Integration of Remote Sensing and GIS: Data and Data Access

Manfred Ehlers
Department of Geoinformatics, International Institute for Aerospace Survey and Earth Sciences (ITC), Enschede, The Netherlands
David Greenlee
EROS Data Center, USGS, Sioux Falls, SD 57198
Terence Smith and *Jeffrey Star*
NCGIA, University of California, Santa Barbara, CA 93106

ABSTRACT: The integration of remote sensing tools and technology with the spatial analysis orientation of geographic information systems is a complex task. In this paper, we focus on the issues of making data available and useful to the user. In part, this involves a set of problems which reflect on the physical and logical structures used to encode the data. At the same time, however, the mechanisms and protocols which provide information about the data, and which maintain the data through time, have become increasingly important. We discuss these latter issues from the viewpoint of the functions which must be provided by archives of spatial data.

INTRODUCTION

THE NEED TO INCLUDE ANCILLARY INFORMATION (e.g., maps and ground surveys) in the process of interpreting remotely sensed data has long been acknowledged by the remote sensing community (Campbell, 1986; Lillesand and Kiefer, 1987). On the other hand, the advantage of using information derived from remotely sensed data to correct, update, and maintain cartographic databases and geographic information systems (GIS) has been amply demonstrated over the last years (Ehlers, 1989; Nellis *et al.*, 1990; Eckhardt *et al.*, 1990).

For remote sensing and GIS to be truly integrated, however, several technical and scientific impediments still need to be overcome. Such integrated geographic information systems (IGIS) must facilitate more than data transfer between separate systems or screen overlays (Ehlers, 1990). One major problem for the integration is caused by the difference in the structures used to acquire, access, and store the data. Consequently, one of the purposes of this paper to identify the key impediments for IGIS that are based on data structures. Another purpose of this paper is to examine the functions of an archive or repository of spatial data in order (1) that information about the data is provided to users in a meaningful way, and (2) that the data are maintained for long periods of time. We see GIS and remote sensing as one entity, concerned with acquiring, managing, and analyzing geographic data.

Many topics that are related to the integration of remote sensing and GIS (such as large spatial databases, error analysis, user interfaces, IGIS processing, or IGIS computing environments) are treated elsewhere in this issue, or are current research projects of other NCGIA initiatives. In this paper, we focus on the acquisition, storage, and dissemination of spatial data. Following this brief introduction, we discuss the data acquisition processes in remote sensing as well as geographic information system applications. In this section, we characterize some of the different types of data traditionally employed in remote sensing and GIS. Next, we examine characteristics of data storage and user access. Various physical data structures are briefly examined, as well as recent ideas for the long-term storage and dissemination of spatial data. Finally, we discuss the key research themes that have come from the NCGIA Initiative 12 meetings.

DATA ACQUISITION IN REMOTE SENSING AND GIS

The term remote sensing is commonly restricted to methods that employ electromagnetic energy (such as light, heat, and radio waves) as the means of detecting and measuring target characteristics, excluding electrical, magnetic, and gravity surveys that measure force fields rather than electromagnetic radiation (Sabins, 1987). As a source of geographic information, digital remote sensing represents more than a simple extension of conventional aerial photography, requiring fundamentally different approaches to the analysis of Earth surfaces (Everett and Simonett, 1976).

The function of an information system, in general, is to improve one's ability to make decisions. Consequently, an information system is a chain of operations from planning the observation and collection of data, to storage and analysis of data, to the use of the derived information and some decision-making process (Star and Estes, 1990). A geographic information system is an information system for spatial data that are referenced by geographic coordinates. It is designed to acquire, store, retrieve, manipulate, analyze, and display these data according to user-defined specifications (Marble *et al.*, 1983).

Based on these definitions, the overlap in emphasis and approach is clear. To identify the key impediments for the integration of remote sensing and GIS, it is necessary to focus on the differences in data acquisition, storage, and access.

REMOTE SENSING DATA TYPES

There is a fundamental dichotomy in the use of remotely sensed data between the original (or raw) data collected by the sensor and the products which are commonly extracted from these raw data. We examine each separately in the next sections.

Raw Remote Sensing Data

In general, remote sensing image data represent the result of a one-dimensional sampling in the time domain and a two-dimensional sampling in the spatial domain. Basic properties of a remote sensor can be summarized as (Davis and Simonett, 1991; Strahler *et al.*, 1986):

- spectral coverage and resolution
- spectral dimensionality (i.e., number of bands)
- radiometric resolution (quantization)
- instantaneous and angular field of view (IFOV)
- point spread function (PSF)
- temporal response function

We briefly discuss some of the fundamental data types which arise from remote sensing systems, from the perspective of the acquisition and recording process.

