GEOGRAPHIC INFORMATION SYSTEMS AND REMOTE SENSING FUTURE COMPUTING ENVIRONMENT - AN UPDATE

ABSTRACT

Recent changes in computing technology are causing dramatic effects in the manner in which image processing and GIS functions can be applied to real world problems. The power of RISC workstations allows analysis and integration functions to be performed that were heretofore impractical because of cost and time constraints. Network capabilities and data compression techniques have matured to the point that long range collaboration across networks is becoming viable.

This paper is an update to a survey article published in 1991 on the future computing environment and its effects on remote sensing and GIS integration and general acceptance.

I. Introduction

In June of 1991, the National Center for Geographic Information Analysis (NCGIA), a funded National Science Foundation Center, coordinated a special issue of Photogrammetric Engineering and Remote Sensing (PERS) on the integration of Remote Sensing (RS) and Geographic Information Systems (GIS). NCGIA has been tasked with developing a research agenda for a number of different problems associated with GIS technology and applications. Each research area is represented by a research initiative designed to develop a prioritized set of potential research topics that focus on the impediments to rapid acceptance of GIS technology into the research, applications, and operational environments that exist worldwide.

NCGIA Initiative 12 - Integration of Remote Sensing and GIS has been undertaken during the years 1990 and 1991. Initiative 12 involved NCGIA personnel along with a number of specialists in the fields of Remote Sensing and GIS. Preliminary meetings were held to define the problems that could affect GIS and Remote Sensing acceptance. An initial list of research topics were developed and presented at a specialists meeting in Sioux Falls, South Dakota in late 1990 to gather peer feedback. The results of the specialists meeting was to solidify a research agenda with prioritized needs. The research agenda was then developed into a set of subtopics that were described in the PERS special issue.

The overall topic of RS/GIS integration is broken down into several discrete directions that were reported on in this same volume (Estis, et al, 1991). Lauer (Lauer, et al, 1991) reported on the impediments to the institutionalization of GIS and Remote Sensing Integration Technology. Data Integrity and Error Sources were investigated by another panel, (Lunetta, et al, 1991). Data Structures and Data Access and their affects on acceptance of GIS/Remote Sensing integration were studied by Ehlers, et al. (Ehlers, et al, 1991). Processing Flow limitations and impediments are discussed by another panel, (Davis, et al, 1991). The impact of the future computing environment is also discussed in this issue (Faust, et al, 1991).

This report is an update to the 1991 article on the future computing environment. A summary of some of the information presented in the 1991 article is presented here, along with an evaluation of the recent trends in the computing environment.

The increase in capability of workstations and their networking capabilities in the past year along with reductions in price make the workstation an optimum tool for RS/GIS analysis. Heterogeneous networks with PC's, a number of different UNIX workstations, and perhaps connection to supercomputing capabilities allows a variety of RS/GIS research, development , and applications to be performed utilizing distributed or central databases. This concept of a loosely connected multi-processing environment has the potential to achieve significant increases in processing capability within a constrained budget.

This report will define the basic functions needed for RS/GIS research and applications, discuss the alternatives in computing to address those problems, and look at future considerations that should be watched to achieve maximum acceptance of the technology in our everyday lives.

II. Identification of functions provided in GIS/RS integration

Raster GIS

GIS systems in today's market may be conveniently be broken down into two types that are indicative of their inherent data structure. Raster GIS systems have multiple variables which normally are coded into cell or raster grids. Attributes for variables are normally coded to a grid cell value with a lookup to extended attributes. While data may be captured in either a raster or vector manner, data analysis occurs in the raster domain. Examples of such raster systems include ERDAS by ERDAS, SPANS by TYDAC, GRASS by the U.S. Army Corps of Engineers, and MAPS by Dana Tomlin.

Basic image processing and analysis tools such as geometric correction and pattern recognition are expected to be available to preprocess Remote Sensing information into a form in which attributes are assigned to raster data instead of raw data values from spectral bands of multispectral imagery.

Once the Remote Sensing information is in a form that can be analyzed as a part of a GIS, systems with both data representations apply a number of the same functions, even though they are implemented in a significantly different manner. We will consider each of the basic functions of a GIS and discuss the differences in the implementation for vector and raster systems (Davis, 1990).

