Change Detection Using Unsupervised Learning Algorithms for Delhi, India

Hemant Kumar Aggarwal* and Sonajharia Minz

1Indraprastha Institute of Information Technology-Delhi, India
2Jawaharlal Nehru University, Delhi, India

Abstract

The effectiveness of the three types of unsupervised learning techniques for change detection in water, vegetation and built-up land cover classes of a part of Delhi region in India has been analyzed. A total of eight images of Landsat TM and ETM+ from year 1998 to 2011 were preprocessed for atmospheric corrections. Subsequently three features, Soil Adjusted Vegetation Index (SAVI), Modified Normalized Difference Water Index (MNDWI), and Builtup from Normalized Difference Built-up Index (NDBI) were extracted at the preprocessing stage. The three clustering algorithms kmeans, fuzzy c mean and expectation-maximization were selected to represent the partition based, fuzzy, and probability based technique respectively. The three algorithms were implemented to cluster the pixels of all the eight images using the features SAVI, MNDWI and NDBI. The Silhouette coefficient was used to evaluate the cluster quality that takes into consideration both intra-cluster and inter-cluster distance between clusters. The outcome of clustering has been quantified in terms of the percentage of total pixels grouped in each of the three clusters indicating vegetation, urban and water. Change detection has been performed comparing the outcomes of clustering done on each of the eight images.

Key words: Change detection, unsupervised learning, silhouette coefficient.

1. Introduction

The land cover refers to the characteristics and surface cover of the earth represented by vegetation, forest, water, bare Earth, desert, urban area and such other physical features of the land. Land cover classification is a well-studied problem in the domain of remote sensing (Ehsani and Quiel (2010); Hung et al. (2011); Mukherjee et al. (2009); Jensen (2005)). Earlier visual interpretation of satellite images were carried out to make various thematic maps. However, many effective methods and techniques for automatic classification enable the researchers to generate thematic maps automatically. Land cover classification is a challenging problem because terabytes of data are acquired every day. Processing such a large amount of data has attracted the research attention as it is difficult and time consuming. It also depends on the knowledge of the individual domain expert. Various Change detection techniques have been reviewed in Coppin and Bauer (1996); Singh (1989). These papers discuss various pre-processing requirements of the images acquired by remote devices, factors which causes changes in the land cover and the different change detection techniques.

The classification of satellite imagery have many applications in real world. Land covers such as high mountains, large seas, dense forests, big deserts, deep valleys etc., all can be imaged using satellites. The image classification can provide more information about characteristics, changes, increase or decrease in these areas. A further analysis of the information has been helpful to study their bearing on mankind. Effective classification techniques for satellite images may even prove to be helpful to detect the forest fires. A time series analysis of the imagery may help develop forewarning system for flood, spread of diseases in crops etc. Most of these applications require ground data for supervised classification but in many cases it may be difficult to collect ground data. Therefore, for such applications unsupervised learning techniques may prove to be useful.
Land covers are not disjoint due to which there is inherent fuzziness in the land cover classes. Based on the resolution of the image a single pixel of satellite image may cover a large area including features of many classes. The mixed pixel problem may be handle by considering fuzzy technique for clustering a pixel to belong to more than one class. A single land cover class such as urban and vegetation generally have high intra-class variability due to which different clusters are yielded even though they belongs to same class. This problem was taken care of by visually identifying such clusters and then merging them together. Land cover areas high spatial auto-correlation which means neighboring pixels more likely get grouped into the same category. Spatial auto-correlation have also been addressed by using Majority analysis.