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING,
Vol. 57, No. 6, June 1991, pp. 669–675.

0099-1112/91/5706–669$03.00/0

Images. The most common form of data acquisition for remote sensing is the creation of two-dimensional raster datasets. Key distinctions arise between passive and active systems, as well as the mechanisms used to collect the data (which strongly influence the geometrical and radiometrical quality of the raw data). Depending on the recording device, we can differentiate four geometric principle recording mechanisms: central perspective (where the sensing system works in a fashion analogous to a camera; Luhmann, 1990), along-track (or line) scanners (such as those used on the French SPOT satellites), cross-track (or mirror) scanners (such as those used on the Landsat Multispectral Scanner (MSS) and Thematic Mapper (TM) sensors), and active sensors (such as radar, sonar).

Profiles. Profiles can be obtained in vertical or horizontal directions. Examples include temperature soundings in atmospheric remote sensing or laser/radar altimetric profiles of the terrain. Also, spectrometers and radiometers can provide point or profile sampling. In combined mission with imaging sensors, they may provide efficient means for calibration and absolute references.

Image cubes. Combining imaging techniques and spectrometric measurements led to the development of imaging spectrometers (Goetz *et al.*, 1985). A specified part of the electromagnetic spectrum is continuously sampled, creating image "cubes" without gaps and with extremely narrow bandwidths in the spectral domain. Examples for this technology are the Airborne Imaging Spectrometer (AIS) and the High Resolution Imaging Spectrometer (HIRIS) for the planned Earth Observing Station (EOS) (Goetz *et al.*, 1985; Butler *et al.*, 1986).

Interpreted Remote Sensing Data

In contrast to the raw remotely sensed data described above, users typically transform these quasi-continuous datasets into a diversity of variables in order to begin to extract information from the raw data. These irreversible transformations create several general classes of intermediate data, which we examine in the next paragraphs. In addition, we mention that some of the datasets are often used as a part of the interpretation process.

Raster Image Data. Raster images derived from remotely sensed data are generally stored as fully represented scan lines (i.e., they are seldom compressed or run-length encoded). Typically, these derived products are the results of a calibration or enhancement process, to correct for such effects as low sun angle or sensor striping, or to generate improved contrast, enhance edges, or eliminate systematic image "noise." Examples of these include the derivation of albedo or greenness (Moik, 1980). In contrast to these quasi-continuous products, there are a number of transformations that create nominal raster datasets. A common approach is that of classification, where the original multivariate data are reduced to a single set of themes or classes. Classification effectively reduces the multi-dimensionality of the raw sensor data to a one-dimensional feature space in which the elements can be labeled with a class name or some other descriptive attribute (Swain and Davis, 1978; Duda and Hart, 1973; Mather, 1987).

Vector Data. In traditional remote sensing applications, vector data are required in a limited number of cases. In fact, many remote sensing systems immediately convert vector data to raster form in order to avoid the complexity of handling other data types. These data often include ground truth (the results of field surveys or Earth-based instruments and measurements) and ancillary map data (which is either raster-scanned or vector-digitized). In a number of applications, linear features are extracted from the remotely sensed images, based on raster primitives such as convolution filtering (Sijmons, 1986) and image segmentation procedures (Pratt, 1978) which are followed by conversion to a vector data structure (Greenlee, 1987).

Terrain information can be preserved in a number of data formats, including raster arrays and vector-like structures (such as sets of contour lines or TIN). Frequently, the primary data source for terrain information has been aerial photography coupled with photogrammetric processing which is still based on human interpretation; efforts have been made since the early 50s to replace the operator by automated procedures (Konecny and Pape, 1981; Ehlers, 1985a; Förstner, 1986; Fuller and Ehlers, 1990). Another primary source for terrain information is field measurements using conventional survey or Global Positioning System (GPS) technology. These are inefficiently stored in simple raster arrays. However, most of the digital terrain information used in spatial data processing and analysis comes from secondary (interpreted) data sources. The most common input is the manual digitization of contour lines from existing maps, with subsequent vector set or raster array construction.

GIS Data Types

In traditional applications, GIS data types span a wider range of sources than those used in traditional remote sensing applications. We break the following discussion along the common raster/non-raster dichotomy.

Raster Data

While some operations (e.g., overlays, spatial statistics, model integration) are simplar and faster in the raster than in the vector domain (Burrough, 1986), there are many datasets which are naturally represented in raster data structures. Possible sources for raster input include transformations of remotely sensed data, rasterized versions of cartographic data, interpolated point or profile measurements, and scanned maps.