Information Display

Display of GIS/RS information is , of course, the most visible part of a GIS/RS system. The information must be provided to a user in a concise , and easy to understand manner so that he may be able to comprehend the potential complex relationships that may exist between multiple spatial variables. The techniques for GIS/RS display will tend towards using all of the capability available in a graphic display system.

Display of raw and processed Remote Sensing data normally occurs in a 'true color' mode in which raster images for three different variables (in this case, 3 spectral bands) may be assigned independently to the individual color guns of a display unit.

Raster GIS data variables may also be displayed in a true color mode, with the visual integration of the three primary colors of the display giving dynamic information showing a three layer analysis.

Normally, however, GIS raster variables are displayed in a pseudo color mode in which each GIS data value may be assigned a color value that is taken from either a set color palette or a user selectable combination of grey scales on the red, green, and blue primary color guns.

Attribute Handling

The simplest case of handling identifiers for raster GIS variables is a code that identifies the value of particular cell in the raster data set with a character string. Attributes for vector data are more complex in that single entities such as vectors may have multiple codes depending on directionality or its relationship to other neighboring entities. This coding is generally developed at the time of data entry via a text manipulation program and is usually carried along with the data variable in the same or an auxiliary file.

When analysis is performed on a number of GIS variables, the resulting attributes of the analysis file should indicate which variables have been combined, and which functions were applied to perform this combination. The resulting attributes may then be a simple concatenation of the input attributes, or it may be a user defined text attribute that describes the process. For example, the combination of soil, vegetation, slope, and land ownership may result in an analysis called 'suitability for siting of land fill areas' or it may contain a history of all processed that were applied without a meaningful description of the final result.

New techniques developed in machine vision and artificial intelligence applications are being investigated for storage and manipulation of GIS entities (points, lines, polygons, attributes, etc.) as objects in an object oriented data base management system with an inheritance hierarchy (Ehlers, 1991). It is yet unclear how this procedure which requires identification of features that are to be treated as objects would interact with the raw data variables such as remote sensing imagery which provide continuous grey tone data values.

Vertical Analysis

In a raster based GIS/RS system, a vertical analysis is implemented easily. A vertical analysis allows the user to combine numerical attributes of a number of independent GIS/RS variables into a resulting analysis variable. Each input GIS/RS variable is preprocessed with user input to create a numeric value that represents each raster cell's relative importance to others within its own GIS layer for the desired analysis.

Once the numeric scaling has occurred for each input GIS/RS variable, a weighted sum of the variable's values for each raster cell is calculated and finally scaled by the user for final output. As mentioned above, the attributes are normally either lost with the only attribute being defined by the user for the analysis output, or they are concatenated to form a more complex textural description of the combination process. In some cases, both final attribute user definition and concatenation occurs.

The vertical analysis above, may be described as a 'point' process in which each raster cell within a GIS/RS database is considered without respect to adjacent values with the same variable or other variables. Image Processing techniques utilize similar 'point' functions. A 'point' function allows the value of each raster cell in one variable to be processed with a mathematical function with respect to the same raster cell in another GIS/RS variable. The mathematical operations may include :'+', '-' , '/' , '*' as well as any other legitimate function such as sin or cosine, logarithms, or exponentials. In the case discussed above, the operator was '+'.

Vertical raster functions are functionally identical to normal image processing functions used on Remote Sensing data sets. Ratios and linear combinations (weighings) of multispectral bands are often used functions for enhancement of multispectral imagery.

Proximity Analysis

Proximity analysis is one of the more powerful tools of GIS/RS systems. A proximity variable is one whose values represent the spatial distance of any point from a data value specified as a search criteria. For example, in a land cover classification, a proximity analysis may be required to show the distance away from all water classes. The resulting GIS file will not have anything to do with the coincidence of a test pixel with any other land cover class, but it will include the euclidian distance of any point from the specified criteria, in this case, the existence of water in the land cover classification.

This technique is extremely useful in assessing the relative spatial importance of a specified search class with respect to user specified criteria for an analysis. For example, in a siting problem, nearness to a water body may be an important factor. The best circumstance might be those areas directly adjacent to a water body, but the importance of a site might be only marginally less if it is less than 100 meters from water. It also might be acceptable if it is within 200 meters, since site modifications might be easily made to provide direct water access.

Proximity to certain criteria is often used as a weighing criteria in a subsequent vertical analysis.

