

Supervised Hyperspectral Image Classification Based on Spectral Unmixing and Geometrical Features

Bin Luo · Jocelyn Chanussot

Received: 30 January 2010 / Revised: 1 May 2010 / Accepted: 7 September 2010 / Published online: 22 October 2010
© Springer Science+Business Media, LLC 2010

Abstract The spectral features of hyperspectral images, such as the spectrum at each pixel or the abundance maps of the endmembers, describe the material attributes of the structures. However, the spectrum on each pixel, which usually has hundreds of spectral bands, is redundant for classification task. In this paper, we firstly use spectral unmixing to reduce the dimensionality of the hyperspectral data in order to compute the abundance maps of the endmembers, since the number of endmembers in an image is much less than the number of the spectral bands. In addition, using only the spectral information, it is difficult to distinguish some classes. Moreover, it is impossible to separate objects made by the same material but with different semantic meanings. Some geometrical features are needed to separate such spectrally similar classes. In this paper, we introduce a new geometrical feature—the characteristic scales of structures—for the classification of hyperspectral images. With the help of the abundance maps obtained by spectral unmixing, we propose a method based on topographic map of images to estimate local scales of structures in hyperspectral images. The experiments show that using geometrical features actually improves the classification results, especially for the classes made by the same material but with different semantic meanings. When compared to the traditional contextual features (such as morpholog-

ical profiles), the local scale provides satisfactory results without significantly increasing the feature dimension.

Keywords Supervised classification · Hyperspectral image · Spectral unmixing · Geometrical feature · Scale

1 Introduction

Hyperspectral remote sensing images provide rich information on the ground with high spatial and spectral resolutions. With the help of hundreds of spectral bands, the ability to detect and identify individual materials greatly improves. However, there are two major difficulties if the spectral bands are directly used to classify hyperspectral images.

The first difficulty is the high dimension of hyperspectral data. Hyperspectral images usually contain hundreds of spectral bands, which considerably increase the complexity of the classifier. In addition, some bands may be severely corrupted by noise or artifact, which will introduce errors of classification. In [9], it has been shown that the reduction of data dimensionality is essential as a preprocessing. For this purpose, both supervised and non supervised methods are proposed. The supervised methods, such as band selection [7, 20], or Decision Boundary Feature Extraction and Non-Weighted Feature Extraction [8], transform the data according to the training set in order to improve the separability of the data. However, the supervised methods depend on the quality of the training set. The unsupervised methods, such as PCA (principle component analysis) or ICA (independent component analysis), optimize some statistical criteria (such as the most

B. Luo (✉) · J. Chanussot
Department of Image and Signal, GIPSA-Lab,
Grenoble, France
e-mail: bin.luo@gipsa-lab.inpg.fr

J. Chanussot
e-mail: jocelyn.chanussot@gipsa-lab.inpg.fr

uncorrelated components or the most independent components) to project the data onto a sub-space with a lower dimension. The application of these methods on hyperspectral data can be found in [5, 6, 16]. However, the components obtained by optimizing the statistical criteria do not necessarily have physical meanings. Moreover, the values of the components extracted by the PCA or ICA are not comparable, which leads difficulties when we compute the geometrical features based on these components. This will be discussed more in detail in Section 3.2. In this paper, we propose to use Vertex Component Analysis (VCA) to reduce the dimension of hyperspectral data [15]. We suppose that the spectrum of each pixel is a linear mixture of the spectra of different chemical species (referred to as endmembers). This linear mixture model is physically valid for the reflectance of the surface of the Earth without being affected by the aerosol. The VCA can separate the spectra of these endmembers and estimate their spatial abundances. Since the number of endmembers is much lower than the number of spectral bands, these abundance maps can be considered as a compact representation of the spectral information provided by the hyperspectral image. For the linear unmixing, the VCA is retained because of its low computation cost. The unmixing of a ROSIS image with 300×600 pixels and with 103 bands takes only a few seconds. In addition, the choice of linear unmixing approach is not crucial for the classification. In practice, we found that if the number of endmembers is the same, the spectra and the abundance maps of the endmembers are quite similar.

Secondly, the spectral information alone sometimes does not allow the separation of different structures. For example, the roofs of some buildings and the roads can be made by the same material (asphalt). Moreover, some classes have very similar spectral property, such as the shadows and the water. Therefore, contextual information, is necessary to improve the classification performance. This can be achieved by extracting geometrical features. In order to extract the contextual information from remote sensing images, one can find the methods based on Markov Random Field (MRF) [19]. However, the MRF based methods provide only statistical information on the neighborhood of the considered pixel. Another family of methods for contextual information extraction is based on morphological operators, which allow to extract descriptive features, such as the geometrical features about the structures [17]. Fauvel et al. [5, 6] extract the extended morphological profiles (EMP) of the principle components of hyperspectral images for the classification. It has been shown that the morphological profiles are related both to the

size and the contrast of the structures [4]. However, since morphological methods usually require a structuring element, the features extracted by such methods depend on the used structuring element. Moreover, in order to describe the contextual information of the structures with different scales, one has to use the structuring element with a long discrete scale range. This fact has two major drawbacks. Firstly, the scale range has to be chosen manually and is not necessarily well suited to the image. Secondly, it increases considerably the feature dimension. For example, in order to describe the geometrical information of one pixel on a principle component, the EMP method in [5, 6] uses eight values. In [10], based on topographic map of gray scale image, the authors define a characteristic scale of each pixel in panchromatic remote sensing images. The main idea is that for each pixel, from all the structures containing it, we extract the structure of which the contrast is the most important; and the scale of this structure defines the scale of this pixel. With the help of the abundance maps obtained by VCA, we try to extend the algorithm presented in [10] in order to extract a local characteristic scale for each pixel on hyperspectral images in this paper. The main advantages of our proposed method when compared to EMP are two-fold. On one hand, no structuring element is required for this method. Secondly, we use only one value (the local scale) to describe the geometrical feature of a spatial position rather than the EMP which needs many values. Preliminary results presented in [11] shows that the integration of the scale feature can improve the classification accuracy of hyperspectral image over urban areas.

Support Vector Machine (SVM) is proven to be very efficient for the classification of remote-sensing images [13]. In particular, [9] shows that the SVMs are very suited for the classification of hyperspectral data when compared to other supervised classifiers (such as neural networks, maximum likelihood, linear discriminant analysis, etc.). Therefore, in this paper, SVM is used to classify the spectral information obtained by the VCA and the geometrical feature extracted on hyperspectral images.

