in overall accuracy mean, where T_0 value 0 and 5 leads to the highest and lowest score respectively. The difference between standard deviation σ of kappa κ resulted from each parameter T_0 are not significant. For standard deviation σ of overall accuracy, the T_0 value 10 leads to the highest value.

T_0	Mean κ	σ κ	Mean Overall Accuracy	σ Overall Accuracy
0	0.932	0.002	94.684	0.168
1	0.928	0.002	94.376	0.139
2	0.928	0.002	94.404	0.123
3	0.928	0.001	94.390	0.094
4	0.929	0.002	94.432	0.149
5	0.927	0.002	94.334	0.177
10	0.928	0.003	94.390	0.256

Table 5.7: Parameter T_0 *estimation of 2 m multispectral image*

Table 5.8 presents the results of parameter estimation T_{upd}. Each parameter resulting similar mean κ and standard deviation σ kappa κ. The difference of the overall accuracy result is also very small. The T_{upd} 0.99 leads to the highest overall accuracy score. On the other hand, T_{upd} value 0.95 leads to the lowest overall accuracy score. The T_{upd} value 0.8 has the highest standard deviation σ of overall accuracy.

T_{upd}	Mean κ	σ κ	Mean Overall Accuracy	σ Overall Accuracy
0.8	0.929	0.003	94.446	0.231
0.9	0.929	0.002	94.46	0.165
0.95	0.929	0.002	94.432	0.188
0.99	0.929	0.002	94.488	0.176

Table 5.8: Parameter T_{upd} *estimation of 2 m multispectral image*

Table 5.9 illustrates the result of experiments for parameter T_0 optimal value. The results show that applying different T_0 would not lead to a significant different on the result measured by kappa κ. All the T_0 values resulting very similar mean kappa κ and its standard deviation σ. The overall accuracy also implies the similar outcomes. The different between one overall accuracy mean to others is not significant.

T_0	Mean κ	σ κ	Mean Overall Accuracy	σ Overall Accuracy
0	0.940	< 0.001	95.323	0.021
1	0.941	0.001	95.406	0.053
2	0.941	0.001	95.426	0.074
3	0.941	0.001	95.405	0.066
4	0.941	0.001	95.424	0.052
5	0.941	0.001	95.423	0.087
10	0.941	0.001	95.421	0.052

Table 5.9: Parameter T_0 *estimation of 0.5 m pan-sharpened image*

The result of the experiments for parameter T_{upd} estimation presented in Table 5.10. Considering the results of each mean kappa κ and standard deviation σ of kappa κ, the T_{upd} value 0.95 has the highest mean kappa κ score and lowest standard deviation σ of kappa κ. Nonetheless, the difference is as low as 0,001. This T_{upd} value also the lowest in term of standard deviation σ of overall accuracy with 0.033. The highest mean overall accuracy was reached by applying T_{upd} value 0.9.

T_{upd}	Mean κ	σ κ	Mean Overall Accuracy	σ Overall Accuracy
0.8	0.941	0.001	95.368	0.097
0.9	0.941	0.001	95.425	0.092
0.95	0.941	< 0.001	95.418	0.033
0.99	0.940	0.001	95.347	0.088

Table 5.10: Parameter T_{upd} *estimation of 0.5 m pan-sharpened image*

The result of the simulated annealing parameters estimation shows that the difference between each parameter is not significant based on the mean of the kappa κ value. The standard deviation σ for all parameter combinations are very small. This estimation result makes all of the parameter combinations are proper to be used in the analysis. But, there are some other results to be considered beside kappa κ value, which are the visual result of the image analysis and the computation time. Parameter T_0 does not affect too much on the accuracy, but a higher value of T_0 leading to a longer computation time. Parameter T_{upd} influencing the computation time quite heavily. Meanwhile, a higher value of T_{upd} leads to slower computation time. Based on all of the considerations above, two different parameter combinations was chosen differently for the multispectral and pan-sharpened image. For the multispectral image, the T_0 value 4 and T_{upd} value 0.9 is chosen. T_0 value 4 and T_{upd} value 0.95 is chosen for the pan-sharpened image.

5.2.3. Smoothness parameter at object level

The smoothness parameter at object level for the multispectral image was estimated in a subset image with size 50 x 50 pixels or 10,000 m². Table 5.11 shows the result of smoothness parameter at object level estimation of the multispectral image. Although the difference is not very significant, the result indicates that the quality of the method in identifying tree crown and kappa value is decreased when the λ_{seg} is increased. The quality and kappa value has the same behaviour by the changes of λ_{seg}. With respect to this result, the λ_{seg} 0.1 was chosen to be applied in the implementation of hierarchical MRF. Figure 5.3 depicts the results of the λ_{seg} optimum value estimation. The λ_{seg}=10 leads to a very smooth result and consists of only three classes.

λ	λ_{seg}	Quality	Kappa κ
0.7	0.1	54.68%	0.750
	0.3	54.24%	0.308
	0.5	53.94%	0.143
	0.7	53.07%	0.143
	1.0	52.96%	0.143
	5.0	52.32%	0.143
	10.0	50.46%	0.100

Table 5.11: Smoothness parameter at object level estimation of 2 m multispectral image

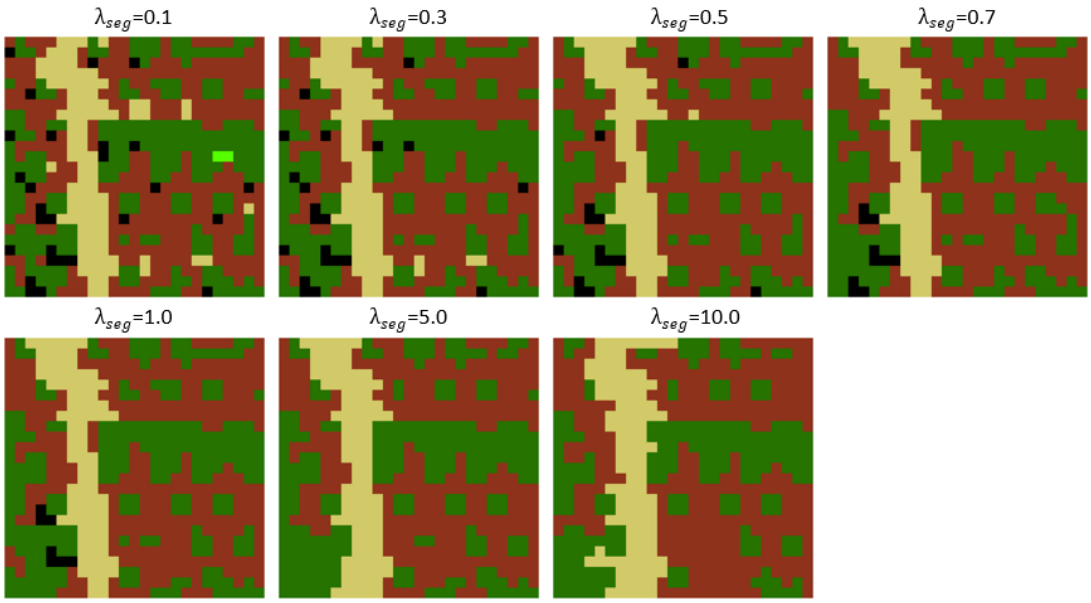

Figure 5.3: Results of the λ_{seg} optimum value estimation for 2 m multispectral image

The smoothness parameter at object level for the pan-sharpened image was estimated in a subset image with size 100 x 100 pixels or 250 m². Table 5.12 shows the result of smoothness parameter at object level estimation of the multispectral image. The result indicates that the quality of the method in identifying tree crown and kappa value is decreased when the λ_{seg} is increased. The quality and kappa value has the same behaviour by the changes of λ_{seg}. With respect to this result, the λ_{seg} 0.1 was chosen to be applied in the implementation of hierarchical MRF. Figure 5.4 showss the results of the λ_{seg} optimum value estimation. The result obtained by employing λ_{seg}=10 eliminated small segments. Consequently, it consists of four classes and the total area of tree crowns became larger.

