OBJECT-BASED RE-CLASSIFICATION OF HIGH RESOLUTION DIGITAL IMAGERY FOR URBAN LAND-USE MONITORING

Stuart L. Barr,
Remote Sensing Unit,
Department of Geography,
University College London,
26, Bedford Way,
LONDON WC1H OAP
England

ABSTRACT

This paper examines the application of object-orientated processing and artificial intelligence techniques to high spatial resolution satellite sensor images for urban land-use monitoring. Although these techniques have been applied to aerial photography for some time, their use in the analysis of digital images acquired by satellite sensors is much less well developed. Within this study, a two-stage approach is adopted to map urban land use from SPOT-HRV multispectral and panchromatic images. Firstly, a conventional per-pixel multispectral classification is performed to derive a map of the principal land cover types present within the scene. The second stage involves the spatial re-classification of these land cover types into land-use categories. Contiguous blocks of pixels with the same class label are grouped into 'objects' to form an object search map of cover types. This is used to derive an extended region-adjacency graph (XRAG). The XRAG contains information not only on the spatial relationships between individual objects within the scene, but also on the attributes of those objects (e.g. class label, size, perimeter), both of which are used in the re-classification. *A priori* knowledge of the previous extent of the urban area can also be used to guide the re-classification procedure. Preliminary results obtained using these techniques are shown to be considerably improved with respect to those obtained using a standard, per-pixel classification of the same scene.

KEYWORDS

Spatial Searching, Object-Orientated Processing, Expert Systems, Rule-Base, Urban Areas

INTRODUCTION

Town planners require various types of information about urban areas, some of which may be derived from remotely-sensed images. The type of information that we might expect to be able to derive from satellite sensors includes the range and spatial distribution of land cover types and land use categories present, as well as the physical extent of the urban area (including some description of its shape and perimeter). Unfortunately, early attempts to derive information of this kind often failed to produce the levels of accuracy and detail required for town planning purposes. At the time, this was ascribed to the relatively low spatial resolution of satellite sensors. However, the use of data from sensors with improved resolving power in more recent years has not always yielded the expected increases in classification accuracy (Townshend 1981). Indeed, in some cases, the levels of accuracy obtained actually worsened. This has generally been referred to in the literature as a problem of increasing 'scene noise' - i.e. increasing spatial heterogeneity in the spectral response of urban areas. In fact, the problem is more properly expressed in terms of the image analysis techniques used and, in particular, to the inappropriateness of standard, per-pixel parametric algorithms for segmenting images of urban areas. Simply stated, it is extremely difficult to derive consistent, representative training statistics for urban land use categories since these are comprised of many land cover types, each with different spectral reflectance properties (Barnsley *et al.* 1991).

Alternative approaches to urban land-use mapping have been explored. The most successful of these make use of other sources of spatially-referenced ancillary data (Sadler and Barnsley 1990, Sadler *et al.* 1991) or image texture measures (Sadler *et al.* 1991, Baraldi and Parmiggiani 1990). The ancillary data set normally forms an additional data plane in the standard algorithms. Increased classification accuracy has been reported using both of these techniques (Sadler and Barnsley 1990, Sadler *et al.* 1991), although improvement has not been consistent across all land use categories.

The fundamental problem involved in producing accurate ***land use*** maps of urban areas arises from the fact that urban areas are complex assemblages of a disparate set of ***land cover*** types - including man-made structures, vegetation types and water bodies - each of which has different spectral reflectance characteristics. In visual analyses of remotely-sensed images, particularly aerial photography, the spatial pattern of these land cover types is often used to identify and to distinguish between categories of urban land use. For example, residential areas can often be recognized by their particular mixture of buildings, roads, grass and trees; by contrast, parkland is primarily composed of grass and trees.

