An object-based information extraction methodology incorporating a-priori spatial information

Thomas Blaschke, Dirk Tiede, Stefan Lang
Z_GIS, Center for Geoinformatics, University of Salzburg

Extended abstract for interactive presentation

Introduction
Remote sensing has become an essential data source for landscape analysis. No other survey technique can operationally provide a regularized survey of landscape with which to assess landscape patterns and change. However, remotely sensed images, like all observations of reality, are an imperfect capturing of patterns, which are themselves an imperfect mirror of the underlying processes. In this paper, we examine some of problems inherent in an analysis of landscape as an array of ‘pixel’ objects, the so-called raster model of geographical phenomena (Fisher 1997). Cracknell (1998) explores the question “What’s in a pixel?” and makes the point that a pixel (the ‘footprint’ or ground instantaneous field of view, GIFOV) of a sensor is often larger than we would like it to be, a penalty imposed by the technology in return for the ability for the sensor to give an overview of a very large area. Unfortunately, sensor GIFOV are often imposed on us by technological or logistical constraints (Steele 1978) and not based on the needs of the ecologist nor on hypotheses about the objects being explored. We discuss later if the increasing spatial resolution resolves these problems or simply shifts the scale or ‘window of perception’ (Hay et al. 2001). In the beginning we structure the available methods from different scientific communities. In particular, we juxtapose image data mining methods in a narrow sense and methods to extract particular knowledge from spatial databases. More recently, imaging methods and GIS-like spatial data analysis techniques are combined to discovering relationships between spatial and non-spatial data aiming for the construction of spatial data bases. Knowledge discovered from spatial data can be obtained and stored in various forms including discriminant rules, prominent structures, spatial associations or topological relationships.

Spatial data infrastructures and image analysis
Spatial data mining has gained significant progress over the last years (Datcu et al. 1999), especially beyond single images, e.g. in regard to multi-temporal images (Datcu 2004). The amounts of data today exceed the human’s ability to analyze them. Only recently, data mining approaches have extended the scope of data mining from relational to transactional databases to spatial databases. This paper provides a somewhat different view compared to the ‘mainstream’ in the information extraction community. Firstly: rather than starting from the pixel perspective we put emphasis on objects. Secondly, we do not regard a remotely sensed image as a black box. We hypothesize that in the 21st century there will always be auxiliary information available. Potential frameworks for an intelligent exploitation of existing Geoinformation are spatial data infrastructures (SDI).

The term “spatial data infrastructure” is often used to denote the relevant base collection of technologies, policies and institutional arrangements that facilitate the availability of and access to spatial data. A spatial data infrastructure provides a basis for spatial data discovery, evaluation, download and application for users and providers within all levels of government, the commercial sector, the non-profit sector, academia.
The word infrastructure is used to promote the concept of a reliable, supporting environment, analogous to a road or telecommunications network. Spatial data infrastructures facilitate access to geographically-related information using a minimum set of standard practices, protocols, and specifications. Spatial data infrastructures are commonly delivered electronically via the internet.

The information from an SDI usually is a meta-information. Therefore, we develop a methodology to utilize spatial databases according to standardizations of SDI’s. Some good examples of a linkage of satellite imagery and web portals exist in various application areas already. For instance, to support relief efforts for natural disasters, such as floods, earthquakes, hurricanes, etc., the capability to look at potential problem areas affected is a feature of several Web services. In the US, the National Map Hazards Data Distribution System (HDDS) provides a dynamic online map interface that can be used to view USGS data sets that are part of The National Map (http://gisdata.usgs.net/website/Disaster_Response). Other examples include the National Disaster Hazard and Vulnerability atlas in South Africa (Sakulski, 2005). Still, to the knowledge of the authors, these applications include images as graphical components but do not analyse or interpret them according to the need of the respective applications. We claim that a fast incorporation of information extraction results into national or regional data sets would be a great asset to these Web services.

