Framework for Hyperspectral Image Segmentation using Unsupervised Algorithm

B. Raviteja¹, M. Surendra Prasad Babu², K. Venkata Rao² and J. Harikiran¹

¹Department of Information Technology, Gandhi Institute of Technology and Management, Beach Road, Gandhi Nagar, Rushikonda, Visakhapatnam – 530045, Andhra Pradesh, India; ravitejabhima@gmail.com
²Department of Computer Science and Systems Engineering, Andhra University, Waltair Junction, Visakhapatnam – 530003, Andhra Pradesh, India;

Abstract
Hyperspectral imaging system contains stack of images collected from the sensor with different wavelengths representing the same scene on the earth. This paper presents a framework for hyperspectral image segmentation using a clustering algorithm. The framework consists of four stages in segmenting a hyperspectral data set. In the first stage, filtering is done to remove noise in image bands. Second stage consists of dimensionality reduction algorithms, in which the bands that convey less information or redundant data will be removed. This deletion will decrease the storage requirement, computational load etc in processing the hyperspectral data. In the third stage, the informative bands which are selected in the second stage are merged into a single image using hierarchical fusion technique. The main goal of image fusion is to combine all the information from the selected image bands to form a single image. This single image is segmented using Fuzzy C-means (FCM) algorithm. The qualitative and quantitative analysis shows that this framework will segment the data set more accurately by combining all the features in the image bands.

Keywords: Dimensionality Reduction, Empirical Mode Decomposition, FCM, Hyperspectral Imaging

1. Introduction
Remote sensing is a technology to acquire information of object on earth with spatial and spectral resolution satellites¹. The satellites will generate huge remote sensing data, which increases the probability of developing new algorithms to analyze this data. One of such remote sensing data is Hyperspectral data containing many narrow bands with a continuous reflectance spectrum². Hyperspectral image segmentation is new research area aiming to produce high segmentation accuracy using efficient algorithms.

The hyperspectral image bands contain noise which is caused by the sensor problems or atmospheric transmission medium disturbance which affects result of image segmentation. To remove noise, a new filter is designed based on Bi-dimensional Empirical Mode Decomposition (BEMD) and Mean filter. The BEMD method³ decomposes the image band into several Intrinsic Mode Functions (IMF), in which the first function is the high frequency component, second function is the next high frequency component and the last function denotes the low frequency component called as residue. The mean filter is applied only to the few first high frequency components leaving the low frequency components, as the high frequency components contain noise. The image band is reconstructed by combining the filtered components and non-filtered components. The same procedure is used for filtering all the image bands.

The next step is dimensionality reduction which can be done in two steps, band selection and feature extraction. The feature extraction methods retrieve the features in the original image bands to create a low dimension feature space. This feature extraction methods change the physical characteristics of the hyperspectral data set. On the other hand, the band selection methods select the best
combination of image bands based on the information in each and every band. The band selection methods are more suitable for hyperspectral dimensionality reduction than feature extraction methods. This paper presents four band selection methods based on the information present in the original data set. The dimensionality reduction step decreases many requirements for processing the hyperspectral data set such as storage space, computational load, communication bandwidth etc., thus increasing the efficiency of segmentation algorithm.

After band selection, the next step is image fusion. The image fusion process will create an image combining all the features in the selected image bands. A hierarchical image fusion technique is used for merging the selected image bands. In this paper, a refinement is done in hierarchical image fusion technique. After getting a single image, the image is segmented using FCM clustering algorithm. The flow diagram of framework presented in this paper is displayed in Figure 1. This method produces more segmentation accuracy when compared with K-means and Moving K-means.

2. Noise Reduction using BEMD + Mean Filter

The filtering of hyperspectral image band using BEMD and mean filter is given below:

a) BEMD is applied to decompose each band in the dataset into IMFs.

b) The first few components are high frequency components which are suitable for de-noising using mean filter. The filtered components are denoted using DIMFs.

c) The filtered image band RI is reconstructed according to the given equation:

\[ RI = \sum_{i=1}^{d} DIMF_i + \sum_{i=d}^{k} IMF_i \]  

Where \( d \) is the number of high frequency bands selected for filtering.

3. Band Selection Methods

In this paper four band selection methods such as Euclidean Distance, Spectral Angle Mapper, Spectral Correlation Mapper and Band Correlation are presented to select the informative bands. The band selection methods between two vectors \( X \) and \( Y \) with data set having \( N_b \) bands is given as follows:

\[ ED(X, Y) = \sqrt{\sum_{k=1}^{N_b} (X_k - Y_k)^2} \]  

\[ SAM(X, Y) = \arccos \left( \frac{X^T Y}{\|X\| \|Y\|} \right) \]  

\[ SCM(X, Y) = \frac{\sum_{k=1}^{N_b} (X_k - \mu_X)(Y_k - \mu_Y)}{(N_b - 1)\sigma_X \sigma_Y} \]  

\[ BC(i, j) = \frac{\sum_{k=1}^{N_b} (x_{ik} - \mu_i)(x_{jk} - \mu_j)}{\sqrt{\sum_{k=1}^{N_b} (x_{ik} - \mu_i)^2} \cdot \sqrt{\sum_{k=1}^{N_b} (x_{jk} - \mu_j)^2}} \]  

4. Hierarchical Image Fusion Technique

In hierarchical fusion technique, the entire hyperspectral data set is partitioned into \( P \) subsets, where \( P \) is given
by $P = \left[ \frac{K}{M} \right]$, $K$ number of bands in data set and $M$ bands in each subset. First image fusion is carried out independently on these $P$ subsets, to form $P$ fused images. These $P$ images are used as input for second stage of fusion, again by dividing into subsets. The same process continues in bottom-up manner until a single image is generated representing the final result of fusion. The flow diagram of hierarchical image fusion is shown in Figure 2.

