

A Land Use Classification Method Based on Region and Edge Information Fusion

Gang WU, Chunhong PAN, Veronique PRINET, and Songde MA

gwu@nlpr.ia.ac.cn

National Laboratory of Pattern Recognition,
Institute of Automation, Chinese Academy of Sciences,
Beijing, P.R. China

ABSTRACT

In this paper, we propose an unsupervised land use classification method for multispectral image. Especially, we present a novel spatial mean shift procedure and an automatic band selection method. We make full use of the edge information from the mean shift procedure to improve the over-segmentation. Experimental results on Landsat TM images validate the efficiency of the proposed method.

1. INTRODUCTION

Land use classification is an important task in remote sensing field. This procedure assigns a class label, such as water, urban, forest, and soil, to each image pixel and results in a thematic map that can be used in environment monitoring and disaster management, etc.

There are many approaches to this problem proposed in the past decades, either supervised ones or unsupervised ones [1]. And some other image segmentation methods are also widely used in this application, such as region-based, edge-based and hybrid method [2]. In recent years, the multi-sensor image fusion and multiple classifiers fusion methods [4] become an active research topic.

The difficulties of land use classification problem mainly lie in two aspects [3]. The first is that the obtained thematic map appears “salt-and-pepper” noise, i.e., isolated mislabeled pixels disperse within another labeled area, due to reasons of sensor noises or atmospheric distortions. The second is the spectral noise. It means that within one pixel, in fact, a variety of response measurements of small ground objects are mixed, which leads to the impure spectral information. Therefore, some pre and post processing steps are necessary.

In this paper, we propose an unsupervised method for multispectral image with the purpose of creating a correct and smooth thematic map. Our input data is Landsat TM

image with 7 bands. During the preprocessing stage, the spectral noise is suppressed by mean shift smoothing; in addition, a new spatial mean shift procedure is presented to generate an image, which is named as spatial accumulation map, and from this map much information, such as homogeneous areas, edges and centers of small regions, can be extracted for the successive processing. Then by seeded region growing an over-segmented image can be generated. During the post-processing stage, the edge information from the spatial accumulation map will be incorporated to get rid of invalid regions, and the different regions with the same land cover will be further assigned the same label by choosing the band from the original multispectral image that responses this land cover best, and then merging them under the selected band. As a result, the salt-and-pepper appearance is effectively eliminated.

2. METHOD DESCRIPTION

2.1. Band reduction

At first, a principle component analysis (PCA) transform is performed. PCA removes the correlation of each band, and it is a general tool for coordinate transform and data reduction in multispectral image processing [1]. After this step, the input seven-band images are condensed into three channels, thus reducing the computational demands and possibly improving performance, for the first three principle components preserve the essential information (>95%) of the original data [6].

2.2. Mean shift procedure

Mean shift procedure is essentially a detector for local maxima of the p.d.f. (probability density function) in the feature space; it can provide reliable solutions for many vision tasks [5].

In d -dimensional feature space, given n points $x_i, i = 1, \dots, n$, the p.d.f. $f(x)$ can be estimated by Parzen window technique as

The work is supported by the project FLOCODS (ICA4-CT2001-10035) and Natural Sciences Foundation of China under grant No. 60121302 and No. 60175007

$$f_{h,k}(x) = \frac{C_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{x-x_i}{h}\right\|^2\right), \quad (1)$$

where $k(x)$ is a kernel function, h is the bandwidth parameter, and $C_{k,d}$ is a normalization term. Exploit the gradient operator on both sides of (1), yielding

$$\nabla f_{h,k}(x) = \frac{2C_{k,d}}{h^2 C_{g,d}} f_{h,g}(x) m_{h,g}(x), \quad (2)$$

$$f_{h,g}(x) = \frac{C_{g,d}}{nh^d} \sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right), \quad (3)$$

$$m_{h,g}(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x, \quad (4)$$

where $m_{h,g}(x)$ is called as mean shift, and $g(x) = -k'(x)$ is also a kernel. Defining the sequence $\{y_j\}$ by

$$y_{j+1} = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{y_j-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{y_j-x_i}{h}\right\|^2\right)}, \quad (5)$$

and $y_1 = x$ be the initial center of the kernel, yields

$$\nabla f_{h,k}(y_c) = 0 \quad (6)$$

according to (2) and (5), which indicates that the convergence point of $\{y_j\}$, if existing and denoted by y_c , is the local maximum of the density. The convergence condition can be referred to [5], and it's satisfied in usual case. In image processing application, for each pixel $x_i, i=1, \dots, MN$ (M and N are the number of image row and column, respectively), if combining its spatial coordinate $x_i^s = (x_{i,row}, x_{i,column})$ with its corresponding range measurements $x_i^r = (x_{i,B1}, x_{i,B2}, \dots, x_{i,Bd})$ (d is the number of bands), a joint spatial-range space is constructed. The mean shift procedure can be generalized to this joint space, where the feature point is $x = (x^s, x^r)$ and the kernel function is

$$g_{s,r}(x) = c \cdot g_s\left(\left\|\frac{x^s}{h_s}\right\|^2\right) g_r\left(\left\|\frac{x^r}{h_r}\right\|^2\right) \quad (7)$$

