DIGITAL MAP GENERALIZATION AND PRODUCTION TECHNIQUES

David D. Selden and Michael A. Domaratz
U.S. Geological Survey
562 National Center
Reston, Virginia 22092

BIOGRAPHICAL SKETCH

David D. Selden received his M.S. in cartography from the University of Wisconsin-Madison in 1979. Since graduating, Mr. Selden has operated a freelance mapping business in Albany, N.Y., worked on an interactive graphics highway safety project for the State of New York, and has done cartographic consulting in the Washington, D.C., area. Currently, Mr. Selden is employed as a cartographer for the U.S. Geological Survey in Reston, Virginia, developing techniques for computer-assisted cartography.

Michael A. Domaratz is a cartographer at the U.S. Geological Survey's Eastern Mapping Center in Reston, Va. He is the team leader for the USGS 1:2,000,000-scale digital cartographic collection project and is involved in various digital cartographic data collection, processing, and display activities. Mr. Domaratz received a B.A. degree in geography from the State University of New York at Buffalo.

ABSTRACT

Two evolving aspects of the cartographic process, digital map generalization and procedures for cartographic manipulation of digital data, will be discussed in this paper. The application of digital techniques to cartographic data has greatly facilitated the cartographer's capability to quickly produce maps of various scales and projections. The hierarchical data structure and unique data content of the U.S. Geological Survey's 1:2,000,000-scale data base offer increased flexibility and enhance these cartographic production capabilities. The explicit encoding of major cartographic categories as well as internal category differentiation (for example, rivers, perennial vs. intermittent; federal lands, national parks vs. national forests) permit cartographic display appropriate for particular scale and theme criteria. The actual cartographic portrayal of these digital data, while based on the coding structure, also requires an efficient interfacing of human/software/hardware components. Finally, compliance with high cartographic standards and exploitation of the data base's potential are dependent on the development of effective digital cartographic production methods.

INTRODUCTION

The availability of digital cartographic data bases is creating a demand for accelerated development of automated map production systems. The content, structure, and unique qualities of these data bases are providing new opportunities for efficient and dynamic cartographic applications. Whether or not these opportunities will be fully exploited depends, in great measure, on the development of an appreciation for the intrinsic value of digital cartographic data designed for mapping purposes.

This paper addresses current work in the area of automated production cartography by focusing on the development of a 1:2,000,000-scale digital cartographic data set at the U.S. Geological Survey. An explanation of the rationale behind its unique data organization and a brief description of its elemental features will be presented. An example of how this data has been used to produce a map product will also be delineated.

DIGITAL CARTOGRAPHIC DATA

Theoretical Considerations

One of the major activities performed during the process of map generalization is data selection. In this activity, the cartographer selects the data to be portrayed on the map. The selection process is governed by two factors: the theme of the map and the scale of the map.

The map theme is the primary idea the cartographer is attempting to communicate on the map. This consideration requires the cartographer to select and portray map data that the user will need to comprehend the theme of the map. For example, a map showing the political subdivisions of an area would probably contain a relatively large amount of data representing the boundaries of the subdivisions and a lesser amount of supporting data (roads, streams, etc.) to help orient the map user. On the other hand, a map emphasizing the hydrography of the same area would have a greater amount of hydrographic data and a lesser amount of supporting transportation and boundary data.

The map scale determines the amount of space available for presenting the data. This factor constrains the space in which data may be legibly portrayed and must be considered by the cartographer in the data selection process. For example, a map emphasizing the transportation network at 1:2,000,000 scale can have more transportation data legibly portrayed than a map of the same area at 1:10,000,000 scale.

Implementation

One major goal of the U.S. Geological Survey's 1:2,000,000-scale digital data set is to provide the cartographer the ability to automatically select the proper combination of data for a map after considering the theme and scale of that map. To achieve this goal, the digital cartographic data are encoded with attribute codes that allow the cartographer to select the data to be portrayed on the map. The cartographer enters these selection criteria decisions into a computer, which processes the data set and extracts those features to be portrayed. By encoding the data with the attribute codes, the data selection process is not dependent on the need for special-purpose map generalization software.

Although the data set contains many different overlays of data (political boundaries, boundaries of federally administered lands, roads and trails, railroads, streams, and bodies of water), a common selection scheme is applied to all of them. The basic scheme consists of organizing the attribute codes hierarchically from the "most significant" feature in the overlay to the "least significant" feature in the overlay. Although the specific criteria for ordering the attribute codes vary with each overlay, the basic scheme allows the cartographer to start with a minimum amount of data for an overlay and to gradually increase map content (within the constraints of the coding scheme) until the cartographer has the data needed to support the theme and scale of a particular map.

