PARCEL-BASED APPROACHES TO THE CLASSIFICATION OF FINE SPATIAL RESOLUTION IMAGERY: EXAMPLE METHODOLOGIES USING HRSC-A DATA

Geoffrey M. Smith* and Andrea Hoffmann**
*Section for Earth Observation, Centre for Ecology and Hydrology, Monks Wood, Abbots Ripton, Huntingdon, Cambridgeshire, PE17 2LS, United Kingdom.
gesm@ceh.ac.uk
**German Aerospace Center (DLR), Institute of Space Sensor Technology and Planetary Exploration, Adlershof, Rutherfordstr. 2, D-12489 Berlin, Germany.
andrea.hoffmann@dlr.de

Working Group IV/4

KEY WORDS: Fine spatial resolution, classification, per-parcel, scene components, HRSC-A

ABSTRACT

Large amounts of remotely sensed data are collected from airborne platforms with spatial resolutions from 2 m down to 15 cm and similar satellite systems are coming on line. Conventionally they have been analysed by manual airphoto interpretation techniques, which are time-consuming, subjective and expensive, therefore it is likely that semi-automated approaches will be attempted.

Data with a fine spatial resolution would appear to be a major advantage for mapping as the proportion of mixed land cover pixels is reduced. It does, however, present a new set of problems related to how these types of data should be analysed using automated and semi-automated techniques.

At finer spatial resolutions the data recorded by an instrument begin to cross a scale boundary where it is related not to the character of objects or areas as a whole, but to components of them. For instance, with 30 m Landsat TM data, pixels within a suburban area will each represent a mix of buildings, vegetation and bare ground etc. that would generate a reasonably unique spectral signature and they could usually be classified by conventional semi-automated techniques. As the spatial resolution gets finer pixels begin to represent components within the suburban area, such as buildings, gardens, paved areas or mixtures of these. This increases the number of classes that must be mapped, reduces their spectral separability and increases the complexity of the result.

This paper addresses two problems associated with the automated analysis of fine spatial resolution data. Firstly, how to classify the images into meaningful surface features when training areas would be difficult if not impossible to find. Secondly, how to integrate the detailed information returned by the classifier to more usable classes and more appropriate scales.

To classify fine spatial resolution data it will be necessary to identify a set of scene components which manifest themselves at that particular spatial resolution. The scene components must then be labelled as real world features after an assessment against the spatial resolution. Recent advances in integrated Geographical Information Systems have allowed the development of automated procedures for analysing remotely sensed data on a per-parcel basis, rather than the conventional per-pixel basis. This allows aggregation of the scene component information within a region, the consideration of context, both within the region and beyond it, and the application of knowledge-based rules at the parcel level to classify the parcels into meaningful land cover classes.

A case study is provided which highlights the problems associated with classifying fine spatial resolution images and also describes the parcel-based approach to the classification with scene components. The case study used data from the High Resolution Stereo Camera-Airborne (HRSC-A) system with an original spatial resolution of 15 cm to map land cover type within forest stand parcels in the Tharandter Forest, Germany.

Although this methodology for image classification makes no radical advances in terms of correspondence accuracy compared to conventional per-pixel classification of coarse spatial resolution data it does hold its own and provides a much richer set of results and an opportunity for further analysis.

1 INTRODUCTION

Large amounts of remotely sensed data are now being collected from airborne platforms carrying digital scanners and digital cameras with spatial resolutions from 2 m down to 15 cm. Spaceborne instruments are now being launched with spatial resolutions between 1 and 4 m (Fritz, 1996; Aplin *et al.*, 1999; Konecny, 1999).

These systems have been and will be used to address mapping issues in locating and identifying objects or areas on the surface. Conventionally, images with such fine spatial resolutions would have been analysed by manual airphoto interpretation techniques using a range of subjective decisions based on tone, texture, shape, size and context. This approach is time-consuming, subjective and expensive. As the majority of image products are now provided in a digital form, they lend themselves to the application of pixel-based, semi-automated analysis and classification techniques.

