HYPERSPECTRAL IMAGE CLASSIFICATION BASED ON SPECTRAL AND GEOMETRICAL FEATURES

Bin Luo and Jocelyn Chanussot

GIPSA-Lab, 961 rue de la Houille Blanche, 38402 Grenoble, France.
(Bin.luo@gipsa-lab.inpg.fr)
(jocelyn.chanussot@gipsa-lab.inpg.fr)

ABSTRACT

In this paper, we propose to integrate geometrical features, such as the characteristic scales of structures, with spectral features for the classification of hyperspectral images. The spectral features which only describe the material of structures can not distinguish objects made by the same material but with different semantic meanings (such as the roofs of some buildings and the roads). The use of geometrical features is therefore necessary. Moreover, since the dimension of a hyperspectral image is usually very high, we use linear unmixing algorithm to extract the endmemebers and their abundance maps in order to represent compactly the spectral information. Afterwards, with the help of these abundance maps, we propose a method based on topographic map of images to estimate local scales of structures in hyperspectral images.

The experiment shows that the geometrical features can improve 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 morphological profiles), the local scale feature provides satisfactory results without considerably increasing the feature dimension.

1. INTRODUCTION

The classification of hyperspectral remote sensing images is a challenging task, since the data dimension is considerable for traditional classification algorithms, typically several hundreds of spectral bands are acquired for each image. These spectral bands can provide very rich spectral information of each pixel in order to identify the material of the objects. However, the spectral information alone sometimes does not allow the separation of structures. For example, the roofs of some buidings and the roads can be made by the same material (alsphalt). Therefore, contextual information, geometrical features for example, is necessary for classification task of hyperspectral images.

The major difficulty to calculate the contextual information of hyperspectral images is the high dimension of the data. To reduce the data dimension, both supervised and non supervised methods are proposed. The supervised methods, such as band selection [1] [2], Decision Boundary feature extraction and Non-Weighted feature extraction [3], 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 training set. The unsupervised methods, such as PCA (principle component analysis) or ICA (independant component analysis), optimise some statistical criteria (such as the most un correlated components or the most independant components) to project the data onto a sub space with lower dimension. The application of these methods on hyperspectral data can be found in [4] [5] [6]. However, the components obtained by optimising the statistical criteria do not necessarily have physical meanings. In this paper, we propose to use Vertex Component Analysis (VCA) to reduce the dimension of hyperspectral data [7]. We suppose that the spectrum of each pixel is a linear mixture of the spectra of different chemical species (refered 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 the endmembers is much less than the number of spectral bands, these abundance maps can be considered as a compact representation of spectral information provided by the hyperspectral image.

In order to extract the contextual information from remote sensing images, one can find the methods based on Markov Random Field (MRF) [8]. However, the MRF based methods provide only statistical information on the neighborhood of the considered pixel. Another family of methods for contextual information extraction are based on morpholigical operators, which allow to extract descriptive features, such as the geometrical features about the structures [9]. [4] [5] extract the extended morphological profiles (EMP) of the principle components of hyperspectral images for the classification. However, since all morphological methods require a structred element, the features extracted by such methods depend on the used structured element. Moreover, EMP increase con-

This work is funded by French ANR project VAHINE.

978-1-4244-4948-4/09/$25.00 © 2009 IEEE

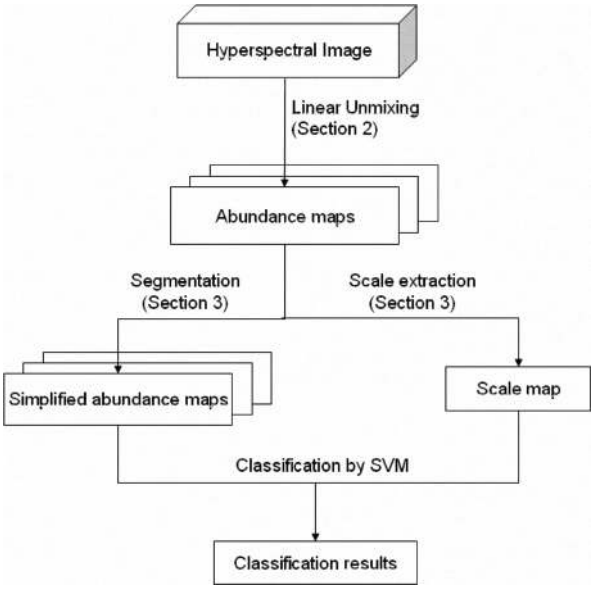


Fig. 1. Scheme of the paper.

