

REGIONAL INFORMATION MANAGEMENT AND EXPLOITATION

C. J. Willis, C. J. Oddy, A. J. Rye

Image Analysis Group
GEC -Marconi Research Centre,
Great Baddow,
Chelmsford
Essex
CM2 8HN

Commission II, Working Group 5

KEY WORDS: database, spatial, map, multi-temporal, image, management, fusion

ABSTRACT

The management of raster and vector spatial data is an important tool for supporting complex image analysis applications. This paper describes a system for handling remote sensing data within a commercial database, Illustra, integrated into a visualisation and system development environment, AVS/Express. This system, spaceBase, provides an environment for developing image analysis algorithms that operate on mixed data types. Some example applications are reviewed and consider the fusion of image and map data. Techniques for identifying features common to a map and image, and features that are present in only one source, are discussed.

1. INTRODUCTION

The management of data generated by remote sensing systems continues to present new challenges to the Geographical Information Systems (GIS) community. Many datasets consist of a number of overflights of the same area on a number of occasions, possibly by a number of different sensors. These sensing systems include both satellite and airborne vehicles with optical and radar sensors, and ground based surveys and ground truth collection campaigns. GEC-Marconi Research Centre (GMRC) have been active in the handling and management of image derived spatial data from remote sensing systems for a number of years. This paper discusses the implementation and application of a system designed to manage both raster and vector spatial data from remote sensing systems. The system is capable of supporting complex queries which may span several images and data types.

The following section discusses the background in spatial database research at GMRC and sets out some of the datatypes which need to be stored in such a system. Later sections go on to provide an overview of the implementation of the system and to describe a number of example applications in the image to map, and image to image data fusion arena.

2. BACKGROUND

GMRC have been active in this development of databases for vector spatial data for many years. Previous developments in this area include the Picture Understanding Database and V/Base systems produced during the 1980's. The V/Base system was the more highly developed of the two and allowed storage and query of point, line and polygon data derived from

single raster images, and of attributes and relationships derived for these features.

The latest development in the Regional Information Management field at GMRC is spaceBase. spaceBase is a new implementation based on concepts developed during the V/Base era. The system is capable of storing all V/Base datatypes along with extra types relevant to the remote sensing discipline. The spaceBase storage types are:

- projections, both map (UTM and UPS) and satellite (orbital) projections are currently handled, but the system may be easily extended to cover coordinate systems from other disciplines (e.g. medical imaging, etc.);
- sensors, band frequencies, polarisations and processing details are handled, but again the system may be easily extended to cover new sensor types;
- coverages, storing the coordinates of the bounds of the dataset (in latitude and longitude if relevant), its acquisition and processing dates, etc.;
- raster image data, for storing pixels imaged in a coverage;
- features, i.e., points, lines and polygons (either by vertices or via a bounding chain code);
- feature relationships such as adjacency and containment;
- attributes, for storage of information derived for features, for example, area, centroid location, mean value for each image band, texture and shape measures, etc.

3. IMPLEMENTATION

A Commercial Off The Shelf (COTS) database, Illustra, has been chosen as the core system for the storage of these types. The Illustra database has features of both conventional relational databases and of the object oriented methodology (notably type and table inheritance). Illustra may also be enhanced by the addition of so-called datablades which add functionality to the basic system. In particular a datablade for the management of 2-dimensional spatial data is used within spaceBase to provide the basic spatial data types and spatial indexing. The standard system has been augmented with a variety of additional basic storage types. In particular, a spatial types capable of storing latitude and longitude coordinates, and quadrangles with latitude and longitude corner coordinates have been added thus moving the two dimensional Cartesian capabilities of the Illustra system into the genuine real world coordinate system.

A C language interface has been developed to provide access to the database tables. This consists of a set of C structures mirroring the tables and an application programmers interface (API) enabling data insertion and selection. The database tables, C structures and API together form the complete spaceBase system and have the potential to be used as stand-alone toolbox for the production of application systems which store, manage and access remote sensing data.

