USING IMAGES WITHIN A GIS FOR SPATIAL ANALYSIS

Eugene Derényi, Honorary Research Professor
David Fraser, Research Assistant
Department of Geodesy and Geomatics Engineering
University of New Brunswick
Fredericton, N.B., Canada E3B 5A3

Commission IV, Working Group 1

KEY WORDS: GIS, Database, Image, Raster, Vector, Spatial Analysis, Image Classification, Database Query

ABSTRACT

A universal geographic information system (UGIS), in which both vector based cartographic data and raster based continuous tone image data can be processed and managed, open new possibilities for performing spatial analysis. The software tools developed in this environment allows to:

- issue location specific database queries through an image display without the superimposition of the vector based geographic data; and
- extract location specific thematic information from images based either on per-pixel classification or on image statistics.

The information generated with the help of these analysis tools can directly be entered into the database.

1. INTRODUCTION

Geographic information is a collection of facts, about a geographic entity, that has been organized so that it conveys some meaning or significance. The information associated with a geographic entity has four major components [Aronoff, 1989]: i) geographic location; ii) attribute(s); iii) spatial relationships; iv) time. In other words, the four components define where the geographic entity is, what it is, what its relationship with its surroundings is, and when was it observed. The two most commonly used form of geographic data representation is the map and the image. There are distinct differences between these two data representations which impacts on their function and utilization in spatial analysis.

During the era when all geographic information was represented in hardcopy form, the two types of products were distinctly separated as cartographic line maps and orthophoto maps. In the softcopy world this distinction in the two forms of geographic data representation exists as digital vector data and raster image data. Advanced geographic information systems (GISs) allow the co-existence of the two data representation forms. Thus, the challenge is to exploit the advantages of both and to devise innovative methods for doing so. The spatial analysis scheme described in this paper is a step in this direction. First, however, a short discussion on the characteristics of these two forms of data representation is presented.

2. GEOGRAPHIC DATA REPRESENTATION

2.1 Maps

Mapping involves several processing operations, two of which are data abstraction (or generalization) and data reduction. These steps are used to create a model of the spatial phenomena of interest which is easier to understand than the real world. As a result, certain portions of the data gathered of the real world are ignored or obscured, while others are enhanced to generate a product which meets the requirements of users. Nevertheless, maps are useful for portraying information about the real world because they can succinctly describe a phenomena, can be easily understood and can provide spatial reference when making decisions regarding a phenomena. Typically, a map depicts either a specific spatial characteristic or a certain theme. In the digital world the information can easily be separated into layers. This approach of isolating the real world into layers, is useful for examining spatial relationships that exist between various themes and the spatial patterns which characterize these themes.

Even though maps have proven to be an effective tool in spatial information management, there are several weaknesses associated with using maps as information sources, for making decisions regarding the real world. These problems have been well documented (e.g. [Unwin, 1981], [Trotter, 1991]) and can be briefly summarized as follows: Maps are static and often outdated even before they are incorporated into a geographic database. Even though each map is a generalization of the real world, displaying all map layers of a multi-layer geographic database at once in superimposition may both complicate visual analysis and either hide or distort spatial patterns that exist. Another deficiency is that the landscape boundaries are generally fuzzy, which means that they rarely conform to the distinct lines shown on a map. finally maps can deceive, or in other words, maps may contain inadequate or misleading information.

2.2 Images

Airborne and especially spaceborne images provide a synoptic view of the landscape and can be collected at a various temporal, spectral and spatial resolutions. The various forms of remotely sensed data, which can be acquired at microscopic and macroscopic levels, provide up-to-date unabridged information of the real world. They are, however, more complex to understand and handle than the vector data.

In *Living Proof*, Petersen [1990] states that there are two data sources about an event: primary sources and secondary sources. Primary sources are actual eyewitnesses to an event and secondary sources are interpretations of the primary source. Images are primary sources. They are complete records of the sensor's view of a particular area, from a particular location, at a specific instance in time, and thus, are very useful especially if there is a question as to the reliability and currency of a secondary source. For example, any ambiguities associated with either the interpreted spatial location or the identification of geographical features may be resolved by querying the original source. Recent advances in computer aided technology have provided for the integration of remotely sensed imagery with maps, within a GIS, to overcome some of the deficiencies associated with using maps only as the primary source of information.

The advantages of using remotely sensing information with other types of geographic data for spatial information management has been well documented. For example, Derenyi and Pollock [1990] discusses several ways in which remotely sensed data may be used to update existing maps in a GIS. Fung et al. [1993] outlines a system which allows the user to incorporate both satellite imagery and conventional GIS data sources for forest inventory management. Price [1995] and Wilson [1995] describe systems which use high resolution digital orthophotos to provide up-to-date, reliable geographic information for facilities management.

Integrating remotely sensed data sources with digital maps in a GIS is not a trivial issue [Trotter, 1991]. Two factors which have impeded the integration process are that [Edwards, 1993] fundamental differences exist between the nature of the information resulting from digital image analysis and digital line maps; and the integration of image analysis technology with GIS has not progressed much beyond data exchange and simultaneous display.

Images can play three distinct roles in a GIS:

1. A passive role, when the image serves as a backdrop to the various layers of information stored in vector form. It aids the viewer of such a composite to understand the geographic entities recorded as points, lines and polygons by portraying the information lost during data abstraction.
2. A stand-alone role, when the image itself serves as a base map. The digital orthoimage maps, which are rapidly increasing in popularity, fall into this category.
3. An active role, when images are incorporated in performing spatial analysis and data queries in a multi-layer geographic database.

A scheme to facilitate the active role of images in a GIS is now presented.

