NATO Advanced Research Workshop "Mapping and Spatial Modelling for Navigation", Denmark, August 21-25, 1989

Tools for Geometric Data Acquisition and Maintenance

Wolf-Dieter Rase
Federal Research Institute
for Regional Planning (BfLR)
Postfach 20 01 30
D-5300 Bonn 2
Federal Republic of Germany

Abstract

The acquisition and maintenance of the geometric data base is an important cost factor for a geographical information system (GIS). The goal of the acquisition process is to encode and store the geometric features as accurately as necessary, and as economically as possible. Operator-guided digitizing and raster-to-vector conversion are the two techniques mainly applied for the capture of geometric data. Both techniques have specific advantages for certain classes of applications with respect to accuracy and economics. Although it is called an "old" technique, operator-guided digitizing can still be the most economical solution for many digitizing tasks. The economics can be improved by technical enhancements, for example the use of the acoustical channel, or by implementing the appropriate strategy for object capture. Additional tools beyond the basic storage and retrieval functions facilitate the management of geometric data.

1. Goals and objectives for geometric data acquisition

The geometric data base is the constituent part of any spatial or geographic information system (GIS). Features on or below the surface of the earth, including the form of the surface, are encoded in numerical form and stored in a computer system. Geometric data is mainly used in most GIS for graphic display, either of the features itself (cadastral, topographic and geological maps), or as a spatial referencing system for demographic, social, economic and ecological data. The geometric data base is also used for various analytical tasks, for example spatial allocation of services such as schools or fire stations, car navigation, route planning, optimal subdivision, zoning, and many others.

Each geometric object has three basic properties, which have to be encoded and stored in the data base:

- Geometry: the location of the object on the surface of the earth or on a map.

- Topology: the relation between the parts forming the object, or between the objects.

- Identification: the name or meaning of the object.

The data acquisition process for geometric objects is usually called "digitizing", although this term is valid for any kind of analog-to-digital conversion. Sometimes digitizing is mistaken for the capture of geometry alone. It should be stressed that the encoding of the topology and the identification is always included when the term "digitizing" is used for the whole geometric data acquisition process. The identification should not be mixed up with the attributes associated with the objects. It is recommended to maintain a strict separation between geometric data acquisition, and encoding of the attributes. The identification serves as the hook for attaching the attributes to the objects.

The cost of digitizing and updating the geometric data base is underestimated in most cases. Up to 90 percent of the overall cost of a GIS may account for the encoding and maintenance of the geometric objects (Goodchild 1982). Thus the techniques, methods and procedures to convert the visible or invisible features into computer-readable form is an important factor for the economics of a spatial information system. On one hand the objects in the data base should represent the features in the real world as closely as possible. On the other hand a finer resolution of the coordinates, for example by using a large-scale map, causes higher investment and production costs: the increased number of coordinates requires more computer resources, such as primary and secondary storage capacity, processing time, and sophisticated storage and retrieval software. Higher costs deteriorate the economics of GIS and hamper their application for financial reasons.

But in many cases the high accuracy is neither necessary nor justified, for example for small-scale thematic mapping. The accuracy requirements and the cost of the GIS have to be in adequate balance. Therefore the acquisition and maintenance of the geometric data base should be

- as accurate as necessary,

- as economical as possible.

The term "accuracy" usually means locational accuracy, how well the coordinates in the data base represent the positions in the real world. This meaning is too narrow: the accuracy of the topological relations and the identification must be included. In a database for international air traffic it is sufficient to encode an airport as a point. For regional planning purposes the parts of the airport, the runways and buildings, must be part of the description. The civil engineer at the airport requires a much finer resolution. He must know where all the pipelines, antennas, traffic guiding lights, and so on, are located. A similar concept is used in cartography. The maps of different scales do not only have different positional accuracy. Also different sets of rules are applied for the graphic transcription of the spatial features.

The seamless, scaleless geometric data base is the ultimate goal. From such a data base, digitized with the highest accuracy possible, all working data bases and maps on smaller scales can be derived (Guptill 1989). But the economical constraints for geometric data acquisition and processing, and many unsolved problems, for example automatic cartographic generalization, will keep alive the current scale-dependent solutions, at least for the foreseeable future.

As we will see, "modern" techniques for geometric data acquisition, such as raster-to-vector conversion, are not necessarily more economical than "old" techniques, such as operator-guided digitizing. It depends both on the requirements of the application and the price the users are willing to pay which digitizing and encoding method should be selected in a specific case.

2. Data structures, features, objects

For representation in a data base the real world has to be reduced to an appropriate level of complexity. Peuquet (1984) defines four layers of abstraction from the real world down to the storage in a computer system:

- Data reality: the features relevant for the needs of the specific application (scope of the data).

- Data model: the sets of components and relationships among the components to represent data reality.

- Data structure: definition of the data model for data processing.

- File structure: implementation of a data structure in a computer system.