λ	λ_{seg}	Quality	Kappa κ
0.7	0.1	60.74%	0.633
	0.3	60.39%	0.612
	0.5	60.35%	0.612
	0.7	60.35%	0.607
	1.0	60.26%	0.601
	5.0	57.42%	0.509
	10.0	55.95%	0.430

Table 5.12: Smoothness parameter at object level estimation of 0.5 m pan-sharpened image

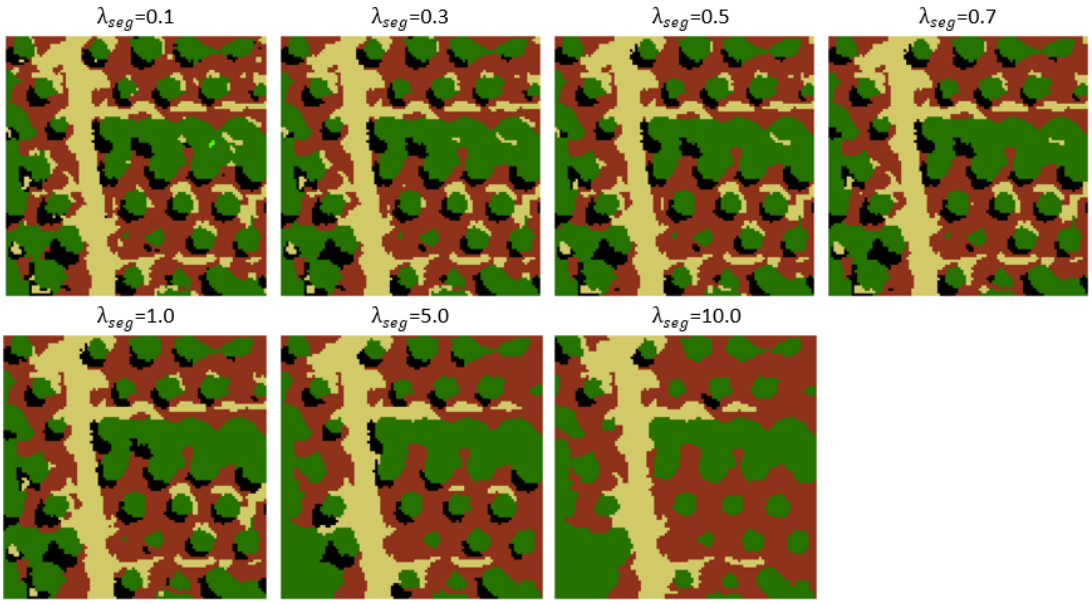

Figure 5.4: Results of the λ_{seg} optimum value estimation for 0.5 m pan-sharpened image

5.3. Maximum likelihood classification

Figure.5.5 presents the output of multispectral image classification using MLC with kappa κ value 0.959. As can be seen, the output has a salt and pepper effect occurs mainly in the *grass* and *bare soil* class. Some grass was misclassified as *trees* and vice versa. Some *bare soil* class was also misclassified as the *unpaved road*. The *unpaved road* class looks nicely classified without any significantly isolated pixels inside. Due to the insufficient of spatial resolution for automatically detecting oil palm *trees*, some small tree crowns was unable to be detected. However, some small tree crowns with size as small as one pixel were successfully identified. Although, bigger tree crowns with size at four pixels has a higher chance to be identified.

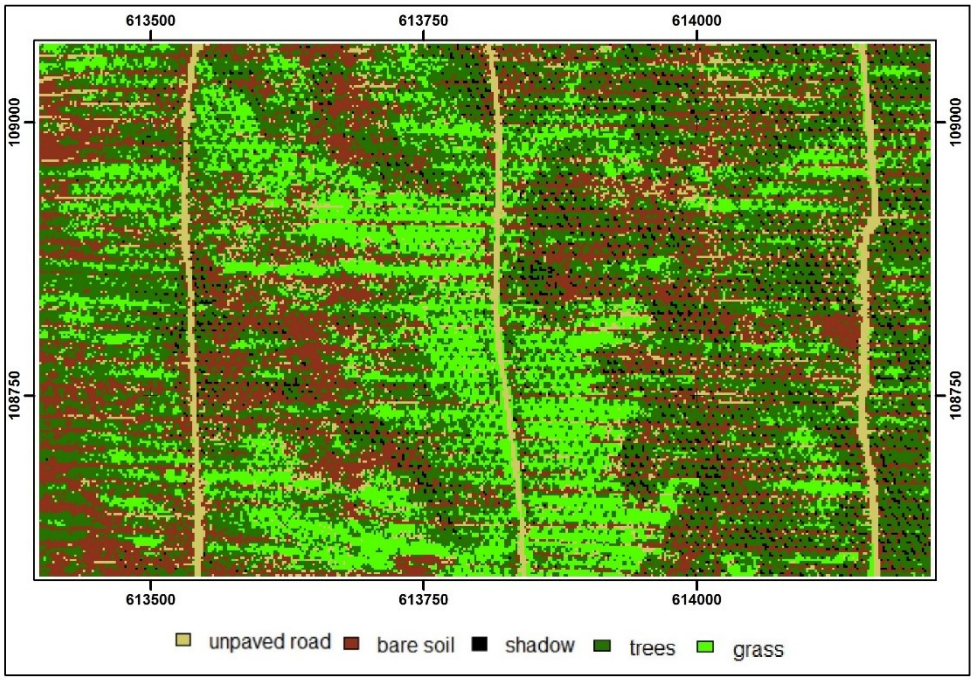

Figure 5.5: Classification result of multispectral image using maximum likelihood classifier

Table 5.13 shows the confusion matrix of the classified image. The lowest user's accuracy is the class tree at 89.36%. Moreover, the class *shadow* has the only 100% user's accuracy. The lowest producer's accuracy is

class *grass* at 91.08%. The highest class in producer's accuracy is the class *unpaved road* at 100% accuracy. The overall accuracy is very high at 96.77%. The highest misclassified occurred between two vegetation classes, which are *trees* and *grass*.

	Unpaved Road	Bare Soil	Shadow	Trees	Grass	User's Accuracy (%)
Unpaved Road	132	1	0	0	0	**99.25**
Bare Soil	0	111	1	0	0	**99.11**
Shadow	0	0	86	0	0	**100**
Trees	0	0	1	168	19	**89.36**
Grass	0	0	0	1	193	**99.48**
Producer's Accuracy (%)	**100**	**99.11**	**97.73**	**99.41**	**91.08**	**96.77**

Table 5.13: Confusion matrix of MLC result from multispectral image

Figure 5.6 presents the object-based accuracy assessment of MLC result from the multispectral image. The accuracy assessment shows that almost half of the area is identified as tree crown. There are 1,226 m² (43.53% of total area) identified as tree crown. The area of true positive is 1,148 m², which covers 81.72% area of total tree crown from the reference data. The overall quality accuracy obtained from this method is 48.37%. The quality is considerably low in term of delineating oil palm tree crown.

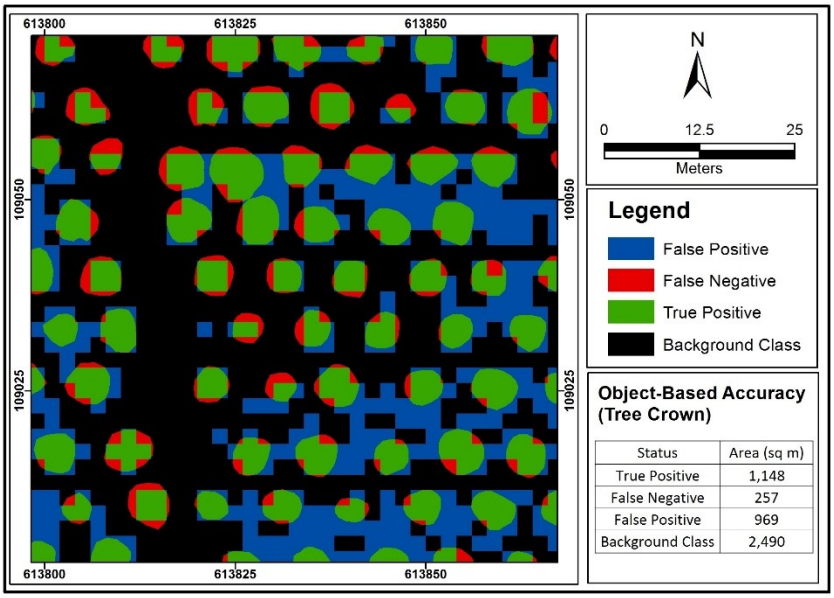

Figure 5.6: Object-based accuracy assessment of MLC result from multispectral image

Figure.5.7 presents the output of pan-sharpened image classification using maximum likelihood classifier with kappa κ value 0.935. The image shows a better result in distinguishing *bare soil* and *unpaved road*. Some grass still classified as tree crown. Some bigger tree crowns merging with their neighbour *trees*. Smaller tree crowns can be detected in this image. A salt and pepper effect occurs in the image. Based on the kappa κ value, the classification of the pan-sharpened image has a lower accuracy than the classification of the multispectral image. But, based on visual interpretation, the result show a better result in detecting oil palm trees.

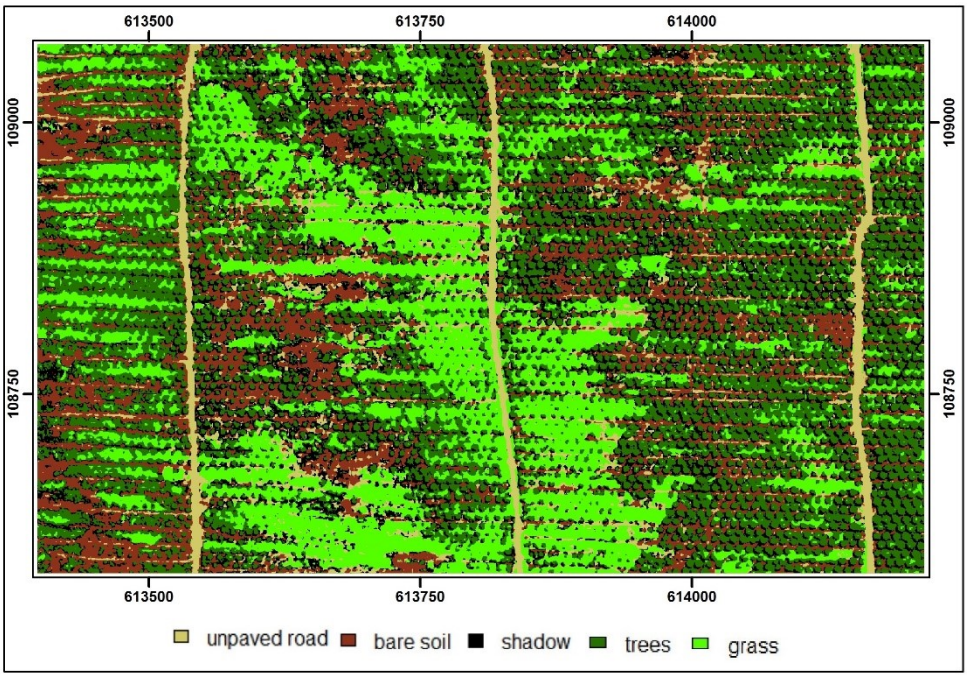

Figure 5.7: Classification result of pan-sharpened image using maximum likelihood classifier

Table 5.14 shows the confusion matrix of the classified image. Similar to multispectral image classification result, the *unpaved road* has 100% producer's accuracy. The class *bare soil* earned the top user's accuracy at 98.42%. The class *trees* is the lowest in user's accuracy with just slightly above 90%. Most of them misclassified to class *shadow* and *grass*. The *shadow* has 90.54% producer's accuracy, which is the lowest among the others. The overall accuracy is 94.96%.