The fine spatial resolution of the data now becoming available would at first appear to be a major advantage for mapping applications compared to conventional satellite systems (spatial resolutions coarser than 10 m). However, it must be remembered that similar reasoning proceeded the launch of Landsat 4 (carrying the 30 m spatial resolution Thematic Mapper (TM)) and the Systeme Probatoire de l'Observation de la Terre (SPOT-1) (carrying the High Resolution Visible (HRV) instrument with a 20 m spatial resolution) in the 1980s (Toll, 1984; Gastellu-Etchegorry, 1990). Work comparing TM and HRV with the established 80 m spatial resolution Landsat Multispectral Scanner System (MSS) found that finer spatial resolutions actually reduced classification accuracy for certain land cover types. The coarse spatial resolution of the MSS smoothed out spatial complexity within heterogeneous land cover types, such as urban, as scene components, such as bulidings and vegetation, become lost within a pixel (Bruniquel-Pinel and Gastellu-Etchegorry, 1998).

It was found that although finer spatial resolutions reduced the proportions of mixed pixels (Markham and Townshend, 1981), the number of detectable classes increased and these classes became less separable spectrally (Gastellu-Etchegorry, 1990). In fact, in some case the classes that are required to be mapped may disappear altogether, for instance, suburban becomes buildings and vegetation as the spatial resolution gets finer.

The current move toward even finer spatial resolution data sets should raise the same questions concerning the interaction of radiation with the surface and how these types of data should be analysed using automated and semi-automated techniques. The first area for consideration is the type of features that influence the response recorded for a pixel. For instance, with 30 m Landsat TM data, pixels within a suburban area will each represent a mix of buildings, vegetation and bare ground etc. that together would generate a reasonably unique spectral signature that could be used in conventional semi-automated classification techniques. As the spatial resolution gets finer (around 5 m) the pixels begin to represent components within the suburban area, such as buildings, gardens, paved areas or mixtures of these. At the spatial resolution of the sensor discussed here (less than 2m) this problem is exaggerated and the pixels may represent facets of buildings, vehicles, pavements, roads, grass, tree crowns and even people (Figure 1). The issues described previously between TM and MSS data are again relevant.

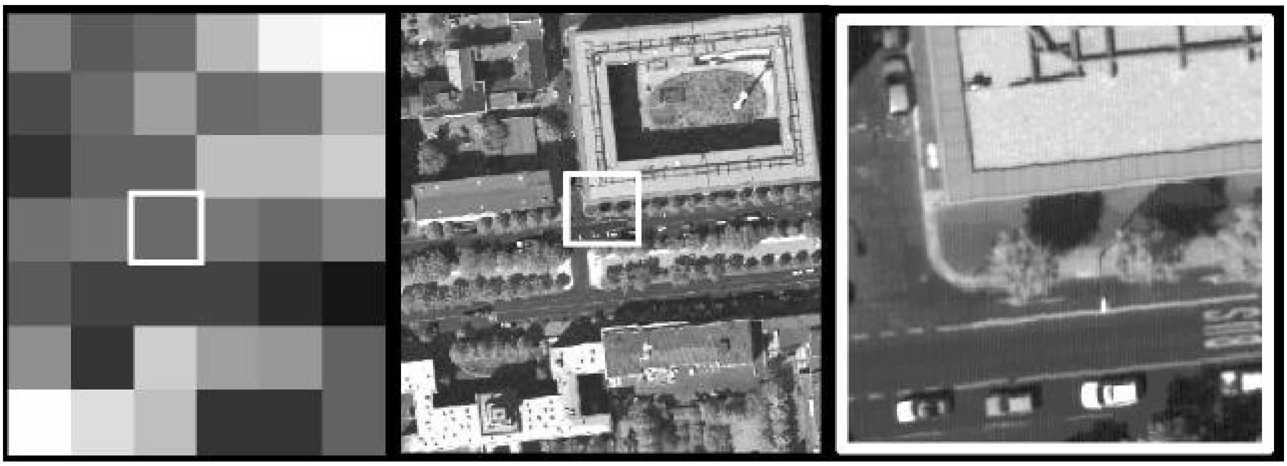

Figure 1. A demonstration of the issues surrounding the spatial resolution of different instruments. Left to right: Landsat TM 30 m data, HRSC-A 30 cm data with TM pixel identified, HRSC-A 30 cm data of the TM pixel enlarged.

