

# ANALYSIS AND DISPLAY OF DIGITAL ELEVATION MODELS WITHIN A QUADTREE-BASED GEOGRAPHIC INFORMATION SYSTEM

Juan A. Cebrian, James E. Mower, and David M. Mark  
State University of New York at Buffalo  
Department of Geography  
Buffalo, New York 14260

## ABSTRACT

Digital elevation models (DEM) are an essential component of many geographic information systems (GIS), especially in natural resources applications. Recently, a number of prototype geographic information systems based on quadtrees have been developed. Digital elevation data create special problems for quadtree systems, since neighboring grid cells seldom have identical values. In this paper, we present a strategy for integrating DEM data into a quadtree-based GIS. The Morton sequence, which is the basis of most linear quadtree systems, is used to provide cell addresses in the DEM file.

## INTRODUCTION

A Geographic Information System (GIS) may be defined as a computerized, spatially-referenced data base organized in such a way that spatial data input, analysis, and output may be accomplished. As is the case for any computer application, the issue of data structures is a critical one. Once a data structure for a GIS has been adopted, it is very difficult to change it; data structures are also a major factor in determining the efficiency with which queries can be answered within a GIS. Geographic information systems frequently involve natural resource data. In most such applications, elevation data (digital elevation models, or DEM) represent an important component, both directly and in the form of derived measures such as slope. Thus, any data structure adopted for a resources GIS must be able to efficiently integrate DEM data with other geographic data.

Recently, quadtrees have received considerable attention as a data structure for GIS applications (see below). Quadtrees appear to have many advantages for handling coherent ('blocky') spatial data, but are inefficient for continuous surfaces such as topography. However, if quadtrees are to be used in natural resources and related applications, it is essential to develop strategies for efficient integration of DEM data into the quadtree environment. This paper presents such a strategy.

### Quadtrees

The quadtree is a data structure which is based on a regular decomposition of an image into quadrants and subquadrants. The region quadtree is relatively new, having been first suggested in 1971 (Klinger, 1971), and

fully developed for the first time in 1976 (Klinger and Dyer, 1976). As an example, we will consider the construction of a quadtree of a soils map. We assume that the soils map exists as a grid cell representation, that is, we begin with an array of square cells, each of which has a soil type. The map is assumed to be enclosed by a square region of side length  $2^n$  cells. The cells are nodes of level 0, and the entire image is a node of level n (note that a node of level L has side length  $2^L$ , and contains  $4^L$  cells).

The 'top-down' construction of the quadtree begins by examining the entire image. If it contains only a single soil type, then the process stops, and the quadtree is a single node representing the whole image. If, however, the region contains more than one value (soil type), it is divided into four quadrants (level n-1 nodes), each of which is referred to as a child of the image square. Each of these quadrants is then examined for uniformity of value, and the process is applied recursively until all subquadrants are uniform. Note that since single cells are of uniform value and are valid subquadrants, the process must stop at the cells if not before. For blocky images, the quadtree typically will occupy much less space than the grid cell representation.

Linear Quadtrees for GIS. Early work on quadtrees used explicit pointers to represent the tree. Recently, however, alternative structures termed linear quadtrees have been adopted by many researchers (Gargantini, 1982; Lauzon, 1983; Abel, 1984; Mark and Lauzon, 1984; Samet and others, 1984; Lauzon and others, in press). In a linear quadtree, each leaf node is given a unique key number (or key and level combination), based on an ordered list of the node's ancestors. Then, the quadtree is simply a list of all the leaves, sorted by key. Neighbor-finding and other tree traversals are accomplished through the use of modular arithmetic or by bit addressing individual bits in the keys. Gray nodes can be omitted. Our research group has adopted the 'Morton matrix' sequence (Morton, 1966) as an appropriate key system for linear quadtrees. Briefly, the Morton number of a cell can be obtained from the row and column coordinates simply by interleaving the bits from the coordinates (coordinates and cell keys begin from 0).