Neighborhood Operations

Neighborhood operations are those functions in raster analysis in which the resulting value or class for a cell at position (x,y) depends not only on cells at the same (x,y) location from different variables, but also on the cell values within a certain distance of the (x,y) location in all data variables. This neighborhood in raster analysis may be symmetric (a 3 x 3 window of cells surrounding the (x,y) cell , a 7 x 7 window, etc.) or it may be asymmetric (a 3 x 10 cell area). A neighborhood operation may be as simple as computing the sum of all values within the designated box, or it may be as complex as the dynamic computation of the mean, mode, median, maximum, and minimum of the box and zeroing values less than zero and greater than 10. The neighborhood function may also be performed on logical variables to indicate presence or absence of a specific criteria.

A neighborhood function may create more than one output variable that represent different functions applied to the window. For example, slope and aspect may be computed using the same window on digital elevation data. Some common neighborhood functions that are applied to GIS/RS data sets are: average value, minimum value, maximum value, mean value, median value, mode value, AND, OR, NOT logical functions, slope, aspect, coincidence, absence, and diversity.

Neighborhood operations are also common in image processing. While the functions are approximately the same, and the implementations are equivalent, the names of the functions often differ. For example, texture analysis of images uses neighborhood functions to define first, second, and third order moments that can be thought of as higher derivatives of an image surface within the specified window.

Raster-Vector Conversion

Raster to vector conversion normally requires a number of user intensive steps to define polygons that will fit into a topologically structured vector network. Initially, clustering or clumping GIS algorithms may be used to define contiguous regions with the same GIS/RS value. The boundaries of the regions may be formed by segmentation algorithms that normally operate on the detection and chaining of edges, or the growing of regions within the raster data set. The perimeter of the regions thus constitute the initial 'polygon' region. Algorithms for smoothing to remove the jagged pixel edges and thinning , to remove short vectors along a consistent path, must be run iteratively until the final vector file is acceptable. Next, the assignment of attributes to vectors, points, nodes, etc. must be assigned in at best a semi-automated manner. The process from raster to vector representations is normally an extremely difficult process. No current totally automated manner is known by the author that can perform without exhaustive preprocessing of the raster input data sets.

Vector GIS Functions

Vector systems, unlike raster systems, capture data in vector format, analyze data in vector format, and produce output in either vector of raster format. Generally, the data are either stored as whole polygons, lines, and points, or in a topologic structure of arcs and nodes . Attributes are coded to each basic representation and are generally stored in a relational database. Examples of current vector based GIS systems include ARC/INFO by ESRI, GENESIS by and TIGRIS by Integraph.

Vector Display

Display of vector GIS information can be accommodated on a relatively simple display system. The display must be capable of a limited number of colors, and software must have the capability of converting the vectors into a raster representation for display on a bit mapped display. Stroke vector storage displays are no longer desirable even though they can portray more points or pixels per inch than raster bit mapped displays.

Attribute Handling

Attributes in a vector system may be stored either in a linked list form, which again must be managed by the application program, or they may be stored in a RDBMS as discussed above. A typical inquiry of such a system might be: find me all areas with at least 60 % pine forest, within Jasper County, containing a stream, with a size of 10 acres or more, and a cost of less than $2000 per acre. Polygonal areas may be defined with multiple attributes, such as the percentage pine cover, ownership, cost per acre, soil type, etc. These attributes may be searched along with multiple attributes from other GIS variables to find candidate areas. The solution to such a request requires several polygon overlays as well as a search into a relational data base management system for attributes satisfying the request.

Vertical Analysis

Vector GIS/RS operations for vertical analysis are substantially more complex than raster vertical operations. To perform vertical operations on vector data, a set of algorithms for point, line, and polygon overlay must be executed. Logical functions, such a union and intersection of polygons , point inside polygon, vector intersection, and clipping must be performed on all data entities within each of the variables with respect to all entities in the second data variable. Clipping and windowing functions hopefully reduce the complexity of the analysis, but normally ,substantial portions of each data variable must be processed. Most of these functions involve floating point logic, and are thus extremely time consuming. For analyses in which two simple GIS/RS variables are overlaid, the process may go fairly fast. However, after a number of combinations have been performed, the number of possible intersections for subsequent analyses may become prohibitive with the current implementation of technology. For simple variables, the storage of the GIS/RS variable in vector format provides a large compression in the amount of memory and disk space. For the storage of complex variables and the results of complex analyses, however, the vector file will often exceed the storage necessary for storage of the file in raster form.