2. Related Work

The change detection in Limpopo province of South Africa has been studied in salmon et al. (2011) on time series data extracted from MODIS imagery. Three unsupervised learning techniques: Ward clustering, k-means, and expectation-maximization, were used for classification of sub-sequences of the time series data. In Mukherjee et al. (2009) the effect of canal on land use and land cover in an area of Punjab, India has been discussed. The authors applied four classification techniques namely: ISODATA, maximum likelihood classifier, NDVI based approach and spectral mixture analysis. Land cover classification efficiency of Landsat ETM+ thermal band has been examined in Ehsani and Quiel (2010) using Sequential Maximum a Posteriori (SMAP) segmentation algorithm. Optimum index factor was used to select the bands with highest information content. The authors observed that classification accuracy increased with the use of thermal band while it has lower spatial resolution. SMAP is a segmentation algorithm that takes in account contextual information by using spatial information of spectral bands and also uses multi scale random field(MSRF), there by increases the accuracy of classification. Many studies are there about spatial and spectral features extraction from Landsat satellite imagery using various transformations, filtering in frequency domain and band ratios The paper Xu (2007) has discussed extraction of urban built-up features from satellite imagery using logic-based method, classification based method and second principle component. Procedure for converting digital numbers to at-satellite reflectance have been described in Chander et al. (2009). This is a two step procedure. In the first step the digital numbers are converted to radiance and radiance is converted to the reflectance values in the second step.

3. Data Description

Eight Landsat TM5 and ETM+ images with path/row 146/40 have been used to carry out the experiments. All images were taken in the month of October month with five images from year 1998 to 2002 and three images from year 1999 to 2011. These images are freely available from USGS and Earth-explorer websites USGS (2012). Total area under analysis is 96.8 km². All the images are of UTM projection with zone 43N and WGS-84 datum. The upper left co-ordinates of each image are 28°39′15″N, 77°11′51.81″E and bottom right co-ordinates are 28°33′46.61″N, 77°17′36″E. Landsat data is already geo-referenced which means that images were prepared as registered with each other. Delhi shape-file was used to extract Delhi from original Landsat image. Thermal band have not been considered in this study because it have high spatial resolution.

4. Experiments

All the experiments have been performed using ENVI 4.8 and Matlab software. The details regarding the experiments, the data preprocessing, feature extraction, clustering, post-processing, and cluster validation have been presented in the related subsections.

4.1 Data Pre-processing

The radiometric calibration of the data was done to reduce the noise due to the differences in atmosphere, sensor calibration, and sun-angle at different dates(Coppin and Bauer (1996)). The following equation in Chander et al. (2009) has been used in the paper for radiometric calibration:

\[ L_\lambda = \frac{L_{max}}{Q_{max}} - \frac{L_{min}}{Q_{min}} (Q - Q_{min}) + L_{min} \]  

where, \( L_\lambda \) is the spectral radiance at the sensor’s aperture in \( W/(m^2sr\mu m) \), \( L_{max} \) is the spectral radiance that is scaled to \( Q_{max} \), \( L_{min} \) is the spectral radiance that is scaled to \( Q_{min} \), \( Q \) is quantized calibrated pixel values in digital numbers (DNs), and \( Q_{max} \) is maximum quantized calibrated pixel value corresponding to \( L_{max} \), \( Q_{min} \) is minimum quantized calibrated pixel value corresponding to \( L_{min} \).

All these values are provided in the header file.hdr associated with each image. After calculating the radiance, the at-satellite reflectance is determined. It reduces the between scene variability through the normalization by solar irradiance:

\[ \rho_\lambda = \frac{\pi L_\lambda d^2}{E_{sun,\lambda} \cos(\theta)} \]

where, \( \rho_\lambda \) is a unitless planetary reflectance at the wavelength \( \lambda \), \( E_{sun} \) is mean solar exo-atmospheric irradiances, \( d \) is Earth-Sun distance in astronomical units, and \( \theta \) is solar zenith angle in degrees. In Figure 1(a) an area of interest in false color composite (FCC) image has been presented. The image is the result of preprocessing the data for at-sensor calibration.

