

# An adaptive framework for spectral-spatial classification based on a combination of pixel-based and object-based scenarios

Amin Zehtabian<sup>1</sup>  · Hassan Ghassemian<sup>1</sup>

Received: 8 August 2016 / Accepted: 2 March 2017 / Published online: 1 April 2017  
© Springer-Verlag Berlin Heidelberg 2017

**Abstract** Remotely sensed image analysis using spectral-spatial information plays a key role in modern remote sensing applications. This article presents a new semi-automatic framework for spectral-spatial classification of hyperspectral images. The proposed framework benefits from a combination of pixel-based and object-based classification scenarios in which the main parameters are adaptively tuned. In order to reduce the complexity of the method, an unsupervised band selection technique is used as well. Meanwhile, the wavelet thresholding is applied in order to smooth the selected bands. The classification results after applying the proposed method to well-known standard hyperspectral datasets are better than those of the most of the other state-of-the-art approaches. As an example, the overall classification accuracy achieved by applying the proposed semi-automatic spectral-spatial classification framework to the Salinas dataset is more than 99% for 10% training samples per class. Moreover, the vital parameters are adaptively set in our approach.

**Keywords** Hyperspectral images · Spectral-spatial classification · Segmentation · Wavelet thresholding · Band selection · Support vector machines

---

Responsible editor: H. A. Babaie

✉ Amin Zehtabian  
amin.zehtabian@modares.ac.ir

<sup>1</sup> Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran

## Introduction

Hyperspectral images possess considerable amounts of useful spatial/textural information that cannot be addressed effectively by the use of traditional pixel-based image analysis approaches. Challenges arise especially when the spatial resolution of images is very high, and as a consequence, neighboring pixels are highly correlated (Fauvel et al. 2013). One solution to overcome this problem is to develop data analysis methods that are able to sufficiently exploit the spectral, spatial and textural information present in the remotely sensed data.

There are numerous spectral-spatial classification methods which have been presented so far in the state-of-the-art. Some useful surveys on the related works including object-based classification techniques as well as other recent advances in remotely sensed data classification approaches can be found in the works of Plaza et al. (2009), Lu and Weng (2007), Liu and Xia (2010), Blaschke (2010), Blaschke et al. (2014), Holbling et al. (2015), Samal and Gedam (2015), and Shi and Mao (2016). Some of the most recent works dealing with spectral-spatial classification of images (which are more similar to the proposed approach) are briefly reviewed in the following paragraphs.

Bernabe et al. (2014) proposed a spectral-spatial classification methodology that was especially suitable for classification of multispectral data with limited spectral resolution. They exploited kernel feature extraction in order to expand the dimensionality of data and then extracted spatial features using extended multi-attribute profiles that were built on the spectral features.

Gaetano et al. (2015) proposed a watershed segmentation approach in order to provide an object-level representation of images to be classified in subsequent stages. They benefited from automatically extracted morphological and spectral

markers to control over-segmentation problems which may arise with watershed.

Ghamisi et al. (2014a) proposed a spectral-spatial classification approach in which two consecutive segmentation levels were applied in order to provide an object-map to be classified later by Support Vector Machine (SVM) classifier. The two segmentation levels were based on fractional-order Darwinian particle swarm optimization and mean shift segmentation. The authors continued their efforts (Ghamisi et al. 2014b) to develop an automatic framework for classification of hyperspectral data which simultaneously utilized both spectral and spatial information. They applied morphological attribute profiles for including the spatial information and utilized decision boundary feature extraction technique as well as discriminant analysis feature extraction method in order to reduce the effect of Hughes phenomenon. In one of their most recent works, Ghamisi et al. (2015) carried out a comprehensive survey on the spectral-spatial classification approaches in which attribute profiles were used.

Kang et al. (2014) proposed a classification approach in which a pixel-wise classification map was represented as multiple probability maps. After applying edge-preserving filters to each of the probability maps, the finalized class-map was achieved by assigning each pixel to one of the classes with regard to the maximum probability.

Readers can also refer to Khodadadzadeh et al. (2014), Mylonas et al. (2015), Mirzapour and Ghassemian (2015), Golipour et al. (2016), Zehtabian and Ghassemian (2015), Samal and Gedam (2015), Zahidi et al. (2015), Machala and Zejdova (2014), Li et al. (2014), for further discussions on object-based remotely sensed image classification.

The main concern of the presented paper is to propose a new framework for development of hyperspectral data classifiers in which more spectral and spatial information is utilized. In order to satisfy this concern, the proposed framework benefits from a combination of the pixel-based and object-based classification scenarios. In this framework, first an unsupervised band selection technique is applied in order to produce a limited number of representative bands. It reduces the complexity of the proposed algorithm. The wavelet thresholding is then exploited in order to smooth the selected bands and produce larger objects with higher level of homogeneity. In the next step, a set of object-maps are produced using a novel Pixon-based segmentation method which is applied to each representative band, separately. The proposed Pixon extraction technique exploits an innovative distance metric. Two different sets of class-maps are then produced, one assigning each object to one of the classes using three different spectral/spatial features, and another made from applying majority voting inside each object with regard to a reference pixel-based classification map. A single class-map is finally achieved by using another level of majority voting among all the individual class-maps. In the proposed framework, all

the vital parameters (i.e. the Pixon extraction parameter used in the segmentation step as well as the soft thresholding parameters used in the wavelet smoothing) are adaptively tuned and hence there is no need to manually set the parameters. The gained results are compared to those of the other state-of-the-art approaches to prove the considerable performance of the proposed framework in terms of classification ratios.

The rest of the presented article is ordered as follows: The proposed spectral-spatial framework section describes the methodology of the proposed spectral-spatial classification framework. Each subsection of the second section provides detailed information about one of the main steps in the proposed process. Experimental results and discussions are reported in Experimental results section, while the conclusions are provided in Conclusion section.

## The proposed spectral-spatial framework

General architecture of the proposed idea for spectral-spatial classification of hyperspectral data is illustrated in Fig. 1. The following subsections explain the presented framework in more details.