The separation of the logical levels facilitates the conceptual definition and implementation of the geometric data base. It is possible to represent a data model by more than one data structure, and a data structure by more than one file structure. Data models and data structures can be considered as different sides of the same coin: the data model is the outside view, the one of the user or designer of the GIS; the data structure is the inside view, the one of the implementer or programmer. It is not by chance that data structures are dealt in computer science textbooks, not data models.

Sometimes the terms "entity", "feature", and "object" are used to describe the sequence of abstraction from reality to the data base. Basically an entity is a phenomenon in the real world which is not subdivided into phenomena of the same kind. A feature is an abstraction of the entity, for example the graphical transcription of a street on a map. An object is the representation of the feature in a data base. There is no room in this paper for further discussion on models and data structures. A short description how boundaries are stored in a geometric data base may suffice to demonstrate the principles (fig. 1).

The entity in the real world is the county. The feature on the map is an area bounded by a line. The object in the data base is a polygon defined by a sequence of coordinates. Some of the counties are direct neighbours with common boundaries. It would be a waste of memory space to store the common boundaries of neighbouring polyons twice in the data base. Thus all lines from node to node (also called segments or arcs) are stored only once, and form a line network. The polygons are defined by a list of numbers pointing to the lines in the network. Many other features can be represented as networks, for example streets, pipes, power lines, and also three-dimensional features, such as the surface of the earth. The principles explained for boundary networks apply analogously.

The data and file structure reflects another objective for data bases, the redundance-free storage of the objects. The topological relationships, the sequence and orientation of the lines of a polygon, are coded as pointers. From the topology in the network it can be derived, for example, which objects are direct neighbours. The experience in implementing large data bases, has shown, however, that complete redundance-free storage is too vulnerable in the case of coding errors, or hardware and software malfunctions. Safety straps should be implemented to enable self-repair, for example, when a pointer is coded wrongly, or is destroyed. The additional storage space for double pointers is necessary to ensure the consistency of the data base.

3. Technical realization of geometric data acquisition

The adequate technique to encode the entities is the geodetic survey, either from registrating terrestrial instruments, or by photogrammetry. In most cases, however, such a survey is too expensive for the purpose of geometric data acquisition for a GIS alone. Instead an existing and readily available database can be used, the map. A map is a database in its own right: the features are coded following an adopted standard for modelling. The difference to a computerized data base is the analog storage technology. Digitizing a map is just a kind of analog-to-digital conversion process. Two main groups of techniques are in common use for digitizing maps:

- Operator-guided (semi-automatic) digitizing,

- Raster-to-vector conversion.

Automatic line following has been applied to a smaller extent for digitzing maps. The operator sets a cursor on the beginning of a line. The device follows the line and registers its coordinates, until the end is reached. If a conflict situation is encountered, for example a X- or Y-crossing in the line, the operator is called for help. Line-following could not prove its superiority over the two other groups of techniques. The main reasons for discarding line following are the high cost for the devices, and the progress made with raster-to-vector conversion.

A new device has been introduced recently for encoding positions directly on the surface of the earth. The device called "Pathfinder" calculates its own position by processing the signals from navigational satellites. The accuracy is stated in the range of 2 to 5 meters, depending on the measuring time. The coordinates are recorded on a computer-readable medium. Tests done in cooperation with a GIS supplier produced promising results (ARC News 1989). The device is a good compromise in cases where surveying would be to expensive, and maps in an appropriate scale are not available. The use is hampered at the moment by the insufficient number of navigational satellites visible at a time at any place. It is also not clear how the device deals with the problems of signal shadowing and reflection in areas with high buildings and in mountainous terrain.

3.1 Operator-guided digitizing

The map sheet with the features to be encoded is mounted on the digitizing table. A human operator positions a pencil or a cursor with a crosshair on a point on the map. Triggered by the push of a button the digitizer senses the location of the cursor. The coordinates of the point are encoded as digital numbers. The numbers are sent to a computer for further processing and storage in primary or secondary memory.

Several technical alternatives for sensing the position of the cursor have been developed. In modern devices the cursor interchanges electromagnetic signals with a mesh of sensing wires (or several overlaid meshes) embedded in the surface of the table. The exact location is derived from form and timing of the signals. Technical principles and design details differ from manufacturer to manufacturer. For the regular user the quality of the digitizer (accuracy and stability) seems to be more dependent on the quality of the production process and the fine-tuning in the factory than on the technical principle used.

Some digitizers provide a "stream mode" besides the regular point mode to facilitate the acquisition of lines. The operator moves the cursor continously along the line. The capture of the coordinates is triggered automatically at preselected intervals. The interval can be a time distance, a distance along the line, or along the axes. The positional accuracy of the crosshair during line following is influenced how smooth the cursor glides on the surface, and by the mass of the cursor. Digitizers for cartographic work are mostly equipped with a magnifying glass to improve the positioning accuracy of the crosshair, and a keypad for function input. The cursor's mass hampers the mobility for line tracing, and unwanted artifacts, such as jitter and spikes, may occur in the digitized line. The irregularities can be filtered out by software, but it is nearly impossible for a program to discriminate between artifacts and real data.