	Unpaved Road	Bare Soil	Shadow	Trees	Grass	User's Accuracy (%)
Unpaved Road	2057	99	6	0	3	**95.01**
Bare Soil	0	1681	27	0	0	**98.42**
Shadow	0	8	1273	70	1	**94.16**
Trees	0	1	100	2422	154	**90.47**
Grass	0	0	0	93	3156	**97.14**
Producer's Accuracy (%)	**100.00**	**93.96**	**90.54**	**93.69**	**95.23**	**94.96**

Table 5.14: Confusion matrix of MLC result from pan-sharpened image

Figure 5.8 presents the object-based accuracy assessment of MLC result from the pan-sharpened image. The 41.9% of the area is classified as tree crown, this percentage is less than the result of the multispectral image. The percentage of true positive is 86.52% of the total area of tree crown from the reference data. The overall quality accuracy obtained from this method is 54.58%. Using a fine spatial resolution, smaller tree crown can be identified and the shape of each class is closer to the real shape and the quality is increased.

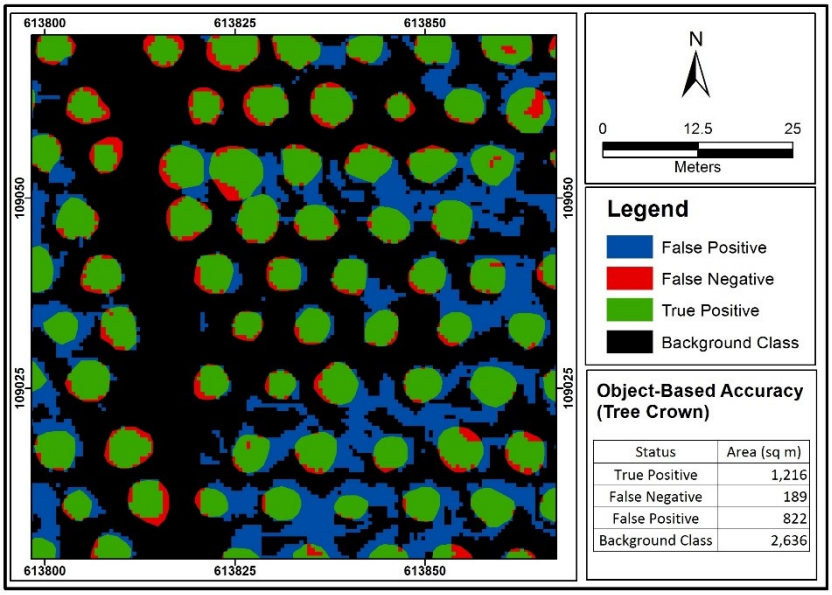

Figure 5.8: Object-based accuracy assessment of MLC result from pan-sharpened image

5.4. Markov random field

5.4.1. Pixel-based Markov random field

The multispectral image classification using MRF output is much smoother than the one using MLC as shown Figure 5.9. The smoothness parameter reduces the salt and pepper effect significantly. Some similarly misclassified case appears in the image that happens between *grass* and *trees* and between *bare soil* and *unpaved road*. Some *trees* are merged with their neighbour *trees*. The kappa κ value of the image is 0.957. There is a slight increase of kappa κ value after introducing MRF for classifying the image.

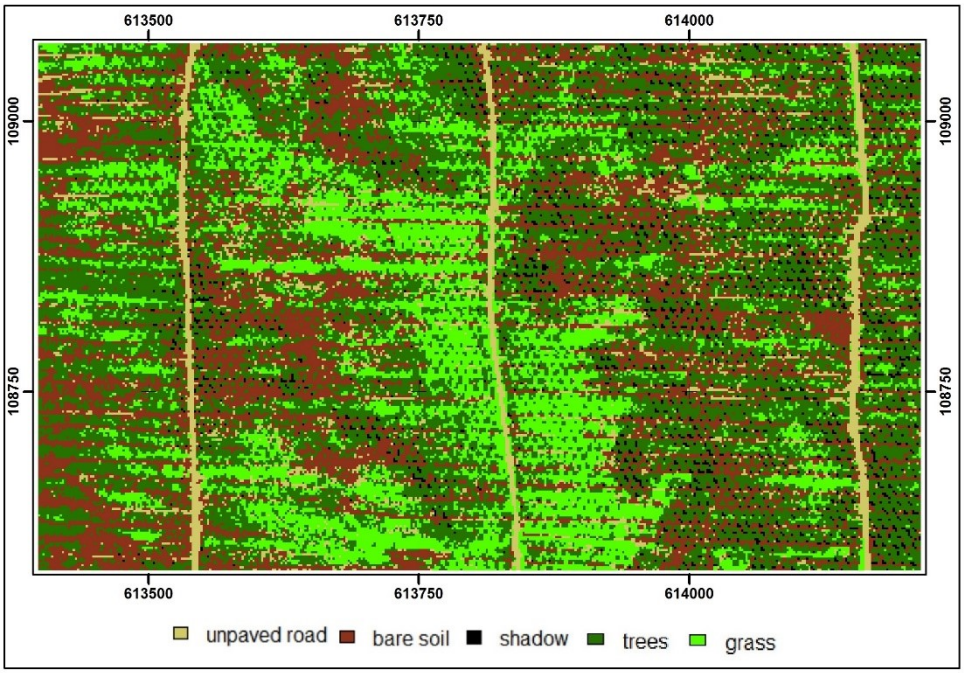

Figure 5.9: Classification result of multispectral image using Markov random field

Table 5.15 presents the confusion matrix of the classified image. Two classes have top accuracy for each producer's and user's accuracy. The class *unpaved road* and *bare soil* have 100% accuracy in producer's accuracy

while the *unpaved road* and *shadow* has 100% in user's accuracy. The class *trees* have the lowest user's accuracy at 88.42%. Most of the wrongly classified pixels are classified as the class *grass*. On the other hand, the class *grass* has the lowest producer's accuracy at 90.09%. The overall accuracy obtained from the applied method is 96.63%.

	Unpaved Road	Bare Soil	Shadow	Trees	Grass	User's Accuracy (%)
Unpaved Road	132	0	0	0	0	**100.00**
Bare Soil	0	112	1	0	0	**99.12**
Shadow	0	0	86	0	0	**100.00**
Trees	0	0	1	168	21	**88.42**
Grass	0	0	0	1	191	**99.48**
Producer's Accuracy (%)	**100.00**	**100.00**	**97.73**	**99.41**	**90.09**	**96.63**

Table 5.15: Confusion matrix of pixel-based MRF result from multispectral image

Figure 5.10 presents the object-based accuracy assessment of pixel-based MRF result from the multispectral image. Compare to the result of the MLC, the MRF approach resulted a slightly lower accuracy of tree crown delineation. The percentage of true positive is 82.14% of the total tree crown area from the reference data. This is a little better than the MLC result. The overall quality accuracy obtained from this method is 48.17%. In this method, some false negative pixels still occur in the middle of the tree crown although smoothness parameter has been applied.

