The CSI Journal on
Computer Science and Engineering
Vol. 11, No. 2 & 4 (b), 2013
Pages 43-54
Regular Paper

Spatial-Spectral Hyper Spectral Classification Based on Statistical Dependence between Adjacent Pixels

Mostafa Borhani Hassan Ghassemian

Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran

Abstract

This paper contributes some spectral-spatial classification methodologies and techniques based on spatial homogeneous regionsfor hyper spectral remotely sensed images. These techniques mainly focus on adequate object segmentation and simultaneous combination of spectral and spatial information. Different segmentation methods such as hyper spectral Robust Color Morphological Gradient (HRCMG), Adequate Expectation Maximization (AEM) and hyper spectral Recursive Hierarchical Image Segmentation (HRHSEG) were introduced and applied in the empirical implementations. This paper also contributes integration of the local weighted Markov Random Fields (MRF) on SVM framework for hyper spectral spectral-spatial classification. Using marginal weighting function in the MRF energy function, which preserves the edge of regions, is a new approach. In this paper, merits and issues of different proposed techniques are examined and compared as well as their classification maps with some known spectral-spatial methods for four real hyper spectral images. Experimental results illustrate our proposed approaches including higher accuracies when compared with elder schemes.

Keywords: Spectral-Spatial Hyper Spectral Classification, MRF, Local Weighted Marginal, SVM, Adaptive Segmentation.

1. Introduction

To take full gain of the amusing information in hyper spectral imaging, new algorithmic developments are required. Currently, in many application fields, hyper spectral remote sensed image classification is a significant mission. Traditional hyper spectral data classification methods were designed to allocate each pixel to one of the classes based on its spectrum. There are spectral based methods, or pixel wise classifiers. Pixel wise classifiers are just applied on spectral features regardless of how classify the bordering pixels. In remote sensing images, adjacent pixels are relevant and interdependent [1]. There are two reasons for neighbor pixels independency; first because the imaging devices are getting significant energy from adjacent pixels and second reason is the structure similarity. These structures are usually greater than a pixel in size. This local information can be used to properly interpret the landscape [2]. So, for improving the classification accuracy, some novel spectral-spatial methods must be developed to assign the correct class of each pixel.

To improve spectral-spatial classification accuracies, some investigation to incorporating the spatial and spectral information are recommended by this paper. The structures in spatial dimensions, sizes and shapes of experimental world are recognized as objects. Therefore, adding information about spatial structures into classifier shall increase the accuracy of classification.

The question is how to describe these structures automatically and effectively [3]. This information is not directly available from camera recordings as the spectral information. The second question is how to combine spectral and spatial information [4].

Possibly the most intuitive method to study the spatial information is to accept that for a given pixel; its bordering samples with a high probability belong to the same class, and therefore classify a central pixel taking into account its neighboring samples with a good approximation[5]. In the previous researches, MRF and contextual features (mean and standard deviation) have been used to combine the information from the nearby samples in classification. The main benefit of this method is its simplicity. However, the bordering fixed area does not always include exact information about spatial structures, especially at the border of areas.

Land grebe and his research group is the inventor of spatial context in multi-band image classification. They presented the well-known Extraction and Classification of Homogeneous Objects (ECHO) [6]. M. Fauve et al. explored the use of Extended *Morphological Profile* (EMP) [7] as a substitute way of execution joint classification. The idea behind a morphological profile contains progressively simplifying the image. In this paper, we used ECHO [6], EMP [7] as standard methods for spectral-spatial classification.

We proposed a new method based on static and adaptive weighted segmentation on the probabilistic outputs for support vector machines [8][9] for hyper spectral image classification. In the first step, the probabilistic SVM classification is applied. The second phase is to use spatial data in order to improve the localized classification results gotten in the first phase. This is succeeded by different approaches such as MRFs, Robust Color Morphological Gradient (RCMG) [1], Expectation Maximization (EM) [18] and Recursive Hierarchical Segmentation (RHSEG) [22]. The important differences with the formerly proposed approaches are in the description and incorporation of weighting function in adaptive local energy function to keep margins in the segmentation. The operational scheme of proposed classification method is shown in figure 1. There is a B-band hyper spectral image as input and which can be seen as a set of pixel vectors of n elements $\mathrm{X} = \{\mathrm{X_j} \in \mathrm{R^B}, \mathrm{j} = 1,2,\dots,\mathrm{n}\}$. The classification involves allocating each pixel to one of the K classes $\{\mathrm{w_1}, \mathrm{w_2}, \dots, \mathrm{w_k}\}$.

In this paper, an advanced spectral-spatial classification method is proposed using adaptive adequate weighted segmentation to explore and analyze the dependencies between the pixels. Some advanced approaches for classification hyper spectral images would be developed by using adaptive spatial vicinity based on probabilistic outputs for support vector machines and MRF which are two powerful tools in hyper spectral data classification and spatial context analysis. Segmentation results can be realized by adaptive neighborhood area. As novelty, after surveying basic segmentation approaches (unsupervised), these segmentation approaches can be generalized to hyper spectral images remote sensing. We develop a new method to combine spatial and spectral information of hyper spectral classification. In this advanced method, we use segmentation maps as spatial information and classification map as spectral information.

The main objective of this paper is to additional developing approaches for classification hyper spectral data using for both spectral and spatial information in adaptive local segment. We have planned and established some general strategies for hyper spectral data classification. These are the contributions of this paper.

Our proposed method use adequate spatial neighborhood area derived from segmentation outcomes. Different segmentation methods (watershed, partial clustering, and hierarchical segmentation) are explored and extended to the case of hyper spectral images. Then, merging the extracted spatial regions with spectral information in a classifier are proposed and established. Our proposed methods are compared with some modern spectral-spatial classification approaches.

The proposed method uses SVM classification with some novel segmentation methods such as hyper spectral RCMG (HRCMG), Adequate EM (AEM) and hyper spectral RHSEG (HRHSEG), which we engaged for integration spatial and spectral labels. Figure 1 shows the proposed scheme of spatial-spectral classification, the Median Voting (MV) has been used for fusion spectral and spatial information. We combine spectral and spatial information in classification. The proposed classification scheme based on probabilistic spatial energy regulation is composed of the following steps:

- Segmentation, what is now presented in section 3.
- Perform probabilistic spatial energy regulation. We suggest it in section 2.
- Then, for every area in the segmentation map, assign all pixels to the median class within this region.

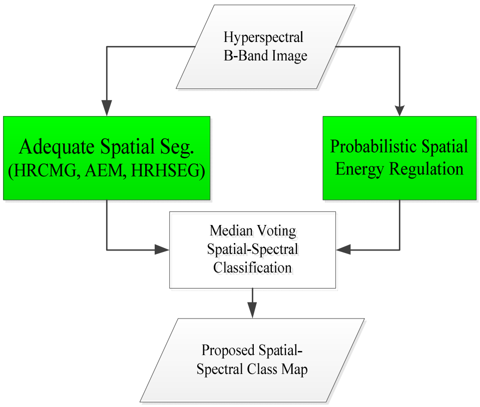


Figure 1. The scheme of proposed approaches

The rest of this paper is organized as follows: the details of probabilistic weighted spatial energy regulation methods using suitable vicinity area are derived from unsupervised segmentation in section 2. Section 3 discusses classification using suitable vicinity area derived from adequate spatial segmentation outcomes. Some new segmentation approaches are also proposed in this section. Clearly RHSEG [22] is extended to Hyper spectral issue (HRHSEG) and then, the EM [18] is mathematically generalized to Adequate version (AEM). We propose the novel spatial-spectral classification technique based on median voting of segmentation maps (as spatial information) and pixel wise classification maps (as spectral information). Then, the proposed approach is examined with realistic hyper spectral Indian Pines [10], Hekla, university and center of Pavia [11] datasets and compared with some earlier methods. Experimental results are presented in section 4. Finally, discussion about experiments and conclusions are outlined in section 5.