Where existing graphic representations of spatial data are used as sources, the fidelity of the output and the amount of preprocessing required to create "sufficiently" accurate data layers depend largely on the quality of the course, the geometric and radiometric accuracy of the scanner, and the capabilities of the associated software (Ehlers, 1985b).

Vector Data

The most common data source of vector data for GIS comes from existing maps. Features of interest are extracted using digitizer tablets and electromechanical scanners; each of these extraction processes has different kinds of error in accuracy and precision. The output depends on the skill of the operator and the quality of the associated systems.

Direct input from digital cartographic databases, CAD systems, or photogrammetric data files is usually possible through data exchange programs. Here, national and international standards have been developed or are evolving that allow data transfers between most of the cartographic/photogrammetric/CAD data files (Billingsley, 1990; Johnson, 1990; Moellering, 1991).

Other Data Types

In addition to the sources mentioned above, there are other common sources of spatial information.

Field Data. Remote sensing data and their derivatives must be related to, and calibrated with ground based measurements (e.g., field samples, ground truth) and calibration models of various kinds. Conventional map data must also be validated against observations in the field.

Deterministic Model Data. Deterministic models attempt to describe quantitatively the physical, chemical, biological, ecological, or economic structures and developments for regions of various size (e.g., global, national, state, district, city, or block). They may be spatial, temporal, or spatio-temporal in nature. Models may range from simple one-parameter logarithmic growth models (for instance, algae bloom as a function of time) to complex atmospheric or weather models with a set of parameters. Integration of these models will tax current capabilities of remote sensing and geographic information systems (Burrough *et* *al.*, 1988; Itami, 1988; Ehlers *et al.*, 1989). The data structures used in these models are rarely the same as those found in GIS.

Survey Measurements. Survey measurements can be used to accurately relate GIS and remote sensing information to a geodetic coordinate system. Usually, survey data (including, for example, GPS measurements) are used to triangulate a geodetic network. Recent developments, however, stress the storage of raw survey data from which a network can be triangulated "on the fly." The advantage of such a survey-based measurement system is that new measurements can easily improve such a network (Hintz and Rodine, 1990).

DATA STORAGE AND ACCESS

Issues of science data management have been at the forefront of many disciplines for a number of years. Seminal documents in this area include the reports of the Committee on Data Management and Computation of the National Academy (CODMAC, 1982; CODMAC, 1986), as well as more recent technical papers such as the 1989 conference proceedings from NCGIA Initiative 5: Large Spatial Databases (Buchman *et al.*, 1990). We direct our comments in this section on those areas that we believe are specific to large *spatial* datasets. First, we examine the physical data structures used to store the data described in the previous section on practical computer peripherals. Then, we examine some issues in providing users with effective access to these large data stores.

PHYSICAL DATA STRUCTURES

For GIS and remote sensing data to be stored in a readily usable form, they have to follow specified data structures. Structures for spatial data may differ in several ways from one another (Frank and Barrera, 1990):

- type of geometric data (point *vs.* region),
- object handling (non-fragmenting *vs.* fragmenting),
- division of space (regular *vs.* data determined), and
- retrieval (direct *vs.* hierarchical).

We continue the organization of previous sections and discuss raster and vector systems separately, followed by some comments on hybrid systems.

Raster Systems

Raster tessellations of space provide explicit information for each location (Burrough, 1986). They may be described as discrete samples of continuous fields, as opposed to object-based representations provided by vector structures, referring to the fact that fields are assigned object attributes in a raster model whereas objects are given locations and attributes in an object model (Ehlers *et al.*, 1989).

Image Files. The most common raster structure is a square lattice whose values are stored as two-dimensional arrays in the computer. This structure essentially follows "naturally" the structure of imaging devices in remote sensing or scanning devices for digitizing. Its advantages include simplicity, ease of display and processing, ease of data aggregation and overlay, and uniform cell size which allows multi-dimensional spatial analysis and modeling.

Compressed Raster Structures. Several methods of data compaction have been developed to store raster data more efficiently making use of coding (sun as run-length algorithms) and hierarchical structures (top-down approaches such as the quadtree; Samet, 1989). Access and processing of such compacted data, however, usually requires decoding.

Vector based systems

CAD Systems. Computer-Aided Design (CAD) systems are often compared with GIS. The data model for CAD systems is generally simpler than is needed for GIS applications. For example, feature topology is generally quite limited in CAD systems. In addition, while there may be several tabular or numeric attributes for the features maintained in a CAD system, they do not generally contain a comprehensive database management and manipulation capability found in many vector GIS.

