Geographic Data Processing

GEORGE NAGY AND SHARAD WAGLE*

University of Nebraska-Lincoln

This survey attempts to provide a unified framework for the constituent elements—originating in numerous and diverse disciplines—of geographical data processing systems. External aspects of such systems, as perceived by potential users, are discussed with regard to extent, coordinate system and base map, range of applications, input/output mechanisms, computer configuration, command and interaction, documentation, and administration.The internal aspects, which would concern the system designer, are analyzed in terms of the type of spatial variables involved and of their interrelationship with respect to common operations This point of view is shown to lead to a workable classification of two-dimensional geometric algorithms and data structures. To provide concrete examples, ten representative geographic data processing systems, ranging from automated cartography to interactive decision support, are described. In conclusion, some comparisons are drawn between geographical data processing systems and their conventional business-oriented counterparts.

Keywords and Phrases: geographic information systems, geographic data processing, management information systems, automated cartography, computational geometry, natural resource management, graphic input and output, picture processing, remote sensing.

CR Categories: 3.53, 3.7, 8.2

INTRODUCTION

Geographic data processing requires the compilation, storage, transformation, and display of information traditionally represented in the form of maps. The problems associated with automating these tasks bridge computer science and the fields of application of spatial analysis such as cartography, demography, and natural resource management. Geographic data processing also draws on other two-dimensional computer applications including image processing and pattern recognition, and on recent advances in business data management. The field is beginning to take on an identity of its own, with annual symposia and a journal, but new developments are still presented at conferences and in journals devoted to topics as disparate as computational complexity, remote sensing, database organization, and sociology. A recent inventory lists over 320 information systems and programs for spatial data handling [BRAS77].

We attempt to describe, categorize, and analyze the principal constituents of this new field. In collecting the information presented in this article, we examined many of the geographic data processing systems in current use, and unravelled the common threads linking the different applications and implementations. We hope to present our findings from a uniform perspective which will prove helpful both to new ar-

* Current address Department of Computer Science, Illinois Institute of Technology, Chicago, Illinois 60616

CONTENTS

rivals and to those whose direct experience has been confined to one or two specialized aspects.

This survey is oriented toward the design and development of geographic data processing systems rather than their operation and use. It is assumed that the reader has a working knowledge of digital computers, programming, and data structures, but no more than a layman's acquaintance with the tools and vocabulary of geographical analysis. For the reader with the complementary blend of skills—the professional geographer who wants to learn how computers can contribute to his field—the sophisticated review of Rhind [RHIN77] is highly recommended.

After disposing of some general considerations in the remainder of the Introduction, we explore in Section 2 the specialized input/output devices and techniques required by the essentially two-dimensional nature of geographic problems. Section 3 is devoted to the closely coupled topics of internal data organization and computational algorithms. Section 4 is a review of ten specific systems which we consider representative of recent developments. In the concluding section some analogies and comparisons are drawn between geographic data processing and "ordinary" (administrative) data processing.

1.1 Sources of Additional Information

The classic work on geographic data processing is undoubtedly the two-volume compendium prepared for the International Geophysical Union by Tomlinson [TOML72]. Tomlinson also edited, under the sponsorship of the United Nations Educational, Scientific, and Cultural Organization (UNESCO), two less detailed collections of papers [TOML70, TOML76]. Automated cartographic systems are discussed, and emphasis is placed on British and American contributions, respectively, in DALE74 and in TAYL76. A bibliography containing about one thousand references was prepared by Peucker for the 22nd International Geographical Congress in 1972 [PEUC72]. A short overview without literature references was published recently in *Datamation* by Dutton and Nisen [DUTT78].

Geographic data analysis is treated from the point of view of the social scientist in a monograph [NORD72], and from the point of view of the technician in a collection of symposium papers [DAVI75]. Algorithms are compared in the proceedings of a seminar on geographic data processing sponsored by IBM-UK [ALDR75]. Already mentioned is the excellent review by Rhind, which provides an insight into the economics of mapmaking and the constraints affecting the automation of the cartographic process. Rhind has also written a useful earlier paper on spatial interpolation techniques [RHIN75].

A directory of individuals in the field of geographical data handling was published by the International Geophysical Union in 1977 [IGU77a]. IGU also publishes a directory of available software [IGU77b]. The software survey is part of a continuing project of the IGU, with the working group currently headed by Duane Marble of the Geographic Systems Laboratory of the State University of New York at Buffalo. The organization of this inventory and a summary of the distribution of programs according to criteria such as applications, cost, size, programming language, host machine, required peripheral devices, and sponsor are described in BRAS77. Another valuable IGU report is *Geographic Information Systems, Methods, and Equipment for Land Use Planning* [CALK77]. Unfortunately, the references in this publication are almost without exception outdated.

Current developments were reported in a 1977 symposium organized by the Harvard University Laboratory for Computer Graphics and Spatial Analysis [HARV77]. A follow-up symposium in July 1978 [HARV78] was aimed more at users than at developers of computer mapping software and geographic information systems.

References to work on specific topics are dispersed throughout the remainder of this survey.

1.2 Systems Considerations in Geographic Data Processing

A geographic data processing system may include any or all of the functions shown in Figure 1. Each of the functions requires both hardware and software elements. As emphasized in the figure, specialized graphic input and output components often play a dominant role in and shape the architecture of the remainder of the system.

Some geographic data processing systems consist only of a handful of FORTRAN programs. At a primitive level they require the user to manually record the coordinates and produce the map display on a line printer. These programs are acceptable for

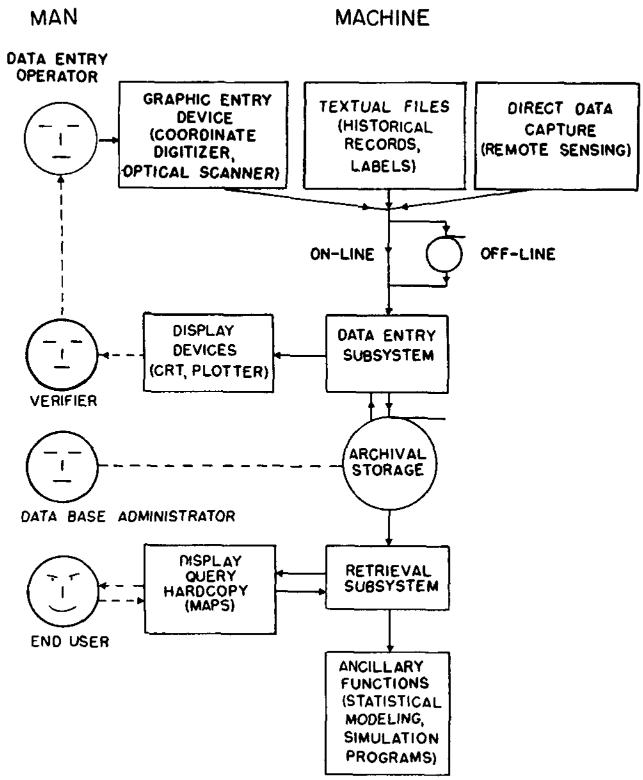


FIGURE 1 Universal GDP system schematic

a given function if the data volume is low and the level of detail does not require finer displays. Their portability is their main attraction, and any processor having sufficient core memory can serve as a host machine. An example of such programs is given in Section 4.1.

Some systems are developed for the environment of an off-line coordinate digitizer and an output plotter which are necessary in any high-volume high-precision application. Use of these systems is contingent on the availability not only of these devices but also of corresponding device-related software support. While standardized support for programming plotters in FORTRAN is available, such has not been the case, until recently, for digitizers. Automated data-entry systems are therefore relatively less portable. Examples of highly specialized systems are discussed in Sections 4.4 and 4.5.

Geographical data entry involves interaction between the user and the computer system. If the interaction is available instantaneously, as in a timesharing or a dedicated system, the total time for data entry can be considerably reduced. Furthermore, if the interaction is interleaved with the data entry, then the correction of errors can be simplified considerably [DUEK77].

Due to the high cost of digitization and display hardware, data transfer requirements, and utilization factors, minicomputers are cost effective for data entry and display tasks. However, they are not well suited for computationally intensive tasks; these are better served by the resources available on large systems (which also provide greater programming convenience). This has resulted in a pattern, which may predominate in the future, of using a minicomputer to perform the data-entry, edit, and display activities (and occasionally data retrieval as well), and using a large computer to perform data manipulation. Currently, the link between the two commonly consists of manual transfer of data via a magnetic tape, but as computer networks evolve, this procedure will be replaced by electronic transfer via a network link. Then it will also be possible to initiate the job on the large system from the minicomputer itself, making the entire operation local from the point of view of the user-operator [BOYL76, MCLE74].

Some systems, especially production-oriented cartographic ones, allow the user very little flexibility to introduce new functions. Some experimental systems, on the other hand, such as GADS (Section 4.9), allow the user to weave his own programs into the command structure, in a manner similar to "user-defined functions" in APL. Command languages may be rigid keyword parameter cards similar to IBM's JCL, or rich, high-level procedure-oriented languages based on FORTRAN or PL/I.

The capabilities of a specialized system can often be augmented at little additional cost by interfacing it with standard statistical and report-generating program packages. In many instances the inexpensive printed output of such packages can also be used to advantage for detecting errors at intermediate processing stages.

1.3 Applications

The intended use of a geographic data processing system is closely related to its *extent*. The extent, or area covered by the system, can be as small as a city. Statewide systems are common. Nationwide systems also exist, and even worldwide systems are being planned. No cosmic system has come to our attention, but stellar atlases are available in computer-readable form.

As noted by Rhind, " . . . the ability to select features and then to treat the world as being an infinitely extensive plane which may be sampled at any point are basic to almost all automated cartographic facilities . . . " [RHIN77]. Practical limits are, of course, imposed by the available storage facilities; up to 100 million bits may be required for the data commonly represented in a single sheet of the 1:24,000 scale US Geological Survey Map [ACSM76].

Some areas of interest have natural bounds; a forest or a river basin may constitute the extent of the corresponding geographic data processing system. Citywide systems are used mainly for urban planning, sociological research, market studies, and municipal administration, while natural resource inventory and planning are typically the principal applications of more extensive systems.

The list of natural and cultural features on and below the surface of the earth which are already encompassed by data processing systems is indeed impressive, yet it is only a beginning [SELI78]. Perhaps the only thread linking the variety of applications is that almost all are financed by governmental agencies.

Figure 2 provides a sampling of current applications. While many of these applications are still currently conducted on an operational basis using conventional cartographic products, the necessity for more frequent updating of the maps, more elaborate spatial analysis, and greater cost effectiveness will increasingly tip the scale in favor of automation.

1.4 Geography, Geometry, and Geodesy

When the extent of the database under consideration is larger than a city or county, the system must allow for the spherical nature of our planet. Although for internal computer processing a three-dimensional model for locating objects in terms of spherical coordinates (latitudes and longitudes) is perfectly adequate, for display and mapping purposes there is no satisfactory alternative to projecting the earth's surface on a two-dimensional plane. The ideal projection would preserve relative distances, angles, and areas, but it can be shown by geometry that no projection can faithfully represent more than two of these three properties. The distortions introduced in the remaining property become more and more significant as the area projected subtends a larger solid angle at the center of the earth [CHAM47].

The most common projections for mapping purposes are the *Lambert Conformal* and the *Transverse Mercator*. The former is based on projecting the earth's surface on a cone centered about the earth's polar axis with its surface secant to the extent under consideration. It is particularly suitable for regions elongated in an east-west direction. The Transverse Mercator projection, on the other hand, is based on a cylinder with its axis perpendicular to the polar axis and its surface tangent to a meridian (longitude) near the center of the projected area. It is generally used for regions with their major dimension in a north-south direction. Both the cone and the cylinder are developed into a flat surface by cutting along a single line segment [RICH72].

The widely used *State Plane Coordinates* consist of a rectilinear grid defined on the plane of projection. The origin is placed outside the extent; hence locations can be specified entirely in terms of pairs of positive integers (Figure 3). In large states, the State Coordinate System is based on more than one projection [CLAI73].

Unfortunately, neither the projected geodetic lines (latitudes and longitudes) nor the State Plane Coordinates are readily perceptible on the surface of the earth. Hence the location of fence lines, roads, and monuments is commonly described by reference to the work of the early land surveyors, expressed in terms of *range* and *township*. Carefully surveyed latitudes and

Subject Area	*Agency*	*Reference*
Topographic cartography	United Kingdom Ordnance Survey	DALE74
Military cartography	Defense Mapping Agency	STRU75
Hydrography	Canadian Hydrographic Service	BOYL74
Geology	Swedish Council for Building Research	NORD72
Agriculture	New York State Office of Planning Coordination	TOML76
Meteorology	United Kingdom Meteorological Office	DALE74
Hydrology	United States Geological Survey	WAGL78
Forest management	USDA Forest Service	RUSS75
Demography	United States Bureau of the Census	SILV77
Highway planning	United Kingdom Durham County Council	ALDR75
City planning	City of San Jose Police Department	CARL74
Regional planning	Canada Land Inventory	TOML76
Remote sensing	National Aeronautical and Space Administration	TEIC78

FIGURE 2 Range of GDP applications. These examples of actual geographic data processing systems created for various government agencies are drawn from the cited references

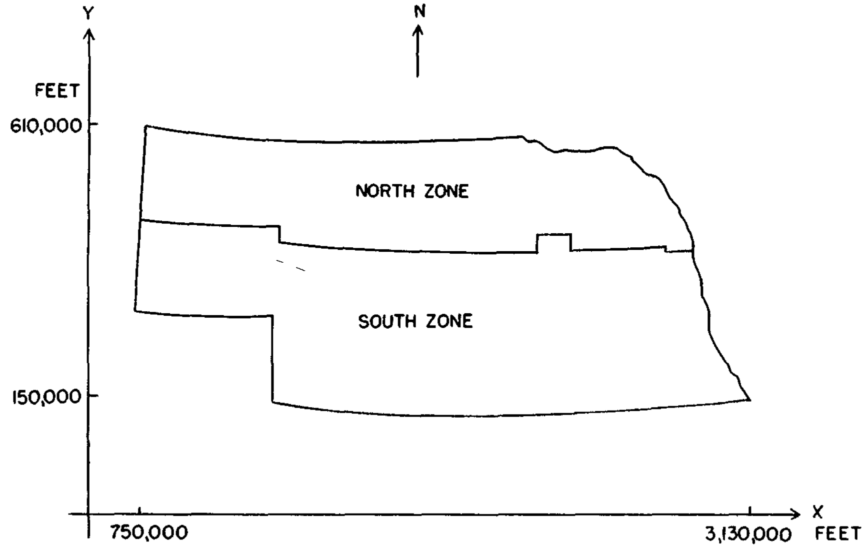


FIGURE 3. State Plane Coordinates. The State Plane Coordinates of Nebraska are based on the Lambert projection. In fact, two projections with a common touching latitude are used, resulting in two zones. The zone-dividing line zigzags to follow county boundaries. For convenience, the entire state is placed in the positive quadrant

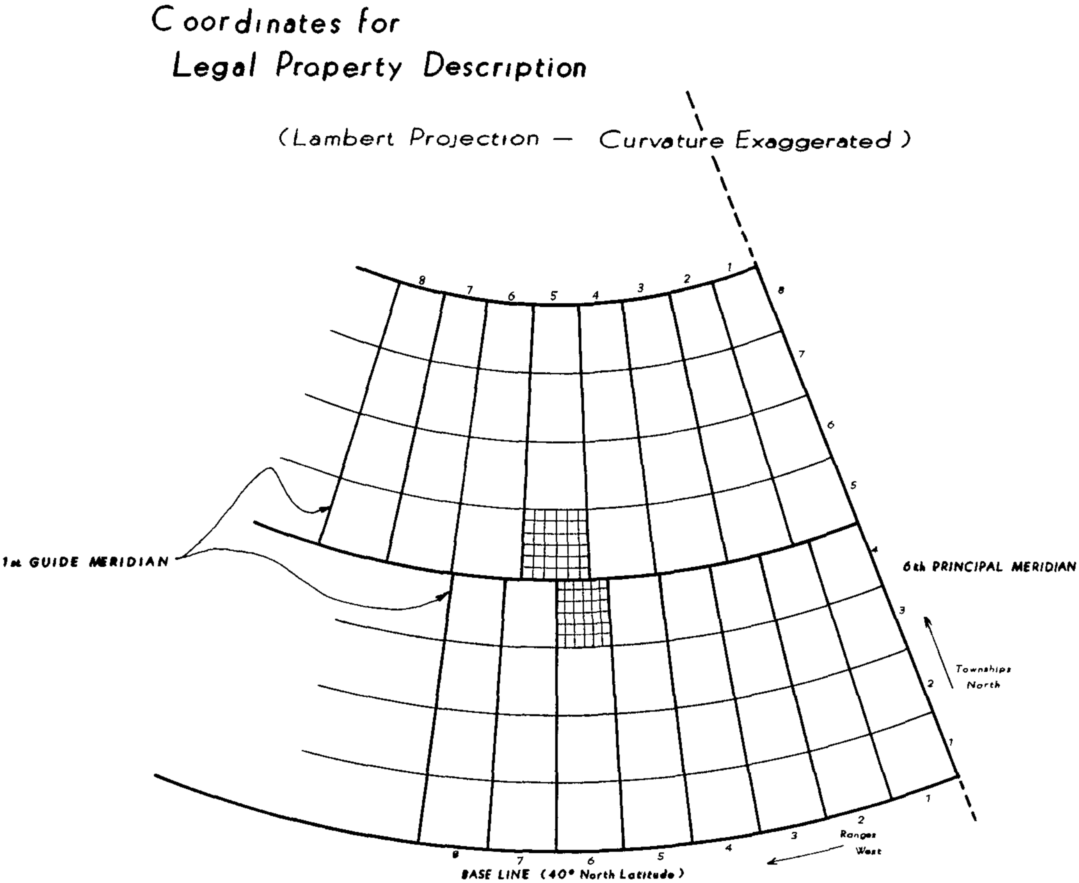


FIGURE 4a. Range and township scheme of geocoding.

