GENERATING INFORMATION FROM SCANNING OHIO MAPS (GISOM): THE CONVERSION MODEL FOR GENERATION OF DLG-3 FILES FROM 7.5' QUADRANGLES FOR THE STATE OF OHIO

J. Raul Ramirez, Ph.D.
GISOM Project Manager
The Ohio State University Center for Mapping
1216 Kinnear Road
Columbus, Ohio 43212

Abstract

The Ohio State University, several Ohio state agencies and the U.S. Geological Survey (USGS) have a 4-year cooperative agreement to convert more than 700 7.5' quadrangles covering the State of Ohio to DLG-3 Standard form. This paper describes the conversion model used in this project.

1. Introduction

On October 1, 1993, The Ohio State University, four Ohio state agencies under the umbrella of the Ohio Geographically Referenced Information Program (OGRIP), and the U.S. Geological Survey (USGS) started a 4-year cooperative agreement to convert 700 7.5' (1:24,000) quadrangles partially covering the State of Ohio to DLG-3 Standard form (the 1:24,000 series is the largest scale national topographic series of the United States; DLG-3 is the digital format developed and used by the USGS). Conversion includes boundaries, hydrography, hypsography, public land survey system and transportation (roads, railroads and miscellaneous transportation). At the end of the agreement, Ohio will be the first state in the nation with such coverage. This project is known as the Generating Information from Scanning Ohio Maps (GISOM) project.

This paper describes a cost efficient conversion model (PC-based) developed for GISOM. Topics discussed include: scanning, warping, plotting, vector data generation and QC/QA. Experiences gained by converting the first 192 quadrangles and the resultant improvements in the production model are discussed.

2. Background of the GISOM Project

The Ohio State University Center for Mapping started a pilot project in 1990 to become familiar with the conversion of 7.5' quadrangles into DLG-3 Standard format files. This pilot project was sponsored by The Ohio State University, four state agencies and the USGS. In 1991, a high resolution scanner (12.5 microns), two workstations and several computer programs were acquired to support this project. A full-time researcher and two graduate students spent the next two years learning about DLG-3 conversion and its improvement. During the second semester of 1992, these personnel plus two volunteers produced the first DLG-3 files. By August 1993, a total of 16 complete quadrangles (117 DLG-3 files) were completed and delivered to the USGS.

During this three-year effort, a great deal was learned about the process of map conversion, and the need for an improved solution was apparent. As a result, a plan was prepared to improve existent conversion methodology. Two types of tasks were planned: short-term tasks (those that could be accomplished in less than six months) and long-term tasks. High pay-off tasks were given higher priorities. During the second semester of 1992 and the first semester of 1993, most of the short term tasks were completed. Long term tasks have been in development since January 1992, and several have been completed. In February 1993, The Ohio State University Center for Mapping presented a proposal to the USGS for the conversion to DLG-3 format of 700 7.5' quadrangles. In September 1993, The Ohio State University Center for Mapping was informed that the USGS had accepted the proposal, and the cooperative agreement started on October 1, 1993.

3. Project Organization

The Ohio State University Center for Mapping, Ohio Department of Development (ODOD), Ohio Department of Natural Resources (ODNR), Ohio Department of Transportation (ODOT), Ohio Environmental Protection Agency (OEPA) and the USGS are participating in the GISOM project. The 700 7.5' quadrangles to be converted to DLG-3 format are part of 793 quadrangles covering the state of Ohio. The 93 quadrangles not included in the agreement already have partial or complete DLG-3 coverage. The conversion will be completed in four years with a cost of $4,900,000. Half of the money will be provided by the state agencies and the rest by the USGS. Another state agency, the Ohio Department of Administrative Services (ODAS), is storing the DLG-3 files once they are accepted by the USGS. ODAS is also the depository and distribution center of these files for the state of Ohio.

At The ohio State University Center for Mapping, the project is organized as follows: there is a project manager responsible for the research and development and the coordination of the conversion effort. Research and development is carried out by the project manager and a group of graduate students. A total of five graduate students are currently working as graduate research associates (GRA) for the GISOM project. The conversion effort is supervised by four full-time senior project leaders under the direction of the project manager. A junior project leader and two student-employees assist the senior staffers. Map conversions are done by 18 independent contractors working from their homes. Their work is reviewed and approved by The Ohio State University Center for Mapping project leaders before final submission to the USGS.

4. Conversion Methodology

Each year of the project, the state agencies submit a prioritized list of quadrangles for DLG-3 conversion. The Ohio State University Center for Mapping then coordinates the selection of quadrangles for a given year. This list is submitted to the U.S. Geological Survey for their revision (against other on-going conversion projects and 7.5' quadrangles updating), and they proceed to approve it or recommend changes. Once the list is approved, The Ohio State Univeristy Center for Mapping orders the corresponding source material (up to 5 chronopaque separates: hydrography, hypsography, cultural features, photo-revisions and fence roads). It can take up to six weeks for The Ohio State Univeristy Center for Mapping to receive the material ordered from the USGS. This material is examined at The Ohio State Univeristy Center for Mapping, and when obvious problems are evident (such as too many broken lines or missing tic marks), new material is requested.