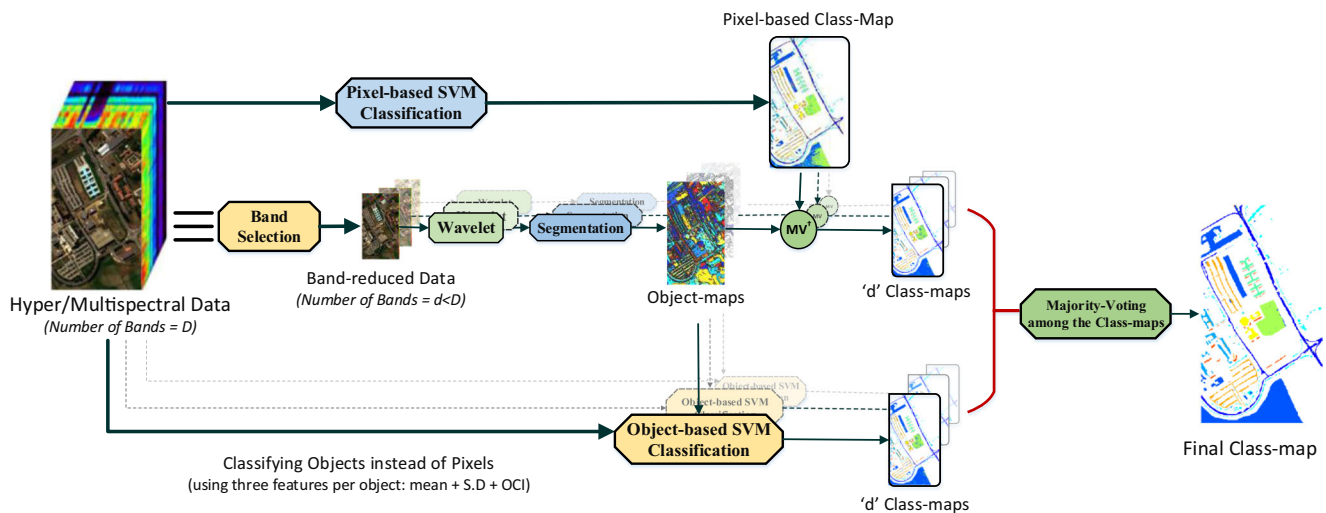
### Pixel-based SVM classification

In order to provide a pixel-wise thematic map as a reference for further analysis of the data, a pixel-based classification is used which is based on support vector machines. In this paper, the SVM implementations are carried out using the library package LibSVM (Chang et al. 2001).

### Band selection step

While dealing with high dimensional remotely sensed data, at least two serious drawbacks may arise: the lack of adequate labeled training samples, and the data redundancy (Fauvel et al. 2013; Sun et al. 2015). The latter is the concern of this subsection. In other words, using modern sensors with high spectral resolutions often leads to hyperspectral datasets with a large number of bands which are highly correlated (Fauvel et al. 2013). Therefore, selecting a reasonable range of spectral bands and eliminating the others may result in reduced information redundancy without losing important details. This leads to a considerable reduction in the computational cost of most of the remote sensing applications (such as data classification) with only a minor degradation in their accuracies (Martinez-Uso et al. 2007).

As shown in Fig. 1, in order to reduce the dimensionality while keeping the main structure of data unchanged, we suggest applying the WaLuMI (Ward's Linkage strategy Using Mutual Information) band selection method (Martinez-Uso et al. 2007). WaLuMI is a fully unsupervised hierarchical



**Fig. 1** The proposed framework for spectral-spatial classification of hyperspectral data. ( $\dagger$  MV means majority voting inside each object with regard to the reference class-map)

clustering approach which utilizes the mutual information based distance in order to form cluster of bands which their intra-cluster variance is minimized (Martinez-Uso et al. 2007). On the other hand, since different subsets of bands are mutually exclusive, the inter-cluster variance is maximized (Ward 1963). In the WaLuMI band selection approach, the most similar bands are hierarchically merged and constitute larger clusters until a specified number of clusters are reached (Cariou et al. 2011).

After forming the band clusters, one representative band for each cluster must be selected and then fed into the next steps for further analysis. Martinez-Uso et al. (2007) suggested two techniques for selection of the cluster representative bands. In their first technique, for a given cluster, the band with highest average correlation (with regard to the other bands in that cluster) is selected as representative of the cluster. In their second technique, the band with the highest average divergence (with regard to the other bands in the cluster) is chosen as representative band (Martinez-Uso et al. 2007). This approach is called WaLuDI (Ward's Linkage strategy Using Divergence). In this paper, we exploit the first approach (i.e. WaLuMI). Indeed, after several sets of experiments we found out that benefiting from the mutual information often results in slightly higher classification accuracies when the SVM classifier is used. After the band selection step, the number of selected bands is ' $d$ ' (i.e., ten bands in the proposed schema) which is much lower than the number of bands in the original hyperspectral data (i.e. ' $D$ ' which is often larger than 100 in a hyperspectral data).

### Wavelet thresholding

In this paper, we propose applying a wavelet-based preprocessing step in order to smooth the selected bands. The smoothing step results in larger objects with higher level of homogeneity. In other words, by eliminating the redundant

details in each band of data, the smoothing preprocessing step reduces the probability of over-segmentation. This consequently leads to smaller number of extracted objects, lower computational time for object-based classification, and more robustness against unwanted environmental noise (Hassanpour et al. 2011, Zehtabian et al. 2015).

Using the wavelet thresholding for data smoothing may itself cause over-relaxation in which some useful details such as edges and boundaries may be relocated, faded or even disappeared in the smoothed image. To avoid these problems, the value of the threshold in the wavelet thresholding technique must be tuned appropriately. Other concerns are choosing a proper mother wavelet as well as determining an adequate number of decomposition levels.

At the first level of the wavelet filtering, each band of data is decomposed into four frequency channels (sub-bands) namely low-low (LL), low-high (LH), high-low (HL) and high-high (HH), each of which with a particular coefficient (Hassanpour et al. 2011). Among these channels, the LL contains the low frequency components of the image which constitute its main structure, while the others possess the high frequency components which can be mainly regarded as redundant details as well as noises (Burrus et al. 1998; Gupta and Kaur 2002). At the next levels of the wavelet transformation, the decomposition process is recursively applied to the low frequency channel (LL) to generate the sub-bands at the next levels. In other words, only the low frequency coefficients are subject to further processing. After performing the decompositions levels, the thresholding algorithm is applied to all sub-bands from each level, except to the LL channels that are exempt from being processed. In this article, we use 'sym6' mother wavelet with four levels of decomposition.

The thresholding techniques used for processing the wavelet coefficient can be categorized into two groups: hard

and soft thresholding. In hard thresholding, the wavelet coefficients, which are smaller than the threshold value, are substituted with zero while the other coefficients are kept unchanged. In the soft thresholding, the coefficients higher than the threshold are reduced as well. This reduction is done in accordance with the amount of the threshold. The soft thresholding function can be stated as follows (Hassanpour et al. 2011):

$$H(c) = \begin{cases} \text{sign}(c)(|c|-\tau) & , |c| > \tau \\ 0 & , \text{otherwise} \end{cases} \quad (1)$$

where  $c$  is a given coefficient,  $\tau$  is the threshold and  $H(c)$  is the soft thresholding function.

For a beneficial wavelet thresholding, a proper threshold value is needed. There are a few threshold estimation techniques that have been proposed. Among them, three methods are more widely used, namely Visushrink, Bayesshrink and Sureshrink (Hassanpour et al. 2011). In this paper, we benefit from the Bayesshrink method in which a Bayesian framework is used to derive sub-band dependent thresholds. In the Bayesshrink approach, it is assumed that the wavelet coefficients in each sub-band can be summarized properly using the generalized Gaussian distribution (GGD) (Chang et al. 2000).

