EXTRACTION INFORMATION OF ENVIRONMENTAL CHANGES FROM SATELLITE IMAGE

Nang Mya Mya Nwe, Myint Myint Sein
[1]Room-23, Mann-cherrry Hostel, Mandalay Technological University
[2]Professor (Visiting), University of Computer Science, Yangon, Myanmar

Abstract

Nowadays, Remote Sensing (RS) technology is widely used in various fields such as: Environmental Studies, Thematic map, water resource management etc. Interested Information such as: the urban area, vegetation, water and bare soil etc. will be extracted from remotely sensed images by using Digital Image Processing.

Hidden Markov Model (HMM) is constructed for training the image based on the RGB color values and identifying the classes of satellite images. In the training stages, HMMs are dynamically assembled according to the class sequence. The basic theory of HMMs is easy to understand. This makes it easier to analyze and develop implementations for the recognition. To extract the information in the multi-date satellite images, both the Maximum Likelihood (ML) classifier and Minimum Distance classifier (MD) are used in training and also in recognition. Therefore, the precise training is received for the database and the recognition is processed accurately. Using ML classification, pixels are assigned to a specific class if they fall within the box regions, which are defined by the training data, and are allocated to the appropriate categories. This classifier reduces the possibilities of any pixels being unclassified. The MD classification uses the mean vectors of each region of interest (ROI) and calculates the Euclidean distance from each unknown pixel to the mean vector for each class. All pixels are classified to the closest ROI class unless the user specifies standard deviation or distance thresholds, in which case some pixels may be unclassified if they do not meet the selected criteria.

The main purpose of this paper is to evaluate and compare the images with results from the supervised training based methods.

Key words: remotely sensed image, extract information, Maximum Likelihood classifier, Minimum Distance classifier, Hidden Markov model

1. Introduction

Remotely sensed data are increasingly used for mapping and monitoring the physical environment. One of the advantages of monitoring with remotely sensed data is that temporal sequences of images can accurately indicate environmental changes, assuming that digital values are radiometrically consistent for all scenes. Factors contributing to the potential inconsistency in measured radiance include changes in surface condition, illumination geometry, sensor calibration, observation geometry and atmospheric condition (Jensen et al. 1995).

Advances in sensor technology for Earth observation make it possible to collect multispectral data in much higher dimensionality. In addition, multisource data also will provide high dimensional data. Such high dimensional data will have several impacts on processing technology: (1) it will be possible to classify more classes; (2) more processing power will be needed to process such high dimensional data (3) with large increases in dimensionality and the number of classes, processing time will increase significantly.

The analysis of remotely sensed data is usually done by machine oriented pattern recognition techniques. One of the most widely used pattern recognition techniques is classification based on maximum likelihood (ML) assuming Gaussian distributions of classes. A problem of Gaussian ML classification is long processing time. This computational cost may become an important problem if the remotely sensed data of a large area is to be analyzed or if the processing hardware is more modest in its capabilities. The advent of the future sensors will aggravate this problem. As a result, it will be an important problem to extract detailed information from high dimensional data while reducing processing time considerably (Chulhee Lee, David Landgrebe).

Efforts to reduce processing time have been pursued in various ways. By employing feature selection/extraction algorithms, the number of features can be reduced substantially without sacrificing significant information (Muasher, M. J. and D. A. Landgrebe). Feature selection/extraction is generally done by removing redundant features or by finding new features in transformed coordinates. This reduction in the number of features has several advantages. First of all, higher accuracies can be achieved in cases where the number of training samples is low, due to the Hughes phenomenon (Hughes, G. F.). Since generally processing time increases with the square of the number of features, a benefit of feature selection/extraction is reduction in processing time.

Another possible approach to reduce computing time can be found in decision tree classifiers (Swain, P. H. and H. Hauska). Though the decision tree classifier can have several advantages depending on the situation, one of the advantages is processing time. However, how to find the optimum tree structure still remains a problem for the decision tree classifier, though many algorithms are proposed for the design of decision tree classifiers (Argentiero, P., R. Chin, and P. Beaudet).

In this paper, ML and MD classifiers are used for the high-resolution satellite image. Using ML classification, pixels are assigned to a specific class if they fall within the box regions, which are defined by the training data, and are allocated to the appropriate categories. This classifier reduces the possibilities of any pixels being unclassified. The MD classification uses the mean vectors of each region of interest (ROI) and calculates the Euclidean distance from each unknown pixel to the mean vector for each class. All pixels are classified to the closest ROI class unless the user specifies standard deviation or distance thresholds, in which case some pixels may be unclassified if they do not meet the selected criteria.

The aim of this paper is to present the preprocessing of the satellite imagery, land cover change detection methods and results and to evaluate for the best classification and compare the classified images with results from the supervised training based methods.