2. Probabilistic Spatial Energy Regulation

The problem is what could be alternative approach to describe the structure element for each pixel. The proposed answer is “to comprehensive dividing of the image into local similar areas”. In a perfect case, each local adaptive homogeneous area would match to one spatial object. Each area in the segmentation map describes a correct proximity area for all the pixels within this local area. Three methods for image segmentation were recognized. The First one is based on the margins, the second method is based on the region, and the third one is based on threshold with marked or grouped. First two approaches are in spatial domain, while the third one is in the spectral domain. This paper would classify by using probabistic spatial energy as define in this section and figure 2.

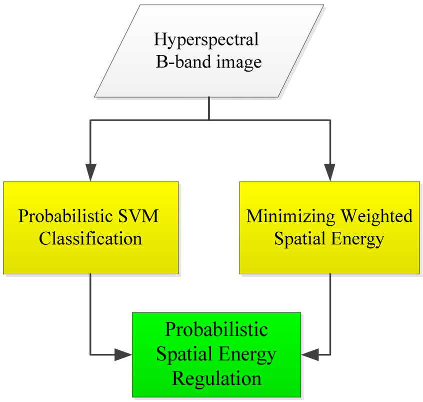


Figure 2. The proposed scheme based on probabilistic SVM, local energy function and spatial suitable segmentation

2.1. Probabilistic Outputs of Classification

The first step of the planned procedure includes performing pixel wise Platt's probabilistic SVM [12][9] in hyper spectral images. Other classifiers can also be investigated. However, SVM is well-matched for classification of hyper spectral data. The results are: 1. Classification map; where each pixel has its individual class label and 2. Approximation of the probability of each pixel belonging to each class.

The standard SVM does not offer estimation of the probability of classes. We endorse using the method that has been realized in the LIBSVM [13] to estimate the probability of each pixel belonging to the target class:

$$p = \{p_k = p(y = k|X), k = 1, \dots, k\} \tag{1}$$

In this paper, the 'one against one' multi-class SVM classification strategy with Gaussian RBF kernel is executed. Pair wise class probabilities $r_{ij} \approx p(y = i|y = i \text{ or } j, x)$ are appliedby using enhanced operations such as Equ. (2).

$$r_{ij} \approx \frac{1}{1 + e^{A\hat{f}+B}} \tag{2}$$

where A and B are gotten by minimizing the negative logarithm of the likelihood function using training data and the value of $\hat{f}$ is which is estimated. In addition, the probabilities of Equ. (1) is calculated by solving followed optimization problem:

$$\min_p \sum_{i=1}^{K} \sum_{j:j\neq i} (r_{ji}p_i - r_{ij}p_j)^2 \; ; \; \sum_{i=1}^{K} p_i = 1, p_i \geq 0 \forall_i \tag{3}$$

This problem has a unique solution that can be obtained by solving a simple linear system.

2.2. Minimizing Weighted Spatial Energy

The regulation step of SVM classification map is founded on the hypothesis of independence among pixels of class, meaning that it is probable that the pixel belongs to the class ω_i, then the nearby pixels are belong to the same class. In this paper, we implement the generalized Metropolis-Hasting stochastic relaxation with annealing algorithm [14] to calculate the MAP (Maximum a Posteriori) estimated from the pixel wise classification maps. This Method is based on Bayesian method and planned to minimize the overall image energy by repeating minimization energy procedure in the spatial domain. We propose that the local energy related with a given pixel $x_i$measured as follow:

$$U(x_i) = U_{Spectral}(x_i) + U_{Spatial}(x_i) \tag{4}$$

where $U_{Spectral}(x_i)$, and $U_{Spatial}(x_i)$ are the spectral and spatial energy function respectively, those are considered on $N_i$local vicinity. (N_i Is the set of neighboring pixels ofx_i). Spectral energy function is defined as:

$$U_{Spectral}(x_i) = -\ln\big(P(x_i|w_i)\big) \tag{5}$$

where $W = \{w_j, j = 1,2, \dots, n\}$is the set of class labels for the image and $P(x_i|w_i)$ is valued by using binary probability estimates from multi-class "one against one" SVM results. Two different equations are planned for the spatial energy function. We start with the standard spatial energy formula that is considered as:

$$U_{Spatial}(x_i) = \sum_{x_j \epsilon N_i} \beta\left(1 - \delta(w_i, w_j)\right) \tag{6}$$

where $\delta(.,.)$ is the Kronecker delta and β parameter controls the importance of spatial energy against the spectral energy in Equ. (4). $U_{Spatial}(x_i)$ is proportional to the number of pixels x_i thatis dedicated to one of the other classes exceptw_i.

This spatial energy equation is good, especially for images with large spatial structure. However, if the object is small and displayed as a single pixel in the image, the model prefers to assign this pixel to the class of objects around it. In order to reduce the earlier spatial energy function bugs and maintain margins and small structures on the classification map, we propose to use weighting function such as Equ. (7). Calculating the correct margin map for hyper spectral image is a challenging task. For example, it may be made by applying a threshold to gradient image $\{p_j \in R, j = 1,2, \dots, n\}$. Therefore, ansuitable threshold should be selected. Instead of calculating the margin map, we advise the definition of next function:

$$\varepsilon(x_i) = 1 - \frac{\rho_i}{\alpha + \rho_i} \quad 0 < \varepsilon(x_i) \leq 1 \tag{7}$$

where, $\alpha > 0$ is a parameter which controls the approximation margin threshold? When $p_j = 0$ (no margin), then$\varepsilon(x_j) = 1$. Increasing p_j leads to smaller and closer to zero$\varepsilon(x_j)$. Thus as novelty of this paper, the spatial energy term is obtainable as:

$$U_{Spatial}(x_i) = \sum_{x_j \epsilon N_i} \beta \varepsilon(x_i)\left(1 - \delta(w_i, w_j)\right) \quad (8)$$

Both of the methods for calculating spatial energy function (Equ. (6) and Equ. (8)) are investigated in experimental implementations of this paper. The proposed procedure, generalizes Metropolis-Hasting algorithm [14] to the weighted margin.

3. Suitable Spatial Segmentation

The goal of this section is to automatically segment a hyper spectral image into acceptable homogeneous areas and to combine the gotten segmentation map with the available spectral information to improve the classification results obtained before. Segmentation of the hyper spectral image is an interesting mission, due to the high dimensionality of data. We examined three segmentation approaches and extend them to hyper spectral subjects.

3.1. Hyper spectral RCMG Segmentation

The first examined method is a watershed transformation. It is a area growing methodbased on the some assumptions such as if we calculate a gradient of the image, each minimum of the gradient agrees to the core of a homogeneous area, and high values match to the borders between spatial structures. Area growing is done from each local minimum, so that one area in the segmentation map contains of the set of pixels linked to one local minimum of the gradient. We mean borders between adjacent areas by watershed lines.

We examined the use of a watershed method for hyper spectral images in this paper. A brief description of the process leading to the best segmentation can be summarized as: a one-band RCMG of the original image is calculated by investigating watershed transformation and using the procedure of ghamisi et al. [15]. The image is divided into a set of areas, and one subset of watershed pixels, i.e., pixels located on the borders between areas. Finally, every watershed pixel is allocated to the spectrally most similar vicinity area.

In this section, Hyper spectral RCMG (HRCMG) is established to raise the structural separation ability in hyper spectral images based on single band gradient calculation procedure. There are some methods to compute the gradient of an image, such as “median centered difference gradient”. For hyper spectral images, one can simply investigate median centered difference gradient to single band PC1 (first component of Principle Component Analysis (PCA)). Median centered difference gradient is defined and calculated as the difference between Dilation and Erosion [16] of image Y by the structure factor E.

$$\rho_E(x) = \delta_E(x) + \varepsilon_E(x) \quad (9)$$

For computing the gradient of hyper spectral images, some methods were assumed; they can be separated into two groups: 1. computing the vector gradient and 2. Computing the multidimensional gradient.