Recently, several studies have attempted to use the spatial mixing of land cover types within urban areas as a means of mapping land use. The studies by Wharton (1982), Whitehouse (1990) and Barnsley *et al.* (1991) have utilised various forms of (per-pixel) spatial re-classification techniques applied to an initial (land cover) segmentation of urban areas. The fundamental basis of these techniques is that it is possible to obtain some measure of the density and distribution of land cover types that is characteristic of a particular urban land use (Barnsley *et al.* 1991). Areas of similar land use can therefore be delineated by grouping pixels with different class (land cover) labels on the basis of these measures. All three studies attempt this through the use of a type of convolution kernel which either sums the density distributions of the constituent cover types (Wharton 1982, Whitehouse 1990) or measures their spatial arrangement (Barnsley *et al.* 1991) within the kernel.

Although promising results have been obtained using these algorithms, the use of a pre-defined kernel places an undesirable restriction on the nature of the spatial searching employed. In particular, it is doubtful whether a single kernel of any size can adequately characterize the complex spatial distribution of the cover types contained in all of the land use categories likely to be found within a typical urban scene.

THE POTENTIAL OF COMPUTER VISION AND ARTIFICIAL INTELLIGENCE TECHNIQUES

Investigations by McKeown using very high resolution panchromatic aerial photography have demonstrated that certain techniques used in computer vision, artificial intelligence and database design may provide suitable tools for analysis of urban scenes (McKeown 1988, 1991). In particular, McKeown makes use of both the observed spatial relationships between objects identified within the images and *a priori* knowledge relating to both general and scene-specific spatial relationships between objects typical of urban areas. This information is used to locate features of interest (Huertas and Nevatia 1988, McKeown 1988, 1991).

This paper explores the use of a similar set of techniques in the segmentation of digital, multispectral images of urban areas obtained by high spatial resolution satellite sensors, for the purpose of land use mapping. Of necessity, the expert systems developed in this study represent spatial information differently from McKeown (1988). Detailed knowledge of the shape and geographic location of a particular building is of lesser use in the segmentation process at satellite sensor resolutions. Information such as the general form of urban areas and the inter-relationships between land cover types within urban areas forms a much more valuable source of *a priori* knowledge at this spatial scale.

This paper outlines one possible approach to urban segmentation and monitoring using some of the techniques described above, implemented within an object-oriented processing environment. Consideration is given to the spatial resolution and information content of spaceborne multispectral imagery of urban areas in the development of these techniques. A new data structure is suggested for use in the spatial processing of urban areas and the results of a preliminary investigation into its incorporation within an expert system to map urban land use are given.

DEVELOPING SPATIAL STRUCTURES FOR COVER TYPE OBJECTS

Given an initial, low-level (land-cover) segmentation derived from a high spatial resolution multispectral image, urban areas can be considered to be composed of many discrete regions of individual land cover types (hereafter, referred to as 'objects'). Each object may have a number of attributes. These may either be *internal* to the object, such as its geometric parameters (i.e. shape and size), or *external* to the object, such as 'adjacency' or 'containment'.

McKeown (1988) discusses the advantages of processing spatial information a variety of levels, ranging from low-level processes, such as region-growing routines, through to high-level process concerned with the recognition of specific features. This paper not only describes work performed at a variety of processing levels, but also two levels of spatial information, a low-level map of image 'objects' and an intermediate data structure containing object information relevant to high-level processing.

Creating an Object Search Map

The fundamental spatial information structure used in this study is the representation of individual land-cover type parcels (objects) by their boundary coordinates, often referred to as *iconic boundary representation*. Given a segmented (i.e. land cover) image, the problem of extracting information on object boundaries becomes one of recognising the initial start pixel of each object, tracing the outer boundary of the object and noting the location of each pixel encountered during the trace. A contour encoding algorithm has been used to achieve this task. Further discussion of this algorithm and its coding considerations may be found in Gonzalez and Wintz (1987). Unfortunately, the contour-encoding algorithm generates large volumes of data when held in image coordinates. Freeman's chain code - both forward and backward conversion routines (i.e. Freeman code to image coordinates and *vice versa*) - has therefore been implemented to allow more efficient storage of object boundaries (Freeman 1961). Figure 1 shows the process of extracting the iconic boundary for a cover object and its representation in the object search map. Figure 2 shows the iconic boundary representation (in image coordinates) of a complex land cover object.