The attribute codes for the political boundary overlay are organized in the following hierarchy: international boundaries, state boundaries, and city and (or) county boundaries. The scheme starts by considering the United States as an area bounded by international boundaries. This area can be decomposed into 50 constituent state polygons bounded by adjacent state boundaries; these state polygons can be decomposed into their constituent city and (or) county polygons bounded by adjacent city and (or) county boundaries.

The coding scheme for the federally administered lands overlay is organized differently. The United States is divided into generic polygons which represent national forests, national parks, national wildlife refuges, Indian reservations, etc. Each polygon is further classified by a size criteria determined by the longest dimension of the polygon. This scheme allows the cartographer to select the various types of federally administered lands that are to be displayed, and to control the amount of detail for the types of lands selected.

The road and trail coding hierarchy is organized to display different densities of connected networks of highways for the United States. The hierarchy of attribute codes for this overlay starts with major limited-access highways and their connectors (nonlimited access links where gaps appear in the limited-access network) and continues through major U.S. routes, minor U.S. routes, major state routes, paralleling U.S. and State routes, and other roads and trails.

The theory behind the attribute codes for the railroad overlay is similar to the theory for the road and trail overlay. The scheme, which is based on annual tonnage, starts with a connected network of the primary rail routes and their connectors, and gradually increases the density of the network.

The stream attribute code hierarchy is organized by length of drain. The scheme begins with the longest drain for an area, and the density is gradually increased by adding drains of shorter length, while retaining a connected network. The stream segments are also identified as being perennial, intermittent, or canal.

The theory supporting the water body attribute codes is similar to the theory of the federally administered lands overlay. The categories of perennial and intermittent lakes, marsh/swamps, dry lakes, and glaciers are subdivided by length-at-longest-dimension criteria. The cartographer may select among the different types of water bodies to be displayed, as well as control the density of water bodies chosen.

The limitations of map theme and map scale are important considerations in the data-selection process of map generalization. Using the criteria encoded in the 1:2,000,000-scale digital cartographic data, the cartographer has the much-needed versatility to generate maps by digital techniques.

CARTOGRAPHIC PRODUCTION

During the past year the U.S. Geological Survey has been experimenting with various procedures for applying the 1:2,000,000-scale digital cartographic data set. One such effort has concentrated on using the data with computer-assisted techniques to produce a facsimile of a general reference sheet from The National Atlas of the United States of America. The techniques utilized in this effort and some of their strengths and weaknesses are described below.

Planning

Planning for the (re)production of the general reference sheet required compliance with communication objectives established prior to its original construction in 1966. These objectives included the portrayal of appropriate geographic/cultural features for a region of the United States at a scale of 1:2,000,000. Unique to this planning process were concerns about the adaptability of the digital data to computer techniques. Such concern was justifiable, based on the knowledge that no known cartographic production system had been developed which fully automated the making of a map.

Design

Cartographic design is the process of graphically transforming data and ideas into map symbols and arranging them effectively to promote successful communication. Traditionally, this activity required imagination, a good understanding of accepted symbology, and a trust in the manual skills of the mapmaker. Even though design specifications for the National Atlas reference sheet already existed, the limitations of automated cartographic techniques required new design assessments. The ability of computer-assisted production methods to implement the design specifications had to be determined. Specifically, data symbolization, the placement and availability of high-quality type, and the automatic creation of open-window negatives were only some of the major areas of concern.

An appraisal of automated capabilities resulted in the revamping of certain previous design assumptions in order to maximize the utilization of computer techniques. A decision not to display type was made near the outset. Although the placement of type could have been accomplished using interactive graphic techniques, this was still considered to be essentially a manual procedure. Line and area symbolization was possible by automated means and standards for their specifications were considered of primary concern. In order to produce a final color proof, which necessitated the creation of open-window negatives, a combination of automated and manual methods was utilized.

The design activity also included the identification of appropriate tools (hardware), the selection of techniques (software and human interactive graphics), and the efficient interfacing of these hardware-software components. Three main pieces of hardware were chosen: a medium-sized mainframe computer which utilized an automated mapping software package to transform and symbolize data; a minicomputer with interactive graphics software for layout, editing, and some minor textual information placement; and a flatbed photoplotter for the plotting of color separates. In addition, appropriate software was selected to link the hardware devices.

Compilation

The recent availability of a cartographic data source in digital form changed the emphasis of the compilation activity. Unlike traditional map compilation where source materials are gathered, scaled, edited, and coded for intended graphic manipulation, the results of most of this activity were already provided. In one sense the 1:2,000,000-scale digital cartographic data base was a compiled source "manuscript."

Compilation thus took on a new form. Instead of collecting source information, a great proportion of the activity focused on the study of data format, structure, and basic definitions. Symbolization and color separation schemes were then devised by grouping the appropriate ranges of attribute codes together. The final step in this phase was the conversion of the digital data to latitude/longitude equivalents compatible with the mainframe symbolization software because the original digitized cartographic data resided in on-line disk storage on a mini-computer in X, Y format.