Table 1. Some of sistance/similarity measures criteria

Measure	Formula
Euclidean Distance	$d_{Euc} = \sqrt{\sum_{k=1}^{L} \left\lvert x_{i,k} - x_{j,k} \right\rvert^2}$
City Block, Manhattan, rectilinear distance, taxicab norm L_1	$d_{CB} = \sum_{k=1}^{L} \left\lvert x_{i,k} - x_{j,k} \right\rvert$
Minkowski Distance L_r	$d_{Mk} = \sqrt[p]{\sum_{k=1}^{L} \left\lvert x_{i,k} - x_{j,k} \right\rvert^p}$
Canberra Distance	$d_{Can} = \sum_{k=1}^{L} \frac{\left\lvert x_{i,k} - x_{j,k} \right\rvert}{x_{i,k} + x_{j,k}}$
Lorentzian	$d_{Lor} = \sum_{k=1}^{L} \ln\left(1 + \left\lvert x_{i,k} - x_{j,k} \right\rvert\right)$
Sorensen	$d_{Sor} = \frac{\sum_{k=1}^{L} \left\lvert x_{i,k} - x_{j,k} \right\rvert}{\sum_{k=1}^{L} \left(x_{i,k} + x_{j,k}\right)}$
Chebyshev, chessboard L_∞	$d_{Cheb} = \max_k \left\lvert x_{i,k} - x_{j,k} \right\rvert$
Gower	$d_{gow} = \frac{1}{d} \sum_{k=1}^{L} \frac{\left\lvert x_{i,k} - x_{j,k} \right\rvert}{R_i}$
Soergel	$d_{sg} = \frac{\sum_{k=1}^{L} \left\lvert x_{i,k} - x_{j,k} \right\rvert}{\sum_{k=1}^{L} \max\left(x_{i,k}, x_{j,k}\right)}$
Kulczynski	$d_{kul} = \frac{\sum_{k=1}^{L} \left\lvert x_{i,k} - x_{j,k} \right\rvert}{\sum_{k=1}^{L} \min\left(x_{i,k}, x_{j,k}\right)}$
Squared Euclidean	$d_{sqe} = \sum_{k=1}^{L} \left(x_{i,k} - x_{j,k}\right)^2$
Pearson χ2	$d_P = \sum_{k=1}^{L} \frac{\left(x_{i,k} - x_{j,k}\right)^2}{x_{j,k}}$
Neyman χ2	$d_N = \sum_{k=1}^{L} \frac{\left(x_{i,k} - x_{j,k}\right)^2}{x_{i,k}}$
Squared χ2	$d_{SqChi} = \sum_{k=1}^{L} \frac{\left(x_{i,k} - x_{j,k}\right)^2}{x_{i,k} + x_{j,k}}$
Probabilistic Symmetric χ2	$d_{PChi} = 2\sum_{k=1}^{L} \frac{\left(x_{i,k} - x_{j,k}\right)^2}{x_{i,k} + x_{j,k}}$
Divergence	$d_{Div} = 2\sum_{k=1}^{L} \frac{\left(x_{i,k} - x_{j,k}\right)^2}{\left(x_{i,k} + x_{j,k}\right)^2}$
Clark	$d_{Clk} = \sqrt{\sum_{k=1}^{L} \left(\frac{\left\lvert x_{i,k} - x_{j,k} \right\rvert}{x_{i,k} + x_{j,k}}\right)^2}$
Additive Symmetric χ2	$d_{AdChi} = \sum_{k=1}^{L} \frac{\left(x_{i,k} - x_{j,k}\right)^2 \left(x_{i,k} + x_{j,k}\right)}{x_{i,k} x_{j,k}}$
Kullback–Leibler	$d_{KL} = \sum_{k=1}^{L} x_{i,k} \ln \frac{x_{i,k}}{x_{j,k}}$
Jeffreys	$d_J = \sum_{k=1}^{L} \left(x_{i,k} - x_{j,k}\right) \ln \frac{x_{i,k}}{x_{j,k}}$
K divergence	$d_{Kdiv} = \sum_{k=1}^{L} x_{i,k} \ln \frac{2x_{i,k}}{x_{i,k} + x_{j,k}}$
Vicis-Wave Hedges	$d = \sum_{k=1}^{L} \frac{\left\lvert x_{i,k} - x_{j,k} \right\rvert}{\min\left(x_{i,k}, x_{j,k}\right)}$

The vector gradient is based on the distance between the pixel vectors and generates a single band gradient from the B-band image. The gradient distance for hyper spectral image [1] is defined as the difference between the supermom and the infimum between x_p and set of vectors ψ as Equ. (10) Where $\psi = [x_p^1, x_p^2, \dots, x_p^e]$ is the set of e neighbor's vectors of pixel x_p (the set ψ does not contain x_p).

$$\nabla_{\psi,d}^{MB}(x_p) = sup_{i\in\psi}\{d(x_p, x_p^i)\} - inf_{j\in\psi}\{d(x_p, x_p^j)\} \qquad (10)$$

In order to evaluate the gradient of Equ. (10), several distances can be investigated such as Euclidean, Mahalanobis and Chi-Squared. Some suitable distances are mentioned in Table 1.

Another type of gradient distance for hyper spectral image is RCMG [1]. This gradient was offered by Evans and Liu [17] for color images. In the experimental implementations of this paper, we use the generalization of RCMG to hyper spectral images (as proposed HRCMG). To explain this new proposed technique, we clarified definition of the Color Morphological Gradient (CMG) [1] and then the RCMG at the end; we generalized it to hyper spectral images as the part of novelty of this paper.

For each pixel x_p; neighborhood set is defined as a vectors in the structure factor E as $X = [x_p^1, x_p^2, \dots, x_p^e]$, the set X contains x_p. The Color Morphological Gradient is defined as Equ. (11) That means the maximum distance between all pairs of vectors in the set X:

$$\nabla_{X,d}^{CM}(x_p) = \mathrm{Max}_{i,j\in X}\{d(x_p^i, x_p^j)\} \qquad (11)$$

Here we can select different distances. If we use the Euclidean distance, Equ. (11) can be printed as:

$$\nabla_{X,d}^{CM}(x_p) = \mathrm{Max}_{i,j\in X}\left\{\left\|x_p^i - x_p^j\right\|^2\right\} \qquad (12)$$

One disadvantage of CMG is that it is very sensitive to the noise. In order to overcome this problem, RCMG had been planned. The CMG technique includes the removal of two pixels that are furthest, then applying CMG to the remaining pixels. This procedure is repeated until a good estimation of the gradient is acquired. RCMG which used Euclidean distance is defined as follows:

$$\nabla_{X,d}^{RCM}(x_p) = \mathrm{Max}_{i,j\in[X,REM_r]}\left\{\left\|x_p^i - x_p^j\right\|^2\right\} \qquad (13)$$

$\mathrm{REM_r}$ is the set of r pairs of vector. The furthest pixels had been removed. Suitable value of parameter r in Equ. (13) depends on the main structure and choice the amount of the impairment in the image.

We express the hyper spectral RCMG (HRCMG) to combine the most spectrally similar adjacent and probably non-adjacent areas. We investigate using of Spectral Angle mapper (SAM) distance between the area mean vectors as the dissimilarity criterion between areas instead of Euclidean distance. Our proposed method can be summarized as (14).

where $x_i = (x_{i1}, \dots, x_{iB})^T$ and $x_j = (x_{j1}, \dots, x_{jB})^T$.

$$\nabla_{X,d}^{HRCM}(.) = \mathrm{Max}_{i,j\in[X,REM_r]}\left\{\mathrm{arcCos}\left(\frac{\sum_{b=1}^{B} x_{ib}x_{jb}}{(\sum_{b=1}^{B} x_{ib}{}^2)^{\frac{1}{2}}(\sum_{b=1}^{B} x_{jb}{}^2)^{\frac{1}{2}}}\right)\right\} \qquad (14)$$

3.2. Adequate Expectation Maximization

Another category of examined segmentation method is partial clustering. Clustering aims at grouping pixels into C clusters, so pixels belonging to the same cluster are spectrally similar. This paper assumes that pixels belonging to the same cluster are drawn from a multivariate Gaussian probability distribution [1]. The distribution parameters estimated by the EM procedure [18]. Spatial structures are well defined by the clustering method [18].However, as no spatial information is used during the clustering process, pixels with the same cluster label can form a connected spatial region within the spatial coordinates, or can belong to disjoint areas. Therefore, in order to get a segmentation map, a connected component labelling procedure is investigated to the image partitioning obtained by clustering; which means that we distinguish and label differently every adjacent group of pixels belonging to the same cluster.