## Object extraction

As can be seen in Fig. 1, after applying the wavelet thresholding to each of the selected bands, the smoothed bands are fed into the next step which is a Pixon-based segmentation similar to what we have recently proposed in (Zehtabian and Ghassemian 2015), but with a major modification. In the traditional version of our proposed Pixon extraction algorithm, adjacent pixels were hierarchically merged together if the Euclidean distance between the pixels in the spectral space (which was expressed by the gray-levels of the pixels) was smaller than a predefined threshold (Zehtabian and Ghassemian 2015). Additionally, the merging procedure followed a predefined order in choosing appropriate pixels to be joined to the current Pixon. After merging the first pair of neighboring pixels, the Euclidean distance between their average spectral intensity and the spectral value of the next pixel was calculated and compared to the threshold. The Pixon extraction process continued until all the pixels in the image were analyzed and the final segmentation map was produced.

In the present paper, however, we apply a modified version of the Pixon extraction technique in which the Euclidean distance is substituted with a new distance metric. The newly proposed distance benefits from a higher level of textural information that exists in an image.

Noting  $S_P$  the average spectral intensity of  $P^{th}$  Pixon, and  $I_p$ , the gray-level of  $p^{th}$  pixel, the proposed distance between Pixon  $P$  and its adjacent pixel  $p$  can be calculated as:

$$dis = \frac{|S_P - I_p|}{G_P + G_p} \quad (2)$$

where  $|\cdot|$  stands for the 2nd norm,  $G_p$  is the sum of gradients around pixel  $p$  along eight cardinal/diagonal directions, and  $G_P$  denotes the average sum of gradients around all the pixels which are located in Pixon  $P$ . Therefore, the denominator of (2) can be regarded as a simple texture descriptor (especially when the sum of gradients is computed within a wider neighborhood) since it characterizes the image texture by addressing the pattern of spectral variations in a particular neighborhood. On the other hand, the numerator of (2) is the simple Euclidean distance between the spectral intensity of Pixon  $P$  and that of pixel  $p$ .

Since the proposed distance simultaneously incorporates the spectral and textural information into a single measure, it provides a more realistic analysis of images. As an example, consider a case in which there is a significant difference between the gray-levels of two adjacent pixels (or a Pixon and its adjacent pixel). In this case, the traditional Euclidean distance is large, reflecting that one of the two pixels may be located in an edge area. However, the Euclidean distance metric cannot make a difference if the considerable spectral contrast between two pixels is not because of the edge and it is caused by intensity variations due to the texture. In such cases, since the proposed distance metric considers the spectral variations in the neighborhood of Pixon  $P$  and pixel  $p$  as well, the denominator of (2) has a large value. Therefore, the value of the proposed distance in this case is considerably smaller in comparison with cases in which there is a real edge with no notable textural variations in its neighborhood.

Once the distance between a given pixel (or Pixon) and its neighboring pixel is computed, it must be compared to a predefined Pixon extraction threshold. This threshold can be either set for each band separately or for the multiband data as a whole. In this paper, we suggest setting this threshold for each of the selected bands. It is due to this fact that the proposed object extraction algorithm is applied to each band, separately.

Moreover, since a semi-automatic object-based classification framework is desired in this work, we suggest adaptively tuning the Pixon extraction thresholds as well. The adaptation technique used in the present article is similar to what we proposed in (Zehtabian and Ghassemian 2015). In our recently published work, in order to tune the Pixon extraction threshold for each band of data, the gradients of each pixel were firstly computed along the cardinal and diagonal directions. Then the differences between gradients in each opposite side were calculated and inserted into a new matrix. By this, four different matrices were produced. The elements of the

matrices were then powered by two and summed together. The square root of the result finally formed a unique matrix for each band of the data. Our research proved that a proper Pixon extraction threshold for each band could be achieved by multiplying the variance of the elements of the final matrix by a constant factor which was fixed for all remote sensing datasets. Further discussions about the proposed adaptation algorithm can be found in (Zehtabian and Ghassemian 2015). The sensitivity of the classification ratios to variation of the Pixon extraction threshold will be also evaluated in the next section.

Figs. 2 and 3 are provided in order to visually evaluate the proposed distance metric as well as some other competing distances. In these experiments, the results of applying nine other well-known distance metrics are plotted, namely Euclidean, Chi-Square, Cosine, Norm-1, Earth Mover, Kolmogorov-Smirnov, Jensen-Shannon Divergence, Kullback-Leibler Divergence, and Jeffrey Divergence. Technical details about the competing distances and metrics are comprehensively addressed in (Rubner et al. 2000). The data utilized in these experiments is the F210 dataset which is a multispectral aerial image with twelve spectral bands. The ground truth map (GTM) of F210 comprises nine different classes.

In order to plot Figs. 2 and 3, first we calculate the distances along horizontal (i.e. east-west) and vertical (i.e. north-south) directions, respectively. In other words, the distances between each spectral vector  $I(i, j)$  and its neighboring spectral vectors (i.e.  $I(i, j + 1)$  for horizontal direction and  $I(i + 1, j)$  for vertical direction) are calculated. Then the outcome values are inserted in the horizontal distance matrices and vertical distance

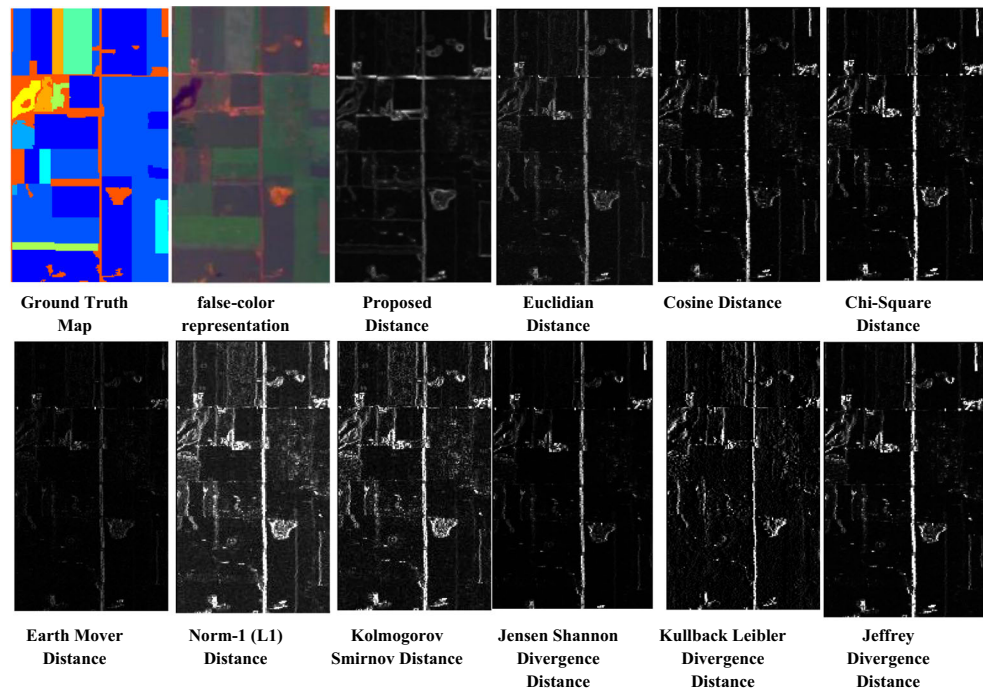
matrices, respectively. These matrices are then plotted as equivalent 2D gray-level graphs in Figs. 2 and 3, respectively for horizontal and vertical directions.

