Geoinformation for European-wide Integration, Benes (ed.)

Simulating the scale dependence of urban land cover inference for the region-based derivation of fine scale urban land use information

Stuart Barr
School of Geography, University of Leeds, Leeds LS2 9JT, United Kingdom

Keywords: urban land use, land cover structure, scale sensitivity

ABSTRACT: Images acquired by very high spatial resolution multispectral satellite sensors are one means by which urban land use information can be derived. It has been argued that this may be achieved using a two-stage region-based approach, where initially derived land cover parcels (regions) are subsequently analysed, in terms of their spatial and morphological structure, to infer land use. Although initial results of the above approach have been promising, little work to date has been conducted into quantifying its scale dependence. This is particularly the case in relation to how inferred region morphology and spatial composition changes as a function of image spatial resolution. In order to address this issue this paper employs fine scale land cover digital map data to generate simulated classified images at resolutions of between 2–6m. Subsequent region-based structural analysis shows that inferred region morphology, particularly for man-made objects (buildings and roads), is severely degraded at spatial resolutions below 2m, making the inference of certain residential land use types problematic. The implication of these findings are that for certain spatially fragmented residential types, characterised by a relatively large number of small detached houses, region-based structural inference of urban land use using multispectral images acquired by sensors such as IKONOS-1 (4m) is likely to be problematic.

1 INTRODUCTION

The inference of accurate and consistent urban information using images acquired by optical Earth observation sensors has traditionally been found to be a problematic task (Forster, 1985). The poor performance of Earth observation in relation to the mapping of urban scenes is commonly attributed to the unsuitability of traditional spectrally based per-pixel classification algorithms in relation to the themes of interest (Gastellu-Etchegorry, 1990). Many urban remote sensing studies are primarily concerned with obtaining information on land use (*i.e.,* information on the different types of residential development present within an urban scene as well as, information on the spatial distribution of industrial, commercial and service-based industries) rather than the more commonly inferred information of land cover (*e.g.,* in the context of urban scenes information on the different roof materials of buildings, road materials and intra-urban vegetation). Traditional classification approaches are unsuitable because urban land use, in terms of the image acquisition process, is often comprised of a complex spatial assemblage of many different land cover types (Barnsley and Barr, 1996). Thus, many urban land use categories often exhibit considerable variability in their spectral response patterns and multivariate multi-modal distributions that break the multivariate normal requirements of parametric classification approaches (e.g., the maximum likelihood algorithm).

In order to address this problem a number of alternatives approaches have been developed for the inference of urban land use information from very high spatial resolution remotely sensed images. One of the most promising of these alternatives is the use of reclassification approaches (Wharton, 1982; Gong and Howarth 1992). Reclassification attempts to infer urban land use information from a multispectral image on the basis of two functions, rather than the traditional single classification function, and can be generically expressed as a composition (Barr and Barnsley, 1999)

$$I \circ S = \{(i \mapsto s) : i \in I, s \in S, \exists f \in F([i \mapsto f] \wedge [f \mapsto s])\} \quad (1)$$

This composition states that the inference of urban land use information is, conceptually at least, achievable by defining a mapping from the multispectral image (I) to the land cover thematic domain (F), and thereafter, by defining a mapping from the

derived land cover to the land use thematic domain (S). The first mapping of Equation 1 ($f : I \mapsto F$) involves the inference of land cover information from the original multispectral image and is commonly achieved by means of traditional per-pixel classification approaches. The second mapping, that from the land cover to the land use thematic domain ($f : F \mapsto S$), may be achieved using either per-pixel reclassification approaches (Wharton 1982; Barnsley and Barr 1996), or region-based structural pattern recognition approaches (Barr and Barnsley, 1999).

In relation to the use of the above multi-function approach to urban land use inference, it has been argued for images acquired by a new series of very high spatial resolution (1–4m) multispectral Earth observation sensors, such as IKONOS (Space Imaging), Quickbird (Earth Watch) and OrbView (Orbimage), that a region-based structural pattern recognition approach is the most appropriate means by which to successfully derive the second mapping (Barr and Barnsley, 1999). The basis for this assertion is that at the spatial resolution of the images acquired by these sensors it should be possible to describe classified land cover in terms of regions whose morphological and spatial pattern facilitates land use inference (Barr and Barnsley, 1999).

However, while some success has been reported in the use of region-based computer vision approaches to the derivation of land use from very high spatial resolution ($<< 1m$) panchromatic aerial photographs (Mahldau and Schowengerdt, 1990), the successful application of very high spatial resolution Earth observed images has yet to be fully realised. One possible reason for this lack of current success may be that while such images represent a significant improvement in spatial resolution compared to those acquired by traditional sensors (*c.f.* SPOT-HRV XS (20m) and Landsat TM (30m)) this may still be insufficient to allow the accurate and consistent characterisation of region morphology and spatial pattern.