The spectral segmentation approaches based on the similarities between the image pixels and group pixels and spatial information of these pixels are not reflected. There are two sets for these approaches: Histogram-based approaches and procedures based on groups. Histogram-based methods allocate pixels to the spectral areas. In hyperdimentional data, these methods need great memory and the results are not accurate at all. Group based segmentation approaches are seeking to find structures in distinct spectral feature space. So, grouping is the comprehensive segmentation pixels to similar groups. Clustering and grouping approaches were evaluated in [19]. Two major sets of grouping procedures can be noted as hierarchical strategies and partial approaches.

Usually the hierarchical approaches provide a tree that it has only one pixel as each group at the lowest level. (ie, each pixel contains a group), and in the higher levels, the most like groups are fused (and the number of groups decreases). Then, the structure which has the wanted group is selected. Hierarchical clustering method was applied to hyper spectral unsupervised classification in [20]. Hierarchical grouping is a wide-ranging method for image segmentation that can be applied to create a set of segmentation outcomes. However, it needs great time and memory requirements. A substitute method is EM. EM, as a segmentation method, is so easier than hierarchical grouping approaches and has some compensation when investigate in hyper spectral datasets in terms of computation and memory requirements. However, the number of groups should be chosen. The other problem is about dependency of outcomes to the initialization values. EM strategy [18] for hyper spectral images is investigated in this paper and extended to adequate version. Our method is to confine data set to local area and apply AEM process at each area.

3.3. Hyper Spectral RHSEG

The third method examined for segmentation in this paper is hierarchical image segmentation. This procedure, established by James Tilton et al. [21], is a combination of area growing and spectral clustering. In the initialization step, each pixel is considered as one area. In each iteration the procedure merges the most spectrally similar adjacent and probably non-adjacent areas. We use SAM as Equ. (15) between the area mean vectors as the dissimilarity criterion between areas.More than 70 distance/similarity measures instead of SAM are identified. Some of them are stated in Table 1.

$$\mathrm{SAM}(\mathrm{x_i}, \mathrm{x_j}) = \mathrm{arcCos}\left(\frac{\sum_{\mathrm{b}=1}^{\mathrm{B}} \mathrm{x_{ib}x_{jb}}}{(\sum_{\mathrm{b}=1}^{\mathrm{B}} \mathrm{x_{ib}}^2)^{\frac{1}{2}}(\sum_{\mathrm{b}=1}^{\mathrm{B}} \mathrm{x_{jb}}^2)^{\frac{1}{2}}}\right) \qquad (15)$$

where$x_i = (x_{i1}, \ldots, x_{iB})^T$ and $x_j = (x_{j1}, \ldots, x_{jB})^T$. For each iteration our proposed method finds the smallest dissimilarity criterion between spatially adjacent areas. Then we merge adjacent areas with the dissimilarity criterion equal to the found minimal value. Now, we merge non-adjacent areas if a dissimilarity value between them satisfies this expression. The process is repeated till convergence, and in the output we have a hierarchical sequence of image partitions.

The Hyper spectral RHSEG (HRHSEG) is examined as part of novelty in this paper. HRHSEG works in both spectral domain and spatial information. It is an extension of RHSEG [22] for hyper spectral images. The HSEG splits image into area classes or objects appropriately. However, the HSEG procedure is computationally intensive for hyper dimensional data sets. The computational requirement of the HSEG procedure can be meaningfully reduced through a recursive approximation to HSEG, called RHSEG [22]. The current RHSEG [22] does not successfully extract the information content from increasingly available high spatial resolution remotely sensed imagery data. Our experiments have exposed this method do not perform well on hyper spectral data. So, this paper extends this method to hyper spectral issue, called HRHSEG. Our proposed technique is developed form original HSEG for images to be used hyper spectral remote sensing implementation. HSEG method is offered to enlarge the area optimization technique based on Hierarchical Step-Wise Optimization (HSWO) [1]. Furthermore, it provides the option of fusion of non-contiguous areas in the spatial by spectral clustering. This paper proposes the use of the SAM for the mean vectors of the most similar pair of regions as a measure of dissimilarity. HRHSEG as RHSEG [23], provides hierarchical string section with high exactness. The parameters are chosen automatically by the method at [24]. Using the SAM criteria in this situation can be done as another novelty of this article.

4. Experimental Results

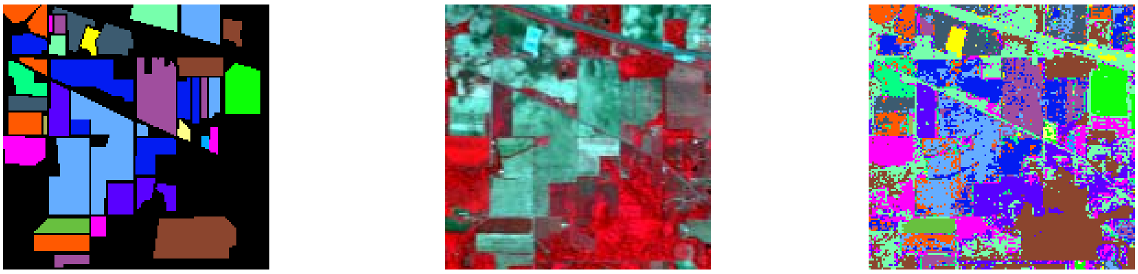
This section deals with experimental outcomes and compares proposed hyper spectral spatial-spectral classification with some existing approaches such as k-Nearest Neighbors (k=3) (3-NN), Maximum-Likelihood (ML), SVM, ECHO [6] and EMP [7]. Four remotely sensed hyper spectral images with different subjects (agricultural, volcanic, and urban areas) are used to provide the experimental results.

- Indian Pines image [10] was collected by AVIRIS sensor in nominally 224 bands. There are about 200valuable bands data. Each image has size of 145*145 pixels. The database is accompanied by a reference map; ground truth, whereby pixels are labeled as belonging to one of sixteen classes of vegetation or other land types. Ten classes have been investigated in this paper. A False composite image and different methods' classification results are depicted in figure 3. The Overall Accuracy (OA), Average Accuracy (AA) and Kappa factor of the proposed methods are compared with preceding methods in figure 4. The comparison of accuracy per class of the proposed methods (HRCMG+SVM, AEM+SVM, HRHSEG+SVM) is depicted in figure 5.
- Hekla image was achieved by AVIRIS sensor over the central volcano in Iceland.157 data channels were used for our tests. Spatial dimension of investigated image is 560×600 pixels. Twelve classes of land cover were examined. Artificial colored image and the classification results of different methods are shown in figure 6. Some metrics for evaluation of accuracy are described in figure 7. In both Indian Pines and Hekla images, 50 samples from reference data are considered as training samples randomly for each class and the remaining samples formed the test set. Accuracy of Hekla for each class for different methods is shown in figure 8.
- University of Pavia dataset [11] was recorded by ROSIS optical sensor over the metropolitan area, university of Pavia in Italy. This image has 340×610 pixels, with a spatial resolution of 1.3m per pixel. The number of captured data channels is 115 (with spectral range from 0.43µm to 0.86µm). Twelve channels were the most impaired deleted and the remaining 103 bands are used for experiments of this paper. The considered task is to assign each pixel to one of the nine classes of interest, knowing to which class some of the image pixels belong. Artificial composite color image and also the classification results of different methods are shown in figure 9. Thirty samples are randomly selected as training samples for each class. The comparison of the overall accuracy, average accuracy and Kappa factor for proposed approaches with SVM, EMP [7] and ECHO [6] is illustrated in figure 10.
- Center of Pavia dataset [11] was recorded from an urban area by the ROSIS sensor. The image used for experiments has 300×900 pixels with 102 spectral channels (13 channels with the most noise were excluded). Nine class labels of reference ground truth map with randomly chosen thirty labeled samples in each class are considered. The comparison of the accuracy, validity, and Kappa factor of different method is depicted in figure 11 for center of Pavia dataset. Classification results are also shown in figure 12.