An Intermediate Data Structure for Spatial Searching and Processing

Nichol (1990) outlines the potential applications of Region Adjacency Graphs (RAG) for spatial analysis within high spatial resolution multispectral imagery. The RAG data structure contains the information on objects

$$Y(y_1, y_2, y_3, \ldots y_n)$$

adjacent to any other object, X. In this form, the data structure is an intermediate representation of spatial information, although a low-level representation of object, X, in question.

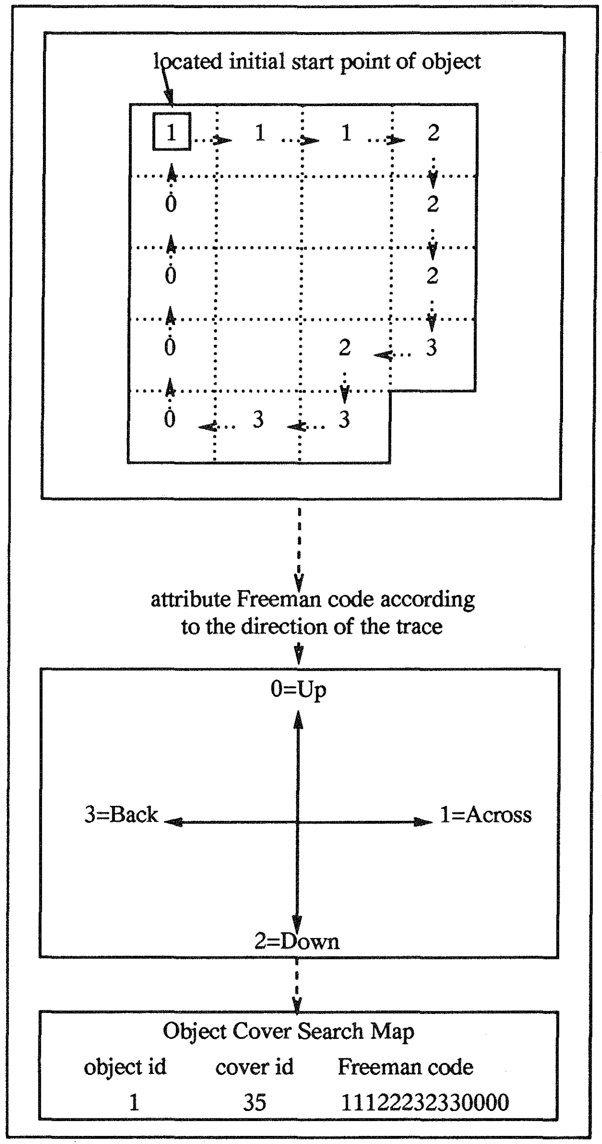


Figure 1 :Diagram Representing the Processing Stages for the Extraction of the Object Cover Search Map.

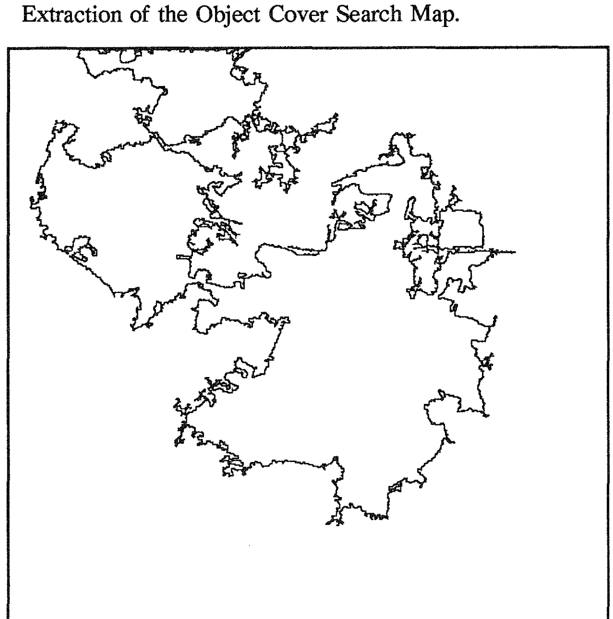


Figure 2 :The Boundary of a Complex Object in Image Coordinates Extracted by the Contour Encoding Algorithm.