In order to investigate this issue, this paper employs fine spatial scale land cover digital map data to generate simulated images at spatial resolutions of between 1–6m for a range of different urban land use types. These images are then classified to the level of land cover and compared in order to quantify how both land cover region morphology and spatial pattern changes and how this influences the degree of separability that exists between the different urban land use types under investigation.

2 METHODS

2.1 *Rasterised 1m land cover data*

Ordnance Survey (OS) 1:1,250 scale Land-Line.93+ digital map data covering a small urban area was used to generate a 1m rasterised urban land cover map. This data set was used for the generation of the simulated multispectral images. The digital map data consisted of up to 32 different digital vector features, which were processed using ARC/INFO to generate a single coverage. This was then topologically structured to produce a coverage that contained polygonal entities of Road, Built, Tree (areas of extensive tree coverage), Open-Space (all non-tree areas of open space and vegetation) and Water-bodies. The resulting coverage was then rasterised (in the Grid module of ARC/INFO), resampled to a spatial resolution of 1m and exported to the image format used in this study.

2.2 *Simulated multispectral images*

The derived 1m spatial resolution land cover map was used in association with field spectroscopy measurements to generate a 'pure' multispectral image corresponding to the known reflectance of the five land cover types for the 4 centre points of the spectral bandwidths that correspond to the IKONOS multispectral instrument ($480\eta m$, $551\eta m$, $665\eta m$ and $805\eta m$). The spectroscopy measurements were acquired at solar noon (solar zenith of 32°, solar azimuth of 175°) and recorded from nadir. For the Built class, reflectance measurements for the 4 wavelengths were obtained by sampling planimetric (flat) slate roof tiles. Spatially scaled multispectral images for spatial resolutions of between 2-6m were subsequently generated by convolving a 2-D Gaussian Point Spread Function (PSF) with the simulated reflectance images (Figure 1). Both across-track (X) and along-track (Y) the convolution was performed by spatially oversampling the desired Ground Instantaneous Field of View (GIFOV) by a factor of 1.25.

2.3 *Classification*

In order to produce a classification of the simulated multispectral images a standard feed-forward backward propagation neural network (FFBP-ANN) classification approach was employed. The FFBP-ANN consisted of 3-layers (input, hidden, output) of 4-12-5 sigmoidal activation function units respectively. For each simulated multispectral image, a random selection of class training pixels were sampled – using the original 1m digital map data – in order to train the corresponding FFBP-ANN. Training involved at each epoch the random, without replacement, presentation of each pixel training-pattern to

the network, with network adjustment being performed for each pattern presented. Training was limited to a maximum of 10,000 epochs or was terminated when the average output unit Sum Square Error (SSE) fell below 0.2 (corresponding to an overall SSE for the training pattern size of ≈120.0). Once an FFBP-ANN had been trained for each simulated image, its entire population of pixels were presented for classification in order to generate a series of classified urban land cover maps at spatial resolutions of 2m, 3m, 4m, 5m and 6m respectively.

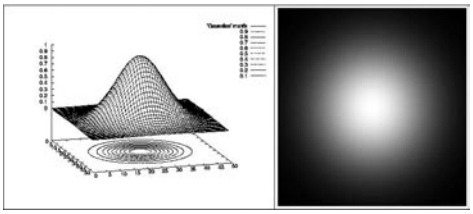

Figure 1. An example 2-D Gaussian Point Spread Function template and its representation as an image suitable for convolution.

2.4 *Post-classification analysis*

Post-classification analysis involved the use of a region-based structural analysis system SAMS (Structural Analysis and Mapping System) which derives structural information about the regions (parcels) present in ordinal or nominal digital map data-sets (Barr and Barnsley, 1997). Region boundaries are identified using a simple contour-tracing algorithm, represented using Freeman chain-codes and stored in a Region Search Map (RSM). The RSM is processed to derive further information about the structural characteristics of the observed scene, including various morphological properties (area, perimeter and various measures of shape), as well as spatial relations (adjacency, containment, distance and direction). This information is used to populate a graph-theoretic data model, known as XRAG (eXtended Relational Attribute Graph) (Barr and Barnsley, 1997).