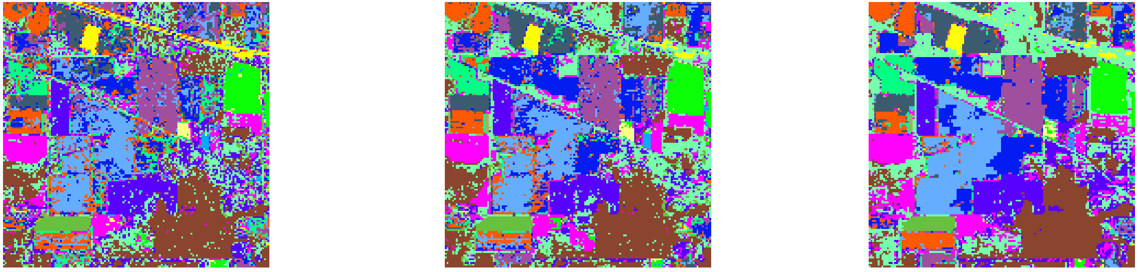
The experimental results show that our proposed method gives higher accuracies than the formerly approaches. Spectral-spatial classification meaningfully increases accuracies for almost all the classes, except for some matters (such as shadows in Indian Pines dataset), where accuracies are non-considerably reduced. As can be seen, all segmentation maps lead to higher classification accuracies and classification maps with more homogeneous areas compared to pixel wise classification. The best global accuracies are reached when performing HRHSEG and also AEM, i.e. performing segmentation both in the spatial and spectral domain.

Several segmentation methods are explored for this purpose. Furthermore, the outcomes of a pixel wise classification and a segmentation map are combined, using the median vote rule. In order to excavate the spatial information, we investigate some novel adaptive segmentation for hyper spectral images such as HRCMG, AEM and HRHSEG. These methods estimate adequate vicinity for all pixels in homogeneous area. The best classification outcomes are realized when performing hierarchical image segmentation which means performing segmentation both in the spatial and spectral domain. Furthermore, the classification accuracies shown in this paper are higher than all earlier results that we have found in the literature for this data set. HRCMG segmentation technique is fast and does not need input parameters, thus, automatic segmentation of the image is probable by this manner. HRHSEG technique produces sectors of image, and thus provides so detailed spatial information.

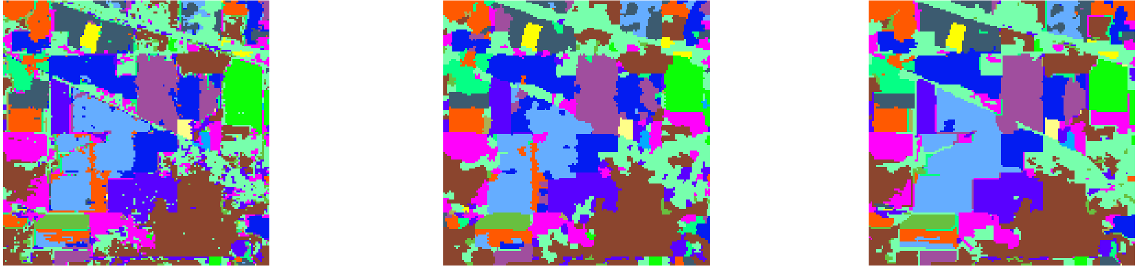
Our results show that the proposed technique outcomes in all cases (Indian Pines [10], Hekla, university and center of Pavia datasets[11]), all three criteria (overall and average accuracy and Kappa factor) are explored better than the previous methods (3-NN, ML, SVM, ECHO [6], EMP[7]), in Indian Pines image [10], according to more information in its texture, the improved criterion is more robust.



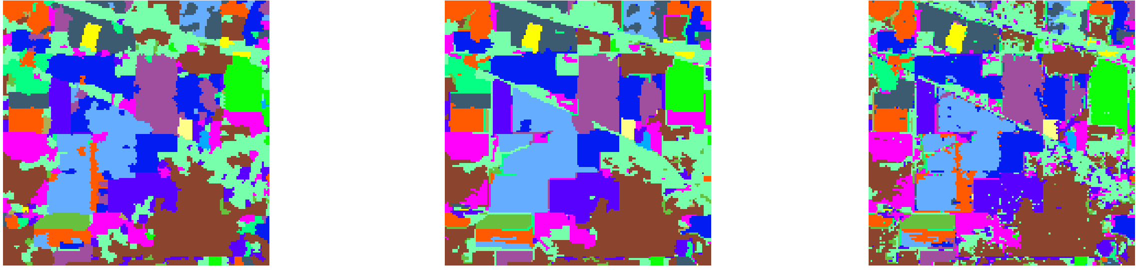
Reference data 3-band false color,837 636, 537nm ML



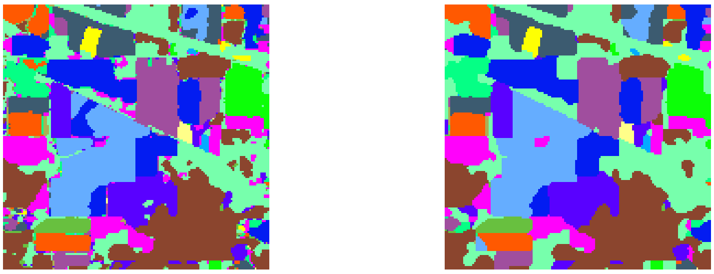
3-NN SVM ECHO [6]



SVM+Partial Clustering SVM+Watershed Segmentation SVM+Hierarchical Segmentation



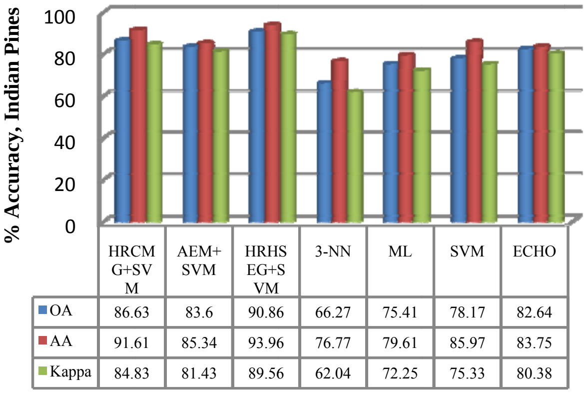
HRCMG+SVM HRHSEG+SVM AEM+SVM



SVM with MRF SVM classifier with locally weighting MRF

Figure 3. Classification map of Indian Pines image

This approach fits and maintains all spatial and spectral information correctly. Therefore, this proposed spatial-spectral hyper spectral classification approach is simple, accurate and fast.



