Geometrical features for the classification of very high resolution multispectral remote-sensing images

Bin Luo, Jocelyn Chanussot

GEOMETRICAL FEATURES FOR THE CLASSIFICATION OF VERY HIGH RESOLUTION MULTISPECTRAL REMOTE-SENSING IMAGES

Bin Luo and Jocelyn Chanussot

GIPSA-Lab, Grenoble Institute of Technology, Grenoble, France
{Bin.luo, Jocelyn.chanussot}@gipsa-lab.inpg.fr

ABSTRACT

In order to extract geometrical features from a multispectral image and derive a classification, an approach based on the topographic map of the image is proposed. For each pixel, the most significant structure containing it is extracted. The classification of this pixel is based on its spectral information and the geometrical features of the corresponding structure (its area and perimeter). The results obtained on multispectral remote sensing images taken by two different sensors show the efficiency of the extracted geometrical features for separating some classes with very similar spectral attributes but of different semantic meanings.

1. INTRODUCTION

Recently, very high resolution multispectral remote sensing images are increasingly used for Earth observation. These images, providing rich spatial and spectral information on the observed scenes, better describe the material attributes than panchromatic images. However, by using only the spectral bands, it is difficult to distinguish some classes with similar spectral attributes, such as water and shadow. Moreover, it is impossible to separate objects made by the same material but with different semantic meanings (such as the roads and the roofs of some buildings). Therefore, contextual information, in particular geometrical features, become important for the classification of very high resolution multispectral images.

In order to extract the contextual information from remote sensing images, one can use models based on Markov Random Field (MRF) [1]. 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 [2]. In [3], it is proposed to extract the extended morphological profiles (EMP) of the principle components of hyperspectral images for the classification. However, since all morphological methods 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 structuring elements with a large discrete scale range. This fact has two major drawbacks. Firstly, the scale range has to be chosen manually and is not necessarily adapted to the image. Secondly, it considerably increases the feature dimension. For example, in order to describe the geometrical information of one pixel on a principle component, the EMP method in [3] uses 8 values.

In [4], for panchromatic remote-sensing images, it is proposed to define the pixel-wise characteristic scale of the data using the topographic map of gray scale images. The main idea is that for each pixel, the most contrasted structure containing it is extracted, and the scale of this structure defines the scale of this pixel. In [4], it was shown that the structures extracted on all the pixels form a partition of the image and that the structure extracted on a pixel is in fact the most significant structure containing it. The geometrical features of these structures are very pertinent for describing the contextual information. In this paper, we try to extend the algorithm presented in [4] for extracting geometrical features (the areas and the perimeters of the structures) of multispectral images. The main advantages of our proposed method when compared to EMP are two-fold. Firstly, no structuring element is required. The proposed approach can extract the geometrical features of structures of any scale. In addition, only a very few values are required to describe the geometrical attributes of a spatial position when compared the EMP.

The outline of the paper is as follow: in Section 2, we briefly introduce and then extend the approach proposed in [4] for extracting geometrical features of remote-sensing images. In Section 3, the classification results obtained with the help of the geometrical features are presented. In Section 4, the conclusion is drawn and some perspectives are given.

2. EXTRACTION OF GEOMETRICAL FEATURES

2.1. Extraction of geometrical features on panchromatic remote sensing images

In [4], 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

This work is funded by French ANR project VAHINE.

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 [5], which can be obtained by Fast Level Set Transformation (FLST) [6], represents an image by an inclusion tree of the shapes (which are defined as the connected components of the level sets). 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.

$$f_{\hat{i}}(x, y) = \arg\max_j \{C(f_j(x, y)\} \quad (1)$$

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 [4] 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. $S(f_{i+1}) - S(f_i) \sim \lambda P(f_i)$, where λ is a constant. It is shown in [4] that the most contrasted shapes extracted in an image form a partition of this image. The area and the perimeter values of the most contrasted shape $f_{\hat{i}}(x, y)$ are very pertinent geometrical features for the classification task.