One manufacturer tries to overcome the line following problem by a cursor with a built-in miniature TV camera (Poiker 1983). The operator has to position the cursor only roughly on the line. The relative center coordinates of the line are extracted from the raster image provided by the camera. The absolute line coordinates are computed by adding the relative coordinates to the position of the cursor. The obvious advantages of the device are counterbalanced by very stringent specifications for the line width, the graphic quality of the lines, and the relatively high price, which is higher than the price of the plain digitizer itself.

3.2 Raster-to-vector conversion

In a first step the map sheet is converted into a digital raster image with appropriate resolution. On a drum scanner the map is mounted onto a rotating drum. A sensor or an array of sensors moves in small steps along the drum and samples the light intensity of the sheet at regular intervals during each turn. The intensity values are stored as a regular grid of picture elements (pixels) for further processing. On a flatbed scanner the map is mounted stationary, and the sensor is moved in both directions over the map by an x-y mechanism.

After rasterizing a specialized pattern recognition program extracts the geometric elements from the raster image. The lines in the image, having a certain width, are thinned to a one-dimensional line by finding the central pixels. That process is also know as "skeletonizing". The absolute line coordinates are computed from the grid positions of the central line pixels (for a complete summary of hardware and software for raster-to-vector conversion see Peuquet and Boyle, 1984).

Dependent on the software used and the requirements of the application additional geometric elements can be extracted from the raster image, for example the position of nodes (where lines end and more than two lines meet), numbers, or cartographic symbols. Only a limited amount of information on topological properties is available, for example the coordinates of line ends and nodes. This information usually suffices to derive the topology of a boundary network, provided the lines are error-free or close to error-free. More complex relationships and object identifications must be added separately, in most cases in an interactive session called "tagging".

3.3 Comparison of the technical alternatives

Table 1 gives an overview of the some aspects of geometric data acquisition with respect to operator-guided digitizing and raster-to-vector conversion.

Geometric accuracy. The overall geometric accuracy with operator-guided digitizing is restricted by the technical accuracy of the device, and the ability of the human operator to position the crosshair. The accuracy should not be mixed up with the resolution, which is usually in the range from 0.01 to 0.025 mm. An accuracy of 0.1 mm should be the minimum for cartographic applications. Sometimes it is not quite clear if the accuracy figure stated by the manufacturer is an average over the whole table, or a maximum value at any point. Technical inaccuracies are caused mainly by inevitable instabilities during manufacturing, errors during measurement and analog-digital conversion induced by electromagnetic noise, the finite resolution of digital numbers and the resulting inaccuracy of transcendental functions, especially at small angles, variations in the environmental conditions (temperature, humidity), and the instability of the medium on which the map is drafted. The humidity from the hands positioning the cursor causes deformations of the paper. Thus it is advisible to copy the map on a stable foil, not on paper.

The accuracy in positioning is limited by the optical arrangement of the cursor and the capabilities of the human eye/brain system. The parallactic deviation between the crosshair lines and the point on the map accounts for some error, especially if the map is covered with a protectional layer or a transparent "scratchpad" sheet. The resolution of the human eye is about 0.1 mm, and may vary considerably from person to person, as well as the ability to estimate the central position of a broader line or a symbol. All these errors may add up to the worst case. Experience has shown that it is reasonable to expect deviations between 0.2 and 0.3 mm under the best conditions for hardware, software and personnel. A good means to reduce the positional error is to digitize from a photographically enlarged map sheet and scale down the coordinates numerically to the scale wanted. The size of the sheet and the table, however, sets the limits where this approach is applicable.

The limiting factor for the geometric accuracy of raster-to-vector conversion is mainly the mechanical stability of the scanner. Of course the proper adjustment and transformation methods must be applied, but this is also valid for operator-guided digitizing. Although a finer resolution is technically feasible a raster width of about 0.05 mm is a good compromise. A higher resolution will increase processing time for rasterization and line extraction. The increase in accuracy is only marginal due to other sources of error, such as the mechanical tolerances, or the instabilities of the medium.

Encoding the object properties. The human eye/brain system is generally far superior to any program for pattern recognition. The operator at the digitizing table is able to recognize and extract the geometry and the topological features of the geometric objects, even in a complicated map scene. The capture of the geometry can be combined with the encoding of the topology and the identifications of the objects in a single operation.

Raster-vector conversion, however, usually requires several stages. The programs for line extraction and thinning are not able to sort out tiny errors in the map a human would not even notice. Examples for such errors are very small holes in the lines, dust particles, cracks in the ink layer or in the medium, blurred line edges, or variations in line width and gray level. Microscopic holes in lines (caused by dust particles on the negative) become tiny islands enclosed by two artificial lines. A fine crack, invisible without a magnifying glass, interrupts the line, and destroys its inherent topological structure. Dependent on the algorithm used line-thinning may also generate artifacts, such as small splinters and spikes, or unwanted bents and deviations at line crossings. Some of the errors can be found and corrected automatically by knowledge-based systems (Pernot 1986). Other errors need human intervention for correction, and many remain undetected forever.