4.2 Feature Extraction

The three features namely, Soil Adjusted Vegetation Index(SAVI), Modified Normalized Difference Water Index (MNDWI), and Bulitup. These features corresponds to vegetation, water, and built-up classes respectively. SAVI (represented by \( \text{SAR} \)), Modified Normalized Difference Water Index (MNDWI), and Bulitup. These features are extracted from MODIS imagery. Three unsupervised learning techniques: Ward clustering, k-means, and expectation-maximization, were used for classification of sub-sequences of the time series data.
The radiometric calibration of the data was done to reduce the noise due to the differences in atmosphere, sensor calibration, and sun-angle at different dates (Coppin and Bauer (1996)). The following equation in Chander et al. (2009) has been used in the paper for radiometric calibration:

\[ L = \frac{\rho_2 - \rho_4}{\rho_5 + \rho_4} \]  

(5a)

\[ B = N - S \]  

(5b)

With the extraction of the three features from the imagery a 50% reduction in data was also achieved. This was achieved as the feature extraction resulted in a triple (SAVI, MNDWI and Builtup) reducing the dimensionality of data to three as opposed to six dimensions (bands) used to represent the raw imagery. The Figures 1(b), 1(c), 1(d) represent the images based on the three extracted features SAVI, MNDWI and Builtup respectively. The white portions in the three images represent high density of the feature whereas darker shade to black portions indicate reduced or no feature values. The FCC image yielded by combining the three features is presented in Figure 1(e). A visual comparison of the Figures 1(a) and 1(e) is indicative of the three extracted features to represent the three land cover classes.

4.3 Clustering

Three extracted features were used for computing the distance measure by the clustering algorithms k-means, expectation-maximization (EM) and fuzzy-c means (FCM) clustering algorithms. The k-means which is a partitioning based algorithm was applied using ENVI. The remaining two algorithms EM and FCM were implemented using Matlab. FCM computes the probability by which each pixel belongs to a class therefore each pixel belongs to each of the classes by certain probability. A single class label was assigned to each pixel based on maximum probability value. EM is a generative algorithm which tries to fit a probability model in

![Calibrated image](image1)

(a) Calibrated image

![Vegetation](image2)

(b) Vegetation

![Water](image3)

(c) Water

![Urban](image4)

(d) Urban

![Features FCC](image5)

(e) Features FCC

![Kmeans](image6)

(f) Kmeans

![FCM](image7)

(g) FCM

![EM](image8)

(h) EM

Figure 1. Figure showing various results of preprocessing, feature extraction and clustering on year 2000 images.
the data. It finds maximum value of log-likelihood in each iteration. Five to seven clusters were found in the data some of which were merged together to finally obtain the three clusters of interest. Visual analysis was carried out to identify the clusters which were initially yielded as distinct but were merged later by the clustering algorithms based on intra-class variability. The Figures 1(f), 1(g), and 1(h) shows the outcome of the results of K-means, FCM, and EM algorithms respectively.

4.4 Post-processing and cluster evaluation

Majority analysis of the clustered images was done with a 3×3 window size. In order to incorporate the spatial auto-correlation between the neighboring pixels majority analysis was implemented. After finding the clusters silhouette coefficient (Rousseeuw (1987)) combining inter-cluster distances between clusters and intra-cluster distances within clusters was computed to evaluate the cluster quality. The silhouette coefficient \( s_i \) for \( i^{th} \) object can be calculated using equation:

\[
s_i = \frac{b_i - a_i}{\max\{b_i, a_i\}}
\]

where \( a_i \) is the average distance of \( i^{th} \) object from all other points within the cluster and \( b_i \) is the minimum of the average distances of \( i^{th} \) object from all other objects in all other clusters.

5. Result Analysis

The results of applying the three clustering algorithms on the eight images between 1998 and 2011 have been presented in Table 1. The images are referred to by the year of acquisition. The three major columns correspond to the three classes of interest namely, vegetation, urban and water. Further the outcome of the three algorithms have been presented in the sub-columns. Each entry in the table represents the percentage of the total area under study occupied by the corresponding class. For example k-means clustering algorithm grouped 52.8% pixels of the area as vegetated area, 41.92% of the total pixels as urban area, and 5.26% pixels of the total study area as water area in the year 1998.

