Comparison of Frequency-based Contextual and Maximum Likelihood Methods for Land Cover Classification in Arid Environment

M.R. Mustapha, H.S. Lim, M.Z. Mat Jafri and S. Syahreza
1School of Physics, Universiti Sains Malaysia, 11800 Penang, Malaysia
2Department of Physics, Faculty of Mathematics and Natural Sciences, Syiah Kuala University, Banda Aceh, 23111, Indonesia

Abstract: The classification accuracy obtained from the classification of satellite images using pixel-by-pixel conventional methods can be improved if the contextual information is considered during the classification process. This study presents a comparison of frequency-based contextual and maximum likelihood approaches to identify the land cover patterns in arid environment of multi-spectral images collected by SPOT-2 satellite. In image classification, in order to obtain a good result, not only the image resolution is considered but the selection of the classifier to be used during decision making process is important as well. In present study, two classifiers have been experimented in order to evaluate their performances which is Maximum Likelihood classifier representing as conventional method whereas contextual approach representing as advanced method. Conventional classification methods commonly cannot handle the complex landscape environment in the image. The result of each method has often “a salt and pepper appearances” which is a main characteristic of misclassification. It seems clear that information from neighbouring pixels should increase the discrimination capabilities of the pixel based measured and thus, improve the classification accuracy and the interpretation efficiency. This information is referred to as spatial contextual information. The experimental results indicated that frequency-based contextual algorithm with 83.7% overall accuracy and 0.693 Kappa coefficient is more reliable than the maximum likelihood algorithm with 72.1% and 0.527 overall accuracy and Kappa coefficient, respectively. The high value of the frequency-based contextual classification is due to the fact that this algorithm could overcome the mixed pixel problem and reduce the speckle error in the image significantly.

Key words: Frequency-based contextual, maximum likelihood, SPOT data, supervised, arid environment, remote sensing

INTRODUCTION

Image classification is an important part in many remote sensing applications especially for detecting the land use/cover types. Remote sensing is a valuable data source from which land cover information can be extracted efficiently (Chen et al., 2008). In the past decades, there has been a growing trend in the development of land cover map using remote sensing data. Land cover is a fundamental variable that impacts on and links with many parts of the human and physical environment (Foody, 2002). Effective classification of remote sensing image data depends upon separating land cover types of interest into sets of spectral classes that suited to the particular classifier algorithm used. In present study, land cover for the selected date was estimated using the supervised classification. To effectively derive reliable information from satellite data, appropriate classification techniques are essential. A number of classification approaches have been developed over the past decades. The classifiers can be categorized as either common or advanced. Some of the common classification algorithms include the K-Means, ISODATA, maximum likelihood classification and minimum distance to means (Mather, 2004; Lillesand and Kiefer, 1999). The advanced classification algorithms include the Artificial Neural Network (ANN), contextual, decision trees, support vector machines and object based image analysis (Lawrence et al., 2004; Keuchel et al., 2003; Verbeke et al., 2004; Lucieer, 2008; Hay et al., 2003; Xu et al., 2003; Mustapha et al., 2010a). But in present study, only contextual classification and maximum likelihood classifier were highlighted. Spectral imagery has been the primary tool for scene classification. In recent years, the progress of computer capabilities makes spatial feature processing techniques practical to implement in pursuit of improvement in classification accuracy (Olsen et al., 2002). Contextual information is one kind of such spatial

Corresponding Author: M.R. Mustapha, School of Physics, Universiti Sains Malaysia, 11800 Penang, Malaysia
relationship and has drawn our particular interest for remotely sensed imagery interpretation shown in present study. In the last years, efforts have been devoted to develop contextual classifier (Arain, 1993; Kontoes and Rokos, 1996). Classifiers that incorporate contextual information into classification have been reported in the literature as well (Chica-Olmo and Abarca-Hernandez, 2000; Atkinson and Lewis, 2000).

One of the advantages of classifying land cover types in arid environment is that it free from cloud. Cloud covers are generally thicker in the equatorial region (Sarkar and Kumar, 2002) and it causes the main problem to the researcher when using remote sensing technique in their research. Although, the cloud free condition in the arid area but there is still challenging to do the classification in this area due to the similarity of spectral characteristics among the land cover types. This situation would create the high possibility of mixed pixels to be occurred. Essentially, mixed pixels problem mostly appeared in urban class due to the large amount of spectral information are extracted in that particular region. Mixed pixels problem would lead the difficulty in decision making process. As stated by Lu and Weng (2006) urban land use/cover classification is still challenge with medium or coarse spatial resolution remotely sensed data due to the large number of mixed pixels and the spectral confusions among different land use/cover types. The mixed pixel problem would require additional physical information for a good classification. For that reason, the classification process needs the context information to be considered instead of depending on spectral information alone. A contextual classifier consistently produces higher classification accuracies than the per-pixel classification (Kontoes and Rokos, 1996; Stuckens et al., 2000). The purpose of present study is to highlight the potentiality of using different classifiers to generate important land cover information for the Mina city area using SPOT data.