2. Change Detection

Change detection technology, which discovers the change information on the surface of the earth by comparing and analyzing multi-temporal satellite images, can be usefully applied to the various fields, such as environmental inspection, urban planning, forest policy, updating of geographical information and the military usage. Especially, change detection methods with high-resolution satellite imagery are very useful for the missions of inspecting on the earth such as environmental monitoring, circumstantial analysis of disaster damage, inspection of illegal buildings, the military use and so on, which cannot be achieved by low- or middle-resolution satellite imagery.

As follows, there are general change detection processes to detect change between multi-temporal images (Ross, 1998).

- Data Acquisition and Preprocessing
- Radiometric/Geometric Co-registration
- Change Detection Analysis
- Accuracy Assessment
- Final Product Generation

To get higher-quality change detection results, we have to minimize any other noise factors by selecting multi-temporal image pairs that have similar photographing conditions, such as atmospheric conditions, variation in the solar illumination angles, and sensor calibration trends. It is, however, very difficult to maintain radiometric consistencies between images due to these different photographing conditions. Therefore, radiometric co-registration should be done to remove these noise effects between multi-temporal images (Yong, 2002).

3. Multispectral Classification

The multispectral images were highly responsible for a good classification of land use. The focus lies on an acceptable reliability of class discrimination and less on the geometric accuracy of small details. Although only wooded areas and settlements were the key classes of the project, one has to introduce additional classes that are able to describe the land use of the area of interest in a not necessarily complete but at least satisfactory way. The classes (and subclasses) defined for the areas under investigation were the following:

- Water (water bodies such as lakes, rivers)
- Forest (coniferous and deciduous forest)
- Settlement (urban areas, villages, hamlets and possibly individual farm houses)
- Field (agriculturally used ploughed and unploughed fields)
- Grassland (meadows, pastures)
- Rock (high alpine areas without vegetations)
- Glacier

The classification has been performed according to the decision rules of the well-known maximum likelihood algorithm. One knows that the quality of a classification increases if the multispectral signatures of the classes involved form a homogeneous and normally distributed cluster. Heterogeneous classes, such as urban settlements that are mixtures of sealed, unsealed areas, vegetation etc., lead to huge and quite often non-normally distributed clusters. The flow diagram is shown in figure 1.

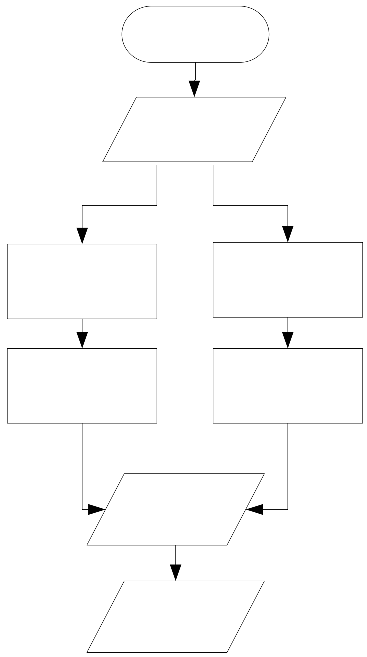


Figure 1. Block Diagram

4. Methods

4.1. Hidden Markov Model

HMM is used because it uses only positive data, they scale well. The basic theory of HMMs is also very elegant and easy to understand. This makes it easier to analyze and develop implementations for. In this paper, the classification has been performed according to the decision rules of the well-known maximum likelihood algorithm plus the minimum distance classifier.

HMMs are efficient at modeling real world processes that exhibit “sequentially changing behavior” (Argentiero, P., R. Chin, and P. Beaudet). i.e. These processes are characterized by:

- a period of steady behavior with minor variations over time, and
- a period of gradual change to another steady plane.

A HMM is one of many signal models used to describe a sequence of observable symbols (discrete or continuous) produced by a real world process. Once the model is successfully built, it can be used to identify (recognize) another sequence produced by the same process. This signal model can potentially tell us a lot about the real life process that created the signal.

In order to define an HMM completely, following elements are needed (PHYBME).

- the number of states of the model, *N*.
- the number of observation symbols in the alphabet, *M*. If the observations are continuous then *M* is infinite.
- A set of state transition probabilities

$$A = \{a_{ij}\} \quad (1)$$

$$a_{ij} = p\{q_{t+1} = j / q_t = i\} \quad (2)$$

Where q_t denotes the current state. Transition probabilities should satisfy the normal stochastic constraints,

$$a_{ij} \geq 0, 1 \leq i, j \leq N \quad (3)$$

and

$$\sum_{j=1}^{N} a_{ij} = 1, 1 \leq i \leq N \quad (4)$$

- A probability distribution in each of the states,

$$B = \{b_j(k)\} \quad (5)$$

$$b_j(k) = p\{o_t = v_t / q_t = j\},\ 1 \leq j \leq N, 1 \leq k \leq M \quad (6)$$

where v_t denotes the k^{th} observation symbol in the alphabet, and o_t the current parameter vector.

