Segmentation Evaluation for Object Based Remotely Sensed Image Analysis

Padraig Corcoran & Adam Winstanley
National Centre for Geocomputation
National University of Ireland Maynooth, Co. Kildare, Ireland.
Tel. +353 1 708 3853
Fax +353 1 708 3848
padraigc@cs.nuim.ie

KEYWORDS: OBIA, Segmentation, Texture, Evaluation.

1. Introduction

Object based image analysis (OBIA) is a relatively new form of remote sensing which aims to overcome the failings of traditional pixel based techniques at providing accurate land-use classification for high resolution data. The failure of pixel based techniques is due to the fact that these techniques are based on the assumption that individual classes contain uniform visual properties. As we increase the spatial resolution of data the intra-class variation increases and this property of class uniformity is broken leading to very poor performance (Blaschke 2003). The human visual system (HVS) can interpret high resolution data very easily and accurately. If a truly accurate and robust automated land-use classification system is to be achieved, it must draw from research in the area of cognitive psychology and attempt to model how we as humans interpret aerial imagery. This is the aim of research in the area of OBIA.

Traditionally OBIA comprises two steps. First a segmentation of the given scene is performed which defines the different objects or land-covers. This is then used as input to a land-use classifier. In a recent paper (Corcoran and Winstanley 2006) we discussed a model to define land-cover objects which agrees more with current theories of visual perception about how we perceive such objects then previous approaches. In this model, segmentation is performed in two stages. First a bottom-up feature extraction of visual properties and segmentation is performed, where each area of uniform properties is represented once within a single segmentation. This is motivated by the principle of uniform connectedness which states that adjacent regions of uniform visual properties are perceived initially as single perceptual units in early vision, and serve as entry level visual stimulus (Palmer and Rock 1994; Pylyshyn 1999). This low level segmentation is then affected by top-down and bottom-up influences where objects are aggregated to form segmentation at a higher scale. In (Vecera and Farah 1997) it was shown that such top-down factors do influence the segmentation process. Therefore different levels of segmentation exist, one where each area of uniform properties is represented as a single object and others where these objects influenced by top-down and bottom-up factors are aggregated to form segmentation at a larger scale. When attempting to evaluate a segmentation algorithm using a particular evaluation method we must be aware of this fact, and ensure the evaluation method targets the corresponding level of segmentation which we are attempting to model.

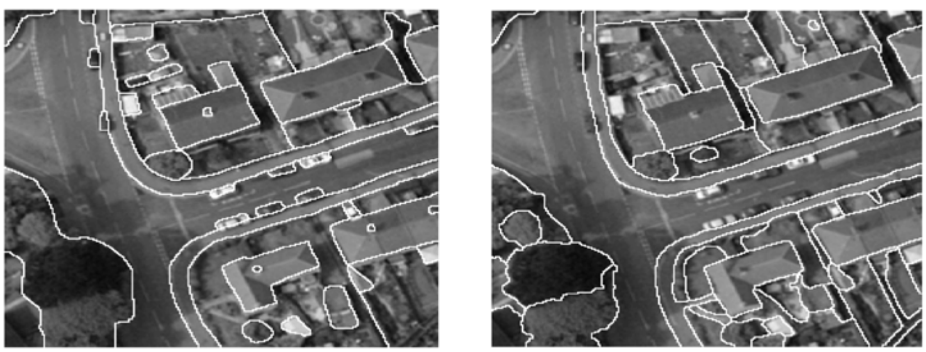
The aim of my overall research is to model the early vision process of representing each area of uniform properties as a single object. The research question I attempt to answer here is how best to evaluate such a model. We now discuss the different forms of supervised and unsupervised evaluation which may be used. A novel form of unsupervised evaluation is introduced which incorporates a set of features which complement as opposed to compete against each other as is standard in previous approaches.

2. Supervised Evaluation

In an experiment to assess the extent of top-down and bottom-up influenced object aggregation in human visual segmentation, we asked a number of subjects who were unaware of the research background to segment a number of aerial scenes. They were untrained photograph interpreters but familiar with aerial photography. The instructions to the subjects were brief and similar to those used in (Martin, Fowlkes et al. 2001) to capture the Berkeley segmentation dataset of natural scenes:

You will be presented with a photographic image. Divide the image into some number of segments, where the segments represent "things" or "parts of things" in the scene. The number of segments is up to you, as it depends on the image.

