Multimodal Image Classification through Band and K-means clustering

Archana M R[1] , Keerthana M M[2]

[1] *Assistant Professor, Dept. of CSE., ATME College of Engineering, Mysuru , Karnataka , India*
[2]*Assistant Professor, Dept. of CSE., ATME College of Engineering, Mysuru , Karnataka , India*

---***---

Abstract - *Multimodal images composed of a very large number spectral channel which ranges from visible to infrared spectrum. Remote sensing involves measurement of energy in various parts of the electromagnetic spectrum. Multimodal images has crucial role in remote sensing as spectral bands are in rich information which is helpful to classify the spectrally similar objects. Multimodal image classification with limited number of labelled pixels is a challenging task. In this paper, we propose a bilayer graph-based learning framework to address this problem. For graph-based classification, how to establish the neighbouring relationship among the pixels from the high dimensional features is the key toward a successful classification. The first-layer constructs a simple graph, where each vertex denotes one pixel and the edge weight encodes the similarity between two pixels. Unsupervised learning is then conducted to estimate the grouping relations among different pixels. These relations are subsequently fed into the second layer to form a hypergraph structure, on top of which, semisupervised transductive learning is conducted to obtain the final classification results.*

Key Words*: Multimodal images, bilayer graph, Unsupervised, semisupervised, classification.*

1. INTRODUCTION

About Spectral Imaging Spectral imaging is a branch of spectroscopy and photography in which a complete spectrum or some spectral information (such as the Doppler shift or Zeeman splitting of a spectral line) is collected at every location in an image plane. Various distinctions among techniques are applied, based on criteria including spectral range, spectral resolution, number of bands, width and contiguousness of bands, and application. The terms include multispectral imaging, hyperspectral imaging, full spectral imaging, imaging spectroscopy or chemical imaging. These terms are seldom applied to the use of only four or five bands that are all within the visible light range. Spectral images are often represented as an image cube, a type of data cube. . Applications include astronomy, solar physics, planetology, and Earth remote sensing [14][15].

The increasing of satellites and airborne devices for Earth observation has resulted in massive amount of Multimodal image data covering the earth surface. As a result, hyper-spectral image processing and analysis has become an active research topic in both the image processing , and the remote sensing societies. There are two prominent challenges which confront Multimodal image classification. The first one is the difficulty in evaluating the similarity of two pixels induced by the high dimensionality. A Multimodal image contains hundreds of spectral bands and correspondingly each pixel is described by hundreds of observed values from these spectral bands. This high dimensional data leads to the difficulties on Multimodal image analysis due to the curse of dimensionality.

To mitigate the curse of dimensionality, some existing works and focused on dimensionality reduction of the high dimensional features extracted from Multimodal images, including classical methods such as Independent Component Analysis , and Principal Component Analysis and recent works such as Kernel Nonparametric Weighted Feature Extraction and Tensor Discriminative Locality Alignment. To mitigate the issue of learning with small amount of training samples in Multimodal image classification, semisupervised learning has shown its superiority. For instance, the neighborhood relationships among all pixels are modeled by a graph structure in and a semi-supervised learning procedure is conducted for Multimodal image classification.

To overcome the challenges of both the complex relationship and the limited labeled samples in Multimodal images, motivated by the superiority of high-order relevance exploration of the hypergraph structure, we propose a Multimodal image classification framework by using a band clustering and k-means clustering along with bilayer graph based learning in this paper.

This bilayer graph is composed of a layer of simple graph as well as a layer of hyper graph, which effectively exploits the underlying structure of the data. In the first-layer, a simple graph is constructed, where each vertex in the graph denotes one pixel and the similarity among vertices is determined by the feature based pairwise pixel distances. Learning is conducted on this layer to estimate the connectivity relationship among pixels. In the second-layer, a hypergraph structure is constructed, where each vertex denotes one pixel and the hyperedges are generated by using the neighborhood relationship produced from the first-layer. Semi-supervised learning is conducted on the hypergraph structure to estimate the pixel labels to achieve Multimodal image classification.

