

# Watershed Segmentation of Remotely Sensed Images Based on a Supervised Fuzzy Pixel Classification

Sébastien Derivaux, Sébastien Lefèvre, Cédric Wemmert and Jerzy J. Korczak  
LSIIT - CNRS - Université Louis Pasteur - UMR 7005  
BP 10413  
67412 Illkirch FRANCE  
Email: {derivaux,lefevre,wemmert,jjk}@lsiit.u-strasbg.fr

**Abstract**— Remotely sensed images are more and more precise (spatial resolution under 1 meter). For these images, objects of interest contains several pixels. Generally a segmentation method is used to cluster pixels that belong to the same objects before classification. The quality of such a segmentation method is crucial to achieve good classification results. In this paper, a new segmentation method is proposed which aims to improve the classical watershed segmentation method based on multispectral gradient. The proposed method uses some labeled samples with classes of interest to induce a new dissimilarity between pixels which defines a new representation space to be used.

## I. INTRODUCTION

Automatic interpretation of remotely sensed images becomes an increasingly active domain. Sensors are now able to get images with a very high resolution (i.e. 1 meter resolution). This increasing precision generates a significant amount of data. Technics of automatic interpretation of remote sensing images which assign to each pixel a land cover class become essential to process all data in reasonable time. Due to the high complexity of these images, methods considering each pixel independently have shown their limits with very high resolution sensors. Considering regions instead of pixels seems the way to solve this problem [1]. These regions are created by a segmentation process, generally based on region growing [2], [3] or watershed [4]–[7]. A recent survey with these two methods and others has been made by Carleer et al. [8].

A watershed limitation is over-segmentation, i.e. it produces a lot of regions for each object of interest. Many methods have been developed to overcome this problem. Some are specific to watershed like gradient thresholding [6], [7], dynamics [4] or volume extinction [9]. Others are more general region merging methods [5], [7]. All these methods need a threshold to be given in order to stop segmentation simplification. A threshold is very difficult to choose and depends on the image.

In this paper, a watershed-based method is proposed. To reduce over-segmentation, a space transformation is made before applying watershed. We assume some knowledge about the processed data. This knowledge consists in definition of the interesting classes and in some pixel examples for each class. As segmentation is often used for supervised classification, this assumption is not too reductive. Labeled pixels are used to

induce a fuzzy pixel-based classifier. The new space describes, for each pixel, its membership for each class given by the classifier.

The first part of this paper describes the classifier induction. Secondly, the watershed transformation is explained. Then, the proposed algorithm is evaluated and compared with a classical watershed algorithm and results are discussed. Finally a conclusion and some research directions are presented.

## II. FUZZY SUPERVISED CLASSIFICATION

A supervised classification algorithm induces a model for classes prediction. To do so, it needs labeled samples for each class the user wants to find in data. If the user gives samples of roads, roofs and vegetation, the classifier will assign to each unlabeled pixel one of these three classes.

Fuzzy classification means that each pixel is not associated with a class label but with a memberships vector. For each class  $i$  a number between  $p_{xi}$  is given which denotes how much the classifier thinks the pixel  $x$  belongs to the class  $i$ . Two relations must be satisfied :

$$0 \leq p_{xi} \leq 1 \quad (1)$$

$$\sum_i p_{xi} = 1 \quad \forall x \quad (2)$$

If a pixel  $x$  has memberships of  $p_{xi} = 1$  for a given  $i$  and  $p_{xj} = 0 \quad \forall j \neq i$ , the classifier is very confident on the classification. On the other hand, a pixel may have a membership  $p_{xi} > 0$  for more than one class. This may indicate that the pixel belongs to a class which is not taken into account by the classifier or that given attributes for the pixel are not sufficient to discriminate these classes.

How can membership information be better than assigning the pixel to the class with the highest membership? The benefit is that we have a finer granularity of information. For example some roofs have a spectral response very similar to road so that a fuzzy classification will give values around 50% membership for road and roof. A such roof has some pixels classified as roof and other as road. The roof will be interpreted as small segments of roads and roofs. In the membership space it is

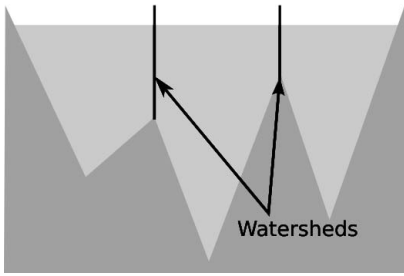


Fig. 1. The watershed transformation

quite clear that the roof is only one object. If this object is extracted it becomes available to find its correct class with its form.