In our case, after PCA transform, $d = 3$, and g_s is defined as the uniform kernel

$$g_s(x) = \begin{cases} 1 & \text{if } \|x\| \leq 1 \\ 0 & \text{else} \end{cases}, \quad (8)$$

thus

$$y_{j+1} = \frac{\sum_{x_i^s \in N(y_j^s, h_s)} x_i g_r\left(\left\|\frac{y_j^r - x_i^r}{h_r}\right\|^2\right)}{\sum_{x_i^s \in N(y_j^s, h_s)} g_r\left(\left\|\frac{y_j^r - x_i^r}{h_r}\right\|^2\right)} \quad (9)$$

where $N(x^s, h_s)$ is the neighborhood of y_j^s with the radius of h_s .

One important application of mean shift procedure is the ‘‘discontinuity preserving smoothing’’ [5]. For each pixel $x = (x^s, x^r)$, initializing $y_1 = x$, computing y_j according to (9) until reaching the convergence point y_c , and then replacing the range component of x with y_c^r , we get the filtered image, where all the pixels that converge to the same point have the same range values.

Furthermore, we consider the spatial location of the convergence point. Extract the spatial component from (9), yielding

$$y_{j+1}^s = \frac{\sum_{x_i^s \in N(y_j^s, h_s)} x_i^s g_r\left(\left\|\frac{y_j^r - x_i^r}{h_r}\right\|^2\right)}{\sum_{x_i^s \in N(y_j^s, h_s)} g_r\left(\left\|\frac{y_j^r - x_i^r}{h_r}\right\|^2\right)}, \quad (10)$$

which is the weighted mean of the spatial location, and we name it as spatial mean shift. Obviously, it indicates that y_{j+1}^s has the tendency of shifting to the region where range values are homogeneous, but away the region where edges lie. For each spatial location x^s , we count the number of points that converge to x^s , and create an accumulation map which we name spatial accumulation map. In this map, generally, value 0 happens in the edge because no point converges to, 1 happens in the large homogeneous area because iteration (10) stops immediately by converging to y_1^s itself, and the big value happens in the center of the small area whose scale is less than h_s , since most of points inside the area converge identically to its center (fig. 3). Thus, the information of edges, small areas, and large homogeneous areas can be distinguished according to the index of the spatial accumulation map, and such information will be used in the following steps.

2.3. Seeded region growing

After the mean shift smoothing in the spectral domain, an initial segmentation can be well performed by means of seeded region growing. Region growing is a simple but practical method for image segmentation; it involves two critical issues, the selection of seed points and the criterion of growing or stopping, respectively.

In this paper, the seeds are selected according to the spatial accumulation map. Those points corresponding to 1-value and big-value pixels, which happen in large homogeneous regions and centers of the small areas respectively, are considered as seeds of region growing. Obviously this selection guarantees that the seed point is representative of the corresponding region in terms of the spectral measurements. In addition, in the strategy of growing, we update dynamically the mean and variance of the already grown region, and compare them with the candidate point x in the eight-neighbor. If satisfying:

$$\begin{aligned} |pc_i(x) - \mu_i(R)| < \max \{k\sigma_i(R), T_i\}, \quad (11) \\ i = 1, 2, 3 \end{aligned}$$

point x will be included into region R and be assigned the same label as the region. Here, pc_i is the i th component of the PCA transformed images; μ , σ are the mean and standard variance of the region, respectively; k , T_i are some constants, and $T_1 > T_2 > T_3$, since the variance of pc_i decreases with i increasing. However, it should be noted that there might exist some points that any labeled region can't expand to. Such points are included into the labeled surrounding region. Similarly, those very small labeled region containing less than T_{area} pixels are also eliminated in such a way. As a result, a region label map is created, but it's just an over-segmentation of the image, and some postprocessing for region removal and merging is needed.