Proximity Analysis

Vector proximity analysis usually entails the definition of a buffer zone area defined by lines and/or polygon features showing areas a certain distance away from a specified criteria. The buffer zone may be used as a mask in a polygon overlay of vector data.

Neighborhood Operations

Since most polygon overlay operations are boolian in nature, there does not seem to be an equivalent function to neighborhood processing in raster GIS and Remote Sensing images. As mentioned above, however, proximity masks may be used to extract regions within a specified distance of a selected value.

If data are stored in an arc/node representation, then directionality and right and left attributes must be dynamically interpreted to form polygonal areas, and access their attributes. Access to the attributes may require a search command with the supporting RDBMS.

Vector-Raster Conversion

Vector to raster conversion is traditionally the easier of the two conversions between data representations. Normally, vector to raster conversion occurs by simply performing polygon fill logic on individual polygons, lines, and points within a database and writing the raster version of a polygon into a raster image buffer. If all polygons and all of the raster image buffer fit into memory, then the task is efficient. If, on the other hand, enough memory cannot be allocated, the procedure must operate on a polygon by polygon basis and an image block basis. Virtual memory normally will eliminate the need for the user to have to manage this partitioning, but disk space accesses will invariably cause longer run times.

One of the principal problems with vector to raster conversion is deciding what to do with the multiple attributes that may have been assigned to the polygons in the analysis or data entry phases.

The integration of Remote Sensing information into GIS occurs naturally in a raster GIS since both data structures are approximately the same. Integration into a vector system requires somewhat more effort, but it has been recently achieved by several GIS and Remote Sensing vendors, at least to the extent of updating vector information by using an image as a backdrop for vector editing (Sperry, 1989,ERDAS, 1989)

III. Computing Technology

Advances in computing technology have been so rapid in the last decade, that it is difficult to evaluate the many alternatives offered in today's and tomorrow's market for high speed interactive computing. The above functions describe the algorithms and techniques which are at work in the current generation of GIS/RS systems. The next decade of computing promises much more in the individual power of computers and the synergistic capability of distributed networks. Jack Dangermond, President of ESRI predicted that for geographic analysis in the next 10 years, CPU's capable of processing 1000 MIPS will be common in large organizations and greater processing power will be packed into compact, less expensive systems. (Dangermond,1988). It is possible that with the emerging technologies in the computing, user interface, and visualization areas, the concept of how GIS/RS analyses are performed will be totally redefined.

Single Processor Computation Power

One of the most dramatic ares of change in the overall computing environment that affects GIS/RS systems is in the raw computation power of computing systems. Initially we will look at changes that have occurred in single Central Processing Unit (CPU) architectures, and later we will expand to consider multi-processing and special purpose architectures.

In an about face from the trend of having longer word lengths to give more power and speed, the Reduced Instruction Set Computer (RISC) was introduced in 1986. The RISC computers achieve high speeds by optimizing the implementation of only a few instructions of the more complex systems. A 32 bit wordlength was used, then, to provide faster data access, and to combine multiple instruction executions in a clock cycle. The state of the art in speed for single processor systems is the MIPS R2000, R3000, R4000, and R6000 RISC chips used in Silicon Graphics and STARDENT products as well as the SUN SPARC RISC chip. Depending on the clock speed for these chips, these RISC processors show a performance of between 8 and 30 MIPS. Most of the workstations in today's market use these two chips as their basis. IBM, on the other hand has recently developed its own RISC chip with very attractive speed figures of between 30 and 50 MIPS. Hewlett Packard recently announced workstations with 50-70 MIPS. Other vendors are also in the process of releasing similar speed systems.

New generation serial processors include the INTEL dedicated graphics processor called the 82786 and its followon general purpose i860. A graphics processor seeks to offload are graphics processing from the host into the high speed special purpose processor. The i860 product boasts 33 MIPS and 66 MFLOPS and is today being imbedded into many imaging and graphics products (Keller, 1990), (Wilson, 1990). Because of the speed of the new generation cpu chips, new memory technology was needed to prevent memory bottlenecks. The Video RAM chips allow direct refresh of video screens from solid state memory, and take the load of memory management from the cpu.

