SEGMENTATION OF VERY HIGH SPATIAL RESOLUTION SATELLITE IMAGES IN URBAN AREAS FOR SEGMENTS-BASED CLASSIFICATION

E. Frauman, E. Wolff

Institut de Gestion de l'Environnement et d'Aménagement du Territoire, Université Libre de Bruxelles,
50 av. F. Roosevelt, CP 130/02, 1050 Brussels, Belgium – evelyne.frauman@ulb.ac.be – ewolff@ulb.ac.be

KEY WORDS: Urban remote sensing, Satellite Image, Very High Spatial Resolution, Segmentation, Image Object, Texture, eCognition Software, Scale Parameter.

ABSTRACT:

High spatial Resolution (VHR) satellite images offer a great potential for the extraction of information for urban areas. In order to derive useful thematic maps from VHR satellite images, other approaches than the standard pixel by pixel mapping are needed. One way to avoid the salt and pepper effect can be to divide the image into homogeneous regions prior to classification instead of classifying individual pixels. These so called segments, or objects, do not necessarily have any meaning and can be considered as image primitives. Once they are created, they can entirely be assigned to the land-cover / land-use classes by any classifier. Multi-resolution image segmentation in the software eCognition is a region growing technique starting from objects with the size of one pixel. It subsequently merges adjacent image objects into bigger ones with a procedure that minimizes the weighted heterogeneity criterion of the newly created image objects. The heterogeneity criterion is set by the user before starting the segmentation process and should be tightly linked to the spatial structure of the image. If a link can be established between the size of the objects contained in the image and the heterogeneity criterion, a users friendly rule could be derived that would allow any user to get over the time consuming stage of finding the best segmentation before any other image analysis. To find that rule, we test the hypothesis that a link between the image structure and the heterogeneity criterion exists thanks to synthesis images. These images are built up from objects of equal sizes, at different resolutions.

1. INTRODUCTION

1.1 Motivations

The accessibility of the commercial very high resolution satellites increases the amount of information on land cover at local to national scales (Aplin et al., 1999). These data provide amazing details of the earth surface but information extraction using computer-assisted classification techniques appear to be more complex (Carleer et al., 2005). A major drawback of pixel-based classification approaches is that when they are applied on urban areas they often produce thematic maps that lack spatial coherence because of spectral heterogeneity and spatial variance in the image. To circumvent the problem of structural clutter the image can be divided into regions of similar pixels prior to classification. These so-called image segments do not necessarily have any semantic meaning and can be considered as image primitives. Once they are created, they can be wholly labelled to a land-cover class by any classifier. Many techniques of image segmentation have been developed (Pal & Pal, 1993). While they all address the problem of the salt-and-pepper effect, many of these segmentation techniques suffer from significant drawbacks (e.g. over and under segmentation or not being useful at all scales) which make them less suitable for urban land-use/land-cover classification (Carleer et al., 2004). A novel approach is multi-resolution object-oriented image analysis, which uses image segmentation to homogenize the spectral variability within segments and performs a nearest neighbour classification based on these segments, also called regions (Baatz & Schäpe, 2000). With this approach, segments not only have spectral properties but also region-based metrics as shape, texture, structure, size and context. Some successful land-cover or land-use classifications in urban areas have been obtained with this approach (van der Sande et al. 2003, Burnett and Blaschke, 2003). The segmentation technique is a region-growing procedure starting at each point in the image with one pixel objects and merging these image objects into bigger ones throughout a pair-wise clustering process. The merging procedure is based on three concepts: colour, smoothness and compactness (Carleer et al., 2005). These criteria are combined in one parameter defined as within-segments heterogeneity. The balance at which these criteria are applied depends on the desired output (Thomas et al., 2003).

$$f = w \cdot h_{Color} + (1 - w) \cdot h_{Shape} \qquad (1)$$

Where f is the heterogeneity criterion, also called **scale factor**, w is the user defined weight for colour with $0 = w = 1$, h_{colour} is the colour criterion and h_{shape} is the shape criterion.

$$h_{Shape} = w_{Cmpt} \cdot h_{Cmpt} + (1 - w_{Cmpt}) \cdot h_{smooth} \qquad (2)$$