In this study, the RAG structure has been extended in such a way that it contains both spatial information and 'internal' object parameters (object attributes). Figure 3 illustrates a simple image with a number of objects, along with its representation as a extended RAG (XRAG) structure. In this case each object is characterised by a object ID, a class label ID, a set of *internal* object attributes (area and perimeter values in 'ground-based metrics' (McKeown 1988)) and, a stream of the adjacent objects with the @ operator denoting 'contained within'. Nichol (1990) uses the basic RAG structure to merge objects within a classified image. However, the interest in this study is to use the XRAG structure for more intricate and complex spatial problems where both internal and external object attributes are used as the basis for high-level analysis and processing.

The basic RAG structure may be extracted by re-tracing the iconic boundaries held in the object search map. As each object is traced, the algorithm makes a 'view' noting the objects encountered along the boundaries. The view taken at any particular trace point is always towards the left, relative to the direction of the trace. For example, if the trace direction is from left to right across the image, then the view taken is up. At each location an additional 'view' is made inwards from the boundary, thus accounting for those objects which are adjacent to, but also *contained within*, the object boundary being traced.

The additional features of XRAG may be extracted by imbedding the required code within the basic RAG extraction program; however this is seen as undesirable for the proposed system. It is argued that information should only be added to the structure via direction from a higher level process. This removes data redundancy and ensures that the XRAG structure is tailored to the high level process which is to use it. In this study, several programs use the same low-level processes for manipulation of the object search map, whilst the main body of the program specifically extracts the attribute in question. Outlined below is a preliminary investigation into the use of the XRAG structure for complex spatial processing.

URBAN BOUNDARY SEGMENTER (URBS)

One of the basic requirements for planners at regional scales is the accurate differentiation of urban and rural areas. Traditionally, this has been carried out by manual digitisation, following a number of set guidelines (HMSO 1984). URBS (Urban Boundary Segmenter) is a high-level expert system which extracts such information through the spatial analysis of the XRAG structure.

Definition of Urban Areas in England and Wales

Although there are many definitions of what constitutes an urban area, the one used in this study is that employed in England and Wales by the Department of Environment. This considers the definition of urban areas to include areas of permanent man-made structures - including transport corridors which have built up sites on one or both sides - and any area completely surrounded by built-up sites (such as playing fields), provided that their areal extent exceeds twenty hectares (HMSO 1984). It was decided to restrict URBS, at this stage, to the development of rules relevant to this definition of urban areas. This allows a number of important points to be analysed, namely:-

- Can the criteria used in manual digitisation of urban areas be encoded and replicated in an expert system
- Is the XRAG structure flexible enough to analyse complex urban areas within such an expert system.

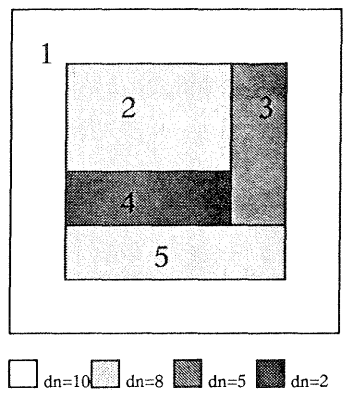
An Overview of URBS

It is often the case that, to represent all possible relationships that may occur during processing, considerable rule-bases have to be written for high-level expert systems (McKeown 1988, Mehldau and Schowengerdt 1990). By contrast, URBS utilises relatively few rules. This is achieved through the use of the XRAG structure which forms an intermediate abstraction of the complex spatial and geometric land-cover segmentation under analysis. Two types of rule are used within URBS. The first comprises the information that is to be extracted within the low-level processing routines to form the XRAG structure. These control which low-level processes are to be implemented, as well as the order in which the desired XRAG structure should be derived. The XRAG structure used within URBS consists of information on *adjacency* and *containment*, as well as the *internal* attribute of object area. The second type of rule is concerned with the segmentation to be carried out. These basically comprise statements of the form 'objects to search for', 'attributes to process', 'operations to be carried out' and 'conditions to be achieved'. Objects which are potentially urban (i.e. the built and large structure cover type) are located within the high-level process. They are then processed according to URBS rules at two intermediate levels described below. Figure 4 shows the overall processing stages within URBS for both rule sets.