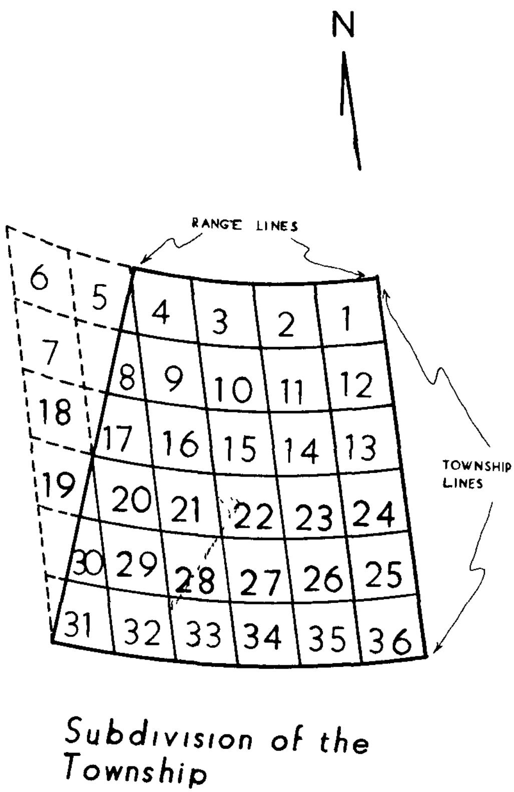


FIGURE 4b.

longitudes did serve as the base line for these surveys, but for more detailed subdivisions the spherical geodetic coordinates were understandably abandoned for a somewhat irregular grid system of six-mile-square townships comprising one-mile-square *sections* subdivided into 40-acre *quarter-quarter sections* (Figure 4b) [BREE42].

The range-and-township grids are irregular because some state boundaries follow natural lines such as rivers, and because the uncompromising convergence of the meridians cannot accommodate square subdivisions. Nevertheless, roads, farms, and even cities are laid out according to this system, and much of the data culled from official records is identified only in terms of it. Furthermore, remotely sensed data from aerial and satellite photographs is often most readily located by reference to roads and highways, which in much of the United States faithfully follow section lines.

Since geographic data processing systems normally combine information from many different sources such as maps, aerial photographs, and tabulated records, they generally include formulas and programs for the interconversion of the various coordinate systems. Coding the geodetic projection formulas is relatively easy; although they generally consist of rather formidable-looking trigonometric expressions, a dozen parameters usually suffice to specify the particular projection used for a given map [TOBL62]. Conversion to and from range and township is far more troublesome; it involves hundreds of arbitrary constants even for a single state, and requires delicate interpolation to dovetail adjoining subdivisions.

In urban areas a grid based on city blocks is often used, but since the extents under consideration are too limited for troublesome manifestations of the earth's curvature, we shall postpone discussing this system until Section 4.3.

Comprehensive studies of place-encoding practices among government agencies can be found in ECU71 and MENE76. Additional information on map projections may be obtained from standard texts such as ROBI69, RICH72, and DEET77.

2. INPUT/OUTPUT

2.1 Sources of Data And Input Devices

The sources of data for geographic data processing may be conveniently divided into three broad categories:

1) alphanumeric information, such as census and crop reports, field notes, and tabulation of historical meteorological data,
2) pictorial or graphic information, such as photographs and maps, and
3) remotely sensed data in digital form, such as that obtained from the LANDSAT and ITOS satellites.

The first category, although important, requires no special consideration here since it is either already available in computer-readable form, or may be converted using conventional alphanumeric data-entry equipment.

Two-Dimensional Data Entry

Two-dimensional data entry is performed most conveniently by means of *digitizers*.

The terms *coordinate digitizer* and *graphic tablet* are usually reserved for semiautomatic devices which allow an operator to trace curves or select individual points on a drawing or photograph. Digitizers may be *on-line,* transmitting the coordinates immediately to a general purpose digital computer, or *off-line,* storing the coordinates on a magnetic tape or other convenient storage medium [KELK74].

The *head* or *stylus* used by the operator may be mounted on a pantograph arm connected mechanically to the table, or it may be *free-moving,* registering the coordinates through a mesh buried in the table to which it is coupled magnetically, electrostatically, or acoustically. The head itself may be a special ballpoint pen or a magnifying glass with crosshairs. The size of the area which may be digitized without remounting the drawing or photograph typically ranges from 20 cm × 26 cm to 75 cm × 100 cm, with a resolution of 0.2 mm. Additional features such as backlighting, projection systems, digital coordinate readouts, monitor displays, alphanumeric keyboards, and echo plotters are available. Prices range from $3000 to $20,000.

The head position is sampled either when a microswitch is activated by depressing the stylus or the pushbutton on the crosshair assembly, or at a predetermined number of times per second (1, 10, or 15). Yet another option is an integrating mechanism which allows the sampling to take place at constant distance (rather than time) intervals. The constant rate (so many samples per second) has the advantage that more points are selected when the operator slows down for exceptionally sinuous curves. Typical tracing speeds vary from 30 cm/minute for relatively straight lines such as railroads to 6 cm/minute for streams in mountainous regions. The recommended reference on digitization procedures is still DIEL72.

To the processing programs the digitized data appears as a stream of coordinates.

While digitizing maps or photographs it may also be necessary to associate each trace with some alphanumeric identifying information, such as the name of a river or a town, or the elevation corresponding to a topographic contour. This information may be entered along with the trace information using an alphanumeric keyboard (or, on small systems, a hex keypad).

Alternatively, a unique identifying number may be generated by the program which processes the trace information, and the appropriate information may be entered subsequently using a computer-generated trace map which displays the identification number associated with each trace.

A third possibility is the use of a *software keyboard* [WAGL78]. This device consists of a grid on which each square is labelled with a number, letter, or other commonly required item (such as a soil type). Rather than alternate between tracing stylus and conventional keyboard, the digitizer operator may then enter ancillary information by simply pointing with the tracing stylus at the appropriate square. The x-y coordinates of the square uniquely identify the desired item, and the processing program can convert these coordinates to the appropriate code. The keys may, of course, be recoded for each application, and the software keyboard itself may be moved at will on the digitizing tablet (provided that its location is entered, perhaps by pointing at two corners, each time it is moved.)

Over the last ten years coordinate digitizers have become available from an ever-increasing number of manufacturers and with sufficient software support to allow attachment to most computer systems. It is not unlikely that soon they will be almost as common as output plotters in multi-user computer facilities.

An entirely different type of digitizer is the *stereo-photogrammetric model* which is a high-precision optical instrument for reconstituting three-dimensional surfaces from pairs of photographs taken from different perspectives. An eyepiece enables the operator to trace contours of constant elevation or to mark the position of special features on the photographs. The output of the device is either a plot (or map) of the surface or a computer-readable file of cards or tape. As yet, the stereo-photogrammetric model is little used directly for data input to geographic data processing systems, but the maps traced with coordinate digitizers are the product of this instrument in the first place [THOM66].

Optical scanners (sometimes also called *scanning densitometers* or simply *digitizers*) perform the conversion of graphic material to computer-readable form automatically. Such instruments either scan the area line by line (*raster scanners*) or they follow the traces on a line drawing in a manner akin to the operator of a coordinate digitizer (*curve followers*). Raster scanners are far more common. They range in resolution from about 400 lines by 400 lines for television camera type scanners to 20,000 lines by 20,000 lines for high-precision drum scanners. Prices for complete systems range from $3000 to $300,000. Some scanners are designed for grey-scale material and produce an eight-bit byte for each picture element (*pixel*), others are restricted to black and white operation. Several have color separation capabilities, enabling them to scan complex multicolor line drawings in one pass. Advances in solid state scanning arrays promise considerable improvements in the price-performance ratio of optical scanners [BOYL72].

To the processing programs the raster-scanned data appears as a row or column ordered matrix of reflectance values.

As a rule, optical scanners are used in geographic data processing only for material specially prepared for this type of input by *tracing* manually certain predesignated features from maps or aerial photographs. Different colors may be used for different features, although multiple overlays may be used instead. The result is a clear, uncluttered line drawing which may be digitized readily with an optical scanner. There is a direct analogy here to the widespread practice of retyping documents in OCR fonts for input to automatic optical character readers.

2.2 Remote Sensing

The third major source of information for geographic data processing is the terrain data produced directly by airborne or satellite sensors. Remote sensing is, of course, a burgeoning technical field in itself, with more than a dozen dedicated conferences each year, the *Journal of Remote Sensing,* a two-volume handbook [ASP75], and a large and active research community. A survey on computer activities in remote sensing published by one of the present authors [NAGY72] is now almost completely out of date, but an up-to-date survey by Eric Teicholz from the Harvard University Laboratory for Computer Graphics appeared in the July 1978 issue of *Datamation* [TEIC78]. For details on current research, the *Proceedings of the Michigan Remote Sensing Symposia* (the Twelfth Symposium was held in 1978) remain the best source of first-hand information. Digital computer processing techniques are reported in a useful collection of reprinted papers [BERN78].

A comprehensive assessment of the impact of remote sensing on geographic data processing is outside the scope of this paper. We must confine ourselves to a brief description of the most important data acquisition instruments, and of the major steps necessary to transform the digitized data into a form suitable for further use.

The two major types of transducers are *multispectral scanners* and high-resolution *television cameras* (return-beam vidicons). With the former, one scan direction is the result of the forward motion of the scanning platform (airplane or satellite) itself, while the other is the result of mechanical oscillation or rotation of the optical system. With the latter, the electronic scan mechanism records "snapshots" of overlapping frames [BERN76].

The analog signals corresponding to terrain reflectance are first converted to digital form either on board or at a ground receiving station. The array of picture elements (*pixels* or *pels*) recorded for each spectral window of the sensor may then be restored, enhanced, filtered, or otherwise transformed using the well established techniques of image processing [ROSE74]. The digitized "signatures" may also be converted to attribute form by automatic or semi-automatic pattern recognition systems or, more commonly, the necessary information may be extracted manually from a two-dimensional plot of the data. Regardless of the manner in which the information is extracted, it is desirable, of course, to start with the most faithful representation possible of the area under observation.

The need for exact (element-by-element) superimposition of two images of the same scene upon one another arises in the preparation of color composites, chronological observations, and sensor-to-sensor comparisons. The spatial, temporal, and spectral aspects of image congruence are discussed in ANUT69 and ANUT77.

The *registration* problem is cumbersome to state mathematically in its entire generality, but the following formulation may help in understanding the work currently in progress.

The scene under observation is considered to be a two-dimensional intensity distribution $f(x, y)$. The recorded image is another (digitized) two-dimensional distribution $g(u, v)$. The image is related to the "true" scene $f(x, y)$ through an unknown transformation T:

$$g(u, v) = T(f(x, y)).$$

Thus in order to recover the original information from the recorded observations, we must first determine the nature of the transformation T, and then execute the inverse operation T^{-1} on the image.

When independent information is available about T, such as calibration data on distortion and degradation, a model of atmospheric effects, or attitude data concerning the angle of view, then the two operations may be separated.

Often, however, only indirect information about T is available, usually in the form of another image or a map of the scene in question. In this case, our goal must be to transform one of the pictures in such a manner that the result looks as much as possible like the other picture. The measure of similarity is seldom stated explicitly, since even if the two pictures are obtained simultaneously, the details perceptible to the two sensors may be markedly different. Thus, for instance, in registering photographs of the same scene obtained simultaneously through different color filters, we would want shorelines and rivers, but not intensity levels, to correspond. On the other hand, if the photographs are obtained years apart with the purpose of observing the erosion of the shoreline or the shift in drainage patterns, then we must expect changes in the location of such features. Seasonal variations also give rise to problems of this type.

The case of known (or derivable) T is sometimes called *image restoration*, as opposed to the classical registration problem where T must be obtained by repeated comparison of the processed image with some standard or prototype. This dichotomy fails, however, when the parameters of T are obtained by visual location of outstanding landmarks followed by automatic computation of the corrected image.

Geometic distortions in electronically scanned imagery are due to changes in the attitude and altitude of the sensor, to nonlinearities and noise in the scan-deflection system, and to aberrations of the optical system.

Geometric distortion affects only the position rather than the magnitude of the grey-scale values. Thus

$$f(u, v) = f(T_c(x, y))$$

where T_c is a transformation of the coordinates.

If the transformation is linear, the parameter vector contains only the six components necessary to specify the transformation, i.e., $c = (A, B, C, D, E, F)$ where

$$u = Ax + By + C$$
$$v = Dx + Ey + F$$
$$f(u, v) = f(Ax + By + C, Dx + Ey + F).$$

Important subcases are pure translation ($A = E = 1$, $B = D = 0$), pure rotation ($C = F = 0$, $A^2 + B^2 = D^2 + E^2 = 1$), and change of scale ($A/B = D/E$, $C = F = 0$). From an operational point of view, the transformation is specified by the original and final location of three noncollinear points. In executing a linear transformation on the computer, it is sufficient to perform the computations for a small segment of the image in high-speed storage, and transform the remainder, segment by segment, by successive table lookup operations. Aside from the saving in high-speed storage requirements, this technique results in an approximately tenfold decrease in computation over direct implementation of the transformation.

A fast algorithm suitable for digital computers equipped with move byte-string instructions has been reported in BAKI71a, BAKI71b, and WILL70. This algorithm is intended for the correction of small distortions, such as those due to the camera characteristics, and is based on the fact that relatively large groups of adjacent picture elements retain their spatial relationship in the corrected picture. The program computes the maximum number of adjacent elements that may be moved together without exceeding a preset error (typically one coordinate increment). Experimental results show that the boundaries between such groups are not visually detectable [BERN71, MARK71, BERN76].

Photometric degradation (occasionally also referred to as "distortion," with questionable propriety) arises from modulation transfer-function defects including motion blur, nonlinear amplitude response, shading, vignetting, and channel noise.

The atmospheric effects of scattering and diffraction, and variations in the illumination, also degrade the picture, but these effects are in a sense part of the scene and cannot be entirely eliminated without ancillary observations.

Once the pictures to be matched have been corrected for these sources of error, resulting in the digital equivalent of perfect orthophotos, the relative location of the pictures must still be determined before objective point-by-point comparisons can be performed. In reality, this is a chicken-or-egg problem since the pictures cannot be fully corrected without locating a reference image, but the location cannot be determined accurately without the corrections.

Tracking and ephemeris data usually provide a first approximation to the position of the sensor at the time of exposure, but for exact registration, more accurate localization is required. In operator-aided systems the landmarks are located by eye, while in fully automatic systems some correlation process is usually employed.

The volume of data produced by satellite observation platforms renders automatic classification of the data particularly attractive. Current research in pattern recognition is directed at the automatic production, using spectral "signatures," of Level I land-use maps, soil-classification maps, snow and ice cover maps, and crop inventories. The reader interested in the range of potential applications investigated under NASA sponsorship may consult FRED74.

The largest single cohesive undertaking to date, aside from military and meteorological applications, is undoubtedly the Large Area Crop Inventory Experiment (LACIE) sponsored jointly by the National Aeronautics and Space Administration, the United States Department of Agriculture, and the National Oceanographic and Atmospheric Agency. This experiment incorporates a database containing historical crop yield data, agromet information, and multitemporal LANDSAT images [LACI78]. Nevertheless, perhaps because of concern over the continuing availability of LANDSAT data and the difficulty of extracting the data from the digital tapes distributed by the EROS Data Center, the use of information produced from remote sensor data by pattern recognition techniques is still the exception rather than the rule in geographic data processing.

2.3 Output Devices and Products

The input devices described in Section 2.1 all have counterparts used as output devices. Thus *incremental plotters* correspond to coordinate digitizers, *vector display* cathode-ray tubes to curve-following scanners, and *raster displays* and *matrix plotters* to raster scanners. Some laboratories have taken advantage of this duality by converting a relatively inexpensive output device into the corresponding input device.

Before we proceed to these specialized devices, we note that ordinary *line printers* are frequently used as output devices for low-cost, portable geographic data processing systems. Almost all installations possess such devices, the software interfaces are relatively simple, and fairly good outline maps can be readily produced at low cost (Figure 12, see p. 165). Furthermore, the overprint feature allows grey-scale printouts with about ten levels of shading.

Incremental or step plotters decompose the picture into a large number of straight-line segments. Software packages are usually available to facilitate the drawing of axes, labels, shading, and variable-width lines. With multicolored pens and acetate plotting surfaces, extremely high quality output can be provided, though a single map sheet may require several hours of plotting. Plotters, like digitizers, may function on-line or off-line. Plot sizes of 25 cm × 25 cm to 100 cm × 100 cm are common, with much larger plot-beds available for specialized applications. Prices range from $2000 to $25,000.

Storage tubes provide the display equivalent of incremental plotters. They are commonly used as interactive computer terminals and are quite suitable for displays. If equipped with a joystick, they also provide limited graphic input capability. Typically, storage tubes provide 1012 × 1012 addressable raster points. Grey-scale displays are just becoming commercially available. Hard copy "slave" units have substantially increased the popularity of the storage-tube output option [BOYL72].

Refreshed *vector-displays* have come down considerably in price since the $100,000 configurations of the late 1960s. Their domain of applicability is similar to the storage-tube displays, except that they provide a faster response since the entire display does not have to be redrawn for each minor change. Lightpens provide a convenient means of interaction with the display. Although they are popular in computer graphics laboratories, refreshed displays have not caught the fancy of the geographic data processing community (perhaps because of their limited resolution).

Matrix plotters have a black and white output similar in quality to that of office facsimile machines. With a resolution of 0.25 mm between adjacent elements and page-sized output, they are best used as a rapid and inexpensive preview option for higher quality incremental plotters. Priced under $5000 and usually provided with incremental-plotter compatible software, they are a bargain for high-volume operations.