Two-dimensional run-encoding (Lauzon, 1983; Lauzon and others, in press) is a technique for compacting a linear quadtree. Whenever the linear quadtree has consecutive leaves of the same color, only the key of the last leaf in such a 'run' is stored. Lauzon discovered that such runs could readily be decoded into maximal leaves; run-encoding usually reduces the number of records in a file by more than fifty percent (Lauzon, 1983; Lauzon and others, in press; Mark and Lauzon, in prep.). Furthermore, since any valid quadtree leaf of level m represents a sequence of  $4^m$  consecutive Morton numbers all of the same color, runs of leaves are simply runs of pixels. A 2DRE file could be constructed simply by visiting the cells of a digital image in Morton sequence, saving the Morton key and color of the last cell in any run.

### Digital Elevation Models

A digital elevation model (DEM) can be defined as any machine-readable representation of topographic elevation data. Digital elevation models are widely used in a variety of fields, including cartography, engineering, mining, land use planning, and defense. A major issue in DEM research is the selection of an appropriate data structure (Mark, 1979). The most common data structure for digital elevation models is a regular square grid; the geographic space is partitioned into square cells of constant size, and an elevation value is associated with each cell. One problem with regular grids is that they are often highly redundant. The grid size must be sufficiently small to sample the smallest feature of interest in the entire study area, and to define the boundaries of larger features to some required level of precision. Then, because a constant cell size is used throughout the area, cells in most of the region will be smaller than needed (i.e., there will be too many cells).

In an attempt to address this problem, alternative data structures for digital elevation models have been designed. The most widely used of these is the triangulated irregular network (TIN), which represents the terrain by a triangulation on a set of points chosen to represent the surface (see Peucker and others, 1978). While this approach appears to have many advantages, it has not made a major inroad into the use of regular grids for digital elevation models. In large part, this is because devices are now available to produce very dense regular grids directly from aerial photographs (Noma and Spencer, 1978; Swann and others, 1978; Allam, 1978).

The quadtree would seem to hold considerable potential here, since it is an adaptive structure with variable resolution. There is, however, a major problem with the strict application of quadtree concepts to topographic data: it is unusual to find sets of four mutually adjacent cells which are of identical height. Of course, this depends on the height resolution of the data, but a typical case is the DEM data distributed by the U.S. Geological Survey (USGS). These data sets report elevations of 30 by 30 meter cells as integers in meters (Elassal and Caruso, 1983). Ground slope has to be less than 3 percent before adjacent points will have the same value.

### Digital Elevation Models and Quadtrees

In existing quadtree-based geographic information systems at the University of Maryland (Samet and others, 1984) and at the University of California at Santa Barbara (Pequet, 1984; Chen, 1984), topographic data has been included by classifying the DEM data and producing quadtrees of the resulting coverage files. Both of the above systems have included in their published examples only maps of inter-contour areas, presented as quadtree files. The topographic component thus consists of as many binary quadtrees as there are inter-contour areas. We assume that slope data would be included in these systems in a similar way, as quadtrees of images of classified slope maps.

Inclusion of such secondary or derived data in a GIS is undesirable. In the slope example, it is unlikely that various users of a GIS will require maps at the same critical slope values. Once a set of critical slopes has been adopted and included in the data base, the flexibility of this aspect of the GIS has been greatly reduced. We propose that it is highly desirable to represent the topography as precisely as possible within the GIS; then, classified images can be derived to fit the user's exact requirements. In this case, there seems to be little alternative to the storing of all cells in the DEM grid.

We propose to do exactly that. However, in our system, the grid is not stored row by row, but rather, the grid cells are stored in Morton sequence. Lauzon (1983) and Lauzon and others (in press) have proposed that digital elevation models and other full matrix data be integrated into the linear quadtree GIS by using the Morton numbers as addresses for cell elevation data in a one-dimensional array or file. The use of the Morton key as a position within the DEM file allows for easy and efficient integration of DEM data into a GIS based on a linear quadtree structure such as 2DRE.

#### DEM DISPLAY WITHIN A LINEAR QUADTREE GIS

DEM display algorithms for grid data can be easily modified to process data in the Morton sequence. Quadtrees of two-dimensional (coverage) data, such as land-use or geology, can then be used to select portions of the DEM for analysis or display. In this section, we describe QDEM, a system based on this approach. Algorithms for DEM display using this strategy will be presented, and examples will be shown. The data display techniques discussed include: elevation band maps; slope steepness maps; slope aspect maps; and analytical hill shading.