WALKABOUT performs the first set of intermediate-level processing. WALKABOUT, through direction form URBS, performs 'walks' between XRAG objects (i.e. it searches through the XRAG structures rather than in image space) evaluating some criteria. In this case, the criterion is to walk through adjacent potential urban cover objects summing their area. Two possible conditions may occur during the 'walk'

- The sum may exceed 20 hectares in which case those objects walked through are considered to be urban,
- Or all possible adjacent potential cover objects have been walked through and the sum is less than 20 hectares, in which case they are considered to be buildings or large man-made structures outside the urban boundary.

The second intermediate-level process, is the recursive analysis of the XRAG objects *contained within* those objects labelled as 'urban' by WALKABOUT. This is performed by LOOKIN. LOOKIN re-labels all objects fully contained within a larger urban object to be urban also. LOOKIN operates recursively to re-label objects contained within objects, and so on. In this way, it accounts for multiple levels of containment. Once recognised as belonging to the urban area, an object is re-labelled as 'urban' by manipulating its class ID within the XRAG structure.



Spatial Resolution=20 meters

XRAG of the above image objects

object id	cover id	area	perimeter	adjacency
1	10	8000	480	@ 2 @ 3 @ 4 @ 5
2	8	2400	200	1 3 4
3	5	1200	160	1 2 4 5
4	2	1200	160	1 2 3 5
5	8	1600	200	1 4 3

Figure 3 :Diagram Showing a Simple Image and its Representation by the XRAG Data Structure.

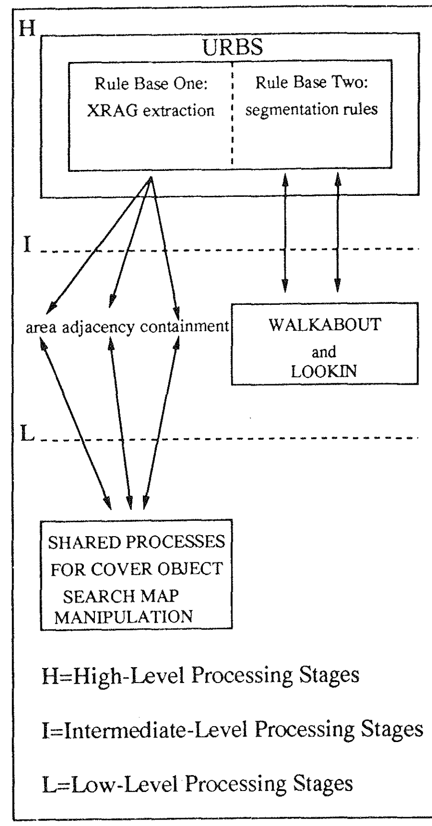


Figure 4 :Diagram Showing the Processing Stages within URBS.