After the source material is accepted, each source document is scanned at 25 microns resolution to generate binary images. A set of pre-defined scanning parameters is used and it takes about 20 minutes to scan each source material (18' x 24'). Raster files are plotted and examined to evaluate the quality of scanning. Three major factors are examined: resolution of the corner tic marks, location of the quadrilateral defined by the tic marks with respect to the map features and quality of the line work. Tic marks must be about 7 pixels wide, the map features must be completely inside the quadrilateral defined by the tic marks and lines must be continuous. In the case of contours, lines must not touch each other. If problems are detected, the source material is scanned again after adjusting the parameters to try to overcome the problems. If better results cannot be obtained, new source material is requested from the USGS.

Once raster files are accepted, each file is warped to the coordinate space of the corresponding quadrangle. Warping is done by a program developed at The Ohio State Univeristy Center for Mapping. The whole warping process is done automatically. The set of files to be warped are placed in a given directory in the main computer unit at The Ohio State Univeristy Center for Mapping, and the program automatically finds the tic marks, computes and performs a coordinate transformation and the pixel resampling.

After files are warped, the hypsographic file is used as input to generate vector lines from a raster-to-vector program. A pre-defined set of rules for data classification, filter values to generate center lines and tolerance values to close gaps and delete line spurs are used. The output of this process is a vector file representing the contours generated from the raster image in the coordinate space of the corresponding 7.5' quadrangle.

Once vectorization is completed, the raster and vector files are copied to The Ohio State Univeristy Center for Mapping main processing unit under the directory of the corresponding independent contractor. An E-mail message is sent to him/her indicating that files are ready for data collection, attributing and editing, and that a copy of the corresponding 7.5' quadrangle plus plots of the raster files are available.

Independent contractors working at their homes download the files into their PC's using the File Transfer Protocol (FTP) software. A telephone modem of 14,400 bps and compress programs make this transfer operation efficient. Independent contractors' work can be divided into three major tasks: (1) Non-hypsography heads-up digitizing and attribute tagging, (2) Contour correction and elevation and attribute tagging, and (3) Error detection and correction. Error detection is accomplished by a program called PROSYS developed by the USGS and other programs developed at The Ohio State Univeristy Center for Mapping. PROSYS resides in the main processing unit at The Ohio State University Center for Mapping.

After vector files are completed (no errors detected), independent contractors send an E-mail message to the senior project leader coordinating their work. The senior project leader finds the latest version of these files at our main processing unit and proceeds to catalog them. Then, the files are plotted and the plot material is compared with the source material. Then, project leaders examine the files (line work, topology and attributes) in an interactive fashion. A sampling approach is used to check these files. The goal is to check at least twenty percent of the features in each file. If errors are detected in the visual quality control or in the interactive evaluation of a file, a more extended evaluation is performed. Depending upon the number of errors detected, files are either given back to independent contractors for correction (in the case of many errors) or are corrected at The Ohio State Univeristy Center for Mapping (when there are few errors). The files are checked again (PROSYS, visual QC, inter-active sampling). Once no errors are found, the files are sent (via FTP) to the USGS for final evaluation and approval.

5. The Conversion Model

One major goal of the GISOM project is the design of a cost-efficient conversion model for 7.5' quadrangles into DLG-3 files. Such a model is needed because there are more than 57,000 7.5' quadrangles covering the United States (excluding Alaska), and in 1992, only 2,850 quadrangles (less than 5%) had DLG-3 files for some or all the layers [1]. For hypsography, specifically, only 1,710 quadrangles (less than 3%) had digital representation. The USGS started the conversion effort in 1978. At the rate of conversion achieved over the last 15 years, more than 300 years would pass before the whole country was finally converted. In order to convert all the 7.5' quadrangles in a reasonable number of years, a massive effort must be undertaken. Of course, conversion cost is a very important issue: the costlier the conversion process, the more difficult it is to undertake a massive effort. Therefore, cost-efficiency is a major consideration in the design of the conversion model.

Great improvements in PC processing power and storage, together with the high degree of sophistication of PC-based CAD software, were very important factors in the development of The Ohio State Univeristy Center for Mapping conversion model. The current model uses an assortment of hardware and software with different degrees of sophistication. Today, this model can be replicated using only PC's. The major components of the conversion model are described below.