On the basis of ground survey and analysis of analogue aerial photography, five principal land use types were recognised in the scene under investigation; namely, 1930s, 1980s, 1990s residential, Hospital and School complexes. Extracts for each of these land use types were derived (Figure 2) from both the original 1m land cover map and the classified simulated images. The land use extracts for the original 1m land cover map and also the classified simulated images were processed using SAMS to characterise their Built region land cover structure. The structural description of Built region morphology was obtained by deriving the properties of area, compactness and geographic centroid. The spatial pattern of Built regions was characterised by deriving Gabriel graphs (Jaromezyk and Toussaint, 1992), the edges of which represent relative neighbours in terms of the distance between the geographic centroids of Built regions (Figure 2).

The resulting structural measurements were subsequently used to generate a multivariate statistical description of each land use type for each of the classified multispectral images. The statistics consisted of the mean structural vector and structural covariance matrix of each land use. These were then used in a transformed divergence analysis, which saturated in the range $0 \le TD \le 100$, in order to characterise the change in statistical structural separability that occurs as a function of changing spatial resolution.

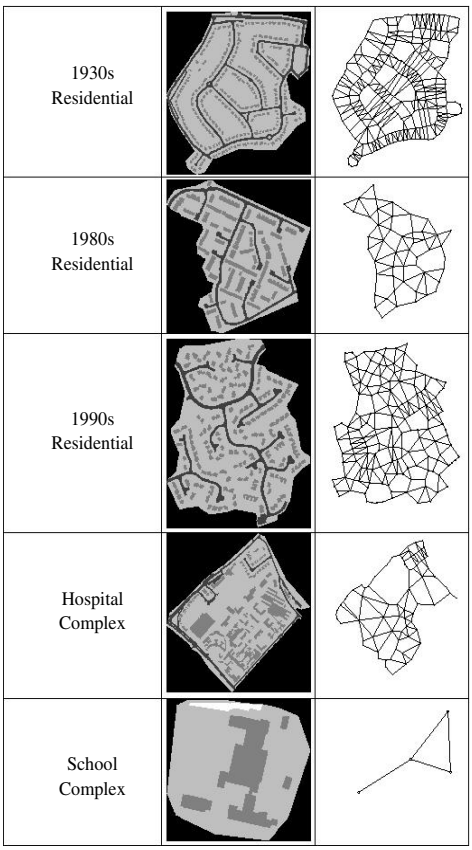

Figure 2. 1m spatial resolution rasterised land cover images showing the urban land use types selected for analysis and their corresponding Gabriel graphs used in the analysis.

3 RESULTS

3.1 *Classifications*

Figure 3 shows for each simulated multispectral image its corresponding FFBP-ANN learning curve. In all cases none of the networks converged to an average output unit SSE of less than 0.2 (overall SSE $\approx$ 120.0), although the training procedure for the 2m image had a SSE very close to this threshold (120.5). As one would expect, the final training SSE error increases as a function of decreasing spatial resolution. This is because, as the spatial extent of the applied Gaussian PSF template is increased one is more likely to obtain pixel vectors that are increasingly different from the original wavelength specific reflectance assigned to a class; due to the derived reflectance being a non-linear function of different class reflectance's. This in turn results in individual classes exhibiting increased spectral variance and increases the likelihood of classes exhibiting feature space overlap. These two features mean that as the spatial resolution is decreased it becomes harder for each FFBP-ANN to converge to a low overall SSE.

The one case where the above monotonic relationship does not hold is for the 6m spatial resolution case, which consistently exhibits a slightly lower training error compared to the 5m case. One possible explanation of this feature may be that at this spatial resolution, spatially proximal but disjoint objects of the same class are included in the same PSF, resulting less class variance being introduced than in the 5m simulated multispectral images.

While each FFBP-ANN does not achieve a particularly low overall SSE, an analysis of Figure 3 suggests, however, that they do mange to converge to a relatively stable network state and are suitable for use in the subsequent classification of each simulated multispectral image.

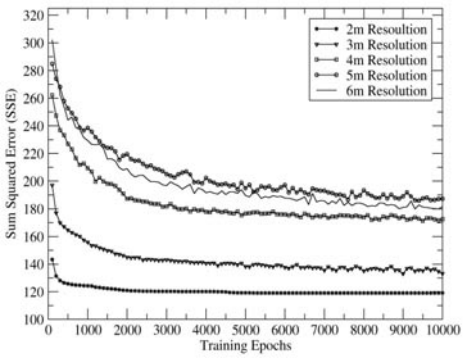

Figure 3. Plot showing the learning curves of the feed-forward backward propagation classification of the full simulated image data-sets at different spatial resolutions.