From Fig. 2, after calculating the distances along the horizontal direction, the vertical edges and boundaries of the data are emphasized. Moreover, by calculating the distances along the vertical direction, the horizontal edges and boundaries are highlighted (Fig. 3). However, as can be inferred from these figures, due to the ability of the proposed distance metric to make use of the textural information gained from the neighborhood of each pixel, it can also extract the horizontal edges while it is horizontally applied to the data, and vice versa. In other words, since larger amount of textural/spatial information is used in the proposed metric, using the horizontal (or vertical) distance per se may be adequate to extract all the horizontal, vertical and even diagonal boundaries. In these figures, in order to have a better assessment, the ground truth map and the false-color representation of the F210 dataset are also shown.

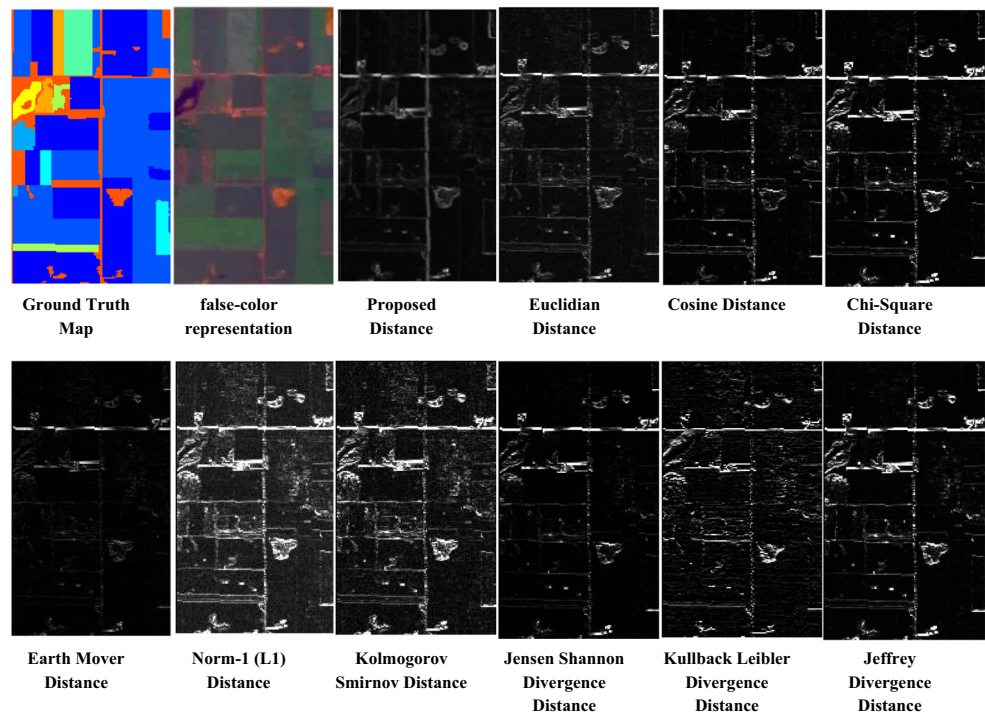
### Majority voting inside each object

After manually setting the thresholds and performing the Pixon-based segmentation, a unique object-map is constructed for each of the selected bands. We suppose the class-map achieved by the pixel-wise SVM classification as a reference. Referring to this reference, majority voting is then applied inside the objects from all of the object-maps to constitute new class-maps. In other words, by assigning the pixels of each object to the most frequent class within that object, ‘ $d$

**Fig. 2** Two-dimensional gray-level illustrations that are plotted according to various distance metrics applied to F210 multi-spectral data. Each image corresponds to distance values between all horizontal pairs of feature vectors in data



**Fig. 3** Two-dimensional gray-level illustrations that are plotted according to various distance metrics applied to F210 multi-spectral data. Each image corresponds to distance values between all vertical pairs of feature vectors in data



class-maps are produced. A visual example of the majority voting process utilized in this step is shown in Fig. 4. As can be inferred from this figure, in objects in which there is no majority for one class rather than the other classes, the majority voting is not applied (however, it rarely occurs in practice). Therefore, the pixel-wise classification result is used for the pixels located in such objects.

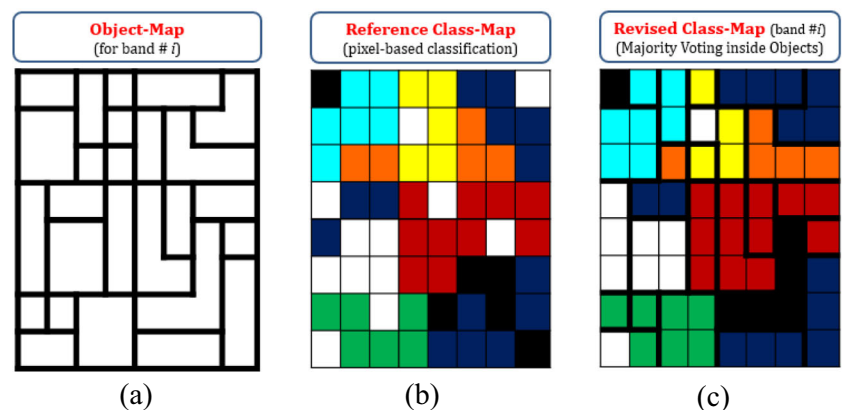
### Classification of objects

The second proposed approach for incorporating the spatial information achieved from the object-maps into the spectral-spatial classification framework is a typical object-based classification step in which each object is completely assigned to one of the classes. In order to classify the objects rather than the

pixels, we need to extract some relevant spectral/spatial features from the objects. In this paper, we suggest applying two widely used spectral features (i.e. mean and standard deviation of the pixels/vectors located in the object) as well as a newly developed spatial feature called object correlative index (OCI).

OCI describes the correlation between a given object and its neighboring objects using a well-defined spectral similarity measure (Zhang et al. 2013). This leads to a new model of spatial information that can be used as a description of the relationship between individual objects and the image as a whole. Finally, the value of the OCI spatial feature for a given object is the sum of the lengths of the correlative lines which are oriented toward various directions from the object's center of gravity to the apogee intersection points (Zhang et al. 2013).

**Fig. 4** A visual example of assigning the pixels located in each object to the most frequent class within that object. A given object-map constructed for one of the selected bands (a), the reference class-map produced by using the pixel-based SVM classifier (b), and the revised class-map after applying majority voting inside each object



After extracting the features from each object, the objects are classified using the stacked features. This results in ‘*d*’ new class-maps which are totally different from the class-maps previously achieved by applying majority voting inside each object.

### Majority voting among class-maps

From Fig. 1, after using two different approaches for incorporating the spatial and spectral information into the classification process, two sets of classification results are produced each with ‘*d*’ different class-maps. Finally, another level of majority voting process is carried out on the class labels in order to achieve a unique class-map. A schematic example of applying the majority voting process to three given class-maps is shown in Fig. 5.