To obtain fuzzy memberships, a five-nearest neighbours classifier [10] is used. For each pixel, the five nearest labeled pixels in feature space are selected. Each of these pixels increases the membership for its class. Weight of each selected pixel is the inverse of the distance between himself and the pixel to classify in the feature space. If  $x$  denotes the current pixel,  $x_n$  for  $n = 1 \dots 5$  the five nearest pixels in feature space and  $d(x, x_n)$  the distance between  $x$  and  $x_n$ ,  $p_{xi}$  are obtained by :

$$p_{xi} = \frac{1}{W} \sum_{n=0}^5 w_{ni} \quad (3)$$

$$\text{where } w_{ni} = \begin{cases} d(x, x_n)^{-1} & \text{if } x_m \text{ is labeled with class } i \\ 0 & \text{otherwise} \end{cases}$$

$$\text{and } W = \sum_{n=0}^5 \sum_i w_{ni}$$

The last criterion involved is the feature space used to compute pixels memberships. In this paper, feature space is composed of spectral values of pixels. Nevertheless, nothing prevents us to use a more elaborated feature space with texture indices or other extractable attributes.

The memberships obtained by this algorithm can be used as a new representation. In this paper, the watershed transformation is used to segment the image using this new representation.

### III. WATERSHED

The watershed transformation [11] is the main segmentation method within mathematical morphology framework. A gradient image is used to represent an elevation map. This elevation is flooded from its minima. Where two catchment basins encounter, a dam is created. An example is given at figure 1.

To obtain a gradient image, a morphological gradient [12] is performed on each channel. The definition of a morphological gradient is  $G(A) = \delta_B(A) - \varepsilon_B(A)$  where  $\delta_B$  and  $\varepsilon_B$  are the erosion and dilation with a structuring element  $B$  (in this case a square of  $3 \times 3$  pixels) and  $A$  the image. The final gradient image is obtained by a combination of the different

channels using an euclidian norm. If the gradient of the  $i^{th}$  channel is noted  $G^i$  and  $C$  the number of channels, the final gradient is :

$$G = \sqrt{\sum_{i=0}^C G_i^2} \quad (4)$$

In this paper, two segmentation methods based on the watershed transformation are used :

- **spectral-based watershed** : This is the classical algorithm where each pixel is represented by its spectral values. A multi-spectral gradient is computed on which the watershed transformation is applied.
- **memberships-based watershed** : This is the new proposed algorithm where pixels are represented by the memberships to each class of interest. The gradient is computed from this representation on which the watershed transformation is applied.

## IV. RESULTS

### A. Data description

To evaluate the proposed method, a Quickbird image of Strasbourg, France has been used. The Quickbird sensors give one panchromatic channel with a resolution of 0.7 meters and 3 color channels of a resolution of 2.8 meters. The panchromatic channel is merged with the 3 other channels using the method proposed by Puissant et al. [13] to obtain a four channel image with a resolution of 0.7 meters. The image size is 900x900 pixels and the spectral resolution is 8 bits for each channel. The image can be seen in figure 2.

To reduce noise and heterogeneity of objects, a median filter is applied on each channel. This kind of filter preserves edges. The filter consists of a sliding window encompassing an odd number of pixels (a square of 5 pixels length in this example). The center pixel is replaced by the median value of the pixels within the window for each channel.

For fuzzy supervised classification learning and segmentation evaluation, some regions are labeled by an expert as shown in figure 3. The *ll* zone corresponds to learning regions. It contains 9 houses and 6 towers labeled as roofs, two regions labeled as vegetation and two segments of roads. The *e1* and *e2* zones are used as two evaluation sets, the first contains houses, and the second towers. Evaluation zones are completely classified as it is needed for the segmentation evaluation. Pixels of these zones that have no label mean that the expert can not give a label.