The outline of this paper is illustrated by Fig. 1. In Section 2, we introduce very briefly the Vertex Component Analysis (VCA) for estimating the abundance maps of endmembers in order to reduce the dimension of hyperspectral data. In Section 3, we introduce the method presented in [10] for extracting the local scale of panchromatic images and extend this method to hyperspectral images. In Section 4, we classify a hyperspectral image by using both the spectral features and the scale feature in order to show the efficiency of the local scale feature.

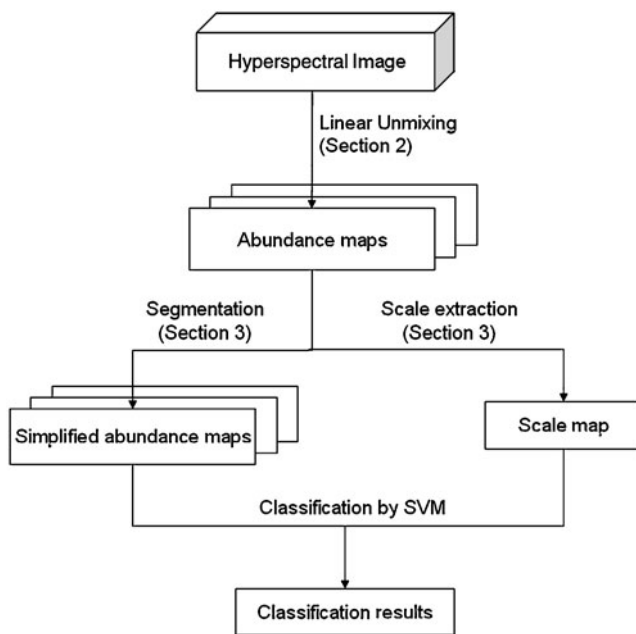


Figure 1 Scheme of the paper.

2 Linear Unmixing of Hyperspectral Data

We note \mathbf{X} the matrix representing the hyperspectral image cube, where $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_a}\}$ (N_a is the total number of pixels in the image) and $\mathbf{x}_k = \{x_{1,k}, x_{2,k}, \dots, x_{N_s,k}\}^T$, $x_{l,k}$ is the value of the k th pixel at the l th spectral band. We assume that the spectrum of each pixel is a linear mixture of the spectra of N_c endmembers, leading to the following model:

$$\mathbf{X} = \mathbf{M}\mathbf{S} + \mathbf{N} \tag{1}$$

where $\mathbf{M} = \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{N_c}\}$ is the mixing matrix, where \mathbf{m}_n denotes the spectrum of the n th endmember. $\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{N_c}\}^T$ is the abundance matrix where $\mathbf{s}_n = \{s_{n,1}, s_{n,2}, \dots, s_{n,N_a}\}$ ($s_{n,k} \in [0, 1]$) is the abundance of the n th endmember at the k th pixel). \mathbf{N} stands for an additive noise of the image.

In order to extract the spectra of the endmembers and to compute their abundance maps, linear unmixing approaches are proposed, such as the N-Finder, PPI and VCA. N-Finder, proposed in [21], attempts to find the simplex with the maximal volume which contains all the hyperspectral data set. N-Finder starts from a set of candidates which are randomly selected. Therefore the results of the N-Finder method depend on the initialization [18]. The Pure Pixel Index (PPI), proposed in [1], generates a large number of random unit vectors through the data set. Every pixel is then projected on these unit vectors. The pixels which represent the extrema of the projections have scores. And the pixels

with highest scores are considered as the pure pixel of endmembers. However, the results of this method also depends on the initialization and the computation if time consuming. In addition, as been discussed in the Section 1, the results of different linear unmixing methods are quite similar for the purpose of classification. Therefore, in this paper, we retain the method VCA which is very efficient and the results do not depend on the initialization.

In [15], the Vertex Component Analysis (VCA) is proposed as an efficient method for extracting the endmembers which are linearly mixed. The main idea is to extract the vertex of the simplex formed by \mathbf{M} which contains all the data vectors in \mathbf{X} . The sum of the abundances of the different endmembers at each pixel is one, i.e. $\sum_{n=1}^{N_c} s_{n,k} = 1$, which is called the *sum-to-one* condition. Therefore the data vectors \mathbf{x}_l are always inside a simplex of which the vertex are the spectra of the endmembers. VCA iteratively projects the data onto the direction orthogonal to the subspace spanned by the already determined endmembers. And the vertices corresponds to the projection of the endmember spectra. The algorithm stops when the requested number p of endmembers are extracted. Even though the number of endmembers is much smaller than the number of spectral bands, i.e. $p \ll N_s$, abundance maps of the p endmembers can provide the same spectral information as the original hyperspectral image with N_s spectral bands. Therefore this linear unmixing step can be considered as a dimension reduction of the data. Moreover, the values of the abundance maps for one given pixel, which represent the proportions of different physical components within this pixel, are comparable. This is an essential issue for estimating the characteristic scale of this pixel. On the contrary, the values obtained for one pixel by other dimension reduction methods (such as PCA or ICA) are not comparable. In the next section, we demonstrate that the comparability of the proportions of different components obtained by VCA is very important for extracting the local scale of a structure.

3 Local Characteristic Scale for Hyperspectral Image

3.1 Local Scales for Panchromatic Images

In [10], the authors propose a method based on the topographic map of the image to estimate the local scale of each pixel in the case of gray scale remote sensing images. The idea is that, for each pixel, the most contrasted shape containing it is extracted, and the scale of this shape defines the characteristic scale of this pixel. The topographic map [3], which can be obtained

by Fast Level Set Transformation (FLST) [14], represents an image by an inclusion tree of the shapes (which are defined as the connected components of the level sets). An example of such inclusion tree is shown in Fig. 2. For each pixel (x, y) , there is a branch of shapes $f_i(x, y)$ ($f_{i-1} \subset f_i$) containing it. Note $I(f_i)$ the gray level of the shape $f_i(x, y)$, $S(f_i)$ its area and $P(f_i)$ its perimeter. The contrast of the shape $f_i(x, y)$ is defined as $C(f_i) = |I(f_{i+1}) - I(f_i)|$. The most contrasted shape $f_i(x, y)$ of a given pixel is defined as the shape containing this pixel, of which the contrast is the most important, i.e.