Topological Arc. Systems that are based on topological arcs employ a data model that requires that polygons can be formed as needed from chains or arcs. While these systems require complex algorithms for creating and maintaining topological structure, they tend to be a robust and comprehensive model for performing spatial analysis tasks.

Full Polygon. Systems that employ a full polygon representation can be used for some GIS applications that require polygons to be cycled for display and analysis. In applications that require such operations as overlay, buffer zone calculation, or line generalization, the fully polygon model is generally lacking. The principal advantage of this data model is its simplicity for display of polygons and their attributes. Because each polygon is stored without regard to adjacent polygons, it is possible to create and propagate errors in the spatial database that cannot be easily detected without special consideration (topological checking or validation).

Network. Systems that employ a network data model generally support different applications from topological arc and full polygon systems. In the network model, the arcs and nodes provide the spatial framework for describing flows and directions. Network analysis and display of flow through a network (e.g., hydrology, traffic, facility siting).

Hybrid Systems

Raster and vector data structures are clearly two and members in the set of spatial data structures. There has been debate as to which is the most effective, from the viewpoints of storage efficiency, processing efficiency, and ability to capture the essence of geographic space for an application. There can be no clear answer to this debate, because the answer depends on a large number of factors, including the nature of the application, the kinds of data and information, and the distribution of the queries. It seems appropriate to note that the main difference between them relates to the degree to which the spatial entities of interest are explicit in the data.

In general, rasters provide a generic form of representation in which most objects and the spatial interrelationships between them are implicit. This is particularly appropriate where the raster is considered a discrete sample of a continuous field. A large set of operators may be used to make the objects and their interrelationships more explicit. On the other hand, vector data structures typically represent an additional "layer" of processing and interpretation on top of a raster structure (e.g., as in the digitizing and topology-producing processes) and objects are made explicit in terms of points, lines, and polygons and their interrelations.

Hence, one may conceptualize a large family of representational schemes characterized by the way in which they explicitly encode information about the objects and relationships of interest; one is characterizing the degree of object orientation involved in the data structure. The key reason for object orientation is to represent objects in a manner that a user in some application will find more efficient, in terms of representation and manipulation. Hence, the essential tradeoff is the cost of processing in order to produce object oriented data structures versus the additional user cost if the preprocessing is not performed.

Based on this, it follows that the optimal choice of data structures for an application mix should relate, at least in part, to the degree to which a user finds it worthwhile to perform vector preprocessing and to store the objects in a more explicit form. For example, in a spatial database system in which one is storing large amounts of image data relating to meteorological phenomena, it is not possible *a priori* to process the data and to obtain polygonal representations of all objects that will be of interest to users, because the cost would be prohibitive even if one could decide upon or even define the objects of interest. On the other hand, it might be of great value in such a system to have a vector representation of certain important objects, such as the outlines of terrain features (e.g., coastlines). For this application, one could design a system with both raster and vector capabilities, as well as tools to interrelate them.

While one can argue that a mix of data structures is the optimal solution in the abstract, it is clear that there is no practical *a priori* optimal solution. There is no technical obstacle to providing procedures that allow one to move easily between various physical implementations of spatial data. The main issue is providing users with efficient and accessible means to do so, which operate with known effects on the data's accuracy and precision.

In order to examine issues such as these, we suggest that it would be useful to explore

- High-level (declarative) languages, in which the users can easily express their requirements concerning processing by indicating, for example, operations that they would like to carry out on certain data sets, independent of whether they are in raster or vector form.
- A database management system, which would then take care of the details of such a transaction with respect to data representation. In a sense, this service is natural for a DBMS in the sense that database systems are specifically designed to hide from the user the various levels below the view level.

A key problem is that users of remotely sensed and GIS data are still thinking in terms of file systems rather than DBMS. In order to merge remotely sensed and conventional GIS data in a single analytic system, we believe it is important to

- provide a high-level language and view support in which the users expresses what they wish to compute in a way that is relatively independent of irrelevant details, such as the physical organization of the data; and
- provide a DBMS that is capable of handling both types of data representation, and optimizing over their use; hence, such a DBMS might include, for example, facilities for converting from one representation to the other and for making objects that are implicit in the raster form and explicit in the vector form in a quasi-automated manner.

More generally, a key issue is whether and how to support more general classes of data structures that involve different degrees of object orientation, and the provision of a DBMS that supports the different representations in an efficient and transparent manner, with the ability to integrate data from different representations. Such problems are good candidates for research.