2.4. Invalid region removal

It is known that the edge and region are dependent on each other. It means that if a region is valid, its edge in the spatial accumulation map will be remarkable. Therefore we assume that along the boundary of a region, if there exist enough edge points (i.e., zero-value points) in the spatial accumulation map, the region is regarded as valid, otherwise as invalid. With the help of the spatial accumulation map, some regions without remarkable edges will be removed.

For each labeled region R , define its boundary as,

$$Bound(R) = \{x | x \notin R \wedge \exists y (y \in N(x) \wedge y \in R)\}, \quad (12)$$

where $N(x)$ is the neighbor of x , and define the edge/boundary ratio as the number of edge points divided by the length of boundary,

$$Ratio(R) = \frac{\sum_{x \in Bound(R)} \delta(AccumMap(x))}{|Bound(R)|} \quad (13)$$

where δ is the pulse function that equals one only at zero position; $||$ means the elements number of a set. If $Ratio(R)$ is less than a threshold T_{ratio} , R would be

removed and relabelled as the surrounding region that shares the longest boundary with it.

2.5. Homogeneous region merging

After invalid region removal, there might exist some homogeneous regions, which in fact belong to the same land use class, however are labeled as different region, since the same class of land cover may present different spectral characteristic in each original band. For example [6], TM4, the near infrared band, is strongly absorbed by the water, so water bodies appear the deep dark and clear contours (fig.1.b). However, TM2, the green band, can penetrate the water deeply into 10-20m and can therefore indicate the underwater object and the turbidness of water, so water bodies appear different gray values (fig.1.a) due to turbidness or underwater objects. Thus, the river and the lake (at the center of fig.2) appear different colors, although they are both water bodies, because the PCA transformed image combines information from each original band.

Therefore a merging procedure for homogeneous regions is necessary, and it should be performed in different band according to different land use class, instead of directly in the PCA transformed image. For instance, TM4 is the best band for water body, so the merging procedure of water body will be performed in this band. Here an automatic selection method for the best band is firstly applied also based on the edge/boundary ratio of each band. We assume that the probability distribution of each region can be modeled as Gaussian distribution with the mean μ_i and the standard variance

σ_i . Then in this selected band, we determine whether mean μ_i of each rest region is in the confidence range of the land use class or not, the procedure is described in detail as follows: a) Generating the spatial accumulation maps $MapImg_i, i = 1, \dots, 7$ for each original band by means of the spatial mean shift procedure; b) Constructing a region list from the label map resulted from section 2.4; c) Popping the largest region R from the list, then computing the edge/boundary $Ratio_i(R)$ similar as (13) (just substituting $AccumMap_i$ for $AccumMap$) for all i s, and choosing the i_0 which maximizes $Ratio_i(R)$; d) For each of the rest region R' in the list, under the i_0 th band, if

$$|\mu_{i_0}(R) - \mu_{i_0}(R')| < k\sigma_{i_0}(R) \quad (14)$$

merging R' into R , i.e., assigning R' the same label as R , and deleting R' from the list; e) Performing step c to e repeatedly until the list becomes empty. As a result, a good thematic map will be created.

3. EXPERIMENTAL RESULTS

Our input data are sub-image of Landsat TM multi-spectral image of the Red River area in Vietnam, with the size of 512 by 512, and resolution of 30m/pixel (fig.1). Fig.2 shows the false color image composed of three channels after band reduction with PCA. In the step of the mean shift procedure, we define $g_r(x)$ as the normal kernel, and results are demonstrated in fig.3, where (a) is the filtered result in the spectral domain, (b) is the spatial accumulation map in the spatial domain, and (c) shows the 0-value, and 1-value points of the spatial accumulation map that indicate the information of edges and large homogenous areas. Fig.4 shows the over segmentation result with seeded region growing method. Fig.5 is the final classification result after region removal and merging. The resulted thematic map consists of four classes that correspond to water body, urban area, soil land and sand beach respectively.

4. CONCLUSIONS

In this paper, a new unsupervised method for land use classification is proposed. In the pre-processing stage, the mean shift filter is extended to spatial domain, a spatial accumulation map as well as a smoothed color image is therefore generated. And in the post processing stage, the edge information from the spatial accumulation map is incorporated to remove invalid regions, and to merge

homogeneous regions in the best band according to different land cover class. Experimental results on Landsat TM images validate the efficiency of the proposed method.

In the future work, more powerful tool or criterion, such as graph cuts, will be considered for region removal and merging. And fusion of texture information also expects to achieve a better result.