### Experimental results

In order to evaluate the proposed spectral-spatial classification approach and compare it to the other state-of-the-art works, two well-known hyperspectral datasets are utilized in this article. The first one is captured by ROSIS-03 sensor over the University of Pavia in Italy, namely, the Pavia University dataset. This data comprises 115 spectral bands, however, 12 channels are eliminated in our work due to the noise problem. Each band of this data is of size 610 pixels by 340 pixels and the spatial resolution of each band is equal to 1.3 m per pixel. The ground truth of Pavia University dataset contains 9 different classes (Fauvel et al. 2013).

The second hyperspectral dataset is the Salinas dataset which has been collected by AVIRIS sensor. It comprises 204 spectral bands. The size of each band in the Salinas

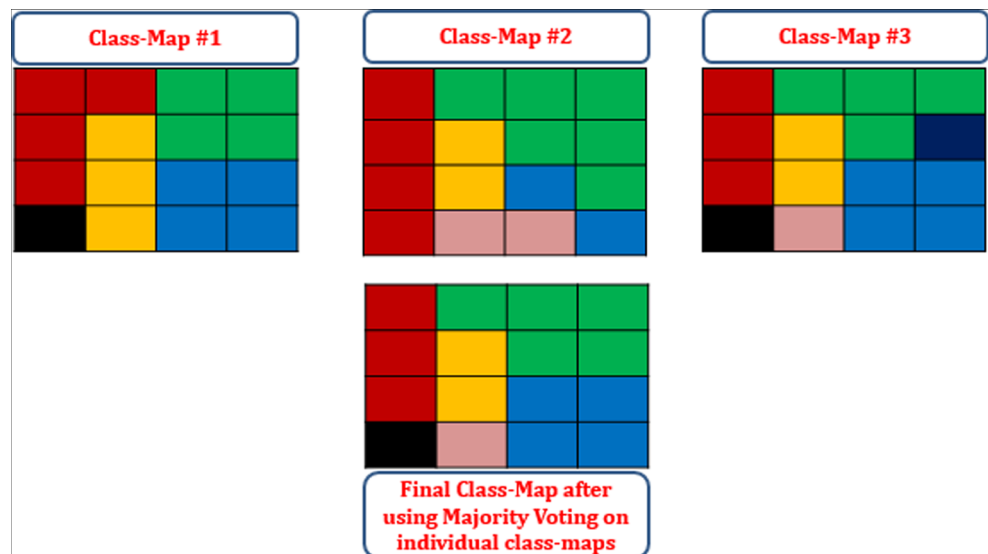
dataset is  $512 \times 217$  pixels and its ground reference map comprises 16 different agricultural classes. The spatial resolution of this data is equal to 3.7 m per pixel (Golipour et al. 2016).

Fig. 6 shows the classification map produced after applying the proposed spectral-spatial approach to the ROSIS-03 Pavia University dataset, when the number of training samples is equal to 50 per class. The GTM and the false color representation of this data as well as the thematic map resulted after performing a pixel-wise classifier are also shown in this figure in order to provide some references to highlight the efficiency of the proposed approach.

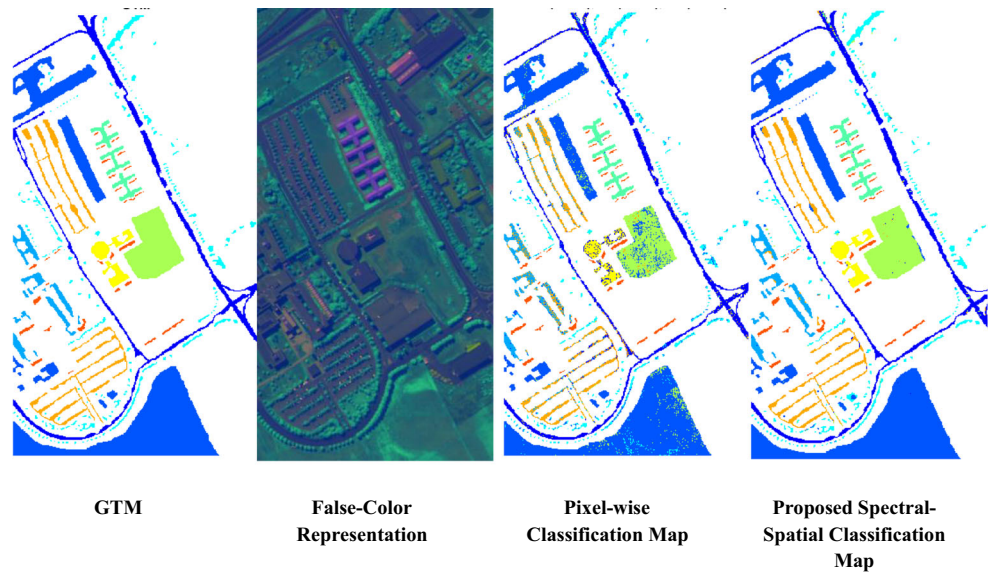
Meanwhile, Fig. 7 illustrates the differences between classification maps after applying two different versions of the proposed approach: one which uses only the traditional mean vector as the extracted feature of each object, and one which benefits from three different features (i.e. mean vector, standard deviation and OCI) per object. The data is still the same (i.e. the Pavia University dataset) and the standard set of training samples is used similar to what suggested by Fauvel et al. (2013). As can be inferred from Fig. 7, in a few number of regions, the classification results of applying the traditional mean vector feature are slightly better than those of applying the three features. Such regions are highlighted by red dashed circles. Moreover, from this figure, in several other regions (which two of them are marked with green circles), the proposed approach in which three features are used outperforms the other version of our method in which only mean vector is applied as the extracted feature of each object. This difference can be also expressed in terms of averaged accuracy (AA) and overall accuracy (OA) since AA and OA increase from 92.63% and 94.19% to 93.87% and 95.08%, respectively.

The quantitative comparisons are also reported in Tables 1 and 2 for Pavia University dataset (with 50 training samples

**Fig. 5** A visual example of applying the majority voting process to three given class-maps. Each colored square stands for one pixel, while each color represents a different class



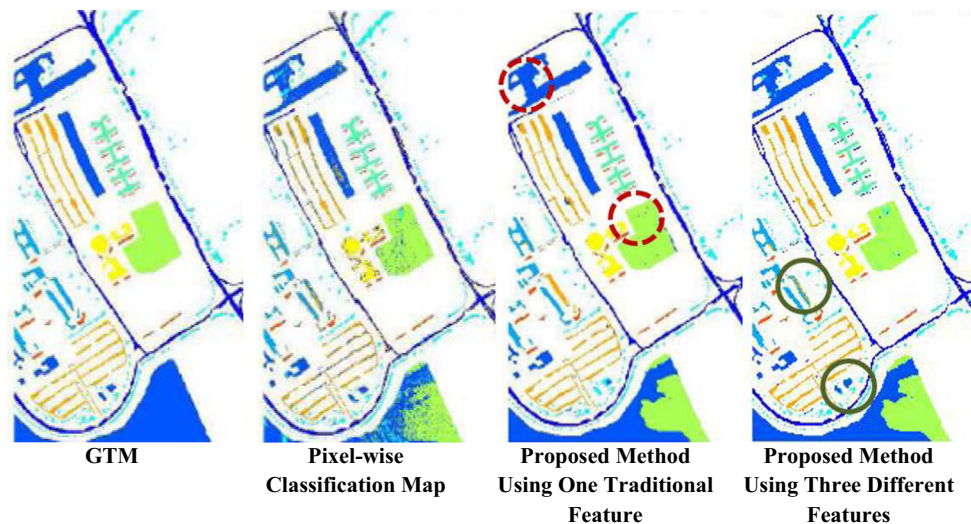
**Fig. 6** The classification maps produced after applying the proposed spectral-spatial classifier as well as a traditional pixel-wise SVM classifier to Pavia University data with 50 training samples per class



per class) and Salinas dataset (with 10% training samples per class), respectively. Since the process of selecting the training samples is random, the results of applying the proposed method are averaged after 30 runs before reporting in these tables. As can be inferred from the tables, the proposed spectral-spatial classification framework excels most of the other competing methods, especially for the Pavia University case.