2.2. Extension to multispectral images

We note $Y_n(x, y)$, $(n = 1, \ldots, N_c)$ the nth channel of the multispectral image and N_c the number of spectral channels. In order to extend the estimation of local scale to multispectral images, we firstly extract the most contrasted shapes for each channel. Therefore, for each spatial position (x, y) in a multispectral image, N_c shapes are extracted. We note $f_{\hat{i},n}(x, y)$ the most contrasted shape extracted by Equation (1) for the pixel (x, y) on the nth channel. The simplest way to use the scale features of a pixel is to use all the scale values of all the channels as features for the classification. The major drawback is that for a given object, if the nth channel is dominant, the area and perimeter values computed on the other channels may introduce classification errors. Therefore, we try to define an unique area value and an unique perimeter value for each pixel by the most significant channel. For a given pixel (x, y), we try to select the most significant shape $f_{\hat{i},\hat{n}}(x, y)$ from the most contrasted shapes $f_{\hat{i},n}(x, y)$ extracted from all the channels of the image. The criterion of the selection is always based on the contrasts of these shapes $C(f_{\hat{i},n}(x, y))$. Remark that, since the dynamic of each channel may be quite different, we normalize the contrast value by the total variation $TV_n = \int \sqrt{\partial_x Y_n^2 + \partial_y Y_n^2}$ of the corresponding channel. Therefore, the most significant shape of the pixel (x, y) is defined as:

$$f_{\hat{i},\hat{n}}(x, y) = \arg\max_n \frac{C(f_{\hat{i},n}(x, y))}{TV_n}. \quad (2)$$

Since the area and the perimeter values of the structures in an image may significantly vary (from several pixels to tens of thousands of pixels), we associate to the pixel (x, y) the logarithmic values of the area and the perimeter of the shape $f_{\hat{i},\hat{n}}(x, y)$.

Therefore, the feature vector $\Theta(x, y)$ of a given pixel (x, y) containing spectral and geometrical features for classification is defined as:

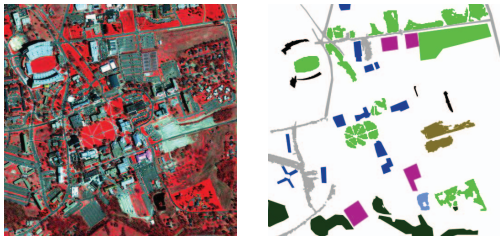
$$\begin{aligned} \Theta(x, y) \quad = \quad & \{Y_1(x, y), \ldots Y_{N_c}(x, y), \\ & \log S(f_{\hat{i},\hat{n}}(x, y)), \log P(f_{\hat{i},\hat{n}}(x, y))\}. \end{aligned} \quad (3)$$

3. EXPERIMENTS AND RESULTS

Geometrical features are more pertinent on very high resolution remote-sensing images, since fine structures (such as buildings, roads, etc.) of urban areas become individually visible. Therefore, in this section, we validate the efficiency of the geometrical features on classification by the experiments carried out on two multispectral remote sensing images taken by two different high resolution sensors: Quickbird (2.4m) and simulated Pleïade (1m).

3.1. Experiments on QuickBird image

The first experiment is carried out on a Quickbird multispectral image of Toulouse (see Figure 1(a)) with four channels (infra-red, red, green and blue). The spatial resolution is 2.4m. The image size is 512×512 pixels. 23% of pixels of this image are manually classified into 8 classes serving as ground truth. The labelled pixels and the definitions of the classes are shown in Figure 1(b).

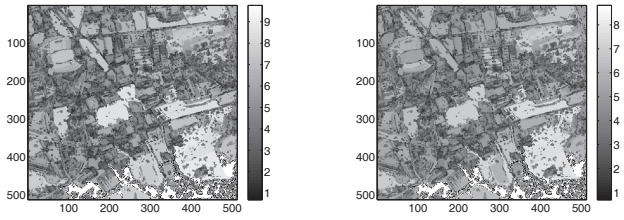


(a) (b)

Fig. 1. (a) Quickbird image on Toulouse. For the visualization, only the infra red, red and green channels are shown. (b) Manually labeled pixels. The definitions of the classes are: Grass, Tree, Shadow, Building, Road, Water, Bare soil, Parking.