	HRCMG+SVM	AEM+SVM	HRHSEG+SVM	3-NN	ML	SVM	ECHO
🟦 OA	86.63	83.6	90.86	66.27	75.41	78.17	82.64
🟥 AA	91.61	85.34	93.96	76.77	79.61	85.97	83.75
🟩 Kappa	84.83	81.43	89.56	62.04	72.25	75.33	80.38

Figure 4. Indian Pines dataset, Comparison of the accuracy, reliability and Kappa factor of the proposed method with previous methods

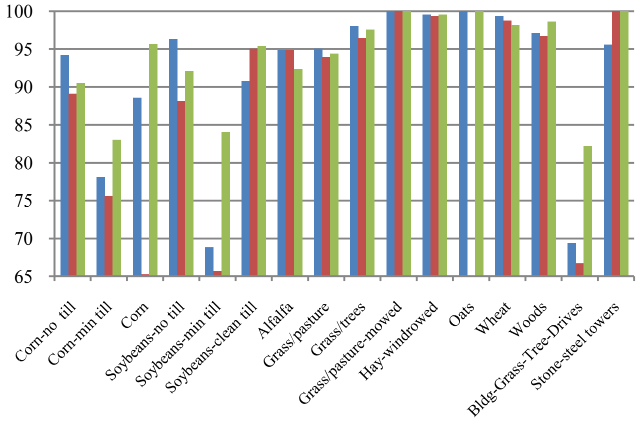
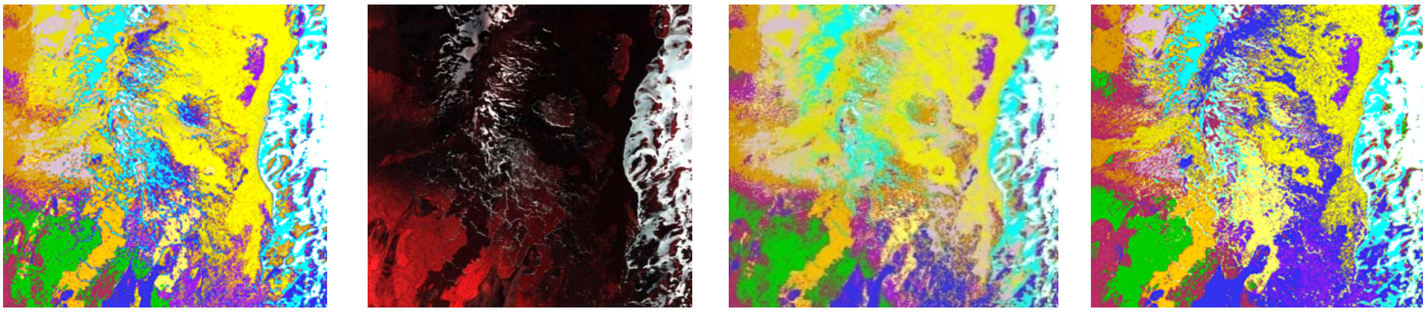


Figure 5. Indian Pines, Comparison of the accuracy, reliability and Kappa factor of the proposed method with previous approaches and accuracy per class (Blue: HRCMG+SVM, RED: AEM+SVM, Green: HRHSEG+SVM)

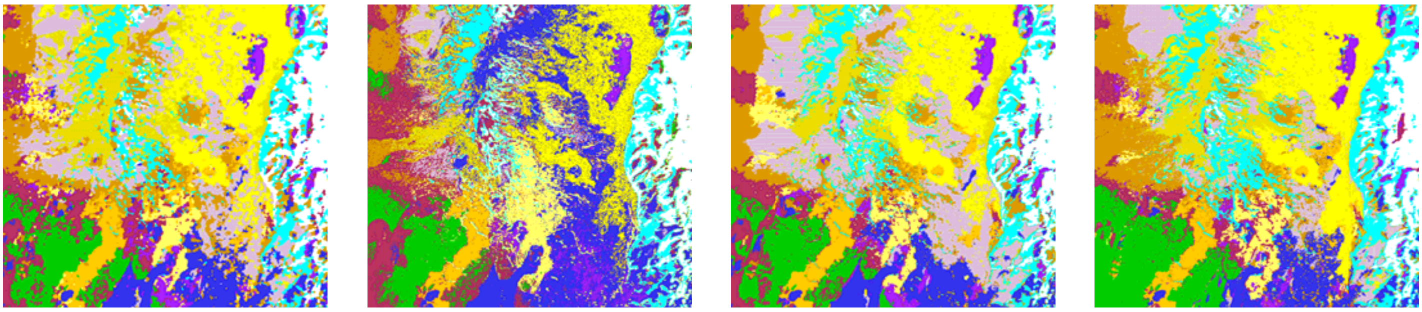


3-NN

3-band false color,1125,636,576 nm

SVM

ML

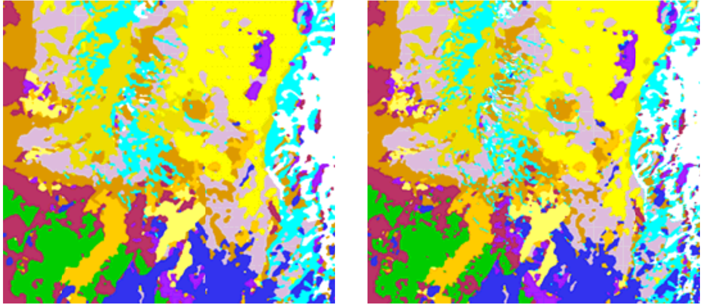


HRCMG+SVM

ECHO [6]

HRHSEG+SVM

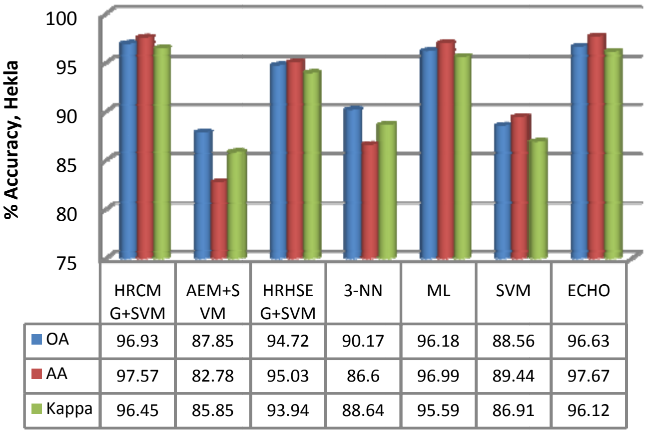
AEM+SVM



SVM classifier with MRF

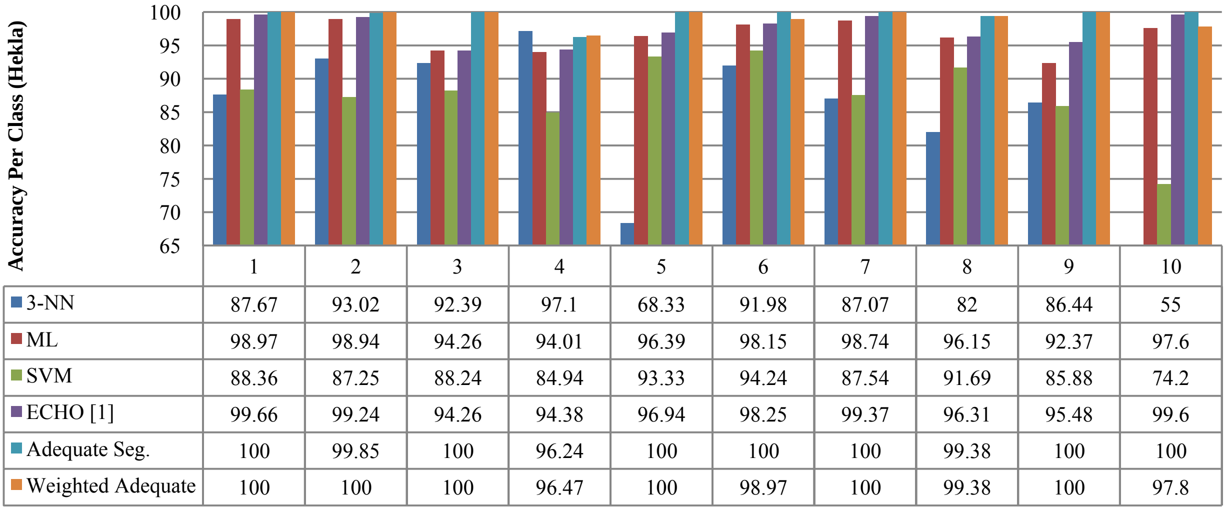
SVM classifier with locally weighting MRF