MATERIALS AND METHODS

Description of study area: The study area is located in the Makkah province which is in the western of the Saudi Arabia (Fig. 1) and was conducted in the year of 2010. Mina city is located between Makkah and Arafah. Mina is the city known as Tent City where the pilgrims travel to nearby mount Arafah to spend the night camping in Mina valley during the Hajj season. Following a series of fatal disasters culminating in a large fire in 1997, the Saudi government decided to replace the existing cotton structure with fireproof Teflon coated glass tents (Grundig, 2002). This selected area was bounded by mountainous and desert terrain in every direction. The study area in the Arabian Peninsula is located on 21°24′N to 21°26′N latitude and 39°50′E to 39°54′E longitude.

Materials: The materials used for the classification were SPOT-2 satellite data of Mina City, Saudi Arabia and its surrounding areas which was acquired on January 2010 under clear weather conditions. It provides better spatial land-cover maps and land-use classification maps for monitoring regional environments. The SPOT-2 image has three reflective bands with 20 m spatial resolution. In this research, all reflective bands were used in the processing and image analysis. This band contains the visible and near infrared band [Band 1: 0.50 to 0.59 μm (green band); Band 2: 0.61 to 0.68 μm (red band) and Band 3: 0.78 to 0.89 μm (near infrared band)].

Pixel based versus contextual classification: Digital image processing is the numerical manipulation of digital image and includes pre-processing, enhancement and classification. Image classification refers to the extraction of differentiated classes or themes, usually land cover and land use categories, from raw remotely sensed digital satellite data. The information contained in a remotely sensed image and can be used to conduct image classification includes spectral pattern, spatial pattern and temporal pattern. Spectral pattern is the combination of Digital Numbers (DNs) for different feature types. Spatial pattern refers to the spatial relationship of the pixels, such as image texture, pixel proximity, features size and shape. Temporal pattern refers to temporal characteristics of the features (Qian et al., 2010).
A wide range of classification algorithms has been developed to derive land use and land cover information from remotely sensed images. Since remotely sensed images consist of rows and columns of pixels, pixel-based classification becoming the conventional method for land cover mapping. This classification method assigns a pixel to a class fundamentally according to the spectral similarities (Gong et al., 1992; Casals-Carrasco et al., 2000). The unsupervised classification approach provides an automated platform for image analysis, mainly based on surface reflectance and generally ignoring basic land cover characteristics (i.e., shape and size) of landforms (Chust et al., 2004). The supervised classification approach can preserve the basic land cover characteristics through statistical classification techniques using a number of well-distributed training pixels. The Maximum Likelihood (ML) classification method is well known for the analysis of satellite images. So far, satellite image interpretation using the ML approach was mostly applied for land cover classification (Huang et al., 2007) and monitoring of land use changes (Shalaby and Tateishi, 2007). Although, the techniques are well developed and many successful applications have been reported, it suffers from ignoring the spatial pattern in decision making process. The ML classifier quantitatively evaluates both the variance and covariance of the category spectral response patterns when classifying an unknown pixel without considering contextual information. ML is a statistical classifier so that it needs many training data for the classification process. Therefore, with very large of training sets, it can generate the statistical distribution and used this data for classifying the images (Mustapha et al., 2010b). ML is a parametric classifier that assumes normal spectral distribution for each feature of interest and an equal prior probability among the classes is also assumed. This classifier is based on the probability that a pixel belongs to a particular class.

Usually, the spectral signature is the main aspect of the classes used to classify the pixels. Unlike traditional pixel-based methods, contextual technique considering both spectral and spatial information in order to perform the classification process instead of depending on spectral component alone. The integration of spatial information in the image classification is expected to improve classification accuracy. The contextual classifier uses pixel-centred window to estimate the density function associated to a pixel. Selection of the window size is very important in contextual classification. Pixel-window size determines the amount of spatial information that can be included in the classification. Since optimal pixel window varies with individual class and image resolution, it is usually difficult to determine before image classification. Therefore, an appropriate window size is usually determined empirically (Huang et al., 2007). In term of the training data, contextual classifier do not required large number of training set as this classifier is not a statistical method.