### B. Evaluation criterions

Segmentation evaluation is not quite defined. This paper uses empirical discrepancy evaluation methods similar that those in Carleer et al [8]. We will use the two evaluation zones to estimate segmentation quality. The first criterion measure over-segmentation (*OV*) for each class:

$$OV_i = \frac{N_{Si}}{N_{Ri}} \quad (5)$$



Fig. 2. Source image



Fig. 3. Labeled regions for learning (l1) and evaluation (e1 and e2)

where  $i$  is the considered class,  $N_{S_i}$  is the number of segmented regions which contain at least one pixel of the  $i^{th}$  class and  $N_{R_i}$  is the number of expert regions for the  $i^{th}$  class. After over-segmentation, under-segmentation must be evaluated. Using only the over-segmentation criterion, a segmentation which produces only one segment for the whole image will be the best so there is a need for another criterion. Under-segmentation occurs when a segment contains pixels from two regions of different classes, this reduces the maximum accuracy of an image classification. Theoric maximum accuracy ( $TMA$ ) of a subsequent classifier represents the under-segmentation. By exemple, if the  $TMA$  is 63%, the classification of the image could not be better than 63%. For this criterion, the label of a segment is the class that has more referenced pixels in this segment. A confusion matrix  $C$  can be extracted from this classification step using  $N_{Ref}$  classes.  $TMA$  is defined as:

$$TMA = \frac{\sum_{i=0}^{N_{Ref}} C_{ii}}{\sum_{i=0}^{N_{Ref}} \sum_{j=0}^{N_{Ref}} C_{ij}} \quad (6)$$

The last evaluation criterion is the empiric accuracy ( $EA$ ). A five-nearest neighbours classifier is used on segments produced by segmentation. Each segment is represented by its average spectral value in each channel. When a segment is over some reference pixels, we label this segment with the class that has the greatest number of pixels. These labeled segments are used to induce the classifier. This classification procedure is quite simplistic, many interesting attributes that can be extracted from segments are not used, but is sufficient to show the benefits of the proposed method. Results are filtered by a sliding window encompassing a square of 5 pixels length, the center pixel is labeled by the more frequent label in the window. This filtering step improves results up to 1% for region-based classifications and up to 4% for pixel-based classifications.

### C. Discussion

A visual comparison is given on figure 4. Membership-based watershed clearly produces less regions for vegetation and roofs, but some portions of roads (left) are more segmented than with spectral-based watershed methods.

Results for the two segmentations and a pixel-based classification are given in table I. The pixel classification simply uses the membership space and assigns the pixel to the class which has the greatest membership.

The first result of interest is that memberships-based watershed gives a twice fewer number of segments than spectral-based watershed. Looking more closely on over-segmentation per class, memberships-based watershed reduces drastically vegetation over-segmentation, is better for roofs segmentation and worse for road segmentation. Under-segmentation is quite low for the two methods with a theoretical maximal accuracy around 99%.

TABLE I  
RESULTS FOR SEGMENTATION

Methods	Nb of segments	Zone	OV [roofs, vegeta- tion, road]	TMA	EA
Spectral	23439	Houses (e1)	[17.93, 29.90, <b>24.66</b> ]	99.1%	71.5%
		Towers (e2)	[29.09, 24.88, <b>26.5</b> ]	99.6%	58.9%
Membership	11886	Houses (e1)	[ <b>13.66</b> , <b>2.63</b> , 31.66]	98.7%	79.2%
		Towers (e2)	[ <b>24.27</b> , <b>2.77</b> , 38.25]	99.3%	<b>73.4%</b>
Pixel	na	Houses (e1)	na	na	<b>79.3%</b>
		Towers (e2)	na	na	68.0%

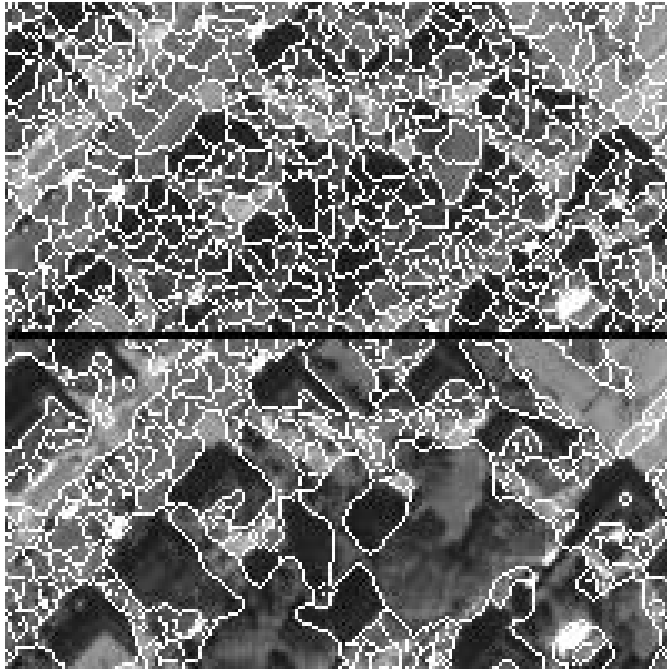


Fig. 4. Watershed on spectral values (top) and memberships (bottom)

