Local urban sprawl accuracy from image segmentation uncertainties simulation

Abstract

Urban sprawl can be monitored using remote sensing. One method is to segment and classify the images to obtain polygons of impervious areas and then derive urban patches. We evaluated the geometric accuracy of the impervious polygons using the boundary distance distribution signature (BDDS) and investigated their propagation to spatial indicators of urban sprawl indicators at municipality scale, south of France, from impervious polygons uncertainties simulation.

Keywords: Geometric uncertainties, Boundary distance distribution signature, Polygon simulation, Error propagation, Urban sprawl.

1. Introduction

Urban sprawl is a crucial issue for land use planners and should therefore be monitored regularly. This can effectively be achieved using remote sensing. Object-based image analysis is especially suited for mapping impervious areas. The segmentation-classification process produces vector maps, which can directly be used to compute spatial indicators of urban sprawl. The impervious polygons, however, suffer from geometric uncertainties due to the pre-processing of the remote sensing image, to the segmentation errors and to the object misclassifications. In this case study in southern France, we focus on the uncertainties of the segmentation-classification step, and investigate how they propagate into the indicator maps using Monte Carlo simulations.

2. Materials and methods

2.1. Study area and data

Languedoc-Roussillon is a French "region" covering 27'376 km², located along the Mediterranean coast. Sixty-five multispectral RapidEye images with 5 m spatial resolution were acquired in spring - summer 2009. They were segmented and classified using eCognition by Dupuy *et al.* (2012). Their impervious polygons, covering 1913 km^2 (including roads), were used as the inputs for this study.

2.2. Segmentation accuracy

For 75 municipalities, covering a total area of 1'237 km^2, an external operator digitized reference impervious polygons on the RapidEye images, ignoring roads. This reference dataset (R, 4'870 polygons) was used to evaluate the accuracy of the segmented dataset (S, 5'076 polygons), using:

1) the confusion matrix of S versus R, and

2) the boundary distance distribution signature (BDDS, Huang and Dom, 1995), which is the histogram of the minimum distance of each polygon vertex of one dataset

to the polygon vertices of the other dataset. This measure is not symmetrical, and two measures are therefore necessary:

$\text{BDDS}_{S,R}$: histogram of $d(S,R) = \min_{r \in R} d(s,r), \forall s \in S$ (1)

$\text{BDDS}_{R,S}$: histogram of $d(R,S) = \min_{s \in S} d(r,s), \forall r \in R$ (2)

where *d* is the Euclidian distance, and *r* and *s* are vertices of R and S.

2.3. Polygon Monte Carlo simulations

Using the three metrics above, equiprobable sets of impervious polygons were simulated based on S. The vertices of each polygon were moved in a random direction by a distance randomly drawn from $\text{BDDS}_{S,R}$. Random fields were used to account for spatial correlations between neighbor vertices. Finally under- and over-segmentation were added to randomly selected polygons using stochastic tessellation, taking care to match the measured overlapping area between S and R. A total of 1000 datasets were simulated in a Monte Carlo set up.

2.4. Impervious patch and spatial indicators

For each simulated dataset, an impervious patch was built by a closing procedure, using a radius of 50 m, according to the methodology defined by Dupuy *et al.* (2012), that is: a dilation (buffer with positive radius) followed by an erosion (buffer with negative radius). This operation connects impervious polygons that are closer than 100 m into larger polygons. The impervious patch was then used to calculate the impervious spread coefficient (I_{spread}) of Balestrat (2011), which was selected as an example. It is calculated as:

$$I_{spread} = \frac{S_{>3ha}}{S_{\leq 3ha}} \quad (3)$$

where $S_{>3ha}$ (resp. $S_{<3ha}$) are the cumulated areas of the patches larger (resp. smaller) than 3 hectares (30'000 m^2).

3. Results and discussion

3.1. Segmentation accuracy

The confusion matrix in Table 1 shows that 40.04 km^2 of the impervious area of the 75 municipalities was correctly segmented. This represents 72% of the total impervious area in R (producer accuracy), and 79% of the total impervious area in S (user accuracy). The error rate for the user is therefore 21%.