The heart of the whole model is the scanning component. The scanning component generates digital raster files from the hard-copy source documents. The high-resolution scanner and corresponding computer programs generate a digital raster file from a hard-copy document. Any scanner of a resolution of at least 32 microns, which is the minimum resolution allowed by the USGS, can be used for the conversion of 7.5' quadrangles. The scanner is controlled by a computer and the corresponding software. High-resolution scanners are driven today by PC's or workstations and corresponding scanning software is available. The format of the output files depends on the scanner used. Today, file format is not a major problem because there is enough information about most of these formats, and therefore, it is relative simple to convert files to the most appropriate format. In the case of The Ohio State Univeristy Center for Mapping, the scanner used is the Optronics 5040, the computer is an Intergraph workstation, the software is I/SRIF and the raster format is RLE (run length encoding).

The scanning component is the most expensive part of the conversion model. Generally, a single high-resolution scanner can be used as part of several conversion efforts. If a scanner is used in two shifts (16 hours/day), you maybe able to scan 8 quadrangles per day or 2080 quadrangles per year at 1024 dpi resolution (the resolution used in GISOM). Therefore, a single scanner driven by a PC could be used to support twelve conversion projects similar to GISOM.

The other components of the conversion model are: warping, plotting, vectorizing and attributing, and QC/QA. Warping is the component that transforms the raster files from the scanning to the map space. In general, scanning is performed in a local coordinate system similar to the local coordinate system of the hard-copy source document (in the GISOM project, the upper left corner is the local origin and the units are inches). Map features usually are given with respect to an absolute coordinate system (in the case of the 7.5' quadrangles, the Universal Transverse Mercator map projection is used). In order to warp a raster image, two distinct operations need to be performed: a geometric transformation using common points and a resampling of the pixels to go from one space to the other. In the GISOM project, a program developed in-house is used. In this program, the affine transformation is used to evaluate the accuracy of the tic control points. If they are found satisfactory, then an eight parameter projective transformation is used in combination with a backward resampling.

Plotting is the component of the production model that generates hard-copies of raster files (after warping) and vector files received from independent contractors. There are two elements of plotting: a high-resolution plotter and the corresponding plotting software. In the GISOM project, a Versatec electrostatic plotter (400 dpi) is used. Two different plotting packages are used: I/PLOT from Intergraph and one developed at The Ohio State Univeristy Center for Mapping. The first requires a Intergraph workstation to run and the second runs from a PC. Today, an ink-jet plotter driven by a PC could perform as well as our plotting hardware at a fraction of our current cost.

Vectorizing and attributing are the components of the conversion model that generate graphic data in vector form with corresponding attributes. Three basic options for vectorization and attributing exist: heads-up digitizing, interactive line following and automated line following. At the GISOM project, only heads-up digitizing and automatic line following are operational at this point. Heads-up digitizing and attribute tagging involve manually selecting DLG-3 codes and tracing vector lines along a particular feature using the raster image as a backdrop. PC-based software developed at The Ohio State Univeristy Center for Mapping is used for heads-up digitizing. This software is based on USGS workstation software. Heads-up digitizing is used for all layers but hypsography. Automated line-following is the process of generating vector lines from the contour raster images without operator intervention. As indicated earlier, a set of rules and parameters and tolerance values are set before running the program. At the GISOM project, the Intergraph I/VEC program has been used for automatic line-following. The processes of contour correction and elevation and attribution tagging improves the vector lines generated by I/VEC and attach the corresponding DLG-3 codes (including the elevation) to each contour line. A tagger program developed at The Ohio State Univeristy Center for Mapping is used for this operation.

QC/QA is the component that controls the map conversion process to assure consistency and quality. Two major tasks of this component are production model evaluation and error detection. Production model evaluation studies the different components of the production model, their inter-relations and performance and makes decisions about their improvement. Error detection finds and fixes geometric, topologic and attribute problems in any layer. Most of these problems are detected by PROSYS and visual means, and they are corrected in an interactive fashion by the independent contractors.

6. Future Improvements to the Conversion Model

By January 1995, The Ohio State Univeristy Center for Mapping has participated in the conversion of 1344 DLG-3 files (192 quadrangles). Since June 1992, we have decreased production time from 400 hours/quadrangle to 200. The conversion of hypsography has dropped from 80% of the overall conversion time to 61%. The goal of the project is to decrease the average quadrangle conversion time to 65 hours and hyposgraphic conversion to 20% of that 65 hours. In order to achieve these goals, four research projects are in progress: (1) development of a highly automated contour conversion and tagging program, (2) replacement of line work visual quality control by computer generated quality control, (3) automation of the conversion of the hydrographic layer, and (4) automation of the conversion of the transportation layer. The first two research projects will be completed during 1995.

Reference

[1] USGS 1992 State Mapping Advisory Committee Workshop. The United States Geological Survey, Reston, 1992.