Finishing

Traditional map finishing concerns the creation of final map artwork. Drafting (or scribing), type and symbol stickup, and final layout of the map components represent the major tasks involved in map finishing. Computer-assisted map finishing utilizing the 1:2,000,000-scale digital cartographic data to re-create a National Atlas reference sheet consisted of two major steps: digital cartographic data symbolization-processing and automatic plotting.

Symbolization-Processing. In this particular case the data was symbolized on a medium-sized mainframe computer using an automated mapping package. A job stream was written which defined the desired symbolization parameters and the color separation scheme. In batch mode, this software extracted the desired level of data base information, symbolized the data, and wrote output commands on magnetic tape, which was used to drive a flat-bed photoplotter. Line-weight selection and various symbolizations (for example, railroad, canal, reefs, etc.) were supported. In addition, commands for the plotting of special symbols were available.

An alternative symbolization method which is currently under development utilizes the graphics capability of an interactive graphics system. Once these capabilities are upgraded to match or supersede those described above, a shift in the production flow would be possible. This shift would be advantageous in terms of bypassing the need to convert the original data source to geographic equivalents for compatibility with batch-symbolization processing. A considerable amount of time and effort would be saved. The only advantage remaining to the batch software would be the ability to transform the data into any one of 18 different map projections. This capability is not immediately planned for the interactive graphics facility.

The quality resulting from either symbolization technique, in terms of consistency and maintenance of USGS specifications, is lacking. Judged on merit alone, current symbolization quality is marginally acceptable. Different types of symbolization are distinguishable, although limited and unsophisticated. For the present, development efforts will be directed at expanding symbolization vocabulary as well as streamlining the procedures for their production. However, the amount of software development required to reach these goals of higher standards and expanded capabilities must not be underestimated.

Automatic Plotting

The automatic plotting of symbolized data was accomplished on a flatbed photoplotter. Both positive and negative film separates can be produced as well as scribecoats and pen drawings. Once a secure system of pin registration was established, a decision was made to plot film negatives. Film was punched prior to plotting, and pins were firmly attached to the plotting table. Precautions taken during plotting maintained registration within the approximate 0.003" accuracy specifications of the plotter. Given this reliable registration system, separation plotting in negative increased production efficiency by not requiring the making of contact negatives from plotted positives (much like manual scribing affected drafting).

Reproduction

Once all negatives were processed, they were submitted to a photographic laboratory for the production of peelcoats in the traditional manner. Peeling was done manually to create open-window negatives. The line negatives and open-window peelcoats were then returned to the laboratory for compositing in preparation for proofing and eventual printing. Essentially, the reproduction step remained a purely manual operation. However, some interesting developments are currently underway.

The implementation of raster processing near the end of the production flow could dramatically increase production efficiency. The potential for the elimination of traditional photomechanical reproduction work appears to be attainable, given certain software interface developments. Instead of photoplotting symbolized digital cartographic data, the data would be read directly into an interactive raster color-editing system. Interactive manipulation of the raster files would create the color specifications. Screened and composited negative separates would then be plotted on a laser plotter. These plots would be submitted to a photographic laboratory for proofing and (or) the generation of printing plates.

A Final Product

Certain impressions come immediately to mind when viewing a color proof of the National Atlas reference sheet produced by the methods described. The lack of type, route shields, point symbols, and any geographic/cultural information surrounding the mapped area stands out. The absence of map type has already been addressed in the design section. At this point it is sufficient to say that name/symbol placement is the most complex of cartographic activities, particularly in terms of automation. The lack of map information surrounding the States of Arizona and New Mexico is attributable to the current format of the 1:2,000,000-scale digital cartographic files. They were originally digitized in regions that were usually clipped along state political boundaries.

Even though it is certainly possible to present the data according to the established National Atlas reference sheet format, this does require a considerable amount of additional work at the present time.

On the positive side, some achievements are worthy of note. A symbolized (line weight and pattern) hierarchy of transportation, hydrography, and political boundaries is clearly portrayed. The line negatives which produced the open-window negatives were plotted automatically. The time required to produce the color separations for this map was much less than would normally have been required to produce the same product using traditional manual techniques. In general, a feel for the geographic/cultural features which make up this region can be quickly derived from the map in its present form, thus fulfilling its basic communication objective.

CONCLUSIONS

The availability of a digital cartographic data base, designed with built-in graphical flexibility, and the development of computer-assisted techniques have greatly influenced the character of the map production flow. Such production, as it is presently configured, results from a fusion of manual, photomechanical, interactive graphics, and digital techniques. It is hoped that as a truly automated map production system emerges, new efficiencies unattainable by traditional cartographic means will be realized, and the ultimate goals of efficient production and effective map communication will be better served.