Vector Processing

Vector processing has been the basis for most "supercomputer" systems in the last decade (Cheng, 1989). Scientific computing has a need for fast computation using floating point or greater precision. A vector processor is generally a single controlling processor which sequences a long data vector through a number of 'pipelined' stages. A pipeline operation, as the name suggests, is the process by which a complex operation is broken into a number of independent sub steps that can be implemented sequentially. Each step in a pipeline operation may be handled by a dedicated processing element (an adder, multiplier, etc) and the results passed as input directly to the next processing element.

Vector pipelining assumes that the same operation is being applied to a large amount of data. The control and execution of the next instruction is sequential in nature, but the CPU must know whether the last batch of data has passed through the total pipeline. A timing interrupt or message passing strategy must provide this information.

Current supercomputer vector architectures have greatly expanded the power of floating point computation, with the Cray X-MP having a performance measure of 235 MFLOPS per processor with up to 4 processors and the Cray Y=-MP having a performance measure of 333 MFLOPS per processor with a maximum of 8 processors for a performance of 2.8 GIGAFLOPS.

Parallel Processing

When multiple cpu's or vector processors are linked together, the major differentiation between systems relates to the method of synchronization between the various processors and their memory. For a synchronous system, all operations are coordinated through a timing clock. The vector processing architectures shown above depend explicitly on timing to send the input data stream to multiple processors and various subparts of a pipeline. Multiple processors may operate through local memories or a global memory that is shared by all processors.

Dasgupta has developed a taxonomy for computing which represents serial and parallel processing alternatives (Dasgupta, 1990).

SIMD

One kind of synchronous parallel processing environment is the Single Instruction , Multiple Data (SIMD) systems (Duncan, 1990). For this type system, multiple processors are required to execute the same instructions on multiple data streams. Image processing systems have been designed to take advantage of this architecture, and the same kinds of SIMD systems may be applied to most GIS/RS data sets. Synchronous timing is used to move data in and out of a SIMD system and to move data within the system.

MIMD

Multiple Instruction, Multiple Data (MIMD) computer systems are the classic case that most of us think of when parallel computing is discussed. MIMD systems have multiple processors which may operate on different instructions and differing data. MIMD machines do not have synchronous timing with the same instruction being performed; therefore, a sophisticated intercommunication scheme is necessary to tell each processor when to execute its instruction and on which data set to operate. Synchronous processing allows each processor to perform a number of different operations on its own local data without concern of the neighboring processors. Each processor acts alone, but when it finishes its process, it must notify the other processors.

In some cases MIMD may be combined with SIMD processors to form a hybrid parallel processing scheme. One of the implementations involves a tree structure topology in which higher level MIMD processors send messages to SIMD subservient processors which then operate on multiple data in a synchronous mode.

The costs of high speed computing closely coupled with a quality graphics system for visualization has dropped dramatically in the late 1980's and early 1990's with the advent of the mini-supercomputer class of machine which is expected to meet the computing and visualization needs of a project scientist with a deskside or desktop system (Diede, 1988).

Parallel Processing Software

One of the greatest challenges in parallel computing is the development of the software that will allow full use of the hardware capabilities of the new hardware systems (Prasanna-Kumar, 1989). Most of the new systems are trying to avoid the development of special languages for implementation of applications code and instead rely on optimization of Fortran an C code. SIMD systems have been developed with reasonably efficient parallelizing compilers since the synchronization between processors is a vital part of the architecture (Little, 1988). MIMD systems, on the other hand, often have to have special tailoring of the application program to achieve speedups. For example, in MIMD programs, the software developer must identify variables and sections of the code that must be kept in global memory with sophisticated access lockout protection. Other portions of the code may have variables in local memory that do not affect the other processors.

IV. Visualization

In the last several years, the computing power available in a workstation environment have made it possible for detailed analysis and incremental simulation to be used at a scientists' desk to help solve nonlinear differential equations or massively parallel problems that operate on huge amounts of geographic or image based data. Systems are now becoming available that fall into a new class of personal supercomputing. Superior graphics on some of these systems allow for instantaneous viewing of time series or multidimensional data (Weigner, 1990). Software for these systems as well as the current generation of desktop workstations is now being developed to harness the power of the hardware for aids in interpretation of multidimensional data sets.