$$f_i(x, y) = \arg \max_j \{C(f_j(x, y))\} \tag{2}$$

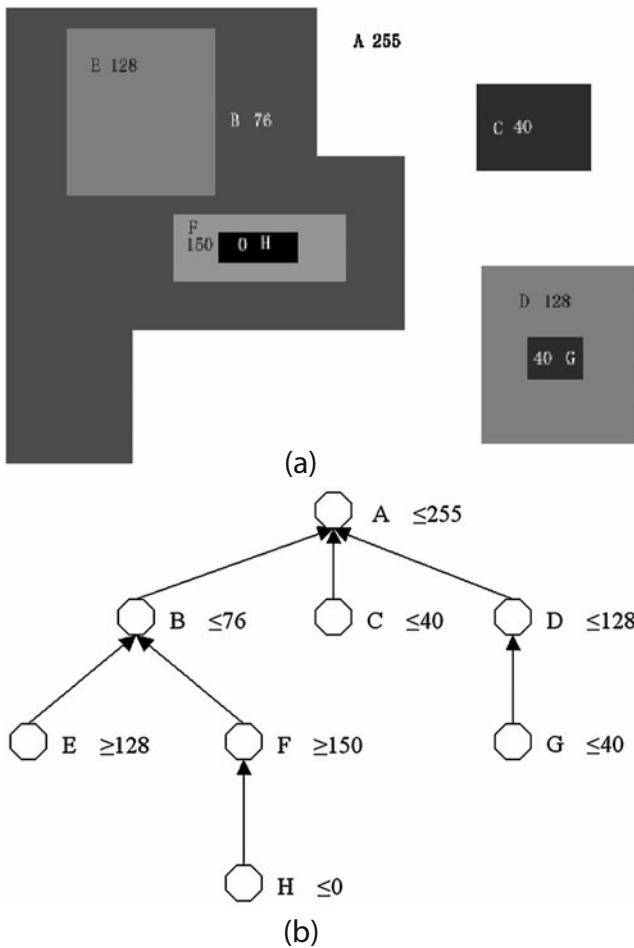


Figure 2 Example of FLST: **a** Synthetic image; **b** Inclusion tree obtained with FLST. All the nodes correspond to the shapes of constant values in the image. The leaves of the tree are the smallest shapes with constant levels. For example, for the pixels in Shape H, the smallest shape containing them is Shape H. As we can see H is contained by F, therefore the parent of H is F. F is included in B; the parent of F is B. And Shape A represents the whole image; A is therefore the root of the tree.

Since the optical instruments always blur the data, several shapes with very low contrasts can belong to the same structure. In order to deal with the blur, the authors of [10] propose a geometrical criterion to cumulate the contrasts of the shapes corresponding to one given structure. The idea is that the difference of the areas of two successive shapes (for example f_i and f_{i+1}) corresponding to one given structure is proportional to the perimeter of the smaller shape, i.e.

$$S(f_{i+1}) - S(f_i) \sim \lambda P(f_i), \tag{3}$$

where λ is a constant, which will be detailed later in this section. This criterion is based on the fact that the boundaries of two successive shapes (f_i and f_{i+1} , and $f_i \in f_{i+1}$) corresponding to the same structure caused by blur can be approximated by two parallel curves. If f_i is convex (which indicates that f_{i+1} is also convex), the area of f_{i+1} is then

$$S(f_{i+1}) = S(f_i) + \lambda P(f_i) + \pi \lambda^2, \tag{4}$$

where λ is the distance between the boundaries of f_i and f_{i+1} . It depends on the quantization of the image and the PSF (Point Spread Function) of the sensor. It is shown in [10] that for most of remote sensing images, λ can be fixed to 1. Therefore if we ignore the term $\pi \lambda^2$ (which is ignorable when compared to the area of the shapes), and if the area of two successive shapes f_i and f_{i+1} satisfy the criterion of Eq. 3, we consider that these two shapes correspond to the same structure and then cumulate the corresponding contrasts. We thus define the cumulated contrast of a shape f_i as:

$$\bar{C}(f_i) = \sum_{k=a(i)}^i C(f_k), \tag{5}$$

where, for all i ,

$$a(i) = \min\{j | \forall k = j + 1, \dots, i, S(f_k) - S(f_{k-1}) \leq \lambda P(f_{k-1})\}.$$

If $a(i)$ is not defined (that is if Eq. 3 is not satisfied), then $\bar{C}(f_i) = C(f_i)$.

Finally the definition of the most contrasted shape and of the local scale are no longer based on the cumulated contrast, i.e.

$$f_i(x, y) = \arg \max_j \{\bar{C}(f_j(x, y))\}. \tag{6}$$

And the scale of this pixel $E(x, y)$ is defined as the area of the most contrasted shape divided by its perimeter, i.e.

$$E(x, y) = S(f_i(x, y)) / P(f_i(x, y)), \tag{7}$$

so that the geometry of $f_i(x, y)$ is taken into account. In particular, long and thin shapes (e.g., the roads) correspond to relatively small scales, even though their area can be quite large.

It is shown in [10] that this local scale corresponds very well to the size of a structure and it is a very significant feature for characterizing a structure in remote sensing images. Remark that the most contrasted shapes extracted in an image form a partition of this image. We can therefore have a simplified image by defining the value of each pixel as the mean value of the gray levels of the pixels in its most contrasted shape, i.e. the value of (x, y) in the simplified image is:

$$\bar{I}(x, y) = \frac{\sum_{(k,l) \in f_i(x,y)} I(k, l)}{S(f_i(x, y))} \tag{8}$$

where $I(k, l)$ is the gray level of the pixel (k, l) in the original image. This value is more significant for characterizing spectrally the pixel than its original gray level since it takes into account the contextual information.

In practice, the algorithm for computing the local scales is implemented by using MEGAWAVE.¹ The computation time for computing the local scales of an image with 340×610 pixels is 5.61 m on a workstation with 2 Xeon processor of 3.33 GHz and 16 GB RAM.

3.2 Extension to Hyperspectral Images

In order to extend the estimation of local scale to hyperspectral images, the scales on all the abundance maps of the p endmembers obtained by VCA are first computed. Therefore, for each spatial position (x, y) in an hyperspectral image, there are p scale values. For the pixel (x, y) on the n th abundance map, we note $E_n(x, y)$ its local scale calculated by Eq. 7, $f_{i,n}(x, y)$ the most contrasted shape extracted by Eq. 6, and $\bar{I}_n(x, y)$ its value of the simplified image defined by Eq. 8. The simplest way to use the scale features of a pixel is to use all the $E_n(x, y)$ values as features for classification. However, the major drawback is that if an object is mainly made by the n th endmember, the scale values calculated on the abundance maps of the other endmembers for this object have no meaning. Therefore, we try to define one single scale value for each pixel which corresponds to the scale of this pixel on the most significant abundance map. More precisely, since the values of different abundance maps are comparable, the charac-

teristic scale at the spatial position (x, y) for a hyperspectral image is defined as

$$\hat{E}(x, y) = E_{\hat{n}}(x, y) \tag{9}$$

where $\hat{n} = \arg \max_n \{\bar{C}(f_{i,n}(x, y))\}$.

