Franz W. Leberl
Technical University and Graz Research Center
A-8010 Graz, Austria
Dale Olson
Mark Hurd Aerial Surveys, Inc.
Minneapolis, MN 55426

Raster Scanning for Operational Digitizing of Graphical Data

Maps and drawings are converted into digital format, and the digital data are edited, stored, retrieved, and generally manipulated.

Introduction

OPERATIONAL RASTER SCANNING is not a new phenomenon, but a widely employed technique in the graphics industry and in digital image acquisition. Its use in cartography, however, is at a beginning. Since the mid-sixties or so there have been developmental studies to demonstrate the capabilities and limitations of automated raster scanning in cartography. A breakthrough to opera-

Digitizing

There is clear evidence that the requirement for digital graphics data is steadily increasing (see, for example, SIGGRAPH '80 (1980) and ISP (1980)). This has led to numerous intense efforts to speed up and automate the task of digitizing maps, drawings, and other graphical data. In its simplest and most widely applied form, this digitization employs an x, y digitizing table on which the map

ABSTRACT: *The current growth of computer graphics has resulted in operational automated systems to digitize existing maps and drawings. The configuration and performance of SysScan, a general purpose hard- and software configuration for converting maps and drawings into a digital format, and to edit, store, retrieve, and generally manipulate the digital data are described. A unique component of SysScan is KartoScan, a raster scanner using white light and a CCD array sensor. In the commercial mapping segment, this system has pioneered the operational application of raster scanning to a number of digitizing tasks, among them the generation of digital elevation models from contour lines, and the creation of digital polygon data sets. Experience in these applications are reported upon.*

tional application had not occurred until recently. Currently, automated cartographic raster scanning is being applied in a commercial photogrammetric and cartographic mapping environment to support digitization operations.

This paper discusses digitization of graphical data, describes the hardware, and then goes on to report in some synoptic form on procedures and on our experiences in applying raster scanning to various digitizing tasks. Emphasis is on the generation of digital topographic elevation models (DEM's) from contour lines and on the creation of digital polygon data. The hard- and software-system on which our experiences and this report are based, is called SysScan, manufactured and distributed by Messerschmitt-Boelkow-Blohm (MBB) of Munich (Germany) and by Kongsberg (Norway).
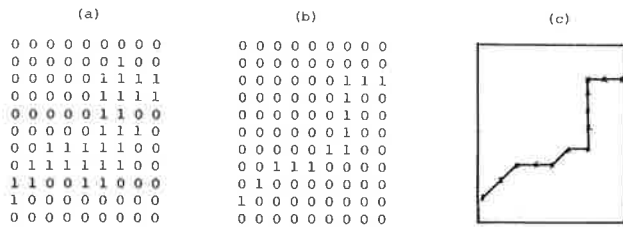

FIG. 1. Example of a manual digitizing table.

Fig. 2. Sketch explaining raster-to-vector conversion: (a) raster data structure, (b) skeletonizing as one example of the conversion, and (c) vector format.
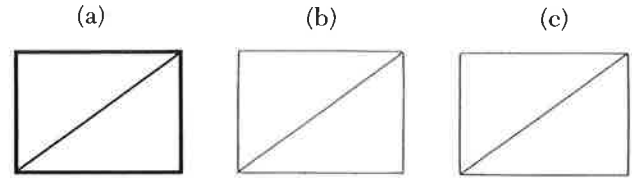

Fig. 3. Principle of manual correction of errors of the automatic process: (a) original analog data, (b) after automatic digitizing and vectorization. Note the problem at corner points (nodes), (c) after manual corrections of the nodes.

is placed (Figure 1). A cursor is moved manually over each point, P, to be digitized. Upon an operator's command, the x, y coordinates of P are sensed and recorded automatically. Pairs of points P_1, P_2 form vectors, and a sequence of vectors forms lines.

A bottleneck in digitizing is the load of required manual work. Different routes have been and are being pursued to speed up data acquisition and reduce cost. Among these are techniques to (a) aid the human operator in his effort to move the cursor quickly from point to point to generate vectors and lines ("aided track cursor," see, e.g., Hunka (1978)); (b) largely automate the movement of the cursor ("automated line following," see, e.g., Howman and Woodsford (1978)); and (c) acquire x, y coordinates in systematic scans of the map along parallel profiles to generate a rastered digital image of the map (raster scanning).

The data produced by scanning represent a digital raster data structure which is thoroughly different from the vector data structure produced in the various forms of manual or automated line following. It is currently a commonly established requirement to store, retrieve, manipulate, and analyze graphical data exclusively in vector format, although other data structures could be used meaningfully as well. Raster scanning is therefore only acceptable if it is followed by a computational

raster-to-vector conversion (see Figure 2). A comparison of the various digitizing techniques leads quickly to the conclusion that it is highly desirable to support the human operator by an automated process. If very large quantities of graphical data must be digitized, it remains a question, as to whether one should employ a technique of line following or of raster scanning together with raster-to-vector conversion. A recent position paper is by Boyle (1980), who clearly favors raster scanning due to superior flexibility and capability.

The digitization task consists of two different elements, i.e., acquisition of x, y coordinates and generation of associated descriptive data. It is fairly obvious that any conversion of map data to a meaningful digital format requires—in addition to machine intelligence—a certain amount of human intelligence. This could be needed to (a) correct machine errors, effects of noise, and ambiguities, as illustrated in Figure 3; (b) create topological relationships among digital data elements, as in Figure 4; (c) add to the graphical elements descriptive, labeling, tagging, and other data, as in Figure 5; or (d) perform the digitization task altogether, for example of alphanumeric data.

This fact leads to the conclusion that automation of the digitizing process is appropriate where it relieves the human from repetitive, tedious, and precise manual work without causing excessive
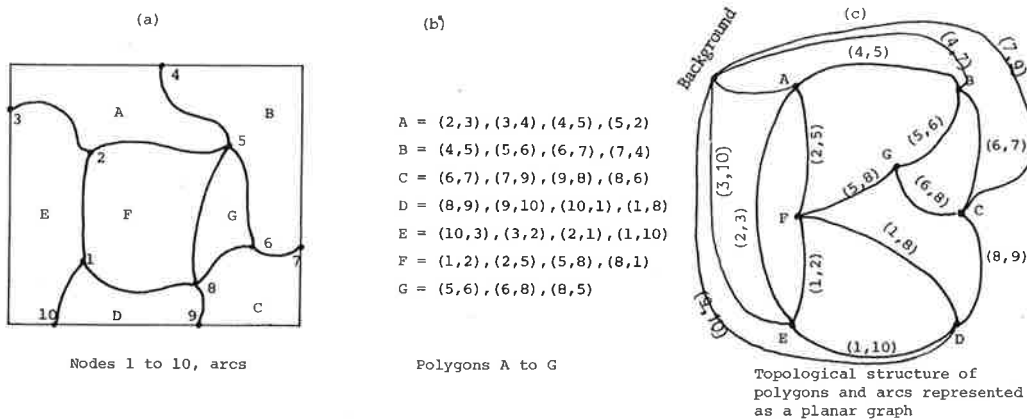