Because only a limited amount of topological information can be derived from the raster image more complex topology and the identification of the polygons in the network must be added in a separate step. The latter can be done, for example, by pointing interactively to the polyon, or by geometric matching of the polygons with coordinates digitized manually, and associated with the polygon key.

Device cost. As a sequence of the general price reduction for electronic equipment digitizer tables are available in the range from 5,000 to 10,000 dollars, including an embedded microcomputer for the basic functions, such as scaling, rotation, data reduction, and communication to a host computer. A scanner on the other hand is a high-precision mechanical device, which has its price. Together with a controller able to manage the stream of pixel data, the host computer, the raster-to-vector conversion software, and the interactive system for correction, editing and tagging the raw data, the price lies in the range from 200,000 to over a million dollars. The inexpensive scanners now available for desk top publishing provide the resolution necessary for raster-to-vector conversion, at least in the upper price range. The problem is their small format which is not sufficient for most cartographic requirements.

Labour cost. At first sight the labour cost for operator-guided digitizing is much higher than for raster-to-vector conversion. This is true if only the capture of geometry is taken into account. The cursor must be positioned on every single point, or the line must be followed very carefully. This is a slow and tedious process. For the raster-to-vector conversion no operator is required. The line geometry is extracted without manual intervention.

In fact the cost for the capture of geometry is only a part of the overall labour cost. Raster-to-vector conversion requires a map sheet with nearly perfect linework. Otherwise too many errors and artifacts are introduced during line extraction. These errors must be detected and corrected interactively. Therefore it is wise to invest labour prior to rasterizing to provide a clean map sheet. If only selected features are to be encoded from a map, a new drawing with these features only should be produced. The elimination of unwanted or unnecessary objects may cost as much as the digitization itself. But if a new map must be drawn it might be more economical to invest the labour directly in operator-guided line capture.

If the map is already coded as a raster programs for pattern recognition can be applied to extract more features than the lines only. For example in German cadastral maps the main use of buildings is encoded as the angle of the hachure lines in relation to the base line of the building. A pattern recognition program is able to find the hachures by recognizing parallels, and to calculate the angle of the parallels and the base line. A first implementation of the search for parallels took five hours of computing time on a DEC Microvax II computer for one cadastral map. With some improvements the time could be reduced to two hours (Illert 1988). On the other hand an operator is able, after a short period of customization, to encode the same information by pointing at the polygon and pressing a button on the cursor in much less time, let's say, in a quarter of an hour. This example illustrates again that the high-tech solution is no always the most economical one. The picture may change with much faster and cheaper computers, of course.

Main application areas. From this overview it can be roughly estimated which of the two possibilities has an economic advantage over the other in specific cases. Raster-to-vector conversion should be applied if the objects consist of very complex lines with many coordinates, and for applications with high accuracy requirements. This is valid for cadastral and topographic maps, for hydrographic and nautical charts. In contrast thematic maps usually have lower accuracy requirements, and the lines are simpler. It is more economical then to use operator-guided digitizing.

The picture becomes blurred for cases where a new map has to be drafted to extract specific features. Sometimes the wanted features, for example contour lines, are already on a separate color sheet used for printing. Only minor finishing is required to provide a clean map, such as closing the holes left in the contours by the height figures. If complete redrafting is necessary operator-guided digitizing can be the less expensive way, even for complex linework.

A certain threshold in the volume of geometric data acquisition must be reached to recover the high investment costs for the hardware and software of a raster-to-vector conversion system. If this threshold is not met operator-guided digitizing is still more economical, even if the labour cost is higher. The economics can be improved by contracting out the line capture by raster-to-vector conversion, and adding the topology and identifications in-house. For the latter step a digitizer and associated equipment is required, also for the maintenance of the geometric data base. If the hardware is necessary anyway, it might be more economical to do the whole job in-house.

An example. The BfLR is the editor of a map of the community boundaries in the Federal Republic of Germany in the scale 1 : 300 000 on 28 map sheets with about 9500 polygons. The map has been digitized by a software/consulting firm. The raster-to-vector conversion was contracted out, and the error correction, topology construction and tagging was done interactively with its own computer system and personnel. The labour cost alone was estimated to be about two man-years (Hansen 1988). The cost for the line extraction (scanning, thinning, vectorization) could not be reconstructed, as well as the cost for computer use to do the postprocessing, because the firm went out of business some time ago.

The same set of map sheets was digitized again some years later by another firm, this time by operator-guided digitizing. The overall locational accuracy, estimated by visual comparison with the original, was even better than in the previous data set. A test with a syntax-checking program revealed also much less topological errors. The labour cost summed up to 4 man-months (Tappert 1988). The cost for the hardware and computer time was certainly lower than for the cost of the raster-to-vector conversion.

The example shows that raster-to-vector conversion is not necessarily superior to the "old" technique of operator-guided digitizing. On the other hand it is no proof that raster-to-vector conversion is less economical in general. The example demonstrates also that comparisons are always a dangerous field. For example the algorithms for transformation and abutment of the different sheets after line extraction were wrong or improperly applied, resulting in large displacements at certain parts of the data base. The original data was already purged when the error was detected. The blame is not on the method of raster-to-vector conversion for providing wrong results, but on the programmers for lack of knowledge, and the project leaders for organizational blunder.