Although all of the devices mentioned in this and the preceding section normally come with a considerable amount of software, the programs included are not usually suitable for the exacting demands of geographic data processing. Some high-precision plotters, for example, are intended primarily for automated drafting applications, so additional routines must be written for spatial registration, for editing the data, and for the identification of every entity. Human factors play an important part in the design of a good graphic data-entry system, a fact which took far too many years to become widely recognized in the much simpler area of alphanumeric input [GILB77]. In regard to output, perhaps the most difficult—and so far unsolved—problem is the judicious placing of labels and names (Figure 5), which constitutes yet another demonstration of the vast distance from heuristics to art. These considerations will be evident in our discussion of specific systems, where input/output programs typically consume several man-years of effort.

2.4 Error Control

The most error-prone component in geographic data processing is data entry. Furthermore, errors not caught at this stage lead to subsequent high costs. In experimental systems the data entry and quality control functions are not usually adequately separated. It is therefore essential to inform the data entry operator of the exact division of labor between him and the computer. The operations manual supplied by the manufacturer of the input device must be supplemented by a careful description of the permissible range of operations and of the manner in which the computer programs will interpret each entry. Consideration should also be given to the preparation of notation conventions for communication between the geographer and the data entry operator. Another desirable aid is an exact, step-by-step protocol for a sample data entry task, including the interpretation of error messages and the correction of errors.

Both the data entry operator and the end user must be provided with a description of the hardware and software environment in which the programs were designed and tested. This is particularly important when the system incorporates numerous custom-built components.

A simple model of computation and data organization within the computer is quite desirable to help the user understand the overall logic of the programs he executes. The user must also be warned of system

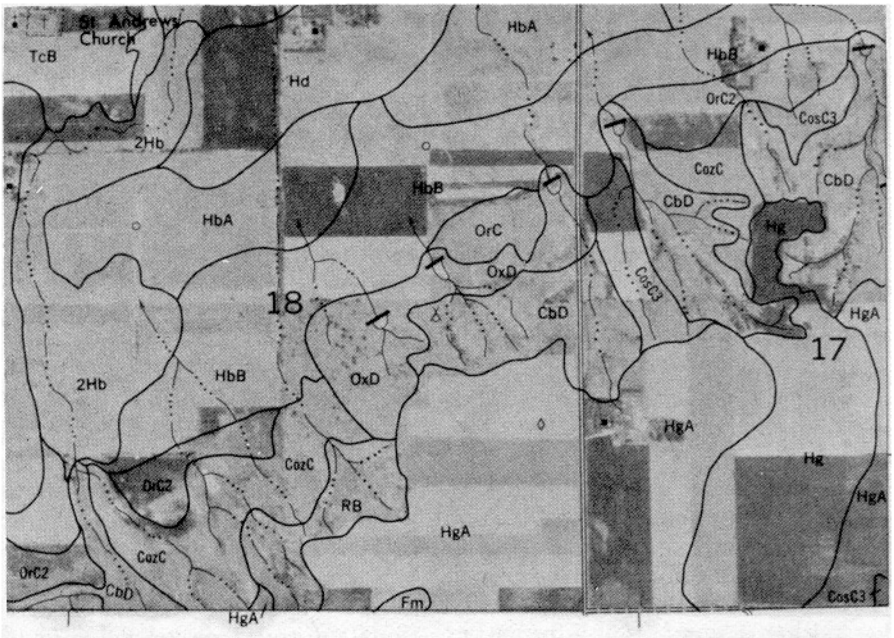


FIGURE 5. The complexity of label placement. Some capabilities required of automatic label placement programs are shown above: variable typesize, slanted labels, avoidance of mapped features, pointers to very small regions, multiple labels for large regions

limitations, such as questions of the accuracy of positional observations or statistical independence of data, to avoid undue faith in the results.

Other administrative aspects related to documentation include indexing and maintaining files of *source materials* such as photographs, maps, field notes, and special-purpose input (i.e., tracings for optical scanning), and of *products* such as maps, internal and external queries, reports, and user documentation.

Logs must be kept of digitization and scanning sessions, computer processing runs, plotting, and other accesses to specialized equipment in order to determine possible bottlenecks. Program and data updates need to be carefully monitored and logged, files of raw and processed data (on tape or disk) indexed, and accurate financial records kept on personnel, equipment, and computer costs. Ideally, all of these functions would be computer based.

Adequate *error control* requires the continuing collection of statistical information from many sources. *Quality control* is usually achieved by the separation, and sometimes duplication, of functions.

In many instances these functions are more difficult to implement than in other data processing environments because of lack of permanent or continuing funding, shortage of experienced administrative personnel, and the high turnover rate in the academic or research environments where most geographic data processing systems still operate. Furthermore, the operations are usually too small to warrant the establishment of a bona fide EDP arm to provide administrative support of the types just discussed.

3. DATA ORGANIZATION AND PROCESSING

The amount of processing required in a geographic data processing system may vary from next to none between input and output (as in simple cartographic applications) to the very complex computations required in some planning and forecasting applications. The former may, for instance, maintain the trace of a street and the location of a house alongside (all in coordinate geocodes) and never associate the two as being proximal. The latter may combine information from stream networks, ground cover maps, and climatological records to indicate areas suitable for wildlife preserves for particular species. In analogy with MISs (Management Information Systems), we might properly call a system capable of *spatial analysis* a GIS (Geographic Information System).

The algorithms required to transform the raw data into meaningful information in the form of cartographic products or answers to queries range from the trivial to the intricate. We have chosen to classify them according to the characteristics of the data on which they operate.

Broadly speaking, we may separate the information attached to geographic entities into *geometric* and *nongeometric* types of attributes. The geometric attributes generally specify the location and shape of the entity. The nongeometric attributes may provide either *nominal* information ("river," "San Francisco," "soybeans") or *scalar* information (elevation, soil permeability, flow rate). Of course, some attributes, such as "population density," may involve combinations of other primitives.

Most of the necessary processing and data organization thus involves either relations between the various attribute types, or means of interfacing with input and output devices. In the following sections we will therefore elaborate on the entity and attribute types encountered in geographic information processing systems, and in the process construct a rude taxonomy for the algorithms which serve as the basic building blocks. We will, however, attempt to refrain from pushing this taxonomy beyond its natural limits.

3.1 Data Types

Geographical entities are conventionally divided into *point*, *line*, and *region* types according to their geometry. This neat division is to some extent illusory, since lines or curves can be represented by a succession of points and regions or areas can be defined by their perimeters. Nevertheless, these three primitives will provide a convenient point of departure for our discussion of *geocoding*.

Examples of point data are irrigation wells, dams, mountain peaks, cities on large scale maps, and street lights on city maps. Lines are used to represent rivers, roads, or topographic contours. Land use, soil classification, crop type, and drainage basins represent region types.

The geometrical entity type refers only to the manner of locating an item. In addition to its location attributes each entity generally carries a number of other important attributes, which may be either nominal, such as a city name or crop type, or scalar, such as well depth or contour elevation. In some systems, such additional information is not stored directly with the locational attributes but is referenced instead by means of elaborate pointer structures.

It should be noted that items of a quasi-geographic nature often appear in conventional information retrieval or management information systems. Indeed, it is possible to refer to entire hierarchies of geographic subdivisions (state, county, township, precinct) without explicit definition of locational coordinates. Here the proper name or label of each subdivision is used as the locational identifier for all the entities within its boundaries. Such systems are, of course, incapable of producing maps of graphic displays, nor can they deal with questions involving relative positions or distances. We shall not concern ourselves with them.

3.2 Geometric Relations and Operations

In this section we will enumerate the functions which can be defined in terms of relations between entity types discussed in the preceding section. Since there are three entity types—point, line, and region—there are six types of binary relations we must consider, including those between entities of the same type (Figure 6).

Point-Point Relations

We may consider *coordinate conversion*, which has already been discussed at length in a previous section, as an example of a relation defined on geometric points. The associated operations range from simple

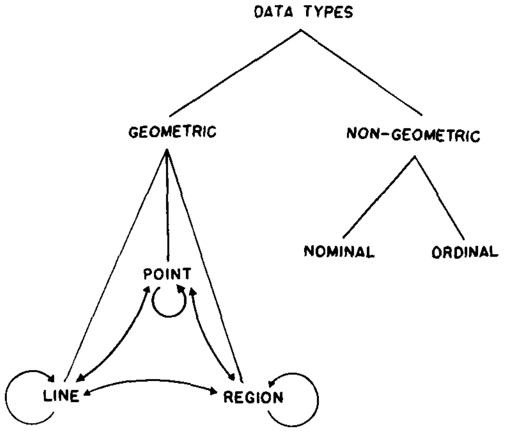


FIGURE 6. Geographic data types. Algorithms may be characterized according to the types of data on which they operate. For example, point inclusion is a *point-region* operation, while line intersection is a *line-line* operation.

scaling to the diverse cartographic projections.

Another relation defined on two points is that of *identity*. When entering point entities such as the location of irrigation wells by means of a digitizer, for example, it is often necessary to determine whether a given entity is the same or different from an entity entered earlier (and possibly from another source). This determination is usually carried out according to a simple threshold criterion on the distance between the two points.

Point-Line Relations

A distance criterion is also used to determine whether a point entity, such as a town, should be directly associated with a line entity, such as a river. This operation is more time consuming than the point-to-point comparison, since the entire line may have to be searched. One possible shortcut is to divide the line into segments and to store the maximum and minimum x- and y-coordinates in each segment. Then a given segment needs to be searched only if the point is within the minimum and maximum coordinates of the segment. We will encounter this notion of the *bounding rectangle* repeatedly [LOOM65, BURT77a].

Connected sets of line segments play an important part in many geographic systems. Examples are stream and road networks. In these instances, the end points and intersections of line segments constitute the *nodes* of the corresponding planar graphs which are often of interest in their own right.

Point-Region Relations

Among the most important relations under this category is that of *inclusion*. In the simplest case, we need to determine only whether a point is inside a given region or not. This can be accomplished through the well-known *plumb-line algorithm*, which can be visualized as dropping a plumb-line from the point under consideration, and noting whether the string crosses the boundary an odd or even number of times [ALDR72] (Figure 7). For continuous curves this algorithm yields the correct result whether the region is convex or not, but needs appropriate modifications for vertical lines, for multiply-connected regions, and for quantized data. It is known to mathematicians as *Jordan's Theorem*.

In the more general case, we wish to decide to which of *several* regions a point or points belong. For instance, we may want to find out in which county a well with given latitude and longitude is located. It is possible to do this, of course, without examining any boundary segment twice. In either case, the bounding rectangle method mentioned above can offer substantial savings over a direct application of the plumb-line algorithm.

In the most general case, we wish to give

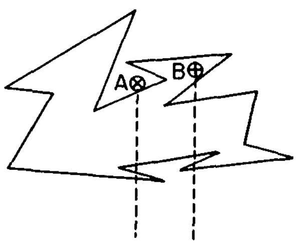


FIGURE 7. The plumbline algorithm. The plumbline from point A crosses the region boundary an even number of times because it is outside the region On the other hand, for a point inside the region, such as point B, the plumbline exhibits an odd number of crossings.

unique labels to all of the points on a uniformly spaced grid which are associated with each region. For convex regions, a single scan of the entire map suffices. When the regions are not convex, cellular propagation algorithms provide one method of solution [ROSE66, ROSE69]. This process commonly arises in converting regions coded in coordinate-chain form to grid-coded form, and is a type of *map coloring* [NAGY79].

Yet another point-region operation is the determination of the *centroid* of a region, for instance for the purpose of printing a label on a map. Here again a single traversal of the boundary is sufficient for the purpose.

The conversion of an entire region to a single point on a smaller scale representation of the information also entails a point-region relation. An example is the reduction of the outline of towns on county maps to dots on state maps. This operation, which is not reversible, is an instance of what geographers call *generalization*. Generalization has many difficult conceptual and practical problems associated with it, some of which are examplified by queries of the type: How many trees make a forest?

Line-Line Relations

As in the case of points, the question of the *identity* of a line arises when data is entered from multiple sources. A threshold criterion may be applied either to the average deviation between the two lines or to the maximum deviation.

The *continuity* of curves may also be in question when one segment of a stream is digitized from one map quadrangle or photograph, and the next segment from the map or photograph covering an adjacent region. Here the usual approach is to determine the intersection of the lines with the map edges, then to use the point-identity algorithm to match the corresponding segments on either side. This operation is called *catenation* or *join*. The complementary operation of *segmenting* a line arises in sectioning a database to prepare map products whose extents do not coincide with that of the source material (see the description of *sectioning* in Section 4.4).

It is often necessary to calculate the *length* of a line or network of lines (or rivers or road networks). For highly involuted curves, such as those representing mountain streams, allowance needs to be made for the apparent reduction in length due to the sampling process on data entry.

Intersections of curves may have to be computed unless they are explicitly marked during digitization (this is possible only with manual data entry). In the worst case, when multiple intersections are expected, every point in one curve must be compared to every point in the other, and the intersections determined on the basis of point identity. Here again, the application of bounding rectangles may result in a significant speedup.

Line-Region Relations

Among the operations requiring a very systematic approach to algorithm and program design is the reconstruction of a set of regions from separately entered boundary segments. This operation really involves entities of all three types: the points of intersections, or nodes; the boundary segments and the complete boundaries; and the regions themselves. The provision of redundant information during digitization, such as a connectivity matrix or the directions in which the boundary segments are traced, facilitates the task of sorting out the boundary segments.

In the case of directed line segments such as rivers, *right/left* relations may be important. Even where line segments have no inherent direction, an arbitrary sense may be assigned internally to facilitate tracing the boundaries of closed areas.

A simpler relation defined on lines and regions is that of *inclusion*. One may, for instance, ask whether a certain river crosses any forests, and if so, which segments are in the forested region. This problem reduces to repeated application of the line-intersection algorithm.

The *medial axis transform*, or *skeleton*, which has proved useful in image processing applications [PFAL67, MONT69], does not seem to have surfaced yet in geography, but is mentioned here as an example of an esoteric line-region relation.

Region-Region Relations

The basic measurement involving regions is that of *area*. The area is usually computed by integrating along region boundaries with respect to one of the coordinate axes. Even if the areas of several regions need to be determined, one pass along each boundary segment suffices [NAGY79].

The most interesting operations involving regions are the logical combinations such as intersection and union. There is no limit, of course, to the number of conceivable logic functions defined through the basic operators AND, OR, and NOT. Boolean algebra is also adequate for the expression of *hierarchical relationships* such as those between states, counties, towns, and precincts, and for the common cartographic operations of *joining* and *sectioning* maps (Section 4.4).

Regions are considered *neighbors* if they share a boundary segment. One region is the *island* of another if it is completely contained in it and the two regions do not share a boundary segment. Islands give rise to multiply-connected regions, which pose unexpected problems in some operations.

All six of the relations just described can be put in a formal framework through *map grammars*, which are the two-dimensional counterpart of the powerful tools developed for linguistic analysis [MILE68, DACE70, ROSE71, PFAL72]. The necessity and usefulness of such formalisms remain, however, to be demonstrated in the context of geographical data processing.

3.3 Computational Geometry

Geographic data retrieval is often based on distance considerations. Examples are finding the point closest to a predetermined set of points, and finding the two nearest neighbors among a set of points. In the one-dimensional analogy to each of these examples, the solution can be obtained by means of standard sorting techniques with a number of operations proportional to the logarithm of the number of points. In the two-dimensional case the optimal solution has been more elusive, and the recent discovery of efficient algorithms has signalled the emergence of the new and exciting field of *computational geometry* [ALDR72, PALM75, SHAM75, DOBK76, LEE76, BENT76, BURT77a, BURT78].

The efficient determination of the *nearest neighbor* of a specified point is based on the computation of the proximal polygon corresponding to each point. The edges of these polygons are the right bisectors of the straight lines joining pairs of points (Figure 8); they are called *Voronoi polygons* in geometry and *Thiessen polygons* in geography. It is sufficient to determine which of these polygons contains the specified point. The nearest neighbor problem is thus transformed into a *point-inclusion* problem. Fast algorithms are then derived by sorting the nodes or edges of the planar graphs formed by the proximal polygons [SHAM75, LEE76].

An elegant solution of the *two-nearest neighbors* problem (finding the closest pair in a specified set) can be derived through a divide-and-conquer approach [BENT76]. The general result is the replacement of an n-point problem in d dimensions by two $n/2$-point problems in d dimensions plus an n-point problem in $(d - 1)$ dimensions. A logarithmic bound ensues by recursion.

Both of these approaches require a certain amount of computation—such as the calculation of the proximal polygons—which is independent of the query and de-

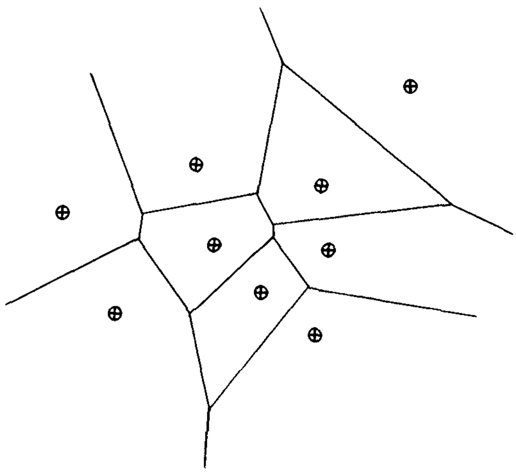


FIGURE 8. Voronoi polygons. This is a geometric construction for "nearest neighbor" queries. Formed by interpoint bisectors, the polygons partition the plane into as many regions as there are data points. Finding the nearest neighbor of a test point is thus reduced to determining its containing polygon.

pends only on the database. Thus in evaluating an algorithm it is necessary to take into account the cost of this *preprocessing* and the amount of *storage* required by the preprocessed data, as well as the number of operations required by the search algorithm itself. Among the problems investigated in computational geometry is the nature of the tradeoffs among these factors.

It should also be noted that the above algorithms are most efficient when the nodes of the underlying planar straight line graphs are of order 3 or greater. In geography, where natural curves must often be approximated by long chains of short segments, the old suboptimal algorithms may actually be faster! Nevertheless, the new developments are bound to have an eventual impact on analytical geography.