For the empirical accuracy criterion, memberships-based watershed shows the best performances. Even with a so simplistic classification scheme (i.e. just an average of spectral values) it performs better (2.6% on average) than a pixel-based classification. Spectral-based watershed give poor results (5.5% worse on average than membership-based).

## V. CONCLUSION

The object-oriented paradigm for remotely sensed images interpretation heavily depends on the segmentation algorithm quality.

This paper proposes to segment an image not using classical attributes, but a fuzzy pixel-based classification. Experiments show that performing a watershed on a membership space is less prone to oversegmentation. A basic region-based classification show that the proposed algorithm lead to subsequent better classification results.

Nevertheless some questions remain open. First, what are the best methods to obtain membership values? Which classes have to be considered? What is the influence of the learning

set? Finally, the presented algorithm has some complementarity with the classical watershed which is still better for some classes. One may want to combine the results of these two algorithms to obtain a better segmentation.

## REFERENCES

- [1] M. Baatz and A. Schape, "Multiresolution segmentation: an optimization approach for high quality multi-scale images segmentation," in *Angewandte Geographische Informationsverarbeitung XI*, 2000.
- [2] M. Mueller, K. Segl, and H. Kaufmann, "Edge- and region-based segmentation technique for the extraction of large, man-made objects in high-resolution satellite imagery," *Pattern Recognition*, vol. 37, no. 8, pp. 1619–1628, 2004.
- [3] J. Tilton, "Image segmentation by region growing and spectral clustering with a natural convergence criterion," in *IEEE Geoscience and Remote Sensing Symposium Proceedings*, 1998, pp. 1766–1768.
- [4] N. Malpica, J. Ortuno, and A. Santos, "A multichannel watershed-based algorithm for supervised texture segmentation," *Pattern Recognition Letters*, vol. 9–10, no. 24, pp. 1545–1554, 2003.
- [5] G. Castilla, "Size-constrained region merging: A new tool to derive basic landcover unit from remote sensing imagery," in *Proceedings of ESA-EUSC 2004*, 2004.
- [6] Q. Chen, C. Zhou, J. Luo, and D. Ming, "Fast segmentation of high-resolution satellite images using watershed transform combined with an efficient region merging approach," in *IWCIA*, ser. Lecture Notes in Computer Science, R. Klette and J. D. Zunic, Eds., vol. 3322. Springer, 2004, pp. 621–630.
- [7] K. Haris, S. N. Efstradiadis, N. Maglaveras, and A. K. Katsaggelos, "Hybrid image segmentation using watersheds and fast region merging," *IEEE Transaction On Image Processing*, vol. 7, no. 12, pp. 1684–1699, 12 1998.
- [8] A. P. Carleer, O. Debeir, and E. Wolff, "Assessment of very high spatial resolution satellite image segmentations," *Photogrammetric Engineering and Remote Sensing*, vol. 71, no. 11, pp. 1285–1294, 11 2005.
- [9] B. Marcotegui and S. Beucher, "Fast implementation of waterfall based on graphs," in *International Symposium on Mathematical Morphology*, ser. Computational Imaging and Vision., C. Ronse, L. Najman, and E. Decencière, Eds., vol. 30. Dordrecht: Springer-Verlag, 2005, pp. 177–186.
- [10] D. W. Aha, D. F. Kibler, and M. K. Albert, "Instance-based learning algorithms," *Machine Learning*, vol. 6, pp. 37–66, 1991.
- [11] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Pattern Recognition and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, 1991.
- [12] P. Soille, *Morphological Image Analysis*, 2nd ed. Springer-Verlag, 2003.
- [13] A. Puissant, T. Ranchin, C. Weber, and A. Serradj, "Fusion of quickbird ms and pan data for urban studies," in *EARSel Symposium*, 2003, pp. 77–83.