The example is one of the very rare cases where the same data has been digitized twice. Normally comparisons rely only on estimates of the method not used. It is understandable that sometimes important cost factors are neglected or forgotten, such as the encoding of topology, or tagging. Thus is it advisable to look twice at figures given in the literature describing the advantage of one over the other technique. Because so many factors have to be taken into account, the decision in favour of one solution has to be made for each application again.

4. Improvements for operator-guided digitizing

Operator-guided digitizing is called quite often the "old" technique. The term "old" is correct in the sense of the historical sequence of development. It is not correct, as we have seen, in the sense of "obsolete". Manual digitizing still has its merits: in many cases it is still more economical than raster-to-vector conversion, even with the high labour cost involved. Digitizer tables are needed in any case for tagging, correction and maintenance of the geometric data base. There is still enough room for improvements in operator-guided digitizing to raise the quality and the economics, some fostered by the development of electronics and computer technology, some by taking alternative strategies for the acquisition process.

4.1 Technical enhancements

The typical configuration of a workstation for geometric data acquisition consists of the digitizer table, a graphic display terminal, and sometimes an additional alphanumerical terminal. The tables used for cartographic purposes come mostly with a cursor equipped with a magnifying glass and an array of pushbuttons. The height and tilt of the table is variable to allow adjustment to the requirements of the operator. Backlighting of the table is recommended. The backlighting should be dimmable to provide the most appropriate light conditions. A color display is preferred over a monochrome display. The alphanumeric terminal and its keyboard are used for command input and messages. The workstation, with or without a certain amount of local intelligence, is connected to a computer system which controls the station and administrates the data.

Ergonomics. Achieving a low error rate is one of the principal goals for geometric data acquisition, because errors in the data base are hard to detect and costly to correct. The error rate during the digitizing session is highly correlated to the amount of stress the operator is exposed to. One stress factor is the muscular pain caused by frequent and unusual hand, arm, head, and eye movements. Although muscular activity is necessary to do the job, the frequency of movements can be reduced. The operator should keep his eyes on the area of the map where he is doing coordinate capture. If operator is forced, for example, to position the cursor on a menu to select an input option, the arm has to be moved, the head turned, and the focus of the eyes changed. By changing the field of view the visual contact to the digitizing area is lost. After returning to the digitizing area the object or the line must be found again. Menu picking is a valuable tool for interaction with a computer, but on a digitizer table used for cartographic applications it is a pain in the neck, quite literally.

The array of pushbuttons available on most high-precision cursors is a more convenient alternative for signaling a function to the computer than a menu. If properly arranged, and after some accustoming the buttons are pressed blindly. To avoid unusual and uncomfortable body movements the size of the map and the table should not exceed certain limits. Beside the muscular stress the parallactic deviations become larger, and the accuracy diminishes considerably. Dividing the map and numerical abutment (sewing the parts together) is more adequate. Another possibility is the aforementioned cursor equipped with a camera and line extraction electronics. It reduces the effort for fine-positioning, especially on the eyes.

Because digitizing requires a high degree of mental concentration it is self-evident that the workstation should be located in a comfortable environment, with respect to noise, climate, illumination, and disturbances. The somatic and mental stress should also cause managers and project-leaders to reconsider the digitizing procedures. In many organizations digitizing is organized like an assembly line, with low-paid clerks, mostly women, for the line capture, and employees on higher salary levels for error correction, topology encoding and tagging. It is more economical if these activitities are united in one hand. The stress on muscles and eyes is reduced by the possibility to alternate between the tiring line capture and the other activities. The motivation for providing better quality is boosted by the responsibility for the whole product, as we know from the example of a Swedish car manufacturer who dropped the assembly line in one factory.

Sound and speech. The use of the acoustic channel is the natural means for messages from the computer to the operator tied up with visual interaction (Rase 1984). The bell of the terminal is usually the only acoustical signal used to get the attention of the operator. Inexpensive PSGs (programmable sound generator) connected to the computer provide a variety of sounds under program control. They convey short messages, for example a request to input a coordinate, a receipt for a successful match operation, or warnings on different levels of severity. The variable sounds can also be used to signal internal activities of the system, not only during digitizing. Sometimes the computer is busy for a longer period without any external sign of activitiy, for example reading a large file. Especially the casual users get nervous because apparently nothing happens. They try to figure out what's going on, start to hit keys on the keyboard, and might ruin the running job. Instead of displaying a message from time to time, such as "still working", a sound could be issued at a regular interval to signal that the program is still executing properly.

Certain messages are too complex to be coded as single sound signals. In digitizing systems with strong operator guidance (see next chapters) the operator is told which object is to be digitized next. Displaying the name of the object on the terminal requires the operator to change the field of vision, with all the problems described above. A small display within the field of vision is already a great improvement. With computer-generated speech the names can be read to the operator in natural language under program control. For some years now inexpensive devices are available producing quite legible speech output. The natural language is approximated by a sequence of the most common sounds called phonemes. The phoneme codes are transmitted to the device which outputs the selected sound patterns with some interpolation between the phonemes. The base frequency (male or female voice), the relative frequency and pitch (language melody), the interpolating function, and other parameters can be influenced under program control.