There are also many older and widely known problems with geometric, and hence geographic, interpretation. Among these are the *transportation*, *facility location*, and *traveling salesman* problems in operations research [THIE75]. A geographical system aiming to provide planning assistance may yield a problem that has already been elegantly modeled in this way. Workable solutions are, however, only approximations, since these problems are generally NP hard or NP complete [HORO78].

3.4 Surface Variables

The nongeometric attributes of entities may also be involved in the determination of the inter-entity relations. Thus, the end point of a road at a map edge may coincide with the end point of a river on the other side, but obviously this does not constitute a proper catenation. Similarly, in the computation of the *urban* population of a county the inclusion relation as well as the attribute "city" are relevant. Even the tolerances used in establishing geometric relations may depend on the entity types. In summary, the distinction between geometric and value-related operations is a matter of the degree to which the respective attributes are used.

It is useful to distinguish operations on data where there are no distinct entities and therefore no underlying geometry. In this case the geometry is *induced* by the values of a *surface variable* of the type $v = f(x, y)$ at every point in the area. When the values v are *nominal* this geometry takes the form of a partitioning of the area into regions of various types, as in a land-use *thematic* map. Political *subdivision* maps illustrate another instance of the use of nominal surface variables. *Scalar* values, such as elevations, on the other hand, can be geometrically viewed as a set of contour lines. These geometries are important as equivalent encodings of such data and as an aid to human perception (Figure 9).

A matrix provides the most convenient nongeometric representation of surface data. It stores the values of the variable at a selected matrix of points in the area. Interconversions between the matrix and the geometric representations (line or region) form an important group of operations. Different data representations are discussed further in Section 3.6.

Two or more surface variables may be combined to give a new variable. Soil erosion, which depends on the soil type, slope, and land use, is an example. Average annual rainfall from two periods can be differentiated to yield a surface which may indicate a pattern of weather change.

When population is given by a matrix, the elements represent the number of persons in the corresponding cell of the area.

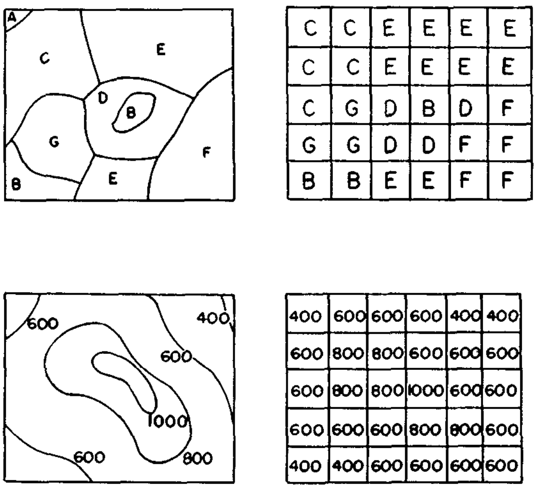


FIGURE 9. Dual representation of surface variables. The cellular representation is obtained from the polygonal representation by determining the dominant value of the variable in each cell.

Using this data it is possible to assign to any point a value which represents the distance within which a given population resides. The resultant *range* variable is an example of geometrical combination of surface variables to obtain new variables.

In most cases the surfaces generated by the scalar-valued variables are continuous, allowing interpolation theory to be applied. In particular, interpolation is used to substitute the requirement of measuring the value of the variable at each matrix point by its measurement at only a relatively *few* data points. The application of interpolation to compute the matrix from these values is called *gridding*. The two most popular methods of gridding are the weighted average method and the trend surface method.

In the *weighted average* method the value of the dependent variable at every grid point is set at the average of a fixed number of data points, the contribution of each of which is weighted inversely to its distance from the grid point under consideration. In order to perform the computation expeditiously, a grid containing the sample points is first constructed, and the search then proceeds in an expanding spiral about each grid point to be filled until the required number of data points are found. The weighting function may be set by the user, with the inverse square law serving as the default option.

The *trend surface* is simply a two-dimensional polynomial computed by a least-squares fit to the data points [WATS71]. The value of the mapped variable at each grid point is then set equal to the value of the polynomial at that grid point. The degree of the polynomial function is chosen by the user. Either a single (global) polynomial is used for the entire extent, or local approximations are made by means of spline functions with fixed or variable knots [BIRK69, PAVL72]. The residual differences between the points and the values of the computed polynomial provides some measure of the degree of fit, but care must be taken to avoid using a polynomial of degree higher than the measurements warrant.

A comparison of these and other methods of gridding is provided in RHIN75. Among the methods Rhind believes to be superior to the above are the *polyhedron method*, which consists of establishing a network of triangles linking all of the data points and assuming that within each triangle the dependent variable is a linear function of the values at the vertices, and the theory of *Kringing*, which is based on two-dimensional filtering of the independent variable in the correlation domain [DELF75].

The reliability of interpolation gridding depends on how randomly spaced the data points are. The user can evaluate the reliability if the system computes for him a *distribution coefficient* of the sample points.

Gridding produces values of a dependent variable, such as elevation, at *uniform increments of the independent variables* (x and y). A common objective of gridding is the production of contour (or *isopleth*) maps. Contour maps and displays provide an improvement over a stark matrix of values for human appreciation of relief (elevation) or similar scalar-valued data [MORS69]. The preparation of contour maps requires, however, the computation of the values of the independent variable at *uniform increments of the dependent variable.*

Contour maps can also be derived directly from the raw data, without the intermediary of gridding. Many sophisticated algorithms exist; established techniques involve the detection of the intersections of contours with edges of a mesh whose nodes correspond to known values of the independent variable. Intersections are then connected using either straight-line segments or interpolated curves [COTT69, CRAN72, MCLA74]. Plotting contours of a scalar function by means of limited-resolution character-oriented output devices such as programmable-font printers or CRT displays is demonstrated in WARD78. The inverse problem—reconstructing a surface from contour data—is discussed in FUCH77.

A contour map, though an improvement over numerical representation, is still not the most agreeable depiction of three-dimensional surfaces. One enhancement consists of *shading* the contours so as to give the impression of viewing the "terrain" from an aircraft under the morning sun

[SPUR75]. Alternatively, *isometric* or *perspective* views from one or several vantage points may be constructed. The representation of three-dimensional surfaces is one of the main concerns of computer graphics; several texts treat the subject in detail, although usually with more emphasis on the display of three-dimensional objects than on surface variables [NEWM73, CHAS78, GILO78].

3.5 Input-Output Operations

The arrangement of data as it is produced by the input device or demanded by the output device is not always compatible with the data organization most suitable for processing. The same may also be true with respect to the storage of data on auxiliary storage media. In addition, a variety of source formats and product requirements may have to be accommodated with respect to scale, projection, extent, etc. The rational way to treat these differences is to localize them in the input and output processing modules.

Coordinate conversion is an integral part of these modules. As an example, every point digitized from a source map, mounted arbitrarily on the digitizer table or scanner platen, has to be converted from its coordinates relative to the digitizer-frame-of-reference format to the internal geographical referencing format (e.g., latitude and longitude).

The conversion is a two-stage process. By rotation and translation, the first stage restores the coordinates relative to the source-map frame of reference. The second stage takes care of scale and projection changes. The parameters for conversion often precede the coordinates in the input file. Their entry and processing constitutes *map registration.*

Line-Matrix Conversion

When a map contains only lines (i.e., no points) it may be entered either by digitization or by optical scanning. (The latter is not suitable for points). At the output stage, line maps may be produced by line plotters or by raster plotters. The scanner and the raster plotter model the map as a picture composed of a matrix of black and white dots. The coordinate digitizer and the line plotter model the map as a sequence of (x, y) coordinates. As a result, the complementary operations of extracting lines from the corresponding bit matrix and of converting lines into matrix form are introduced (Figure 10).

There are differences between these line-matrix conversions and similar conversions involved in the case of the dual surface variable representations (*gridding*). The source area represented by each bit in this matrix is an order of magnitude smaller than the cells of a surface matrix. Further, in a bit matrix only the cells covered by the lines are marked, while in the conversion of polygon maps to thematic cell maps all the cells are marked with the type of the regions to which they belong and the boundaries between the regions are implicit in the transitions between types.

The line-to-bit-matrix conversion is relatively easy. One starts with a zero ("white") matrix corresponding to a blank picture, and merely sets to "black" the points covered by each line in turn. Unfortunately, without knowing the line origins and directions, this process in reverse does not yield a line-extraction algorithm. Instead, the origins have to be globally searched and then the eight possible neighboring cells have to be examined for each extension of a line [ROSE76].

Other factors also contribute to the complexity of this operation. Optical scanners actually record the grey level of each point and depending on the quality of source material the black/white threshold has to be determined. Noise in scanning may re-

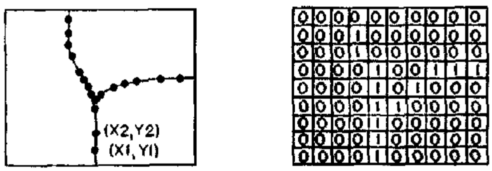


FIGURE 10. Bit matrix representation. These dual representations of line maps arise when raster devices are used for input or output. Since the segment representation on the left is more economical to store, interconversions between these representations become necessary.

sult in black/white reversals. In some cases this manifests itself as a line break, which upsets the line-extension mechanism. More disturbing is the fact that a line on the source map causes a whole band of dots, rather than a simple string of connected cells, to be set to black. The line-thinning process aimed at restoring the center-line of cells in this case depends on the direction of the line, which must be estimated. To date, only polygonal maps in which the lines are available as separate overlays have been entered successfully by means of optical scanners. Even with such specially prepared maps, considerable editing is usually necessary, and this cannot be performed conveniently without a coordinate digitizer or lightpen display.

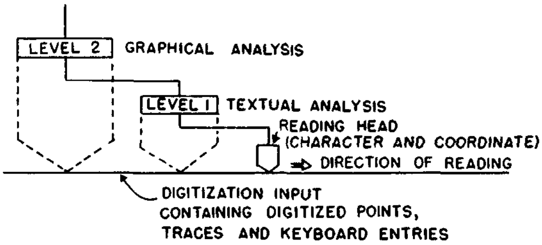


FIGURE 11a. Lexical analysis schematic.

alphabetic ::= letter | letter alphabetic
numeric ::= digit | digit numeric
delimiter ::= delimiter-character
letter ::= A | B | ··· | Y | Z
digit ::= 0 | 1 | ··· | 8 | 9
point ::= coordinate coordinate

FIGURE 11b Regular grammar used in the textual analysis

Coordinate Digitizer Input

The line and dot maps entered using a digitizer appear to the program as a stream of characters with the coordinate and command-language strings (entered from the keyboard) intermixed. The two are distinguishable only by certain delimiting characters. An elegant structure can be imposed on the input module that processes this stream by viewing the data as "text" in a digitization language and writing a compiler for it. The grammar that characterizes that language then also becomes useful as a model of implementation and digitization procedure. Other techniques in compiler writing are also applicable. A lexical analyzer, for example, can isolate tokens and take care of local editing, elimination of duplicate coordinates, and coordinate conversion [WAGL78].

A reduction in the complexity of the lexical analysis of the digitization input is achieved by separating the textual considerations from the graphical considerations. The lexical analyzer can be viewed as having a hierarchical set of three reading heads (Figure 11). The heads at levels 0, 1, and 2 read character-coordinates, point-tokens, and graphic-tokens, respectively. The reading heads in Figure 11 are staggered with respect to each other due to the look-ahead at the lower level which is required in any lexical analysis.

The textual analysis at level 1 uses conventional techniques based on the use of regular grammars [LEWI76]. The digitization input is a string of letters, digits, delimiters, and coordinates. These are grouped into tokens using the grammar in Figure 11. Only the point tokens among these are further analyzed at level 2. The remainder are also tokens at level 2 and are passed to the parsing routines without further modification. The remainder of this discussion on the lexical analysis describes the various functions performed at level 2.

In most digitizers, digitizing an isolated point causes a mini-trace to be recorded. The mini-trace consists of a set of points clustered around the point entered by the operator. The digitizer samples the position of its pen at regular intervals of time, and therefore, the number of points in the mini-trace depends on the length of time the operator rests the pen on the point. A function of the graphical analysis at level 2 is to recognize a mini-trace as a "discrete point" and reduce it to a single pair of coordinates. The clustering property is used to differentiate a mini-trace from a genuine trace.

The remaining functions at level 2 are related to reading the digization of a trace. The first among these, called "filtering," eliminates any duplicates in the sequence of points representing a trace. Furthermore, additional points in the input are skipped in such a manner that the length of each "link" (the line between two preserved points) is approximately equal to the value of a parameter set by the parser. Filtering reduces the number of points to be stored and processed, but this is offset by a possible loss of accuracy.

A major advantage of processing the digitization tokens with a lexical analysis module is that it renders the system relatively independent of the specific model of digitizer used. For a new digitizer, only the symbol table must be recoded.

Plotter Output

If the programming language supports the basic line-drawing and character-writing operations, the task of generating a map by using an incremental plotter is relatively straightforward. Unless the output is an exact reproduction of the source, the additional operations of drawing the subextent boundary and deleting the lines which lie outside it are involved. Other important procedures include drawing captions and geographical grid marks, labeling entities, showing directions with arrows (as in the case of rivers), and shading. Symbol placement generally requires sophisticated heuristics for producing maps of acceptable aesthetic quality. Alternatively, the position, size, and orientation of the lettering and symbols may be specified by the operator.

Display Output

In an interactive system the input and output modules are strongly coupled. One way of taking advantage of this coupling is to set up an *interaction context.* The context keeps information about what the user is currently viewing and delimits his possible input. It may also maintain a brief history of the interaction. In geographical interactive systems, extent, scale, and similar information would be included in the context. The use of context to simplify the editing commands is in this case extended to the pictorial entities.

3.6 Internal Data Organization

There appear to be only two major approaches to internal data organization in geographic systems; all of the examples that we have examined make use of variations or combinations of these two prototypical structures. We shall call the two approaches *cellular* organization and *linked* organization although, as is commonly the case with database terminology, these terms are not entirely satisfactory. In some respects the cellular and linked organizations correspond to discrete and continuous models, and are therefore analogous to integer and real number representations respectively.

Either organization is suitable for describing all three geometric types—point, line, and region—of geographic data. Each organization can also be applied to any entity type such as roads, towns, or soil classifications. Nor is the degree or kind of cross referencing between entity types uniquely related to the essential difference between the two structures. Furthermore, the entity attributes may be stored either with the geographic (coordinate) attributes or referenced indirectly by means of a pointer structure. We now describe the two organizations, with a view to discovering the strengths and weaknesses of each.

Cellular Organization

In the cellular organization, the extent under consideration is divided into a matrix of cells of uniform size, and some computer storage is set aside for each cell. The storage reserved for each cell contains either the information associated with the cell (presence or absence of a point or line entity, nominal or ordinal attribute of a region entity), or a pointer to that information. A linear ordering of the matrix elements thus implicitly specifies the storage address of each item.

Among the advantages of this structure is its direct correspondence to the format of raster-scanner input and matrix-plotter output. Processing related to distance is also easily accomodated since the separation of any pair of points can be computed directly from their storage addresses. Likewise, retrieval based on location can be accomplished with no additional overhead. In general, software development for almost any application is easier for the cellular approach than for the alternative linked organization. The cellular organization tends to be most suitable for the representation of surface variables. Powerful arguments for its exclusive use are given in PEUQ77.

The major disadvantage of the cellular organization is its wasteful use of computer storage for spatially sparse data. The resolution of the grid lines underlying the cellular structure is normally set equal to the highest coordinate resolution required for any of the entity types under consideration; hence the presence of a single attribute entailing great detail will affect the total storage requirement disproportionately. The resolution at input may range from a few meters (as might be obtained from optical scanning of large-scale topographic map overlays) to hundreds of meters (as in the case of satellite imagery). The internal resolution will generally be lower than the input resolution and higher than the output resolution. The storage grows linearly with the extent and number of entity types, and quadratically with the cell resolution.

It is possible to improve the storage efficiency of the cellular structure by varying the resolution of subextents of differing characteristics. This is commonly accomplished by subdividing grid divisions in such a manner that the number of cells increases by an even power of 2. This procedure, however, carries considerable overhead in terms of software development costs.

Linked Organization

The second structure is the linked organization. Here point entities are described directly by their coordinates, line entities by a chain of uniformly or nonuniformly spaced coordinates, regions with nominal attributes (thematic or subdivision maps) by their boundaries, and regions with scalar attributes by contour curves. The coordinate information corresponding to a single entity usually follows a *header* which references the entity.

A linked organization can be applied to line entities at three levels: segments, chains, or polygons. A *segment* is the smallest possible line entity in a map and consists of a straight line between two points. The DIME files described in Section 4.3 are an example of this. In a typical map, most points simply join two segments. The *chain* representation of a map reduces the number of distinct entities by recognizing only those points, called *nodes*, where more than two segments meet. In this case the sequence of segments between a pair of nodes is represented as a single chain. The *polygon* view of a map groups the chains further into boundaries of disjoint polygons. It may be noted here that in the GDP literature these three basic entities have other names also, which in some cases conflict with the definitions given here.

The linked structure corresponds directly to the format of data entered through a coordinate digitizer or reproduced through an incremental plotter. Except when the information is extremely dense, with variations of the same scale as the coordinate resolution, this organization is highly storage efficient. It is eminently suitable for processing *polygonal* (region) maps composed of non-intersecting line segments and for *contour* maps which have the property that lines do not meet.

Topological relations between data items such as "inclusion" can be represented by abstracting the appropriate properties of the coordinate chains in the form of graphs [BURT77b]. The well-developed methods of processing graphs by computer—incidence matrix, spanning tree, and lists of node pairs—can then be readily applied. Furthermore, relations between the entity types, such as those between rivers and their drainage basins, also prove easily tractable [DEO74].

Disadvantages of the linked method include the complexity of the software necessary for editing and updating the data, and the increased difficulty of performing set algebra and distance-related operations.

At the lowest level of geometric representation, the data storage required for a given coordinate string may be decreased by chain encoding where each unit increment is represented by a single digit denoting its direction, or by delta encoding where straight line segments of variable length (depending on the local radius of curvature) represent each line [FREE74].