4.2 Maximum-likelihood Algorithm

Classification of data is an important task in many fields. Parametric approaches to classification use models of data, for which a set of parameters can be estimated. A popular technique is maximum likelihood estimation (MLE) (Richard O.Duda and Peter E. Hart). A classifier using estimates of model parameters calculates the probabilities that each new datum to the class for which the posterior predictive probability *P(c/x, X)* is the greatest. This is the approximate Bayes classifier (Richard O.Duda and Peter E. Hart). The parameters of the model are $\Theta = \{\Theta_D, \Theta_p\}$ where Θ_D refers to the data, and Θ_P refers to the class priors. Θ may be estimated by maximizing the joint likelihood

$$\begin{aligned} L_j(\Theta) &= \log p(X,C/\Theta) \\ &= \log p(X/C,\Theta)P(c/\Theta \\ &= L_n(\Theta_D) + P(c/\Theta_p) \end{aligned} \quad (7)$$

Where C are the class labels of the class-labeled training data:

$$X = \{X_c\}_c^n = 1 \quad (8)$$

$$X_c = \{x_{jc}\}_c^{N_c} = 1 \quad (9)$$

such that each datum x_{jc} has an accompanying class label c, where c Є 1…n and there are N_c data labeled as belonging to class c. The set of parameter estimates Θ is the only information available to the classifier about the source distribution of the data, it is clearly desirable that the estimate be a good representation of the data in some sense. On the other hand (PHYBME) , point out the "we want the estimate that leads to the best classification performance" and "no single estimate possesses all of the properties one might desire". The motivation for using a discriminative likelihood function is the notion that selecting parameter estimates which maximize $L_d(\Theta)$ is likely to result in similarly good performance in classifying previously unseen data from the same source distribution. This will be tested by comparing relevant attributes of $L_n(\Theta)$ and $L_d(\Theta)$, with some results from applying the two methods to suitable data sets.

Using ML classification, pixels are assigned to a specific class if they fall within the box regions, which are defined by the training data, and are allocated to the appropriate categories. This classifier reduces the possibilities of any pixels being unclassified.

4.3. Minimum Distance Algorithm

The MD classification uses the mean vectors of each region of interest (ROI) and calculates the Euclidean distance from each unknown pixel to the mean vector for each class. All pixels are classified to the closest ROI class unless the user specifies standard deviation or distance thresholds, in which case some pixels may be unclassified if they do not meet the selected criteria. This method calculates the center for each group of pixels and measures the distance from the center of each group to the pixel being considered. The pixel is classified to the group with the nearest center. The center of the group then is recalculated each time a pixel is added or taken away. Feature vector is computed for these samples again and a distance Euclidean and Mahalanobis classifiers are used to classify the unknown samples.

Minimum Euclidean distance:

$$d_j(x) = \sum_{q-1}^{Q} (x_q - m_{i,q})^2 \tag{10}$$

Mahalanobis distance

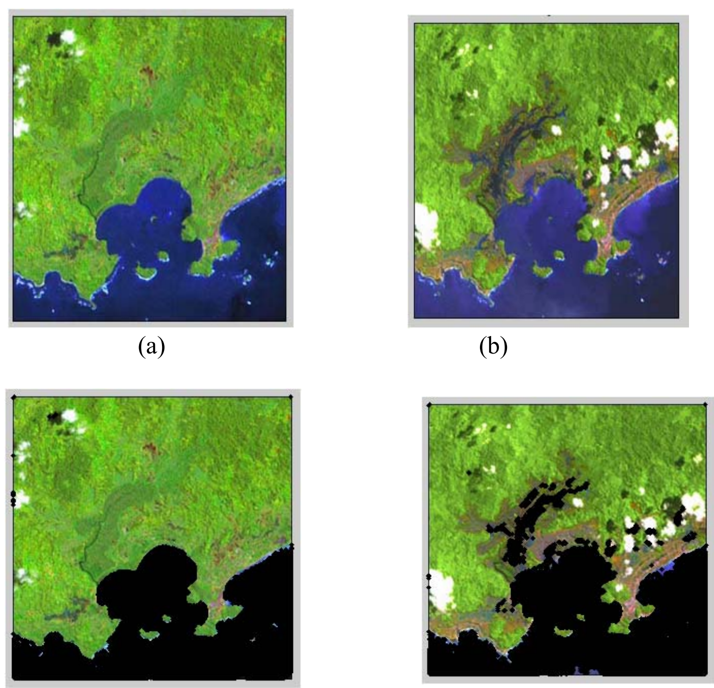
$$d_j(x) = \{(x - m)^2 C_j^{-2} (x - m_j)\} \tag{11}$$

x is the feature vector to be classified, m is the mean value of the training vectors, m is the transpose, C is the covariance matrix. The classifiers are used to recognize 32 different unknown samples from each of the 6 color images. The results are shown comparing performance of Euclidean classifier and Mahalanobis classifier with and without smoothing. The Euclídean classifier performs poorly compared to Mahalanobis with or without smoothing. The results are better with Mahalanobis classifier which is also computationally more efficient.