Tools are provided to "register" remote sensing imagery, possibly in CEOS or other similar form, with the database. This registration process lets the database know of the existence of the image, where to find it on the disk system, and records its locational and sensor based parameters in the database. Typical locational parameters might be latitude and longitude of the image corner coordinates and the current projection of the image. Typical sensor based parameters are the acquisition and processing date and time, sensor type and frequencies of the image bands.

When a number of images are registered with the database it may then be queried to find, for example, all images collected within a certain date range, and containing a given point on the ground, or perhaps, for all images from infra-red sensors which overlap a given area of interest.

In order to make spaceBase available to users and to demonstrate its capabilities a second COTS tool, AVS/Express, has been used to provide a visualisation and application rapid prototyping and development environment. A selection of AVS/Express modules, macros and applications have been developed. These tools allow both insertion of data into the database and selection and querying of results.

4. USE OF THE DATABASE

One of the most frequently used operations when preparing an image for analysis is the segmentation of the incoming raster data into homogenous regions. A variety of methods are available for this operation, with some being generally applicable, whereas others are suitable for certain sensor types only. We don't consider the segmentation process here, only the representation and storage of the segmented data. This data is generally most efficiently represented through the boundaries of the derived regions as so-called vector data. The spaceBase system allows the storage of the boundary vectors of these regions within the database as polygons, either by their vertices or, more efficiently for data derived from raster imagery, as chain codes. Once derived the region boundaries may be used

to calculate shape attributes for the regions such as area, moments and fractal dimensions and so on. The boundaries may also be used with the raster data to calculate intensity related attributes of the regions such as mean, standard deviation, texture measures, weighted moments, and many others. The relationships between regions may also be stored including the neighbours of a region and any contained or containing regions.

The vector image data and the associated attributes may be used to enhance the visualisation and analysis of the imagery. For example, one simple operation which may assist in the interpretation of the image is to display the raster image data with the region vector overlay. In a similar way map data may be stored in the database directly as vectors. This too may be displayed as an overlay over the raster image data. Results demonstrating the basic data storage and access are presented in later sections.

The data, particularly attributes and relationships, stored in the database may be used to enhance the analysis of the imagery in many ways. At the most basic level the attributes derived from a single image may be used to generate classifications of the objects contained therein. This process is similar to that carried out by many image analysts using existing tools. The quality of classifications may be enhanced if a number of images, either from a range of sensors or spread over time, are available. In this mode the intensity signatures, for example, of agricultural crops over the whole of a growing season, may be used to generate classifications with a very high confidence.

The identification and monitoring of change through image sequences is also possible when using the database as a source of data. Several examples of this type of analysis are possible including: the appearance or disappearance of objects, possibly temporary buildings or vehicles; the movement of region boundaries, possibly associated with deforestation, and; the appearance or disappearance of boundaries, possibly associated with the aggregation of agricultural land. This work is allied to the classification of regions in the database in that either vector map data or segmentations of the imagery are required to derive good features, and change, either of the imaged areas from the maps, or between segmentations of different images of the same area, must be resolved. This is an area of active research at the moment with techniques being developed which test the hypothesis that the segmentation currently being examined differs from a base vector data set in a significant way, and go some way toward modifying the vector data sets to account for this change. Results showing identification and resolution of change through image sequences are presented in later sections.

5. APPLICATIONS

The following sections describe work performed using the spaceBase database in support of image analysis operations. The study has concentrated on the investigation of techniques for combining information from a map and an image of the same area. Three possible methods for fusing the information have been devised and tested using the system. These include: incorporating new vector information from an image into an existing map; using map information to assist in detecting edges in the image; editing the map based on the intensity values from the image. The techniques are complementary and could be used separately, or in combination in the spatial part of a data fusion system.

All software written for this study has been developed within the AVS/Express environment [1] which enables rapid

prototyping of algorithms in a modular way with the ability to transfer data between modules easily and to visualise the data.