3.2 *Region-based structural analysis*

In order to compare the region-based structural pattern of the five land use extracts under investigation for each simulated multispectral spatial resolution, the transformed divergence was first derived for the original 1m land cover data set. Table 1 shows the results of this process. This shows for the original regions, without any structural ambiguity being introduced as a function of Gaussian PSF scaling and subsequent classification, that the majority of classes exhibit a high pairwise statistical structural difference in terms of the area, compactness and proximity of Built regions. Indeed, the only pairwise combination that falls below a value of 85 – values above which are commonly cited as being indicative of good separability (Mather, 1999) – is the 1930s/1990s pairwise combination.

Table 1. Transformed divergence results for the 1m spatial resolution urban land use types of Figure 1. The results have been derived from the statistical distribution of Built region area, compactness and Gabriel graph edge length.

Type	1930s	1980s	1990s	Hos	School
1930s	0.0	93.5	64.1	100.0	100.0
1980s	93.5	0.0	100.0	99.9	100.0
1990s	64.1	100.0	0.0	100.0	100.0
Hos	100.0	99.9	100.0	0.0	100.0
School	100.0	100.0	100.0	100.0	0.0

A similar Built region transformed divergence analysis for each classified simulated multispectral image is shown in Table 2 in terms of the average transformed divergence obtained for all pairwise land use combinations. This shows that on average good separability is still maintained as the spatial resolution of the images is reduced. Indeed, these results suggest that reducing the spatial resolution is likely to have relatively little affect, at least statistically, on the potential to quantitatively discriminate between the different land use types.

Table 2. Transformed divergence results for each classified simulated high spatial resolution image data-set. The results have been derived from the statistical distribution of Built region area, compactness and Gabriel graph edge lengths of the urban land use types.

Resolution (m)	Average TD
1m	95.7
2m	93.4
3m	89.5
4m	87.7
5m	88.6
6m	91.1

However, while decreasing spatial resolution does not seem to significantly change the average separability of the land use types under investigation, this is not because they retain their original structural pattern. This is shown by Table 3 which expresses for each land use type the pairwise trans-

formed divergence between the structure of the Built regions in the original 1m land cover digital map data and the Built regions at each simulated spatial resolution. As a value of 0 shows that a pairwise combination are, in multivariate statistical terms, identical and a value of 100 shows that a pairwise combination exhibit no statistical overlap, this table shows that as spatial resolution decreases the Built region structure of each land use changes significantly. At a spatial resolution of 5m all of the land use types, apart from the Hospital complex, exhibit a Built region structure that is entirely separable. This would clearly seem to indicate that there is little similarity in the Built region structure of images at low spatial resolutions compared to that present in the land use extracts of the original 1m land cover digital map data.

Table 3. Transformed pair-wise divergence results between the base 1m spatial resolution land cover map and the classified simulated high spatial resolution image data-sets expressed in terms of the individual urban land use types under investigation. The results have been derived from the statistical distribution of Built region area, compactness and Gabriel graph edge lengths of the urban land use types.

	1-2m	1-3m	1-4m	1-5m	1-6m
1930s	32.9	97.3	99.9	100.0	100.0
1980s	37.7	89.3	96.5	100.0	100.0
1990s	94.9	100.0	100.0	100.0	100.0
Hos	9.1	40.4	61.9	64.1	89.4
School	99.5	99.9	100.0	100.0	100.0

4 DISCUSSION

In combination the results presented in Table 1 and Table 2 suggest that changes in spatial resolution have relatively little effect on the ability of quantitatively separate different land use types on the basis of their Built region structure. Given the high transformed divergence achieved for the original 1m digital map data and the relatively small decreases experienced as function changing spatial resolution one may be tempted to concluded that their is no fine scale sensitivity to the ability of structurally distinguish different urban land use types.

However, such a conclusion ignores the fact that as the spatial resolution is decreased individual land use types seemingly exhibit significant changes in their structural composition, as shown by the results presented in Table 3. For example, Table 3 reveals for the 1990s residential class at 2m that the derived Built regions share virtually no similarity to those for this class in the original 1m digital map data. While less dramatic, the 1930s and 1990s also experience a noticeable loss of structural similarity between the 1m digital map data and the Built regions derived from the simulated 2m spatial resolution multispectral image.

A visual appreciation of just how dramatic the change in structural composition is as function of changing spatial resolution can be ascertained from Figure 4, which shows the 1930s land use extract for spatial resolutions of 2m, 4m and 6m. A comparison between the classification results and Gabriel graph structure of Figure 4 and the original structure of this class in Figure 2, reveals that by a spatial resolution of 4m both Built region morphology and Built region proximity has changed significantly. Visually at a spatial resolution of 6m Built region morphology and proximity bears very little resemblance to that of the 1m digital map data and quantitatively these two structural compositions are found to be 100% separable.