Fig. 4. Creation of a topological relationship among digital data elements as shown in (a) by combining arcs and nodes to polygons (see (b)) and generating a planar graph (c).
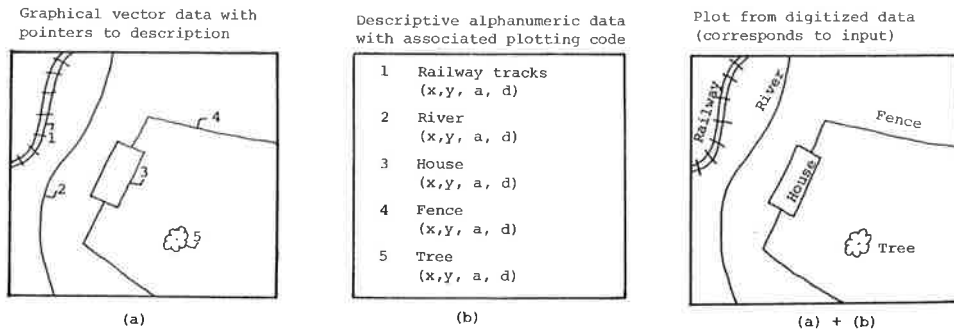
Graphical vector data with pointers to description

(a)

Descriptive alphanumeric data with associated plotting code

1	Railway tracks (x,y, a, d)
2	River (x,y, a, d)
3	House (x,y, a, d)
4	Fence (x,y, a, d)
5	Tree (x,y, a, d)

(b)

Plot from digitized data (corresponds to input)

(a) + (b)

FIG. 5. Sketchs illustrating the addition of descriptive information to the graphical elements. In Figure 5 (b), the descriptive data would have a pointer to the graphical elements (1, 2, . . ., 5), and a reference to position (x,y) and to orientation (a) as well as to scale (d) for plotting.

demands on required human participation. Applicability is thus assured when digitization implies large quantities of graphical (vector-type) data and a lesser amount of descriptive data.

The range of tasks to which automated digitization can be applied cannot be definitely delimited. It depends on both the performance of the automated process and the specifications for the ultimate product. Both these factors are greatly variable and subject to continuous development. However, as we will show in the remainder of this paper, raster scanning and digital data requirements have evolved today to a point where this technology can be very successfully applied to several typical and important tasks. We are certain that this range will enlarge in the future as techniques, in particular software and man-machine interactions, improve and requirements develop further.

THE DIGITIZATION SYSTEM SYSSCAN

OVERALL CONFIGURATION

Figure 6 is a diagram illustrating the components of SysScan as they are used in our work. Figure 7 presents a view of the KartoScan raster scanner and one interactive graphics work station, whereas Figure 1 already had shown a manual digitizing station that is part of SysScan. The raster scanner employs white light and a geometrically stable charge-coupled device (CCD) for scanning.

The hardware must be and is complemented by extensive software to convert raster scan data to vector format (vectorization). In SysScan there are several different options provided for this task, depending on the data type to be processed. Similarly, there are programs to support conventional manual digitizing. There is software to employ the interactive graphics work station to edit, tag, ma-
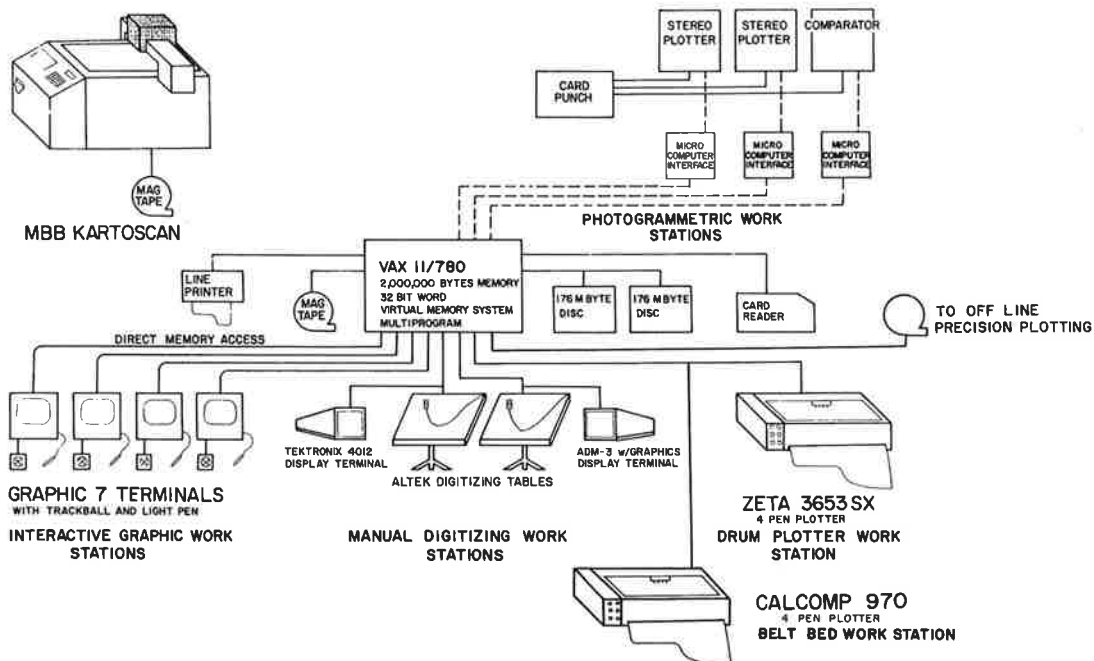

FIG. 6. Overall diagram of the SysScan configuration on which this report is based.

FIG. 7. KartoScan raster scanner in operation, recording output on magnetic tape.

nipulate, and group the digital data in various levels. Again, different programs are used for different types of data.

Features that are of particular importance are

- use of raster scanning with a CCD array for variable resolution and for white light, so that color separation is feasible and no rotating elements are required;

- use of refresh type, vector oriented (calligraphic) interactive graphics work stations for data editing and manipulation (Sanders' Graphic 7);
- incorporation of manual digitizing capabilities with storage tube graphics monitor;
- configuration around a large mid-size computer (in the current configuration a virtual memory, 32-bit VAX 11/780);
- modular, basic software consisting of parts where each one is optimized for a specific task; and
- various applications programs.

The technical data for the raster scanner are summarized in Table 1.

GEOMETRIC PERFORMANCE

For cartographic work, geometric performance of the digitizing equipment is of great significance. Figure 8 illustrates the result of a geometric evaluation of the scanner. A grid was measured first on a precise coordinatograph in four orientations and then scanned. Figure 8 presents the difference vectors in a grid of 600 by 1100 mm² after a four-parameter linear conformal transformation. These geometric errors amount to an RMS value of ± 16 μm in x and ± 20 μm in y. They are smaller than the instantaneous field-of-view of each CCD detector or the size of each smallest digital picture element (pixel)* in the raster.