Data Archive Systems

Archive systems for spatial data are relatively rare; the exceptions are often the large, homogeneous government holdings. Further, users are currently required to understand the use of each of a number of archives to be able to locate and examine potential sources of spatial information. This is exacerbated when trying to integrate remotely sensed and traditional cartographic data, because they are often collected and maintained by different agencies. Irrespective of the source of the data, there are a number of requirements which are imposed on a long-term archive of spatial data, in order that these data are accessible to users and maintained in a useful fashion over a long period of time. Issues in the design and operation of archives include volumetric efficiency of data storage, stability of the media, appropriateness of the user interface, and cost-effectiveness with high performance. In the following discussion, we group these into two areas: issues which are based on a consideration of the physical storage medium, and those which are based on considerations of the user's interaction with the archive.

Media

Conventional removable magnetic media have serious limitations as a means to archive large datasets for long periods of time. Industry standard 9-track computer tapes have limited lifetimes (on the order of 5 years between maintenance), limited data storage (on the order of 100 Mbytes in a volume of 4500 cm^3), restrictions on access modes (sequential, in contrast to random), a relatively high cost to copy, and relatively expensive hardware. The storage industry has recently provided two new tape formats based on tape formats and transports used in mass-market entertainment systems: 4-mm DAT (digital audio tape) and 8-mm video tape. Each of these has excellent storage density (for example, 2 or more Gbytes in 200 cm^3 for 8-mm tape), and intermediate costs for the drive and interface. Serious concerns in using these technologies include the serial nature of the medium, relatively slow I/O speeds (affecting copying time and cost), and lifetime of the medium. These tapes may be suitable for cost-effective distribution of large volumes of spatial data to users.

CD-Rom disks are being used in a number of experiments as a means to store and distribute spatial datasets. While their lifetime is under debate, their ability to hold on the order of 600 Mbytes of data in 200 cm^3, their fundamentally random access nature, and extremely low cost to make copies (on the order of a few dollars each) make them very attractive. Mastering costs for producing a CD-Rom have dropped dramatically in the past 24 months, making it possible for smaller agencies to think about creating their own CD-Rom disks. Availability of low-cost drives for a wide range of computers and operating systems make these disks appealing as well.

There are at least three other optical disk technologies that must be considered as well. WORM (write once, read many) disks, erasable (often using magneto-optical technology), and video disks are all in use to varying degrees in the computer industry. WORM and erasable disks at this time are expensive (disks more than $100, drives on the order of several thousand dollars), copying times are long, and standards to permit data written on a disk to be read on multiple systems are lacking. The precision of recalling data from video disks, a fundamentally analog medium, is perhaps not known. However, the low costs of players and duplication may make video disks extremely attractive for browse or other index datasets.

Archives of the future will need to consider not only these kinds of issues of storage costs and data reliability, but also the media which are used to distribute information to the users.

User Interface and Access

An important set of concerns for spatial data archives involve the ways in which users identify potentially interesting datasets, and then obtain copies of these datasets. The former includes modes of the user interface and information which describes the datasets. The latter includes consideration of distribution mechanism and media, price versus performance, and data format. The single most important criterion for data selection, however, is almost certainly the content of the data. Access by content is an extremely interesting, important, and difficult problem to solve, and it is currently an area of increasing research and development activity.

A number of studies have examined portions of the user interface to collections of spatial data (Star *et al.*, 1987). In general, these have shown a need for

- graphics (an electronic map for specifying the spatial region of interest, for example);
- hierarchical levels of detail (perhaps including a general descrip-

tion of a collection, documentation of the kinds of information stored and how they were gathered, and a detailed inventory of the collection);
- browse (interactively accessible samples of the data, perhaps reduced in spatial precision as well as other characteristics);
- interactive query of data characteristics, with intelligent aids; and
- ability to store user profiles and use them to optimize its response in the future.

These capabilities are not present in the spatial data archive systems now in operation. Concerns about obtaining copies of the identified datasets typically revolve around more practical issues. These include

- ability to request data in "standard" formats;
- ability to acquire documentation about the datasets, including information about other uses of the data and available software for processing;
- distribution mechanism (electronic vs. delivery of media);
- cost trades between various mechanisms for delivery; and
- time probably required to satisfy the request.

These issues clearly impact users when trying to locate datasets of potential interest from a universe of potential suppliers.

KEY RESEARCH ISSUES

Although the potential for the integration of remote sensing and GIS is evident and some success has been achieved in several areas, improved integration of remote sensing and GIS require that a number of research questions be addressed. One key impediment to a total integration is the different concepts of space on which GIS and remote sensing are based. Investigation of this topic is seen as one fundamental research question.