5. REFERENCES

- [1] Bradt Tso and Paul M. Mather, *Classification Methods for Remotely Sensed Data*, London: Taylor & Francis Inc, 2001.
- [2] Carolyn Evans, et al, "Segmentation Multispectral Landsat TM Images into Fiel Units", *IEEE Trans. Geoscience and Remote Sensing*, vol. 40, no. 5, pp. 1054-1064, May 2002.
- [3] Bassel Solaiman, Marie-Catherine Mouchot and Alain Hillion, "An Information Fusion Method for Multispectral Image Classification Postprocessing", *IEEE Trans. Geoscience and Remote Sensing*, vol. 36, no. 2, pp. 395-406, March 1998.
- [4] Gunnar Jakob Briem, et al, "Multiple Classifiers Application to Multisource Remote Sensing Data", *IEEE Trans. Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2291-2299; October 2002.
- [5] Dorin Comaniciu and peter Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, May 2002.
- [6] Yingshi Zhao, *Principles and Methods for Remote Sensing Application and Analysis*, Beijing: Science Press, 2003.

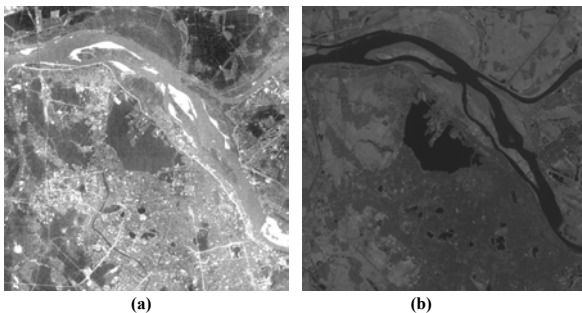


Fig. 1. Input Images (Landsat5 TM, 512 by 512, 30m/pixel, in the Red River of Vietnam) (a) band 2, (b) band 4

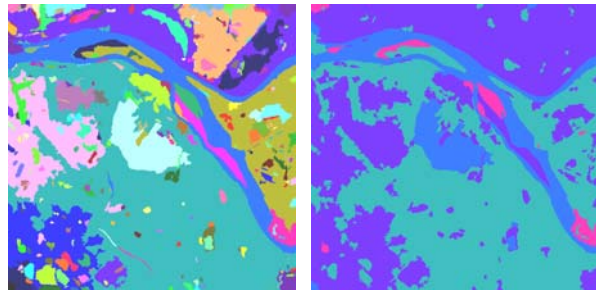


Fig. 4 Over-segmentation result, $(T_1, T_2, T_3, T_{area}) = (30, 27, 25, 20)$

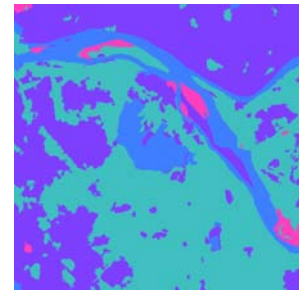


Fig. 5. Final thematic map, water (blue), urban (green), soil land (purple), sand beach (red) $(T_{ratio} = 0.6, k = 1.5)$

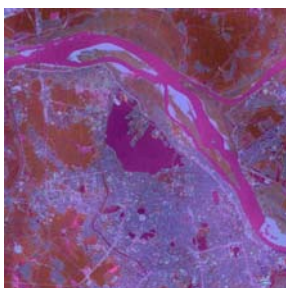


Fig. 2. PCA transformed false color image

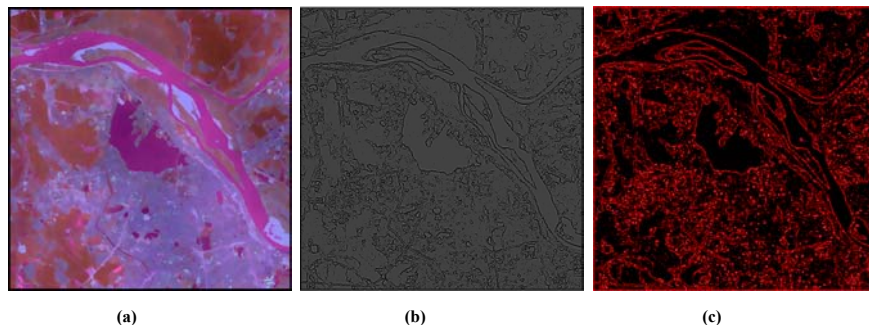


Fig. 3. Mean shift procedure $(h_s = 4, h_r = 16)$ (a) mean shift smoothing result, (b) spatial accumulation map, (c) 0-value points (red), 1-value points (black) and big-value points (white)