Figure 1 show examples of the segmentation results returned from two individuals for the same image. From these images we can see that aggregation of objects with uniform properties has played a major factor in the segmentation process. Although large scale conceptual objects such as trees, buildings, roads and sidewalks are segmented, a large amount of individual objects of uniform properties are not represented. Although only two images are shown here, this property was uniform across the whole data set captured.



(a) (b)

Figure 1: Segmentation results returned by two individuals are shown in (a) and (b). Object boundaries are represented by the colour white.

Traditional remote sensing classification algorithms are evaluated in a supervised manner using ground truth (Congalton and Green 1998). Supervised evaluation requires the capturing of ground truth and a method for comparing a segmentation result to this ground truth. As can be seen in figure 1, when asked to produce ground truths, human subjects tend to produce segmentation at a larger scale, where individual objects are merged, then we desired. Capturing of ground truth where each object of uniform properties is represented would require a great effort on the interpreter's part and may not be even possible due to the fact that we tend to merge objects unconsciously. To date no such data set exists. Using existing land-use classifications as ground truth is also not an option because these will also contain aggregated objects. Another point to note is that although each individual's segmentation shares similarities with others, they also have significant differences. This is due to the fact that each individual's interpretation of a given scene and how objects should be merged will vary. Another form of supervised evaluation which is widely used in remote sensing literature is visual assessment of the segmented images. This approach has the advantage of not requiring any implementation but is subjective and time consuming.

3. Unsupervised Evaluation

In order to overcome the difficulties of supervised segmentation evaluation, we propose to perform evaluation in an unsupervised manner. In unsupervised evaluation a measure of uniformity of intra-class features and/or a measure of separation of inter-class features is used to measure the quality of segmentation (Rosenberger, Chabrier et al. 2006). Although this approach has received must attention, most methods in the area are based on the assumption that each individual object is of uniform intensity properties. In cases where images contain areas of uniform texture such as remotely sensed images, these approaches will fail.

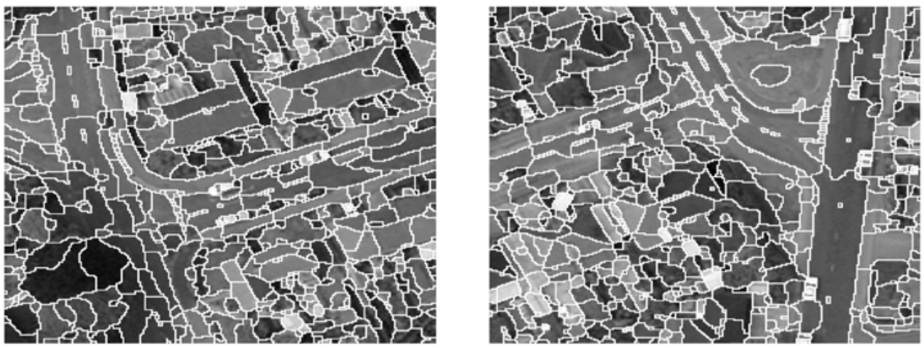
Our unsupervised evaluation technique builds on the work of (Chabrier, Emile et al. 2006) and (Rosenberger, Chabrier et al. 2006). For unsupervised evaluation Chabrier (Chabrier, Emile et al. 2006) defined a set of texture and intensity features which cannot complement but actually competed against each other. In this approach each region is defined as being either predominately uniform intensity or texture, and only the corresponding subset of features is used in segmentation evaluation. Most objects will have unique average intensity and texture values so failing to use both feature sets leads to reduced discrimination strength. In this paper we propose a set of texture and intensity features for segmentation and unsupervised evaluation that actually complement each other, thus giving greater discrimination power. These features consist of the average intensity and texture features for each land-cover, for further details see (Corcoran and Winstanley 2006) and (Corcoran and Winstanley 2007). Given a feature set and segmentation result derived using these features, we must then define the cost function of the features which measures the accuracy of segmentation. We use a novel measure of the ratio between inter and intra object features as such a function.