The shape criterion consists in two parameters: the smoothness and the compactness criterions. They have to be mixed using the user defined weights with 0 = w_{cmpt} = 1 being the user defined weight for the compactness criterion (Definiens Imaging, 2003).

$$h_{Smooth} = n_{Merge} \cdot \frac{l_{Merge}}{b_{Merge}} - \left(n_{Obj1} \cdot \frac{l_{Obj1}}{b_{Obj1}} + n_{Obj2} \cdot \frac{l_{Obj2}}{b_{Obj2}}\right) \quad (3)$$

$$h_{Cmpt} = n_{Merge} \cdot \frac{l_{Merge}}{\sqrt{n_{Merge}}} - \left(n_{Obj1} \cdot \frac{l_{Obj1}}{\sqrt{n_{Obj1}}} + n_{Obj2} \cdot \frac{l_{Obj2}}{\sqrt{n_{Obj2}}}\right) \quad (4)$$

Where n is the object size, l the object perimeter and b the perimeter of the bounding box.

If the smallest growth exceeds a heterogeneity tolerance defined by the user, the process stops (Brunett and Blaschke, 2003, van der Sande et al., 2003). This procedure was developed by Definiens Imaging as the eCognition software.

1.2 Aims

The heterogeneity criterion has to be threshold by the user prior to the segmentation. This is done by choosing a s**cale factor** (which is equivalent to the threshold) and by fixing the weights of the colour and shape criterions, and the smoothness and compactness criterions. Woodcock and Stahler (1987) showed the relation between the local variance of a digital image and the resolution size. By analogy, we believe that a rule giving the heterogeneity criterion as a function of the image structure probably exists. Indeed, preliminary studies showed how the classification assessment of a digital image changes as the scale factor increases in the case of real VHR images in urban areas (Van de Voorde et al., 2004). If a link can be established between the size of the objects contained in the image and the heterogeneity criterion, a users friendly rule could be derived that would allow any user to get over the time consuming stage of finding the best segmentation before any other image analysis. The colour and shape criterions weights are left at their default values as a first step.

2. METHODS

2.1 Synthesis images

We decided to work first on synthesis images. In view of the equation of the heterogeneity criterion (eq (1)) and as we decided to keep the colour and shape criterion at their default values, we chose to represent the best case possible: simple shape (squares) with high contrast (black and white on a 256 grey levels scale). We built images containing repeated squares of the same size from 5x5 pixels (size of the smallest house in the real QuickBird image used in this study) to 100x100 pixels (size of the biggest building) by steps of 5 pixels. We made two more images made of squares of 12x12 and 65x65 pixels as these were the mean sizes of the houses and the buildings in the real QuickBird image of the centre of Brussels that we worked with. The synthesis images should allow us to have the threshold of the heterogeneity criterion, the so-called scale parameter or scale factor, as a function of the size of the objects (in pixels). We will then use the real image as a test of these results.
The "best" threshold for the heterogeneity criterion is the first one for which the objects of the image are not over-segmented anymore. It is the first scale parameter at which the number of regions given by segmentation equals the number of objects of the image.

2.2 Real images

The real image we used in this study is part of the centre of the town of Brussels, Belgium from a QuickBird satellite image. It has been taken in 2003, and its spatial resolution is of 0.71m for the panchromatic layer and 2.84m for the multi-spectral ones. The image is divided in two sites with different morphologies by the railway. At the West of the railway, the landscape is mostly industrial with large and high buildings. At the Eastern part of the image the landscape is more residential with dense housing and annexes and garden.

Figure 1: QuickBird Image of the centre of Brussels (2003)
© Digital Globe

2.3 Segmentation tests

The synthesis images have all been submitted to the same treatments: segmentations in the eCognition software, then exportation of the segments with, as attribute, the mean of the spectral information for the first layer. We were then able to calculate the number of segments that were built-up for the whole image and, knowing the number of real objects, i.e. squares, that were contained in the image we could know exactly when the process was not over-segmenting anymore. In fact, as the contrast is high and the shapes simple, the under-segmentation was almost never occurring. An important step is then when the software does not divide each object in more than one segment.

In the case of the QuickBird real satellite image, to find the best scale parameter we completed the classification and we assessed it. In this case the parameters to be compared were the classification results: the overall accuracy and the Kappa parameter. We are not comparing the number of segments and the number of objects anymore but there were no possibilities to do so.