As can be seen on Figure 1, $\text{BDDS}_{S,R}$ and $\text{BDDS}_{R,S}$ have similar profiles starting with high frequencies at short distances, and rapid decrease towards longer distances. As expected, the number of vertices is much lower in R than in S. This difference is intrinsic to the production of the dataset, the manual approach yielding much fewer

vertices than the automatic segmentation. The two histograms, however show similar Poisson-like distributions.

High producer and user accuracies and high frequencies at short distances indicate that S matches R reasonably well. Many commission errors are due to confusion of bare soil with impervious areas.

Table 1: Confusion matrix of the OBIA impervious classification (S) versus the reference dataset (R). The areas are indicated in km^2.

		S			
		Impervious	Non-impervious	Sum	Producer accuracy
R	Impervious	40.04	15.39	55.43	72%
	Non-impervious	10.73	1 171.06	1 181.79	99%
	Sum	50.77	1 186.45	1 237.22	
	User accuracy	79%	99%		

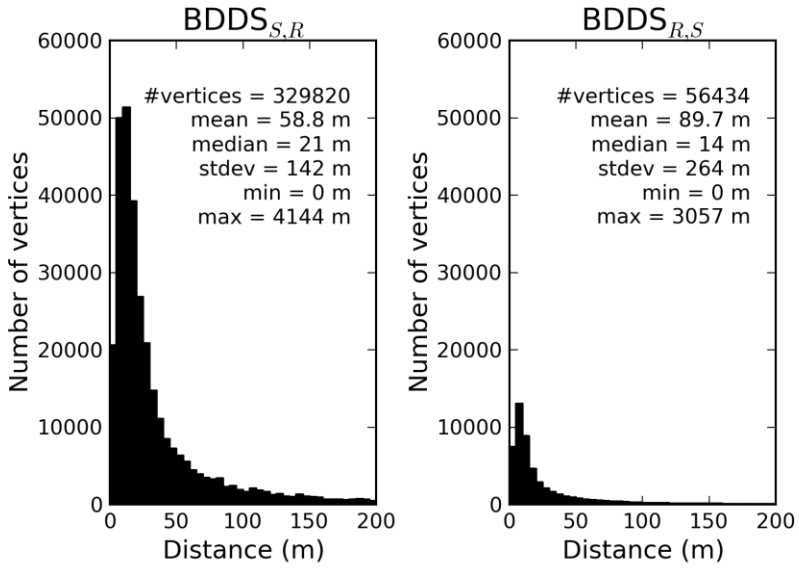

Figure 1: The two BDDS for the 75 municipalities, restricted to the distance range 0 - 200 m. The maximum distance value is indicated in the plot text.

3.2. Polygon simulations

Starting from the S dataset, we simulated 1'000 segmentations, using the accuracy measures above. The vertices of the polygons in S were moved in a random direction, by a distance randomly drawn from $BDDS_{S,R}$, accounting for spatial correlations between neighboring vertices of the same polygon using random fields. Segmentation errors (over/under-segmentation and omission/commission) were implemented as detected on the analysis of $BDDS_{S,R}$ and $BDDS_{R,S}$, and using constraints to match the user and producer accuracy values of Table 1.

3.3. Impervious patch and spatial indicators

For each simulated segmentation, the impervious patch and the map of I_{spread} were produced. Based on the 1'000 I_{spread} maps, an accuracy map as well as confidence interval maps were produced, to accompany the original I_{spread} map (based on the S segmentation of Dupuy *et al.*, 2012).

3. Conclusion

Realistic simulation of the segmentation errors, the Monte Carlo analysis allowed producing accuracy and confidence interval maps for the I_{spread} indicator. These maps will be useful for the local authorities.

References

Balestrat, M. (2011), *Système d'indicateurs spatialisés pour la gouvernance territoriale : Application à l'occupation des sols en zone périurbaine languedocienne*. PhD thesis, Université Montpellier III and Cemagref, France.

Dupuy, S., Barbe, E., Balestrat, M. (2012), "An Object-Based Image Analysis Method for Monitoring Land Conversion by Artificial Sprawl Use of RapidEye and IRS Data". *Remote Sensing*, Vol. 4(12):404–423.

Huang, Q., Dom, B. (1995), "Quantitative methods of evaluating image segmentation". *International Conference on Image Processing*, Washington D.C., USA, pp. 53–56.