The problem with speech output is the translation from written text to phonemes. In some languages, for example in French, Spanish, Italian, and Dutch, the translation rules are relatively simple because the written text is a more or less phonetic transcript of the spoken language. In others, such as English and German, the rules are either complicated, or have so many exceptions that the use of phoneme libraries for the most common words is mandatory to achieve good speech quality. What can be done already with a microcomputer and 8 kbyte of memory demonstrates the Microvox device, unfortunately with English text-to-speech translation only (Ciarcia 1982). The speech synthesizer can also be used for transmitting error messages. It proved to be practical to issue warnings with the PSG, and severe errors in natural language.

The opposite direction in speech interaction, transmitting commands and messages to the program by computer-controlled input of speech, can be realized as well. The voice input is a pattern-matching operation. The device has to "learn" the sound patterns of a specific speaker. When the sound pattern is entered during a session, it is compared with the stored pattern. If a certain level of similarity is recognized the command is accepted and processed by the program. Speech input imposes some problems, for example speaker dependance, sensitivity to environmental noise, uncomfortable headset with microphone. Thus the keypad on the cursor is a better alternative to speech input in most cases.

4.2 Strategies for digitizing boundary networks

Improvements in the ergonomics of operator-guided digitizing by technical enhancements will increase the qualitity of the result, and subsequently, the economics of geometric data acquisition. The procedure of the data acquisition, how and in what sequence the geometric and topological data is encoded, has also a certain influence on the cost for the acquisition of the data base. The system designers should always keep in mind that the users of GIS are usually not interested in data structures. Users are primarily interested to retrieve and display spatial objects, not how these objects are coded and stored in the data base. This should be taken into account for the implementation of a digitizing system. Different approaches have been tried for geometric data acquisition. The most common strategies for the acquisition of line (boundary) networks are presented in short.

Spaghetti approach: The points and lines of the map are digitized straightforward, without any regard to objects, data structures, or topology. In the following step a batch-type program checks all lines for intersections, calculates the intersection coordinates, and generates a new set of lines. The topology of the network is derived from the new lines and nodes (Douglas 1989). The identification of the polygons is added, either interactively, or by the "meatball" technique: the polygons are matched geometrically with points supposed to be located inside a polygon, and bearing a name (Chrisman 1986).

The advantage of the approach is the fact that inexpensive (dumb) equipment can be used, and the labour cost for line capture geometry is low. Line intersection, however, is both a costly and intricate process. Because of the inevitable inaccuracies during digitizing the coordinates of different lines meeting at a node are never consistent. The program tries to fix the inaccuracies by setting a circular tolerance for a node. If the catch circle is too small, the larger errors are not caught. If it is too large, neighbouring nodes are included as well. If the operator has made a severe mistake, for example he has forgotten a line, the map has to be returned to the digitizer. Besides the problems with readjustment of the map on the digitizer the postprocessing step has to be repeated. The opposite case, lines digitized twice, cause problems as well. Because of the inaccuracies mentioned the same lines, although logically identical, have seldom identical coordinate values. Comparing the coordinates is not sufficient to detect the similarity. Algorithms for line matching have been developed, but are rather complex and time-consuming (Becker and Ottmann

1988). Line intersection has its own problems due to the finite resolution of digital numbers (Franklin 1984).

Segment approach: To avoid the processing time and the problems with line intersection the segments (arcs) of the network are digitized. In the basic version matching of the node coordinates (line ends), topology construction and identification is done after line capture. On the next level, possible by more powerful computers, the nodes are matched during digitizing. If a coordinate is outside the catch circle, a warning is issued to the operator. The topology can be added "on the fly". The segments of a polygon are digitized in the same sequence as they appear in the data structure. If a segment has been digitized already for a previous polygon the program completes the actual polygon by searching through the linework until the end node for the polygon is found (Türke 1976).

The comfort can be further enhanced by constant plausibility checks and rigid operator guidance. For example the names of the polygons are stored in a file. The operator is told which polygon is to be digitized next. After completion of the polygon the name is added to the data structure. Lines and polygons are displayed on a terminal for casual visual inspections. The display window can be changed, either to get an overview, or to look at a small part of the map. Computer-controlled sound generation and speech output have been successfully implemented and used in a system for acquisition of boundary networks (Rase 1984).

Polygon approach: Although the segment approach is an improvement over the spaghetti approach there are still some unsolved problems. If the operator, for example, overlooks a node and encodes two consecutive segments as one line the program is not able to set up the network topology correctly. The unexpected high frequency of that error is an indicator that the object (the polygon) and not the data structure (the segments and their relationships) is visually recognized. The proper solution is to digitize the polygon, and leave the construction of the data structure (segment generation and polygon definition) to the program. The same principles of operator guidance, plausibiltity checks and use of the acoustic channel are applied.