2. RELATED WORK

All The literature review of Multimodal image classification fall under three categories supervised hyperspectal image classification,unsupervised hyperspectal image classification, semisupervised hyperspectal image classification to handle the various issues which are faced while classifying Multimodal images such as large number of spectral channels, acquisition of labeled data etc. The task of acquisition of labeled data is time consuming and costly.

Supervised Classification Methods to Multimodal Image Classification :

Bands are selected using mutual information (MI). Mutual information term calculate the statistical dependence between two random variable form which it easy to understand relevance of that particular band to classification. Those most relevant bands are selected for further analysis of image which in turns handles the issue of high dimensionality [1]. Supervised Kernel nonparametric weighted feature extraction (KNWFE) method is proposed in [2] to extract the relevant features. This method combines kernel methods and nonparametric weighted feature extraction method to possess both linear and nonlinear transformation. In [4] supervised method based on a stochastic minimum spanning forest (MSF) approach to classify Multimodal data is proposed. In this method a pixel wise classification is first performed on hyperspectal image .From this classification map, marker maps are created with random selection of pixels and labeling them as markers for the purpose building of MSFs. MSF is built from each of the marker maps and final classification map generated with maximum vote decision rule.

Unsupervised Classification Methods to Multimodal Image Classification :

In [5] unsupervised method based on fuzzy approach which uses linear 1-D discrete wavelet transform (DWT) for reducing dimensionality of Multimodal data. In this approach segmentation of Multimodal images by applying fuzzy c-means (FCM) clustering as well as its extended version Gustafson Kessel clustering (GKC). Image categorization is done with the help of hypergraph partition [6]. Hypergraph has advantages over simple graph. Complex relationship between unlabeled is represented with help of hypergraph. In this procedure unsupervised method is conducted to select the Region of Interests (ROIs) of the unlabeled images. Based on the appearance and shape descriptors extracted from the ROIs to measure two types of similarities between images from which two kinds of hyper edges are formed and compute their corresponding weights based on these two kinds of similarities, respectively. As discussed above all the unsupervised methods are insensitive to the number of labeled data since these methods work on the whole image, but the relationship between clusters and classes is not guaranteed. The use of semi-supervised classifiers e in these situations can help to improve the classification accuracy.

Semisupervised Classification Methods to Multimodal Image Classification :

In semi-supervised methods the algorithm is provided with some available labeled data in addition with unlabeled data. In literature three different classes of semi-supervised learning algorithms are introduced.

1. Generative models-In these types of algorithm conditional density p(xy) (e.g. expecta-tion maximization (EM) algorithms with finite mixture models are calculated.
2. Low density separation These algorithms, maximize the hyperplane between labeled and unlabeled samples simultaneously (e.g. Transductive SVM [7]).
3. Graph-based methods-Each sample spreads its label information to its neighbors until a global stable state is achieved on the whole data set.

Semisupervised version of neural network introduced to overcome limitations of TSVM such as falling under local minima by adding a regularizer to the loss function which issued for training neural networks [8].

3. SYSTEM ARCHITECUTURE

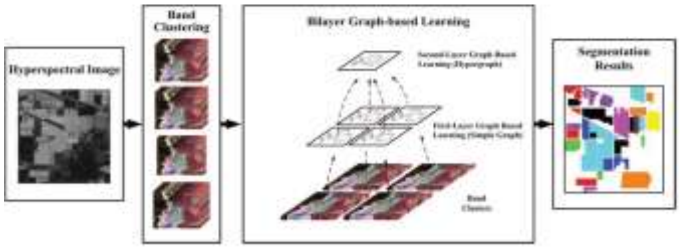

Figure 1 System Architecture

Figure 1 shows the schematic illustration of the proposed method. Due to the high dimensionality of Multimodal images, we need to find effective relevance estimation method.

In our proposed framework, we first conduct an unsupervised learning to estimate the relevance between each two pixels based on the original spectral data. This procedure can also be regarded as a feature transformation process, in which a subspace of the full feature space between each two pixels is used to estimate the relevance among the pixels. In the subsequent semi-supervised learning procedure, all pixels are modeled in a multimodal, upon which the learning is conducted to estimate the pixel labels.