Concept of Space in GIS and Remote Sensing

GIS may be viewed as object-based data representations whereas remote sensing employs a field-based model. Other differences include the level of abstraction, the level of accuracy and precision, scale, metric, and temporal abstraction. Within this concept, there are several research issues which need to be addressed:

- Is there a unifying theory (e.g., a model of the world) that would allow one to see remote sensing and GIS as different representations in this "model of space"?
- What are the predominant parameters in this representation model that differentiate GIS from remote sensing data structures?
- Is it possible to define transformations between the different representations?
- Do transformations without loss of information exist? If not, can this loss be quantified?
- To what degree are these transformations reversible? If not, can an inverse transformation be approximated?

Data Conversion and Exchange

Given the wide range of data types and formats which are currently in use in spatial analysis, there are a set of unanswered questions about the conversion and exchange of spatial data. These include

- What are the requirements for remote sensing data (temporal, spatial and spectral resolution, timing, areas, etc.) to be incorporated in GIS applications (e.g., urban, regional planning, global monitoring, etc.)?
- What data structures and data management strategies are appropriate for GIS guided image interpretation?
- What are the data exchange/conversion standards within GIS and remote sensing to provide best access to large distributed cartographic, GIS, and remote sensing databases? Are these standards appropriate?
- Are there alternative concepts for integration, i.e., information exchange rather than data transfer?
- Is there a need to set up an integrated test data set containing remote sensing data, cartographic data, field measurement data, and possibly model data?

Integrated Databases

Following the raster/vector discussions above, we make the additional comment that remote sensing systems and GIS are not fundamentally different, in the sense that users are interested in analyzing the objects and relationships that are encoded in some set of spatially referenced data. This is particularly the case of raster-based GIS; hence, the research problems in both remote sensing systems and GIS are closely related.

As noted above, the key to better integration may be the construction of DBMS that provide services that are specific to spatial data handling. To this end, it is necessary to determine user requirements with respect to such services, and then design and implement systems that embody them. Several research issues of major importance in this area are

- Determining the user requirements with respect to the degree of object orientation required in a system that handles remote sensing data, and deciding on the degree and nature of object orientation that is required in such a system in order to satisfy user needs.
- Deciding whether a single DBMS architecture is the correct approach to follow, or whether a more modular approach is desirable, in which one can build a system of different components in order to handle the varieties of data models that one may be required by users to handle.
- If a large DBMS is a reasonable approach, to design and implement a DBMS that can handle different data models in an efficient and (where necessary) transparent manner. Extensibility and efficiency are key issues for such a system.

Display and User Interface

A number of organizations are at the technological frontiers of user interface and data visualization. Particularly with respect to spatial data of a variety of kinds of formats and scales, we identify a number of research topics that we believe are unique to IGIS. These include

- Is it possible to establish a consistent set of terms that can serve as a standard for GIS and remote sensing systems? The lack of consistency may be illustrated in any number of comparisons, including lines and samples in digital images versus latitude and longitude, x and y, and Eastings and Northings in geographic information systems, or display coordinates usually presented in the first Cartesian quadrant in GIS versus fourth quadrant in remote sensing.
- Is it possible to establish a complete and consistent processing treatment of features? For example, a description of spatial objects based on complex polygons with island topology is more complete than simple convex polygons with no allowable inclusions; the latter is typical in remote sensing. Another example of restrictions on analysis involve systems that do not permit a complete set of data elements within raster structures (including byte, integer, and floating point elements of various dynamic ranges).

Inventory mechanisms that provide descriptions of, and access to, existing GIS and remote sensing data are clearly inadequate to the next decade of Earth science (Earth Systems Sciences Committee, 1988), as well as the technological demands of storing terabytes of data per day from decade-long remote sensing missions (Chase *et al.*, 1986). These lead to a number of priority concerns, including

- Can we develop (and facilitate widespread acceptance of) spatial data exchange standards?
- Can we define minimum standards for the encoding and digitizing of spatial data (to minimize redundant data collection and inadequate quality control)?
- Are there databases that should be made available (or more convenient) to users? (e.g., DTED).

CONCLUSIONS

The lists of key research topics presented in the previous section could provide the focus for many a Ph.D. dissertation, and keep many of us busy for our entire careers. While these lists of research topics present a non-exhaustive list of areas which were considered of high priority, we conclude with several others which may be of longer-term interest, or are perhaps of a more theoretical nature.