In the next experiments, we analyze the sensitivity of the proposed approach to variation of its main operational parameter: the Pixion extraction threshold. Since, a band-by-band analysis is exploited in this article, the needed parameters have

been adaptively tuned and used for each band of the hyperspectral data, individually. To be more clear, for a hyperspectral data with “ $d$ ” bands, “ $d$ ” Pixion extraction thresholds as well as “ $2d$ ” wavelet parameters (i.e. the thresholds ‘ $\tau_{(b)}$ ’ and the number of levels in soft thresholding ‘ $N_{(b)}$ ’, while  $b$  is the band index) should be automatically set and then used in the proposed algorithm. However, it is almost impossible to report the effects of the variations in the parameters of each band on the final classification ratios. Alternatively, in order to assess the sensitivity of the proposed object-based classification method to the Pixion extraction threshold



**Fig. 7** The classification maps produced after applying the proposed spectral-spatial classifier with two different sets of extracted features as well as a traditional pixel-wise SVM classifier to Pavia University data with the standard training samples. In regions highlighted by *red dashed circles*, the classification result of applying the traditional mean vector

feature are slightly better than that of applying the three features (i.e. mean, S.D. and OCI). The regions marked with *green circles* are two examples among many of other regions in which using the three features leads to better classification results compared to the other version of our method in which only mean vector is used as the feature of each object



**Table 1** The classification results after applying the proposed spectral-spatial classification method as well as the other competing approaches to the Pavia University dataset with 50 training samples per class

	Performance measures		
	OA	AA	Kappa
Object-based classification methods			
Proposed spectral-spatial classification approach	95.33%	94.23%	0.92
Three-band EMAP (with KPCA $\sigma = 1$ ) [Bernabe et al. 2014]	93.22%	-	-
RPDE-based Pixon extraction [Zehtabian and Ghassemian 2015]	92.80%	91.02%	0.90
SVM-MLR <sub>sub</sub> -MRF (M = 2) [Khodadadzadeh et al. 2014]	92.68%	91.93%	0.90
AMICA [Ghassemian and Landgrebe 1988]	91.03%	89.63%	0.87
SVM-MRF [Khodadadzadeh et al. 2014]	83.96%	91.80%	0.80
Traditional pixel-based SVM	77.85%	85.87%	0.72
MLR <sub>sub</sub> [Khodadadzadeh et al. 2014]	65.47%	73.81%	0.57

parameter, we assign a unique value to this parameter for all the spectral bands, at each step of the experiments. The gained results are then reported as a function of the varying Pixon extraction parameter.

In Fig. 8, the classification ratios are plotted along with various values for the Pixon extraction threshold that has been simultaneously varied in the segmentation steps of all bands. The pixel-based classification results (which are independent from the values of this threshold) are also illustrated in this figure. In this experiment, the values of the wavelet parameters are kept fixed to the optimum values (i.e.  $\tau_{(average)} = 120.40$  and  $N_{(average)} = 4$ ) that have been previously set for each band, using the Bayesshrink method. The Pavia University dataset is used in these experiments while the number of training samples is set to 10% of the available samples.

Meanwhile, Fig. 9 shows the relation between the value of the Pixon extraction threshold parameter and the time needed for object extraction as well as object-based classification.

From Fig. 9, the number of the extracted objects is directly proportional to the Pixon extraction parameter. As this parameter increases, the object-to-pixel ratio and consequently the time spent for object-based classification monotonically decrease. In other words, the level of data compactness increases while the Pixon extraction threshold gets larger. It is due to this fact that larger values of this threshold result in smaller number of objects with larger sizes. On the other hand, it may increase the possibility of under-segmentation.

By decreasing the value of the Pixon extraction threshold, the object-to-pixel ratio as well as the object-based classification time gets increased (Fig. 9). It is not surprising that for extensively small values of this parameter, the number of objects reaches its maximum possible value which is the number of pixels. In other words, for very small Pixon extraction thresholds, the objects are in size of pixels. Therefore, it can be deduced that the probability of over-segmentation increases for small thresholds.

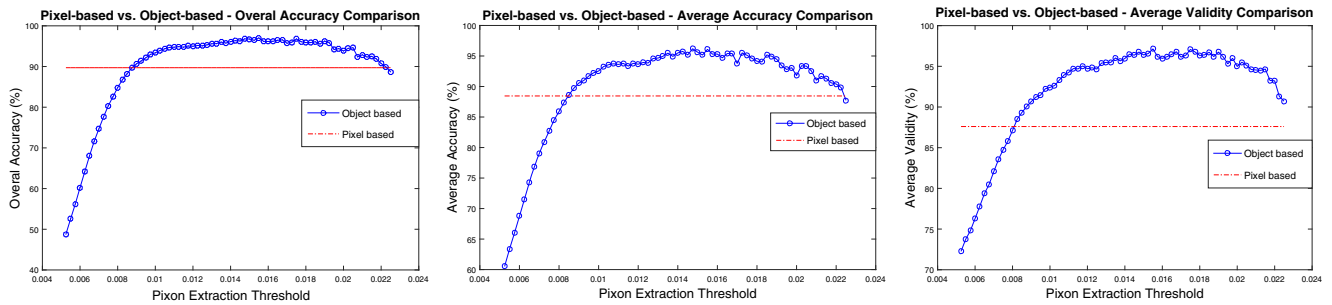
There is no meaningful relationship between the Pixon extraction threshold and the time needed for object extraction (Fig. 9). On the other hand, the time spent to classify the objects is much less compared with the time needed to classify the pixels. It is due to this fact that the number of extracted objects is considerably smaller than that of the pixels, especially for larger values of the Pixon extraction parameter.

While increasing the Pixon extraction threshold, the classification accuracies (i.e. OA, AA, and AV) first increase to reach their maximum levels and then decrease again (Fig. 8). This is due to the under-segmentation phenomenon that is caused by relatively large values of the Pixon extraction parameter.