As the spatial resolution of a sensor is equivalent to the scale of the observations it makes, then an appropriate spatial resolution, as with scale, is related to the environment under investigation and the information to be extracted

(Woodcock and Strahler, 1987). At finer spatial resolutions a scale boundary is crossed where the data recorded for each pixel is related not to the character of object or area as a whole, but to components of it and this requires a re-definition of the information that can be extracted. As the spatial resolution of an instrument is fixed, then simple pixel-based analyses may be invalid for certain environments and tasks. In fact, Woodcock and Strahler (1987) estimated appropriate spatial resolutions for mapping different land cover types (Table 1). The fine spatial resolution data now becoming available fall well below the lower spatial resolution for forest mapping and even further below that of urban mapping which is being suggested as a major application.

Application	Spatial resolution (m)
Forest	7 - 8
Urban	10 – 15
Agricultural	≈ 250

Table 1: Spatial resolutions suggested by Woodcock and Strahler (1987) for different applications.

The above discussion is borne out by the lack of information concerning the classification and analysis algorithms which are intended to be used for producing value added products from the new fine spatial resolution satellite systems.

2 OTHER STUDIES

A number of approaches have been used and developed for the semi-automated analysis of fine spatial resolution data.

Gong *et al.* (1994) classified scan digitised aerial photography with a conventional supervised maximum likelihood classification procedure. The results were reasonably successful, but the site under investigation had a simple structure with sparsely distributed dark pine trees on a light background. This facilitated the easy selection of training areas for the spectrally distinct classes.

As a demonstrator for the new fine spatial resolution spaceborne sensors Aplin *et al.,* (199?) analysed 4 m spatial resolution data recorded with an airborne scanner. The data were classified using supervised maximum likelihood classification into conventional land cover classes (urban, arable crops, grassland etc.) using training areas. As would be expected, the agricultural land produced the cleanest results with problems occurring within the urban areas. The urban problems were most likely the result of the true complexity of the urban landscape at 4m spatial resolution. To improve the per-pixel results in the urban areas low pass and texture filters were applied, but this had the disadvantage of affecting the whole image. The dominant land cover type from the classification within a land parcel was then determined to produce a final parcel-based classification.

Barr and Barnsley (1999) used 2 m spatial resolution airborne scanner data to analysis urban areas. The data were classified into 10 classes by conventional per-pixel maximum likelihood classification. The classification accuracy was acceptable, but the results contained considerable ‘clutter’ due to classification error and the real complexity of the urban environment. A structural filter was then used to reduce the clutter by recoding pixels within a moving window based on their spatial context.

3 AIMS

This paper describes a methodology for classifying fine spatial resolution data to land cover types. The problem faced was two fold; firstly, how to extract meaningful information from the pixels within the fine spatial resolution images and secondly, how to integrate this detailed information at the pixel level to usable classes and appropriate scales.

4 PROPOSED METHODOLOGY

Each pixel within an area to be classified will contain some useful information, although this may be difficult to interpret as a definite land cover class in isolation and can therefore only be referred to as an arbitrary scene component. Neighbouring pixels, even though they belong to the same area, land cover type or real world object, may be completely different classes. Some mixed pixels will be present, however, the fine spatial resolution of the data does reduce their number. Therefore, even within a small area the number of classes and mixtures could be enormous making training for supervised classifications virtually impossible.

The proposed methodology begins therefore with an unsupervised per-pixel classification to identify spectral clusters which represent the range of scene components present rather than a set of *apriori* land cover types. The actual number and types of scene components present within the scene will most likely be unknown and will depend on the target environment and the spatial resolution of the sensor. To separate out subtly different scene components the number of spectral clusters identified by the unsupervised classification should be significantly larger than the expected number of scene components. This will consequently result in some scene components being represented by a number of spectral clusters.

As the number of spectral clusters exceeds the number of scene components, an aggregation phase is required to merge spectral clusters which effectively represent the same scene components. The control of the aggregation of spectral clusters may be considered as training, but not against land cover types in the conventional sense. Merging is required where there has been an over division of spectral feature space, but also where two or more variants of the same scene component, e.g. slate roofs with different slopes/aspects, occur. The aggregation can be performed manually by comparison between the raw and classified images or in an automated fashion by spatial/spectral clustering or the use of spectral libraries. All spectral clusters that represented the same scene components are identified and re-coded to a single scene component.