**Methodology:** Present study focuses on extracting land cover by adopting ML and contextual approaches for SPOT-2 data. The reason for implementing this mechanism is to perform comparison between the traditional and advanced approaches in land cover classification. The results of these methods were evaluated using the available field information on land cover and by visual interpretation. In the Mina city, 4 classes were selected to represent and classify the image namely Urban, Mountain, Mina Tent and Shadow. Although the shadow appears mostly in the mountainous area, the authors decided to separate them as a new class due to their obviously different in their spectral information against Mountain. The description for each of the land cover category is stated in Table 1. The ground resolution for this satellite is 20 m and the image was acquired using High Resolution Visible (HRV) sensors carried on the SPOT satellite. It has three bands ranging from visible to near infrared portion. The total work for present study has been executed using PCI Geomatica 10.3 image processing software packages. The size of data set used in the image

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mina tent</td>
<td>Fireproof Teflon coated glass tents (white color) in Mina City.</td>
</tr>
<tr>
<td>Mountain</td>
<td>Hills, large rocks, rugged terrain</td>
</tr>
<tr>
<td>Shadow</td>
<td>Appearing due to the high mountain or building</td>
</tr>
<tr>
<td>Urban</td>
<td>Residential, commercial, building, roadway, infrastructures, concrete and any developed areas</td>
</tr>
</tbody>
</table>

Fig. 2: The raw data of the study area (SPOT-2)
processing is 450x350 pixels and is presented in Fig. 2. Prior to the image analysis, image pre-processing step was applied to make the images comparable before classification. A composite colour has been done on this set. The aim of this pre-processing is to have a better visual interpretation of the scene and to be able to identify representative areas which will constitute a training base for the supervised process. Two sets of training data were used in order to perform the classification. One set of training data was used for ML approach and the other set for Frequency-based Contextual (FBC) classifier. The use of different set of training data is due to the ML approach needs more training samples as it is a statistical classifier so that it can generate the statistical distribution of each class for classification. Unlike ML approach, contextual do not required to have many training data as this classifier is not a statistical approach. The supervised classification of digital data is carried out after creating training sites and the task was performed and the classification was carried out using ML and FBC classification methods (Gong and Howarth, 1992). Basically, there are many algorithms (e.g., minimum distance, neural network, decision tree etc.) can be used in image classification but in this study only these two classifiers were chosen for the classification process. The selection of the classifier is very important in order to obtain good classification accuracy. The classification accuracy is most important aspect to assess the reliability of maps, especially when comparing different techniques. During present study the accuracy has been estimated on the training set samples and the results of all the classification techniques is summarized in the table. A total of 190 points were selected randomly in order to test the accuracy assessment of the classification result. A statistical classification assessing is carried out by means of error matrix establish between truth ground and obtained classifications. Performance of the obtained classifications is evaluated by calculating kappa parameters which is a usual indicator in land cover classification analysis. The error matrices related to the classification of Fig. 4a and b are given in Table 2 and 3. The methodology used in this project is summarized in Fig. 3.

**Statistical analysis**

**An accuracy assessment:** The accuracy assessment sites were used to provide a statistically sound assessment of the accuracy produced by each of the automated mapping approaches tested for this project. The accuracy assessment sites were set aside until the map was completed, and accuracy assessment was performed. This process insured that the accuracy data were completely independent of the training data (Thomas et al., 2003). A common method for classification accuracy assessment is through the use of an error matrix (also called confusion matrix). For a classified image, an error matrix can be made by comparing the classification results with reference data. In this matrix, the reference data are represented by the columns of the matrix while the classified data are represented by the rows. The major diagonal of the confusion matrix indicates the agreement between these two data sets. It is typical to extract several statistics from the error matrix: overall accuracy, producer's accuracy, and user's accuracy. Congalton (1991) used the terms Producer Accuracy (PA) and User Accuracy (UA) to describe within class measures of classification accuracy and thus provide a breakdown of the figure for overall accuracy in error matrix. PA gives the analyst an estimate of how successful the classification procedure is in the different classes. UA gives the user an estimate of how reliable the thematic map is as a predictive device in the different classes.