The proposed Multimodal image classification framework by using a bilayer graph based learning in our project. This bilayer graph is composed of a layer of simple graph as well as a layer of multimodal, which is effectively exploits the underlying structure of the data. In the first-layer, a simple graph is constructed, where each vertex in the graph denotes one pixel and the similarity among vertices is determined by the feature based pairwise pixel distances. Learning is conducted on this layer to estimate the connectivity relationship among pixels. In the second-layer, a multimodal structure is constructed, where each vertex denotes one pixel and the hyperedges are generated by using the neighborhood relationship produced from the first-layer.

To construct hyperedge in feature space the pixels which are close to each other in feature space are connected to form hyperedge .This closeness is measured in distance metric. The pixels with small distance are considered as close to each other. The pixels which are close to each other has same label. At first simple graph constructed and unsupervised learning conducted over simple graph to identify grouping relation while in second step multimodal constructed from previous step and semisupervised learning conducted to achieve desired classification result.

4.Pseudo code and Algorithm

a. Band Clustering and K-means clustering

Pseudocode For Band Clustering

```
[file,path] = uigetfile('*.jpeg;*.jpg;*.png;*bmp;*.gif','Pick an image');
if isequal(file,0) || isequal(path,0)
  warndlg('user press cancel');
else
  image=double(imread(file));
end
 Qlevels=2.^(8:-1:0);
[maps,images]=srm(image,Qlevels);
imseg = images;
mapList = maps;
precision=numel(mapList);
Iedge=zeros([size(imseg{1},1),size(imseg{1},2)]);
 quick_I1 = cell(precision,1);
quick_I2 = cell(precision,1);
%% FREQUENCY BAND %%
 k =1;
  quick_I2{k} = imseg{k} ;
  figure;
  imagesc(uint8(quick_I2{k}));axis off;
k =2;
  quick_I2{k} = imseg{k} ;
  figure;
  imagesc(uint8(quick_I2{k}));axis off;
 k =3;
  quick_I2{k} = imseg{k} ;
  figure;
  imagesc(uint8(quick_I2{k}));axis off
k =4;
  quick_I2{k} = imseg{k} ;
  figure;
  imagesc(uint8(quick_I2{k}));axis off;
k =5;
    quick_I2{k} = imseg{k} ;
  figure;
  imagesc(uint8(quick_I2{k}));axis off;
 k =6;
  quick_I2{k} = imseg{k} ;
  figure;
  imagesc(uint8(quick_I2{k}));axis off;
 k =7;
  quick_I2{k} = imseg{k} ;
  figure;
  imagesc(uint8(quick_I2{k}));axis off;
 k =8;
  quick_I2{k} = imseg{k} ;
  figure;
  imagesc(uint8(quick_I2{k}));axis off;
 k =9;
  quick_I2{k} = imseg{k} ;
  figure;
  imagesc(uint8(quick_I2{k}));axis off;
```

b. Pseudo code for Multimodal Image Classification through Bilayer Graph Based Learning

```
%% segmented image %%
  map=reshape(mapList{k},size(Iedge));
  quick_I1{k} = srm_randimseg(map) ;
  figure;   imagesc(quick_I1{k});axis off;
```

5.Simulation Results

Figure 4.1 Original Multimodal Image

Here we use a original Multimodal image which is sampled from hundreds or thousands of contiguous and narrow spectral bands by Multimodal sensors. Using Multimodal makes it easier to unmix pixels, thus improving confidence in classification results.

Figure 4.2 Input Image

The original Multimodal image is taken as the input. If the file has no image then the warning dialogue is displayed else the input image undergoes band clustering.

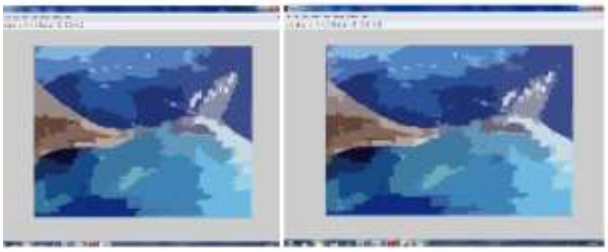

Figure 4.3 Band Image

The first two band clustering images are shown in the above figure. Each band consist of different quantization value. First band has 256 quantization value and the second band consist of 128 quantization value. Band clustering occurs based on these quantization values.