The integration of remote sensing into GIS systems is an application that may be suitably shown in three dimensional perspective (Long, 1989). Landscape planning is one application which can benefit from dynamic interaction of a 3 dimensional GIS along with CAD models and the capability to instantaneously render a new view of a database showing the effects of potential changes in the GIS information (Robertson, B., 1990).

The huge amounts of data that will need to be analyzed to allow understanding of the spatial and temporal dynamics of the global environment make visualization techniques not only useful but absolutely necessary (Hibbard, 1989). New remote sensing instrumentation is being developed that will dwarf the data volumes of all previous remote sensing systems. These data sets must be integrated into multidimensional GIS systems which provide expanded capabilities for time and space modeling.

A number of vendors of single and multiprocessor computing systems have developed "Visualization environments" that attempt to provide sophisticated display of scientific data without detailed programming knowledge by the potential user (Bishop, 1990), (Myers, 1990).

V. Factors to Consider

A number of factors should be considered in estimating the effects of the dramatic changes occurring in computing on the acceptance of RS/GIS technology. The effects of escalating workstation cpu power, affordable implementations of parallel computing, advances in data storage technology, improvement in networking capabilities, and increases in software usability all affect the way in which RS/GIS systems are perceived by the potential user of such systems.

CPU Power/Cost

Workstation cpu's have greatly increased in power during the last year while costs for workstations have dropped dramatically. Entry level UNIX systems include those from DEC, HP,IBM, and SUN with a low of a diskless DEC Station 5000/ model 20 at $3995 (Buxbaum, 1992) and 16.3 SPECMARKS and Sun SPARCstation ELC at $3995 with SPECMARKS of 20.3. The HP Apollo Series 705 price is $4995 with 34 SPECMARKS. The Silicon Graphics IRIS Indigo starts at $7995 with 26 SPECMARKS and 8 bit color. IBM announced the RISC System/6000 Powerstation 220 with 25.9 SPECMARKS which comes in at $6345.

Medium cost single processor configurations for each of the vendors reach 19.1 SPECMARKS for a DEC 5000 Model 25, 49.7 SPECMARKS for an HP Model 710, 26 SPECMARKS with 1 million 3D vectors/sec rendering for the Silicon Graphics Indigo, 43.4 SPECMARCS for the IBM RISC 6000 Model 320, and 24.4 SPECMARKS for the SUN SPARCstation IPX.

The maximum SPECMARK performance was given by IBM RISC System/6000 Model 560 which came in at 89.3 SPECMARKS.

In addition to single cpu configurations, several vendors now offer multiple cpu's with multiprocessing capability. The Silicon Graphics PowerVision series has up to 8 cpu's working in a shared memory multiprocessing environment. Sun is delivering a SPARCserver 690 MP with 4 processors in a multiprocessing environment.

While these multiprocessor systems can increase the capability of a server and possibly a workstation, the operating systems that support the multiple processors only support multiprocessing, not true parallel operation.

Wavetracer is offering a massively parallel computer as an adjunct to a workstation that implements SIMD parallelism. The processors for the Wavetracer system can be dynamically configured into a 2-D or 3-D mesh for image processing or volumetric processing. Pricing depends on the number of processors with a low end cost of $85,000 and a$350,000 cost for 16,000 processors. Each processor, however, is not as powerful as the RISC processors above.

Storage

Dramatic changes also have occurred in storage technology. Choices for on line storage may be magnetic disk, optical READ/WRITE disk, or optical WRITE ONCE disk. Off line storage includes magnetic tapes in addition to the above. Prices for disk storage have dropped radically recently to the point where 1 GIGABYTE of magnetic disk can be purchased for 2-3 thousand dollars with a SCSI interface. Now it is possible to keep a large GIS or Remote Sensing database on line without continual backups and restores to get temporary disk space.

Several companies are now offering optical disk "jukebox" systems that mechanically load disks from an archival disk holder and place them on line when information is requested from the archived disk. Epoch Systems has a total storage management system that uses all three types of disk storage in a hierarchical manner. A catalog is kept which has the location of each file, no matter where it is in the system. Files that are continually accessed are kept on the magnetic disk, files that are less often accessed are kept on the READ/WRITE optical disk, and other files are placed on the WRITE ONCE optical disk. Files that are seldom accessed may be written to magnetic tape. When a file is requested, the software locates which medium the file is on and initiates retrieval by staging. The software for operating the storage function is a sophisticated database management system.