The meaningful scene components still produce a classified raster image that is poorly structured and difficult to interpret. Even after merging and identification there will be a certain amount of mis-classification between the scene components, especially where they are spectrally similar. To provide meaningful land cover type information the results must now be aggregated to a suitable scale or within selected regions. The final results could therefore be either estimates of scene component fractions or a classification not available directly from the scene components themselves. Recent advances in integrated Geographical Information Systems have allowed the development of automated procedures for analysing raster-based data on a per-parcel basis, rather than the conventional per-pixel basis, using a vector data set of land parcels (Smith and Fuller, 2000). This allows aggregation of information within a region, the consideration of context, both within the region and beyond it, and the application of knowledge-based rules at the parcel level driven by other data sets. The land parcel data set identifies the subdivision of the landscape into reporting units and can be derived either from conventional digital cartography or by the segmentation of a spatially degraded version of the original image.

In this methodology the per-parcel analysis extracts the pixels classified as scene components within the land parcel under examination and calculate the total numbers and fractions for each scene component present. This information may be all that is required if the scene components themselves are being assessed. To identify land cover types not represented by scene components at the land parcel level, it is necessary to process the scene component information and infer relationships between the scene components present and land cover type. A set of rules can be devised to identify a range of land cover types from the mixtures of scene components found within each land parcel. These rules could be in the form of a threshold for each component, ratios between components or indices that combine a number of components. The fractions of the scene components per-parcel can also be used as inputs to more complex analyses. In a similar way as the data from each of the image bands are used in conventional per-pixel classifications the component fractions can be used in clustering algorithms or statistical classifiers and the results applied to the whole region.

5 THE HIGH RESOLUTION STEREO CAMERA - AIRBORNE

This study used a data set from the High Resolution Stereo Camera - Airborne (HRSC-A) digital camera system. The HRSC-A is a multiple line pushbroom instrument (Table 2a) originally developed for planetary exploration and will be flown onboard the ESA Mars Express mission to be launched in 2003 (Neukum *et al.*, 1999). The instrument acquires nine image lines simultaneously with nine CCD line sensors mounted in parallel and behind one single optics (Table 2b). Five of these are panchromatic sensors are arranged at specific viewing angles and provide multiple stereo and photometric capabilities. The four other CCD lines are covered with different filters for the acquisition of multispectral images. The data sets from the camera combine the advantages of digital acquisition, very high resolution and the acquisition of both multispectral and elevation information.

The camera is mounted on a stabilised platform (ZEISS T-AS) to damp mechanical vibrations and to enforce near-nadir viewing. Position and orientation during flight navigation are measured continuously by means of differential GPS and INS. To get high accuracy measurements of the exterior orientation an APPLANIX integrated navigation system including a GPS receiver and a strap-down inertial navigation system are used. The INS is mounted directly on top of the sensor unit. With this navigation system the camera position can be determined during post-processing with an accuracy of ±2-3 cm (Wewel and Scholton, 1999).

A completely automated procedural software system has been built up for the use of HRSC-A in airborne application. It makes use of a set of systematically preprocessed image, orientation and calibration data.

High resolution multispectral orthoimages and DEM data have been acquired simultaneously for applications such as volcano monitoring, mapping of urban areas, forestry and agriculture, mapping of flood hazards, open coal mines and coastal zones. The potential of the HRSC-A system for photogrammetric surveys in urban areas has been shown in its operational use at several European cities (Renouard and Lehmann, 1999).

Technical Parameters	HRSC-A
Focal Length	175 mm
Total Field of View	37.8° x 11.81°
Number of CCD Lines	9
Stereo Angles	±12.8°; ±18.9°
Swath	11.8°
Pixels per CCD Line	5184
Pixel Size	7 µm
Radiometric Resolution	10 bit reduced to 8 bit
Read-out Frequency	450 lines/s
Mass	32 kg
Platform	Zeiss T-AS- gyro stabilized
Data Recording	SONY Digital Tape Recorder
Flight navigation	PC and GPS

Table 2a: Technical specifications of the HRSC-A.

CD Line	Band	Filter	Spectral Range [nm]	Viewing Angle [Degrees]
5	Stereo	Pan	585-765	+18.9
4	Red	Red	730-770	+15.9
3	Stereo	Pan	585-765	+12.8
2	Blue	Blue	395-485	+3.3
1	Nadir	Pan	585-765	0
2	Green	Green	485-575	-3.3
3	Stereo	Pan	585-765	-12.8
4	Infrared	Infrared	925-1015	-15.9
5	Stereo	Pan	585-765	-18.9

Table 2b: Waveband specifications of the HRSC-A.

6 FOREST CASE STUDY

The case study describes the application of this methodology for the classification of forest stand parcels into a range of forest land cover types using a supervised classification of the scene component fractions.