3. IMAGE SUPPORTED SPATIAL ANALYSIS

Spatial analysis means to study the relationships between geographic features using spatial and non-spatial (attribute) data and to answer questions about the real world.

A universal GIS (UGIS), one which can process and manipulate both vector based geographic data and raster based continuous tone image data, and is interfaced with a relational database management system (DBMS) provides the best environment for this purpose. Positional and attribute information can be accessed and cross-referenced quickly and efficiently. New thematic information can be generated easily or existing information updated and entered into the database.

A usual procedure for a database query is to display the polygon layers of interest and initiate a search in the database by pointing on the entities whose attributes are sought or submitting a list of feature codes. The DBMS then returns the information requested.

Two difficulties may occur in relying solely on a vectorized data representing geographic features:

1. Displaying several data layers at once in superimposition complicates the location and selection of the entities visually. Therefore the queries are processed a few layers at a time.
2. In the absence of suitable landmark features on the data layer queried or if the map layer is out of date, it may be difficult to find the correct geographic location of the information needed. This is especially true in exploratory investigations.

The image based data query can alleviate these difficulties. The concept is rather simple. An image of suitable scale is displayed at the GIS workstation without any superimposition of vector data. An image provides full details of the land cover not only the outlines. It also contains qualitative information which can be assessed visually before the actual database query is issued, to localize the query area and thus reduce the processing time. The analyst then points with the cursor on the location in the image where information

is required and enters the code of the data layers to be queried. After a search through the database, the attributes found for that location are displayed. The polygons surrounding this location in each layer may also be displayed if so requested.
An alternate version of this procedure is to localize the query by delineating a window in the image. The database search then returns the attributes of all polygons which are inside or intersected by this window and the polygons in question are displayed.

Another situation that could arise in spatial analysis is that new information is sought within certain geographic boundaries or the current information has to be verified or updated. These boundaries are either stored already in the database or may be delineated by the analyst. Digital image classification techniques can then be employed for this purpose. A prerequisite is, however, that suitable training samples can be located for each information class searched for.

Two procedures can be followed in this type of query [Derenyi and Turker, 1996]. In the first case, a per-pixel image classification is performed covering the entire study area, using one of the standard algorithms such as the parallelepiped or the maximum likelihood classifier. Thereafter, a polygon-by-polygon assessment of the result is performed to establish the class distribution within the specified geographic boundaries. The individual polygons are located through the DBMS for assessment.

In the second case, the image classification is bypassed entirely and only image statistics are generated within the polygons to be analyzed. Mean, standard deviation and median values are possible statistics selected. A comparison of these statistics with those obtained for the training samples forms the class assignment of each polygon. This approach is especially suitable for monitoring and change detection.

Database queries may be issued through any kind of imagery, be it airborne or spaceborne, monochrome, colour or multispectral. The image must of course be geometrically registered to the geographic database. Image query through classification is primary intended for spaceborne multispectral imagery. Decision making based on image statistics only is especially efficient for analyzing single or small groups of polygons scattered through a study area. The polygons do not have to be pre-established but can be drawn by the analyst in an interactive mode. All information generated through the image supported spatial analysis can, of course, immediately move into the database. This spatial analysis scheme was successfully implemented and tested in the Computer Aided Resource Information System (CARIS), GIS which supports both vector and raster data handling [Derenyi and Pollock, 1990; Derenyi, 1991].

4. CONCLUSIONS

Incorporating images within a GIS for spatial analysis is a powerful tool with several advantages over the conventional, vector data only, technique.

- Image analysis can be performed within known geographic boundaries, which are stored in the GIS and can be queried through the DBMS.

- Database queries can be issued through image displays which provide a more complete and detailed view of the real world than graphical displays do.
- The new information generated through digital image analysis can directly be entered into the database.
- The interactive query location selection feature of this scheme is especially attractive for exploratory, browsing type spatial analysis.

ACKNOWLEDGMENT

The authors wish to acknowledge the contribution of Mustafa Türker who has implemented some of the concepts presented in this paper.

REFERENCES

Derenyi, E. and R. Pollock, 1990. Extending a GIS to support image-based map revision. *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 11, pp. 1493-1496.

Derenyi, E., 1991. Design and development of a heterogeneous GIS. *CISM Journal ACSGC*, Vol. 45, No. 4, pp. 561-567.

Derenyi, E. and M. Turker, 1996. *Polygon based image analysis in a GIS.* Proceedings of the 19th Canadian Symposium of Remote Sensing, Vancouver, British Columbia.

Edwards, G., 1993. *The integration of remote sensing and GIS: Fundamental questions and new approaches.* Proceedings of the 16th Canadian Symposium of Remote Sensing, Sherbrooke, Québec, pp. 873-878.

Petersen, P., 1900. *Living Proof.* Colorado Springs, U.S.A., NAVPRESS.

Price, K., 1991. "Spatial modeling using geographical information systems." *The ABC's of GIS.* American Society of Photogrammetry and Remote Sensing.

Pries, R., 1995. A system for large scale image mapping and GIS data collection. *Photogrammetric Engineering and Remote Sensing*, Vol. 61, No. 5, pp. 503-511.

Trotter, C.M., 1991. Remotely sensed data as an information source for geographical information systems in natural resource management: A review. *International Journal of Geographical Information Systems*, Vol. 5, No. 2, pp. 225-239.

Unwin, D., 1981. *Introductory Spatial Analysis.* New York, U.S.A., Metheun and Co. Ltd.

Wilson, J., 1995. Client-vendor partnership advances digital aerial photography system. *GIS World*, Vol. 8, No. 9, pp. 44-47.