#### Hardware and Operating System

QDEM is written in the C programming language, under Venix, a UNIX-like operating system, on a Terak microcomputer. The hardware configuration includes a Terak 8510/b Graphics Computer System and an 8600 Color System. The 8510/b uses the DEC LSI-11/23 16-bit microprocessor (and thus shares the architecture of the DEC PDP-11), and has 256k bytes of random access memory. The 8600 is based on the Intel 8086 16-bit microprocessor, and accesses up to 32k of display processor memory as well as 256k of frame buffer memory. The frame buffer is logically divided into four banks, each controlling 64k of RAM and three bit planes. As configured, the user can control two visually interlaced 640 by 240 pixel displays and 6 bit planes, providing a virtual resolution of 640 by 480 pixels and a palette of 64 colors. The 8600 provides a default palette, and also allows the user to define customized palettes. The red, green, and blue color guns each have 8 possible values (0-7), giving 512 possible colors.

Within a program, the user selects a color for drawing via a color index (0-63). Specifying a positive color causes

the new drawing index to alter all bit planes, setting them to '0' or '1' according to the binary representation of the color index. Specifying a negative color index causes a bitwise OR to be performed on the current color index (ignoring the sign) and the color index already present in the pixel. In effect, only the bit planes equivalent to the bits which are '1' in the current color index will be changed to '1's. Thus, users may overlay and manipulate up to six distinct binary images (or, as described below, two distinct 8-value images).

The Terak Color Graphics Software package provides both ACM-SIGGRAPH compatible routines and Terak extensions. Many of the Terak extensions are primitives which allow display addressing at the physical device coordinate level. Since quadtrees are based on pixels, we have found that the primitive physical device routines are much more useful than the SIGGRAPH ones, and have written our own quadtree-based higher level graphics procedures.

#### Data Preparation

The first procedure developed reads a grid DEM, such as one of the U.S. Geological Survey's 7-1/2 minute Quadrangle series of digital data sets (Elassal and Caruso, 1983), and creates two files. The first one, which is called the 'header' file, contains such information as the map title, location, maximum and minimum elevations, grid cell size, and the parameters necessary for the conversion of the data to screen coordinates. The second file (termed the 'Morton' file) contains only the altitudes of the grid cells (30 by 30 meter squares in the case of the USGS data), which will be represented by single pixels in the display device. As the heights in the original file are usually processed in row-column order, the pre-processing routine calls the primitive 'morton' to generate the Morton key of the cell, which can be easily matched with several 2DRE thematic files by the system. Morton is a procedure which converts x-y (column, row) coordinates to a Morton key, using an algorithm based on modular arithmetic (see Lauzon, 1983). Then, procedure 'put\_height' stores the height of the point in the appropriate location in the Morton file. Given the Morton number of a pixel, its height can then be directly accessed by key, using procedure 'get\_height'. The procedures get\_height and put\_height have been implemented in C, and use the procedure 'lseek' (Kernighan and Ritchie, 1978, p. 164), which goes to a position in a file which is a specified offset (number of bytes) from the beginning of the file. Each integer height occupies two bytes, and thus the offset is just twice the Morton key.

In a 2DRE file of some type of coverage data, two Morton numbers  $K_1$  and  $K_2$ , with  $K_1 < K_2$ , define a homogeneous region. As a consequence of the organization of the Morton file containing the heights, the current system can very quickly retrieve all the topographic information concerning such a uniform area. A program can retrieve the heights of all points which are on, say, a certain soil type simply by reading the soil type file and examining the soil value. Whenever this corresponds to the type being examined, the

program can read the heights of all points with keys from the previous (in the 2DRE file) record's key, plus one, to the current record's Morton key.

### Software Configuration

As noted above, QDEM is a package of procedures written in C. The program is interactive and menu driven, and manages a number of other routines which perform user-requested graphic and other analytical tasks. Dialogue between the system and users is provided by means of an alphanumeric keyboard and a graphic cursor. Within this dialogue, users select appropriate DEM and thematic files, and define analysis and/or display tasks and associated parameters. The data base is not yet accessible through queries based on coordinates.

The principal characteristics of each user query are resolved in a three stage dialogue. The first selects the type of analysis to be performed, the second determines whether or not thematic data are to be used to window the DEM, and the third defines the type of output (graphic or other).

Immediately after this stage, a second level of dialogue is established to determine the parameters which will drive the chosen process; this will depend on the type of analysis and/or display selected. There are two different sets of parameters: those which refer to the classification of the pixels (number of intervals, amplitude of the intervals, and so on) and those which will characterize the graphic output, if required. The first ones are introduced from the keyboard.