3. RESULTS AND DISCUSSIONS

3.1 Synthesis images

The result of the first experiment, which purpose was to find how the threshold of the heterogeneity criterion, the so-called scale parameter or scale factor varied with the size of the objects in the image is given in figure 2. The size of the side of the square objects is given in abscissa, i.e. the square root of the size of the object. As predicted the scale parameter increases when the size of the object increases almost in a linear way. The plateaus can be explained by the fact that we looked to integer values for the scale parameter but this latter may also take decimal values. The plateaus are probably smoothed when decimal scale parameters are used.

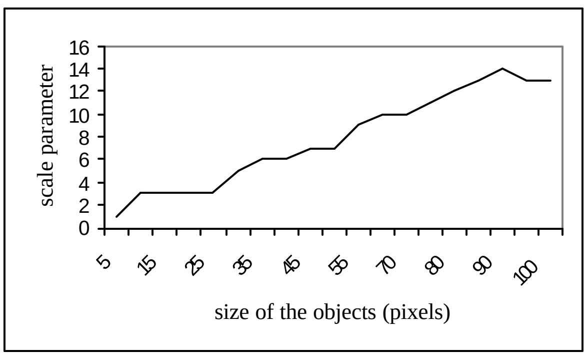

Figure 2: variation of the scale factor as a function of the size of the objects in the image.

This kind of curve could be used to predefine valour for the threshold knowing the mean size of the objects in the image. But to be more realistic we should include noise, more complex shapes, a scale for the contrasts, etc.

In a second step, we worked with images containing objects we wanted to be as close as possible as real objects: houses and buildings for example.

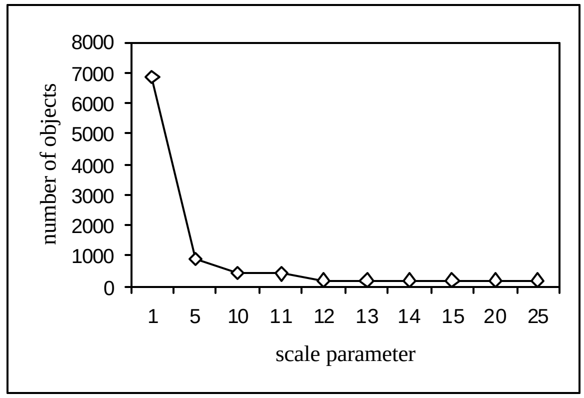

Figure 3: Number of objects produced by segmentation of the synthesis image "houses" as a function of the scale parameter. There are 221 objects in this image.

We used measures of the buildings and houses of an image of the centre of Brussels. At the resolution of 0.71 m (Panchromatic layer) the mean size of the houses was a square of 12 pixels side and the mean size of the buildings was a square with 65 pixels sides. In this case we include noise. This is supposed to give us a situation closer to reality. And this is also done to see if there is any difference in the coming of the best scale factor, i.e. the scale parameter at which there is no over-segmentation. The figure 3 gives one of the results: the objects are not over-segmented anymore at a threshold (i.e. scale parameter) of 12 (versus about 5 in the case without noise) for the houses, 45 (versus about 10 for the noiseless case) for the buildings. In this case the number of 221 objects, which is the number of objects in the synthesis image, is obtained at a scale factor of 12 and under-segmentation never occurs. This is due to the simplicity of the shapes and to the high contrast. These results could explain the threshold we obtained in the study of the real image.

3.2 Real images

To find the best threshold for the heterogeneity criterion in this case we had to complete the segmentation and the classification so that the parameters to be compared are the kappa parameter or the overall accuracy. In the figure 4 we see that there are two maximums of these parameters: one at a scale factor of 5 and another at a scale factor of 35. These two numbers are low compared to the awaited ones, 12 and 45, but they remain in the same order.

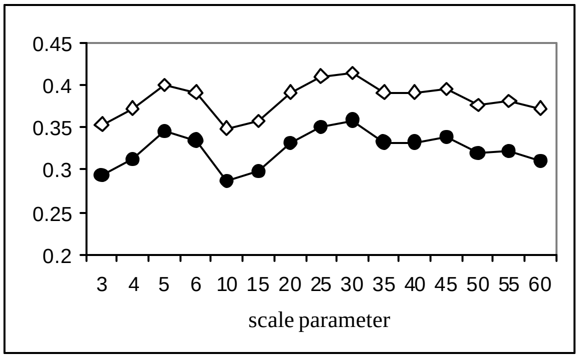

Figure 4: overall accuracy (white) and overall Kappa parameter (black) obtained by classifying the QuickBird image of Brussels using various thresholds for the heterogeneity criterion.