A short description of the digitizing process may illustrate the approach. After the opening dialogue the operator is requested to digitize a specific polygon. He traces the boundary line completely, without respect to any nodes or segments. For the neighbouring unit only those parts of the boundary are digitized which are not part of a previously digitized polygon. The program checks for line intersections, calculates intersection coordinates, if any, divides existing lines, corrects the line ends, and updates the pointers from lines to polygons. Warnings are issued in case of suspected errors or ambiguities. Points, lines and polygons can be erased or changed, either the last one, or the one under the crosshair. A check is done for double lines;

in the present version of the program, however, it is rather crude, and is not able to catch the more subtle problems.

5. Software tools

The digitizing system must be able to do the first data capture as well as the update and maintenance of the geometric data base. A number of functions (not necessarily included in the program for digitzing and editing) supports the user in administrating the data base. The packaged systems from the shelf on which many GIS are based include some of the functions mentioned here, in more or less comfortable implementations. Of course the list is far from complete, because certain applications require tools too special for every-day use in a GIS. A clear-cut separation between data acquisition and processing is not possible, thus some functions can also be counted as processing functions.

5.1 Data management functions

The basic data base functions - storage and retrieval of the geometric objects - are done by either a general data base management system, or special data base software. The relational model and its implementations are not considered to be the most efficient way to manage geometric objects. It is hoped that the next generation of data base management systems using an object-oriented model will have better performance characteristics (Schek 1986), but it will take a few years until such systems will appear on the market.

As we have seen, there are several technical alternatives to encode geometric objects. The GIS must not only provide the necessary tools for the method chosen but also must be ready to accept geometric data from other sources. Large data bases are available from various authorities, for example the World Data Bank files for international boundaries and other features, the TIGER system of the US Bureau of the Census, and the DLG-E files of the US Geological Survey; in Germany the ALK data base, and the ongoing ATKIS developments (Barwinski and Brüggemann 1986). The GIS must provide the appropriate entries to accept the data from outside sources. Format conversion is the lowest level of functionality. In some cases a conversion of data structures is required as well.

Other functions facilitating the management of geometric data can only be listed here. Some of the functions are part of the programs for geometric data acquisition, generalization and intersection, but should be also available as a separate utility to increase the flexibility of the GIS.

- Syntax checker: a program checks the geometric data for violations of the syntactical rules set up for the model or data structure, maybe combined with a repair function.

- Mesh generator: the program constructs the topology of a network from line data, for example originating from raster-to-vector conversion, or from an outside source.

- Line intersection: the function is needed for the "spaghetti" approach, but also for error checking, plausibility checks, and other functions.

- Similarity and containment tests: line matching, point-in-polygon, line-in-polygon, polygon-in-polygon tests.

- Transformation and abutment: the coordinates are transformed in a different axis system, in a different projection; map sheets have to be sewed together, with exact matching of the lines at the seams.

The boundary between management and processing of geometric data is rather unsharp. The following functions fit into both categories:

- Overlay and intersection of networks,

- Set operations on geometric data,

- Creating areas of influence (buffering).

The list is far from complete, especially for digital terrain data where additional functions should be available in the GIS.

5.2 Data reduction and generalization

A scaleless, seamless data base is the dream of the implementers of GIS. The geometric data is digitized and stored in the highest accuracy possible. From the data base all maps at smaller scales, or geometric data set with lower accuracy requirements can be derived. The main obstacles are the immense investments for the data capture, and the problems involved for generalization. A change of scale is not a simple division of the coordinate values by a scale factor. A reduction in the amount of data is required, for graphic display also different rules for visual transcription. Generalization implies, in addition to data reduction, the preservation of characteristic features.

Data reduction. The amount of data is reduced to save storage space and processing time. Data reduction is applied on different properties of the object. The number of points in a line, for example, can be reduced by eliminating all points inside a virtual tube along the line. Several algorithms have been published for data reduction in lines (Douglas and Peucker 1973, Robergé 1985, Thapa 1989). Data reduction on the topology is performed in a boundary network by assembling the base polygons to a higher level of hierarchy, for example from communities to counties. The now unused lines for the municipality boundaries within a county can be removed from the file.

Generalization. The preservation of characteristic features or important information is the main issue in generalization. For a coast line, for example, data reduction is too simple, because characteristics important for the user of the map can be lost. But what is a characteristic feature? How can it be determined from the object defintions in a data base? These two questions summarize the problems which must be solved first. Generalization is a modelling operation which requires more information than the geometric and topological properties of the object. Generalization also requires process knowledge, and is strongly dependent on experience, and also on the cultural environment. To place a cartographer, geographer or geologist in front of a graphic display and let him select the characteristic features is no solution, for two reasons: it takes too long, and the results will be different from person to person, from discipline to discipline. A solid theoretical framework for generalization must be developed before any programming should be done (Müller 1989, Shea and MacMaster 1989).

Wether rule- and knowledge-based systems, such as the approach of Nickerson and Freeman (1986), can be applied sucessfully for the development of the theory is an unanswered question at the moment. The concept of neural networks might a potential tool for the implementation of generalization procedures (Obermeier and Barron 1989).