Networks

Possibly some of the greatest benefits to the Remote Sensing/GIS community may come from networking and the technologies associated with it. Public networks such as INTERNET, SURANET, BITNET, and PEACHNET (in Georgia) are available to most RS/GIS users through government or university gateways. National and international access is allowed with BITNET. The ability exists now to log onto a computer at Georgia Tech via INTERNET from UCSB in California , run a GIS visualization analysis, and display the results through a remote X window in California. This capability , though not instantaneous, is fast enough to entertain thoughts of cross country collaboration. Higher speed networks are being implemented including a proposed 1 Mbit/sec NSFNET. With these transmission speeds , a realistic collaboration could occur nationally and internationally.

Along with networks themselves are innovative technologies and applications that make up Multimedia. Multimedia is the ability to use all forms of communication , images, voice, music, text, and graphics to convey a message. Some of the techniques needed for Multimedia have driven research in communications for the last few years. Data/image compression can be used positively to reduce storage for single or sequences of images. Different compression techniques will be applied to motion sequences than will be applied for still frames. The JPEG (Joint Photographic Experts Group) algorithm is one of the compression techniques that is often applied to still imagery, while MPEG (Motion Picture Experts Group) algorithms are often used for live video or motion sequences. Video conferencing is one of the prime motivators for motion compression.

A network environment might soon exist for a group of scientists to be synergistically working on a RS/GIS analysis with individual video windows showing live video of each of the participants and another window showing an image or a database that all participants can operate on. Pointers across the network may be used to indicate areas of interest.

Networks also bring up another issue that could easily become the most important step in the continual transformation of the RS/GIS industry. A number of entities have begun investigating the potentials for shared and distributed spatial databases. If a number of issues relating to data ownership, data integrity, and data update procedures can be worked out, it is possible that we might begin building a national or global data set while eliminating duplicate data collection and analysis. If we begin to gather GIS and RS data in an unbiased manner with sufficient information to show the lineage of the database and the restraints necessary for database use, we might be able to achieve the goal of building and updating databases at the local level, and aggregating the information up for smaller scales for state and national use. The Spatial Data Transfer Standards that are currently being reviewed are an attempt to standardize the way in which we transfer data to one another.

Usability

Over the past year most RS/GIS systems have gravitated towards a somewhat consistent user interface. An X windows/motif user interface has scroll lists, buttons, sliders, and other widgets that allow a developer of a system to easily configure his user interface and tailor it to his customer's needs. The user interface is used to gather information from a user that will be passed into the application program that he selects. The information gathering aspect of a particular program is thus separated from the program itself. With the required information passed by the interface, the program may then proceed to execute the algorithm on the data files which were also selected.

With the interfaces looking similar, the algorithmic capabilities of the vendors will become much more important in the evaluation of RS/GIS systems in the future.

Many of the tools that are extremely useful in image processing of Remote Sensing and GIS data sets are being developed in parallel by the Desktop Publishing(DP) industry. Since the volume of DP systems that will be purchased is much greater than the volume of RS/GIS systems, the amount of investment in software development will likely be much greater than in our field. It would benefit the RS/GIS community to be aware of the techniques that are being used in DP and look for their applications in RS/GIS.

The Future

Some of the techniques that have been developed for the entertainment industry may also be applied to RS/GIS analysis. Three dimensional visualization is mentioned above as a technique that can be very beneficial in the analysis of spatial data. A technique known as Virtual Reality (VR) is being developed in computer graphics that allows one to immerse oneself within a three dimensional environment. Since the 3-D environment is spatial in nature, it makes sense that the RS/GIS community would be a major participant in this technology. One can imagine the RS/GIS system of the future in which the user can dynamically find himself within the multilayer database and move to any point within the data while visualizing what he could see from any position. Three dimensional perspective images are a natural way for a human to perceive the world around him. He should be able to interact with the database from above, as we are doing today, or from any position within or without the database. At any point he should be able to point to a 3 or 2 D feature and query it for its attributes. The RS/GIS system of the future would remove the artificial constraints on the ability of humans to perform spatial analyses.