QUALITY

In order to evaluate the overall digitizing performance of the automated raster scanning/vectorization process, a test sheet was employed

* "Pixel" is a standard expression from image processing and is a "picture element" of the scanner, or the instantaneous field of view of the single element of the CCD-array (scanner resolution).
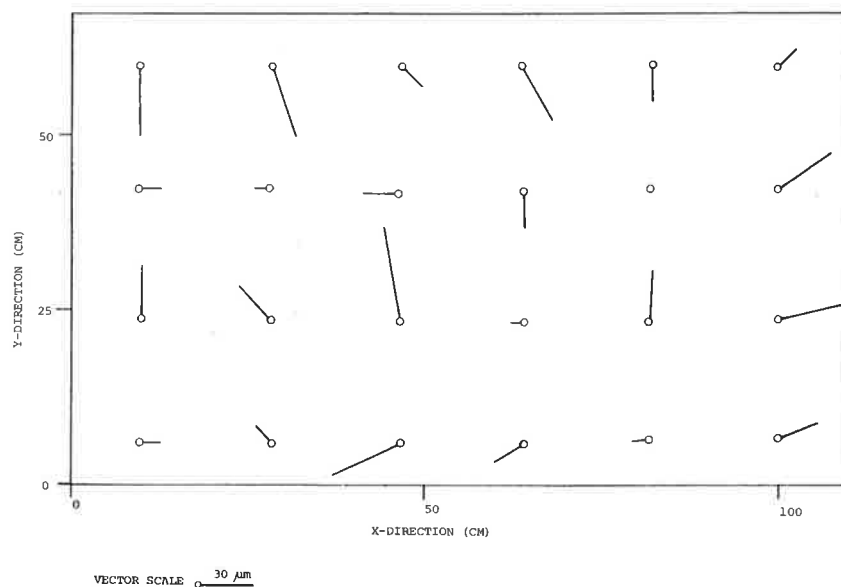

FIG. 8. Geometric performance of the KartoScan using a grid of about 900 by 550 mm². RMS error is ± 16 μm in x and ± 20 μm in y.

TABLE 1. SPECIFICATIONS OF THE KARTOSCAN RASTER SCANNER

Resolution (pixel size) (mm)	0.025	0.05	0.1	0.2
Width of scan strip (mm)	12.5	25	50	100
Scan velocity (mm/sec)	4	8	16	32
Scan area	600 mm × 1010 mm			
Alignment accuracy	± 10 μm			
Reproducibility	± 10 μm			
Spectral range	White light or use of different color filters			

(Figure 9). The quality of digitization is evident as a function of line width, line type, spacing, and curvature. No manual touch-up of the automatically acquired data took place. When analyzed in detail, this sheet confirms that digitized line axes are always within one pixel of the actual line center, so that at no occasion does the digitized line deviate from the original. Manual touch-up can be required at line intersections (nodes).

GENERATION OF DIGITAL ELEVATION MODELS FROM CONTOUR SHEETS

GENERAL

Early studies of raster scanning were in the context of digitization of contour lines for subsequent computation of square grid digital elevation models (DEM's), for example, at the U.S. Army Engineer Topographic Laboratories (Clark, 1980). However, operational, commercial DEM generation with raster scanning has been pioneered only recently using SysScan. An overall flow of operations is described in Figure 10.

DIGITIZATION AND ASSIGNMENT OF HEIGHTS

The procedure is based both on contour lines and on drainage-ridge data in order to strengthen the quality of the resulting DEM. The digitization is split into an automatic step of scanning/vectorizing and producing strings of x, y coordinates, and a separate interactive height tagging step adding z values to each x, y. Manual tagging of each digitized contour line can be supported by the computer. For example, each contour line need not be tagged separately, but a sequence of lines can be assigned a height using an auxiliary straight line on the refresh screen (Figure 11). The lowest and highest contour are assigned their height manually while intermediate contours that are intersected by the auxiliary line are assigned automatically. In this way one can avoid inconsistencies arising from intermediate auxiliary contour lines. Completely automated height tagging is conceptually feasible, but noise and perturbations in the data make human participation necessary.

The drainage-ridge lines remain unassigned; instead, they are automatically intersected with contour lines and therefore determine heights automatically, first at the intersection points and subsequently at all intermediate points by interpolation.

GRIDDING

Height gridding is done in a procedure that can
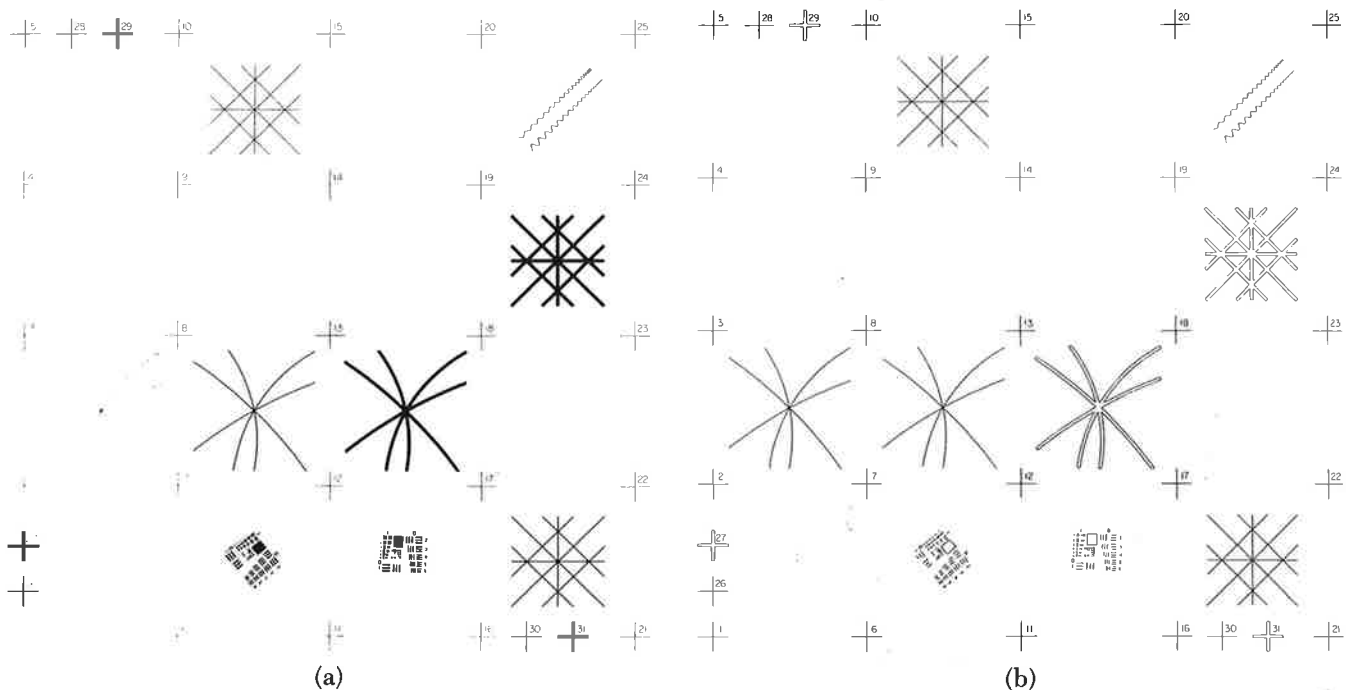