Another elegant data organization for gridded data is the *tightly closed boundary* (TCB) structure described by Merrill. In this organization, all of the region boundary coordinates are sorted and partitioned into sets so that each set contains only points which have the same y-coordinate. The associated x-coordinates are ordered monotonically. Merrill shows that the TCB structure leads to rapid computation of inclusion relations, areas, distance of a point to the boundary, and unions and intersections of regions. The method is also extended to contour representation, and its application to an experimental terrain information handling system is demonstrated [MERR73]. It should be noted, however, that the computational algorithms provided by Merrill, while efficient, are not optimal in the sense discussed in Section 3.3.

The rigidity that normally accompanies the selection of data organization early in the game can sometimes be avoided by employing the *information hiding* principle. It is based on the duality of sorts that exists between a data representation and a set of operations to access the data. Furthermore, for a given set of access operations there generally exists more than one representation that is its dual. The use of an operational definition of a data organization leads to a significant improvement over its representational definition. A data organization module in which these operations are realized becomes a natural component of the system which localizes and monopolizes the access to the information. A related advantage is the increased flexibility of changing the representation at a later time (possibly to achieve better performance) without having to modify the remainder of the system [PARN72].

3.7 External Data Organization

There is no doubt that the design of the external data management is as important as that of the internal data organization. Specifically, factors and constraints related to *implementation*, *operation*, and *data characteristics* need to be examined.

Implementation constraints manifest themselves in the costs of machine execution, maintaining the data on auxiliary storage, and software development itself. If the system or a part thereof is implemented on an interactive system, then the response time requirements may add another constraint.

Each operation carried out on the data influences the design of its organization. The input and output operations influence organization through the equipment characteristics; archiving considerations, through the nature of the data management software and the cost and availability of secondary storage (tape, disk, or mass storage). The processing functions (Section 3.2) are also biased towards certain organizations. The question of whether to preprocess and enrich the archived information, or to store the data as entered and process it when required, is also intimately linked to the auxiliary storage organization.

The most complex operation with most organizations is that of *updating*. Mistakes invariably occur during data entry, and a data organization that allows incremental corrections is therefore desirable. This suggests avoiding any preprocessing. The retrieval operation, on the other hand, requires such preprocessing for efficient operation.

Most designs aim at achieving good overall (average) performance. They do so by using problem-specific information such as the frequency of occurrence of various cases and the existence of additional problem-specific relations. In the geographical context, the relevant data characteristics may be spatial density of entities, number of entity types, source and product formats, volume of information, and geometric type of the entities. Also important are permissible tolerances, which govern the amount of checking which must be performed on the input. The information on auxiliary storage (disk or tape) may be organized either as a data bank or as a database.

Data Banks

A *data bank* is "separation" oriented. The information is segregated either on the basis of type of entity or, in the geographical case, on the subextent basis, or on both. Data banks are simple to implement and are suited for input, output, and processing that follow along the lines of division.

Such systems generally maintain on disk or tape files many maps that span a large extent. The files contain the data in simple sequential order with little or no cross referencing. The *input subsystem* accepts the input and updates the data, and the *output subsystem* generates output based on information currently available in the bank. As a rule, the more complex operations discussed in the previous sections are not part of the data bank, which serves merely as a *repository* of information. For a typical example, see BANS78.

The preparation of input from several source maps usually takes a very long time, and it is unacceptable to withhold any of the data available at any stage. Consequently, the data bank is operated in a state where some subextents are completely available, some require editing corrections, and the remaining subextents are not available at all. The input in any run, therefore, may correspond to an originally empty subextent or may contain incremental editing corrections to one already available.

Selective retrieval of entities allows the user to specify an arbitrary extent in his output request, rather than mere reproduction of input maps. In some cases the data bank's division into subextents may not be the same as the source maps; thus, soil maps that follow county boundaries may not form a good division for a larger extent. Also, the extent may be covered by more than one series of maps based on differing subextent divisions. In all these cases *selective updating* of more than one data bank subextent becomes necessary as a complementary operation. These operations are equivalent to the *join* and *sectioning* operations (see Section 4.4).

Data banks are typically established by administrative divisions responsible for large extents, and hold vast volumes of data. Hence the efficient management of that data becomes important. Techniques for data compression also become relevant in this context. The administrative aids for management have already been described in Section 2.3.

Databases

Databases, on the other hand, are integration oriented. The integration consists of adding inter-entity relations to the information. The augmented relational information increases the usefulness and versatility of the information. When the relations are restricted to be part-whole type, the database is said to be *hierarchical*; otherwise it is called a *network* database [DATE75, MART77]. Inclusion within a region, tributary of a river, political subdivision, and node representation of lines are examples of hierarchical geographical relations [ALVA74]. Neighbor, or confluence of two streams, are examples of more general relations. One of the vexing problems in developing a database model of an information world is the determination of relevant relations. These are seldom obvious for geographical information.

Databases are also important in geographical data processing in another context. If the nominal codes of entities are used as a surrogate for their locational information, the GDP applications reduce to EDP, where numerous database systems are already available. If the locational information and other attributes are separated, with the nominal identifier appearing in both, then a better organization results. Thus the non-geometrical information may be organized into a database, while the geometric information may be organized as a data bank.

While database systems in which the physical data organization is transparent to the applications programmer have already made their debut in business, they are still waiting in the wings in geographic applications. The trend towards database systems will be accelerated by the unpredictable, heterogeneous requirements of the researchers, administrators, and planners who constitute the bulk of the current users, and by their need for access to a variety of data organized into a single integrated structure. For example, to determine erosion patterns, soil type information, land ownership, and crop management practices are simultaneously needed. Furthermore, in many situations, the user cannot foresee all of his information requirements, which are revealed only incrementally in the course of the interaction with the database.

4. SPECIFIC SYSTEMS

This section contains a sketch of ten specific systems which we consider representative of the entire span of geographic data processing. The information about each system was either culled from the literature or obtained through correspondence with the authors or caretakers.

We have not been able to find a completely rational linear ordering for presenting these systems (except possibly by alphabetic order of their villainous acronyms), but we shall attempt to proceed from the simple to the complex, starting with portable program packages, through digital cartographic systems, to the closest that we have been able to find to true geographic information processing systems.

4.1 ODYSSEY

No survey of the GDP field can be complete without mention of the pioneering and innovative work carried out by the Laboratory of Computer Graphics and Spatial Analysis at Harvard University. Over the years the Laboratory has developed a number of general purpose GDP programs which can be bought from the Laboratory at modest prices. Equally important, the Laboratory is a rich source of standard cartographic maps in digital form. Many papers have been written by the staff documenting their work and explaining the algorithms used in the programs. An overview of the work and of the ODYSSEY system can be found in DUTT77.

The well-known SYMAP program, designed by H. T. Fisher of the Northwestern Technological Institute, was the earliest contribution of the Laboratory [SCHM75]. After more than 15 years the SYMAP program is still widely used today, its popularity stems from its applicability to many fields [LIEB76]. The program uses cards and printer as its input and output media and can be readily implemented in most computer installations. SYMAP has also been extended to other input-output media such as digitizers and graphic terminals.

The user of SYMAP can obtain either a contour (*isopleth*) map or a conformant map of a surface variable. In either case, the user provides the values of the variable at a selected set of sample locations. The geographical area under study by the user is treated by SYMAP as a matrix of print positions. In the contour option, the program uses standard gridding techniques to extrapolate the values at the sample locations to each print position (Section 3.4). The value at each print position is then converted to a print-character corresponding to the interval into which it falls. The intervals used for this purpose and into which the range of the surface variable is divided are user-selectable. The final printmap produced by SYMAP for this option displays interval bands instead of the usual lines on a contour map (Figure 12).

In a variation of the contour map called the *proximal map*, the gridding operation is replaced by assigning to each print position the value of the nearest sample point. A proximal map produced by SYMAP looks like a zone map with each zone corresponding to the area of influence of each sample point. This option can, therefore, be used to produce "neighborhood" maps.

In a *conformant map* the study area is divided into zones and the value at a sample location represents the value at each point within its zone. When using SYMAP for producing a conformal map, the user also provides the coordinates of the boundaries of the zones. In contrast with the two previous options, the program assigns the value of each sample point to all the print positions in its zone. Conformant maps printed by SYMAP appear as shaded zone maps with each zone shaded according to its value.

ODYSSEY represents a more recent effort by the Harvard Laboratory to bring all its programs into a single system and make them compatible with each other. The former goal is achieved simply by designing a "navigator" through which the user of ODYSSEY accesses the programs using simple commands. The latter requirement has been realized by defining a standard sequential file structure for storing geographical data and forcing every program in the system to accept and produce such files. The file structure itself is flexible and can be used to represent a variety of geographical data. The flexibility is achieved by introducing a file header which identifies the data elements appearing later in the file and describes their format. All programs in the ODYSSEY system are written in FOR-

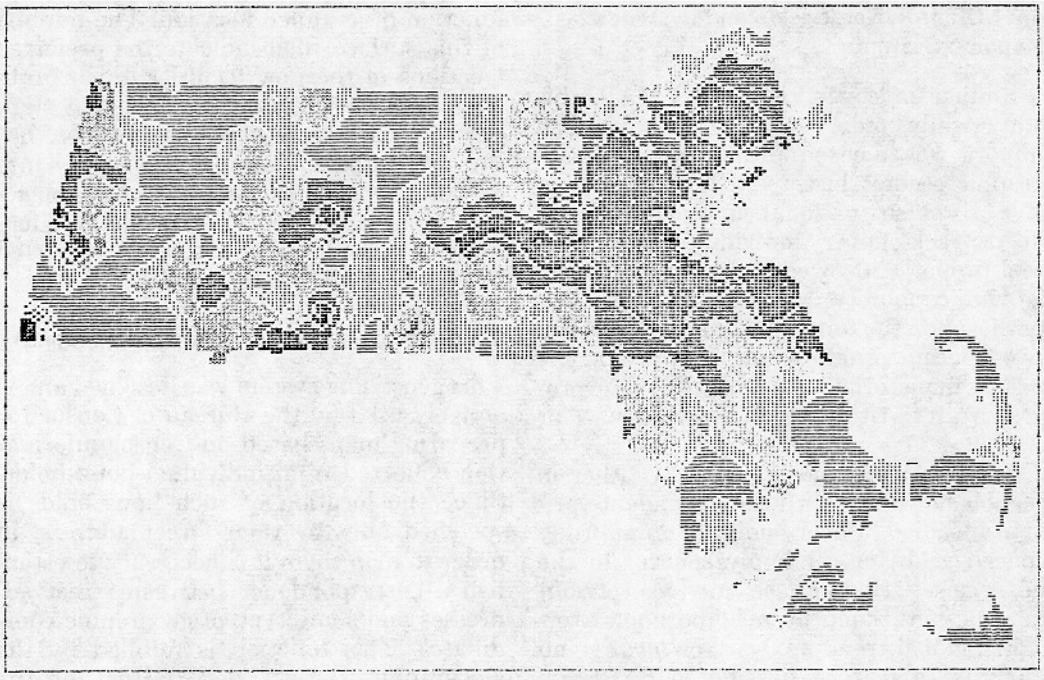


FIGURE 12 A SYMAP-produced contour map Scale considerably reduced, each dot was generated as a character on a line printer [SCHM75].

TRAN, making it a portable system. An I/O module providing common operations on the map files is shared by the programs and is useful in developing additional programs.

Most of the programs in ODYSSEY perform conversions between different representations of a map. Additionally, there are programs which perform utility tasks that include conversion of format, change of projection, entry of a map through a digitizer, output of a map as a plot or a print-map, and interactive editing of a map file. Programs producing 3-D views of maps are also available, and can be used to enhance the visual impact of an output map.

Probably the most significant contribution by the Harvard Laboratory is the technique of *local processing* which has been incorporated in some of their programs. The entities in a map file—segments, chains, or polygons—are first stored according to their leading x coordinate to obtain a "spatially coherent" map file. Processing such a file sequentially is equivalent to scanning the original map from left to right. If the generation of an output map by the program is also treated as a sweep across the study area, then it is clear that an input entity cannot contribute anything to the output after the input sweep has passed its leftmost x coordinate. This allows the data elements to be purged from the data area systematically, as the input sweep represented by the reading of the input file progresses from left to right. The same data is then shared by subsequent data elements. The output sweep, at any time, stands at the rightmost x coordinate in the data area and progresses as the data elements are discarded.

The local processing technique has reduced the storage requirement of many GDP programs. It is therefore now possible to process a much larger volume of data than before. The technique has also improved the detection of errors in the input, which is a difficult and time consuming task in programs that operate on the entire map simultaneously.

The lack of pure cartographic capabilities (see Sections 4.4 and 4.5) in ODYSSEY programs is disconcerting. It would be worthwhile to explore the extension of this system for cartographic purposes.

4.2 NORMAP (Nordbeck's and Rystedt's Mapping System)

Similar in conception to SYMAP, this statistically more sophisticated Swedish import was devised primarily for demographic studies based on the comprehensive Swedish national databank, which keeps track of every individual and piece of real property in Sweden. Its signal feature is the capability to test statistical hypotheses on the functional relation between two spatial variables and to present the results in pictorial form [NORD72]. The programs are written in either FORTRAN or ALGOL.

NORMAP accepts input data either in regular net form (with the dependent variable specified at every point of a uniformly spaced grid), or randomly spaced. In the latter case, gridding (see previous section) may be performed by neighborhood averaging, global trend analysis, or *local* trend analysis (a feature missing in SYMAP). Gridding may be avoided altogether if the assumptions involved in data-point triangulation are acceptable, with contour lines computed by direct linear interpolation along the edges of the triangles.

A statistical hypothesis regarding a functional dependence between two experimentally observed variables is tested by plotting the residual difference between one variable and the postulated function of the other. If the function is a spatial shift or translation, then this provides a means of determining the degree of spatial correlation in the data. Complete spatial autocorrelation functions may also be plotted.

Estimates of two-dimensional probability density functions may be computed with NORMAP either through a Parzen-window approach or a k-nearest-neighbors algorithm [DUDA74]. With specific reference to population densities these plots are called *square net population* and *range* maps, respectively.

Means are also provided for determining the optimal location of service facilities (called *central places*) for a given population distribution. For instance, a cost function corresponding to the sum of the distances each person would have to travel to the nearest facility can be plotted as a function of assumed location. The minima of this surface then indicate the preferred locations of the new facilities. Such problems can also, of course, be tackled by standard operations research techniques, but the map outputs allow planners to take into account the constraints imposed by the actual topography and existing cultural features, a consideration too often neglected by purely mathematical models.

4.3 DIME (Dual Independent Map Encoding)

This geocoding system was designed and is actively used by the Bureau of Census for preparing maps based on Census information collected from individual households. Since the location of such households is specified only by their street address, in order to map them it is necessary to establish a correspondence between street addresses and some type of geographic coordinates. This function is fulfilled by the Geographic Bases Files (GBFs) of the DIME system, now available for over 250 metropolitan areas. The base files provided by the Bureau of Census are available to outside users and are widely used for urban spatial analysis applications; although, according to law, the Census data itself cannot be distributed below certain minimum levels of aggregation (census tracts). In the following we shall not be concerned with the Census data, but only with the means of encoding the street address [SILV77].

In each standard metropolitan area (usually extending beyond the actual city boundaries), each intersection is arbitrarily numbered and its geographic coordinates are recorded relative to a rectangular grid. Fortunately, the accuracy of the measurements does not require taking into account the curvature of the earth. The intersections are considered the *nodes* of a graph, with the street segments between intersections corresponding to *links*. Each block is also numbered and identified with its census tract. For each *link* a record is created listing the nodes to which it is incident, the blocks on either side, and the range of street addresses between corners. Curved streets are broken down into straight line segments, with pseudo-nodes inserted at the junction points (Figure 13).

Lincoln, Nebraska, a city of 165,000 inhabitants, has about 2400 links and nodes, and therefore requires that number of records for complete characterization. Each record requires, on the average, 300 bytes of data. The entire Geographic Base File for Lincoln fits on a 1200-foot reel of 800 bpi magnetic tape.

The records have room for additional information, such as the boundaries of political precincts, school districts, and other administrative subdivisions. Although the Geographic Base Files are maintained and updated by the Bureau of Census, the responsibility for ancillary information rests with the users.

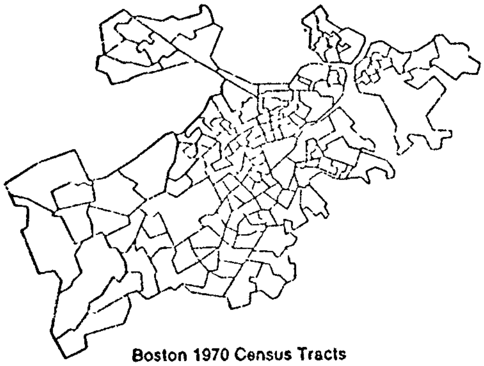
SYMAP and DIME GBF represent two extreme cases of geographic data processing systems, in the sense that the first consists only of a set of programs with no inherent geographic information, while the second consists only of geocode information and no processing capability whatever.

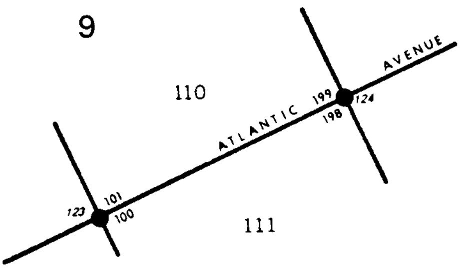
4.4 ECS (Experimental Cartographic System)

The avowed objective of the ECS, developed over the last ten years at the Rome Air Development Center, is to provide the opportunity to investigate automated cartographic techniques in a production environment similar to that of field elements of the Air Force and of the Defense Mapping Agency. As shown in schematic form in Figure 14, ECS is indeed a complete and advanced cartographic system which includes most of the features of interest for general purpose automated cartography [AYER74]. It offers the further advantage of excellent documentation which provides interesting insights into the problems encountered in computerizing an operation hitherto considered in the realm of craftsmanship or even art [RADC72, DIEL72].