siderably the feature dimension. For example, in order to describe the geometrical information of one pixel on a principle component, the EMP method in [4] [5] uses 8 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, the contrasted structure containing it is extracted, and the scale of this structure is defined as the scale of this pixel. In this paper, we try to extend the algorithm presented in [10] for extracting a local characteristic scale for each pixel on hyperspectral images with the help of the abundance maps obtained by VCA. The main advantages of our proposed method when compared to EMP are two-fold. At one hand, no structured 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 need many values.

The plan of this article can be illustrated by in Figure 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 effeciency of the local scale feature.

2. LINEAR UNMIXING OF HYPERSPECTRAL DATA

We note $\mathbf{X}$ the matrix representing the hyperspectral image cube, where $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_a}\}$ and $\mathbf{x}_k = \{x_{1,k}, x_{2,k}, \ldots, x_{N_s,k}\}^T$, $x_{l,k}$ is the value of the kth pixel at the lth 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:

$$\mathbf{X} = \mathbf{MS} + \mathbf{n} \tag{1}$$

where $\mathbf{M} = \{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_{N_c}\}$ is the mixing matrix, where $\mathbf{m}_n$ denotes the spectral signature of the nth endmember. $\mathbf{S} = \{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_{N_c}\}^T$ is the abundance matrix where $\mathbf{s}_n = \{s_{n,1}, s_{n,2}, \ldots, s_{n,N_a}\}$ ($s_{n,k} \in [0,1]$ is the abundance of the nth endmember at the kth pixel). $\mathbf{n}$ stands for the additive noise of the image. In [7], the Vertex Component Analysis (VCA) is proposed as an effecient 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 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 extreme of this projection is the new endmember signature. The algorithm stops when all the p endmembers are extracted, where p is the number of endmembers. Even though the number of endmembers is much smaller than the number of spectral bands, i.e. $p << N_s$, abundance maps of the p endmembers can provide the same spectral information as the hyperspectral image with N_s spectral bands. Therefore this linear unmixing step can be considered as a dimension reduction of data. Moreover, the values on the abundance maps for one given pixel which represent the proportions of different chemical species on this pixel are comparable, which is essential for estimating the characteristic scale of this pixel. On the contrary, values of one pixel obtained by other dimension reduction methods (such as PCA or ICA) are not comparable. In the next section, we can see 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 OF HYPERSPECTRAL IMAGE

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 as the characteristic scale of this pixel. The topographic

map [11], which can be obtained by Fast Level Set Transformation (FLST) [12], represents an image by an inclusion tree of the shapes (which are defined as the connected components of level sets). An example of such inclusion tree is shown in Figure 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 $f_{\hat{i}}(x, y)$ of a given pixel is defined as the shape containing this pixel, of which the contrast is the most important, i.e.

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

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

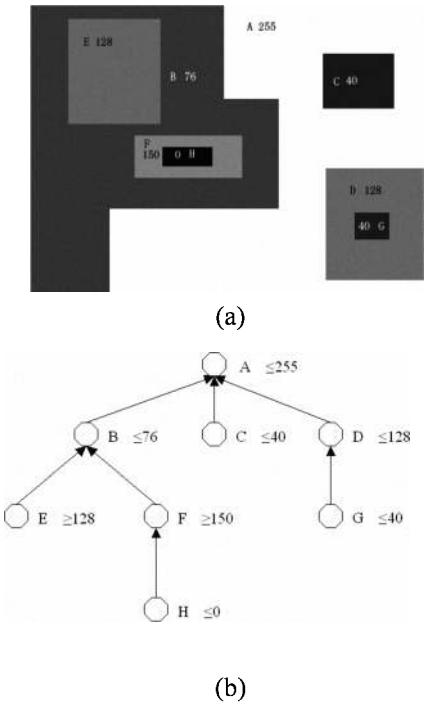
$$E(x, y) = S(f_{\hat{i}}(x, y))/P(f_{\hat{i}}(x, y)). \tag{3}$$