**RESULTS AND DISCUSSION**

From the matrices, the kappa coefficient, overall classification accuracy and accuracy of each class is easily calculated. The resulted classified imagery using context is finding to reveal meaningful patterns. Visual interpretation between classified images reveals that
image produced by contextual approach is better than ML. The major improvement of the overall accuracy between the conventional ML classifier and the FBC classification method is the speckle error. The high content of noise in Fig. 4a produces a less precise of the classification accuracy than in Fig. 4b. Overall accuracy of contextual classification was around 11 percent better compared to the ML classification. It was shown that the involvement of the spatial information, speckle error (salt and pepper appearances) can be reduced significantly. This visual interpretation is confirmed by

the statistical information given on Table 2 and 3. Overall accuracy of FBC classification with the integration of spatial information leads to impressively improved results, up to 83% of accuracy with 0.693 kappa value is achieved in comparison with the output derived from traditional ML classifier where only around 72% of accuracy with 0.527 kappa is obtained. This result confirms the experience that the incorporation of contextual information tends to become more accurate than statistical ML approach (Gong and Howarth, 1992; Stuckens et al., 2000; De Jong et al., 2001; Jackson and Landgrebe, 2002). For instance, Gong and Howarth (1992) reported that an improvement of classification accuracy of 15-20% can be achieved using the frequency-based classifier over the ML approach. In the meantime, the incorporation of contextual information in the classification process improved accuracy by 5.8% as stated by Stuckens et al. (2000) and Jackson and Landgrebe (2002), they could increase 13% of accuracy when using contextual technique compared to ML approach. In addition, De Jong et al. (2001) resulted 80.1 and 92.7% of classification accuracy for ML and contextual, respectively, hence, the contextual outperform ML classifier by approximately 12%.

In addition, the accuracies for each class were presented in the error matrix tables and can be found by evaluating the user and producer accuracies column. For ML approach, the user accuracy varied between 60.8% for Urban and 85.3% for Mountain. Mina Tent and Shadow has an accuracy of 75.0 and 83.3%, respectively (Table 2). In the meantime, the user accuracy for the FBC method varied between 73.9% for Urban and reached the

Table 2: Confusion matrix table derived from Maximum Likelihood classifier

<table>
<thead>
<tr>
<th>Class</th>
<th>Mina tent</th>
<th>Mountain</th>
<th>Urban</th>
<th>Shadow</th>
<th>Total</th>
<th>UA(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mina tent</td>
<td>9</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>12</td>
<td>75.0</td>
</tr>
<tr>
<td>Mountain</td>
<td>0</td>
<td>64</td>
<td>11</td>
<td>0</td>
<td>75</td>
<td>85.3</td>
</tr>
<tr>
<td>Urban</td>
<td>1</td>
<td>36</td>
<td>59</td>
<td>1</td>
<td>97</td>
<td>60.8</td>
</tr>
<tr>
<td>Shadow</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>93.3</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>102</td>
<td>72</td>
<td>9</td>
<td>190</td>
<td>-</td>
</tr>
<tr>
<td>PA (%)</td>
<td>59.0</td>
<td>62.7</td>
<td>81.9</td>
<td>83.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Overall:</td>
<td>72.1%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.527</td>
</tr>
</tbody>
</table>

Table 3: Confusion matrix table derived from Frequency-based Contextual classifier

<table>
<thead>
<tr>
<th>Class</th>
<th>Mina tent</th>
<th>Mountain</th>
<th>Urban</th>
<th>Shadow</th>
<th>Total</th>
<th>UA(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mina tent</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>100.0</td>
</tr>
<tr>
<td>Mountain</td>
<td>0</td>
<td>99</td>
<td>11</td>
<td>1</td>
<td>111</td>
<td>89.2</td>
</tr>
<tr>
<td>Urban</td>
<td>5</td>
<td>13</td>
<td>51</td>
<td>0</td>
<td>69</td>
<td>73.9</td>
</tr>
<tr>
<td>Shadow</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>7</td>
<td>85.7</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>113</td>
<td>62</td>
<td>7</td>
<td>190</td>
<td>-</td>
</tr>
<tr>
<td>PA (%)</td>
<td>87.5</td>
<td>87.6</td>
<td>82.3</td>
<td>85.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Overall:</td>
<td>83.7%</td>
<td>85.3</td>
<td></td>
<td></td>
<td>85.7</td>
<td>0.693</td>
</tr>
<tr>
<td>*PA: Producer accuracy, UA: User accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4 (a-b): Classified land cover map over Mina area using (a) Maximum Likelihood classification and (b) Frequency-based Contextual classification
highest value of 100\% for Mina Tent, 85.7 and 89.2\% were obtained from Shadow and Mountain classes (Table 3). This statistical evaluation shows that all classes have better accuracy when performed by FBC method compared to ML approach. For producer accuracy, the accuracy for each class using ML approach was as follow: 90.0\% for Mina Tent, 62.7\% for Mountain, 81.9\% for Urban and 83.3\% for Shadow (Table 2) whereas the results obtained from FBC method were 37.5, 87.6, 82.3 and 85.7\% for the same sequence (Table 3). Again FBC has a better result compared to ML approach in all classes except the Mina Tent class. For this class, ML classifier performed better than contextual approach as shown by the result from visual interpretation as well as from error matrix table. In FBC method, most of the Mina Tent class was classified as urban class due to the high pixel confusion among the classes. Overall, the classification result performance of the ML method is poor, especially in terms of delineating built up area from surrounding features in the arid environment. In arid zone of Saudi Arabia, the urban areas are often surrounded by mountain. However, they may also confuse with nearby bare soil and stony desert which present very similar spectral characteristics as construction materials such as concrete. The urban environment represents one of the most challenging areas for remote sensing analysis due to the high spatial and spectral diversity of surface materials.