6. ADDING IMAGE VECTOR INFORMATION TO THE MAP

A supplied map will contain features of a number of known types, e.g. roads, rivers, field boundaries. These features will often only give partial information on the true ground cover as some feature types will not be included e.g. sub-divisions of fields containing more than one crop; tracks; temporary features such as vehicles, floods. In addition the map will not contain information on recent changes such as the erection or demolition of buildings; removal of hedges; new roads etc.. A recent image of an area can be used to augment the map information. This method matches the map vectors with the vectors from a segmented image. If a vector is present in the image but is not found in the map then it may be added to the map representation.

The inputs for this process include: the map, in the form of a spaceBase coverage containing polygon and line features; the segmented image data, also as a spaceBase coverage, and; algorithm parameters, which are the minimum area of the image regions and distance and angle thresholds for line matching.

On completion the output takes the form of a spaceBase coverage containing an augmented map with all of the original map features plus additional lines and polygons derived from the segmented images.

To create a new augmented map which contains all the existing map features the following steps must be carried out:

- for each image region which has area greater than the specified minimum:
 - compute a polygonal approximation to the region boundary
 - find map features which overlap with the region
 - compare each polygon side with all the overlapping map lines and polygons using the specified distance and angle thresholds
 - insert any parts of the polygon which are not found into the augmented map (these may be lines, consisting of one or more straight line segments, if part of the region boundary was matched, or polygons if none of the original region boundary was matched)

The method described has been applied to two datasets. The first consists of a simple map containing only a small number of features which had been generated manually, and an image from the Spot satellite. Figure 1 shows the original map features (in white) overlaid on the image.

The second dataset consists of a more complete map, which was derived from a segmentation of the Spot image, and a Landsat image of the same area. Figure 2 shows the Landsat image (left) and the map overlaid onto the Spot image from which it was derived (right).

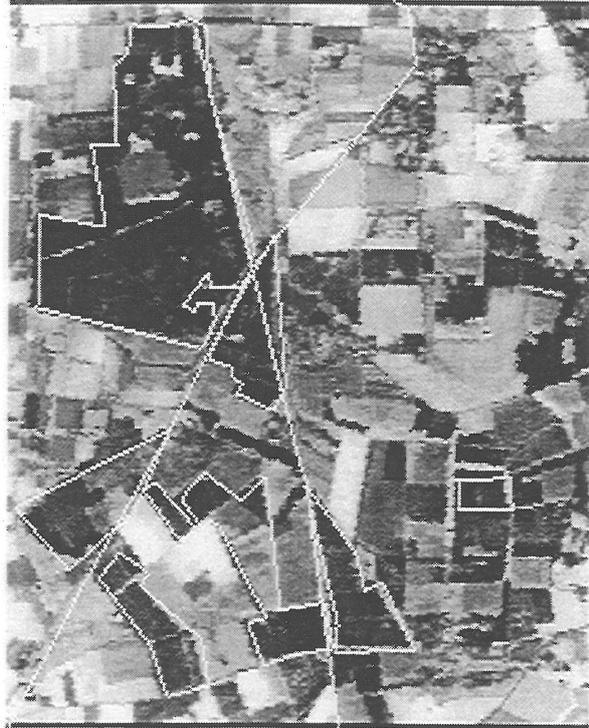


Figure 1. Simple map overlaid on image

Figure 3 shows the results of adding the Spot image to the simple map, firstly with the thresholds set to consider image regions larger than 50 pixels, and matching lines which are within 2 pixels in distance and 0.2 radians in angular difference, and secondly with the line matching thresholds increased to 5 pixels and 0.4 radians and only image regions with area greater than 100 pixels considered. The lower threshold set produces a more detailed map and inserts some image region boundaries which clearly correspond to map edges but are in a different position due either to change or registration errors (for example the triangular forest region to the right of the bottom of the image is inserted from both the map and the image). The higher threshold set removes such edges but gives less detail. Selection of appropriate value for the thresholds will depend on the application and on what further analysis of the data will be performed.