A HRSC-A data set of Tharandter Forest was acquired in October 1998 from a flight altitude of 3000 m producing a pixel size for the nadir image of 15 cm. From the full nine bands the multispectral data were extracted (green, red and infrared) and the data set was resampled to a spatial resolution of 30 cm. Tharandt is located in Saxony, near to Dresden (ca. 20 km south-west). The Tharandter Forest is a heterogeneous region, consisting mainly of mixed deciduous and coniferous forest stands with with large diversity and differences in stand ages. The area under investigation consisted mainly of spruce, fir, ash, beech, birch, larch and some recently reforested stands.

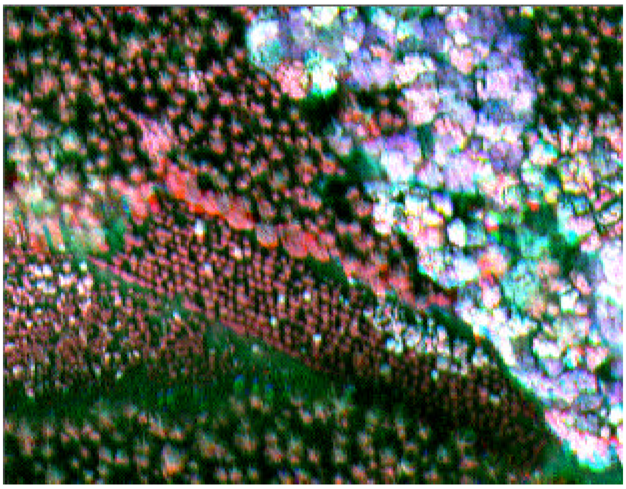

Figure 2. Example HRSC-A data for the Tharandter Forest.

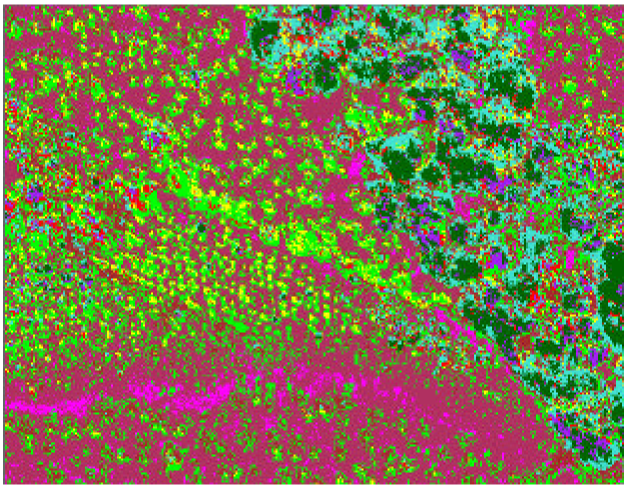

Figure 3. HRSC-A data classified as meaningful scene components.

The HRSC-A data were classified into 80 spectral clusters with an ISODATA unsupervised classification algorithm. By manual interpretation of a small area of the image the 80 spectral clusters were merged to create 9 meaningful scene components (Table 3). The land parcel data set was used to extract scene components fractions for each of the forest stands (For example, Figure 2).

By comparing the scene component fractions with the land cover types (Table 4) of the forest stand parcels it was possible to get a clearer understanding of their relationships (Figure 5). The deciduous scene components appeared to be positively related to the amount of deciduous canopy in the land cover types. The coniferous scene components did not show a similar relationship with conifer canopy amount, but fractional shadow appeared to be more indicative. This situation can be explained with reference to the canopy form of these two canopy types. Deciduous canopies tend to be larger and flatter than the smaller, but taller cone shaped coniferous canopies. These features result in coniferous canopies producing greater amounts of shadow which increase as the stands mature.