**Beyond pixels**

Cracknell (1998) divides his critical examination of the ‘pixel’ into geometry, mixed pixels, point spread functions and resampling, and concludes that the ‘pixel’ is a more complicated entity than is generally acknowledged, and we must approach landscape analysis using EO data critically. Increasing spatial resolution does not solve this problem but decreases the effects to the ‘mixed pixel’ problem. Depending on the objects of interests and the target scales of analyses this may lead to the fact that the ‘pixel problem’ may be more and more neglectable, at least statistically.

The traditional method for analysis of EO data in landscape research is the classification of pixels based on pixels in the same land cover class being close in spectral feature space. Put different way, one assumes that the classes are relatively pure or at least spectrally separable. Although not always a statistically sound method (the popular maximum likelihood classification method assumes a normal distribution of data, which is unlikely in EO images), the methodology there has had some success. However, the reference to land cover classes is instructive, because with GIFOV of 1 km to 30 m, only the broadest land cover classes can be spectrally differentiated. Separability has been improved by the incorporation of spatial information, such as local measures of texture and autocorrelation, but is not assuaged. Spectral separability aside, there is still problems and the scale issue is just one example. A final complication to the pixel approach is termed conceptual error by Jong (1990). This error arises from errors in the design of the classification system that is generated for the landscape analysis. Finally, the pixel-centred view is usually uni-scale in methodology, exploring the pixels of only one scale of imagery and of only one scale within the image.

A variety of techniques have been proposed to wrestle the pixel approach into addressing pixel / ecological object discontinuities. These include spectral ‘unmixing’ using linear mixing models, and the use of fuzzy sets and neural nets. However, by ignoring concepts of hierarchy and scale in the landscape processes driving pattern creation, these approaches are still overly ‘pixel-centred’ (Townshend et al. 2000). They adhere to a concept of the pixel as a spatial entity (Fisher 1997) that is assumed to have a de facto relationship to objects in the landscape. Uni-scalar, pixel-based monitoring methodologies have difficulty providing useful information about complex multiscale systems. If we accept that the reality we wish to
monitor and understand is a mosaic of process continuums, then our landscape analysis must make use of methods which allow us to deal with multiple, yet related scales within the same image and with multiple images of landscape. Increasingly used multiscale methods in landscape ecology include semivariance analysis (Faber & Förstner 1999), wavelet analysis (Sheikholeslami, 2000), fractal analysis (Milne 1991, Nikora et al. 1999), and lacunarity analysis (Plotnik et al. 1993). O’Neill et al. (1992) have also expanded percolation theory to hierarchically structured landscapes.

The pixel-centred approach to landscape analysis is constrained by a relatively weak relationship to ecosystem theory. Advances have been made in exploring hierarchy in image analysis, for example in the nested scene models and image segmentation.

The multiscale segmentation/object relationship modelling (MSS/ORM) methodology suggested by Burnett and Blaschke (2003) segments information (usually remote sensing images plus any georeferenced information). Generally, an advantage of segmentation to classification of pixels is that the resulting division of space tends to involve fewer and more compact subregions. The multiscale segmentation based approach is designed to utilize information in the scales inherent in our spatial (image) data sets in addition to a range of auxiliary data sets, including for airborne and satellite data, but also to the scales of information inherent in single images. Technically, segmentation is not new (Haralick & Shapiro, 1985) but only since around the year 2000 we can observe a rapidly increasing number of applications with is often associated with the advent of commercially available high resolution satellite imagery (Ikonos: 1999) and a commercial software package for object-based image analysis – eCognition – (Blaschke & Strobl 2001, Blaschke et al. 2004, Benz et al. 2004). The main underlying concept is to somehow mimic how a human operator works: to create regions instead of points or pixels as carriers of features which are then introduced into the classification stage. The conceptual idea is that each of these regions corresponds exactly to one and only one object class (Schiewe 2002). Furthermore, segmentation algorithms are able to handle multiple data and information sources, thus performing a fusion on feature level (Ehlers et al. 2006).

Over the last very few years, many new segmentation algorithms and applications have been tested in GIScience applications, but few of them lead to qualitatively convincing results while still being robust and operational (Blaschke & Strobl, 2001; Blaschke et al. 2004, Hall & Hay 2004, Benz et al. 2004, Neubert & Meinel 2005).