Figure 6. Classification map of Hekla volcano, Iceland



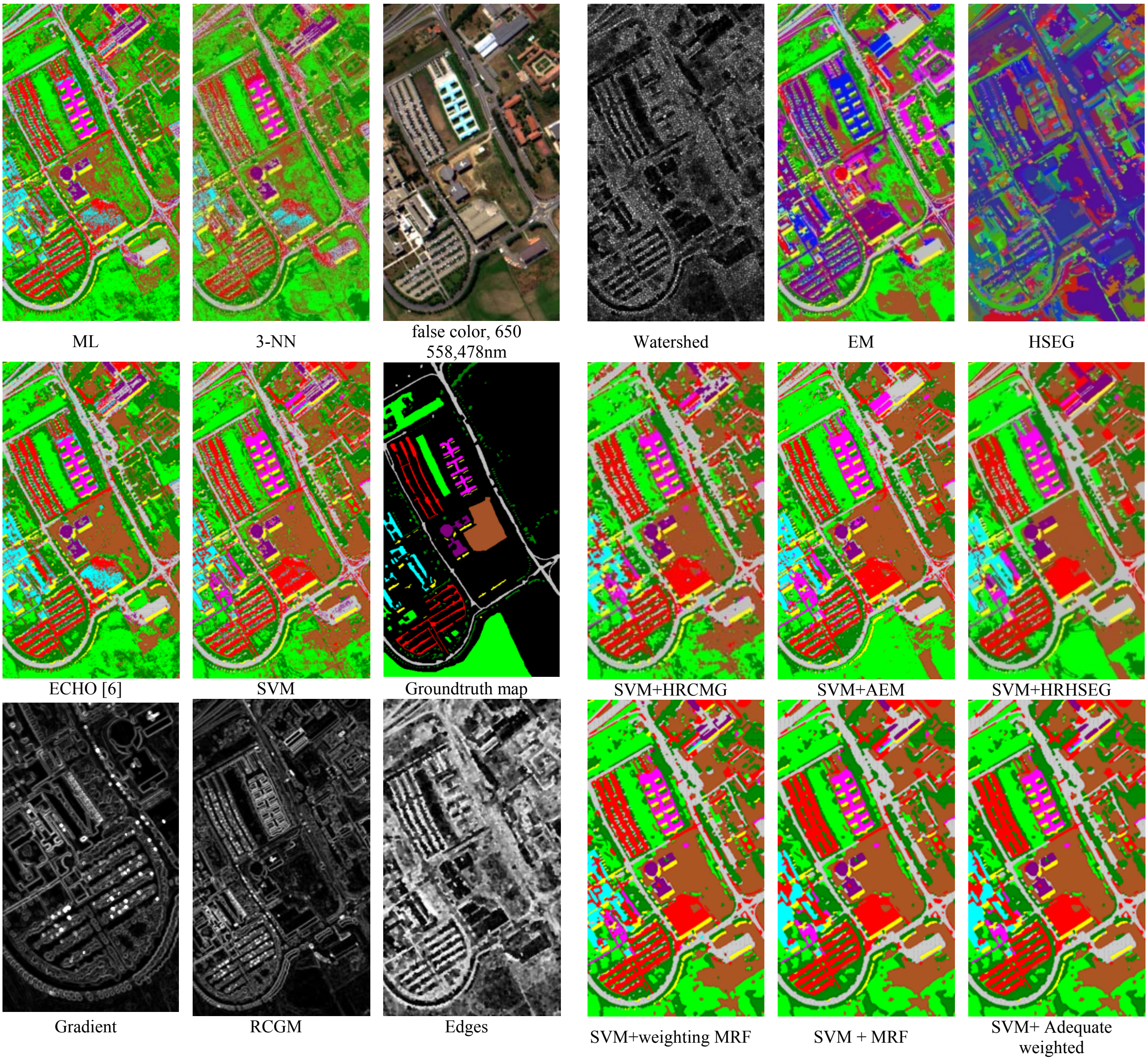
	HRCMG+SVM	AEM+SVM	HRHSEG+SVM	3-NN	ML	SVM	ECHO
🟦 OA	96.93	87.85	94.72	90.17	96.18	88.56	96.63
🟥 AA	97.57	82.78	95.03	86.6	96.99	89.44	97.67
🟩 Kappa	96.45	85.85	93.94	88.64	95.59	86.91	96.12

Figure 7. Hekla dataset, comparison of the accuracy, reliability and Kappa factor of the proposed method with previous methods



	1	2	3	4	5	6	7	8	9	10
🟦 3-NN	87.67	93.02	92.39	97.1	68.33	91.98	87.07	82	86.44	55
🟥 ML	98.97	98.94	94.26	94.01	96.39	98.15	98.74	96.15	92.37	97.6
🟩 SVM	88.36	87.25	88.24	84.94	93.33	94.24	87.54	91.69	85.88	74.2
🟪 ECHO [1]	99.66	99.24	94.26	94.38	96.94	98.25	99.37	96.31	95.48	99.6
🟦 Adequate Seg.	100	99.85	100	96.24	100	100	100	99.38	100	100
🟧 Weighted Adequate	100	100	100	96.47	100	98.97	100	99.38	100	97.8

Figure 8. Accuracy of Hekla for each class for different methods



ML | 3-NN | false color, 650 558,478nm | Watershed | EM | HSEG

ECHO [6] | SVM | Groundtruth map | SVM+HRCMG | SVM+AEM | SVM+HRHSEG

Gradient | RCGM | Edges | SVM+weighting MRF | SVM + MRF | SVM+ Adequate weighted

Figure 9. Classification map of university of Pavia dataset

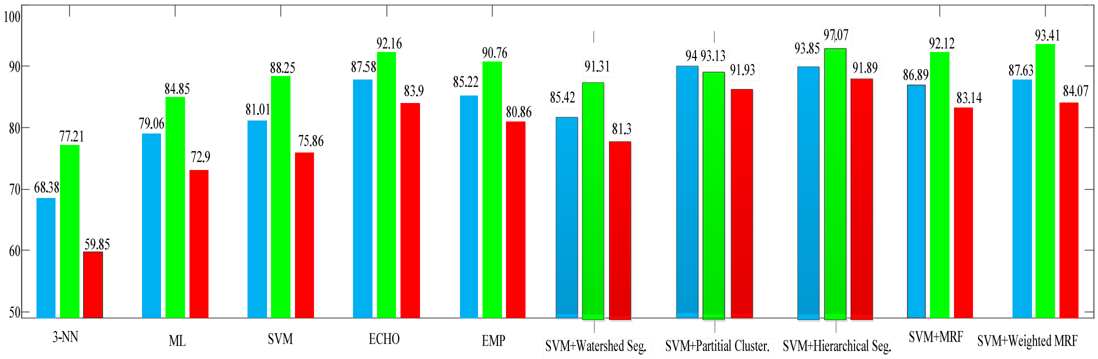


Figure 10. The comparison of the overall accuracy, average accuracy and Kappa factor for SVM, SVM+hyperspectral watershed segmentation, SVM+adequate partial clustering, SVM+hyperspectral hierarchical segmentation, EMP [7] and ECHO [6]

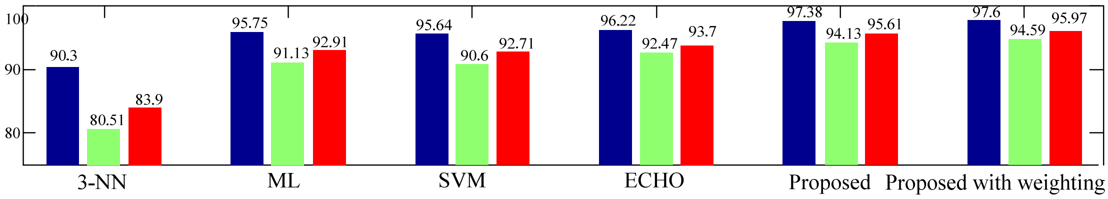


Figure 11.Center of Pavia dataset; the comparison of the accuracy, validity, and Kappa factor of different method. (Blue: overall accuracy, Green: average accuracy and Red: 100*Kappa factor)

5. Conclusions

A new spectral - spatial hyper spectral classification method is offered in this paper. We suggest use of spatial-spectral information simultaneously. Our planned spectral-spatial classification increases accuracies when compared to some older classification approaches. It is suggested to perform segmentation in order to use every spatial region as a proper proximity area for all the pixels within its area.