Figure 4 shows the results of adding the Landsat image to the more complete map with the thresholds set to consider regions larger than 50 pixels, and matching lines which are within 2 pixels in distance and 0.2 radians in angular difference and with the line matching thresholds are again increased to 5 pixels and 0.4 radians and the minimum area to 100 pixels. This example illustrates the need for some post-matching processing of the data in order to remove short unconnected lines which are inserted where there are minor differences in the region boundary. Such map editing might involve reference back to the image to consider the evidence for change.

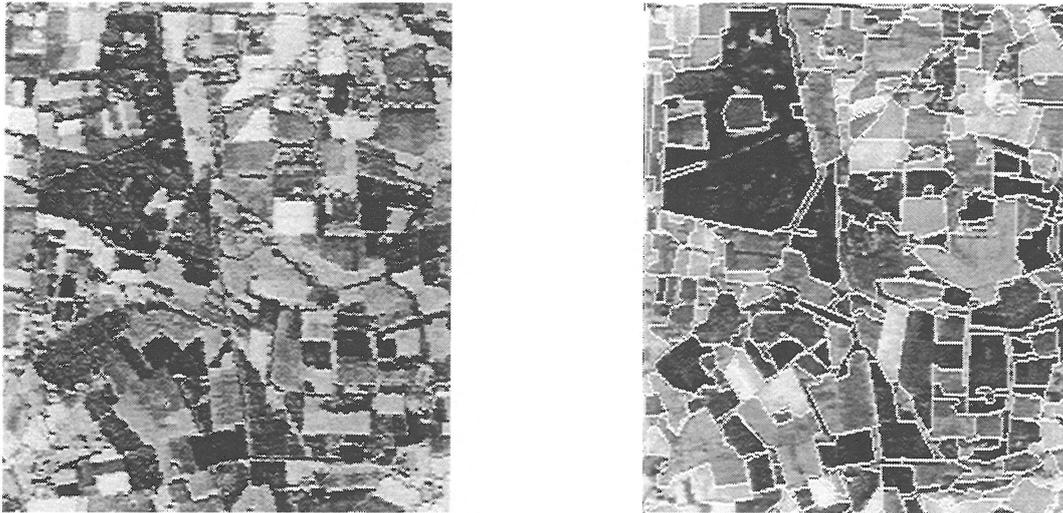


Figure 2. Landsat Image and Spot Image with Full Map

This technique works especially well when the map contains only a few features and provides a means to add other features without duplication. Although there are three user-defined thresholds (minimum region area and distance and angular thresholds for line matching), these are easily understood and their values can be chosen according to the application and using known or estimated values for expected position errors.

There is a tendency for a number of short, unconnected lines to be inserted which are the result of small unmatched parts of a region boundary and may be due to segmentation errors and noise in the image. Some post-insertion editing of the augmented map would be beneficial to remove these lines or to reconcile them with existing features.

7. MAP-BASED EDGE ENHANCEMENT

In many applications it would be beneficial to be able to restrict the processing of an image to particular areas of interest.

Perhaps using map information to guide the extraction of image features. In order to discover the potential of this process a preliminary investigation has been carried out into the possibility of using the line and polygon features in a map to assist in extracting edges in the image. Two alternative possibilities were investigated. In the first case the map information is used to reinforce the detection of edges in the vicinity of map edges, which provides a means of checking for minor changes in the position of an edge while suppressing edges which do not coincide with map edges.

The second case addresses the converse problem of detecting new edges in the image while suppressing detection of edges which are already present in the map.

Once again the process is initiated by a spaceBase coverage containing map information. The output of the process is an enhanced edge image showing amplified edge details either close to, or away from, the map features.