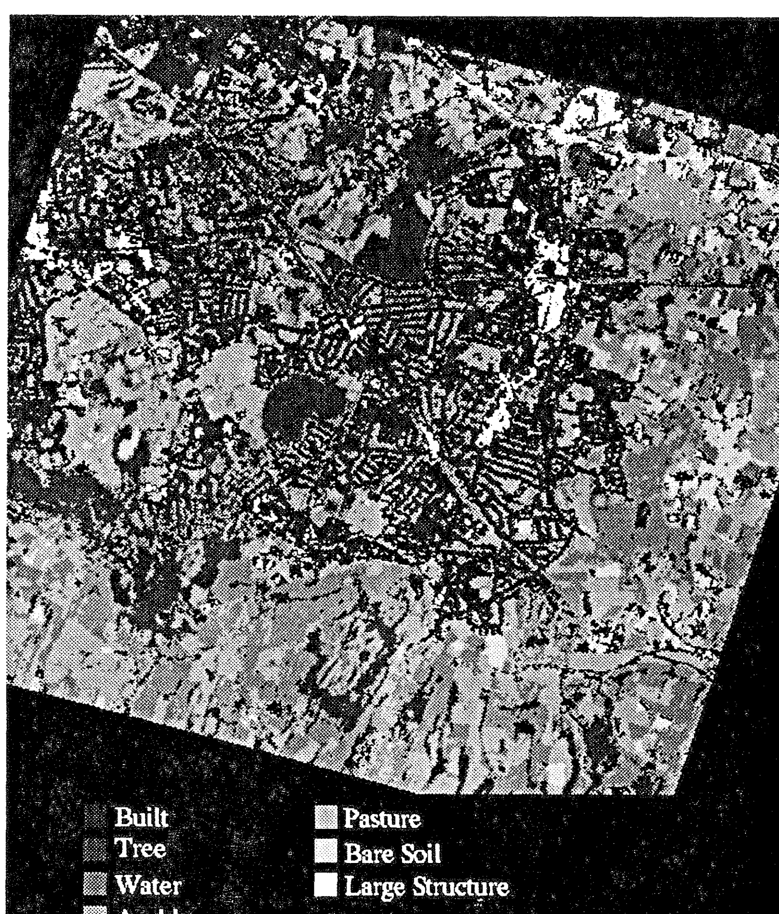


Figure 5 :Initial Land-Cover Segmentation of the Study Area.

Figure 6 :Manually Digitised Urban Areas for Part of the London Borough of Bromley.

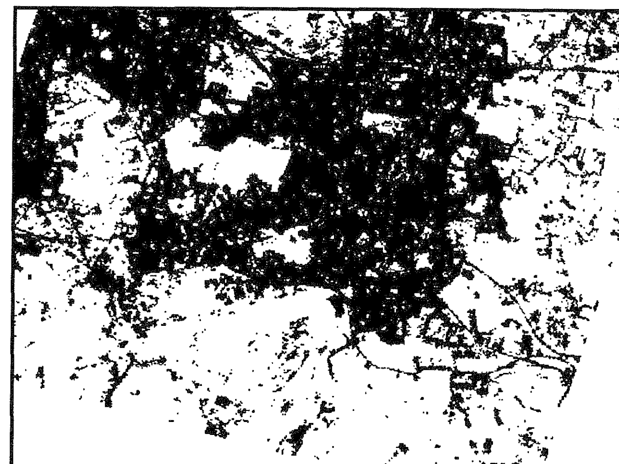


Figure 7 :Classified Urban Areas using a Standard Per-Pixel Algorithm for Part of the London Borough of Bromley.

Figure 8 :Urban Areas Segmented by URBS for Part of the London Borough of Bromley.

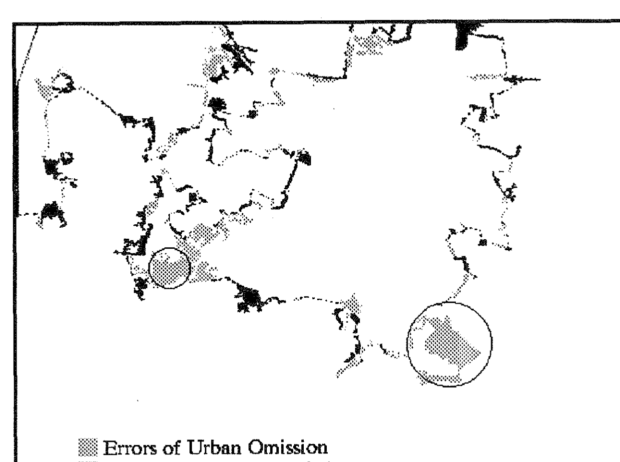


Figure 9 :Urban Areas of Non-Correspondence for The Manually Digitised Urban Areas and the Results from URBS.