(a)

(b)

FIG. 9. Test sheet for quality evaluation of raster digitizing and vectorization. No manual work involved. Scanning resolution was 50 μm. (a) Input sheet; (b) Output sheet, where heavy lines are treated as areas so that the line boundaries are vectorized.
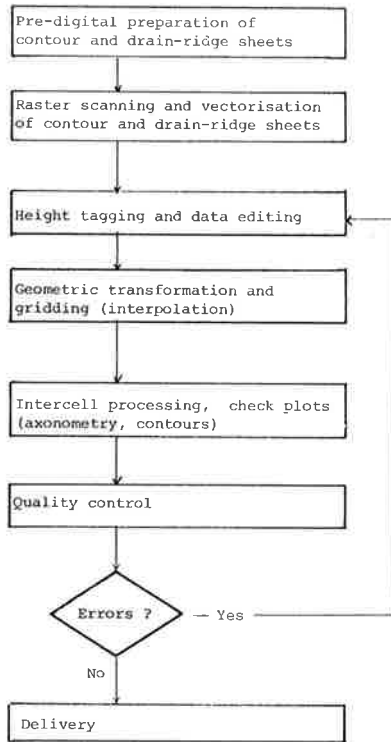
FIG. 10. Flow of operations for DEM generation using contours and drain-ridge lines.

HEIGHT H OF A POINT P:

$$H_P = \frac{H_1 - H_5}{(1,5)} (P,5) + H_5$$

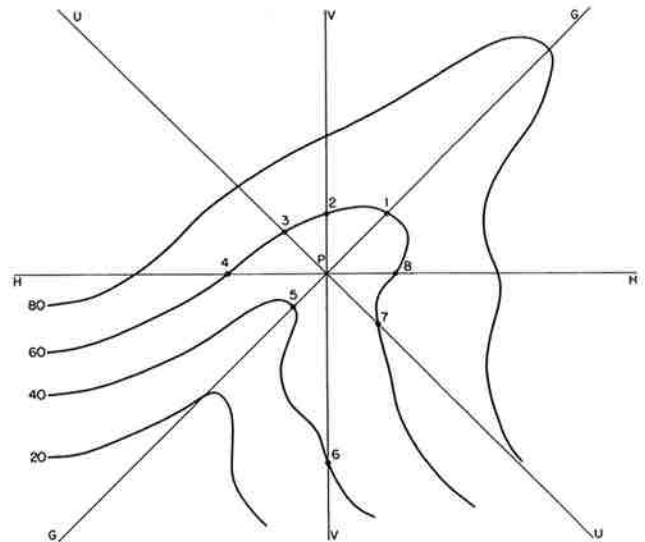SINCE SEARCH LINE (GG) IS STEEPEST AND INTER-SECTION POINTS 1 AND 5 ARE ON THE LINE OF STEEPEST DESCENT.

FIG. 12. Principle of gridding with contour lines, using eight search directions.

be described as a sequential steepest slope algorithm, modeled after manual interpolation with contour lines. One aims at linear interpolation of heights at each raster point along the line of steepest descent. This is achieved using a set of four or more search lines (eight or more search directions) as shown in Figure 12. In each point one selects the steepest line and finds a height through (linear) interpolation along that line, using its intersection with the nearest contour lines. A similar algorithm was proposed by Lauer (1972).

RESULTS

Figures 13, 14, and 15 illustrate the result of DEM generation using a dense contour sheet with

FIG. 11. Assigning heights to a number of contour lines in one step using a straight auxiliary line (locator). The lowest and highest contour get their height manually; intermediate contours are assigned automatically.

about 10,000 inches of line data. Figures 16, 17, 18, and 19 present an enlarged section of the same data to illustrate the original and vectorized data and then the gridded data presented in the form of re-interpolated contour lines, once with and once without the use of drain-ridge data.

We found the overall process to be very competitive when compared with previous manual methods of digitization. The use of drain-ridge data was found to be helpful to maintain the correct shape of certain flat valleys and gentle ridges. Overall discrepancies between manually and automatically determined heights of grid points amounted to RMS values up to ±2 m with a contour interval of 20 m and with auxiliary 10 m contours in flat areas. The discrepancy is within the error budget provided by the U.S. National Map Accuracy Standard (NMAS) for portraying the terrain in the original contours.