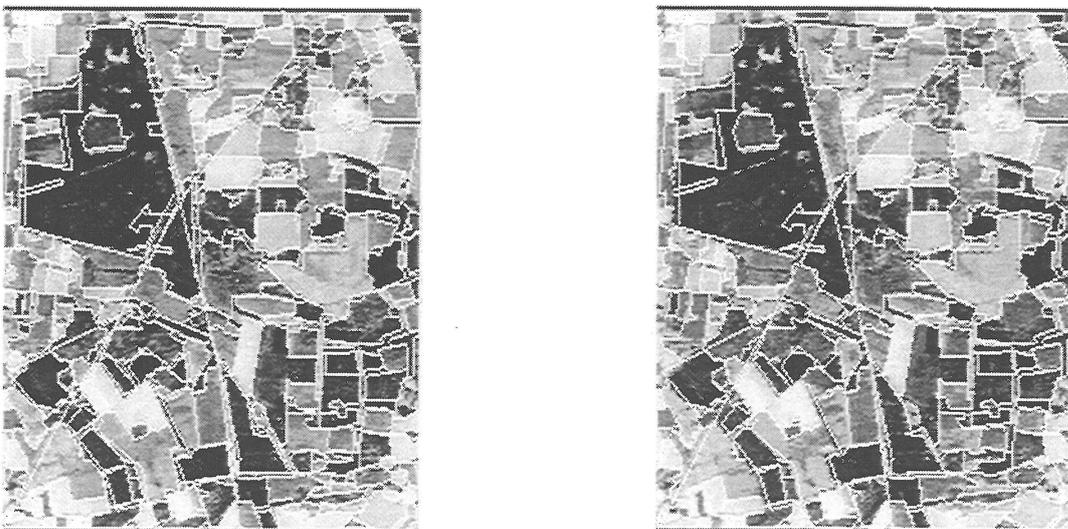


Figure 3. Spot Image added to Simple Map with low threshold values (left) and high threshold values (right)



Figure 4. Landsat Image added to Full Map with low threshold values (left), high threshold values (right)

The algorithm used to complete this process has the following steps:

1. The first stage in the edge-enhancement process is to convert the spaceBase vector map data into a raster representation. A raster image is created in which each pixel is set to the minimum distance (in pixels) to the nearest line or polygon boundary plus a constant value. The pixel size is set to be the same as that of the input image.
2. Secondly the image is pre-processed using appropriate smoothing and edge-detection algorithms to give an edge image in which the pixel intensity depends on the strength of any edges in the vicinity.
3. The edge image and distance image are then cropped to select the area of overlap between the two coverages.
4. Finally the images are combined arithmetically in one of the following ways:
 - Multiply the images to enhance edges which are not present in the map.
 - Compute the ratio of edge image to distance image to enhance edges which are close to the map edges.

This edge enhancement has been applied to the same two datasets described above (fig. 3 and fig. 4).

The original edge image for the Spot data is shown at the top of figure 5. Bottom left shows the results of multiplying the edge and distance images for dataset 1 and bottom right shows the results of computing the ratio of edge to distance image.

The original Landsat edge image is shown at the top of figure 6. Bottom left shows the results of multiplying the edge and distance images for dataset 2 and bottom right shows the results of computing the ratio of edge to distance image.

Further processing is required to extract the enhanced edge information and insert it into an updated version of the map. The edge enhancement could probably be improved by altering the distance image to a binary image in which pixels are classified as either close to, or distant from, an existing map edge or, alternatively by thresholding the distance image so that

pixels which are further than a specified number of pixels from a map edge are set to a capped value.

In order to update map information using the image, we decide whether map polygons are homogenous using intensity attributes, computed for map features using the image pixels. If features are not homogeneous, then these features are split by applying an edge detection algorithm to the image within the polygon.

For input the process requires a spaceBase coverage containing vector map data, and a corresponding raster image against which the map is to be compared. As parameters the process needs a defined minimum standard deviation for regions to be split, and a minimum area for any region which is created by the process.

The output of the process is an updated vector map in which the features have attributes (mean, standard deviation, etc.) computed from the raster imagery.