For the classification, we define two feature vectors, namely $\Theta_1(x, y)$ and $\Theta_2(x, y)$. For one given pixel, $\Theta_1(x, y)$ contains the values of the simplified image of all the abundance maps and $\Theta_2(x, y)$ is defined as $\Theta_1(x, y)$ concatenated with the local scale (x, y) defined by Eq. 9, i.e.

$$\Theta_1(x, y) = \{\bar{I}_1(x, y), \dots, \bar{I}_{N_c}(x, y)\}, \tag{10}$$

$$\Theta_2(x, y) = \{\bar{I}_1(x, y), \dots, \bar{I}_{N_c}(x, y), \hat{E}(x, y)\}. \tag{11}$$

The feature vector Θ_1 can be considered as a compact and regularized representation of the spectral information of the image. And feature vector Θ_2 contains the same spectral information but also one geometrical attribute.

It has to be remarked that this extension can only be applied on the abundance maps derived from the hyperspectral images by using linear unmixing approaches. The extraction of geometrical features on the components extracted by the PCA or the ICA is impossible, since the values of these components are not comparable.

4 Experiments

In this section, we will use the feature vectors Θ_1 and Θ_2 to classify two hyperspectral images captured respectively by ROSIS and HYDICE instruments. For comparison, the classification results are compared with the results obtained by using the Extended Morphological Profiles (EMP) computed on the components extracted by the PCA and the KPCA (Kernel PCA) [6]. The results obtained by using EMP computed on the abundance maps extracted by the VCA are also shown. The global classification results are evaluated by using the Overall Accuracy (OA). The OA computed for the class c is defined as:

$$OA_c = \frac{\#\{\mathbf{x} | L(\mathbf{x}) = \hat{L}(\mathbf{x}); \hat{L}(\mathbf{x}) = c\}}{\#\{\mathbf{x} | \hat{L}(\mathbf{x}) = c\}}, \tag{12}$$

where $L(\mathbf{x})$ is the class assigned by the classifier for the pixel \mathbf{x} , $\hat{L}(\mathbf{x})$ is the real class of the pixel \mathbf{x} . The OA computed for all the image is defined as:

$$OA = \sum_c \frac{\#\{\mathbf{x} | L(\mathbf{x}) = \hat{L}(\mathbf{x}); \hat{L}(\mathbf{x}) = c\}}{\#\{\mathbf{x} | \hat{L}(\mathbf{x}) = c\}}. \tag{13}$$

¹<http://megawave.cmla.ens-cachan.fr/>

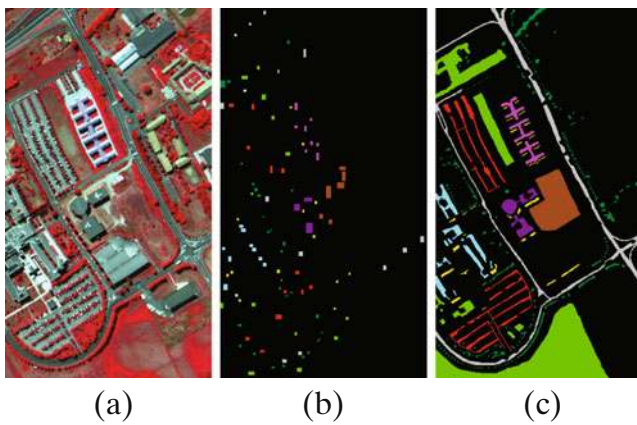


Figure 3 **a** Image of Pavia University (R-band 90, G-band 60, B-band 40); **b** Training set; **c** Test set.

4.1 Results on ROSIS Data

In this section, we classify a hyperspectral image taken by the instrument ROSIS (Reflective Optics System Imaging Spectrometer) over the University of Pavia, Italy (see Fig. 3a). The image (with a spatial resolution of 1.3 m) contains 340×610 pixels and 103 spectral bands covering visible and near infrared light. The image is manually classified into nine classes and the definitions of these classes are shown in Fig. 4. For classification purpose, the training set contains 3,921 pixels (see Fig. 3b) while the test set contains 4,2776 pixels (see Fig. 3c).

According to the experiment in [5], a sub-space of dimension 3 is necessary to retain 99% of the variance of the data. In addition, based on observation, there are mainly three endmembers present (vegetation, bare soil and metal roof). We extract hence three endmembers and their abundance maps by using VCA (see Section 2), which are shown in Fig. 5a–c. Afterwards we compute the simplified images of these abundance maps by using Eq. 8. The simplified images are shown in Fig. 5d–f.

According to Eq. 9, we estimate the local scale for each pixel of this image. The scale map is shown in Fig. 6a. Therefore the feature vector $\Theta_2(x, y)$ of a pixel

Class	Training Set	Test set	Thematic Colour
1. Trees	524	3064	
2. Asphalt	548	6631	
3. Bitumen	375	1330	
4. Gravel	392	2099	
5. (painted) metal sheets	265	1345	
6. Shadow	231	947	
7. Self-Blocking Bricks	514	3682	
8. Meadows	540	18649	
9. Bare Soil	532	3029	

Figure 4 Definition of the classes.

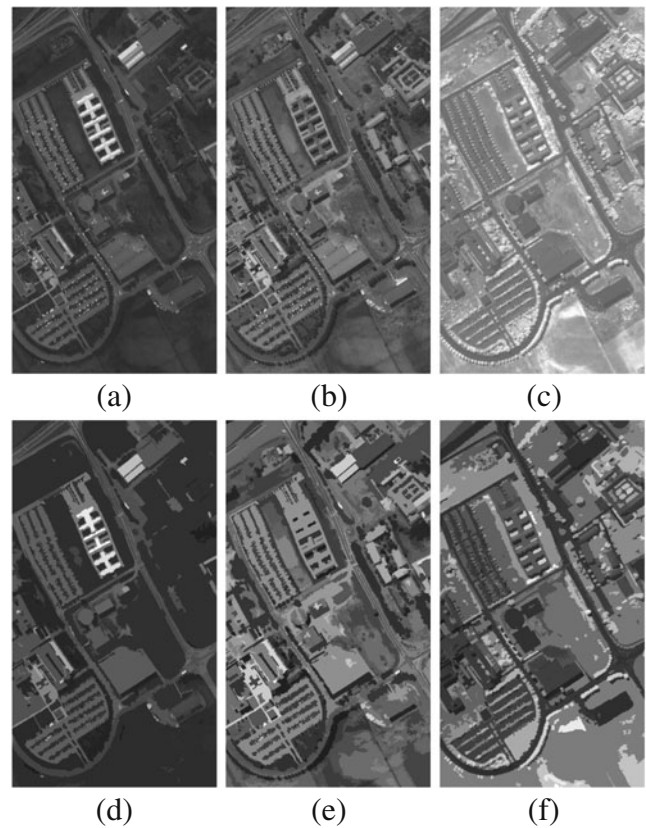


Figure 5 **a–c** Abundance maps of the 3 endmembers extracted by VCA; **d–f** the corresponding simplified images.