It would seem, therefore, that while it may be possible to quantitatively distinguish between different urban land uses as one changes the spatial resolution of observation, this would seem to be clearly at the expense of changes in the original land cover structural composition. If the interest lies not only in the actual inference of land use but also in understanding this inference process in terms of the structure of the underlying land cover primitives (regions) used then it can be argued that one requires a spatial resolution of at least 2m. This is because, beyond this spatial resolution the original structural composition of land cover is predominantly lost.

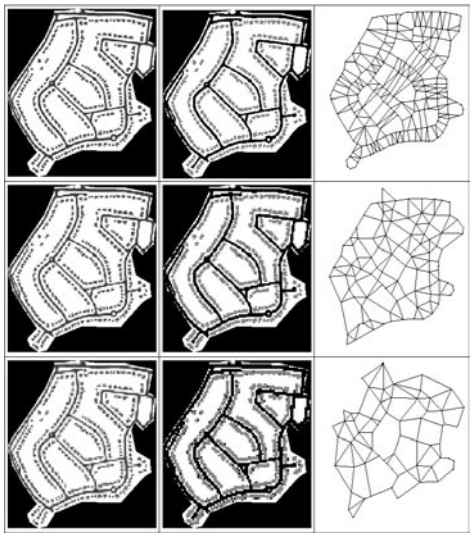

Figure 4. Simulated (scaled) images, classification results and derived Gabriel graph spatial structure for the 1930s urban land use type at spatial resolutions of 2m, 4m and 6m respectively.

5 CONCLUSIONS

This paper has presented preliminary results on the scale sensitivity of a multi-function region-based approach to the inference of urban land use information from very high spatial resolution multispectral Earth observed images. The results, generated on the basis of a simple simulation of a series of very high spatial resolution images of an urban scene, suggest that while it is still possible to accurately distinguish between different urban land use types as spatial resolution is decreased from 1m–6m, the nature of the region-based land cover structure that facilitates this changes significantly. On the basis of the transformed divergence analysis presented, it would seem that a spatial resolution of at least 2m is required if one wishes to retain the original intrinsic land cover structure within such an inference chain.

ACKNOWLEDGEMENTS

The author is grateful to the Ordnance Survey (GB) for supply of the 1:1,250 scale Land-Line.93+ digital map data used in this study. The field spectrometer reflectance measurements of man-made materials used in this study were originally collected by Louise Mackay, School of Geography, University of Leeds.

REFERENCES

Barnsley, M.J. & Barr, S.L. 1996. Inferring urban land use from satellite sensor images using kernel-based spatial reclassification. *Photogrammetric Engineering and Remote Sensing* Vol. 62 (No. 8): 949-958.

Barr, S.L. & Barnsley, M.J. 1999. A syntactic pattern-recognition paradigm for the derivation of second-order thematic information from remotely sensed images. In Atkinson, P.M. & Tate, N.J. (eds) *Advances in Remote Sensing and GIS Analysis.* Chichester: John-Wiley: 167-184.

Barr, S.L. & Barnsley, M.J. 1997. A region-based graph-theoretic data model for the inference of second-order thematic information from remotely sensed images. *International Journal of Geographical Information Science* Vol. 11 (No.6): 555-450.

Eyton, J.R. 1993. Urban land use classification and modelling using cover type frequencies. *Applied Geography* Vol 13: 111-121.

Forster, B.C. 1985, An examination of some problems and solutions in monitoring urban areas from satellite platforms. *International Journal of Remote Sensing* Vol 6: 139-151.

Gastellu-Etchegorry, J, P. 1990. An assessment of SPOT XS and Landsat MSS data for digital classification of near-urban land cover. *International Journal of Remote Sensing* Vol 11: 225-235.

Gong, P. & Howarth, P.J. 1992. Frequency-based contextual classification and grey-level vector reduction for land use identification. *Photogrammetric Engineering and Remote Sensing* Vol 58: 423-437.

Jaromezyk, J.W. & Toussaint, G.T. 1992. Relative neighborhood graphs and their *relatives. Proceedings of the IEEE* Vol 80: 1502-1517.

Mather, P.M. 1999. *Computer processing of Remotely-Sensed Images*. Chichester: John-Wiley.

Mehldau, G. & Schowengerdt, R.A. 1990. A C-extension for rule-based image classification systems. *Photogrammetric Engineering and Remote Sensing* Vol 56 (No. 6): 887-892.

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