Today remote sensing and GIS mainly deal with two-dimensional data structures. As the Earth is essentially three-dimensional, extensions of two-dimensional models need to be studied (Raper, 1989). Also, integration of models may require higher dimensional data structures. Current GIS do not address the time domain whereas remote sensing is a sampling in both space and time. Specific research topics include

- How can GIS and remote sensing data models be extended to handle three-dimensional data (DEMs, geological data, atmospheric data, marine data, hydrological data, and so forth)? Possible approaches include independent layers, 2.5D, and full 3D.
- How can time be treated in GIS and remote sensing data models (explicit parameter versus time samples)?

We appreciate the opportunity to present these priorities, and a glimpse of their underlying thought processes. They are not intended to denigrate other topical areas, nor to constrain debate. Rather, they are designed to focus attention on problems which may be tractable in the next few years, and to sharpen discussion of priorities in the community.

REFERENCES

Billingsley, F. C., 1990. Language Requirements for Understanding retrieved Data. *Proceedings, ISPRS Commission II Symposium*, Dresden, GDR, pp. 89–95.

Buchman, A., O. Gunther, T. R. Smith, and Y. -F. Wang (eds.), 1990. *Design and Implementation of Large Spatial Databases*. Lecture Notes in Computer Science, Vol. 409. Springer-Verlag, Berlin.

Burrough, P. A., 1986. *Principles of Geographic Information Systems for Land Resource Assessment*. Clarendon Press, Oxford.

Burrough, P. A., W. van Deursen, and G. Heuvelink, 1988. Linking Spatial Process Models and GIS: A Marriage of Convenience or a Blossoming Partnership? *GIS/LIS '88*, San Antonio, Texas, Vol. 2, pp. 598–607.

Butler, D. M., *et al.*, 1986. *From Pattern to Process: The Strategy of the Earth Observing System*. NASA Eos Steering Committee Report, Vol. II, 140 p.

Campbell, J. B., 1987. *Introduction to Remote Sensing*. The Guildford Press, New York, USA.

Chase, R.R.P., *et al.*, 1986. *Data and Information System: Report of the EOS Data Panel*. NASA TM-87777, Vol. IIa.

Committee on Data Management and Computation, 1982. *Data Management and Computation. Volume 1: Issues and Recommendations*. Space Science Board, Commission on Physical Sciences, Mathematics, and Resources, National Research Council. National Academy Press, Washington, D.C.

———, 1986. *Issues and Recommendations Associated with Distributed Computation and Data Management Systems for the Space Sciences*. Space Science Board, Commission on Physical Sciences, Mathematics, and Resources, National Research Council. National Academy Press, Washington, D.C.

Davis, F. W., and D. Simonett, 1991. Remote Sensing and GIS. *Geographic Information Systems: Principles and Applications* (D. Maguire et al., editors), Longman LTD., in preparation.

Duda, R. O., and P. E. Hart, 1973. *Pattern Classification and Scene Analysis*, New York, John Wiley and Sons, Inc., pp. 10-256.

Earth Systems Sciences Committee, 1988. *Earth System Science: A Closer View*. National Aeronautics and Space Administration, Washington, D.C.

Eckhardt, D. W., J. P. Verdin, and G. R. Lyford, 1990. Automated Update of an Irrigated Lands GIS using SPOT HRV Imagery. *Photogrammetric Engineering & Remote Sensing*, Vol. 56, No. 11, pp. 1515–1522.

Ehlers, M., 1985a. The Effects of Image Noise on Digital Correlation Probability. *Photogrammetric Engineering & Remote Sensing*, Vol. 51, No. 3, pp. 357–365.

———, 1985b. Large Format Raster Scanners for the Geosciences. *International Yearbook of Cartography*, pp. 37–50.

———, 1989. The Potential of Multisensor Satellite Remote Sensing for Geographic Information Systems. *Technical Papers 1989 ASPRS/ACSM Annual Convention*, Baltimore, Maryland, Vol. 4, pp.

———, 1990. Remote Sensing and Geographic Information Systems: Towards Integrated Spatial Information Processing. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, no. 4, pp. 763–766.

Ehlers, M., G. Edwards, and Y. Bedard, 1989. Integration of Remote Sensing with Geographic Information Systems: A Necessary Evolution. *Photogrammetric Engineering & Remote Sensing*, Vol. 55, No. 11, pp. 1619–1627.

Everett, J., and D. S. Simonett, 1976. Principles, Concepts and Philosophical Problems. *Remote Sensing of Environment*. (J.L. Lintz and D.S. Simonett, editors.) Addison-Wesley Publishing Co., Reading, Massachusetts, pp. 85–127.

Frank, A. U. and R. Barrera, 1990. The fieldtree: a data structure for geographic information systems. *Design and Implementation of Large Spatial Databases* (A. Buchmann, O Gunther, T.R. Smith, and Y-F Wang, editors). Lecture Notes in Computer Science, Vol. 409. Springer-Verlag, Berlin, pp. 29–44.