Typically, a sheet at scale 1:50,000 can be scanned and vectorized in about one hour of elapsed time in the current system configuration when employing a general purpose computer (no array processor) and is largely independent of the data density. Further throughput in subsequent steps is, however, greatly dependent on the density and quality of original analog data, operator performance, and accuracy requirements. With a configuration such as that of Figure 6, a routine overall throughput in excess of ten sheets per week at 1:50,000 scale has been maintained over long pe-
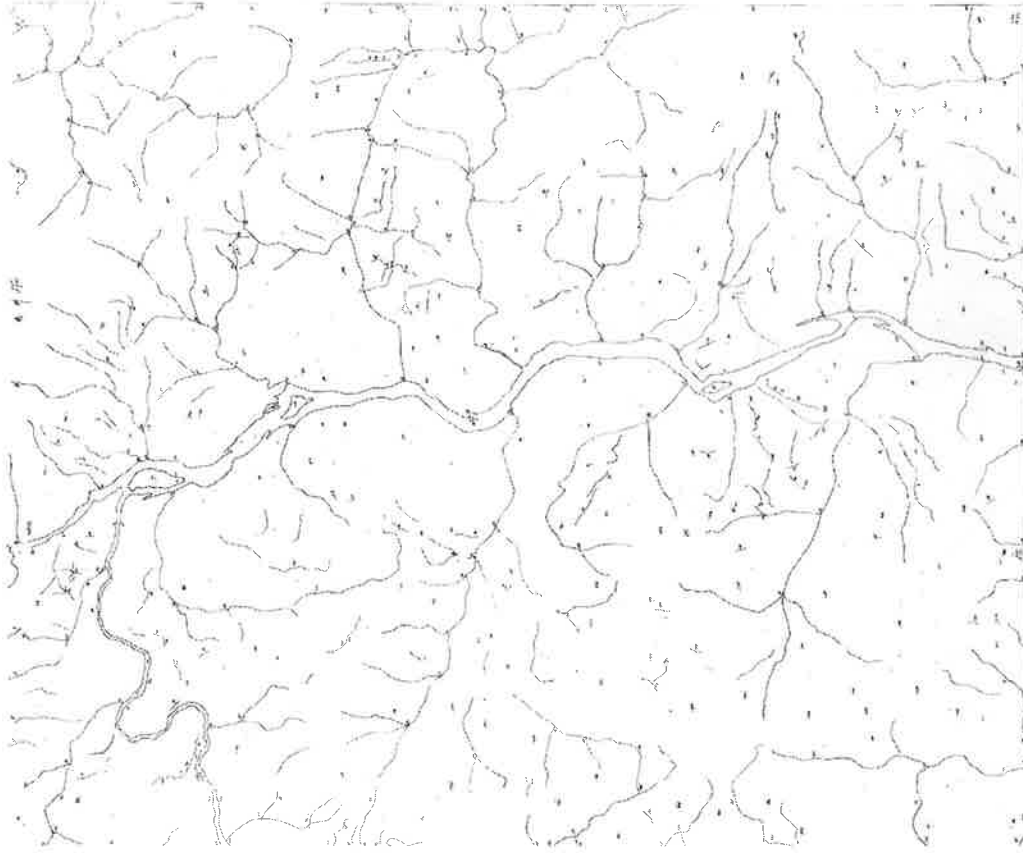
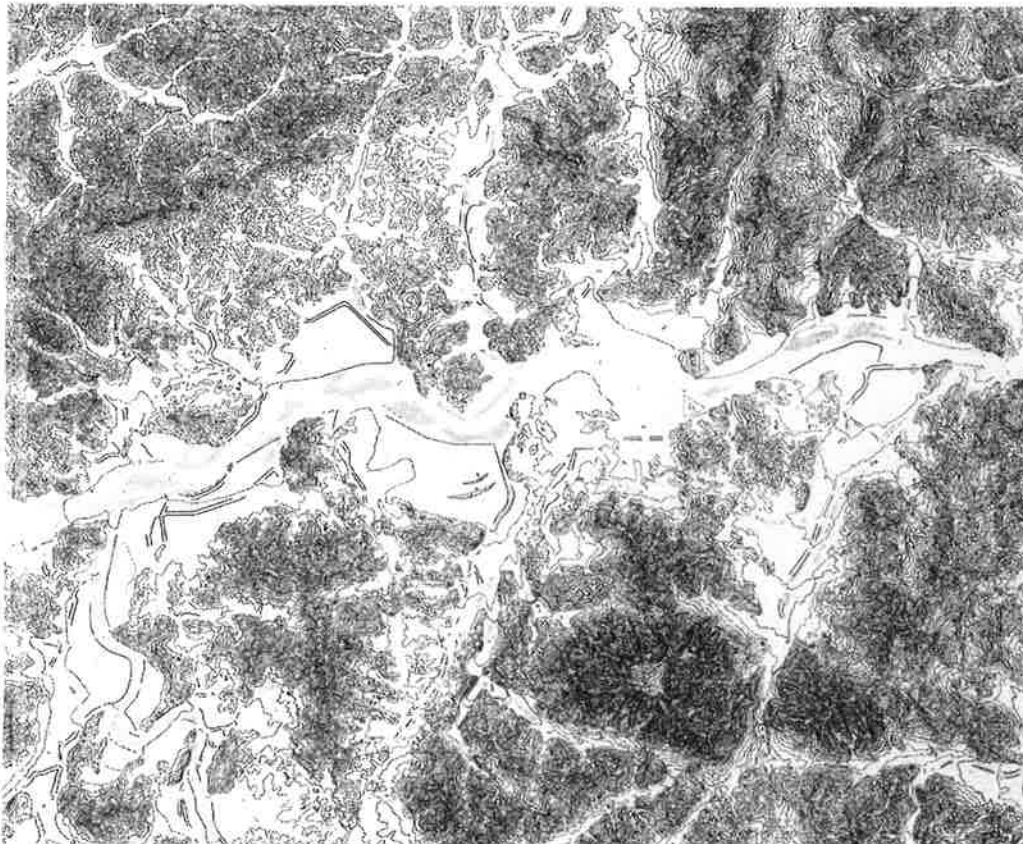FIG. 14. Drain-ridge line sheet associated with the contours of Figure 13.

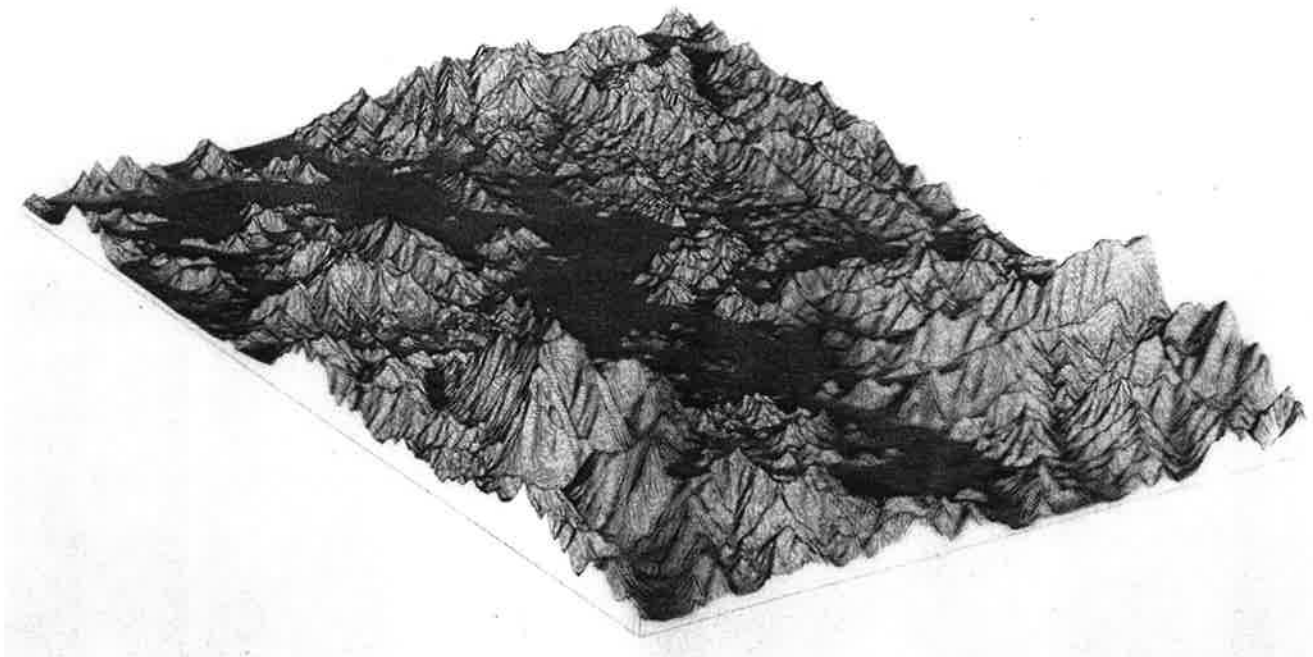FIG. 13. Example of a raw contour sheet originally at scale 1:50 000, 20 m contour interval.

FIG. 15.   Axonometric view of the height grid using the data of Figures 13 and 14.

riods of time when employing the production scheme shown in Figure 10.