Since the optical instruments always blur remote sensing images, 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 succesive shapes (for example f_i and f_{i+1}) corresponding to a same structure is proportional to the perimeter of the smaller shape, i.e. $S(f_{i+1}) - S(f_i) \sim \lambda P(f_i)$, where λ is a constant. 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

$$\bar{I}(x, y) = \frac{\sum_{(k,l) \in f_{\hat{i}}(x,y)} I(k, l)}{S(f_{\hat{i}}(x, y))} \tag{4}$$

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 context pixels.

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 nth abundance map, we note $E_n(x, y)$ its local scale calculated by Equation (3), $f_{\hat{i},n}(x, y)$ the most contrasted shape extracted by Equation (2), and $\bar{I}_n(x, y)$ its value of the simplified image defined by Equation (4). 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 nth 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 characteristic scale at the spatial position (x, y) for a hyperspectral image is defined as



(a)

(b)

Fig. 2. Example of FLST : (a) Synthetic image ; (b) Inclusion tree obtained with FLST.

$$\hat{E}(x, y) = E_{\hat{n}}(x, y) \tag{5}$$

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

The feature vector $\Theta(x, y)$ of a given pixel (x, y) used for classification contains the values of the simplified image of all the abundance maps and its local scale defined by Equation (5), i.e.

$$\Theta(x, y) = \{\bar{I}_1(x, y), \ldots \bar{I}_{N_c}(x, y), \hat{E}(x, y)\}. \tag{6}$$

4. EXPERIMENT

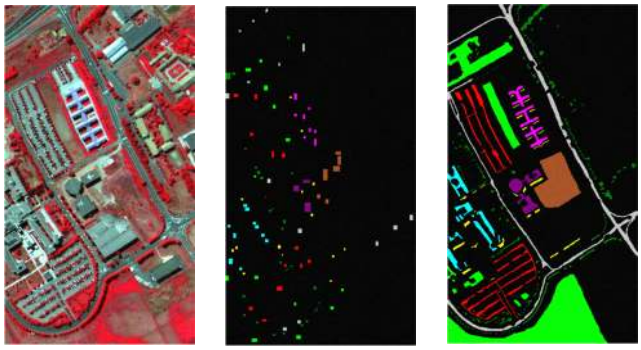
In this section, we use the features defined by Equation (6) to classify a hyperspectral image taken by the instrument ROSIS (Reflective Optics System Imaging Spectrometer) over the University of Pavia, Italy (see Figure 3(a)). The image (with a spatial resolution of $1.3m$) contains 340×610 pixels and 103 spectral bands covering visible and near infrared light. The image is manually classified into 9 classes

and the definitions of these classes are shown in Figure 4. For classification purpose, the training set contains 3921 pixels (see Figure 3(b)) while the test set contains 42776 pixels (see Figure 3(c)). Since in this image, there are mainly 3 endmembers present (vegetation, bare soil and metal roof), we extract 3 endmembers and their abundance maps by using VCA (see Section 2), which are shown in Figures 5(a)-(c). Afterwards we compute the simplified images of these abundance maps by using Equation (4). The simplified images are shown in Figure 5(d)-(f). According to Equation (5), we estimate the local scale for each pixel of this image. The scale map is shown in Figure 6(a). Therefore the feature vector $\Theta(x, y)$ of a pixel (x, y) defined by Equation (6) contains only 4 values: 3 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) and the feature vector Θ defined by Equation (6). Kernel methods are proved to be effecient algorithms for hyperspectral image classification. [13] 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 Figure 6(b) and (c), the classification results obtained on the whole image is shown. In [5], 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 (3 principle components are extracted) or KPCA (12 principle components are extracted) as features for classifying the hyperspectral image. In Table 1, we show the classification results obtained in [5] on the same data set by using SVM with Gaussian kernel.