6. Conclusion

A lot more could be said about geometric data acquisition. I hope it became clear that operator-guided digitizing still has its merits in comparison with raster-to-vector conversion, mostly for economical reasons. There is enough room left for improvements in technical and procedural respects. The basic functions for storage and retrieval of geometric objects must be enhanced by a selection of tools to facilitate data management and exchange with outside sources. Commercially available GIS packages provide most functions mentioned. In some GIS they should be implemented to make the product more useful.

References

ARC News (1989), environmental Research Institute, Vol. 11, No. 1, Winter 1989

Barwinski, K., Brüggemann, H. (1986), Development of digital cadastral and topgraphic maps - requirements, goals and basic concept. Proceedings of Autocarto London, Vol. 2, 1986

Becker, B., Ottmann, Th. (1988), Identitätsprüfung von Polygonzügen aus kartographischen Datenbasen. Proceedings Austrographics 1988, Wien. Informatik-Fachberichte 183, Berlin

1988

Chrisman, N. R., (1986), Effective digitizing: Advances in software and hardware. Technical Papers, American Congress on Surveying and Mapping, 46th Annual Meeting, Washington 1986

Ciarcia, S. (1982), Build the Microvox text-to-speech synthesizer. BYTE, September/October 1982

Douglas, D. H. (1989), New solutions to the boundary network problem. Paper presented at Auto-Carto 9, April 2-7, 1989, Baltimore, MD

Douglas, D. H., Peucker, T. K. (1973), Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. The Canadian Cartographer, Vol. 10, No. 2, 1973

Franklin, Wm. R. (1984), Cartographic errors symptomatic of underlying algebra problems. Proceedings of the International Symposium on Spatial Data Handling, Zürich 198

Goodchild, M. (1982), Accuracy and spatial resolution: critical dimensions for geoprocessing. In: Douglas and Boyle (ed.), Computer Aided Cartography and Geographic Information Processing: Hope and Realism, Univ. of Ottawa, 1982

Guptill, S. C. (1989), Speculations on seamless, scaleless cartographic data bases. Proceedings of Auto-Carto 9, April 2-7, 1989, Baltimore, MD

Hansen, H. (1988), personal communication

Illert, Andreas (1988), unpublished paper given at the University of Bonn, November 1988

Müller, J.-C. (1989), Theoretical considerations for automated map generalization. Unpublished discussion paper, ITC, Enschede

Nickerson, B. G., Freeman, H. (1986), Development of a rule-based system for automatic map generalization. Proceedings Second Symposium on Spatial Data Handling, Seattle 1986

Obermeier, K. K., Barron, J. J. (1989), Time to get fired up. BYTE August 1989 (other articles on neural networks in the same issue)

Pernot, M. (1986), Correction automatique de la géometrie après scannage. International Yearbook of Cartography XXVI, Bonn, 1986

Peuquet, D. J. (1984), A conceptual framework and comparison of spatial data models. Cartographica, Vol. 21, No. 4, 1984

Peuquet, D. J., Boyle A. R. (1984), Raster scanning, processing and plotting of cartographic documents. SPAD Systems, Williamsville 1984

Poiker, T. K. (1983), An intelligent cursor. Computer Graphics World, 10/83

Rase, W.-D. (1984), Computer-controlled input and output of speech: utilizing the acoustic channel for geometric data acquisition. Proceedings of the International Symposium on Spatial Data Handling, Zürich 1984

Robergé, J. (1985), A data reduction algorithm for planar curves. Computer Vision, Graphics, and Image Processing, Vol. 29, 1985

Schek, H.-J. (1986), Datenbanksysteme für die Verwaltung geometrischer Objekte. Proceedings GI-Jahrestagung 1986. Informatik-Fachberichte 126, Berlin 1986

Shea, K. S., McMaster, R. B. (1989), Cartographical generalization in a digital environment: when and how to generalize. Proceedings of Auto-Carto 9, April 2-7, 1989, Baltimore, MD

Tappert, (1988), personal communication

Thapa, K. (1989), Data compression and critical points detection using normalized symmetric scattered matrix. Proceedings of Auto-Carto 9, April 2-7, 1989, Baltimore, MD

Türke, K. (1976), A system for interactive acquisition and administration of geometric data for thematic map production. Computer Graphics, Vol. 10, No.2, 1976

Weber, W. (1987), Massendigitalisierung. International Yearbook of Cartography XXVII, Bonn 1987

Table 1

Comparison of digitizing techniques

	Operator-guided digitizing	Raster-to-vector conversion
Geometric accuracy	low 0.2 - 0.3 mm	high < 0.1 mm
Encoding of		
- Geometry	medium	good
- Topology	good	restricted
- Identification	good	bad
Device cost	low $ 5,000 - 10,000	high $ 300,000 - 1,000,000

Labour cost - Cost factors	high geometry acquisition	high preparation of input, editing, tagging
Main Application areas	editing tagging and encoding of topology	acquisition of geometry
	small tasks	high-volume production
	simple lines	complex linework
	thematic maps	topographic, cadastral maps