GENERATION OF DIGITAL POLYGON DATA

GENERAL

Polygon data need to be converted to a digital format with great accuracy, because their use in-cludes the joint evaluation of separately digitized sheets with data of a different type. Such data can be, e.g., land use, political and property bound-aries, vegetation types, etc. Digitization is not al-lowed to contribute a significant component in the error budget for each graphical element. The de-mands in accuracy and the required density of data
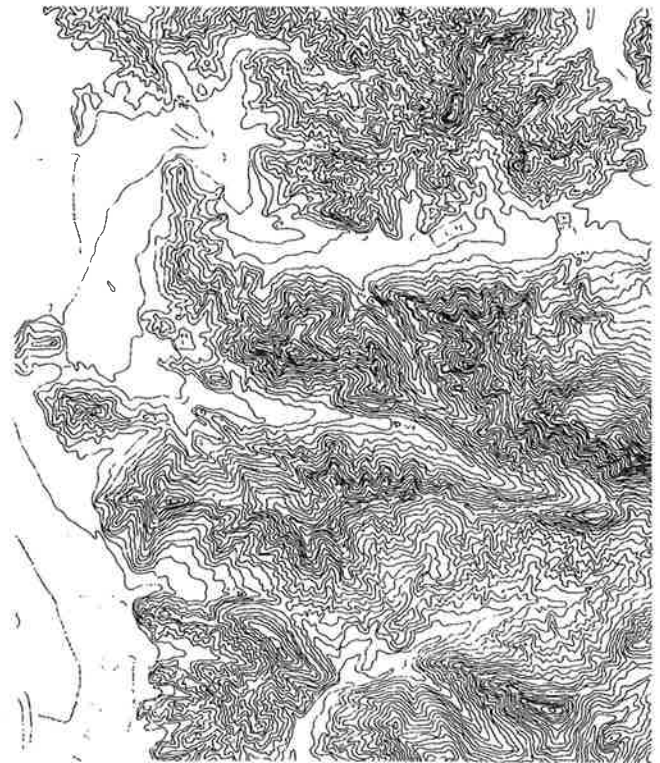

FIG. 16.   Detail of Figure 13.

FIG. 17.   Scanned and vectorized reproduction of the contents of Figure 16.

Fig. 19. Same as Figure 18, but based on contours together with drainridge lines.

Fig. 18. Quality control of the height grid, using a rapid simple contour interpolation routine. The grid was generated using only contours.
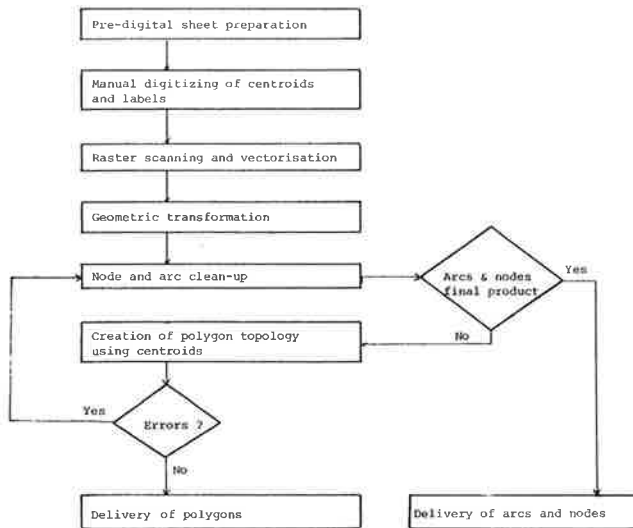
FIG. 20. Flow of operations for generation of digital polygon data.

(sometimes more than 25,000 arcs on a single sheet or more than 200 metres of lines; compare Mitchell *et al.* (1977)) make it particularly promising to automate the process, especially in view of the fact that the amount of descriptive data required is limited. Automation can avoid human errors and produce a specified accuracy.

Figure 20 describes one specific flow of operations for polygon digitization based on raster scanning.

DIGITIZATION AND ASSIGNMENT OF LABELS

We currently separate precision digitization on the KartoScan raster-scanner from the digitization of points within each polygon ("centroids"). The latter is done manually and does not require precision. Also, it includes the assignment of labels or tags.

The high precision polygon data are merged with the low precision centroids and subjected to automated generation of a data structure appropriate for interactive clean-up of ambiguities in the polygon nodes, and possible noise in the arc data*.

The result of the digitizing process consists of three files: an arc file, a node file, and a file of centroids with labels. These data can now be entered into any number of vector-oriented polygon data banks and evaluation systems, for example, into the U.S. Geological Survey's GIRAS (Mitchell *et al.*, 1977) or one of many other systems in existence today (Marble, 1980).

---

* "Arcs" are the line segments connecting adjacent polygon nodes; "nodes" are points at which arcs merge: at a node of order m we have m arcs merging (compare Figure 4).
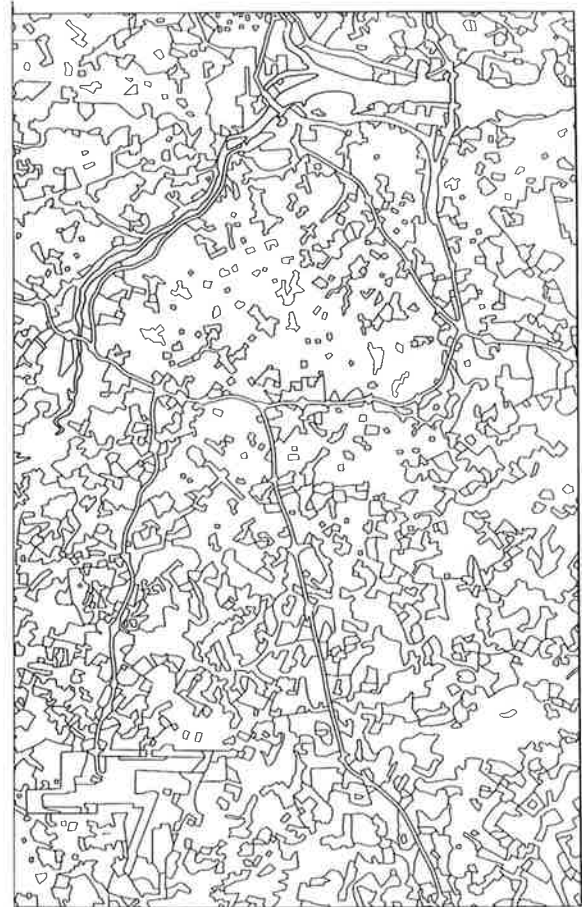

FIG. 21. Portion of an original, dense polygon sheet.

RESULTS

Figures 21 and 22 present a section of a dense polygon sheet, before and after raster scanning/vectorization. No manual work has been done to Figure 22. Manual effort would not be visible in a graphical presentation at this scale, since it could mainly affect small logical ambiguities in the node levels, but not, however, the appearance of arcs.

Labels normally appearing on a polygon sheet can be removed either in a predigital preparation step or during interactive data clean-up subsequent to digitization. Figures 21 and 22 were done after predigital label removal. Figure 23 shows an example where labels (Figure 23a) were digitized, vectorized (Figure 23b), and then removed (Figure 23c). The first alternative appears superior.

The accuracy of the digitized data is best described by the fact that everywhere the vectorized line arcs are within one pixel of the axis of the original lines. Therefore, no error is greater than half a line width or one pixel. This limit can only be violated if one were to smooth the data after vectorization, for example, with the aim to reduce the data volume.

FIG. 22. Scanned and vectorized data of Figure 21.