(x, y) defined by Eq. 11 contains only four values: the three values of the simplified abundance maps ($\bar{I}_1(x, y), \bar{I}_2(x, y), \bar{I}_3(x, y)$) and one scale value ($\hat{E}(x, y)$).

We have classified the image by using the original hyperspectral data (all 103 bands), the feature vectors Θ_1 and Θ_2 defined by Eqs. 10 and 11 respectively. Kernel methods are proved to be efficient algorithms for hyperspectral image classification [2]. Therefore, we use the Support Vector Machine (SVM) with Gaussian kernel as classifier. The optimal scale parameter of Gaussian kernel is selected by 5-fold cross validation on the training set.

The overall accuracies and the classification accuracies of each class for both cases are shown in Table 1. In [6], the authors proposed to use PCA and Kernel PCA (KPCA) to reduce the dimension of hyperspectral images. Extended Morphological Profiles (EMP) are then extracted from the principle components obtained by PCA (three principle components are extracted) or KPCA (12 principle components are extracted) as features for classifying the hyperspectral image. In addition, we have computed the EMP on the abundance maps extracted by VCA. For each abundance map, we have computed the EMP at four scales, for each scale,

Table 1 Classification results on ROSIS data.

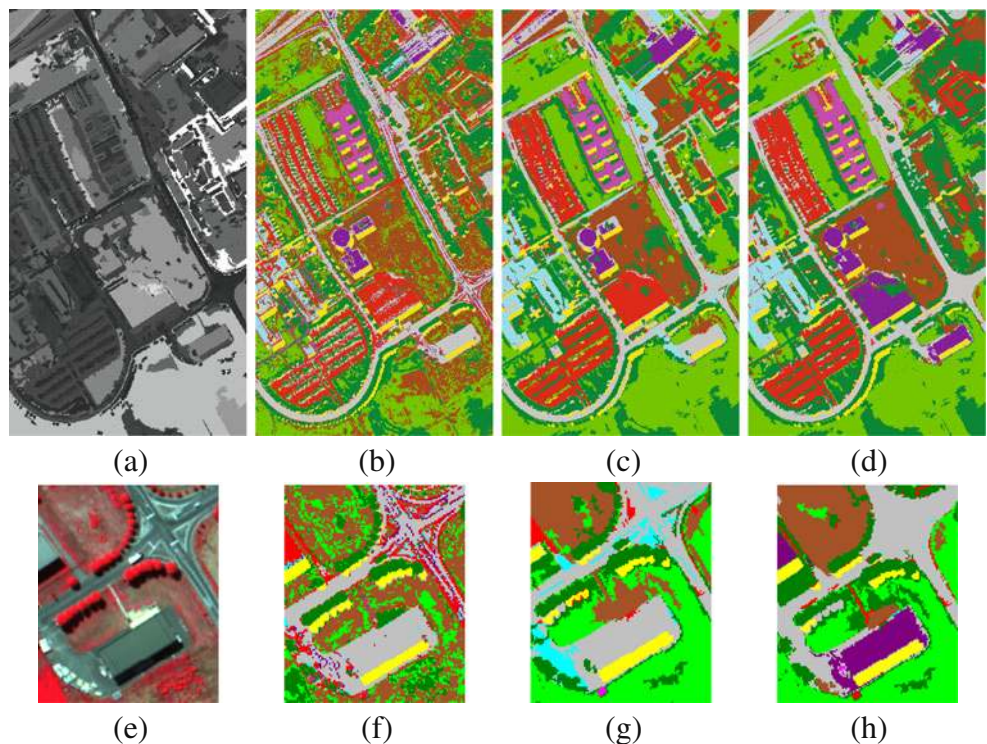
Feature	Original data	Feature vector Θ_1	Feature vector Θ_2	EMPPCA [6]	EMPKPCA [6]	EMPVCA
Number of features	103	3	4	27	108	27
Overall accuracy	76.01%	88.60%	91.54%	92.04%	96.55%	83.44%
Tree	98.59%	93.80%	94.52%	99.22%	99.35%	98.86%
Asphalt	78.16%	88.31%	96.27%	94.60%	96.23%	89.55%
Bitumen	89.02%	83.61%	99.32%	98.87%	99.10%	92.48%
Gravel	64.55%	95.43%	84.61%	73.13%	83.66%	71.84%
(Painted) metal sheets	99.47%	93.61%	99.55%	99.55%	99.48%	99.55%
Shadow	99.89%	98.10%	96.30%	90.07%	98.31%	94.93%
Self-blocking bricks	91.01%	99.48%	99.70%	99.10%	99.46%	99.02%
Meadows	64.23%	87.77%	85.80%	88.79%	97.58%	80.11%
Bare soil	82.72%	78.37%	96.56%	95.23%	92.88%	62.95%

there are two values of EMP which respectively correspond to the profiles of the two complementary morphological operators. The original abundance value of each pixel is used for the classification, too. Therefore, the feature dimension of the EMP extracted on the abundance maps of the VCA is $(4 \times 2 + 1) \times 3 = 27$. In Table 1, we show the classification results obtained in [6] on the same data set by using SVM with Gaussian kernel.

In Fig. 6b–d, the classification results obtained on the whole image are shown. It can be seen that by using Θ_2 (which contains not only the spectral information but also the geometrical information), the classification

results improve considerably when compared to the results obtained using the original hyperspectral data only. Recall that the length of the feature vector Θ_2 for each pixel is only 4 which is much less than the number of spectral bands (103). Moreover, the results obtained by using PCA and EMP are very similar with the results obtained by the features proposed, since we have previously mentioned that the major information described by EMP on a pixel is its local scale. Indeed, the local scale of a pixel can be considered as the width of the morphological filter by which the absolute differential EMP value reaches its maximum [4, 10]. However, the number of features extracted by PCA

Figure 6 **a** Scale map of the ROSIS hyperspectral image, bright pixels correspond to large scale while dark pixels correspond to small scales; **b** Supervised classification results obtained on original data set; **c, d** Supervised classification results obtained by using the feature sets Θ_1 (defined by Eq. 10) and Θ_2 (defined by Eq. 11); **e–h**: Zoom around a building on the: **e** original image; **f** classification results of **b**; **g** classification results of **c**; **h** classification results of **d**.



and EMP is 27, which is much larger than the number of features in Θ_2 . Even though the results obtained by using KPCA and EMP are better than the results obtained by using Θ_2 , the feature dimension, which is 108, is even higher than the dimension of the original data (103). And the improvement of classification accuracies is mainly concentrated on the class of **Meadows**, which is not well defined in the ground truth. In contrary, the classification results obtained by the EMP computed on the abundance maps extracted by the VCA are worse than using Θ_2 , even though the feature dimension of the EMP is much larger (27).