It can be seen that by using the features proposed in our article, the classification results improve considerably when compared to the results obtained on original hyperspectral data. Recall that the length of the feature vector Θ 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. However, the number of features extracted by PCA and EMP is 27, which is much larger than the number of features in Θ. The results obtained by using KPCA and EMP are better than the results obtained by using Θ and the results by using PCA and EMP. However, the feature dimension, which is 108, is even higher than the dimension of original data (103). And the improvement of classification accuracies is mainly concentrated on



(a) (b) (c)

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

CLASSES	Points in Training Set	Thematic Colour
Trees	524	
Asphalt	548	
Bitumen	375	
Gravel	392	
(painted) metal sheets	265	
Shadow	231	
Self-Blocking Bricks	514	
Meadows	540	
Bare Soil	532	

Fig. 4. Definitions of the classes.

the class of **Meadows**, which is not well defined in the ground truth.

It has to be remarked that by using only the spectral information (i.e. the spectrum of each pixel), it is very difficult to distinguish the class **Asphalt** and **Bitumen**, since they are made by the same material. According to the ground truth of Figure 3(c), 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 Figure 6(a) 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 are much more accurate. This phenomenon can be illustrated by Figure 6(d)-(f), where we have zoomed the classification results obtained by using original data and the feature vectors Θ on a building roof made by bitumen. It can be seen that the building roof and the road are classified as the same class in Figure 6(e). In contrary, by adding the local scale feature, these two structures are well separated.

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

Feature Number of features	Original data 103	Feature vector Θ 4	EMP$_{PCA}$ [5] 27	EMP$_{KPCA}$ [5] 108
Overall accuracy	76.01%	91.54%	92.04%	96.55%
Tree	98.59%	94.52%	99.22%	99.35%
Asphalt	78.16%	96.27%	94.60%	96.23%
Bitumen	89.02%	99.32%	98.87%	99.10%
Gravel	64.55%	84.61%	73.13%	83.66%
(painted) metal sheets	99.47%	99.55%	99.55%	99.48%
Shadow	99.89%	96.30%	90.07%	98.31%
Self-Blocking Bricks	91.01%	99.70%	99.10%	99.46%
Meadows	64.23%	85.80%	88.79%	97.58%
Bare Soil	82.72%	96.56%	95.23%	92.88%

Table 1. Classification results

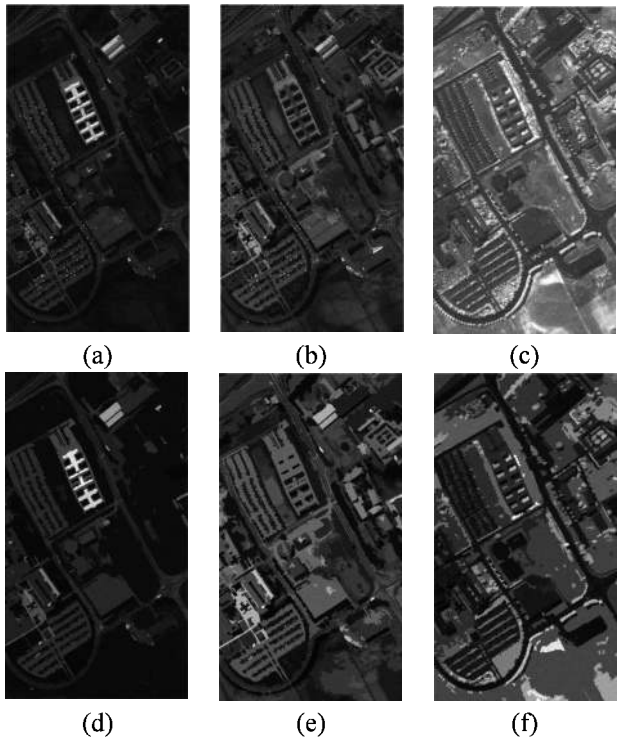


Fig. 5. (a)-(c) Abundance maps of the 3 endmembers extracted by VCA; (d)-(f) the simplified images of these abundance maps.