We use SVM (Support Vector Machines) with Gaussian kernel for classification. 10% of the pixels in each class are

selected for training the SVM. The optimal scale parameter of the Gaussian kernel is selected by 5-fold cross validation on the training set. For comparison, we classified this image by using two different feature sets. The first feature set $\{Y_n\}$, $(n = 1, \ldots, 4)$ contains only the four spectral bands. The second feature set Θ defined by Equation (3) contains not only the spectral bands but also the logarithmic values of the area and the perimeter computed on this image, which are shown in Figure 2.



(a) (b)

Fig. 2. Logarithmic values of (a) the areas and (b) the perimeters of the most contrasted shapes computed on all the pixels of the Quickbird image.

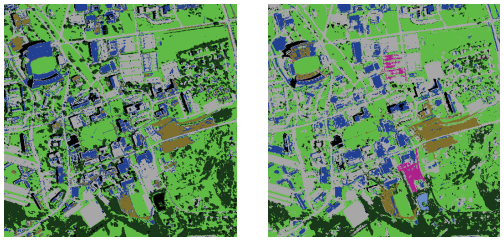
The overall accuracy and the class accuracies obtained by using the two feature sets on the Quickbird image are shown in Table 1. Figure 3 shows the classification results on the whole image with the help of the SVM trained by 10% of the labelled pixels by using the two feature sets. It can be seen that the addition of geometrical features improves the classification results, especially for the classes of Shadow and Water. It is very difficult to separate these two classes when we use only the spectral channels (there isn't even any pixel which is classified as Water). The geometrical features can separate these classes, since the area and the perimeter of the water region in this image (which is in fact a lake) are much larger than for the shadows. Moreover, the geometrical features considerably improve the classification results of the Road and Building classes. The color of the roads and the color of some building roofs are exactly the same, since they are all made of asphalt. Their geometrical attributes are different, since the ratio between the area and perimeter of the buildings is larger than for the roads. The same phenomenon can be observed for the Parkings. By using only spectral channels, no pixel is classified to this class, because the color of the parkings is the same as for the roads.

3.2. Experiments on simulated Pleïade data

The second experiment is carried out on a simulated Pleïade image (see Figure 4(a)) taken on Strasbourg with four spectral channels (near infra red, red, green and blue). The spatial resolution is $1m$ and the image size is 1024×1024 pixels. 38.9% of the pixels are manually labelled into 7 classes serving as ground truth. The labelled pixels and the definitions of the classes are shown in Figure 4(b).

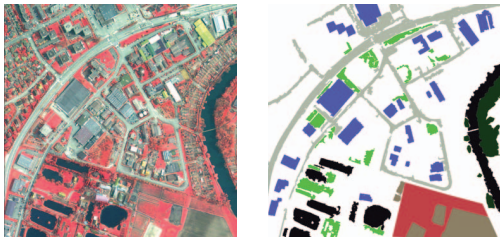
Feature set Feature dimension	$\{Y_n\}$ Original channels 4	Θ (see Eq. (3)) 6
Overall Accuracy	77.27%	84.84%
1. Grass	96.94%	97.96%
2. Tree	88.75%	88%
3. Shadow	74.44%	75.07%
4. Building	51.97%	74.82%
5. Road	77.54%	87.68%
6. Water	0%	71.5%
7. Bare soil	94.05%	97.83%
8. Parking	0%	22.86%

Table 1. Overall accuracy and class accuracies obtained on the Quickbird image.



(a) (b)

Fig. 3. Classification results obtained on the Quickbird image (a) by using the original 4 spectral bands only, and (b) by using feature vector Θ (see Equation (3)).