OTHER GRAPHICAL DATA

Automatic raster scanning offers also a capability to digitize other graphical data to an accuracy commensurate with the drawing's own dimensional stability. Where intensive human interpretation is required, manual digitizing is today the most common technique. Figure 24 presents an example of a portion of an original sample drawing manually generated. These data consist of comparatively simple geometric figures that need to be interpreted. One could for the purpose use raster scanning, but symbols or alphanumeric data would be presented in vector format; in the current example this is not meaningful. Therefore, symbol data have to be appropriately entered into the digital file by a manual method. The control plot (Figure 24b) is made from the digital data using the manually entered code.

We anticipate that raster scanning will be successful also for these and many other similar digitizing tasks, if appropriate interactive programs are developed. These must help to add in an efficient manner the intelligence that is needed.

CONCLUSION

A versatile hardware and software configuration for the acquisition of digital graphics data using either manual digitizing or automated raster scanning has been described. Typical applications of the automated scanning component are in the generation of digital elevation models from contour lines, of digital polygon data sets, and of other dense graphical data. Typical applications of the manual digitizing component are those for sparse
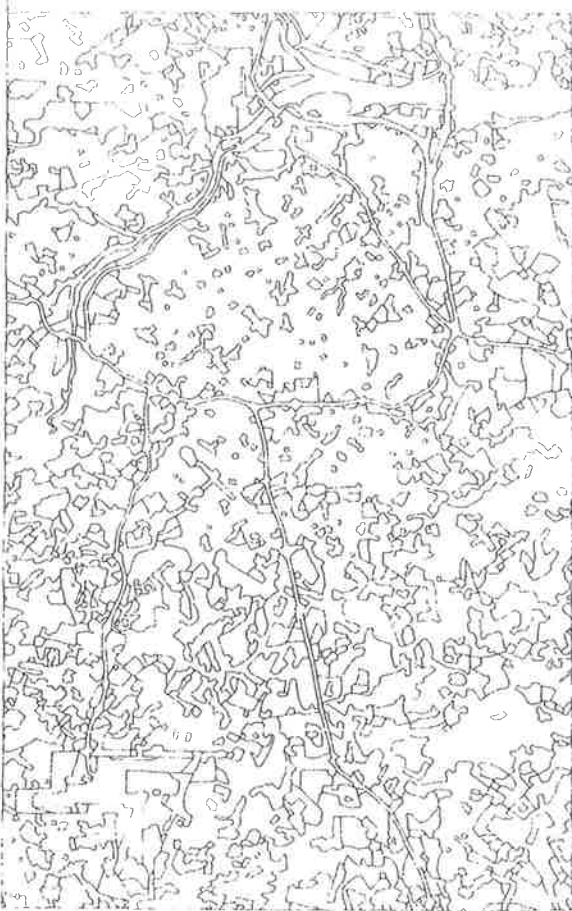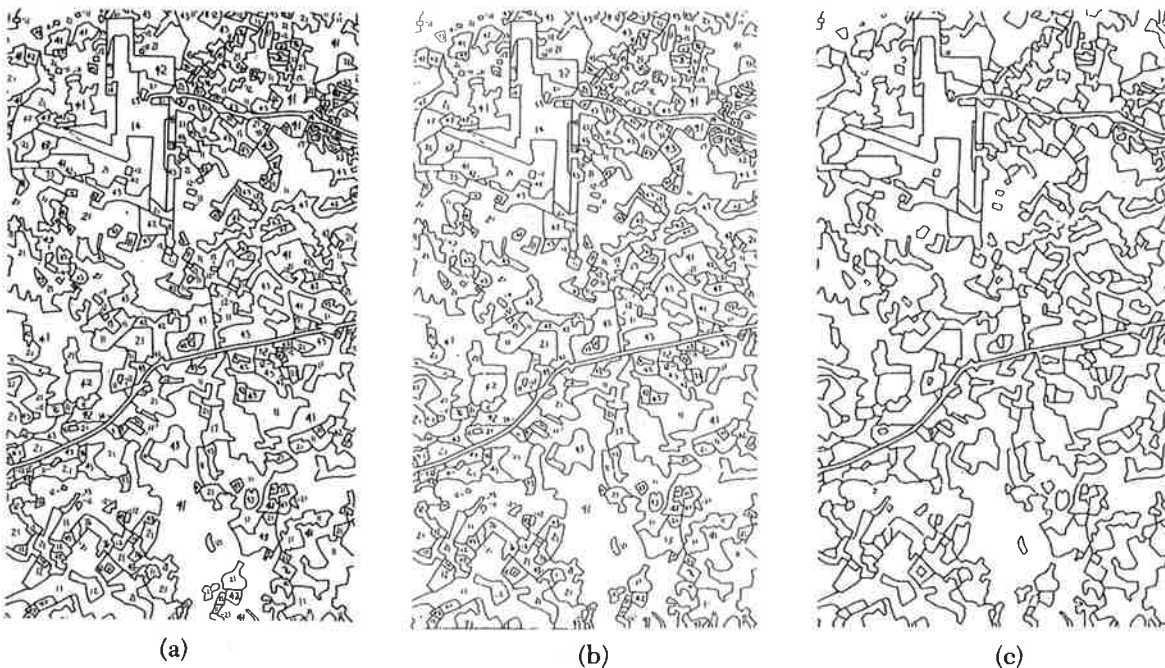

(a) (b) (c)

FIG. 23. (a) Portion of Figure 21, however with labels; (b) scanned and vectorized version of (a); and (c) after interactive clean-up of the data in (b).

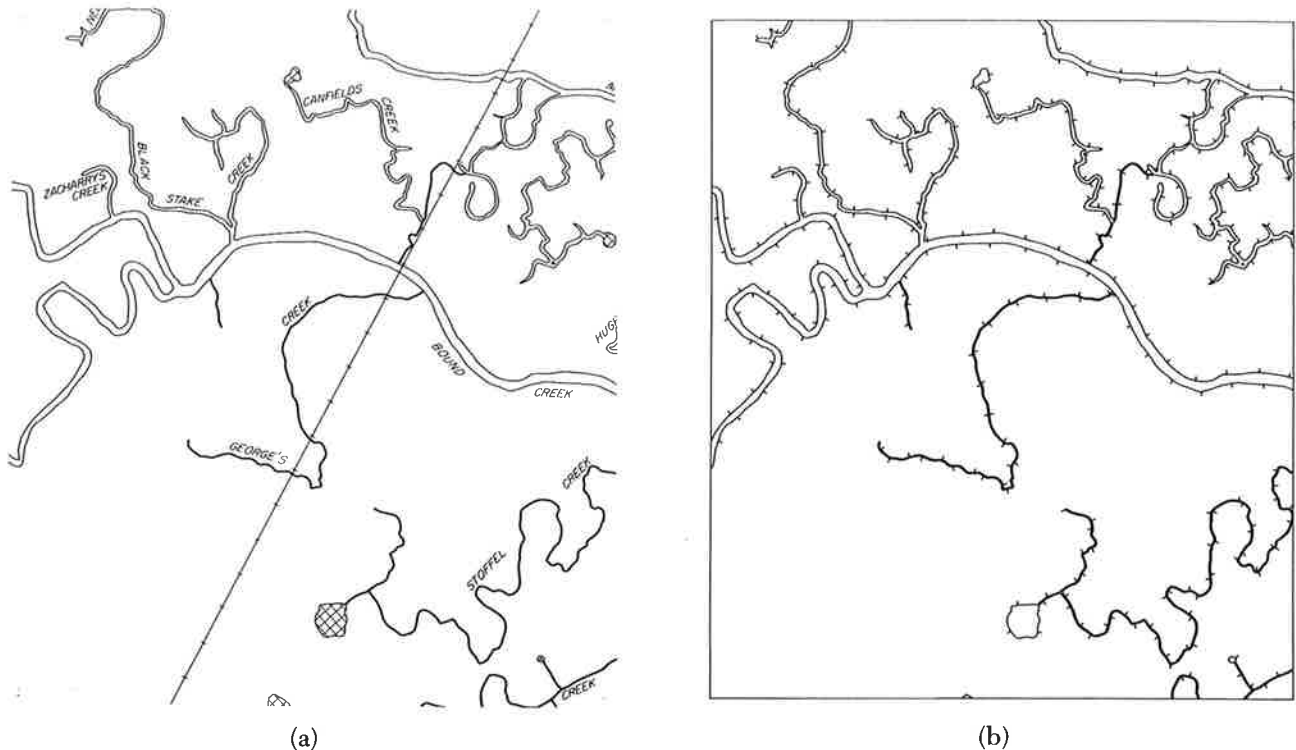(a)   (b)