One should remark that in the two cases where only spectral information is used (i.e. the original spectrum of each pixel and the feature vector Θ_1), the accuracies of the classes **Asphalt** and **Bitumen** are relatively low, since they are made by the same material. According to the ground truth of Fig. 3c, the only difference is that **Asphalt** is used for the roads while **Bitumen** is used for the building roofs. However, it can be seen from Fig. 6a that the scale values of the pixels on the building roofs are very different from the scales of the pixels on the roads. Therefore the classification results of these two classes by using the feature vectors Θ_2 are much more accurate. This phenomenon can be illustrated by Fig. 6e–h, where we have zoomed the classification results obtained by using the original data, and the feature vectors Θ_1 and Θ_2 on a building made by bitumen which is surrounded by a road. It can be seen that the building roof and the road are classified as the same class in Fig. 6f and g. In contrary, by adding the local scale feature, these two structures are well separated (see Fig. 6h).

4.2 Results on HYDICE Data

The HYDICE sensor (Hyperspectral Digital Imagery Collection Experiment) has collected hyperspectral data from Washington DC Mall covering 210 spectral bands in the range 0.4–2.4 μm (see Fig. 7a). The spatial resolution of this image is approximately 3 m. For these data, we have discarded 67 spectral bands which contain either zero value or only noise due to water absorption. The image contains $1,280 \times 307$ pixels, of which 280 pixels are selected as training set and 6,929 pixels are selected as test set. These pixels are manually classified into seven classes (see Fig. 7b and c and Table 2).

For HYDICE data, we use the approach proposed [12] for determining the number of significant endmembers, which is 5. The authors of [12] suppose that the differences of the eigenvalues of the covariance and the correlation matrix are distributed by a Gaussian

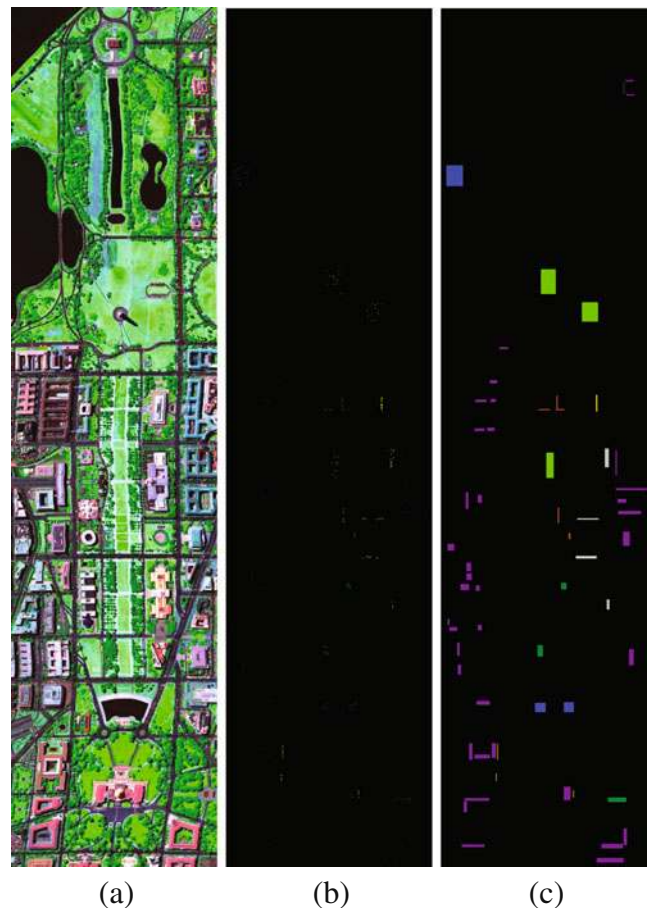


Figure 7 **a** HYDICE data on Washington (R-band 30, G-band 90, B-band 150); **b** Training set; **c** Test set.

distribution centered at zero, if these eigenvalues correspond to noise. If the eigenvalues do not correspond to noise, the differences of the eigenvalues are positive. Therefore the *maximum* of the likelihood function of the distribution of the eigenvalues correspond to the number of endmembers. We then use VCA to extract five endmembers. Figure 8 presents the abundance maps of the five endmembers extracted by using VCA and the simplified abundance maps.

Table 2 Number of pixels and thematic color of each class in the ground truth of HYDICE data.








Class	Training set	Test set	Color
1. Roof	40	3,794	
2. Road	40	376	
3. Trail	40	135	
4. Grass	40	1,888	
5. Tree	40	365	
6. Water	40	1,184	
7. Shadow	40	57	

Figure 8 **a–e** Five abundance maps obtained by linear unmixing on HYDICE data; **f–j** the corresponding segmented abundance maps. For visualization, the gray levels of each figure are linearly normalized to [0, 255].