ECS accepts data from three sources: existing maps, stereo photographs, and panoramic photographs. Features from maps and photographs are entered through a coordinate digitizer; stereo photos through a stereo plotter with digital output; and tracings from panoramic photos or maps through an optical scanner. Editing facilities are provided through a keyboard and a high-precision plotter arranged to function as a coordinate digitizer with the stylus under direct operator control. Both incremental and raster plotters are available for output. All of the input/output equipment is served by a dedicated minicomputer, but the heavy intermediate processing (see below) is accomplished on a large, general purpose machine.

Each item of source data is tied to abso-





FOR EACH STREET SEGMENT A DIME RECORD CONTAINS:

From Node	123
To Node	124
Street Name	Atlantic
Street Type	Avenue
Left Addresses	101-199
Right Addresses	100-198
Left Block	110
Left Tract	9
Right Block	111
Right Tract	9

FIGURE 13 Dual independent map encoding The top right figure was extracted from the geographic base file (GBF) for the Boston metropolitan area. The bottom left figure shows the cross-referencing between street addresses, intersections, city blocks, and census tracts [SILV77].

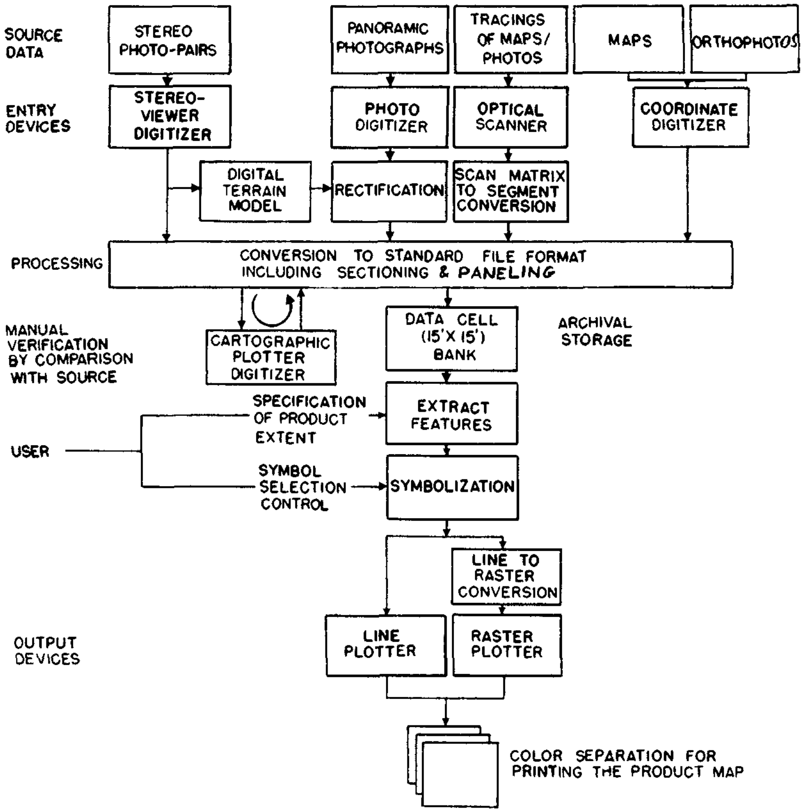


FIGURE 14 ECS schematic. This figure illustrates the relations between the components of a complete cartographic system Adapted from DIEL72

lute geodetic coordinates through *control points* whose relative location is recorded precisely by the digitization operator (control points are prominent natural or cultural features which can be readily identified and the location of which is kept on file in national mapping agencies). After digitization, each source item is plotted and checked, and additional editing entries are made if necessary. The digitized data is then converted to a device-independent code and recorded on magnetic tape for transmittal to the major processing programs. It should be noted here that the system also maintains a three-dimensional surface model of the area under consideration, called the Digital Terrain Model (DTM), which serves to correct the location of objects obtained from panoramic photographs.

The basic ECS unit is the *data cell,* comprising a 15 minute × 15 minute area. In compiling the information for a data cell, the first step is *sectioning,* which consists of eliminating all the elements in the source material (digitized maps and photographs) that extend beyond the boundary of the data cell. The next step is *paneling,* which is the difficult process of piecing together the material from different source items across the "seams" within the data cell. After sectioning and paneling, which are time consuming operations even on a large computer, another round of editing completes the data compilation. The information corresponding to the data cell is then stored in the library on one or two reels of tape.

When an order for a product arrives, paneling and sectioning of the data cells take place in reverse order to assemble the information for a map of arbitrary extent (overlapping either partially or fully one or more data cells). This is followed by *sym-*

bolization and *coloring*, where, for example, a keyboard operator enters the code to have a highway printed in black, single or double line.

The final step is the plotting (really etching) of each color on an emulsion-coated plastic surface suitable for direct multicolor offset printing. For a typical map (a mythical entity like a typical face), the entire process is estimated to require about 100 man-hours.

For an appreciation of two decades of progress in automated cartography, the reader is urged to consult TOBL59 and TOBL75. The parallel path traced at the British Experimental Cartographic Unit (ECU) is documented in BICK64 and THOR77.

4.5 SACARTS (Semiautomated Cartographic System)

Similar in scope and capabilities to the ECS, the Defense Mapping Agency Topographic Center's SACARTS is in routine use for the production of standard, 1:50,000 scale military topographic maps [STRU75]. As in ECS, all of the compiled data is transformed to digital form prior to the preparation of the color overlays for printing the maps. The work flow, from source material to final product, is shown in Figure 15. We note, for the benefit of readers struggling to bring an 8 by 10 inch data tablet or plotter online, that SACARTS comprises nine 100 cm × 150 cm backlighted digitizing tables with 0.025 mm resolution, several minicomputers, two "coarse" tape-driven plotters (really automatic drafting machines) for verification purposes, and a high-speed, ultra-precise Concord Coordinatograph for final output on paper-coated plastic (*scribecoat*), or photographic film. The photographic subsystem consists of an automatic photoscribe head with 25 apertures for varying lineweights, and a projector head for exposing silhouettes of standard map symbols.

The basic source material for SACARTS consists of 1:50,000 scale orthophotographs (vertical view) and of digital relief information on magnetic tape from the stereo equipment. In fact, the contouring programs constituted the earliest functioning portion of the system, becoming operational in the late 1960s. The map projections and the necessary grid overlays are computed on a large general purpose computer as in ECS, with the plotters and digitizers operating either offline or under minicomputer control. A unique feature of the system is the Symbols and Names Placement System (SNAPS), which uses a drum type digitizer and photocomposer keyboard to set the type for all features, including marginal information.

The time required on the various types of equipment to prepare an entire map is shown in Figure 16. (Sample maps can be obtained from the Defense Mapping Agency.)

4.6 WRIS (Wildland Resource Information System)

The Wildland Resource Information System is intended for storing and manipulating data about land areas and, as such, serves as a transition between the purely cartographic systems described in the previous two sections and the more elaborate geographic information systems described below. The building blocks of the system are nonoverlapping polygons which represent timber stands, soil types, alternatives treatments, or other types of *coverage* information. Developed primarily for forest management, WRIS's designers hope to extend applications of the system to other fields [RUSS75].

The programs comprising WRIS are written in standard FORTRAN and documented in an exemplary fashion. The primary input device is a 230 mm × 230 mm flatbed scanning densitometer, although for very simple coverages (comprising few polygons), the use of a coordinate digitizer is recommended. Printer plots and a standard incremental plotter are used for output. Extensive housekeeping programs are provided to keep track of information for as many as 100 different areas and 1000 different attributes.

Each kind of attribute, such as timber type, for instance, corresponds to a polygon *layer*. Each layer of polygons forms a *map*.

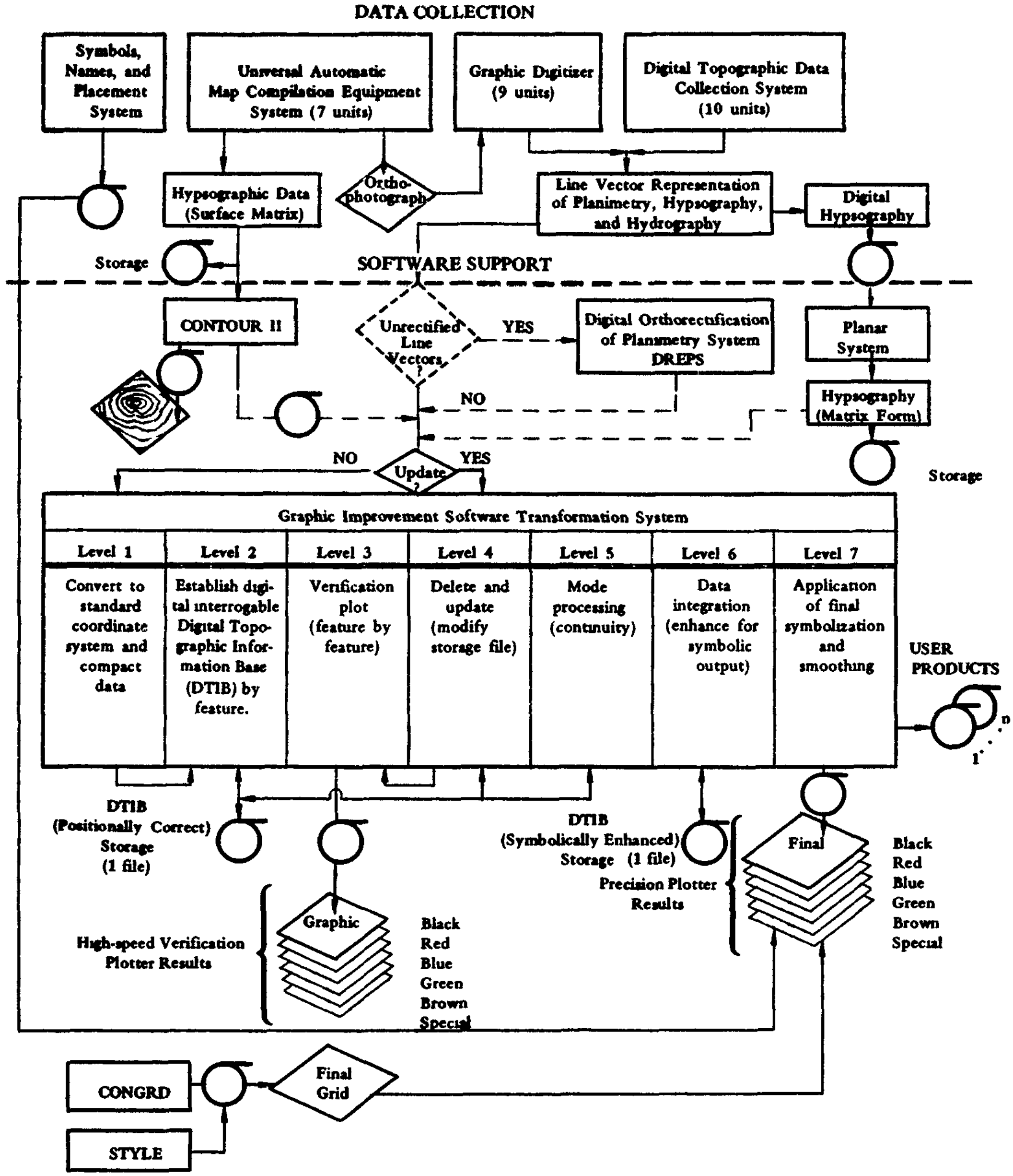


FIGURE 15 SACARTS workflow. The steps necessary to produce a topographic map from the original source material are shown [STRU75]

The specific attribute ("fir") corresponding to each polygon is specified by a *label*. The labels are not unique, and may occur many times on a given map, though the polygons themselves are assigned unique *item* numbers. The polygons in different layers need not, of course, correspond to each other in size or shape.

The polygons for each layer are carefully traced in ink on transparent paper to provide the input to the optical scanner or the coordinate digitizer. For scanner input, the

Equipment Hours

Digitizers	—	CALMA C985 with CAELUS disk	46 hours
Computers	—	UNIVAC 1108	2 hours, 8 min.
		NOVA 1200	Online
Plotters	—	CALCOMP 602	6 hours
		CALCOMP 618	13 hours
		CONCORD II	8 hours
SNAPS	—	locator system	20 hours
Man Hours			80 hours

FIGURE 16. Map preparation times The times are for a 15′ × 15′, 1:50,000, multicolor, topographic map of quality quite comparable to those produced by manual drafting techniques The table is reproduced from an actual map supplied (free on request) by the Defense Mapping Agency in Washington, D.C.

polygon maps are photographically reduced to conform to the size limitations of the densitometer. The 10-bit grey-scale scanner output is converted to a binary matrix by thresholding, and the boundaries reduced to a thickness of one pel (picture element) by means of a simple line-thinning program. Gaps and spurs are then manually deleted by reference to a printplot of the boundaries.

When the overlays are digitized manually instead of optically, the polygons are still converted to a matrix representation, thus eliminating any sequence information provided by the order in which the boundaries were digitized.

The next step is the polygon extraction procedure. This procedure is guided by the operator who enters the appropriate label for each polygon and the x-y coordinates corresponding roughly to the center of the polygon. The program moves outward from the label until it encounters the boundary and then traces the boundary in a clockwise direction until it encounters the starting point. The program also checks for dead ends, islands, and multiple labels in a single polygon, and prints explanatory error messages if necessary. The polygon extraction procedure is followed by another round of editing.

The output routines supported by WRIS include means of combining the information from two overlays (intersections and unions), accumulating acreage totals, and producing thematic maps on the plotter.

WRIS demonstrates that with careful design and a little ingenuity it is possible to develop a workable and useful geographic resource management system with a very modest investment of equipment and software.

4.7 CGIS (Canada Geographical Information System)

Measured by the number of articles which reference it, CGIS is a highly successful system. Its quality is further attested to by the fact that in the United States its adoption is being contemplated by the Geological Survey. The reasons for this lie in the system's long evolution, the commitment by a government to its development and use, the technical firsts it has to its credit, and the large volume of geographical data encoded in its data bank. Initiated in 1963, operational in 1971, its current holdings are shown in Figure 17 [TOML76]. The long development time and high cost (many millions of dollars) are due, to some extent, to the system designers' desire to work with rapidly changing state-of-the-art hardware and software.

Although CGIS accomodates the input, storage, and display of *coverage* maps, such as land use or soil type, it differs from purely cartographic systems in its processing capability to overlay two or more coverages for a region and to compute the areas of simple or compound coverages. These operations are most important in any use of coverage data and are quite expensive to perform manually at the optimal level of detail.

The original plans for CGIS included point and network data as well, but the high cost of additional program development and the relatively low demand for

Class	Coverage	Available	Not Available	For Input	Completed	%
Census	015	219	1	218	172	78.5
Watershed	025	196	25	171	131	66.8
Shoreline	105	229	1	228	183	80.3
Agriculture	205	220	54	166	151	91.0
Forestry	305	220	74	146	110	75.3
Wildlife—ungulates	405	221	7	214	165	77.1
Wildlife—waterfowl	505	221	23	198	149	75.2
Recreation	605	220	4	216	155	71.8
Sport fish	615	220	87	133	55	41.3*
Present land use	705	224	5	219	174	79.4
Totals		2190	281	1909	1445	75.7

* Production temporarily stopped.

FIGURE 17. CGIS holdings as of July, 1975. The values in the table refer to the number of source maps [TOML76].

these types of data caused postponement of this aspect of the project. Also awaiting additional funding is the development of an interactive data retrieval facility. By and large, the accumulation of data has taken far longer than expected; in fact, in 1971 the scale of the source maps had to be changed from 1:50,000 to 1:250,000—since at the former scale, 10 to 15 years would have been required to complete the work!

CGIS was the first major system to use optical scanners operating on transparent tracings of the maps as an input device. More recently, however, provisions have been added to substitute lower cost coordinate digitizers for the scanner input. So far, no definite word is available on the relative merits of the two methods.

The use of boundary encoding by means of coordinate chains in lieu of grid encoding is another distincitve feature of the system. Little has been published on this method of compiling and updating the data.

Hopes of widespread use of the system have not been realized so far. The CGIS' staff has also suffered from lack of direct feedback through involvement in the use of their data bank (in land use planning, for example). Nevertheless, CGIS has provided operational and production experience which will be invaluable in the design of other systems. The large volume of literature pertinent to CGIS has been cogently surveyed by Tomlinson, who has been associated with the system since its inception [TOML76].

4.8 STANDARD (Storage and Access of Network Data for Rivers and Drainage Basins)

STANDARD is an experimental special purpose system for automating the compilation of a certain type of report on stream drainage, hitherto prepared manually by the US Geological Survey. The source data for the system consists of 7½-minute or 15-minute topographic quads, and the final product is a report—identical in form to USGS Water Resources Division Report 7404—rather than a map [WAGL78].

The principal items in the report are the lengths of stream segments between certain points of interest and the areas of drainage basins above important cultural features, such as road crossings and county boundaries or above the confluence of major streams. The elevations of certain features, and the points at which topographic contour lines cross the streams, are also reported.

The programs, consisting of about 3000 PL/1 statements, run in batch mode on an IBM 360/65 computer. The only special auxiliary devices required are a coordinate digitizer and an incremental plotter.

As shown in Figure 18, STANDARD is a seven stage process: five are manual, while two involve use of the computer via programs referred to as "Quad Processor" and "Basin Integrator" respectively. Each of the seven stages requires an operator to man it.

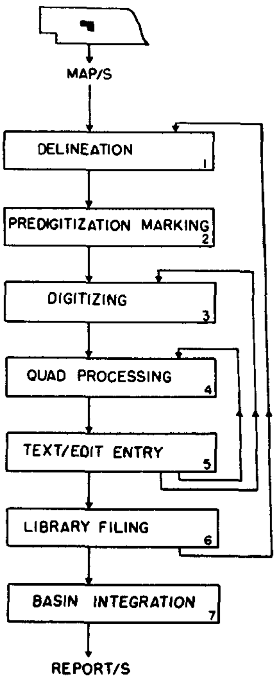


FIGURE 18. STANDARD schematic. Only the Quad Processing and Basin Integration stages are strictly automatic. Digitizing and Text/Edit Entry are interactive, while the other boxes represent manual operations.