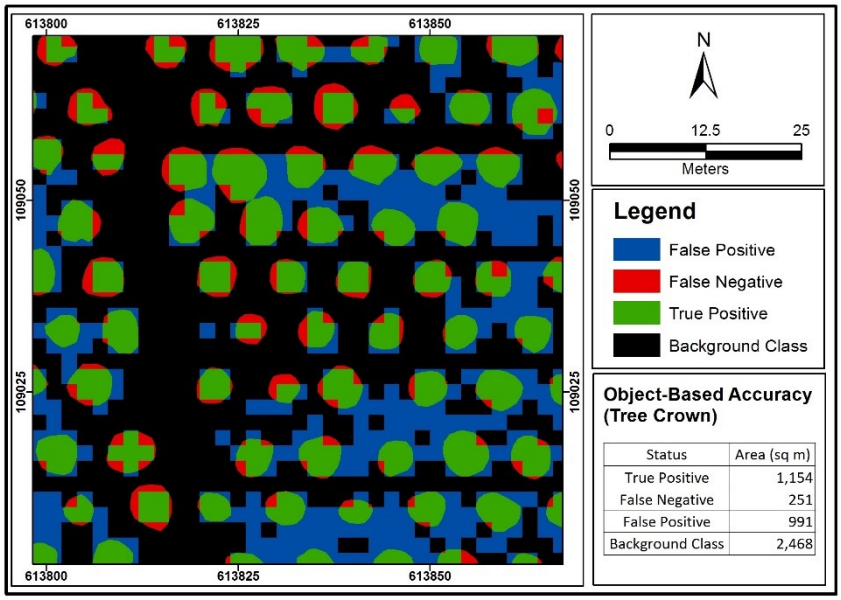

Figure 5.10: Object-based accuracy assessment of pixel-based MRF result from multispectral image

Figure 5.11 presents the output of pan-sharpened image classification using MRF with kappa κ value 0.941. Some bigger trees are merged and some grass classified as tree crowns, especially on the west where they spectrally similar to trees. The *unpaved road* on the road section lined quiet smoothly. After applying MRF, the kappa κ value was slightly increased. Although it was still difficult to distinguish vegetation classes. Due to the smoothness effect, some small tree crowns were disappearing. On the other hand, some bigger tree crowns are merged.

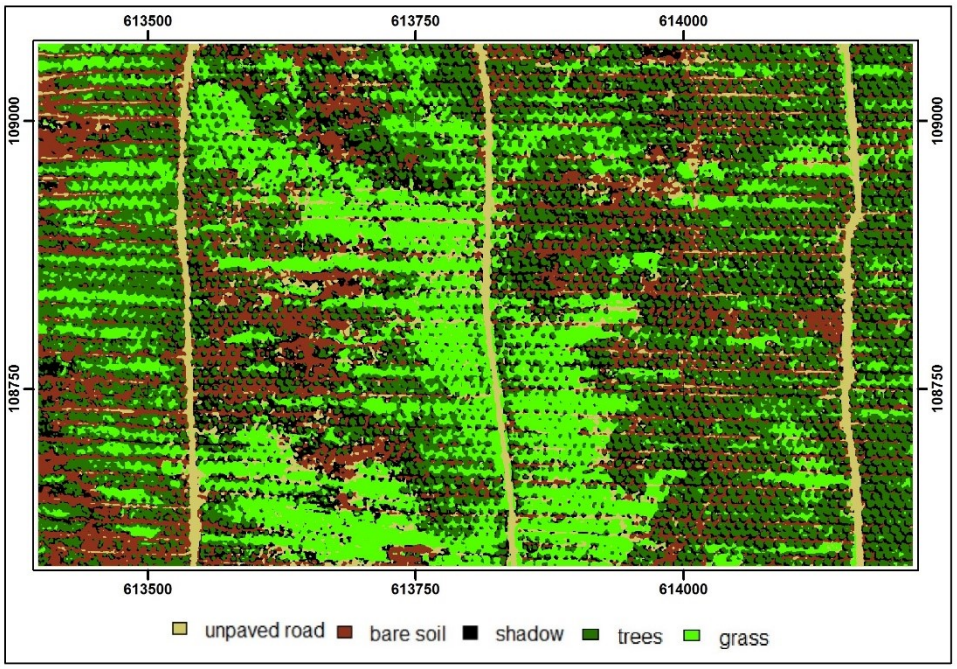

Figure 5.11: Classification result of pan-sharpened image using Markov random field

The error resulted from the pixel-based MRF result from the pan-sharpened image is shown in Table 5.16. The class *trees* has the lowest user's accuracy at 90.65%. Most of the misclassified pixels are classified to class *grass* and *shadow*. The class *shadow* has the lowest producer's accuracy at 90.83%. Most of the wrongly classified pixels are classified as *trees*. The *unpaved road* has the only class with an impeccable result in producer's accuracy. The class *bare soil* has the highest user's accuracy at 98.61%. The overall accuracy is 95.42%.

	Unpaved Road	**Bare Soil**	**Shadow**	**Trees**	**Grass**	**User's Accuracy (%)**
Unpaved Road	2057	74	5	0	2	**96.21**
Bare Soil	0	1704	24	0	0	**98.61**
Shadow	0	10	1277	64	2	**94.38**
Trees	0	1	100	2443	151	**90.65**
Grass	0	0	0	78	3159	**97.59**
Producer's Accuracy (%)	**100.00**	**95.25**	**90.83**	**94.51**	**95.32**	**95.42**

Table 5.16: Confusion matrix of pixel-based MRF result from pan-sharpened image

Figure 5.12 presents the object-based accuracy assessment of pixel-based MRF result from the pan-sharpened image. Compared to the result of MLC, the percentage of true positive is 87.57%, which is increased. The overall quality accuracy obtained from this method is 54.31%, which is quite similar compared to the MLC result. The smoothness parameter applied in this method can reduce some of the isolated pixels inside the tree crown. But, it also increased the number of false positive.

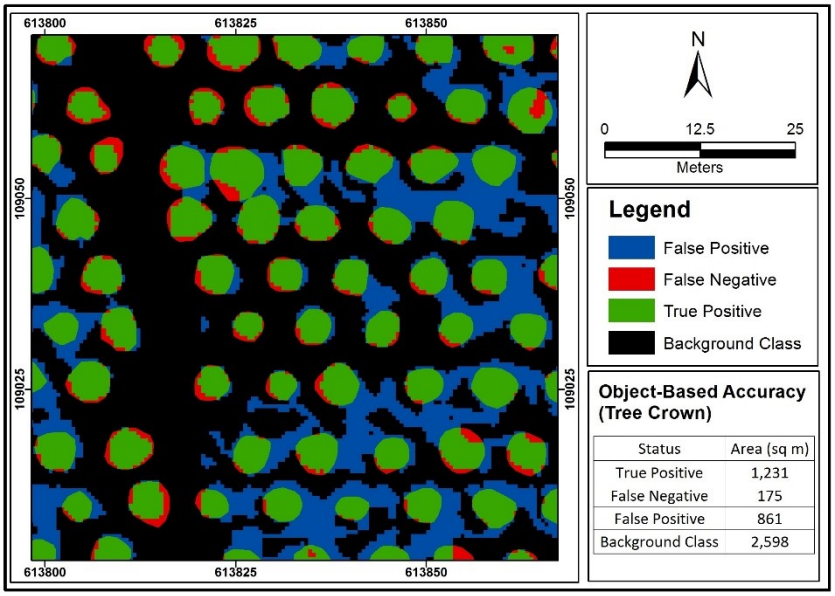

Figure 5.12: Object-based accuracy assessment of pixel-based MRF result from pan-sharpened image

5.4.2. Object-based Markov random field

Figure 5.13 shows the GLCM texture measurements in lag 1 for each training set reflecting each landcover class. This measurement is based on the same training set used in pixel-based image classification. The x-axis show the different angles and y-axis shows the GLCM mean. The different between classes is quite narrow and it is difficult to legitimately distinguish one class from other classes. The dependence of the GLCM measurements of each class is not really dependent on the angle. The class *unpaved road* might be the only class that has a dependency on the angle, despite it is not significant. In this research, the GLCM dissimilarity is used to measure the texture contribution.

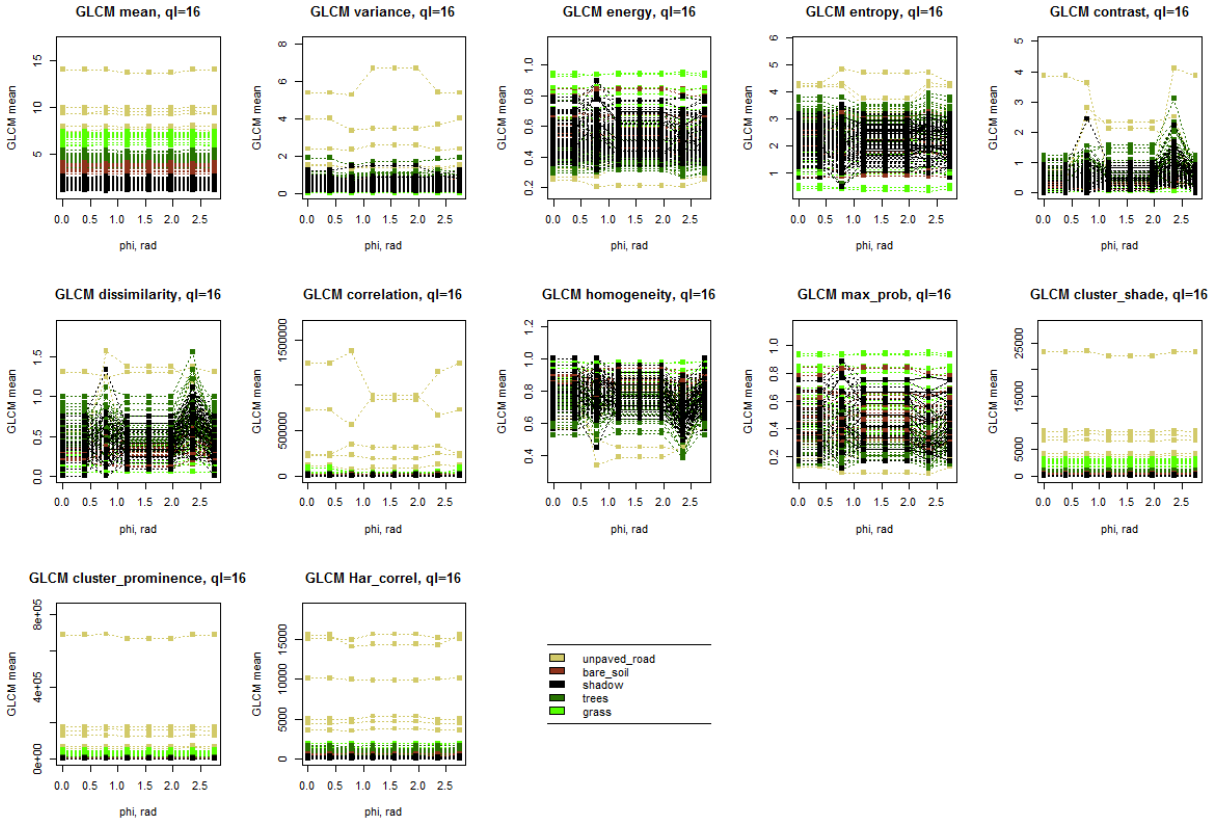

Figure 5.13: GLCM measurements in lag 1. The figure shows the dependence of the GLCM texture measurement on the angle.