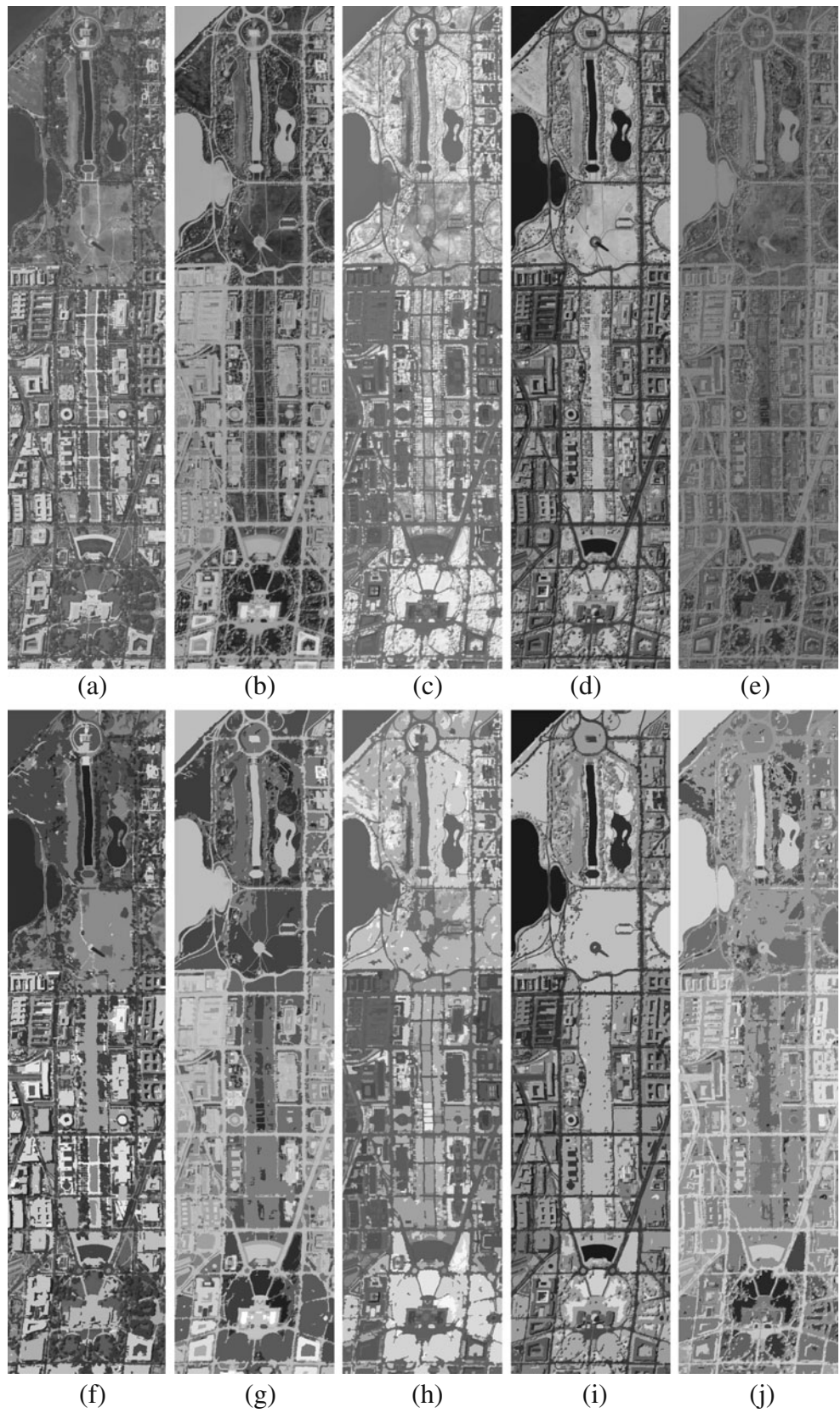


Table 3 Classification results obtained on HYDICE data.

Feature	Original data	Feature vector Θ_1	Feature vector Θ_2	EMPPCA [6]	EMPKPCA [6]	EMPVCA
Number of features	191	5	6	360	108	45
Overall accuracy	99.43%	99.43%	99.43%	98.64%	98.73%	98.62%
Roof	99.92%	99.09%	99.09%	97.52%	97.52%	97.63%
Road	97.11%	98.56%	98.56%	99.52%	98.80%	99.04%
Trail	100%	100%	100%	100%	100%	100%
Grass	99.63%	100%	100%	100%	100%	100%
Tree	97.53%	100%	100%	99.51%	99.51%	100%
Water	99.91%	100%	100%	99.92%	99.92%	99.92%
Shadow	87.62%	94.85%	94.85%	89.69%	100%	84.54%

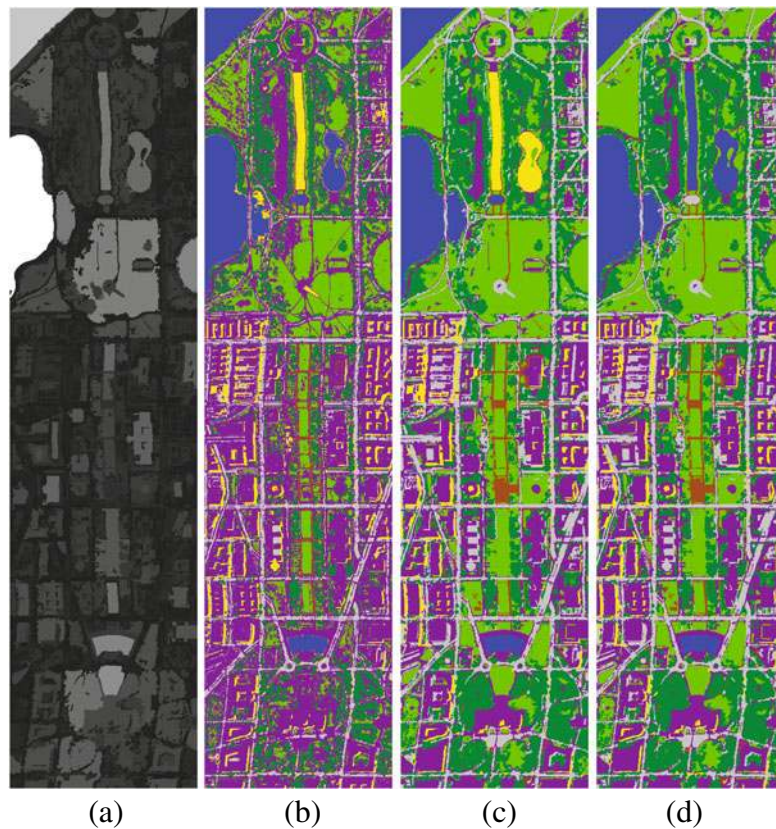
The training set contains 280 pixels and the test set contains 8,079 pixels. We use SVM with Gaussian kernel to classify this image, and 5-fold cross validation is used for determining its optimal scale parameter.

In Table 3, the classification results obtained on the HYDICE data by using respectively the original spectra, the feature vector Θ_1 (see Eq. 10) and the feature vector Θ_2 (see Eq. 11) are shown. For comparison, the results obtained in [6] on the same data set by using EMP on the principle components computed by PCA and KPCA are also shown. The results obtained by using EMP computed on the abundance maps of

the five endmembers are also shown. It can be seen from Table 3 that all the classification accuracies are very similar. This is because the test set is too small. However, the classification results obtained by using feature vectors Θ_1 and Θ_2 are slightly better than the results obtained in [6], even though the feature length of Θ_1 and Θ_2 is much lower.

In order to show the efficiency of the scale feature for classification, in Fig. 9, the scale map computed by Eq. 9 as well as the classification results on the whole HYDICE image by using the original spectra, the feature set Θ_1 and Θ_2 are shown. From Fig. 9b and c where

Figure 9 **a** Local scale values computed by Eq. 9 on HYDICE data, bright pixels correspond to large scales while dark pixels correspond to small scales; **b** Classification results obtained with only the original spectral information; **c, d** Classification results obtained by using feature vectors Θ_1 and Θ_2 , respectively.



only spectral information is used for classification, it can be seen that the class of Water and the class of Shadow are very often mixed, since their spectral properties are quite similar. However, since usually water areas occupy much larger surface than shadow, the scale values of the class of Water are larger than the class of Shadow (see Fig. 9a). Therefore, if scale values are added for classification, these two classes are better separated in Fig. 9d.