STUDY AREA AND SATELLITE SENSOR DATA

For the purpose of this study, an extract from a cloud free, multispectral SPOT-HRV image of London (scene 32, 246; +22.46) acquired on 30th June 1986 has been used. The extract, centred on Orpington in the London Borough of Bromley, is 512x512 pixels (approximately 10km x 10km) in size. The area exhibits a complicated spatial pattern of land cover and land use both within the urban areas and the surrounding rural areas, therefore providing a stringent test of the XRAG structure and the URBS spatial rule-base.

RESULTS AND DISCUSSION

Figure 5 shows an initial land-cover segmentation of the SPOT-HRV sub-scene performed using a standard, maximum-likelihood (per-pixel) classifier. Six broad cover types have been identified, namely 'built' (i.e. houses and roads), 'structure' (mostly large, flat-roofed factories and office blocks), 'tree', 'crop', 'grass', 'soil' and 'water'. The decision to identify two separate classes for man-made structures (i.e. 'built' and 'structure') was taken in view of the pronounced difference in the reflectance properties of these two types of surface. The accuracy of this low-level segmentation (>90%) was considered to be good enough to test the potential of URBS. The segmented image was found to contain 13,400 unique objects, a number that gives some idea of the complex spatial pattern of land cover within this particular scene.

Figure 6 shows a manually digitised coverage of urban areas for this part of the London Borough of Bromley. Figure 7 shows the results obtained using a standard parametric classification algorithm used to map urban areas in the study area. Figure 8 shows the results obtained from URBS. These diagrams show that URBS performs considerably better than the standard, parametric algorithm at segmenting urban areas within the study area (using the manually digitised map as the reference plane). For instance, the standard algorithm has no means of incorporating intra-urban open space (such as parks and other public open spaces) into the 'urban' land use category. In addition to this, many of the small villages and roads outside the main urban area are included into the urban category due to the reliance solely on spectrally assigned properties of the standard parametric algorithm. The output from URBS has far fewer segmentation errors, giving a more precise indication of the extent of the urban areas within the study area. When compared visually, the manually digitised urban areas and those detected by URBS have a strong correlation, with many boundaries seeming to be almost identical, although several notable errors do occur. Figure 9 shows the areas of non-correspondence between URBS and the manually digitised data. The dark areas represent errors of commission, whilst the light represent errors of omission, by URBS. Considerable errors of omission have occurred in two localities (the circles on figure 9). Most of these are due to the poor performance of the initial low-level land-cover segmentation, particulary omission of pixels from the 'built' and 'structure' cover-type classes. Consequently, they are not considered by the intermediate algorithms used in URBS, and cannot therefore be recognised as urban areas. Errors of this type may be overcome in the future through an extension of URBS spatial rule-base to allow a probabilistic technique to be developed based around the spatial information contained within the XRAG structure.

It is considered that the XRAG structure may additionally be used for limited spatial querying of the segmented information. The development of multiple inheritance (Hu 1990) into the XRAG structure would allow objects to belong to a variety of classes at any one time. By allowing objects to have multiple class ID's, it should be possible to interactively query them in any order and for different spatial relationships to be extracted through the segmentation algorithms.

CONCLUSIONS

The preliminary results obtained from this study suggest that the extended-RAG (XRAG) data structure seems to have the potential to perform the type of complex spatial searching required for land-use mapping in images of urban areas. It is particularly useful when incorporated within an expert system, using *a priori* knowledge to formulate the spatial rule-base. It is believed that the XRAG structure also holds the potential for a limited level of interactive, spatial querying - typically performed within Geographic Information Systems - of the identified objects.

Future work will concentrate on the development of a more extensive spatial rule-base for URBS in a attempt to overcome the type of errors outlined. The use of the low-and intermediate-level processing algorithms, in addition to the results from URBS to develop an expert system to segment urban land-use. Also, work will be undertaken to explore the potential of the XRAG structure for simple spatial querying and analysis within an interactive environment.