Figure 5.14 show the histogram plot of the GLCM contrast frequency for each class. This histogram is modelled by a simple normal distribution to estimate the probability. The histogram of class *trees* shows the maximum frequency is in the middle of the histogram. The minimum frequencies are both in the left and right side.

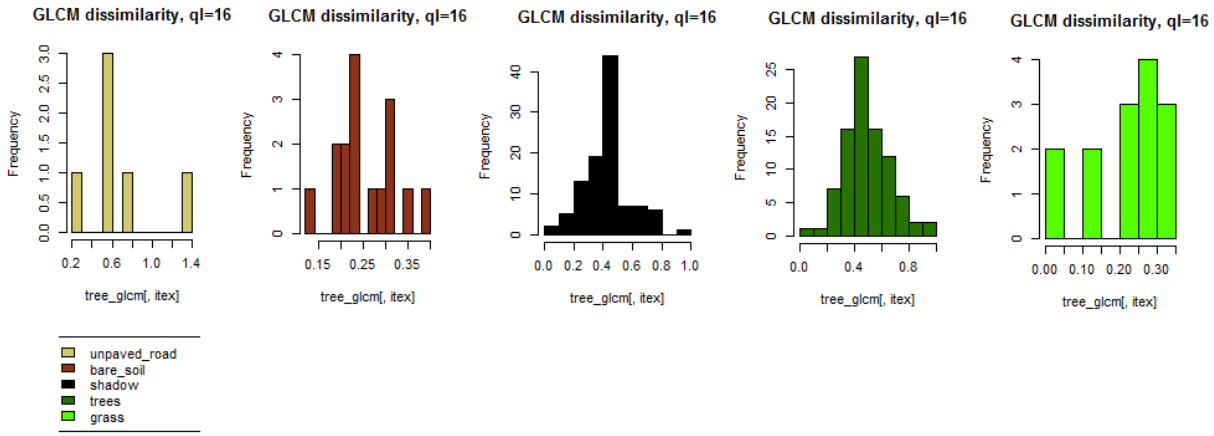

Figure 5.14: Histogram of GLCM contrast for each class in quantization level 8

Figure 5.15 presents the output of multispectral image classification using object-based hierarchical MRF. The result is very smooth because only some small segments occur in this result. The main issue is the area of tree crown is large. Most of the tree crown merged into a larger segment. They cover the part of grass segment. Some grass also identified as *bare soil* class. Some small segments of *unpaved road* class is appeared where it should not. The kappa κ value of this method is 0.75. This kappa value is not comparable to the kappa obtained from the pixel-based methods above because of the difference in the size of the image.

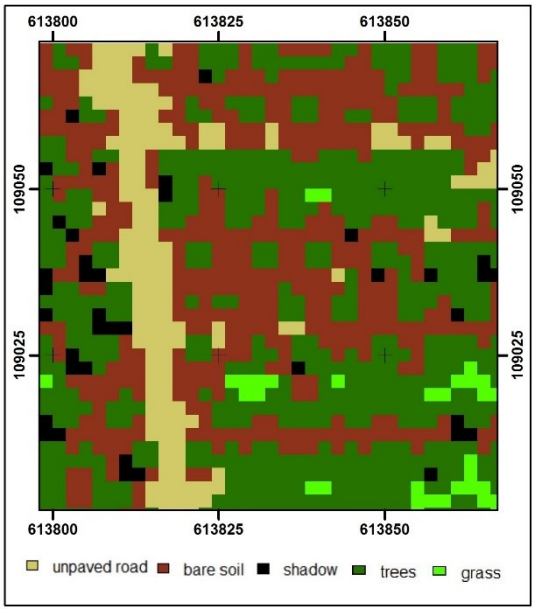

Figure 5.15: Classification result of multispectral image using object-based hierarchical Markov random field

Figure 5.16 presents the object-based accuracy assessment of object-based hierarchical MRF result from the multispectral image. The percentage of true positive is 87.02% of all tree crown in the reference dataset. The overall quality accuracy obtained from this method is 53.08%, which is higher than the result of both pixel-based and object-based MRF method. The number of false positive is increased after applying the object-based method, but the increasing of true positive and the decreasing of false negative leads to the highest quality of obtained from this image.

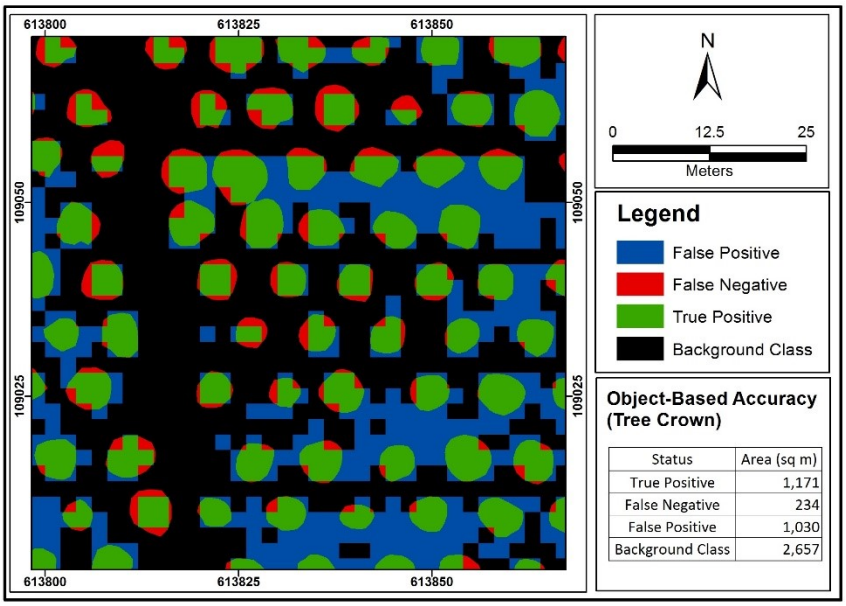

Figure 5.16: Object-based accuracy assessment of object-based hierarchical MRF result from multispectral image

Figure 5.17 presents the output of pan-sharpened image classification using object-based hierarchical MRF. The result of this method is not too smooth and still suffered the same problem as another class. Some small trees that identified in MRF method was disappeared because they merged into the segment in their neighbourhood system. The segments of *unpaved road* are identified, which is unnecessary. The *unpaved road* should be only occurs with a large segment. The kappa κ obtained from this method is 0.623. This result is lower than the result of the same method applied to the coarser spatial resolution.

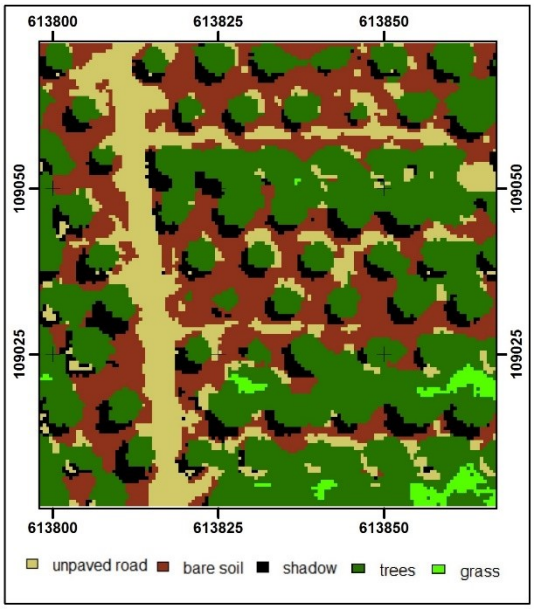

Figure 5.17 *Classification result of pan-sharpened image using object-based hierarchical Markov random field*

Figure 5.18 presents the object-based accuracy assessment of object-based hierarchical MRF result from the pan-sharpened image. The percentage of true positive is 90.15% of all tree crown in the reference dataset. The overall quality accuracy obtained from this method is 55.32%, which became the highest quality resulted from the 0.5 m pan-sharpened image. The application of object-based increased the true positive and decreased the false negative.