The fused image $F$ at any stage is calculated as follows:

$$ F(x, y) = \sum_{k=1}^{M} w_k(x, y)I_k(x, y) $$

and

$$ \sum_{k=1}^{M} w_i(x, y) = 1_i \forall (x, y) $$

where $w$ denotes normalized weight and $F$ denotes the fused image.

This paper presents a refinement to hierarchical fusion method. Instead of grouping the images into subsets of fixed size, in this paper the grouping is done based on the similarity metric $M$ between the images. The metric used is Average Pixel Intensity (API). The images subsets are created based on API, i.e., the images in a subset have same API. The same procedure is continued at each and every stage, until a single image is formed. The API of an image is calculated according to the given equation:

$$ API_i = \frac{1}{M*N} \sum_{x=1}^{M} \sum_{y=1}^{N} I(x, y) $$

and

$$ M = |I_{ref} - I| $$

where $M$ is the similarity metric and $I_{ref}$ is the first image of a group.

5. Fuzzy C-Means (Unsupervised) Algorithm

The FCM algorithm for segmentation of hyperspectral image is described below:

1. Take randomly $K$ initial clusters from the $m*n$ image pixels.
2. Initialize membership matrix $u_{ij}$ with value in range 0 to 1 and value of $m=2$.
   Assign each pixel to the cluster $C_j \{j=1,2,...,K\}$ if it satisfies the following condition [$D(\cdot,\cdot)$ is the Euclidean distance measure between two values].

$$ u_{iq}^m D(I_i, C_j) < u_{iq}^m D(I_i, C_q), q = 1, 2,..., K $$

$$ j \neq q $$

The new membership and cluster centroid values as calculated as

$$ u_{ik} = \frac{1}{\sum_{j=1}^{K} \left( \frac{D(C_i, I_k)}{D(C_j, I_k)} \right)^{1/m-1}} $$

and

$$ C_j = \frac{\sum_{j=1}^{n} u_{ij}^m I_j}{\sum_{j=1}^{n} u_{ij}^m} $$

3. Continue 2-3 until each pixel is assigned to the maximum membership cluster.

6. Experimental Results

The proposed framework is executed on Pavia University hyperspectral data collected from containing 103 spectral bands. The Pavia University data contains nine clusters. K-means, Moving K-means and FCM algorithms are used for segmentation of hyperspectral image. The qualitative analysis of presented framework on Pavia University hyperspectral data set is shown in Figure 3.
Framework for Hyperspectral Image Segmentation using Unsupervised Algorithm

Band selection is done using Band Correlation method. Quantitative analysis is carried out using Mean Square Error function which is defined as:

$$MSE = \frac{1}{N} \sum_{j=1}^{k} \sum_{i \in c_j} \|y_i - c_j\|^2$$  \hspace{1cm} (10)

Table 1 shows the quantitative evaluation of clustering algorithms used for hyperspectral image segmentation. The results confirm that Fuzzy C-means algorithm produces the lowest MSE value for segmenting the hyperspectral image, indicating higher segmentation accuracy.

7. Conclusions

In this paper a framework for hyperspectral image segmentation is presented. The framework is carried out in four stages. First stage contains noise reduction algorithm in each image band, second stage contains dimensionality reduction using band selection methods to select informative bands leaving the bands that convey less descriptive information, third stage contains hierarchical image fusion to generate a single informative band and in the fourth stage, segmentation using FCM algorithm. Existing methods for hyperspectral data sets is

<table>
<thead>
<tr>
<th>Band 100</th>
<th>IMF1</th>
<th>IMF2</th>
<th>IMF3</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Filtered Image" /></td>
<td><img src="image2" alt="Fused Image using proposed method" /></td>
<td><img src="image3" alt="Segmented using FCM" /></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. MSE values of segmented images using three clustering algorithms

<table>
<thead>
<tr>
<th>Method</th>
<th>MSE Values (Segmentation before noise removal using EMD)</th>
<th>MSE Values (Segmentation after noise removal using EMD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-means [KM]</td>
<td>346.678</td>
<td>322.781</td>
</tr>
<tr>
<td>Moving K-means [MKM]</td>
<td>302.986</td>
<td>286.392</td>
</tr>
<tr>
<td>FCM</td>
<td>222.346</td>
<td>198.327</td>
</tr>
</tbody>
</table>

Figure 3. Hyperspectral image segmentation.
done by selecting limited number of bands normally less than seven. The accuracy of any segmentation algorithm decreases if the number of spectral bands increases. The framework presented in this paper provides a methodology for segmenting the hyperspectral data set by incorporating all the information existing in the original bands rather than selecting some spectral bands. The framework segments the hyperspectral data set more accurately than other segmentation methods such as K-means and Moving K-means.

8. References

9. Sweet JN. The Spectral similarity scale and its application to the classification of hyperspectral remote sensing data. Proceedings of Institute of Electrical and Electronics Engineers (IEEE) International Conference on Image Analysis and Processing (ICIAP); 2009.
10. Gharati E, Nasri M. A New Band Selection Method for Hyperspectral Images based on Constrained Optimization", proceedings of Institute of Electrical and Electronics Engineers (IEEE) Conference on Information and Knowledge Technology (IKT); 2015.