In terms of various metrics, the proposed object-based classifier clearly exceeds the pixel-based classifier, at least for a relatively wide range of Pixon extraction threshold values (Fig. 8). However, this must be noted that these experiments have been carried out under specific conditions in which similar parameters

**Table 2** The classification results after applying the proposed spectral-spatial classification method as well as the other competing approaches to the Salinas dataset with 10% training samples per class

	Performance measures
	OA
Object-based classification methods	
MP-Gabor-GLCM [Mirzapour and Ghassemian 2015]	99.26%
Proposed spectral-spatial classification approach	99.04%
SVM-CK [Camps-Valls et al. 2006]	98.87%
HMRFE-SVM [Ghamisi et al. 2014]	97.25%
RPDE-based Pixon Extraction [Zehtabian and Ghassemian 2015]	96.90%
Diverse AdaBoost SVM [Ramzi et al. 2013]	94.87%
Traditional pixel-based SVM	85.37%



**Fig. 8** The effects of applying various values of Pixon extraction threshold on the proposed object-based classifier performance in terms of overall accuracy (OA), average accuracy (AA) and average validity (AV). The wavelet parameters are fixed at their most likely optimum

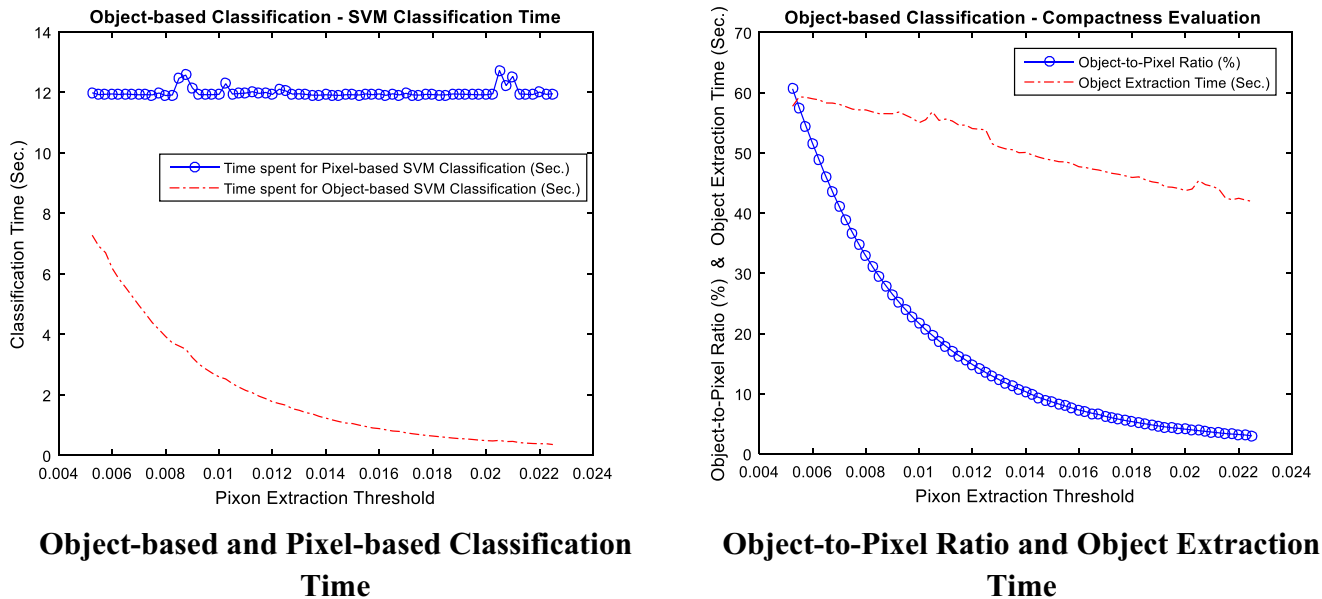
values. In all the figures, the *dotted red lines* denote the results associated with the traditional pixel-based method while the *blue curves* correspond to the object-based approach. The experiments are implemented on the Pavia University dataset

are used for all bands, simultaneously. Therefore, the classification improvement is more likely if the parameters are adaptively tuned for each band, separately, as suggested in the present article.

From Fig. 8, using a grid search over the supplied range for the Pixon extraction parameter, the best classification results in terms of overall accuracy, averaged accuracy, and overall validity are obtained with Pixon extraction thresholds equal to 0.0155, 0.0145 and 0.0155, respectively. These values are relatively close to the average of all adaptively tuned parameters (each for a single band), which is equal to 0.0159. This figure also shows that by setting the value of the Pixon extraction threshold to 0.015 for all spectral bands, an accurate classification (in terms of OA, AA and AV) is more likely. However, as discussed previously, if the parameters are individually tuned for each band, the classification results significantly increase.

**Conclusion**

In the present article, a spectral-spatial classification framework has been developed in which both pixel-based and object-based classification scenarios are utilized. In the proposed method, first a pixel-based classification using support vector machine is applied to the hyperspectral data. It results in a pixel-wise classification map. Using an unsupervised band selection technique that is based on Ward’s linkage strategy using mutual information (WaLuMI), a smaller number of bands are then selected from the spectral bands of the original hyperspectral data. The selected bands are then smoothed using wavelet thresholding in which the parameters are tuned using the Bayesshrink technique. The smoothed bands are fed into the segmentation step which is a modified version of our previously developed Pixon-based algorithm. To be more



**Fig. 9** The effects of applying various values of the Pixon extraction threshold on the time needed for object extraction as well as object-based classification. These experiments are implemented on the Pavia University dataset

clear, the Euclidian distance in our traditional Pixon extraction algorithm has been now substituted with an innovative distance metric. Since it simultaneously incorporates the spectral and textural information into a single measure, the proposed distance often leads to efficient segmentation results. The success of the proposed Pixon-based segmentation algorithm depends on properly tuning a parameter named Pixon extraction threshold. We suggest applying an adaptation technique which has been recently developed by us in order to automatically set proper values for this threshold.

Once image segmentation is performed, two different sets of classification maps are produced with regard to the extracted object-maps. The first set is achieved by performing majority voting inside each object. In other words, for each selected band, a class-map is obtained by assigning the pixels of each object to the most frequent class within that object, using the reference pixel-wise class-map. Moreover, the second set of classification maps is produced by labeling each object from the segmentation map to one of the classes using the spectral and spatial information extracted from each object. The final thematic map is then resulted after applying majority voting among all the available class-maps.

There are a few segmentation/classification approaches which have tried to automatically tune their parameters, however, one of the most important aspects of the proposed framework is that the most important parameters (i.e. the Pixon extraction threshold and the wavelet parameters) are automatically tuned. Therefore, it does not need to manually set the parameters.

The experimental results of applying various classification methods to two widely-used hyperspectral datasets prove the efficiency of the proposed spectral-spatial classification framework in terms of the classification ratios as well as the object to pixel ratio and the computational time.