Figure 4.4 Band Image

The third and fourth band clustering images are shown in the above figure. Each band consist of different quantization value. Third band consist of 64 quantization value and the fourth band consist of 32 quantization value. Band clustering occurs based on these quantization values.

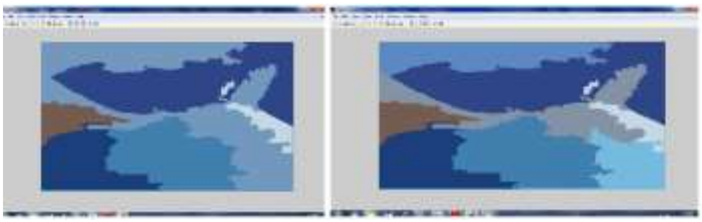

Figure 4.5 Band Image

The fifth and sixth band clustering images are shown in the above figure. Each band consist of different quantization value .Fifth band consist of 16 quantization value and the sixth band consist of 8 quantization value. Band clustering occurs based on these quantization values.

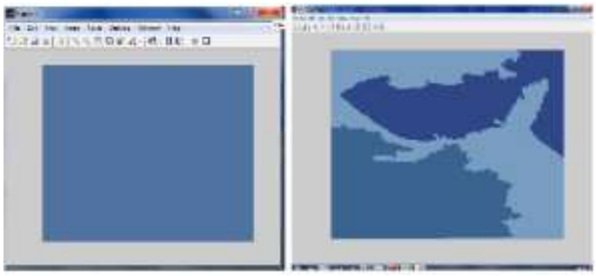

Figure 4.6 Band Image

The seventh and eighth band clustering images are shown in the above figure. Each band consist of different quantization value. Seventh band consist of 4 quantization value and the eighth band consist of 2 quantization value. Band clustering occurs based on these quantization values.

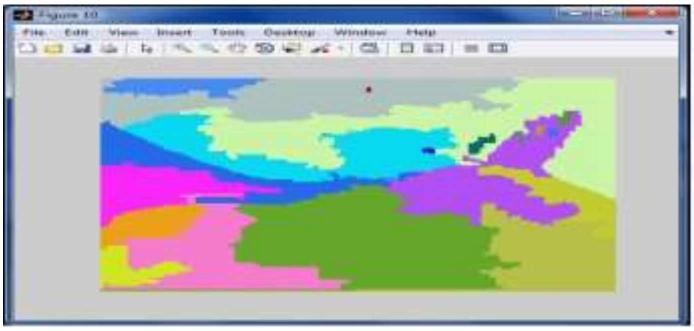

Figure 4.7 Output Segmented Image

The final classified segmented image is shown in the above figure. For the band clustering images bi-layer graph based learning method is applied to get the classified output image.

Figure 4.8 Final Classification of Multimodal Image

The final classification of Multimodal image is shown in the above figure using GUI. It consists of input image, band clustering image and output image shown in same window using GUI.

6. Conclusion and future enhancement

In this paper, we have proposed a bilayer graph based learning framework for Multimodal image classification. In the first-layer, an unsupervised learning is conducted to estimate the pixel relevance for feature transformation. In the second-layer, a semi-supervised learning is conducted to explore the relationship among all pixels. This paper gives some well known techniques based on how training samples are used to classify the Multimodal data. Experimental results on the datasets are provided to validate the effectiveness of the proposed method. As shown in the results, the proposed bilayer framework is able to achieve better results in comparison to the state-of-the-art methods. We have also evaluated the computational cost of the proposed method. As all pixels in the Multimodal image have been involved in the learning process, the increasing image size will lead to high computational cost in terms of both memory consumption and CPU usage. To scale up our approach for large datasets, we will further investigate region based classification method and hierarchical learning schemes in our future work. The larger size of the testing dataset leads to higher computational cost. Therefore, how to deal with such high computational cost is one important issue. There are two possible solutions for this challenge. On one hand, the large dataset can be first split into small regions, and then multimodal image classification is conducted on each of these regions. In this direction, how to conduct the image splitting to minimize the degradation of classification performance is the key issue. On the other hand, a hierarchical graph learning scheme would be effective on reducing the computational cost.