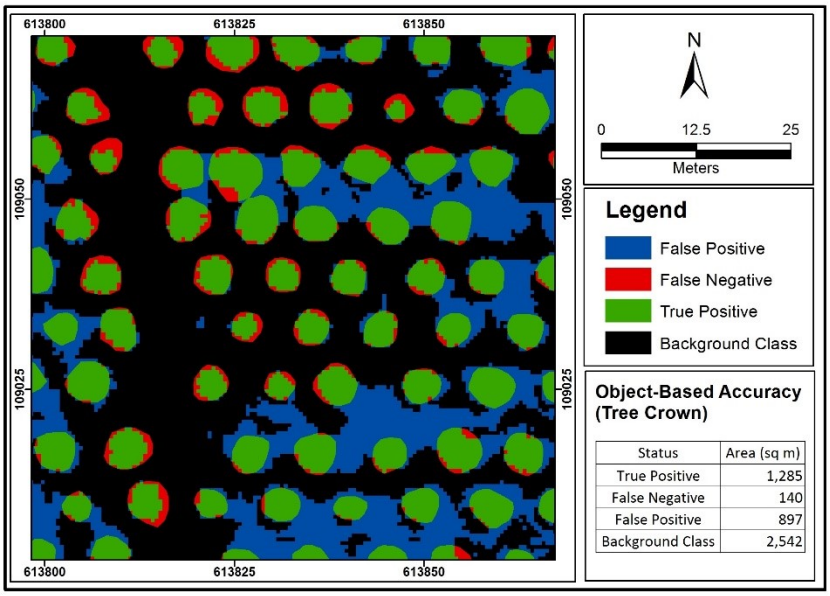

Figure 5.18: Object-based accuracy assessment of object-based hierarchical MRF result from multispectral image

5.5. Accuracy assessment

5.5.1. Pixel-based accuracy assessment

Table 5.17 presents the pixel-based accuracy assessments of all pixel-based method results applied in this research. The results is assessed with respect to their kappa κ value and overall accuracy. Overall, the result of MLC method on the multispectral image obtained the highest kappa κ and overall accuracy. The

multispectral image with regards to the applied methods obtained higher accuracy than the pan-sharpened image. The MRF method applied to multispectral image obtained slightly lower accuracy than MLC method. On the other hand, the MRF method applied to pan-sharpened image obtained higher accuracy than MLC method.

Results	κ	Overall Accuracy
MLC-Multispectral	0.959	96.77%
MLC-Pansharpened	0.935	94.96%
MRF-Multispectral	0.957	96.63%
MRF-Pansharpened	0.941	95.42%

Table 5.17: Pixel-based accuracy assessment of pixel-based method results

5.5.2. Object-based accuracy assessment

Table 5.18 illustrates the error detection at object level of all results. The three methods applied in this study lead to the same behaviour of errors in the result for both 2 m multispectral and 0.5 m pan-sharpened image. The false negative error will decrease when the result is smoother. On the other hand, the false positive will increase when the result is smoother. The result obtained from the pan-sharpened image has a less error.

Results	Error (m^2)	
	False Negative	**False Positive**
MLC-Multispectral	257	969
MLC-Pansharpened	189	822
MRF-Multispectral	251	991
MRF-Pansharpened	175	861
Obj-Multispectral	234	1,030
Obj-Pansharpened	140	897

Table 5.18: Error detection at object level of all results

Table 5.19 presents the quality obtained from all methods. The overall quality is decreasing when the pixel-based. After the introducing of the object-based MRF, the quality reached its highest value. The quality of the results obtained from the pan-sharpened image is better than the results from the multispectral image except the object-based MRF results. The results of pixel-based MRF have the lowest quality compared the other methods.

Results	Quality
MLC-Multispectral	48.37%
MLC-Pansharpened	54.58%
MRF-Multispectral	48.17%
MRF-Pansharpened	54.31%
Obj-Multispectral	53.08%
Obj-Pansharpened	55.32%

Table 5.19: Quality of all results

6. DISCUSSION

This chapter discusses the results from all the obtained results in chapter 5. The first section covers the discussion about the results obtained from each experiment to find the optimal parameters. The second section discusses the results of the classified images acquired from all methods.

6.1. Parameter estimation results

The objective of parameters estimation is to find the optimal value of parameters for each method applied in this research. Experiments were done for MRF-based image classification for both 2 m multispectral and 0.5 m pan-sharpened image separately. In general, each image shows different behaviour for each parameter combinations.

The smoothness parameter λ gave different effect for the results of each multispectral image and pan-sharpened image. In the multispectral image, the rise of λ should lead to a lower kappa κ value and overall accuracy. Higher λ should lead to a less stable result of the classification. The computational time also increases when λ is increasing. In the pan-sharpened image, the kappa κ and overall accuracy will increase when the λ increases until a peak point and decrease significantly. The kappa κ and overall accuracy result are highly dependent on the quality of the test data. In this research, the test data was derived from both pan-sharpened and multispectral image. This may have been influencing the statistical result of the image classification. Besides, the lack of spatial resolution of the multispectral image can also affect the classified image accuracy. In addition to the statistical analysis of the result, the effect of the smoothness parameter λ also explored visually. A less smooth classified image allows more isolated pixels exists in the classified image. Isolated pixels are not preferred to have existed in the classification result because it is unlike the fact on the ground. Visually, higher λ leads to a more preferred classification result.

The results show that the parameter T_0 and T_{upd} do not affect significantly the classification result in respect to their kappa κ value and overall accuracy. It was expected that higher T_{upd} would obtain higher accuracy results because of the slow cooling temperature. In the multispectral image, the lowest T_0 value leads to the highest kappa κ and overall accuracy. But, the increasing of T_0 value does not always leads to a lower accuracy of classification result. The highest T_0 leads to the most unstable result. The statistical analysis of the shows that the each of the parameter T_{upd} value leads to a similar classification result. The stability of the result also similar, but the lowest T_{upd} value leads to the most unstable result. In the pan-sharpened image, the lowest T_0 value leads to the lowest accuracy and the most stable result. Both parameters contribute to the computational time. Higher T_0 and T_{upd} should lead to a more expensive computational time.

In general, the results of λ_{seg} parameter estimation shows similar behaviour in both the multispectral and pan-sharpened image. In the multispectral image, the quality was gradually decreased when the λ_{seg} is increased. The difference between the result obtain from λ_{seg}=0.1 and λ_{seg}=10 is 4.22%. This difference is not too significant. The difference between the higher and the lowest quality percentage in the pan-sharpened image is also very narrow, which is 4.79%. Meanwhile, the results of λ_{seg} optimum value estimation shows a significant different visually. The result obtained by employing λ_{seg}=0.1 still allow some small segments to be appeared. On the other hand, λ_{seg}=10 does not allow small segments in the result. The result is very smooth. Consequently, some classes are eliminated because their segments are too small.

It also make the tree crown area became larger. Therefore, a higher λ_{seg} leads to a lower quality and κ value.

6.2. Pixel-based and object-based results comparison

In this research, three different approaches are applied to two images at two different spatial resolution, which is the 0.5 m and 2 m resolution. Although the focus is identifying oil palm tree crown, the overall image classification is also explored. Because of the difference in the spatial resolution, the results of 2 m multispectral and 0.5 m pan-sharpened image is not comparable. But, the application of object-based image classification for the 2 m multispectral image raises the potential of identifying tree crown.

Figure 6.1 presents the comparison of each method applied in this research. The result of image classification using MLC method for 2 m multispectral image is not good enough to identify tree crown. In some areas tree crowns can be identified clearly. The tree crowns that can be identified are surrounded by background classes that have contrast colour compare to tree crown colour, such as *bare soil* and bright grass. Moreover. The small and medium size of tree crown is easier to be identified. Bigger tree crowns tend to merge with their surrounding background classes and other bigger trees. The background class and size of tree crowns are affecting the performance of MLC in identifying tree crown from 2 m multispectral image. The MRF-based image classification for 2 m multispectral image shows similar behaviour to the MLC one. But, it reduces the salt and pepper effects on the result that can make the classified image smoother. This does not give an extremely positive impact on tree crown identification since there are only view isolated pixel inside tree crown area. The accuracy of the result by applying MRF is lower than MLC. The main issues of tree crown identification in this spatial resolution are difficulties in identifying very small tree crown and grass that identified as tree crown. These main issues are indicating the effect of the lack of spatial resolution and similarities in spectral value. This was expected beforehand as the spectral separability and feature space shown that tree crown and *grass* are close to each other. This was also proved by the confusion matrix in Table 5.11. Based on the object-based accuracy performance, the result of MLC and MRF cannot be separated. The overall quality performance is similar, although the MRF performed slightly better in delineating tree crown. The results show the difficulties in delineating tree crown located around class *grass*. Most of the cases, the grass also identified as class *trees*. Both methods experience the same problem.

The identification of tree crown using MLC for 0.5 m pan-sharpened image is not satisfactory. But, from this spatial resolution, very small tree crown can be identified. Especially, those that surrounded by *bare soil* and bright *grass*. In this image, a considerable amount of shadows appeared around trees compared to the result from the lower spatial resolution. This classified image also suffered a high confusion between *trees* and *grass*. Some pixels inside trees class has identified as *grass* and *shadow* since their spectral value are closer to grass class and there are some shadows inside tree crown. These problems were predicted earlier by the class spectral separability, which shows that the separability of *trees* from *shadow* and *grass* are lower than the others. These isolated pixels inside tree crown can be reduced by employing MRF. The result of applying MRF was considerably higher than MLC since its smoothness and its higher kappa κ value. Based on the object-based accuracy assessment, the MLC performed better with respect to the overall quality performance. But, the MRF is better in delineating tree crown. Both methods suffer the same difficulties in distinguishing tree crown and grass. But, using the finer spatial resolution, the number of false positive can be significantly reduced.