The steps in the process are as follows:

1. The map polygon information is used in combination with the image intensity data to compute numerical attributes, mean and standard deviation, for each polygon.
2. The image is pre-processed using appropriate smoothing, edge detection and bonding algorithms to produce a bonded image in which each pixel is linked to those of its neighbours which are considered to belong to the same region.
3. For each polygon for which the standard deviation of the image pixels exceeds a user-supplied threshold the following steps are taken:
 - Extract the pixels from the bonded image which are contained within the polygon to be split.
 - Extract region boundaries within the polygon and compute the area, mean and standard deviation for each region.
 - Merge any regions with area less than a specified threshold with the adjacent region which is closest in intensity.

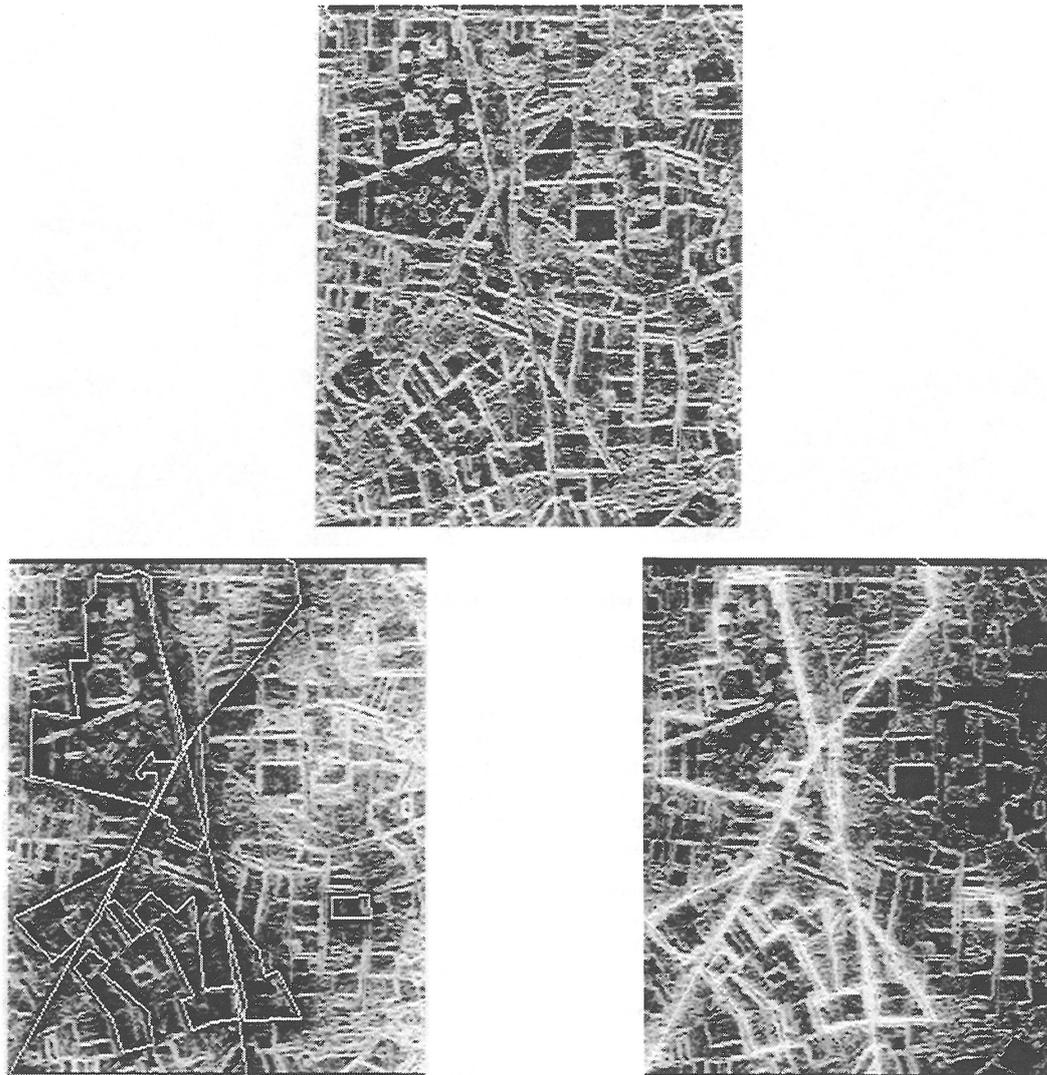


Figure 5: Spot Edge Image (top). Enhanced Edges: Edges not present in Map (left), Map Edges (right)

- Compute a polygonal approximation for the region boundaries.
- Insert the new polygons and their attributes into the updated map to replace the original polygon.

Dataset2 has been used to assess the results of updating the map, which had been derived from Spot data, using Landsat image information. dependant. The minimum area threshold, in particular, will depend on the intended use of the updated map. If changes such as construction/demolition of buildings are of interest then a very low area threshold will be required but if the application is concerned with land cover a high threshold will be appropriate.

The current work on updating the map information relies solely on the standard deviation of the map polygon. In some cases this may not give a good indication of whether a region should be split - for example, it is not possible to discriminate between

a region with a high standard deviation due to texture or noise within the region, and one in which the region is divided into distinct areas which have low standard deviation locally. In the former case it would not be desirable to split the region and the result of attempting to do so might be to extract a large number of small regions, which would then be merged until the area exceeded the threshold. This would result in a split of the map polygon which did not reflect any genuine structure from the image. This situation could be improved by applying a more intelligent merging rule based on the difference in mean intensity and/or the strength of the edge separating the regions. It might also be possible to use rules based on the polygon classification from the map to determine the expected mean and standard deviation of a feature.

The technique could be extended to allow adjacent map features to be merged according to rules based on the mean intensity and the strength of the edge separating the regions.

Again this extension would benefit from the use of rules based on the map classification and knowledge of the resolution and imaging capability of the image sensor. For example, if adjacent map features were classified as forest and field but no difference was visible in the image then the conclusion might be that the forest had either been felled or extended to include the field area; however if both regions are classified as field then it would be difficult to tell whether the fields had been merged (e.g. by removal of a hedgerow) or whether the boundary between them was not detected at this resolution.

8. CONCLUSIONS AND FURTHER WORK

This paper describes three complementary techniques for combining information from a map and an image, all of which have been applied with considerable success to the available datasets. The use of image data to extend and update an existing map has been demonstrated along with the possibility of highlighting areas of change. Possible improvements and

extensions to these techniques have been described in the appropriate sections.

To date the datasets used have consisted of specially generated maps (manually or by image segmentation) and optical data. It would be useful to assess the techniques further using data from other sources, such as digital maps and SAR images.

The study has concentrated on the specific algorithms required to find corresponding features in the image and map and to find features for which there is no correspondence. At this stage these features have then been combined without reference to their classification or reliability. In a practical application it is likely that a more complex decision making process will be required to determine the correct course of action where there is no feature correspondence or where matched features display differences in their spatial or classification information. Decisions might be based on a large number of factors, many of which will be application dependant. Such factors would include: the likelihood of spatial change for features with a particular classification (for example, roads are unlikely to

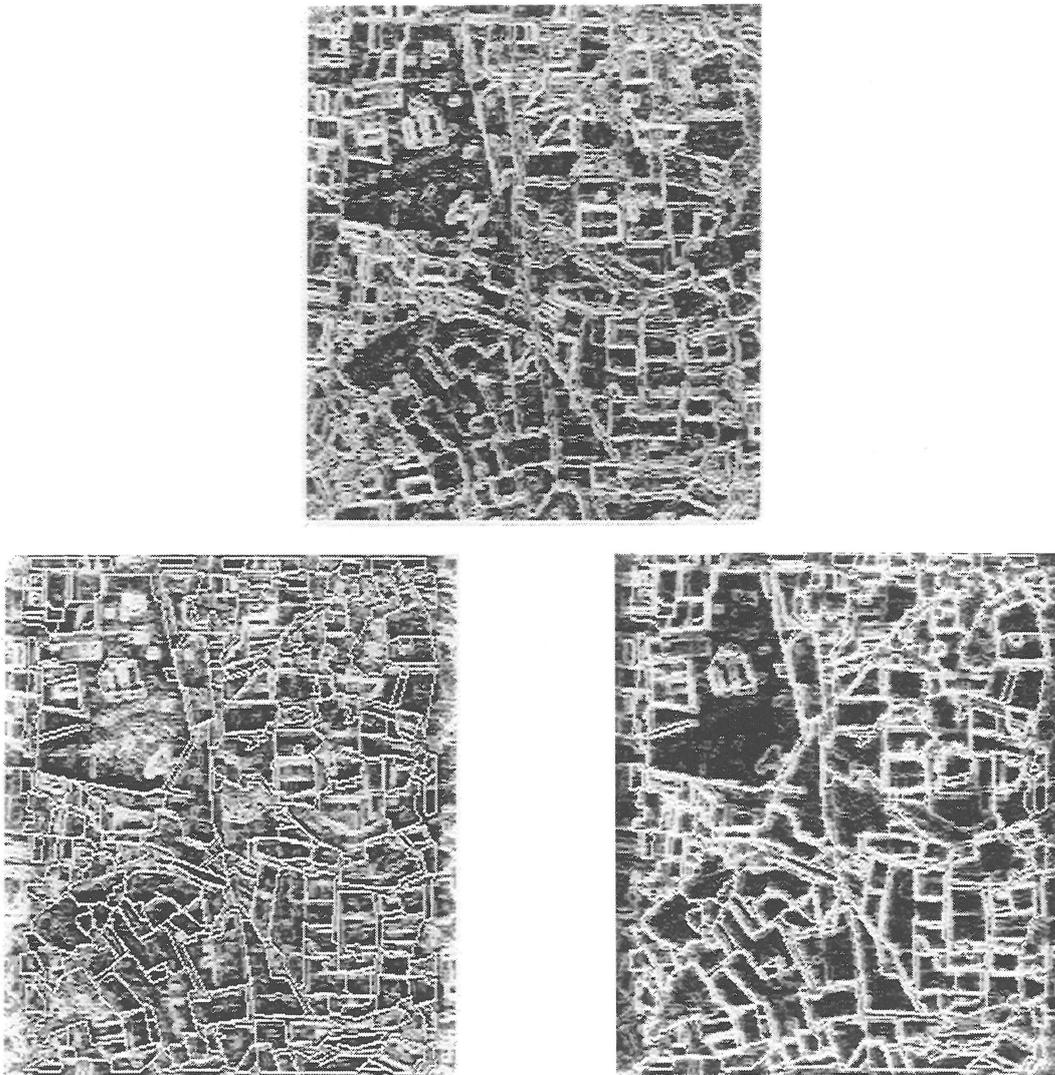


Figure 6: Landsat Edge Image (top),
Enhanced Edges: Edges not present in map (left), Map edges (right)

disappear whereas boundaries of fields or woodland are more likely to change); the probable permanency of spatial change (changes in forest boundaries are likely to be more permanent than extents of a flooded river); the likelihood of classification change (a change from field to forest in a short time scale is unlikely whereas the reverse change might occur); the initial probability of misclassification (if an automatically classified image is input then the correspondence of image to map classes and a classification probability might be used).

The techniques described could be valuable in a number of applications including the automatic updating of maps (especially in the future when high resolution satellite images will allow more detailed classification than could economically

be achieved manually) and the detection of change for military and civilian surveillance applications.

The example applications have demonstrated the capability of the integrated raster and vector spatial data handling capabilities of the spaceBase system.

The copyright of all imagery is acknowledged.

9. REFERENCES

[1] AVS web sites: <http://www.com.avs> and <http://www.iavsc.org>

[2] Management and Visualization of Spatial Data with spaceBase", C.J.Willis, MTR97/18C