The detail explanation of the each class can be explained by analyzing the error matrix table. Table 2 presents the error matrix for the ML classifier. The matrix shows that 64 out of 75 points of the Mountain class had been correctly classified. Only 11 points had been classified wrongly which is all points were misclassified as Urban. For the Urban category, it is shown that out of 97 pixels that had been tested, 59 pixels were correctly classified and most of the remaining pixels were misclassified as Mountain. In addition, the same pattern was obtained for FBC method where Mountain and Urban classes were confusion among each other. For the Mountain, 99 out of 111 observations had been correctly classified and 11 points were misclassified as Urban. A total of 51 out of 69 observations had been correctly classified for Urban and 18 points were wrongly classified with 13 points were misclassified as Mountain (Table 3). Although, FBC has a better result but both of the classifiers have the difficulty to classify these classes due to the spectral confusion among them is high. This situation happens due to the fact that the mountainous area in the arid environment is not cover by tree but it is filled by stones and rocks which is has a similar spectral characteristic of urban area. On the other hand, both classifiers perform well for the Mina Tent and Shadow classes as their spectral characteristics significantly different against other features. For Mina Tent, 9 out of 12 observations were correctly classified for ML (Table 2) whereas all points were correctly classified for FBC (Table 3). From 6 points that have been tested for Shadow class on ML approach, 5 points were correctly classified and only a point was misclassified for FBC method (6 out of 7 points) for this class (Table 3).

The traditional pixel-based classification typically yielded large uncertainty in the classification results as it cannot handle the complex landscape environment. Aim to this problem, the contextual image classification was processed. In contextual classification method, the analyst can adjust the classification process by changing the growth of the kernel window. In the meantime, 11×11 windows were used to classify the images for this assessment. Selection of the window size is performed by ‘trial and error’ basis and the 11×11 window was chosen as the best window to perform the classification. Essentially, selection of the window size is depending on the complexity of the image. If the image is more complex, hence, FBC needs the larger window for classification process. Otherwise, smaller window is enough for the smooth images.

The analysis presented in present study also shows that mixed pixels cause spectral confusion and that such pixels are sometimes assigned to the wrong thematic land cover class. The contextual method makes it possible to identify mixed pixels in an image and assign these mixed pixels to the proper land cover class. Basically, mixed pixels occur at the border of the classes and in the complex landscape environment such as in the urban/built up area. In a per pixel classification, these pixels are often wrongly labelled due to the mixed spectral signature of the pixels. By using ML classifier, some of the Urban is classified as Mountain as their spectral signature resembles. Results of the classified image by using contextual approach show that image classification can significantly improve by capturing spatial information in the classification procedure for the same area.

**CONCLUSIONS**

Present study underlines the performances of two widely used classification techniques for classification of Mina area using SPOT-2 satellite image. The performance of FBC classification technique is reliable assessing high accuracy followed by ML. The landscape of Mina city is diverse and complex, comprising both homogeneous and heterogeneous surface features, causing problems of spectral variability in the satellite image data. In general, the use of contextual information indicated larger
improvements in the classification accuracy. For present study, the use of the contextual information could be improved the classification accuracy up to approximately 11%. The accuracy assessment report showed that FBC algorithm can predict land use and land cover of the complex urban environment more accurately. It also means this classifier shows great potential for dealing with heterogeneous surface features in urban areas. In conclusion, overall performance of FBC is better than ML but in certain homogeneous surface such as Mina Tent, ML performed better than FBC.

ACKNOWLEDGMENTS

This research was funded by USM-RU-PRGS grant (account no: 1001/PFIZIK/831021) and Hajj Research Cluster USM grant from Universiti Sains Malaysia. The authors would like to thank the technical staff that participated in this project. Thanks are extended to USM for support and encouragement.

REFERENCES