It is evident from Table 1 that in thirteen years from year 1998 to 2011, the areas covered by water and vegetation areas indicate a decrease while urban area shows an increase. Since each algorithm is a representative of a different category of clustering algorithm therefore the total area in each cluster varies for different algorithms. However all the three algorithms display almost the same pattern change in terms of increase or decrease in areas covered by the classes of interest in this study.

Table 2 presents the mean silhouette coefficient for each of the three algorithms for all eight years. Since all the values of silhouette coefficients are positive, indicating that clusters obtained were good. However, across the eight image data, the silhouette coefficient of k-means is greater than that of EM with least value corresponding to the FCM algorithm.

6. Conclusions

The three effective features extracted from all eight images enabled 50% reduction in terms of dimensionality of the data. The clustering algorithms using the three features yielded good clusters corresponding the three major land cover classes. All three algorithms clearly display that vegetation has decreased every year whereas urban area has increased. Based on the measure of silhouette coefficients, partition based clustering algorithm is more effective in comparison to probabilistic and fuzzy based clustering techniques and thus for change detection. Exact change from year 1998 to 2011 have not been reported as that requires

<table>
<thead>
<tr>
<th>Year</th>
<th>Vegetation</th>
<th>Urban</th>
<th>Water</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kmeans</td>
<td>EM</td>
<td>FCM</td>
</tr>
<tr>
<td>1998</td>
<td>52.8</td>
<td>76.6</td>
<td>45.8</td>
</tr>
<tr>
<td>1999</td>
<td>35.16</td>
<td>55.43</td>
<td>43.25</td>
</tr>
<tr>
<td>2000</td>
<td>37.318</td>
<td>56</td>
<td>42.89</td>
</tr>
<tr>
<td>2001</td>
<td>36.016</td>
<td>38.72</td>
<td>51.85</td>
</tr>
<tr>
<td>2002</td>
<td>34.283</td>
<td>64.61</td>
<td>63.25</td>
</tr>
<tr>
<td>2009</td>
<td>37.867</td>
<td>35.025</td>
<td>38.02</td>
</tr>
<tr>
<td>2010</td>
<td>35.717</td>
<td>52.87</td>
<td>40.39</td>
</tr>
<tr>
<td>2011</td>
<td>34.662</td>
<td>58.54</td>
<td>36.16</td>
</tr>
</tbody>
</table>

Table 2. Silhouette Coefficients.

<table>
<thead>
<tr>
<th>Year</th>
<th>K-means</th>
<th>EM</th>
<th>FCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>0.55</td>
<td>0.34</td>
<td>0.28</td>
</tr>
<tr>
<td>1999</td>
<td>0.44</td>
<td>0.35</td>
<td>0.36</td>
</tr>
<tr>
<td>2000</td>
<td>0.57</td>
<td>0.47</td>
<td>0.36</td>
</tr>
<tr>
<td>2001</td>
<td>0.61</td>
<td>0.58</td>
<td>0.48</td>
</tr>
<tr>
<td>2002</td>
<td>0.59</td>
<td>0.52</td>
<td>0.55</td>
</tr>
<tr>
<td>2009</td>
<td>0.59</td>
<td>0.51</td>
<td>0.50</td>
</tr>
<tr>
<td>2010</td>
<td>0.59</td>
<td>0.51</td>
<td>0.58</td>
</tr>
<tr>
<td>2011</td>
<td>0.59</td>
<td>0.36</td>
<td>0.59</td>
</tr>
</tbody>
</table>
ground truth information to validate. As satellite data is quite large so there is need of more scalable and parallel data mining algorithms which can take into account spatial features of the land cover. Since some features of satellite image are mixed with each other, they can be enhanced using signal processing techniques.

References