5 Conclusion

In this article, we have proposed to integrate geometrical feature, the characteristic scales of structures, for the classification of hyperspectral images. In order to reduce the dimension of the data, we used a linear unmixing algorithm to extract the endmembers and their abundance maps contained in a hyperspectral image. The abundance maps can be considered as a compact representation of spectral information of this image, since the number of endmembers contained in a hyperspectral image is much smaller than the number of spectral bands. With the help of these abundance maps, we extend the method proposed in [10] to hyperspectral images to estimate the characteristic scales of the structures. SVMs are then used for classifying two hyperspectral images over urban areas with the help of the abundance maps and the characteristic scales. The experiments show that with the use of the scale feature, the classification results improve considerably, especially for the objects made by the same material but with different semantic meanings. By using the features proposed, the classification results are very similar to the results obtained by using the methods based on PCA and EMP, even though the number of features proposed is much lower.

Acknowledgement The authors would like to thank Pr. Paolo Gamba, University of Pavia, for providing the ROSIS data and the corresponding ground truth.

References

- Boardman, J. W., Kruse, F. A., & Green, R. O. (1995). Mapping target signatures via partial unmixing of aviris data. In *Proc. JPL airborne earth sci. workshop* (pp. 23–26).
- Camps-Valls, G., & Bruzzone, L. (2005). Kernel-based methods for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 43(6), 1351–1362.
- Caselles, V., Coll, B., & Morel, J.-M. (1999). Topographic maps and local contrast changes in natural images. *International Journal of Computer Vision*, 33(1), 5–27.
- Chanussot, J., Benediktsson, J. A., & Fauvel, M. (2006). Classification of remote sensing images from urban areas using a fuzzy possibilistic model. *IEEE Geoscience and Remote Sensing Letters*, 3(1), 40–44.
- Fauvel, M., Benediktsson, J. A., Chanussot, J., & Sveinsson, J. R. (2008). Spectral and spatial classification of hyperspectral data using SVMs and morphological profile. *IEEE Transaction on Geoscience and Remote Sensing*, 46(11).
- Fauvel, M., Chanussot, J., & Benediktsson, J. A. (2009). Kernel principal component analysis for the classification of hyperspectral remote-sensing data over urban areas. *EURASIP Journal on Advances in Signal Processing*, 2009, 1–14.
- Guo, B., Gunn, S. R., Damper, R. I., & Nelson, J. D. B. (2006). Band selection for hyperspectral image classification using mutual information. *IEEE Geoscience and Remote Sensing Letters*, 3(4), 522–526.
- Landgrebe, D. A. (2003). *Signal theory methods in multispectral remote sensing*. New Jersey: John Wiley and Sons.
- Licciardi, G., Pacifici, F., Tuia, D., Prasad, S., West, T., Giacco, F., et al. (2009). Decision fusion for the classification of hyperspectral data: Outcome of the 2008 grs-s data fusion contest. *IEEE Transaction on Geoscience and Remote Sensing*, 47(11), 3857–3865.
- Luo, B., Aujol, J.-F., & Gousseau, Y. (2009). Local scale measure from the topographic map and application to remote sensing images. *SIAM Multiscale Modeling and Simulation*, 8(1), 1–29.
- Luo, B., & Chanussot, J. (2009). Hyperspectral image classification based on spectral and geometrical features. In *IEEE MLSP, Grenoble, France*.
- Luo, B., & Chanussot, J. (2009). Unsupervised hyperspectral image classification by using linear unmixing. In *IEEE ICIP, Cairo, Egypt*.
- Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transaction on Geoscience and Remote Sensing*, 42(8), 1778–1790.
- Monasse, P., & Guichard, F. (2000). Fast computation of a contrast-invariant image representation. *IEEE Transactions on Image Processing*, 9(5), 860–872.
- Nascimento, J. M. P., & Dias, J. M. B. (2005). Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Transaction on Geoscience and Remote Sensing*, 43(4), 898–910.
- Palmason, J. A., Benediktsson, J. A., Sveinsson, J. R., & Chanussot, J. (2005). Classification of hyperspectral data from urban areas using morphological preprocessing and independent component analysis. In *IEEE IGARSS'05—International geoscience and remote sensing symposium, Seoul, Korea* (pp. 176–179).
- Pesaresi, M., & Benediktsson, J. A. (2001). A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Transaction on Geoscience and Remote Sensing*, 39(2), 309–320.
- Plaza, A., & Chang, C.-I. (2006). Impact of initialization on design of endmember extraction algorithms. *IEEE Transaction on Geoscience and Remote Sensing*, 44(11), 3397–3407.
- Poggi, G., Scarpa, G., & Zerubia, J. (2005). Supervised segmentation of remote sensing images based on a tree-structure MRF model. *IEEE Transaction on Geoscience and Remote Sensing*, 43(8), 1901–1911.
- Serpico, S. B., & Bruzzone, L. (2001). A new search algorithm for feature selection in hyperspectral remote sensing images. *IEEE Transaction on Geoscience and Remote Sensing*, 39(7), 1360–1367.

21. Winter, M. E. (1999). N-FINDER: An algorithm for fast autonomous spectral endmember determination in hyperspectral data. In *Proceedings of SPIE* (Vol. 3753, pp. 266–277).



Jocelyn Chanussot received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (INPG), Grenoble, France, in 1995, and the Ph.D. degree from Savoie University, Annecy, France, in 1998. In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA–French National Defense Department). Since 1999, he has been with INPG, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of signal and image processing. He is currently conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA–Lab, INPG). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing. Dr. Chanussot is the founding President of IEEE Geoscience and Remote Sensing French chapter (2007) which received the 2010 IEEE GRS-S Chapter Excellence Award “for excellence as a Geoscience and Remote Sensing Society chapter demonstrated by exemplary activities during 2009”. He is a member of the IEEE Geoscience and Remote Sensing AdCom (2009–2011). He is the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He is the Chair (2009–2011) and was the Cochair of the GRS Data Fusion Technical Committee (2005–2008). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006–2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing, (2009). He was an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (2005–2007) and for Pattern Recognition (2006–2008). Since 2007, he is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He is a Senior Member of the IEEE (2004).



Bin Luo received the M.Sc. degree from ENS Cachan, France and the Ph.D. degree from ENST Paris, France in 2007. He worked as post doctoral researcher in GIPSA-Lab, France from 2008 to 2010. He is now an associate professor in the LIES-MARS, Wuhan University, China. His research interests include hyperspectral data analysis, high resolution image processing and indexation of images at different resolutions.