The application of λ_{seg} should make the hierarchical MRF result is smoother. But, in order to maintain the quality of the result, a low value of λ_{seg} was chosen in this research. Consequently, the object-based result is not extremely smooth. The results still show some small segments. But, these small segments helped the

result to obtain a higher quality. If the result is too smooth, then it will neglect the small segments, which might be the properly classified segments. The introduction of object-based approach applied for both images with optimum parameters leads to the highest quality. It increased the total area of the tree crown and decreased the false negative. The object-based MRF still difficult to penalize the wrong size value of each class. The tree class still merged into a large size and some small segments of *unpaved road* still appeared.

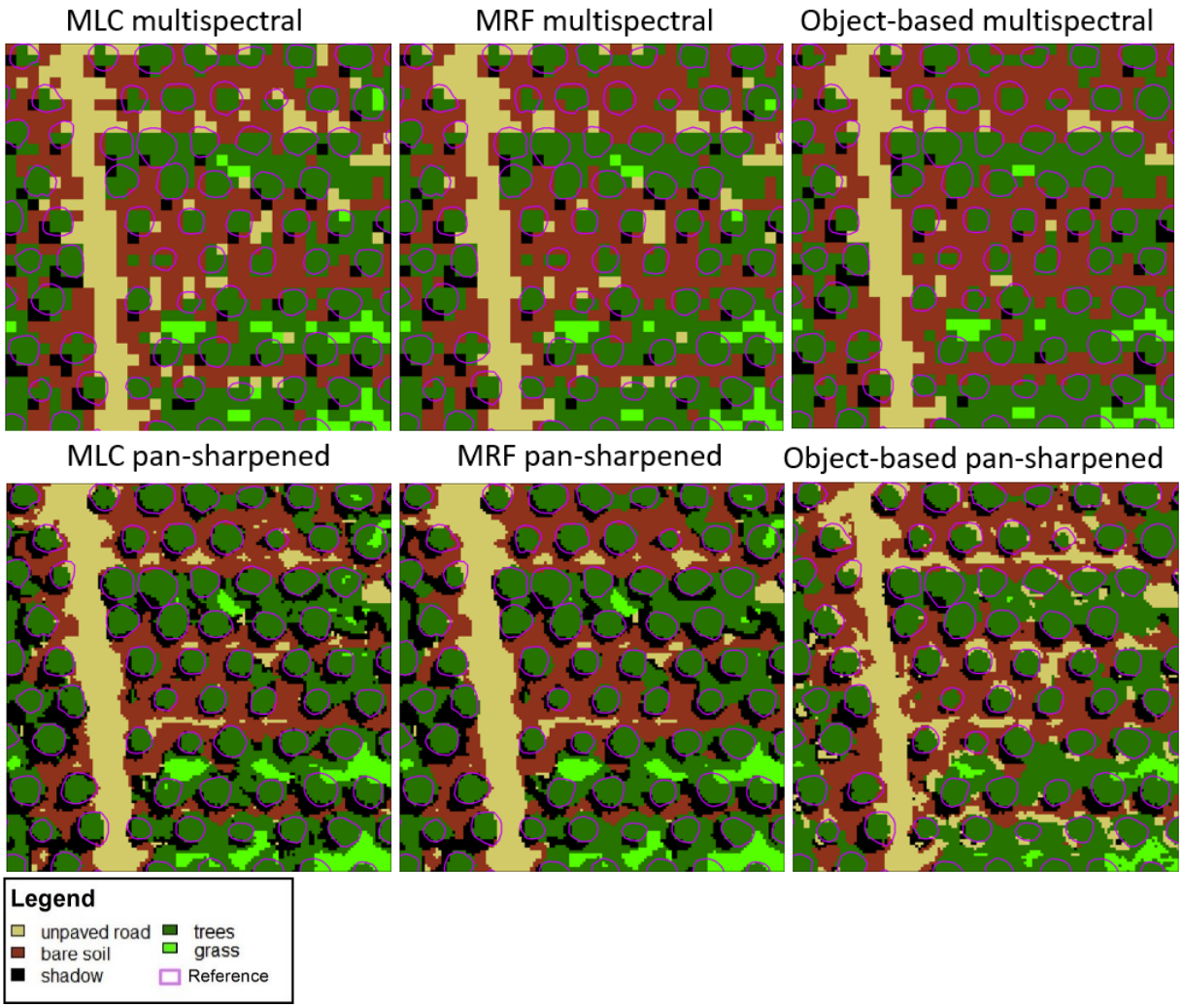

Figure 6.1: Comparison of classification results using: MLC, pixel-based MRF and object-based MRF of multispectral and pan-sharpened image.

7. CONCLUSIONS AND RECOMMENDATIONS

7.1. Conclusions

- How to design MRF probability models?
  The MRF probability models were designed by examining the distribution of each pixel level and object level characteristics.

- How pixel level and object level can be combined using MRF?
  The pixel level and object level were combined by applying hierarchical MRF at two different levels.

- How combinations of smoothness parameter for pixel and segment affect the accuracy?
  A higher value of smoothness parameter at pixel level tend to give a higher accuracy. On the other hand a higher value of smoothness parameter at object level leads to a lower accuracy.

- How the most appropriate parameters for MRF should be determined?
  The most appropriate parameters for MRF were determined by doing trial and error experiments.

- How the assessment of the results should be performed?
  The assessment was performed in both pixel and object level. The pixel level assessed by using error matrix to calculate the κ accuracy and overall accuracy. At object level, the results were assessed by identifying true positive status on tree crown.

- Which size of oil palm tree crowns can be successfully identified?
  The tree crown with size more than 4 m^2 is easier to be identified. The tree crown smaller than 4 m^2 might be identified if it is located around *bare soil*.

7.2. Recommendations

- Incorporate the satellite imagery data with visual interpretation of people who have experience and expertise in analysing oil palm plantation. The visual interpretation may improve the quality of training dataset. This may become more important since ground truth dataset is not available.

- This research does not introduces ground truth data, because the scope of work, limited time and the old acquired imagery. It is required to boost the quality in making the training area and assessing the quality of the result. It is recommended to integrate ground truth data in analysing the result of the applied method. This may increase the quality the research.

- The method developed in this research has become one of methods to handle high spectral variation image. It is recommended to apply this method to a more heterogeneous landcover. The urban area would be an appropriate area to explore further this method because it contains more landcover types and smaller objects.

- The computational time of the developed method is still too expensive for the application in oil palm plantation area. Oil palm plantations usually have a large size of area. Applying the object-

based image classification in thousands of square kilometres of VHR imagery is a heavy work. So, a more efficient method is necessary.

- This research shows that the spatial resolution of 2 m satellite imagery is not sufficient to detect small oil palm trees. Another alternatives are needed to overcome the limitation of image pan-sharpening method. Employing scale factor to obtain finer spatial resolution result in super-resolution mapping would be an interesting choice. It could be applied in hierarchical MRF developed in this research.

LIST OF REFERENCES

Aarts, E., Korst, J., & Michiels, W. (2005). Simulated Annealing. In *Search Methodologies* (Vol. 697, pp. 187–210).

Adams, R., & Bischof, L. (1994). Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *16*(6), 641–647.

Adler, D., & Murdoch, D. (2016). rgl: 3D Visualization Using OpenGL. Retrieved from https://cran.r-project.org/package=rgl

Ardila, J. P., Tolpekin, V. A., Bijker, W., & Stein, A. (2011). Markov-random-field-based super-resolution mapping for identification of urban trees in VHR images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(6), 762–775.

Atkinson, P. M. (2009). Issues of uncertainty in super-resolution mapping and their implications for the design of an inter-comparison study. *International Journal of Remote Sensing*, *30*(20), 5293–5308.

Bellon, O. R. P., & Silva, L. (2002). New improvements to range image segmentation by edge detection. *IEEE Signal Processing Letters*, *9*(2), 43–45.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen, M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *58*(3-4), 239–258.

Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistics Society*, *36*(2), 192–236. Retrieved from

Bins, L. S., Fonseca, L. M. G., Erthal, G. J., & Ii, F. M. (1996). Satellite Imagery Segmentation : a region growing approach. In *Anais VIII Simposia Brasilerio de Sensoriamenta Remoto, Salvador, Brasil, 14-19 abril 1996, INPE* (pp. 677–680).

Bivand, R., Keitt, T., & Rowlingson, B. (2015). rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 1.1-3. Retrieved from https://cran.r-project.org/package=rgdal

Blaschke, T. (2005). Towards a framework for change detection based on image objects. *Göttinger Geographische Abhandlungen*, *113*, 1–9.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16.

Brooks, S. P., & Morgan, B. J. T. (1995). Optimization using simulated annealing. *Journal of the Royal Statistics Society*, *44*, 241–257.