Landscape researchers now have the benefit of working with the next generation EO data sets, comprising 1) images of a significantly finer spatial resolution, and 2) multiple scales of data simultaneously, thus opening up the potential for analysis methodologies that are better adapted to the self-organized complexity of landscapes.

**Image objects**

As stated above, the strong motivation to develop techniques for the extraction of image objects stems from the fact that most image data exhibit characteristic texture which is neglected in common classifications although many authors started to tackle this problem already in the 1970ies (Kettig & Landgrebe 1976, Haralick et al. 1973). In addition to spectral aspects in images, GIS principally introduces topology as a new dimension to map the relations between n-dimensional entities. We speak of objects if we can attach a meaning or a function to the raw information. Generally, the object is regarded to be an aggregation of the geometric, thematic and topologic properties. The topologic relations between the cells the object consists of can be examined once the user has defined his or her objectives, classification scheme and scale of analysis. For a recent overview on image segmentation we refer to Blaschke et al. (2004).
Most researchers applying a segmentation approach argue that image segmentation is intuitively appealing. Human vision generally tends to generalize images into homogeneous areas first, and then characterize focal areas more carefully as required (Gorte 1998). Following this observation, we hypothesize that by creating multiple scales of segmentation, by successively grouping the pixels of an image into homogeneous image objects, a more intuitive and hierarchical partitioning of the image results. Among the most promising are Markov image segmentation, multi-fractals based segmentation (Vehel & Mignon 1994), or segmentation based on representativeness measures (Hoffmann & Boehner 1999).

In this paper we use the region-based, local mutual best fitting segmentation approach (Baatz & Schäpe, 2000) as being implemented in the software eCognition (Benz et al., 2004) and build on the MSS/ORM approach (Burnett & Blaschke, 2003) as described above. The procedure for the multi-scale image segmentation presented is a region merging technique. It starts with 1-pixel image objects. Image objects are pairwise merged one by one to form bigger objects. In this conceptualisation the procedure becomes a special instance of an assignment problem, known as pairwise data clustering. In contrast to global criteria, such as threshold procedures, decisions are based on local criteria, especially on the relations of adjacent regions concerning a given homogeneity criterion. In such an optimisation procedure each decision concerning a merge is based on all previous decisions or merges at least in its local vicinity. Therefore such a procedure includes to a certain degree historicity which can cause problems for reproducibility. The solution for this problem is the optimisation procedures and the homogeneity criteria which are maximizing the constraints in the optimisation process (Baatz & Schäpe, 2000). The growth is based on a high number of pairwise merges. The segmentation process can be seen as a crystallisation process with a big number of crystallization seeds. The requirement for the maintenance of a similar size of all segments in a scene is to let segments grow in a simultaneous or simultaneous-like way.

The collecting of semantic relationships brings up the issue of a priori knowledge. For some landscape situations, it has been hypothesized that, for different scenes the similarity of object scales and object characteristics will enable nearly automated and highly accurate classification of land-use. To achieve this, these analysis systems will have to optimize the data collected and the relationship rules applied. Urban landscapes are an example and the use of 3D models from scanning LIDAR or digital stereo orthophotography combined with spectrometer data is likely to achieve useful results. The varieties of objects in a more natural landscape add additional challenges, and the importance of semantic (human) knowledge of each particular scene may play a more important role in these studies.

Application

In the empirical part we demonstrate the incorporation of external spatial databases in the classification process. We conclude from some ongoing studies within the EU network of excellence GMOSS that these examples are satisfactory concerning the principle linkage capacity. In a cooperation with a commercial company we could already demonstrate that single objects created from a multi-resolution segmentation can further be flexibly broken down to pixel-sized objects and rebuilt to meaningful objects based on local object parameters (Tiede & Hofmann, 2006). The results are promising, especially concerning the problems which occurred in earlier studies on object-based delineation of single trees. These new possibilities are extending a software package to a sort of modular, process oriented programming language. Consequently the research taken herein can illuminate one aspect of a new and complex context-based classification methodology. This reflects our theoretical approach described earlier and opens ways for a hierarchical description and subsequent classification of whole landscapes.