In addition to the above, an independent investigation is being conducted into use of artificial intelligence and computer vision techniques for the segmentation and mapping of urban land cover in high spatial resolution multispectral imagery. The results of these procedures may improve the input into URBS and future algorithms, allowing for fully automated segmentation of urban areas.

ACKNOWLEDGEMENTS

This work was carried out as part of the ESRC/NERC Joint Programme of research into Geographical Data Handling. Thanks to Dr. Mike Barnsley for his supervision and patience during this work. Thanks to my colleagues at UCL, Kevin Morris, Asad Ahamid and Paul Schooling for interesting and helpful conversations into coding considerations. Thanks also to Graham Sadler (Geography, Birkbeck College) for advice and input at various stages of this work.

REFERENCES

Baraldi, A., and Parmiggiani, F., 1990. Urban area classification by multispectral SPOT images. *IEEE Trans. on Geoscience and Remote Sensing*, **28(4)**: 674-680.

Barnsley, M.J., Barr, S.L., and Sadler, G.J, 1991. Spatial Re-Classification of Remotely Sensed Images for Urban Land Use Monitoring. *Proc. Spatial Data 2000*, 17-20 September 1991, Oxford University, (Nottingham : Remote Sensing Society), 106-117.

Freeman, H., 1961. On the encoding of arbitrary geometric configurations. *IEEE Trans. Elec. Computers*, **10**: 260-268.

Gonzalez, R.C, and Wintz, P., 1987. *Digital Image Processing*. (New York:Addison-Wesley), 275-287.

HMSO 1984. Key Statistics for Urban Areas. *Great Britain, Office of Population Censuses and Surveys Registrar General, Census 1981*. (HMSO : London), 203p.

Hu, D., 1990. *Object-Oriented Environment in C++*. (Oregon:MIS Press), 1-109.

Huertas, A. and Nevatia, R., 1988. Detecting Buildings in Aerial Images. *Computer Vision, Graphics and Image Processing*, **41**: 131-152.

Mckeown, D.M., 1988. Building Knowledge-based systems for detecting man-made structures from remotely sensed imagery. *Phil. Trans. R. Soc. Lond, A*. **324**: 423-435.

McKeown, D.M., 1991. Feature extraction and image data for GIS. *Proc. Spatial Data 2000*, 17-20 September 1991, Oxford University, (Nottingham:Remote Sensing Society), 3-11.

Mehldau, G., and Schowengerdt, R.A., 1990. A C-extension for rule based image classification systems. *Photogrammetric Engineering and Remote Sensing, 56(6)*: 887-892.

Nichol, D.G., 1990. Region adjacency analysis of remotely-sensed imagery. *International Journal of Remote Sensing*, **11(11)**: 2089-2101.

Sadler, G.J., and Barnsley, M.J., 1990. Use of population density data to improve classification accuracies in remotely-sensed images of urban areas. *First European GIS Conference (EGIS'90)*, Amsterdam, The Netherlands, 10-13 April 1990, (Utrecht:EGIS Foundation), 968-977.

Sadler, G.J., Barnsley, M.J., and Barr, S.L., 1991. Information extraction from remotely-sensed images for urban land analysis. *Second European GIS Conference (EGIS'91)*, Brussels, April 1991, (Utrecht:EGIS Foundation), 955-964.

Townshend, J.R.G., 1980. Spatial resolving power of earth resources satellites. *Progress in Physical Geography*, **5**, 32-55.

Wharton, S.W., 1982. A contextual classification method for recognising land use patterns in high resolution remotely-sensed data. *Pattern Recognition*, **15**, 317-324.

Whitehouse, S., 1990. A spatial land use classification of an urban environment using high resolution multispectral satellite data. *Proc. 16th Ann. Conf. Remote Sensing Society, Remote Sensing and Global Change*, University Collage Swansea, 19-21 September 1990, (Nottingham: Remote Sensing Society), 433-437.