Although not shown in the diagram, this process is preceded by a planning stage. During the planning stage the area of interest is defined and the set of rivers in the area to be included in the drainage report is listed. The specific maps that will be used to cover the area of interest are also listed. Each of these maps then passes through stages 1 to 5 individually, possibly in parallel.

Delineation

The *segments* of rivers from the reporting set that lie in each map are identified. Each of these segments has an associated basin. Using topographic contour information the boundaries of these basins are deduced and marked on the mapsheet.

A segment of a river is that part which lies between two end-of-segment points. The source, river mouth, points of confluence with tributaries, points on map-edges, and certain special *basin features* come under the category of end-of-segment points. Basin features are those features along a river, such as major highway crossings, considered significant enough to warrant determination of the drainage area upstream to them.

Predigitization Marking

The maps contain a considerable amount of information irrelevant to the system. The purpose of this stage is to improve the contrast between relevant and irrelevant information, thus simplifying the design and operation of the next stage of digitization.

Specifically, arrows are marked pointing out the positions of those features along the rivers which are to be entered. Arrows are also provided to point out the elevation markers along the rivers. The discrete (usually in tens of feet) topographic contour crossings on the rivers and the mouths of the rivers are together called *elevation markers*. The values of elevation markers are also marked at the tail of their arrows. The river and basin boundary traces are numbered for easy sequencing and enumeration during digitization. The start and end of these traces are distinctly labeled with these numbers.

Digitization

In this stage an offline digitizer unit is used to record the relevant information on magnetic tape. The delineated and marked map forms the input to this stage. The operator first enters the map identification (mapname, position on table, etc.), then traces the rivers and basin boundaries as indicated, and finally enters the positions of the features and elevation markers—the latter with the elevation values.

Quad Processing

Here the tape files which correspond to a machine-readable form of map information are processed by the Quad Processor. The next stage of Text/Edit Entry also provides the input to Quad Processor in the second and following passes through the loop formed by stages 3, 4, and 5. The first pass through this loop creates a partial database for this map. In subsequent passes, this partial database is used in conjunction with edit entries to generate an updated version of the database. To compute the geophysical coordinates of the map (latitude and longitude) given its name, a map-index is needed in the first-pass run.

A variety of output that is controllable through a set of options is generated by the Quad Processor. Besides the database just mentioned, this includes schematic maps, a local drainage area report, and details of the database, among others. An important output is a questionnaire aimed at the Text/Edit entry operator, soliciting the outstanding names of the rivers and features. These names appear as text in the report, whereas the other details (river length, elevation, basin area, etc.) are deduced from the digitizer input.

Text/Edit Entry

With the aid of the original marked map and the Quad Processor generated schematic maps, the questionnaire just mentioned is completed. The need for any editing via edit entry or digitizer input is assessed. If additional digitization is necessary, the editing instructions are passed to the digitization stage. If the database is deemed adequate, the marked map, the database, the schematic maps and other output from the Quad Processor are passed to the next stage for deposition in the library.

Library Filing

The filing operator is responsible for maintaining the mapsheets and the databases in an organized form. In the current implementation, the library serves as a database holding stage for the individual maps covering the area of interest until all of them are ready for the Basin Integrator stage.

Basin Integration

In this last stage of STANDARD, the databases of the constituent maps covering the area of interest are assembled into an equivalent database for the entire area. The Basin Integrator generates outputs similar to the Quad Processor including schematic maps and reports.

It is estimated that the manual portions of this process offer a roughly fivefold savings in time and superior accuracy, as compared to the former planimeter-and-dividers operation.

4.9 GADS (Geo-Data Analysis and Display System)

In terms of flexibility and interactive problem-solving capability, GADS, under development at IBM since 1973, is unquestionably the most ambitious of the geographical information systems we have examined. Aimed at unstructured problem solving—where the user does not know, ahead of time, either his data requirements or the specific steps he needs to take to approach the solution, GADS provides advanced information retrieval, programming, statistical analysis, and display facilities for urban applications. The system has been subjected to limited operational testing by municipal agencies interested in the optimal reconfiguration of police beats and school districts, with further applications planned in urban forecasting and for solving certain network-related (as opposed to area-related) problems [CARL74, MANT74, NIEM 78].

The GADS user interface accomodates a variety of devices such as keyboards, a refreshed color display scope, joystick and cursor, lightpen, storage tube, and incremental plotter. There is also provision to project source maps over a display screen, either to provide a detailed backdrop for a computer display or to digitize map information [GIDD73].

The basic geographic units in GADS are called *zones*. Zones may be defined either as strictly geometric entities such as uniform squares, or they may correspond to units natural to the application at hand, such as city blocks or school districts. In particular, the DIME Geographic Base File (see Section 4.3) is often used to provide the necessary locational framework.

Attractive and colorful as the GADS terminals are, the system's major contribution is its sophisticated data management facility. This may be no accident, since for several years the development team occupied quarters next door to IBM's database research group!

As shown in Figure 19, the database is organized in two levels. The large, complete database is never directly accessed by the user; instead, support personnel extract a mini database for each user. The data manipulation and display subsystems operate solely on the extracted data. This has several advantages:

1) The integrity of the main database is protected from user error (and, possibly, user curiosity).

2) The extracted mini database may be small enough to allow the analysis to be performed on a small computer instead of the large—and possibly distant—host machine [CARL75].

3) The efficiency of retrieval is improved and the software for managing the ex-

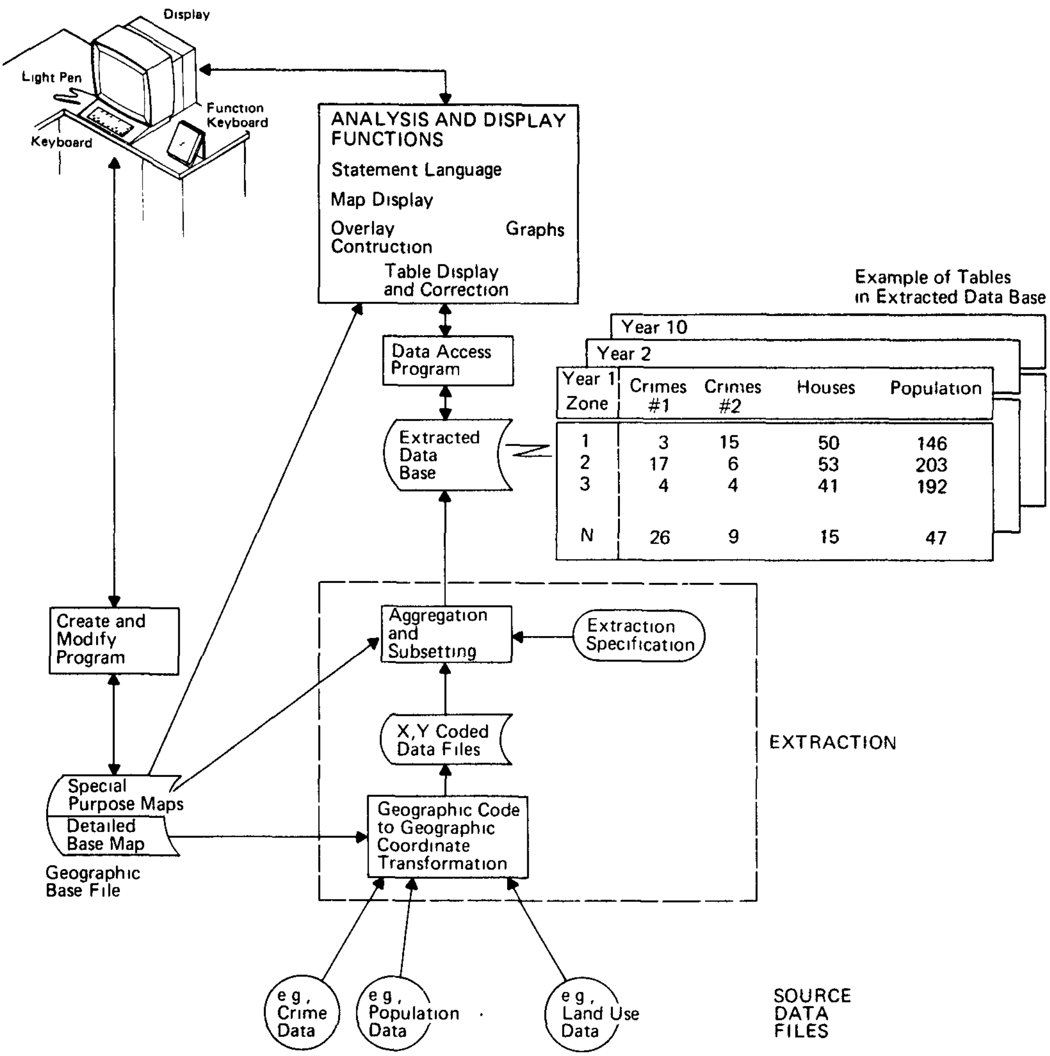


FIGURE 19. GADS schematic. The most notable feature of GADS is the insulation of users from the main database through the extraction of a specialized mini database for each user [CARL74].

tracted database can therefore be simple-minded.

4) The user may view the extracted database as a set of named tables. This view corresponds to the relational model of database organization [CODD70].

Because of heavy computer processing requirements, the relational model has not been implemented in the large original database. This need not, however, trouble the user, since his activities are confined to the mini database. Nor does he have to worry about coordinate transformation, aggregation, and subsetting operations necessary to represent his data according to the zone structure specified by his base map; that is all taken care of during the extraction process!

He can, however, using a simple command language reminiscent of BASIC, display the attributes associated with each zone by symbol, color, or density; create new attributes by arithmetic or logical operations; examine and update any table; and combine zones to create *superzones*. He may also plot one variable against another, obtain tabular displays of the data, and perform statistical computations. Finally, he can save the sequences of commands necessary to perform these actions, and gradually build up his own "program" library.

It is easy to see that human-factors considerations played a paramount role in the design of GADS, and no effort was spared to develop a user-friendly interface. In spite of this, some guinea pigs, recruited from outside the realm of computers, find the system demanding and capricious, and left to themselves revert to primitive means of decision making.

4.10 NIMS

Aided by the availability of records systematically kept by the parish and local authorities on the locations of individuals and real estate, Sweden gained an early lead in geographical data processing [NORD72, SALO69]. This lead is being maintained through systems such as NIMS [RONN76].

Unlike GADS, where street addresses have to be translated to geographical coordinates, the objects in NIMS already carry their coordinates. The NIMS data bank contains information on individuals and real estate, called "outlets." The data on individuals is characterized by age, sex, occupation, and income, among other attributes. The outlets can be classified on the basis of their type (e.g., store) and capacities. The system also keeps nonpoint information such as permanent geographical areas (e.g., towns and housing areas) and road networks. The user can, on a temporary basis, enter additional areas (polygons) or segments to suit his application.

The user at a terminal formulates his query interactively with guidance from the system. He then has the option of having the extracted data listed, tabulated or displayed in the form of maps or graphs. In any case, the query becomes the input to a program which executes in batch mode, and the user receives the output later in the day. While the lack of interactive execution does not allow data exploration as in GADS, this is compensated by the ability to formulate quite complex queries, as described below.

There are twelve basic types of retrieval options available in NIMS. Included in these are retrieval of objects of a specified type that lie within a predefined areal unit and retrieval based on the distance from a set of segments, such as a road or pipeline network. This already powerful set is further enhanced by a feature which allows two independent basic retrieval options to be specified simultaneously, the result being some overlay of the two sets. In effect, the user has a rich set of 144 different types of retrieval available to him. Although this set is more general than that available in any other system, it seems to be biased towards planning applications where the facilities (outlets and roads), treated as resources, have to be optimally distributed among their potential customers.

"Dependent outlet in topologically delimited areas with distance zones along routes" is a type of complex query and is coded 0509, since it is a combination of basic retrieval options of type 5 and 9. An example query of this type is, "Which pupils, living over 2 km from their nearest primary school area (05), are also over 1.5 km from a bus route (09) and how are they distributed in distance classes?" On most systems this query would require development of a special program. On NIMS it can be formulated by the end user. In this respect, at least, it can be said that NIMS provides a query language for geographical data.

CONCLUSION

What this paper purports to show is that geographical data processing has much in common with conventional business data processing, and yet enough distinctive features of its own to merit consideration as a separate field. The similarities include:

1) Both are in a period of rapid development, though the total volume of business data processing is now, and likely to remain, orders of magnitude larger than geographical data processing.

2) Both are characterized by vast amounts of data and complex relationships between data elements. Both will be affected and enhanced by current developments in database technology.

3) Common to both is the importance of a stable, reliable production environment in order for day-to-day operations to be incorporated into the mainstream of organizational activity (the GDP lagging, of course).

4) There exists an obvious yet neglected need for more detailed planning, better management, and more thorough documentation, with the GDP again lagging in adopting modular program development techniques such as top-down design and structured programming.

5) Both are tending towards interactive information retrieval with intelligent terminals or maxi-mini configurations.

Because of these similarities, GDP has much to learn from the best of the business DP systems and from relevant research in computer science. The features which render GDP distinctive are:

1) In geographical applications the locational attributes play an important role. The inherently two-dimensional nature of GDP profoundly influences data organization, algorithm design, and input-output. These considerations also lead to exciting new research problems in computational geometry, operations research, and image processing.

2) There is no known sufficient set of primitive operations (such as +, −, *, /, , =, , and assignment for integers) for geographic data types.

3) Geographic entities tend to require an indeterminate amount of storage.

4) Most GDP systems are in government rather than private hands because major applications are in large area planning, including natural resource management and facility location. Government funding has not yet however resulted in significant standardization (except in cartography).

5) There is a lack of large commercially developed software systems which could establish the cost-effectiveness of GDP systems.

6) Contributors are drawn mainly from the academic and scientific community (earth sciences and social sciences) rather than the business world, resulting in slower than expected conversion of prototype designs to operational systems [RHIN77].

7) GDP involves the use of specialized and expensive input-output equipment. Coordinate digitizers and incremental plotters have probably reached a plateau with regard to price and performance, but optical scanners and graphic display devices are still subject to considerable technical development and possible price reductions.

8) The maturing technology of remote sensing is likely to have considerable impact on GDP.

The number of disciplines and fields which bear on geographical data processing ensure its continuing vigor and progress. Although we have tried to present a state-of-the-art survey, we hope and believe that new developments will soon render most of our material obsolete.

ACKNOWLEDGMENTS

This work represents part of a continuing project supported by the Nebraska Water Resources Center, Natural Resources Commission, Department of Administrative Services, the Old West Commission, the Conservation and Survey Division, and the Maude Hammond Fling Fellowship Fund of the University of Nebraska The help of LuAnn Murray, Marjorie Reid, and Audrey Schardt in the preparation of the manuscript, under National Aeronautics and Space Administration grant NGL 28-004-020, is also gratefully acknowledged.

On the technical plane we are indebted to Drs. George Corliss and Frank Burton for valuable suggestions on the original manuscript, and to the ACM referees for a number of detailed recommendations and additional references

REFERENCES

ACSM74 AMERICAN CONGRESS SURVEYING & MAPPING (ACSM), *Proc. Int. Conf Automation in Cartography*, 1974.

ALDR72 Aldred, B K. *Points in polygon algorithms*, Rep UKSC 0025, IBM (UK) Scientific Center, Peterlee, 1972.

ALDR75 Aldred, B. K. (Ed) *Proc. IBM UK Scientific Center Seminar Geographical Data Processing*, Rep. UKSC 0073, IBM (UK) Scientific Center, Peterlee, 1975.

ALVA74 ALVAREZ, D. T. AND TAYLOR, M. L. *Cartographic data base hierarchy*, Vols. I, II and III, PRC Information Science Co, McLean, Va., Sept 1974.

ANUT69 ANUTA, P. E. "Digital registration of multispectral video imagery," *SPIE J* **7** (1969), 168–175

ANUT77 ANUTA, P. E. "Computer-associated analysis techniques for remote sensing data interpretation," *Geophysics* **42,** 3 (1977), 468–481.

ASP75 AMERICAN SOCIETY OF PHOTOGRAMMETRY (ASP), *Manual of remote sensing*, Vols. 1 and 2, ASP, Falls Church, Va., 1977.

AYER74 AYERS, L. F. *A review of cartographic automation*, Defense Mapping Agency, Aerospace Center, Washington D.C, 1974.

BAKI71a BAKIS, R., WESLEY, A., AND WILL, P. M. "Digital image processing for the Earth Resources Technology Satellite data," in *Proc. IFIP Congress 71*, Elsevier North-Holland, Inc., New York, 1971.

BAKI71b BAKIS, R. "Digital correction of geometric and radiometric errors in ERTS data," in *Proc 7th Int. Symp Remote Sensing of Environment*, 1971, Center for Remote Sensing Information and Analysis, Willow Run Laboratories, Univ. of Michigan, Ann Arbor, 1427–1436.

BANS78 BANSAL, M. *Nebraska natural resources data bank information system*, Nebraska Natural Resources Commission, Lincoln, Neb., March 1978.

BENT76 BENTLEY, J L., AND SHAMOS, M. I. "Divide-and-conquer in multidimensional space," in *Proc. 8th Annual ACM Symp Theory of Computing*, ACM, New York, 1976, 220–230.

BERN71 BERNSTEIN, R., AND SILVERMAN, H F. "Digital techniques for earth resource image data processing," presented at the AIAA 8th Annual Meeting, Sept. 1971, also IBM-FSD Tech. Rep. FSC 71-6017, Gaithersburg, Md.

BERN76 BERNSTEIN, R. "Digital image processing of earth observation sensor data," *IBM J. Res. Dev.* **20,** 1 (Jan. 1976).

BERN78 BERNSTEIN, R (Ed) *Digital image processing for remote sensing*, IEEE Press, New York, 1978.