Förstner, W., 1986. A Feature Based Correspondence Algorithm for Image Matching. *Proceedings, ISPRS Commission III Symposium*, Rovaniemi, Finland.

Fuller, M., and M. Ehlers, 1990. Automated System for Image Integration in Remote Sensing. *Proceedings, ISPRS Commission VII Symposium*, Victoria, B.C., Canada (in press).

Goetz, A. F. H., G. Vane, J. E. Solomon, and B. N. Rock, 1985. Imaging Spectrometry for Earth Remote Sensing. *Science*, Vol. 228, pp. 1147–1153.

Greenlee, D.D., 1987. Raster and Vector Processing for Scanned Linework. *Photogrammetric Engineering & Remote Sensing*, Vol. 53, No. 10, pp. 1383–1387.

Hintz, R.J., and C.J. Rodine, 1990. Automation and Precision in a Cadastral Surveying Environment. *Technical Papers, 1990 ASCM-ASPRS Annual Convention*, Denver, Colorado, Vol. 1, pp. 124–133.

Itami, R.M., 1988. Cellular Automations as a Framework for Dynamic Simulations in Geographic Information Systems. *GIS/LIS '88*, San Antonio, Texas, Vol. 2, pp. 590–597.

Johnson, J.A., 1990. A Standard Method for Creating Self-defining Data Structures for Information Archive and Transfer. *Proceedings, ISPRS Commission II Symposium*, Dresden, GDR, pp. 82–88.

Konecny, G., and D. Pape, 1981. Correlation Techniques and Devices. *Photogrammetric Engineering & Remote Sensing*, Vol. 47, No. 3, pp. 323–333.

Lillesand, T.M., and R.W. Kiefer, 1987. *Remote Sensing and Image Interpretation*. John Wiley & Sons, Canada.

Luhmann, T., 1990. Aufnahmesysteme für die Nahberichsphotogrammetrie. *Festschrift Gottfried Konecny zum 60. Geburtstag*, Institut für Photogrammetrie und Ingenieurvermessung, Universität Hannover, FRG, pp. 209–220.

Marble, D.F., D.J. Peuquet, A.R. Boyle, N. Bryant, H.W. Calkins, and T. Johnson, 1983. Geographic Information Systems and Remote

Sensing. *Manual of Remote Sensing*. (R.N. Colwell, editors). American Society of Photogrammetry, Falls Church, Virginia.

Mather, P.M., 1987. *Computer Processing of Remotely-Sensed Images*. John Wiley & Sons.

Moellering, H. (ed.), 1991. Spatial Database Exchange Standards: Current International Status (in press)

Moik, J.G., 1980. *Digital Processing of Remotely Sensed Images*. National Aeronautics and Space Administration, NASA SP-431. 330 p.

Nellis, M.D., K. Lulla, and J. Jensen, 1990. Interfacing Geographic Information Systems and Remote Sensing for Rural Land-Use Analysis. *Photogrammetric Engineering & Remote Sensing*, Vol. 56, No. 3, pp. 329–331.

Pratt, W.K., 1978. *Digital Image Processing*. John Wiley & Sons, New York.

Raper, J. (ed.), 1989. *Three Dimensional Applications in Geographical Information Systems*. Taylor & Francis, London, New York, Philadelphia.

Sabins, F.F., 1987. *Remote Sensing, Principles and Interpretation*. W.H. Freeman and Company, New York.

Samet, H., 1989. *The Design and Analysis of Spatial Data Structures*. Addison-Wesley Publishign Company, Inc.

Sijmons, K., 1986. *Computer-Assisted Detection of Linear Structures from Digital Remote Sensing Data for Cartographic Purposes*. Inaugural Dissertation, Freie Universität Berlin.

Star, J.L., M.A. Friedl, D.M. Stoms, and J.E. Estes, 1987. Browse Capability for Spatial Databases, *Proc. GIS '87 Conference*, San Francisco, Calif., 26–29 October, pp. 196–205.

Star, J., and J. Estes, 1990. *Geographic Information Systems: An Introduction*. Prentice Hall, Englewood Cliffs, New Jersey.

Strahler, A.H., C.E. Woodcock, and J.A. Smith, 1986. On the Nature of Models in Remote Sensing. *Remote Sensing of Environment*, Vol. 20, pp. 121–139.

Swain, P.H., and S.M. Davis, 1978. *Remote Sensing: The Quantitative Approach*, New York, McGraw-Hill, pp. 137–187.