Chemura, A., van Duren, I., & van Leeuwen, L. M. (2015). Determination of the age of oil palm from crown projection area detected from WorldView-2 multispectral remote sensing data: The case of Ejisu-Juaben district, Ghana. *ISPRS Journal of Photogrammetry and Remote Sensing*, *100*, 118–127.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*(1), 35–46.

Cross, G. R., & Jain, A. K. (1983). Markov Random Field Texture Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-5*(1), 25–39.

Drăguţ, L., Csillik, O., Eisank, C., & Tiede, D. (2014). Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing : Official Publication of the International Society for Photogrammetry and Remote Sensing (ISPRS)*, *88*(100), 119–127.

Eddelbuettel, D., & Francois, R. (2011). Rcpp: Seamless R and C++ Integration. *Journal of Statistical Software*, *40(8)*, 1–18. Retrieved from http://www.jstatsoft.org/v40/i08/

Ediriwickrema, J., & Khorram, S. (1997). Hierarchical maximum-likelihood classification for improved

accuracies. *IEEE Transactions on Geoscience and Remote Sensing*, *35*(4), 810–816.

Fan, J., Zeng, G., Body, M., & Hacid, M.-S. (2005). Seeded region growing: an extensive and comparative study. *Pattern Recognition Letters*, *26*(8), 1139–1156.

Fisher, P. (1997). The pixel: A snare and a delusion. *International Journal of Remote Sensing*, *18*(3), 679–685.

Geman, S., & Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-6*(6), 721–741.

Graffigne, C., Heitz, F., Perez, P., Preteux, F. J., Sigelle, M., & Zerubia, J. B. (1995). Hierarchical Markov random field models applied to image analysis: A review. In E. R. Dougherty, F. J. Preteux, & S. S. Shen (Eds.), *SPIE's 1995 International Symposium on Optical Science, Engineering, and Instrumentation* (pp. 2–17). International Society for Optics and Photonics.

Gunal, S., & Edizkan, R. (2008). Subspace based feature selection for pattern recognition. *Information Sciences*, *178*(19), 3716–3726.

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *3*(6), 610–621.

Haralick, R. M., & Shapiro, L. G. (1985). Image segmentation techniques. *Computer Vision, Graphics, and Image Processing*, *29*(1), 100–132.

Hay, G. J., & Castilla, G. (2006). Object-Based Image Analysis : Strengths , Weaknesses , Opportunities and Threats (SWOT). *ISPRS Journal of Photogrammetry and Remote Sensing*.

Hay, G. J., & Castilla, G. (2008). Geographic object-based image analysis (GEOBIA): A new name for a new discipline. In *Lecture notes in geoinformation and cartography* (pp. 75–89). Berlin, Heidelberg: Springer.

Hu, R., & Fahmy, M. M. (1992). Texture segmentation based on a hierarchical Markov random field model. *Signal Processing*, *26*(3), 285–305.

Karatzoglou, A., Smola, A., Hornik, K., & Zeileis, A. (2004). kernlab - An S4 Package for Kernel Methods in R. *Journal of Statistical Software*, *11(9)*, 1–20. Retrieved from http://www.jstatsoft.org/v11/i09/

Kasetkasem, T., Arora, M., & Varshney, P. (2005). Super-resolution land cover mapping using a Markov random field based approach. *Remote Sensing of Environment*, *96*(3-4), 302–314.

Kato, Z., & Pong, T.-C. (2006). A Markov random field image segmentation model for color textured images. *Image and Vision Computing*, *24*(10), 1103–1114.

Kirkpatrick, S. (1984). Optimization by simulated annealing: Quantitative studies. *Journal of Statistical Physics*, *34*(5-6), 975–986.

Li, S. Z. (2009). Markov Random Field Modeling in Image Analysis.

Mallinis, G., Koutsias, N., Tsakiri-Strati, M., & Karteris, M. (2008). Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. *ISPRS Journal of Photogrammetry and Remote Sensing*, *63*(2), 237–250.

Maselli, F., Conese, C., & Petkov, L. (1994). Use of probability entropy for the estimation and graphical representation of the accuracy of maximum likelihood classifications. *ISPRS Journal of Photogrammetry and Remote Sensing*, *49*(2), 13–20.

Mertens, K. C., Verbeke, L. P. C., Ducheyne, E. I., & De Wulf, R. R. (2003). Using genetic algorithms in sub-pixel mapping. *International Journal of Remote Sensing*, *24*(21), 4241–4247.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2015). e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071). TU Wien. Retrieved from https://cran.r-project.org/package=e1071

Muad, A. M. (2011, December 15). Super-resolution mapping. Retrieved from http://eprints.nottingham.ac.uk/12309/1/Phd_Thesis_2011.pdf

Muad, A. M., & Foody, G. M. (2012). Impact of Land Cover Patch Size on the Accuracy of Patch Area Representation in HNN-Based Super Resolution Mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(5), 1418–1427.

Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, *115*(5), 1145–1161.

Otukei, J. R., & Blaschke, T. (2010). Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, *12*, S27–S31.

Peña-Barragán, J. M., Ngugi, M. K., Plant, R. E., & Six, J. (2011). Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sensing of Environment*, *115*(6), 1301–1316.

Pontius, R. G., & Millones, M. (2011). Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing*, *32*(15), 4407–4429.

Ramezani, M. (2015). *Urban Tree Species Classification Based on Spectral Profile of Crowns and Texture*. Retrieved from http://www.itc.nl/library/papers_2015/msc/gfm/ramezani.pdf

Richards, J. A. (2012). *Remote Sensing Digital Image Analysis: An Introduction*. Springer Berlin Heidelberg.

Ryherd, S., & Woodcock, C. (1996). Combining Spectral and Texture Data in the Segmentation of Remotely Sensed Images. *Photogrammetric Engineering Remote Sensing*, *62*(2), 181–194. Retrieved from

Salehi, B., Zhang, Y., & Zhong, M. (2011). Combination of object-based and pixel-based image analysis for classification of vhr imagery over urban areas. In *American Society for Photogrammetry and Remote Sensing Annual Conference 2011* (pp. 454–460).

Senthilkumaran, N., & Rajesh, R. (2009). Edge detection techniques for image segmentation–a survey of soft computing approaches. *International Journal of Recent Trends in Engineering and Technology*, *1*(2), 250–254.

Sha, Z., Bai, Y., Xie, Y., Yu, M., & Zhang, L. (2008). Using a hybrid fuzzy classifier (HFC) to map typical grassland vegetation in Xilin River Basin, Inner Mongolia, China. *International Journal of Remote Sensing*, *29*(8), 2317–2337.

Shackelford, A. K., & Davis, C. H. (2003). A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(10), 2354–2364.

Shih, F. Y., & Cheng, S. (2005). Automatic seeded region growing for color image segmentation. *Image and Vision Computing*, *23*(10), 877–886.

Srestasathiern, P., & Rakwatin, P. (2014). Oil Palm Tree Detection with High Resolution Multi-Spectral Satellite Imagery. *Remote Sensing*, *6*(10), 9749–9774.

Strahler, A. H. (1980). The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sensing of Environment*, *10*(2), 135–163.

Su, Y. F., Foody, G. M., Muad, A. M., & Cheng, K. S. (2012). Combining pixel swapping and contouring methods to enhance super-resolution mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*(5), 1428–1437.

Swain, P., & King, R. (1973). Two Effective Feature Selection Criteria for Multispectral Remote Sensing. *LARS Technical Reports*.

Tatem, A. J., Lewis, H. G., Atkinson, P. M., & Nixon, M. S. (2001). Super-resolution target identification from remotely sensed images using a Hopfield neural network. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(4), 781–796.

Tatem, A. J., Lewis, H. G., Atkinson, P. M., & Nixon, M. S. (2002). Super-resolution land cover pattern prediction using a Hopfield neural network. *Remote Sensing of Environment*, *79*(1), 1–14.

R Core Team. (2015). R: A language and environment for statistical computing (version 3.2.2) [Computer Software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Tolpekin, V. A., & Stein, A. (2009). Quantification of the Effects of Land-Cover-Class Spectral Separability on the Accuracy of Markov-Random-Field-Based Superresolution Mapping. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(9), 3283–3297.

Tso, B., & Mather, P. M. (2009). *Classification methods for remotely sensed data* (Second). Boca Raton: CRC Press.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S. Fourth Edition*. New York: Springer.

Villa, A., Chanussot, J., Benediktsson, J. A., & Jutten, C. (2011). Spectral Unmixing for the Classification of Hyperspectral Images at a Finer Spatial Resolution. *IEEE Journal of Selected Topics in Signal Processing*, *5*(3), 521–533.

Voisin, A., Krylov, V. A., Moser, G., Serpico, S. B., & Zerubia, J. (2013). Classification of Very High Resolution SAR Images of Urban Areas Using Copulas and Texture in a Hierarchical Markov Random Field Model. *IEEE Geoscience and Remote Sensing Letters*, *10*(1), 96–100.

Yamazaki, T., & Gingras, D. (1995). Image classification using spectral and spatial information based on MRF models. *IEEE Transactions on Image Processing : A Publication of the IEEE Signal Processing Society*, *4*(9), 1333–9.

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., & Schirokauer, D. (2006). Object-based Detailed Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery. *Photogrammetric Engineering & Remote Sensing*, *72*(7), 799–811.

Zhang, Y. (1999). Optimisation of building detection in satellite images by combining multispectral classification and texture filtering. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*(1), 50–60.