The extracted spatial information can be merged with spectral information in the classification. Finally, the proposed method has been used with probabilistic spatial energy regulation and median voting rule for fusion labels of classification and segmentation of hyper spectral images and the outcomes are presented in some imperial cases.

The proposed tactic of this paper is the usage probabilistic spatial energy regulations within the neighboring vicinity based on the extending some segmentation approaches to adaptive adequate subjects. Our proposed technique involves of performing probabilistic outputs for SVM classification followed by a weighted MRF based spatial regulation. One of the significant novelties is the integration technique of weighting function in the energy function in Markov random fields, which seeks to preserve margins during spatial uniformity. The proposed scheme with and without weighting method is examined on four real datasets and is compared with pixel wise approaches such as 3-NN, ML and SVM classification procedures and some spectral – spatial classification approaches such as ECHO [6] and EPM[7]. Innovative proposed segmentation approaches such as HRCMG, AEM and HRHSEG are examined successfully.

References

[1] J. A. Richards, and X. Jia, *Remote Sensing Digital Image Analysis: an Introduction*, 5th edition, 2013.

[2] M. Borhani, and H. Ghassemian, "Novel Spatial Approaches for Classification of Hyper Spectral Remotely Sensed Landscapes," *Proc. IEEE Intl Conf. Artificial Intelligence and Signal Processing*, pp. 41-51, 2013.

[3] M. Borhani, and H. Ghassemian, "Hyper Spectral Spatial-spectral Feature Classification based on Adequate Adaptive Segmentation," *Proc. IEEE Intl Conf. Artificial Intelligence and Signal Processing*, pp. 4-6, 2014.

[4] M. Borhani, and H. Ghassemian, "Kernel Grouped Multivariate Discriminant Analysis for Hyper Spectral Image Classification," *Proc. IEEE Intl Conf. Artificial Intelligence and Signal Processing*, pp. 3-13, 2013.

[5] M. Borhani, and H. Ghassemian, "Hyper Spectral Image Classification based on Non-uniform Spatial-spectral Kernels," *Proc. IEEE Intl Conf. Artificial Intelligence and Signal Processing*, pp. 1-5, 2014.

[6] R. L. Kettig, and D. A. Landgrebe, "Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects," *IEEE Trans. Geoscience Electronics*, vol. 14, no. 1, pp. 19-26, 1976.

[7] A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, J. A. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent Advances in Techniques for Hyper Spectral Image Processing," *Journal of Remote Sensing of Environment*, vol. 11, no. 3, pp. 110-122, 2009.

[8] M. Borhani, and H. Ghassemian, "Hyper Spectral Image Classification based on Spectral-spatial Features using Probabilistic SVM and Locally Weighted Markov Random Fields," *Proc. IEEE Intl Conf. Artificial Intelligence and Signal Processing*, pp. 4-10, 2014.

[9] H. X. Li, J. L. Yang, G. Zhang, and B. Fan, "Probabilistic Support Vector Machines for Classification of Noise Affected Data," *International Journal of Information Sciences*, vol. 22, no. 1, pp. 60-71, 2013.

[10] AVIRIS Indian Pines Hyper Spectral Dataset, https://engineering.purdue.edu/~biehl/MultiSpec/hyperspectral.html, May 2009.

[11] AVIRIS Salinas Valley, ROSIS Pavia University Datasets, http://ehu.es/ccwintco/index.php/Hyper spectral _Remote_Sensing_Scenes, June 2008.

[12] J. C. Platt, *Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods*, Advances in Large Margin Classifiers, 1999.

[13] C. Chang, and C. Lin, "LIBSVM: A library for Support Vector Machines, " http://www.csie.ntu.edu.tw/~cjlin/ libsvm, August 2007.

[14] J. P. Wang, "Stochastic Relaxation on Partitions with Connected Components and its Application to Image Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp. 619-636, 1998.

[15] P. Ghamisi, M. S. Couceiro, F. M. L. Martins, and J. Atli Benediktsson, "Multilevel Image Segmentation based on Fractional-order Darwinian Particle Swarm Optimization," *IEEE Trans. Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2382-2394, 2014.

[16] N. Young, and A. N. Evans, "Median Centered Difference Gradient Operator and its Application in Watershed Segmentation," *Electronics Letters*, vol. 47, no. 3, pp. 178-180, 2011.

[17] A. N. Evans, and X. U. Liu, "A Morphological Gradient Approach to Color Edge Detection," *IEEE Trans. Image Processing*, vol. 15, no. 6, pp. 1454-1463, 2006.

[18] Z. Liang, and S. Wang, "An EM Approach to MAP Solution of Segmenting Tissue Mixtures: A Numerical Analysis," *IEEE Trans. Medical Imaging*, vol. 28, no. 2, pp. 297-310, 2009.

[19] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data Clustering: A review," *ACM Computing Surveys*, vol. 31, no. 3, pp. 264-323, 1999.

[20] S. Lee, and M. M. Crawford, "Unsupervised Multistage Image Classification using Hierarchical Clustering with a Bayesian Similarity Measure," *IEEE Trans. Image Processing*, vol. 14, no. 3, pp. 312-320, 2005.

[21] J. C. Tilton, Y. Tarabalka, P. M. Montesano, and E. Gofman, "Best Merge Region-Growing Segmentation with Integrated Nonadjacent Region Object Aggregation," *IEEE Trans. Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4454-4467, 2012.

[22] J. M. Beaulieu, and M. Goldberg, "Hierarchy in Picture Segmentation: a Stepwise Optimization Approach," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, no. 2, pp. 150-163, 1989.

[23] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, *Spectral-spatial Classification of Hyper Spectral Images using Segmentation-derived Adaptive Neighborhoods*, ISTE Ltd and John Wiley and Sons Inc, 2009.

[24] A. J. Plaza, and J. C. Tilton, "Automated Selection of Results in Hierarchical Segmentations of Remotely Sensed Hyper Spectral Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, no. 2, pp. 4946-4949, 2005.

Mostafa Borhani received the B.Sc. and M.Sc. Degrees in Electrical Engineering from Sharif University of Technology, Tehran, Iran in 2002 and 2004 respectively. He is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering at the Tarbiat Modares University, Tehran, Iran. His research interests include the hyper spectral image analysis, kernel machines and pattern recognition in remote sensing applications.

E-mail: m.borhani@modares.ac.ir

Hassan Ghassemian received the B.S.E.E. degree from Tehran College of Telecommunication in and the M.S.E.E. and Ph.D. degree from Purdue University, West Lafayette, USA in 1984 and 1988 respectively. Since 1988, he has been with Tarbiat Modares University in Tehran, Iran, where he is a Professor of Computer and Electrical Engineering. His research interests focus on Multi-Source Signal/Image Processing and Information Analysis and Remote Sensing.**E-mail:** ghassemi@modares.ac.ir

Paper Handling Data:

Submitted: 16.06.2014
Received in revised form: 22.01.2015
Accepted: 25.01.2015
Corresponding author: Mostafa Borhani,
Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran.