The Image Processing Link to Geographic Information Systems

Barbara A Harrison, Kevin K Mayo, David L B Jupp
& Paul G Hutton
CSIRO Division of Water Resources

and

Waqar Ahmad
University of Tasmania

Abstract

Geographic Information Systems (GIS) provide a means to register, modify and compare geographical data from a wide range of sources, including remote sensing. This paper considers the application of image processing in GIS in terms of data format conversions and processing functions.

A GIS was developed for forest resources in Tasmania using remotely sensed data and mapped based data as input. This system was implemented in a microcomputer environment using the microBRIAN image processing software, MAPEDIT/ATLAS and LOTUS 1,2,3 packages. In this system, Landsat MSS, digital terrain and administrative boundary data were integrated to provide a framework for analyses. These data, with differing formats, projections and spatial resolution, were geo-referenced to a common, spatially congruent base of latitude and longitude co-ordinates and resampled to a two second grid. This raster based data matrix was then analysed to obtain results for use in forest inventory and management. The results illustrate the development of a GIS in the PC environment using commercially available processing systems with an emphasis on the role of remote sensing and image processing.

1. Introduction

Much reference is made in resource management literature to the concept of a Geographic Information System (GIS). The objective of such systems is to register, modify and compare geographical data from a wide range of sources. The systems must therefore have the capability of utilising data stored in a variety of formats.

A number of commercially available systems purport to achieve this; however the enormous range of data sources with their inherent accuracy problems, and the complexity of the operations involved dictate that such systems are necessarily expensive and/or impose significant limitations in terms of processing algorithms and/or the volume of data that can be handled.

This paper presents a simplified view of a GIS as a composite of data processing systems which are readily available in the microcomputer environment.

1.1 The Geographic Information System Framework

Geographic data formats fall into three basic types:

- raster - such as gridded data or remotely sensed imagery
- vector - such as linear map features
- tabular - such as attribute tables relating to map features.

Although these data types are acquired and processed differently, PC-based processing systems operating in one or two of these modes are commercially available.

In the GIS context, vector data, such as cadastral boundaries or contour lines, are typically collated from maps. This data format is produced by digitising linear features; the scale and dimensions of the feature are represented as a sequence of co-ordinates relative to a common origin. Typically, for map data, these co-ordinates are latitude/longitude or Northing/Easting within a standard UTM zone. This allows the digitised features to be directly related to a ground location for comparison with other data. Point features, such as field sites, may also be recorded as vectors. A range of processing functions may then be applied to the vector data such as updating or merging features to form new entities. A digital elevation model (DEM), for example may be derived by grid sampling contour strings.

Raster data represent spatial information as a gridded data set or image, with each grid cell, or pixel, being assigned a numeric value to indicate its status for a particular attribute. In the case of remotely sensed imagery, these values represent the reflectance/emission of radiation from specified portions of the electromagnetic spectrum as recorded by transducers for each pixel. Raster data planes may, however contain data from non-remote sensing sources such as thematic maps or a DEM. Data in a raster form are readily compared, reduced and registered using available image processing routines.

Attribute data are linked to locational data but do not contain independent geographical referencing in the same way as the vector or raster formats. This data type may summarise attributes relating to a particular geographical location (be it polygonal or grid based) and can be manipulated independently of the spatial data. For example, while the boundary of a management zone may be stored in the vector database, crop production figures, soil types and access status information for that zone would be recorded in a tabular database. Such zonal attributes may be derived from remotely sensed or other data. A range of processing functions, including table or matrix manipulation, statistical analyses, graphical presentation and key-based queries may be used to access, analyse and summarise the data. The set of attributes in this format comprise a tabular or attribute database.

Different formats are preferred for different applications. As the preferences relate to the mechanics of data storage, retrieval and analysis, an information system is fully integrated only if all data from a region can be moved between and hence processed within the different formats.

The spatial patterns inherent in geographical data differentiate the GIS from other database systems. The key information obtained from a GIS require analysis of the spatial characteristics of the data. Topological operators may be applied to vector data to determine the degree of inclusion or adjacency of different regions but these functions are limited in their utilisation of the spatial structure of the data. Spatial analyses can be performed most efficiently using raster data and image processing techniques.

Conversion from vector to raster formats requires certain basic functions which locate the digitised regions or points within the image grid. An additional image data plane, or channel, may be created whereby each pixel is assigned a value indicating its zonal location, or an existing channel may be segmented using the digitised or zonal boundaries as a mask. Similarly, point data, possibly representing field measurements, may be embedded in an existing or new channel. Image processing routines may be used to interpolate missing pixel values where the data do not provide continuous coverage over the image.

Functions converting between different formats are currently available in a number of PC-based data processing systems. Variables represented by image channels may be cross compared once the remotely sensed and ancillary data are converted to a common raster grid base. This cross-comparison may be used to stratify the remotely sensed data, to resolve ambiguities that occur with interpretations of remotely sensed imagery or to derive image-based results stratified by zone. Alternatively, an interpretation of remotely sensed data interpretation for a sample set of pixels in an image may be compared with a interpretation based on another data source to determine the "accuracy" of the image-based results. The resulting cross-tabulation matrix may then be transferred to a tabular database for further manipulation.

2. The microBRIAN System

MicroBRIAN is a microcomputer-based image processing system developed by CSIRO (Australia) and MPA Pty Ltd (Jupp et al., 1987; MPA, 1988). The system contains algorithms for a wide range of remote sensing and other applications. While these algorithms are raster based, the system interfaces to the GIS environment defined above by importing various vector data formats, importing and exporting a range of raster data formats and exporting tabular data.

Vector polygon data may be integrated with remotely sensed data as a stratification mask or as an additional image channel. Polygon, line or point data in vector format may also be embedded in existing image channels as a contrasting image value. The system can accept vector data in the USGS DLG-3 format (used by Arc/Info), AUTOCAD .DXF format (used by a number of mapping authorities in Australia), AS 2482 (the Australian digital mapping standard) and MAPEDIT .BNA format (used in the example application below).

Raster data may be converted to the microBRIAN image format from a wide range of standard remote sensing data tape formats, ASCII data files, Arc/Info .SVF (Single Variable Format) files and ERDAS GIS format. MicroBRIAN image data may also be exported in ASCII, Arc/Info SVF or ERDAS GIS formats. Conversion to the Arc/Info SVF file format also facilitates raster to vector conversion with Arc/Info.

Tabular data are produced by cross-tabulation of image channels or samples of image values and user-defined labels (Jupp et al., 1985). Such data may be recorded in Lotus 1,2,3 compatible or any user-defined ASCII format. These tabular results may be manipulated in microBRIAN or exported to a tabular database package for report preparation or record updating.

The composite system shown in Figure 1 provides an inexpensive, but effective, environment for processing, integrating and comparing geographical data (Jupp et al., 1988). The complete system uses hardware and software packages which are economically accessible to land managers and which may be used for a variety of purposes.

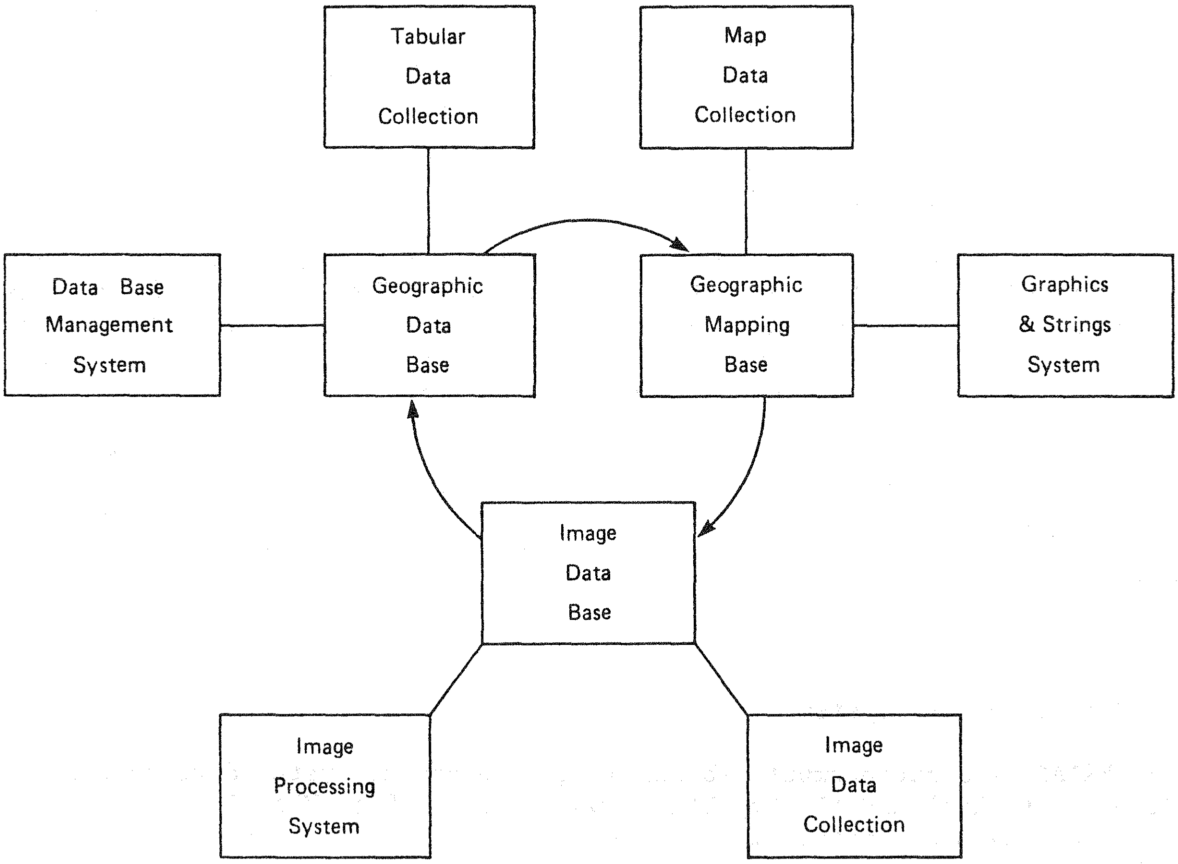

Fig. 1. Resulting composite framework.

3. Example Application

3.1 Introduction

The above GIS framework (Fig. 1) was used to develop a forestry information system for the Scottsdale Forestry District in northeast Tasmania, Australia (Ahmad, 1987). The results reported here form part of a larger study concerned with developing a methodology for extracting landform and land cover information in complex terrain from Landsat MSS data and using such data to monitor changes in land cover. This involved the integration of multiple Landsat MSS scenes, elevation data and management zone boundaries into a common raster file. These data were processed and compared to produced tabular data files for annotation and manipulation with other attribute data. A summary of this process is given in Fig. 2. The details of collection, processing and integration of each data format type are given below.

3.2 Methods and results

The study site contains a wide range of landforms, including mountains, hills and plains and is part of the Wesley Vale forestry reserve. Vegetation ranges from forest to grassland. Some sections have been cleared for agriculture and others planted to softwoods plantations. Forestry is the dominant land use with the native hardwoods being mainly being used for the production of woodchips.

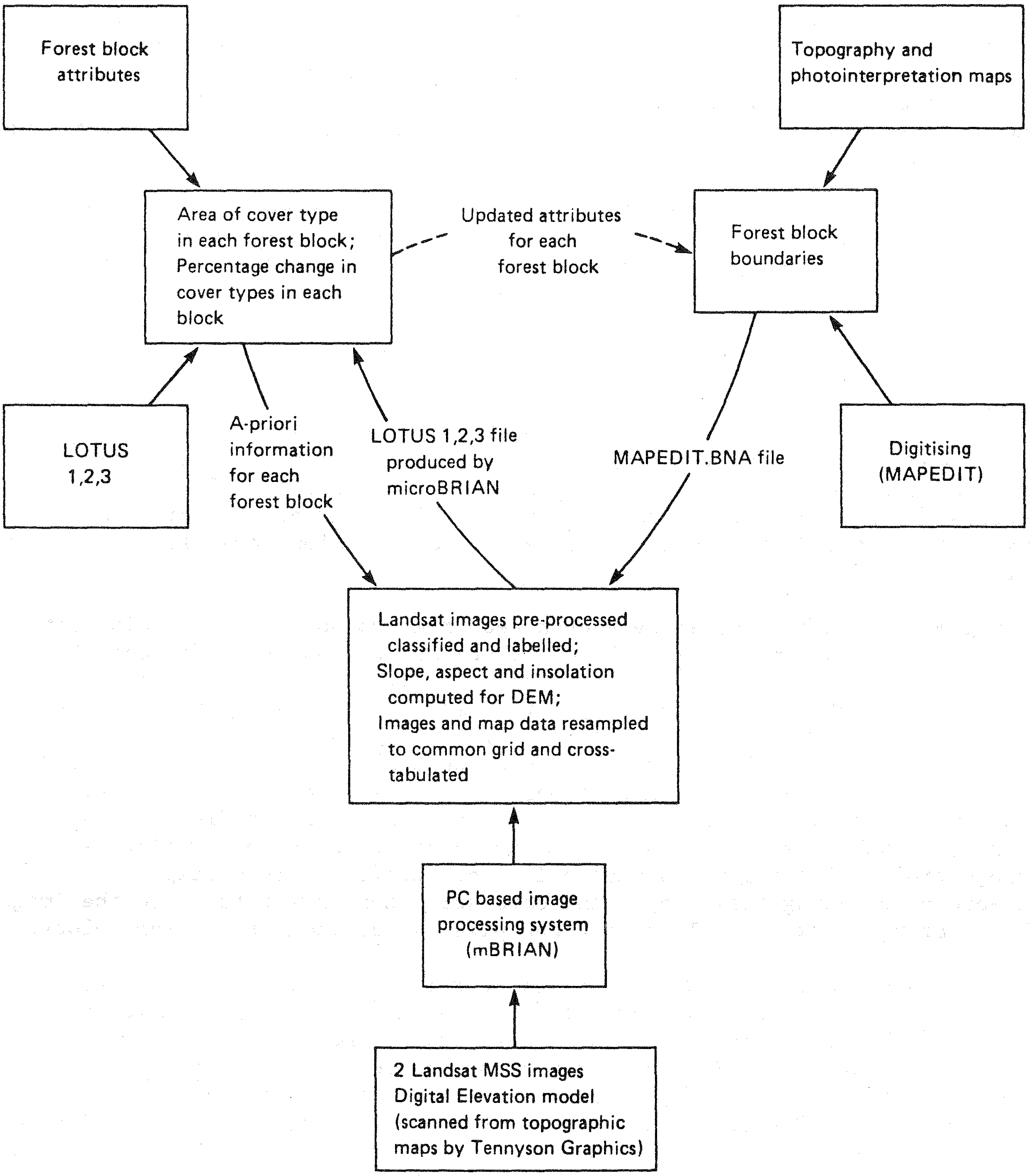

Fig. 2. GIS framework for example application.

Complex topography is an obstacle when developing forest inventory from remotely sensed data. Vegetation can be in full shadow or be completely sunlit, depending on relief. In this study, effects due to landform were separated from those due to land cover by registering elevation data with the image. Subsequently, techniques were developed to reduce the effects due to topography using image data alone. Land cover information extracted from these data was merged with management zone boundaries to resolve ambiguities. This provided statistics on the proportion of land cover types within each zone.

Cloud free satellite imagery is rarely obtained in Tasmania. Only two suitable Landsat MSS scenes were available for this study, the aquisition dates being 9 November 1980 and 4 May 1984. The microBRIAN system (MPA, 1988) was used for image processing tasks.

Elevation data were derived from 1:100 000 topographic map sheets for the study area (Mapping Division, Lands Department, Tasmania). Five map sheets were commercially scan digitised by Tennyson Graphics Pty. Ltd, Melbourne. A Digital Elevation Model was produced by gridding the resulting data to a two second grid (approximately 50 m pixel) with each cell containing a elevation value. This pixel size provided sufficient detail without excessive data volume. Elevation data in raster format were processed using the microBRIAN system to derive insolation, slope and aspect for each pixel.

Management zone boundaries were obtained from two sources. The district boundary was obtained from 1:100 000 topographic maps and individual forest block boundaries from 1:25 000 orthophoto maps. These boundaries were line digitised together with control points which could be accurately located on both map and image. The MAPEDIT software package was used for digitising (SLP, 1986). A model was established using these control points to convert between map digitiser and image co-ordinates. In the microBRIAN system, this modelling process incorporates corrections for known distortions in the imagery, such as sensor or platform characteristics and underlying map projections. The digitised boundaries were converted into image co-ordinates using this model, and data were then integrated into the image by creating a new channel wherein the pixel value indicated forest block.

The Landsat images were pre-processed to remove radiometric striping, oceans and cloud. Log ratio transformations were computed for each image to "remove" the topographic effects. These transformed images were then classified using the BRIAN classification methodology (Jupp et al. 1985a). The classifications were spectrally transferred to a log image and the image re-classified. The difference between the log and log ratio classifications then effectively represented landform or topography, the land cover effects having been "removed" by the log ratio classification. The two effects were combined by splitting each land cover into sunlit and shaded components using the first principal component of the topography image. This procedure resulted in 168 and 188 classes in the 1980 and 1984 images respectively which were then labelled as 28 detailed land cover groups in the 1980 image and 29 in the 1984 image. The 29 land covers were reduced to 12 broad groups for verification of the labelling (Ahmad, 1987).

A number of change detection methods were applied to the imagery. After radiometric calibration of the two images, spectral changes were determined using Principal Components Analysis (PCA) and image differencing techniques. These processes produce images in which changes in land cover were indicated by bright and dark tones.

The images were also compared on a post-classification basis by cross-tabulating the labels for each pixel. The 12 broad land cover types were selected for this purpose and comparisons were made for each forest block separately. The resulting cross-tabulation matrices indicate the area of land in each cover type within forest blocks for each image and the percentage change between sampling dates.

The cross-tabulation matrices between derived and actual land covers were manipulated in LOTUS 1,2,3 for final report presentation. Cross-tabulation included an analysis of the accuracy of each classification and a comparison between them.

The results indicate that using remotely sensed data in conjunction with ancilliary information, it is possible to map broad land cover classes accurately in areas of complex topography, to monitor land cover changes and to incorporate the resulting information into a GIS. While not providing the detailed level available from aerial photographs, it does provide a broad level of resource information with a high degree of spatial accuracy.

4. Conclusions

Image processing constitutes an essential component in the development of an integrated GIS. The raster format is ideal for analysing multi-attribute data. Sources of data may include any gridded data or vector data which can be converted to raster format. Results may be presented in tabular or raster format, and the latter may be subsequently converted into the vector form for processing within another GIS sub-system.

A logical link exists between image processing and GIS because image processing is a fundamental tool for remote sensing. A functional link with the GIS environment facilitates image interpretations as it allows ancillary data to be used in the interpretation process. The application described here illustrates that the necessary links between image processing and GIS already exist in software packages available for microcomputers.

Hitherto, access to such integrated data processing required a significant investment in hardware and software. This has changed with the availability of low cost microcomputer-based systems and the development of a large range of application software. Microcomputer systems have the advantage of being multi-purpose and low cost; the microcomputer provides a very affordable and effective pathway in the linking of image processing with GIS.

Acknowledgements

The microBRIAN system is the result of work by many people since 1980. Firstly, Stuart Kendall was involved in the development of the BRIAN system which was eventually converted to microBRIAN. Geoff Pickup, Vanessa Chewings, Ron Kreymborg, Wayne Petschack, Dennis MacDonell and Peter Booth all provided essential skill and effort in the microBRIAN development. MicroBRIAN is marketed exclusively by MPA (Australia) Pty Ltd.

The authors also wish to acknowledge the financial support given by the Tasmanian Forestry Commission, APPM, ANM and Forest Resources companies for the project reported as the example application in this paper.

References

Ahmad, W. (1987). A Forest Inventory Using Landsat MSS data in NE Tasmania, Australia. Ph.D. Thesis. University of Tasmania, Hobart.

Jupp, D.L.B., Mayo, K.K., Heggen, S.J., Kendall, S.W., Harrison, B.A. and Bolton, J.L. (1985). The BRIAN Handbook - An Introduction to Landsat and the BRIAN (Barrier Reef Image ANalysis) system for users. CSIRO (Australia) Division of Water and Land Resources Natural Resources Series No. 3. Melbourne.

Jupp, D.L.B., Mayo, K.K., Harrison, B.A., Pickup, G., Chewings, V. and Kreymborg, R. (1987). Applications Oriented, PC-based Image Processing Systems : with particular reference to the CSIRO (Australia) microBRIAN system. Space Technology : Industrial and Commercial Applications 7 (3) pp 257-263.

Jupp, D.L.B., Mayo, K.K., Harrison, B.A., Hutton, P.G. and Ahmad, W. (1988). Remote Sensing and Image Processing on Personal Computers : the CSIRO (Australia) microBRIAN system. In "Desktop Planning : Advanced Microcomputer Applications for Physical and Social Infrastructure Planning". Hargreen Pub. Co., Melbourne.

MPA (Microprocessor Applications) Pty Ltd (1988). microBRIAN User Manual. 101-107 Whitehorse Rd, Melbourne.

SLP (Strategic Locations Planning), Inc. (1986). MAPEDIT User's Manual. 4030 Moorpark Avenue, Suite 123, San Jose, California.