References

1) Cheng, H., "Vector Pipelining, Chaining, and Speed on the IBM 3090 and Cray X-MP", COMPUTER, pg 31, Sep 1989.

2) Dangermond, J., "GIS Hardware Trends", COMPUTER GRAPHICS WORLD, Dec 1988.

3) Dasgupta, S., "A Hierarchical Taxonomic System for Computer Architectures", COMPUTER, pg 64, Mar 1990.

4) Davis, F. , Quattrochi, D. , Ridd , M. , S-N Lam, N., "Research Needs in Processing and Analysis of Remote Sensing and GIS Data", NCGIA Initiative 12 Report, Jan 1991.

5) Diede, T., Hagenmaier, C., Miranker, G., Rubinstein, J., and Worley, W., "The Titan Graphics Supercomputer Architecture", COMPUTER, pg 13, Vol 21 , No. 9 , Sep 1988.

6) Duncan, R., "A Survey of Parallel Computer Architectures", IEEE, pg 5, February 1990.

7) Duncan, R., "Choosing Parallel Architectures", Military and Aerospace Electronics, pg 19, Aug 1990.

8) Ehlers, M. , Greenlee, D., Smith , T., Star, J., "Data Structures and Access", NCGIA Initiative 12 Report, Jan 1991.

9) Estis, J., Star, J., Davis, F., Maquire, D., "Overview of Research Issues in the Integration of Remote Sensing and GIS", NCGIA Initiative 12 Report, Jan 1991.

10) Faust, N., Star,J., Anderson,W., "Geographic Information Systems and Remote Sensing Future Computing Environment", Photogrammetric Engineering and Remote Sensing, Vol 57, No. 6,

10) Hibbard, W., and Santek, D.,"Visualizing Large Data Sets in the Earth Sciences", COMPUTER, pg 53, Aug,1989.

11) Hogan, B., "High Performance Image Processing on a Massively Parallel Computer", ADVANCED IMAGING, pg 42, Oct 1990.

12) Hornstein, V., "Parallel Processing Attacks Real-Time World", MINI-MICRO SYSTEMS, pg 65, Dec 1986.

13) Keller, J., "The Rise of Image Processing", Military and Aerospace Electronics, pg 17, Feb, 1990.

14) Lauer, D., Estis, J., Jenson, J., Greenlee, D., Mace, T., "Institutional Issues Affecting the Integration of Remotely sensed Data and Geographic Information Systems", NCGIA Initiative 12 Report, Jan 1990.

15) Little , J., Blelloch, G., and Cass, T., " Algorithmic Techniques for Computer Vision on a Fine-Grained Parallel Machine", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 11, No. 3, pg 244, Mar 1989.

16) Long, L., "GIS Goes 3D", COMPUTER GRAPHICS WORLD, pg 38, March, 1989.

17) Lunetta, R. , Congalton, R., Fenstermaker, L., Jenson, J., McGwire, K., Tinney, L., "Remote Sensing and Geographic Information System Data Integration: Error Sources and Issues", NCGIA Initiative 12, Jan 1991.

18) Meng, "Parallel Processor Gets Data Intensive", ESD, pg 17, Jun 1987.

19) Myers, B., Giuse, D., Dannenberg, R., Vander Zanden, B., Kosbie, D., Pervin, E., Mickish, A., and Marchal, P., "GARNET : Comprehensive Support for Graphical , Highly Interactive User Interfaces", COMPUTER, Vol 23, No 11, pg 71, Nov, 1990.

20) Prasanna-Kumar, V., and Reisis, D., " Image Computations on Meshes with Multiple Broadcast", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 11, Nov 1989.

21) Robertson, B., "Sculpting the Scenery", COMPUTER GRAPHICS WORLD, pg 48, Jun 1990.

22) Tilley, S., and Sperry, S., " Raster and Vector Integration", COMPUTER GRAPHICS WORLD, Aug 1988.

23) Tucker, L., and Robertson, G., "Architecture and Applications of the Connection Machine", COMPUTER, pg 26, Aug 1988.

24) Wiegner, K., "Data and Vision", Forbes Magazine, pg 193, Oct, 1990.

25) Wilson, R., "New Long-word Architecture Threatens to Outshine RISC", COMPUTER DESIGN, pg 26, Sep, 1990