The difference might be explained by the fact that in this case what we look at are the kappa parameter, and overall accuracies, not the number of objects anymore. It is more than probable that a better classification is obtained when the objects are somewhat over-segmented.

The flatten shape of the maxima is due to the fact that the objects in the image are not as simple as the objects in the synthesis image. Houses and buildings are not simple squares. These are not the only objects contained in the real image. In fact, there are trees, roads, etc. But the two maximums are still very significant as they might be the proof that there is a clear link between the size of the objects contained in an image and the scale factor to be chosen.

4. CONCLUSIONS

As awaited there is a link between the size of the object and the threshold for the heterogeneity criterion, the so-called scale-parameter. We could also see that the introduction of noise, which made the tests closer to the real case, did increase the best result of the threshold. In the real case study, the two maxima awaited appear but the scale parameters at which they come are lower than what was predicted by the synthesis images. These results definitely need further investigations. It would be important to consider more complicate patterns: differences in the contrast, shapes more complicate for the objects, different sizes in a same image (e.g. in this case, mix objects with buildings-sizes and houses-sizes)…

REFERENCES

References from Journals:

Aplin, P., Atkinson, P. M., Curran, P. J., 1999, Fine Spatial resolution simulated satellite sensor imagery for land cover mapping in the United Kingdom, *Remote sensing of Environment*, 68, pp. 206 - 216.

Brunett, C., Blaschke, T., 2003, A multi-scale segmentation/object relationship modelling methodology for landscape analysis, *Ecological Modelling*, 168, pp. 233 - 249.

Carleer, A. P., Debeir, O., Wolff, E., 2005. Assessment of very high spatial resolution satellite image segmentations, *Photogrammetric Engineering and Remote Sensing*, in press.

Thomas, N., Hendrix, C., Congalton, R. G., 2003, A comparison of urban mapping methods using high-resolution digital imagery, *Photogrammetric Engineering and Remote Sensing*, 69 (9), pp. 963 - 972.

Van de Sande, C. J., de Jong, S. M., de Roo, A. P. J., 2003, A segmentation and classification approach of IKONOS-2 imagery for land cover mapping to assist flood risk and flood damage assessment, *International Journal of Applied Earth Observation and Geoinfromation*, 4, pp. 217 - 229.

Pal, N.R., Pal, S.K., 1993, A review on image segmentation techniques, *Pattern Recognition*, 26(9), pp. 1277 - 1294.

Woodcock, C. E., Strahler, A. H., 1987, The factor of scale in remote sensing, Remote sensing of environment, 21, pp. 311-332.

References from Books:

Definiens Imaging, 2003, *eCognition Professional user guide 3*, (URL: http://www.definiens-imaging.com/), Definiens Imaging, München.

References from Other Literature:

Carleer, A. P., Debeir, O., Wolff, E., 2004, Comparison of very high spatial resolution satellite image segmentations, *Proceedings of SPIE Image and Signal Processing for Remote Sensing IX*, Vol. 5238, ed. Lorenzo Buzzone, pp. 532-542.

Baatz, M., Schäpe, A., 2000. Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation, in Strobl, Blaschke & Greisebener (eds), *Angewandte Geographische Informationsverarbeitung XI*. Beiträge zum AGIT-Symposium, Salzburg 1999.

Van de Voorde, T., De Genst, W., Canters, F., Stephenne, N., Wolff, E., Binard, M., 2004, Extraction of land / use land / cover – related information from very high resolution data in urban and suburban areas. In: Gossens, R. (ed.), *Remote Sensing in Transition, Proceedings of the 23rd Symposium of the EuropeanAssociation of Remote Sensing Laboratories*, Ghent, Belgium, pp. 237 – 244.

ACKNOWLEDGEMENTS

The Belgian Science Policy is gratefully acknowledged for funding the work that is presented in this paper (project SR/00/02).