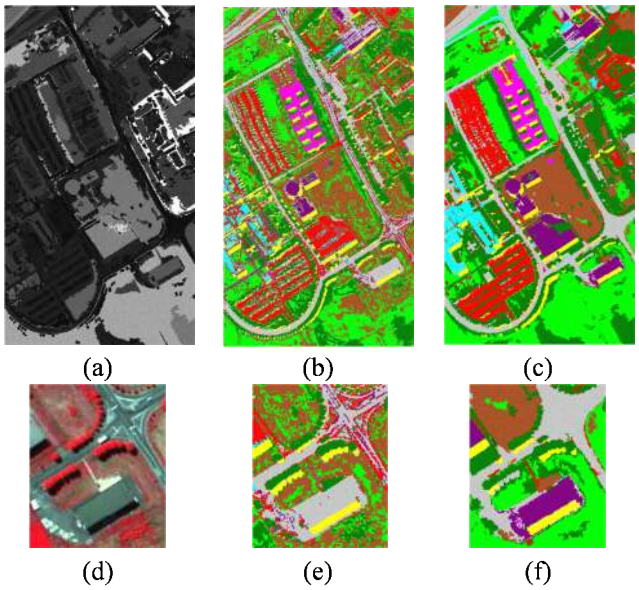


Fig. 6. (**a**) Scale map of the hyperspectral image, bright pixel corresponds to big scale while dark pixel corresponds to small scale; (**b**) Supervised classification results obtained on original data set; (**c**) Supervised classification results obtained on the feature set obtained by Equation (6). (**d**)-(**f**): Zoom around a building on the: (**d**) original image; (**e**) classification results of (b); (**f**) classification results of (c).

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. The experiments show that 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 less.

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

6. REFERENCES

[1] S. B. Serpico and L. Bruzzone, "A new search algorithm for feature selection in hyperspectralremote sensing images," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 39, no. 7, pp. 1360 – 1367, July 2001.

[2] B. Guo, S.R. Gunn, R. I. Damper, and J.D.B. Nelson, "Band selection for hyperspectral image classification using mutual information," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 4, pp. 522–526, 2006.

[3] D.A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*, John Wiley and Sons, New Jersey, 2003.

[4] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profile," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 46, no. 11, November 2008.

[5] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Kernel principal component analysis for the classification of hyperspectral remote-sensing data over urban areas," *EURASIP Journal on Advances in Signal Processing*, 2009, to appear.

[6] J.A. Palmason, J.A. Benediktsson, J.R. Sveinsson, and J. Chanussot, "Classification of hyperspectral data from urban areas using morpholgical preprocessing and independent component analysis," in *IEEE IGARSS'05 - International Geoscience and Remote Sensing Symposium*, Seoul Korea, 2005, pp. 176–179.

[7] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898 – 910, April 2005.

[8] G. Poggi, G. Scarpa, and J. Zerubia, "Supervised segmentation of remote sensing images based on a tree-structure MRF model," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 43, no. 8, pp. 1901–1911, 2005.

[9] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 39, no. 2, pp. 309–320, Febrary 2001.

[10] B. Luo, J.-F. Aujol, and Y. Gousseau, "Local scale measure from the topographic map and application to remote sensing images," *SIAM Multiscale Modeling and Simulation*, to appear.

[11] V. Caselles, B. Coll, and J.-M. Morel, "Topographic maps and local contrast changes in natural images," *Int. J. Comp. Vision*, vol. 33, no. 1, pp. 5–27, 1999.

[12] P. Monasse and F. Guichard, "Fast computation of a contrast-invariant image representation," *IEEE Trans. on Image Processing*, vol. 9, no. 5, pp. 860–872, may 2000.

[13] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1351–1362, 2005.