REFERENCE

[1] G. Moser, S. B. Serpico, and J. A. Benediktsson, Land-cover mapping by Markov modeling of spatial-contextual information in very high- resolution remote sensing images, Proc. IEEE, vol. 101, no. 3, pp. 631651, Mar. 2013..

[2]] S. Schweizer and J. Moura, Efficient detection in Multimodal imagery, IEEE Trans. Image Process., vol. 10, no. 4, pp. 584597, Apr. 2001.

[3] C. Li, T. Sun, K. Kelly, and Y. Zhang, A compressive sensing and unmixing scheme for Multimodal data processing, IEEE Trans. Image Process., vol. 21, no. 3, pp. 12001210, Mar. 2012..

[4] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, Advances in spectral-spatial classification of Multimodal images, Proc. IEEE, vol. 101, no. 3, pp. 652675, Mar. 2013.

[5] J. Li, J. Bioucas-Dias, and A. Plaza, Semisupervised Multimodal image segmentation using multinomial logistic regression with active learning, IEEE Trans. Geosci. Remote Sens., vol. 48, no. 11, pp. 40854098, Nov. 2010..

[6] G. Bilgin, S. Erturk, and T. Yildirim, Unsupervised classification of Multimodal-image data using fuzzy approaches that spatially exploit membership relations, IEEE Geosci. Remote Sens. Lett., vol. 5, no. 4, pp. 673677, Oct. 2008.

[7] O. Eches, N. Dobigeon, C. Mailhes, and J. Tourneret, Bayesian estimation of linear mixtures using the normal compositional model application to Multimodal imagery, IEEE Trans. Image Process., vol. 19, no. 6, pp. 14031413, Jun. 2010.

[8] L. Zhang, L. Zhang, D. Tao, and X. Huang, A multifeature tensor for remote-sensing target recognition, IEEE Geosci. Remote Sens. Lett., vol. 8, no. 2, pp. 374378, Mar. 2011..

[9] K. Bernard, Y. Tarabalka, J. Angulo, J. Chanussot, and J. A. Benediktsson, Spectral-spatial classification of Multimodal data based on a stochastic minimum spanning forest approach, IEEE Trans. Image Process., vol. 21, no. 4, pp. 20082021, Apr. 2012..

[10] H. Du, H. Qi, X. Wang, R. Ramanath, and W. E. Snyder, Band selection using independent component analysis for Multimodal image processing, in Proc. 32nd Appl. Imagery Pattern Recognition Workshop, Oct. 2003, pp. 9398.

[11] B.-C. Kuo, C.-H. Li, and J.-M. Yang, Kernel nonparametric weighted feature extraction for Multimodal image classification, IEEE Trans. Geosci. Remote Sens., vol. 47, no. 4, pp. 11391155, Apr. 2009.

[12] A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, Multimodal image classification with independent component discriminant analysis, IEEE Trans. Geosci. Remote Sens., vol. 49, no. 12, pp. 48654876, Dec. 2011.

[13] A. Hyvrinen and E. Oja, Independent component analysis : Algorithms and applications, Neural Netw., vol. 13, nos. 45, pp. 411430, 2000.

[14]. R. Ji, Y. Gao, R. Hong, Q. Liu, D. Tao, and X. Li, "Spectralspatial constraint hyperspectral image classification," IEEE Trans. Geosci. Remote Sens., vol. 52, no. 3, pp. 1811–1824, Mar. 2014.

[15] G. Moser, S. B. Serpico, and J. A. Benediktsson, "Land-cover mapping by Markov modeling of spatial-contextual information in veryhigh-resolution remote sensing images," Proc. IEEE, vol. 101, no. 3, pp. 631–651, Mar. 2013.

BIOGRAPHIES

Mrs. Archana M R is an assistant professor in Dept. of computer science and engineering in ATME College of Engineering Mysuru, Karnataka, India. She received her master degree in computer network and engineering from NIE, Mysuru . She has 11 year of teaching experience and her field of interest is image processing.

Ms. Keerthana M M is an assistant professor in Dept. of computer science and engineering in ATME College of Engineering Mysuru, Karnataka, India. She received her master degree in computer science and engineering from VKIT, Bangalore affiliated to VTU University, Belagavi. she has 3 years teaching experience and her field of interest is cloud computing .