## References

- Bernabe S, Marpu PR, Plaza A, Dalla Mura M, Benediktsson JA (2014) Spectral–spatial classification of multispectral images using kernel feature space representation. *IEEE Geosci Remote Sens Lett* 11(1): 288–292
- Blaschke T (2010) Object based image analysis for remote sensing. *ISPRS J Photogramm Remote Sens* 65:2–16
- Blaschke T, Hay GJ, Kelly M et al (2014) Geographic object based image analysis—towards a new paradigm. *ISPRS J Photogramm Remote Sens* 87:180–191
- Burrus CS, Gopinath RA, Guo H (1998) Introduction to wavelets and wavelet transforms. Prentice Hall, New Jersey
- Camps-Valls G, Gomez-Chova L, Munoz-Mari J, Vila-Frances J, Calpe-Maravilla J (2006) Composite kernels for hyperspectral image classification. *IEEE Geosci Remote Sens Lett* 3(1):93–97
- Cariou C, Chehdi K, Le Moan S (2011) BandClust: an unsupervised band reduction method for hyperspectral remote sensing. *IEEE Geoscience and Remote Sensing Letters* 8(3):565–569
- Chang SG, Yu B, Vetterli M (2000) Adaptive wavelet thresholding for image Denoising and compression. *IEEE Trans Image Processing* 9: 1532–1545
- Chang CC et al (2001) LIBSVM: a library for support vector machines, Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- Fauvel M, Tarabalka Y, Benediktsson JA, Chanussot J, Tilton JC (2013) Advances in spectral-spatial classification of hyperspectral images. *Proc IEEE* 101(3):652–675
- Gaetano R, Masi G, Poggi G (2015) Marker-controlled watershed-based segmentation of Multiresolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 53(6): 2987–3004
- Ghamisi P, Benediktsson JA, Cavallaro G, Plaza A (2014a) Automatic framework for spectral–spatial classification based on supervised feature extraction and morphological attribute profiles. *Selected Topics in Applied Earth Observations and Remote Sensing* 7(6): 2147–2160
- Ghamisi P, Couceiro MS, Fauvel M, Benediktsson JA (2014b) Integration of segmentation techniques for classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters* 11(1):342–346
- Ghamisi P, Dalla Mura M, Benediktsson JA (2015) A survey on spectral-spatial classification techniques based on attribute profiles. *IEEE Transactions on Geoscience and Remote Sensing* 53(5):2335–2353
- Ghassemian H, Landgrebe DA (1988) Object-oriented feature extraction method for image data compaction. *IEEE Cont Syst Mag* 8(3):42–48
- Golipour M, Ghassemian H, Mirzapour F (2016) Integrating hierarchical segmentation maps with MRF prior for classification of hyperspectral images in a Bayesian framework. *IEEE Transactions on Geoscience and Remote Sensing* 54(2):805–816. doi:10.1109/TGRS.2015.2466657
- Gupta S, Kaur L (2002) Wavelet based image compression using daubechies filters. 8th National conference on communications, I.I.T. Bombay 88–92
- Hassanpour H, Yousefian H, Zehtabian A (2011) Pixon-based image segmentation. In: Pei-Gee Ho (ed) Image segmentation. InTech pub., Rijeka, p 496–516
- Holbling D, Friedl B, Eisank C (2015) An object-based approach for semi-automated landslide change detection and attribution of changes to landslide classes in northern Taiwan. *Earth Sciences Informatics* 8(2):327–335. doi:10.1007/s12145-015-0217-3
- Kang X, Li S, Benediktsson JA (2014) Spectral-spatial hyperspectral image classification with edge-preserving filtering. *IEEE Transactions on Geoscience and Remote Sensing* 52(5):2666–2677
- Khodadadzadeh M, Li J, Plaza A, Ghassemian H, Bioucas-Dias JM, Li X (2014) Spectral–spatial classification of hyperspectral data using local and global probabilities for mixed pixel characterization. *IEEE Transactions on Geoscience and Remote Sensing* 52(10): 6298–6314
- Li J, Zhang H, Zhang L (2014) Supervised segmentation of very high resolution images by the use of extended morphological attribute profiles and a sparse transform. *IEEE Geo. Sci. & Remote Sensing Letters* 11(8):1409–1413
- Liu D, Xia F (2010) Assessing object-based classification: advantages and limitations. *Remote Sensing Letters* 1(4):187–194
- Lu D, Weng Q (2007) A survey of image classification methods and techniques for improving classification performance., *Int. jour. Of. Remote Sens* 28(5):823–870
- Machala M, Zejdova L (2014) Forest mapping through object-based image analysis of multispectral and Lidar Aerial data. *European Journal of Remote Sensing* 47:117–131
- Martinez-Uso A, Pla F, Sotoca JM, Garcia-Sevilla P (2007) Clustering-based hyperspectral band selection using

- information measures. *IEEE Trans Geosci Remote Sens* 45(12):4158–4171
- Mirzapour F, Ghassemian H (2015) Improving hyperspectral image classification by combining spectral, texture, and shape features. *Int Jour Remote Sens* 36(4):1070–1096
- Mylonas SK, Stavrakoudis DG, Theocharis JB, Mastorocostas PA (2015) Classification of remotely sensed images using the GeneSIS fuzzy segmentation algorithm. *IEEE Transactions on Geoscience and Remote Sensing* 53(10):5352–5376
- Plaza A, Benediktsson JA, Boardman JW et al (2009) Recent advances in techniques for hyperspectral image processing. *Remote Sens of Environment* 113:110–122
- Ramzi P, Samadzadegan F, Reinartz P (2013) Classification of hyperspectral data using an AdaBoostSVM technique applied on band clusters. *IEEE Jour of Selected Topics in Applied Earth Observations and Remote Sens* 7(6):2066–2079
- Rubner Y, Tomasi C, Guibas LJ (2000) The earth Mover's distance as a metric for image retrieval. *Int Jour of Comp Vision* 40(2):99–121
- Samal DR, Gedam SS (2015) Monitoring land use changes associated with urbanization: an object based image analysis approach. *European Journal of Remote Sensing* 48:85–99
- Shi W, Mao Z (2016) Building extraction from panchromatic high-resolution remotely sensed imagery based on potential histogram and neighborhood Total variation. *Earth Sciences Informatics*. doi:10.1007/s12145-016-0262-6
- Sun W, Li W, Li J, Lai YM (2015) Band selection using sparse nonnegative matrix factorization with the thresholded Earth's mover distance for hyperspectral imagery classification 8(4):907–918. doi:10.1007/s12145-014-0201-3
- Ward JH (1963) Hierarchical grouping to optimize an objective function. *Amer Stat Assoc* 58(301):236–244
- Zahidi I, Yusuf B, Hamedianfar A, Shafri HZM, Mohamed TA (2015) Object-based classification of QuickBird image and low point density LIDAR for tropical trees and shrubs mapping. *European Journal of Remote Sensing* 48:423–446
- Zehtabian A, Ghassemian H (2015) An adaptive Pixton extraction technique for multispectral/hyperspectral image classification. *IEEE Geo. Sci. & Remote Sensing Letters* 12(4):831–835
- Zehtabian A, Nazari A, Ghassemian H, Gribaudo M (2015) Adaptive restoration of multispectral datasets. *The European Journal of Remote Sensing* 48:183–200
- Zhang P, Lv Z, Shi W (2013) Object-based spatial feature for classification of very high resolution remote sensing images. *IEEE Geo Sci & Remote Sensing Letters* 10(6):1572–1576