An important issue with unsupervised evaluation approaches, is how best to evaluate if we are using the correct feature set, segmentation algorithm and optimizing the correct segmentation evaluation cost function. One approach is to treat a supervised evaluation method as the optimal solution and attempt to generate an unsupervised method which has high correlation with the supervised method (Chabrier, Emile et al. 2006). Generation of the ground truths for aerial images for use in supervised evaluation presents a number of issues which we discussed previously. Another option is to use synthetic images where ground is known (Rosenberger, Chabrier et al. 2006). Unless synthetic images of similar complexity, containing areas of uniform properties and shapes similar to aerial images are used, then this form of evaluating unsupervised methods is not accurate. To create images of this complexity would require an understanding of the underlying processes generating aerial images which is unknown. This view is also shared in (Usamentiaga, Garcia et al. 2006). Therefore visual assessment is used to evaluate the accuracy of our unsupervised evaluation.

That is, the goal of this work is to produce a feature set, segmentation algorithm and evaluation cost function which targets segmentation where individual objects of uniform visual properties are accurately represented. Visual assessment of the optimal segmentation in terms of the given cost function will evaluate this. The need to perform visual assessment of the unsupervised evaluation procedure begs the question of why do we not simply perform visual assessment of the segmentation algorithm and abandon unsupervised evaluation. The motivation for having an unsupervised evaluation procedure is that it will generalize to allow the evaluation of different segmentation algorithms on different data sets therefore removing the need for future visual assessment work.

4. Results & Conclusions

Figure 2 show some examples of the optimal segmentation result in terms of the segmentation cost function for the given feature set. We can see from these images that our unsupervised evaluation strategy favours a scale of segmentation much smaller then that produced by human subjects. It returned a mean of 799 objects with standard deviation of 20 objects when applied to our dataset. On the same dataset the human interpreters returned an average of 48 objects with standard deviation of 6 objects. This optimal segmentation in terms of the unsupervised criteria is quiet close to the segmentation result we desire where individual areas of uniform visual properties are represented as single objects. Since this evaluation method targets the desired segmentation result, it may be used to aid the choice of segmentation algorithm or algorithm parameterization for the given feature set extracted from different data. Close inspection of segmentation results reveals some errors; some of objects are over-segmented and in some cases the boundary localization is not exact. Future work will aim to reduce these errors and improve segmentation performance.



(a) (b)

Figure 2: Some examples of the optimal segmentation result in terms of evaluation cost function for the given feature set are shown in (a) and (b).

Acknowledgements

Support for this work is acknowledged to the Irish Research Council for Science Engineering and Technology (IRCSET). Ordnance Survey UK is acknowledged for providing data for use in this work.

Biography

Padraig Corcoran of the National Centre for Geocomputation (NCG), is in his Third year of PhD study. He has presented at GISRUK on three previous occasions.

References

Blaschke, T. (2003). *Object-based contextual image classification built on image segmentation*. IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data. 113-119.

Chabrier, S., B. Emile, et al. (2006). *Unsupervised performance evaluation of image segmentation*. Special issue on performance evaluation in image processing. EURASIP journal on applied signal processing.

Congalton, R. G. and K. Green (1998). *Assessing the accuracy of remotely sensed data: principles and practice*, Lewis Publishing, U.S.

Corcoran, P. and A. Winstanley (2006). *Using Texture to Tackle the Problem of Scale in Land-Cover Classification*. International conference on object based image analysis, Salzburg.

Corcoran, P. and A. Winstanley (2007). *Removing the texture feature response to object boundaries*. International conference on computer vision theory and applications, Barcelona.

Martin, D., C. Fowlkes, et al. (2001). *A database of human segmented natural images and its applications to evaluating segmentation algorithms and measuring ecological statistics*. International conference on computer vision, Vancouver. 416-423.

Palmer, S. E. and I. Rock (1994). *Rethinking perceptual organization: The role of uniform connectedness*. Psychonomic Bulletin and Review 1(1): 29-55.

Pylyshyn, Z. W. (1999). *Is vision continuous with cognition? The case for cognition impenetrability of vision perception*. Behavioural and brain sciences 22(3): 341-365.

Rosenberger, C., S. Chabrier, et al. (2006). Unsupervised and supervised image segmentation evaluation. Advances in image and video segmentation. Y.-J. Zhang, IRM Press.

Usamentiaga, R., D. F. Garcia, et al. (2006). *A method for assessment of segmentation success considering uncertainty in the edge positions*. Special issue on performance evaluation in image processing. EURASIP journal on applied signal processing.

Vecera, S. P. and M. J. Farah (1997). *Is visual image segmentation a bottom-up or an interactive process?* Perception & Psychophysics 59(8): 1280-1296.