BICK64 BICKMORE, D P., AND BOYLE, A. R "An automated system of cartography," in *Tech Symp. Int Cartography Assoc.*, Edinborough, 1964

BIRK69 BIRKHOFF, G. "Piecewise bicubic interpolation and approximation in polygons," in *Approximation with special emphasis on spline functions*, I. J. Schoenberg (Ed.), Academic Press, New York, 1969

BOYL72 BOYLE, A. R. "Equipment for spatial data processing," in *Geographical data handling*, Vol. 2, R. F. Tomlinson (Ed.), IGU Commission on Geographical Data Sensing and Processing, Ottawa, Ont., 1972.

BOYL74 BOYLE, A R "Geographic information systems in hydrography and oceanography," *Can. Cartogr* **II,** 2 (Dec 1974), 125–141.

BOYL76 BOYLE, A R. "The requirements of an interactive display and edit facility for cartography," *Can. Geogr.* **13** (1976).

BRAS77 BRASSEL, K *A survey of cartographic display software*, Rep. R-77/3, Geographic Information Systems Laboratory, State Univ. of New York, Buffalo, 1977.

BREE42 BREED, C. B. *Surveying*, John Wiley & Sons, Inc., New York, 1942

BURT77a BURTON, W. "Representation of many-sided polygons and polygonal lines for rapid processing," *Commun. ACM* **20,** 3 (March 1977), 166–171

BURT77b BURTON, W. "Efficient retrieval of geographical information on the basis of location," in *Proc. Advanced Study Symp. Topological Data Structure for Geographic Information Systems*, G. H. Dutton (Ed.), Harvard Univ. Laboratory for Computer Graphics and Spatial Analysis, Cambridge, Mass., 1977.

BURT78 BURTON, W "A fixed grid curve representation for efficient processing," in *Proc SIGGRAPH '78, Fifth Annual Conf. Computer Graphics and Interactive Techniques*, ACM, New York, 1978, 64–69. Published as *Comput Gr* **12,** 3 (Aug. 1978)

CALK77 CALKINS, H. W, AND TOMLINSON, R F *Geographic information systems, methods, and equipment for land use planning*, IGU Commission on Geographical Data Sensing and Processing, Ottawa, Ont., Aug. 1977.

CARL74 CARLSON, E. D., ET AL. *The design and evaluation of an interactive geo-data analysis and display system*, Rep. RJ 1342, IBM Research Laboratory, San Jose, Calif, 1974.

CARL75 CARLSON, E. D. *Using large data bases for interactive problem solving*, Rep. RJ 1585, IBM Research Laboratory, San Jose, Calif., 1975.

CHAM47 CHAMBERLIN, W. *The round earth on* *flat paper*, National Geographic Society, Washington, D.C., 1947.

CHAS78 CHASEN, S. H. *Geometric principles and procedures for computer graphic applications*, Prentice-Hall, Inc., Englewood Cliffs, N.J , 1978

CLAI73 CLAIRE, C N *State plane coordinates by automatic data processing*, C&GS Pub. 62-4, US Dept. Commerce, Environmental Science Administration, Washington, D.C., 1973.

CODD70 CODD, E. F "A relational model of data for large shared data banks," *Commun. ACM* **13**, 6 (June 1970), 377-387.

COTT69 COTTAFAVA, G., AND LEMOLI, G. "Automatic contour map," *Commun ACM* **12**, 7 (July 1969), 386-391.

CRAN72 CRANE, C. M. "Contour plotting for functions specified at nodal points of an irregular mesh based on an arbitrary two-parameter coordinate system," *Comput. J.* **15** (1972), 382-384.

DACE70 DACEY, M. F. "The syntax of a triangle and some other figures," *Pattern Recogn.* **2** (1970), 11-31.

DALE74 DALE, P F. (ED.) *Automation in cartography*, British Cartographic Society, 1974.

DATE75 DATE, C. J. *An introduction to database systems*, Addison-Wesley Publ. Co., Reading, Mass., 1975.

DAVI75 DAVIS, J. C , AND MCCULLAGH, M J. (EDS.) *Display and analysis of spatial data*, John Wiley & Sons, Inc., New York, 1975.

DEET77 DEETZ, C H., AND ADAMS, O. S. *Elements of map projection with applications to map and chart construction*, Greenwood Press, Inc., Westport, Conn., 1977.

DELF75 DELFINER, P., AND DELHOMME, J P "Optimum interpolation by Kriging" in *Display and analysis of spatial data*, J. C Davis and M. J. McCullagh, (Eds), John Wiley & Sons, Inc., New York, 1975.

DEO74 DEO, N. *Graph theory with applications to engineering and computer science*, Prentice-Hall, Inc , Englewood Cliffs, N.J , 1974

DIEL72 DIELLO, J *Experimental cartography at the Rome Air Development Center*, Rome Air Development Center, Rome, N. Y., 1972.

DOBK76 DOBKIN, D , AND LIPTON, R. J. "Multidimensional searching problems," *SIAM J Comput* **5** (1976).

DUDA73 DUDA, R., AND HART, P. *Pattern classification and scene analysis*, John Wiley & Sons, Inc., New York, 1973.

DUEK77 DUEKER, K. J., AND NOYNAERT, J E. *Interactive digitizing, editing, and mapping: software and data structure considerations*, Tech Paper 92, Inst Urban and Regional Research, Univ. of Iowa, July 1977.

DUTT77 DUTTON, G. H. "Navigating Odyssey," in *Proc. Advanced Study Symp Topological Data Structures for Geographical Information Systems*, G. H Dutton (Ed), Harvard Univ Laboratory for Computer Graphics and Spatial Analysis, Cambridge, Mass., 1977.

DUTT78 DUTTON, G. H., AND NISEN, W. G. "The expanding realm of computer cartography," *Datamation* **24**, 6 (June 1978), 134-142.

ECU71 EXPERIMENTAL CARTOGRAPHY UNIT, (ECU) *Automatic cartography and planning*, Architecture Press, London, 1971.

FRED74 FREDEN, S C., ET. AL., (EDS.) "Discipline summary reports," in *Proc. Third Earth Resource Technology Symp* , Rep. SP-357, Scientific and Technical Information Office, NASA, Washington, D C., 1974.

FREE74 FREEMAN, H. "Computer processing of line drawing images," *Comput Surv.* **6**, 1 (March 1974), 57-97.

FUCH77 FUCHS, H., KEDERN, Z M., AND USELTON, S P "Optimal surface reconstruction for planar contours," *Commun. ACM* **20**, 10 (Oct. 1977), 693-702

GIDD73 GIDDINGS, G M , AND CARLSON, E G. *An interactive system for creating, editing and displaying a geographic base file*, Rep RJ 1288, IBM Research Laboratory, San Jose, Calif , 1973.

GILB77 GILB, T., AND WEINBERG, G *Humanized input*, Winthrop Publ. Inc., Cambridge, Mass., 1977.

GILO78 GILOI, W. K *Interactive computer graphics*, Prentice-Hall, Inc , Englewood Cliffs, N.J , 1978

HARV77 *Proc Advanced Study Symp Topological Data Structures for Geographical Information Systems*, G. H Dutton (Ed), Harvard Univ. Laboratory for Computer Graphics and Spatial Analysis, Cambridge, Mass., 1977.

HARV78 *Int Users' Conf. Computer Mapping Software and Data Bases. Application and Dissemination*, Harvard Univ. Laboratory Computer Graphics and Spatial Analysis, Cambridge, Mass , 1978

HORO78 HOROWITZ, E , AND SAHNI, S. *Fundamentals of computer algorithms*, Computer Science Press, Woodland Hills, Calif., 1978.

IGU77a INT GEOGRAPHICAL UNION, *Directory of individuals in the field of geographical data handling*, The Lands Directorate, Dept of Fisheries and Environment, Government of Canada, Ottawa, 1977.

IGU77b INT GEOGRAPHICAL UNION, *Computer software for geographical data handling*, UNESCO Press, New York, 1977

KELK74 KELK, B. "Hardware for automated cartography," in *Automation in cartography*, British Cartographic Society, 1974

LACI78 *Proc. LACIE Symp.*, J.SC-14551, 14700, 14557, NASA LBJ Space Center, Houston, Texas, 1978.

LEE76 LEE, D. T , AND PREPARATA, F P "Location of a point in a planar subdivision and its applications," in *Proc 8th Annual ACM Symp Theory of Computing*, ACM, New York, 1976, 231-235.

LEWI76 LEWIS, ET AL. *Compiler design theory*, Addison-Wesley Publ. Co., Reading, Mass., 1976.

LIEB76 LIEBENBERG, E "SYMAP its uses and abuses," *Cartogr. J.* **13** (1976).

LOOM65 LOOMIS, R. G. "Boundary networks," *Commun. ACM* **8,** 1 (Jan 1965), 44–48.

MCLA74 McLAIN, D. H. "Drawing contours from arbitrary data points," *Comput. J.* **17** (1974), 318–324.

MCLE74 McLELLAN, W. G., AND MORITZ, B. K. *An interactive on-line multi-minicomputer configuration for application to cartography,* PRC Information Sciences Co., McLean, Va., 1974.

MANT74 MANTEY, P. E. *Transportation planning applications of the geo-data analysis and display system,* Rep. RJ 1459, IBM Research Laboratory, San Jose, Calif., 1974.

MARK71 MARKARIAN, H., BERNSTEIN, R, FERNEYHOUGH, D. G., GREGG, C. E., AND SHARP, F. E. "Implementation of digital techniques for correcting high resolution images," presented at the *ASP-ACSM Fall Convention,* 1971, also IBM-FSD, Tech. Rep. FSC 71-6012, Gaithersburg, Md

MART77 MARTIN, J. *Computer data-base organization,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1977.

MENE76 MENELEY, G. J. *Place coding· inventory, analysis, and recommendations,* Tech. Rep. No. 73, Inst. Urban and Regional Research, Univ. of Iowa, 1976.

MERR73 MERRILL, R. D. "Representation of contours and regions for efficient computer search," *Commun. ACM* **16,** 2 (Feb. 1973), 69–82.

MILL68 MILLER, W. F, AND SHAW, A. C. "Linguistic methods in picture processing—a survey," in *Proc. AFIPS 1968 Fall Jt. Computer Conf.,* Vol. 33, Part 1, Thompson Book Co., Washington, D.C., 279–290.

MONT69 MONTANARI, U. "Continuous skeletons from digitized images," *J. ACM* **16,** 4 (Oct. 1969), 534–549.

MORS69 MORSE, S P. "Concepts of use in contour map processing," *Commun. ACM* **12,** 3 (March 1969), 147–152

NAGY72 NAGY, G. "Digital image-processing activities in remote sensing for earth resources," *Proc. IEEE* **60** (1972), 1170–1200.

NAGY79 NAGY, G., AND WAGLE, S. "Approximation of polygonal thematic maps by cellular maps," *Commun ACM,* to appear.

NEWM73 NEWMAN, W. M., AND SPROUL, R. F. *Principles of Interactive computer graphics,* McGraw-Hill, New York, 1973.

NIEM78 NIEMEYER, G. "Dispatching by graphical terminals," *Comput. Gr.* **11,** 4 (March 1978), 58–66.

NORD72 NORDBECK, S., AND RYSTEDT, B. *Computer cartography,* Studentlitteratur, Lund, Sweden, 1972

PALM75 PALMER, J. A B. "Computer science aspects of the mapping problem," in *Display and analysis of spatial data,* J C Davis and M. J. McCullagh (Eds.), John Wiley & Sons, Inc., New York, 1975.

PARN72 PARNAS, D. "On the criteria to be used in decomposing systems into modules," *Commun. ACM* **15,** 12 (Dec. 1972), 1053–1058.

PAVL72 PAVLIDIS, T. "Piecewise approximation of functions of two variables through regions with variable boundary," in *Proc. ACM National Conf.,* 1972, ACM, New York.

PEUC77 PEUCKER, T. K. *Computer cartography: a working bibliography,* Dept. Georgraphy, Univ. of Toronto, 1972

PEUQ77 PEUQUET, D. J. *Raster data handling in geographic information systems,* Rep. R-77/7, Geographic Information Systems Laboratory, State Univ. of New York, Buffalo, 1977.

PFAL67 PFALTZ, J. L, AND ROSENFELD, A. "Computer representation of planar regions by their skeletons," *Commun. ACM* **10,** 2 (Feb. 1967), 119–122.

PFAL72 PFALTZ, J. L. "Graph structures," *J. ACM* **19,** 3 (July 1972), 411–422.

RABI76 RABIN, M. O "Probabilistic algorithms," in *Algorithms and complexity,* J. F. Traub (Ed), Academic Press, New York, 1976.

RHIN75 RHIND, D. W. "A skeletal overview of spatial interpolation techniques," *Comput. Appl.* **2** (1975)

RHIN77 RHIND, D. W. "Computer aided cartography," *Trans. Inst. British Geogr. New Series* **2,** 1 (1977), 71–96.

RICH72 RICHARDUS, P., AND ADLER, R. K *Map projections for geodetists, cartographers, and geographers,* American Elsevier, New York, 1972.

RONN76 RONNBERG, A. O. S., AND SELANDER, K. *NIMS methods of area analysis,* Working Paper No. 8, SORSA—Int. Segment Oriented Referencing System Organization, Stockholm, 1976.

ROBI69 ROBINSON, A. J., AND SALE, R. D *Elements of cartography,* John Wiley & Sons, Inc , New York, 1969.

RADC72 ROME AIR DEVELOPMENT CENTER, *A compendium of technical papers on experiments in cartography at Rome Air Development Center,* Rome, N.Y , 1972.

ROSE66 ROSENFELD, A., AND PFALTZ, J. L. "Sequential operations in digital picture processing," *J. ACM* **13,** 4 (Oct. 1966), 471–494.

ROSE69 ROSENFELD, A *Picture processing by computer,* Academic Press, New York, 1969.

ROSE71 ROSENFELD, A. "Isotonic grammars, parallel grammars, and picture grammars," in *Machine intelligence VI,* American Elsevier, New York, 1971.

ROSE76 ROSENFELD, A. AND KALE, D. *Digital picture processing,* Academic Press, New York, 1976

RUSS75 RUSSELL, R. M., SHARPNACK, D. A., AND AMIDON, E L. *WRIS: a resource information system for wildlife management,* Res. Paper PSW-107, USDA Forest Service, 1975.

SALO69 SALOMONSSON, O., ALFREDSSON, B., AND SELANDER, K. *FRIS A· a spatial information system,* Central Board for Real Estate Data, Sundbyberg, Sweden, 1969.

SCHM75 SCHMIDT, A. H., AND ZAFFT, W. A. "Programs of the Harvard University Laboratory for Computer Graphics and Spatial Analysis," in *Display and analysis of spatial data,* J. C. Davis and M. J McCullagh (Eds), John Wiley & Sons, Inc., New York, 1975.

SELI78 SELIGMANN, J., FOOTE, J., AND BEGLEY, S. "Mapping with a computer," *Newsweek*, (Sept. 11, 1978), 87.

SHAM75 SHAMOS, M. I. "Geometric complexity," in *Proc. 7th Annual ACM Symp Theory of Computing*, ACM, New York, 1975, 224-233.

SILV77 SILVER, J. *The GBF/DIME system: development, design and use*, US Bureau of the Census, Washington, D.C., 1977.

SPRU75 SPRUNT, B. F. "Relief representation in automated cartography: an algorithmic approach," in *Display and analysis of spatial data*, J. C. Davis and M. J. McCullagh (Eds.), John Wiley & Sons, Inc., New York, 1975.

STRU75 STRUCK, R. L. *Making a map with digital data*, Defense Mapping Agency, Topographic Center, Washington, D.C., 1975.

TAYL76 TAYLOR, D. R. F. (ED.) *Proc Symp. Geographic Information Processing*, Dept. of Geography, Carleton Univ., Ottawa, Ont., 1976.

TEIC78 TEICHOLZ, E. "Processing satellite data," *Datamation* **24** (June 1978), 117-132.

THIE75 THIERAUF, R J, AND KLEKAMP, R. C. *Decision making through operations research*, John Wiley & Sons, Inc., New York, 1975

THOM66 THOMPSON, M M. (ED.) *Manual of photogrammetry*, American Society of Photogrammetry, Falls Church, Va, 1966.

THOR77 THORPE, L. W. "Cartographic data structures in production use at experimental cartographic unit," in *Proc. Advanced Study Symp. Topological Data Structures for Geographical Information Systems*, G. H. Dutton (Ed.), Harvard Univ. Laboratory for Computer Graphics and Spatial Analysis, Cambridge, Mass., 1977.

TOBL59 TOBLER, W. R. "Automation and cartography," *Geogr. Rev.* **49** (1959), 526-534

TOBL62 TOBLER, W. R. "A classification of map projections," *Annu. Assoc. Am Geogr.* **52** (1962), 167-175.

TOBL75 TOBLER, W. K "Analytical cartography," *Am. Cartogr.* **3** (1975), 21-31.

TOML70 TOMLINSON, R. F. (ED.) "Environment information systems," in *Proc. UNESCO/IGU First Symp. Geographical Information Systems*, 1970

TOML72 TOMLINSON, R F. (ED.) *Geographical data handling*, Vols. 1 and 2, IGU Commission on Geographical Data Sensing and Processing, Ottawa, Ont., 1972.

TOML76 TOMLINSON, R. F. (ED.) *Computer handling of geographical data*, UNESCO Press, Paris, 1976.

WAGL78 WAGLE, S. G. "Issues in the design of a geographical data processing system: a case study," PhD Dissertation, Dept. Computer Science, Univ. of Nebraska-Lincoln, 1978.

WARD78 WARD, S A. "Real time plotting of approximate contour maps," *Commun. ACM* **21**, 9 (Sept 1978), 788-790.

WATS71 WATSON, G. S. "Trend-surface analysis," *J. Math Geol* **3**, 3 (1971).

WILL70 WILL, P. M., BAKIS, R., AND WESLEY, M. A. *On an all-digital approach to Earth Resources Satellite image processing*, IBM Tech. Rep. RC-3027, IBM, Yorktown Heights, N.Y, 1970

RECEIVED APRIL 1978, FINAL REVISION ACCEPTED FEBRUARY 1979