Code	Scene component
C1	Bright conifer canopy
C2	Dark conifer canopy
D1	Bright deciduous canopy
D2	Open deciduous canopy
D3	Dark deciduous canopy
G	Grass
B1	Dark background/understorey
B2	Open mixed canopy
S	Shadow

Table 3. Scene components identified after merging spectral clusters

Code	Land cover types
D	Mixed deciduous
L	Birch dominated
CL	Mixed birch coniferous
C	Coniferous dominated
N	Reforested coniferous

Table 4. Land cover types to be mapped in the Tharandter Forest

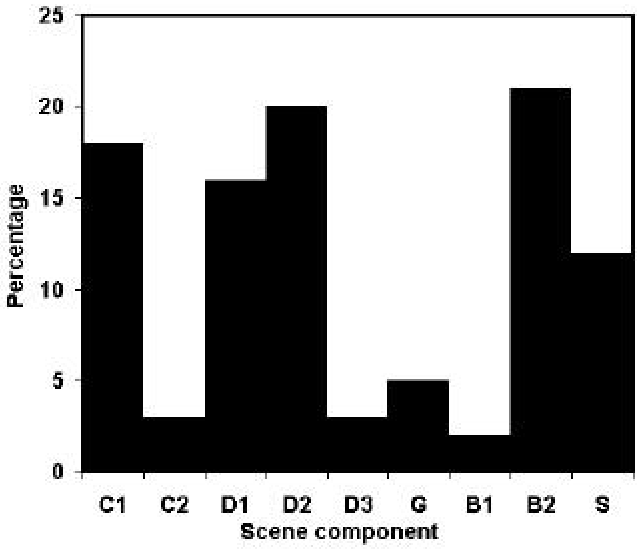

Figure 4. An example set of scene component fractions for a land parcel.

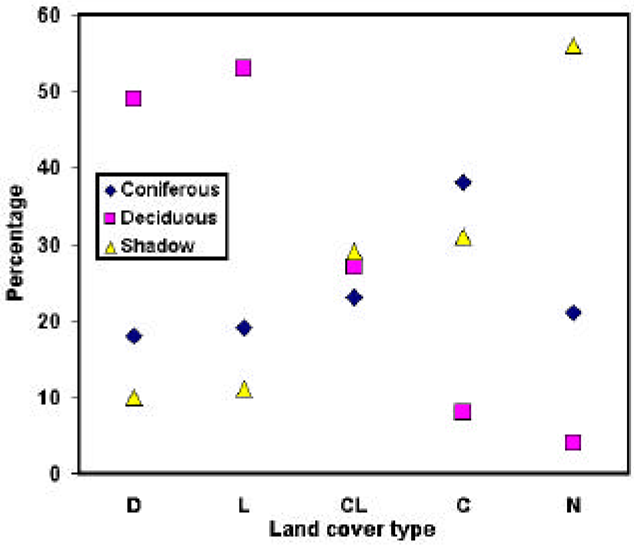

Figure 5. The relationships between selected scene component fractions and land cover type.

The information required in this case study was the forest type within the land parcels. Due to the lack of a simple relationship between the scene components and the land cover type and the presence of a mixed land cover class the scene component fractions were used as the inputs to a supervised classification procedure to identify land cover type. All nine of the scene components fractions from 15 selected training land parcels (3 for each land cover type) were used to parameterize a standard minimum distance classification algorithm. The remaining land parcels were then classified and the results compared to their known land cover type (Table 5).

		Predicted class					
		D	L	CL	C	N	Total
Actual class	D	11		4		3	18
	L		4	3			7
	CL			15	5	1	21
	C			2	18	5	25
	N				3	9	12
							83

Table 5. Correspondence matrix for land cover type classification, showing a direct correspondence of 70 %.

The direct correspondence for the forest land cover types was approximately 70 %. One of the forest land cover types used was a mixed class which are difficult to define, both in ground survey and when training the classification procedure. A number of the mis-classifications were associated with the mixed class and the classes representing the components of the mixed class. If these land parcels were included in the correspondence calculations then in excess of 80 % of the land parcels were classified correctly. There also appears to be a certain amount of confusion between the conifer and new conifer classes, which is again a difficult distinction to make. Although this methodology for image classification makes no radical advances in terms of correspondence accuracy compared to conventional per-pixel classification of coarse spatial resolution data it does hold its own and provides a much richer set of results and an opportunity for further analysis.

7 DISCUSSION

The use of conventional per-pixel approaches for land cover classification with fine spatial resolution data are invalid where the observational scale of the data falls below the environmental scale of the land cover type that is to be mapped. In this case, the concept of the forest does not occur in 15 cm spatial resolution data where a pixel may only represent a facet of a single canopy.

The methodology described here allows the true information content of the data to be extracted, provided in a meaningful form and exploited. The use of unsupervised classification at the pixel level maximises the available spectral information within the image.