The minimum distance method is used for the second classifier for the classification. Mahalanobis distance is the most commonly used and is a very useful way of determining the similarity of a set of values from an unknown sample to a set of values measured from a collection of known samples. One of the main reasons the Mahalanobis distance method is used is that it is very sensitive to inter-variable changes in the training data. In addition, since the Mahalanobis distance is measured in terms of standard deviations from the mean of the training samples, the reported matching values give a statistical measure of how well the spectrum of the unknown sample matches (or does not match) the original training spectra.

5. Experiments and Results

The image is firstly trained via the expert with the appropriate training setting. From the trained image, both classifiers are used to extract the contents in the image. This experience can estimate four feature types of remotely sensed image. These four features are water, urban, bare soil and vegetation. It would be very difficult to find a threshold, or decision surface, with which to segment the images into training classes (e.g. spectral classes which correspond to physical phenomena such as cloud, bare soil, water, etc.). In this paper, water region is classified in both two-date satellite images. Figure 2 (a) and (b) show the original image. These two images are firstly trained with ML and MD algorithms and classified through those classifiers. The water region is selected only as the classified data. Figure (c) and (d) show the classified water region. Figure 3. (a) and (b) show the image of Indonesia island before and after Tsunami affect. (c) shows the change reigion of the two images.



(a) (b)

(c) (d)

Fig 2. The Original Image (a) the previous date image (b) the last date image (c) (d) The detection of water region

(a) (b) (c)

Fig 3. The detection of environmental change

4. Result

The result from the classified image is estimated as percentage of the whole image area. This is shown in the table 1.

Table 1. The water region classified from the two-date Satellite Image

	Water Region	Other (Forest,ect.)	Change Region
Image 1(Fig 2)	33.67 %	66.33 %	-
Image 2(fig 2)	37.05 %	62.95 %	-
Images (fig 3)	-	-	43.63%

5. Conclusion

Results of this work could be help scientists to look and study environmental changes that could affect the globe in the future conditions like pollution, global climate change, natural resource management, urban growth, and much more and trend across large geographic areas. The limitation for the system is that the user must train the image before the classification is processed. The training data is only for a single application.

6. References

[1]. Argentiero, P., R. Chin, and P. Beaudet, *"An Automated Approach to the Design of Decision Tree Classifiers,"* IEEE Trans. Pattern Analyses and Machine Intelligence, Vol. PAMI-4, No. 1, pp. 51-57, January, 1982.

[2]. Chulhee Lee, David Landgrebe, *Feature Extraction And Classification Algorithms For High Dimensional Data*, School of Electrical Engineering Purdue University West Lafayette, Indiana 47907-1285, TR-EE 93-1, January, 1993.

[3]. Hughes, G. F., *"On the Mean Accuracy of Statistical Pattern Recognizers,"* IEEE Trans. Information Theory, vol. IT-14, No. 1, pp. 55-63, 1968.

[4]. JENSEN, J.R., RUTCHEY, K., KOCH, M.S. and NARUMALANI, S., 1995, Inland wetland change detection in the Everglades Water Conservation Area 2a using a time series of normalized remotely sensed

data. Photogrammetric Engineering and Remote Sensing, 61, pp. 199–209.

[5]. L. R. Rabiner and B. H. Juang, “An introduction to hidden Markov models,” *IEEE ASSP Mag*., vol. 3, no. 1, pp. 4-16, 1986.

[6]. Muasher, M. J. and D. A. Landgrebe, *"The K-L expansion as an effective feature ordering technique for limited training sample size,"* IEEE Trans. Geoscience and Remote Sensing, Vol GE-21, pp. 438-441, 1983.

[7]. PHYBME. Hidden Markov Model used in satellite image processing (node 3). http://jedik.phybme.hu/`gerjanos/HMM/node3/ (accessed Dec 3, 2006)

[8]. Richard O.Duda and Peter E. Hart. *Pattern Classification and Scene Analysis.* Wiley-Interscience, New York, 1973.

[9], Ross S. Lunetta, et al., 1998. *Remote Sensing Change Detection, Ann Arbor Press*, USA, pp. 1~19.

[10]. Swain, P. H. and H. Hauska, *“The decision tree classifier: Design and potential,”* IEEE Trans. Geoscience Electronics, vol. GE-15, pp. 142-147, 1977.

[11]. Yong Du, et al., 2002. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sensing of Environment*, 82, pp.123~134.