FIG. 24.   Example of section of a special purpose, symbolised map, in (a) the original analog form on paper; and (b) manually digitized and reproduced on an *x,y* plotter.

data or graphical data with which a high percentage of descriptive information is associated.

Data acquisition, be it automatic or manual requires usually an interactive step of editing, labeling, and manipulating of the data. This is provided for in SysScan by using refresh type graphics work stations with track balls, light pens, and keyboards.

Performance of the entire system has been demonstrated in this paper for the creation of digital elevation models (DEM's) from contour lines, for dense polygon data, and for special purpose drawings. However, applications are not limited to these data. In fact, any type of digitization of graphical data can be performed, using either the automated scanning or the manual digitizing capability with interactive editing/labeling on the refresh-type screen. Also, digital data from sources other than maps or drawings can be acquired or compiled into a digital data set, e.g., from photogrammetric stereo models, or from a data base containing the elements of the graphical presentation (symbol table for computer aided map design). An essential limitation of the applicability of the digitized product relates to geometric accuracy: The digitized data are not more accurate than the input data.

With the increasing need for updated, correct graphical data in various presentations and with the concurrent trend toward increasingly affordable large computer memories and mass storage, we face today a much discussed growth of computer graphics tasks. This has led to a breakthrough of commercial applications for numerous related technologies.

This paper provides evidence that raster scanning and interactive refresh-type graphics screens are valuable tools in a versatile, universally applicable, and operational system for the acquisition of digital graphics data and for the support of graphical information systems.

REFERENCES

Boyle, A. R., 1980. The Present Status and Future of Scanning Methods for Digitization, Output Drafting and Interactive Display and Edit of Cartographic Data, *Archives, 14th Congress of ISP,* Hamburg, Germany, Comm. IV, pp. 92-99 (see ISP 1980).

Clark, R., 1980. Cartographic Raster Processing Programs at USAETL, *Proceedings, 40th Annual Meeting of the Am. Congress on Surveying and Mapping,* St. Louis, MO. U.S.A., March 1980.

Howman, C., and P. A. Woodsford, 1978. *The Laser-Scan Fastrack Automatic Digitizing System,* 9th Conference of the Intl. Cartographic Association (ICA), Maryland, U.S.A.

Hunka, G. W., 1978. Aided-Track Cursor for Improved Digitizing Accuracy, *Photogrammetric Engineering and Remote Sensing,* Vol. 44, No. 8, pp. 1061-1066.

ISP, 1980. *Archives of the International Society for Photogrammetry,* Vol. XXIII, 17th Congress, Hamburg, Germany FR, to be obtained from: Institute of

Photogrammetry, Univ. Bonn, D-5300 Bonn, Germany FR.

Lauer, S., 1972. Anwendung der skalaren Prädiktion auf das Problem des digitalen Geländemodells, *Nachrichten aus dem Karten- und Vermessungswesen*, Series 1, Nr. 51, Institute for Applied Geodesy, Frankfurt am Main, Germany FR.

Marbel, D., and others, 1980. *Computer Software for Spatial Handling;* Intern. Geographical Union, Comm. on Geographic Data Handling, 17, Kippewa Rd., Ottawa, Ontario K1S 3G3, Canada.

Mitchell W., S. C. Guptill, K. E. Anderson, R. G. Fegeas, and C. A. Hallam, 1977. *GIRAS: A Geographic Information Retrieval and Analysis System for Handling Land Use and Land Cover Data,* Geological Survey Professional Paper 1059, U.S. Government Printing Office, Washington D.C. 20402, 16 pages, Stock Nr. 024-001-03007-1.

Siggraph '80, 1980. *Proceedings of Siggraph '80,* Association for Computing Machinery, P.O. Box 12205, Church Street Station, New York, U.S.A.

New Sustaining Member

Martel Laboratories, Inc.

7100 30th Avenue North, St. Petersburg, FL 33710; (813) 345-0100

M ARTEL's Cartography Laboratory in St. Petersburg, Florida is a full service mapping production facility. It is engaged in remote sensing, mapping, and cartography. The company, founded in 1968, provides project planning, data acquisition and interpretation, and reduction through final manuscripts or proofs for maps and charts of all types.

Martel's professional staff of 200 is based in a 30,000 square foot facility in St. Petersburg, the center for its national and international mapping services.

The staff includes environmental analysts, geologists, geographers, photogrammetrists, foresters, soil scientists, biologists, zoologists, bathymetric compilers, planners and systems analysts, and graphic arts photographers and technicians. Personnel includes 22 with graduate degrees, 59 undergraduate degrees, 30 associate degrees, and 24 with specialized technical training.

Martel's thematic mapping experience includes demography, geology, geophysics, ecology, bathymetry, agronomy, soils, and forestry. Martel has also performed cadastral surveys, both site specific and regional. For engineering projects Martel performs mapping and profiling.

In computer oriented mapping, Martel provides the link of primary information for systems/hardware operations. These services include data transfer, digitizing, data entry and manipulation, and data acquisition covering low level, medium level, high level, and quad centered programs.

Aerial photography and interpretation include panchromatic, color, black-and-white infrared, and color infrared films. Aerial photography is obtained with the latest type of precision, high resolution, low distortion cartographic cameras. A subsidiary operates internationally.

Graphic arts services are undertaken in two fully equipped photographic laboratories. Equipment includes copy cameras, processors, and a full complement of support equipment.

Martel also operates analytical laboratories with a staff of 25 in Baltimore, Maryland. This work includes analyses of oil, water, liquid and solid wastes, hazardous wastes, and physical testing.

For professional consultation and planning of mapping needs, write Martel; 7100 30th Avenue North; St. Petersburg, Florida 33710; or call (813) 345-0100.