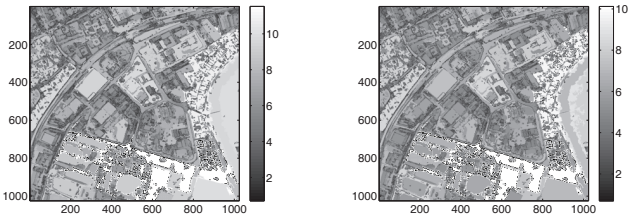


(a) (b)

Fig. 4. (a) Pleïade image on Strasbourg. For the visualization, only the infra red, red and green channels are shown. (b) Manually labelled pixels. The definitions of the classes are: Road, Building, Water, Field, Bare soil, Grass, Tree.

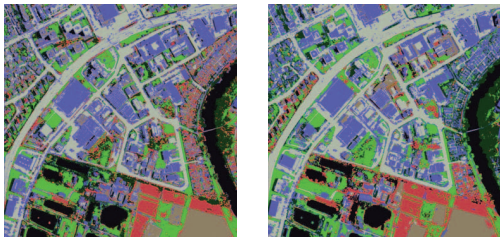
SVM with Gaussian kernel is used, its optimal scale parameter is selected by 5-fold cross validation. 10% of the pixels of each class are used for training the SVM. For comparison, the image is classified by two feature sets: the original channels $\{Y_n, n = 1, \ldots, 4\}$ and the feature set Θ (see Equation (3)). Figure 5 shows the computed logarithmic values of the areas and perimeters computed.

The overall accuracy and the class accuracies obtained by using the two feature sets are shown in Table 2. Figure 6 shows the classification results of the whole Pleïade image



(a) (b)

Fig. 5. Logarithmic values of (a) the areas and (b) the perimeters of the most contrasted shapes computed on all the pixels of the simulated Pleïade image.



(a) (b)

Fig. 6. Classification results obtained on the simulated Pleïade image (a) by using the original 4 spectral bands only, and (b) by using feature vector Θ (see Equation (3)).

with the help of the SVM trained by 10% of the labelled pixels by using the two feature sets. It can be seen that, again, the classification results are considerably improved by adding the geometrical features. The overall accuracy increases by nearly 9.14% and the accuracy of each class increases too. In particular, the accuracies are significantly improved for the Building, Field and Tree classes. For the class of Building, the roads and some building roofs are made by the same material, and hence have the same color. However, the ratio between the area and the perimeter of buildings is larger than for the roads. The colors of the fields and of the trees are quite similar. However, the fields have larger areas and much more regular contours. Therefore, geometrical features can better separate these two classes. Moreover, it can be observed from Figure 6 that when only the spectral channels are used, lots of errors occur on the top left corner of the image. Many pixels corresponding to buildings (mainly small houses) are classified as Bare soil. Again, by adding geometrical features, these pixels are better classified, because the Bare soil class in this image corresponds to larger area and perimeter.

Feature set Feature dimension	$\{Y_n\}$ Original channels 4	Θ (see Eq. (3)) 6
Overall Accuracy	83.87%	93.01%
1. Road	91.84%	96.88%
2. Building	77.37%	84.86%
3. Water	94.14%	97.93%
4. Field	63.23%	92.36%
5. Bare soil	89.17%	90.78%
6. Grass	88.39%	93.03%
7. Tree	63.2%	89.96%

Table 2. Overall accuracy and class accuracies obtained on the simulated Pleïade image.

4. CONCLUSION AND PERSPECTIVES

In this paper, we describe an approach for automatically extracting the geometrical features (the areas and the perimeters of the structures) of multispectral remote sensing images based on topographic map. The results on both Quickbird and simulated Pleïade images show the efficiency of these features: by adding only two features, it is possible to separate some classes which are difficult to separate by using spectral bands only.

However, it can be seen that, in this paper, the geometrical features are simply appended to the feature set. It will be interesting to apply a feature ponderation stradegy (such as the method proposed in [7]) in order to avoid potential classification errors introduced by the geometrical features.

5. REFERENCES

[1] 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.

[2] 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.

[3] 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*, vol. 2009, pp. 1–4, 2009.

[4] 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*, vol. 8, no. 1, pp. 1–29, 2009.

[5] 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.

[6] 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.

[7] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Decision fusion for the classification of urban remote sensing images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 44, no. 10, pp. 2828–2838, October 2006.