The merging and identification of spectral clusters into scene components provides the information in terms of real world objects which can be related to land cover type over areas larger than the pixel. These real world objects may themselves be useful in certain environmental applications as long as they are within a suitable spatial context.

The aggregation of scene components on a per-parcel basis provides a meaningful context for reporting their presence, abundance and distribution. In this way mis-classifications, which are often minor components within a land parcel, can be excluded. Also, the final result may not necessarily be a single hard classification, but a mixture of scene components and their relative dominance. Some of the operations at the parcel level are akin to smoothing or filtering of per-pixel data, but in this case the process can be tailored by each land parcel and the contextual integrity of the data within the land parcel is retained.

The methodology, even in the simple form described here, produces a much richer data source than conventional per-pixel approaches. The processing of scene components can be taken further by assessing the pattern/distribution of scene components within the land parcel. For example, an urban area with 50 % paved and 50 % built could be either industrial urban if the scene components are clustered into large areas (factories/warehouses and parking/roads) or residential urban if they are evenly distributed (houses and roads). The results of the methodology are not just a single layer map, but a data storage framework rich in information and a good starting point for further analysis.

8 CONCLUSIONS

The current developments in airborne and spaceborne remote sensing instruments promise to provide enormous amounts of fine spatial resolution data to the user community. Previous reductions in the spatial resolution available from digital remote sensing instruments have resulted in the re-examination of the data characteristics and processing requirements. These issues must again be considered to make effective use of these new data sources. This paper has described a basic methodology for processing fine spatial resolution data for land cover mapping.

ACKNOWLEDGEMENTS

This work was supported by science funding within the Centre for Ecology and Hydrology, Natural Environment Research Council of the UK.

REFERENCES

Aplin, P., Atkinson, P.M. and Curran, P.J., 1997, Fine spatial resolution satellite sensors for the next decade. International Journal of Remote Sensing, 18, 3873-3881.

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

Barr, S. and Barnsley, M.J., 1999, Improving the quality of very high spatial resolution remotely-sensed land cover maps for the inference of urban land use maps. *Proceeding of the 25th Annual Conference of the Remote Sensing Society: RSS '99 - From Data to Information*, University of Wales, Cardiff, UK, 8 - 11 Sep 1999, 111-118.

Bruniquel-Pinel, V. and Gastellu-Etchegorry, J.P., 1998. Sensitivity of texture of high resolution images of forest and acquisition parameters. Remote Sensing of Environment, 65, 61-85.

Fritz, L.E., 1996. The era of commercial Earth observation satellites. Photogrammetric Engineering and Remote Sensing, 62, 39-45.

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

Gong P., Miller, J.R. and Spanner, M., 1994. Forest canopy closure from classification and spectral unmixing of scene components – multisensor evaluation of an open canopy. IEEE Transactions on Geoscience and Remote Sensing, 32, 1067-1080.

Konecny, G., 1999. The Impact of High Resolution Satellite Data on the Operationalisation of Remote Sensing for Mapping from Space. In: Proc. 2nd International Symposium on Operationalisation of Remote Sensing Conference and Exhibition, CD.

Markham, B.L. and Townshend, J.R.G., 1981. Land cover classification accuracy as a function of sensor spatial resolution. Proc. Of the 15th International Symposium on Remote Sensing of Environment, Ann Arbor, MI, 1075-1090.

Neukum, G., 1999. The Airborne HRSC-A: Performance Results and Aplication Potential. In: Photogrammetric Week, 99, eds. D. Fritsch and D. Spiller. Wichmann, Heidelberg, Germany.

Renouard, L. and Lehmann, F., 1999. High Resolution Digital Surface Models and Orthoimages for Telecom Network Planning. In: Photogrammetric Week, 99, eds. D. Fritsch and D. Spiller. Wichmann, Heidelberg, Germany, pp. 241-247.

Smith, G.M. and Fuller, R.M. 2000. Per-parcel classification of land cover in Jersey. *International Journal of Remote Sensing*, In press.

Toll, D.L., 1984. An evaluation of simulated Thematic Mapper data and Landsat MSS data for discriminating suburban and regional land use and land cover. Photogrammetric Engineering and Remote Sensing, 50, 1713-1724.

Wewel, F. and Scholten, F., 1999. High Resolution Stereo Camera (HRSC): Multispectral Data Acquisition and Photogrammetric Data Processing. In: Proc. 4th International Airborne Remote Sensing Conference and Exhibition, Vol. I, pp 263-271.

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