Users can select colors for the graphic representation from any of several palettes. With the aid of the graphic cursor, users can easily pick colors by moving over the palette menu. If none of the palettes available on the system are suitable, users can build their own palettes by calling the appropriate routines.

### Types of Analysis and Display Procedures

The basic DEM displays described by Kikuchi and others (1982) in their DEMGS system have been implemented here; we refer the reader to that paper for further details of the algorithms. From each of these types of analysis, users may wish to produce graphic output, to compute summary statistics, or to output a linear quadtree (2DRE) file. We have decided that at present only the first and last option will be available, and that statistical analyses will be performed on the 2DRE files rather than on the DEM directly.

Contour Band Map. This procedure simply classifies the pixels according to their heights. The parameters of this classification are requested from the user; all parameters do, however, have default values. The maximum and minimum elevations are already recorded in the corresponding header file. After users determine the number of class intervals, the uniform height range for all classes is then derived from the other parameters. Users can also define varying

height intervals. Elevation maps would seldom use varying intervals, but in GIS queries some users may be interested in a specific partition of the height dimension in order to match this result with some other spatial phenomenon.

Slope Steepness Map and Slope Aspect Map. These images are produced using simple linear algebra operations (see Kikuchi and others, 1982). Users can determine the number of slope classes, their limits, and the colors to be used on graphic output. The pixels are processed in Morton order. The elevations of two non-adjacent neighbors (in this case, arbitrarily the cells to the east and south) are used to establish a cross-product vector orthogonal to the surface. The length of the z-component of the unit vector orthogonal to the surface equals the cosine of the slope angle; the slope aspect can be obtained from the x- and y-components of the orthogonal vector (Kikuchi and others, 1982). Note that, for every cell, two neighbors must be examined. Thus, in the current implementation, each elevation is actually read three times, reducing the speed of the algorithm. We intend to remedy this by modifying the program to read contiguous blocks of data, 512 bytes at a time.

Analytical Hill Shading. This procedure is based on the same linear algebra operations as the slope maps. However, in this display, the classification is done automatically. The cosine of the angle between the surface normal vector (see above) and the light source is determined as the dot product between those two unit vectors. All the pixels are then classified into one of 8 colors: black for surfaces directed away from the light source, white for surfaces perfectly orthogonal to the light source, and 6 gray shades for the intermediate cases. The user can, however, select the location (azimuth and altitude) for the light source.

#### Relating Thematic Attributes to the DEM

The system has been designed in order to optimize the time response of special queries about the topographic characteristic of specific zones inside the area covered by the DEM. As the areas under consideration are described by linear quadtrees, and the DEM is sorted in Morton order, the windowing of the DEM is straightforward. Only the pixels in the area of interest are examined and classified. There is no need to deal first with the DEM within a rectangular area which encloses the search polygon and then later clip the rectangular area following the edges of such polygons.

Two different overlay procedures have been devised and implemented. The first one displays topographic properties (heights, slopes, slope aspects, etc.) within a definite zone (such as a certain land cover or soil type). It is also possible to prepare statistical reports about these properties. The search area is always described by a 2DRE file. Algorithms are available to convert vector descriptions of polygons into linear quadtree representations (see Mark and Abel, 1984).

The second overlay procedure produces graphic output. We

		HUE									
		Red	Blue	Magenta	Green	Yellow	Cyan	White			
Lightness	1 00	(7,7,7)	(7,0,0)	(0,0,7)	(7,0,7)	(0,7,0)	(7,7,0)	(0,7,7)	(7,7,7)	111	Lower Three Bit Planes
	0 85	(6,6,6)	(6,0,0)	(0,0,6)	(6,0,6)	(0,6,0)	(6,6,0)	(0,6,6)	(6,6,6)	110	
	0 71	(5,5,5)	(5,0,0)	(0,0,5)	(5,0,5)	(0,5,0)	(5,5,0)	(0,5,5)	(5,5,5)	101	
	0 57	(4,4,4)	(4,0,0)	(0,0,4)	(4,0,4)	(0,4,0)	(4,4,0)	(0,4,4)	(4,4,4)	100	
	0 42	(3,3,3)	(3,0,0)	(0,0,3)	(3,0,3)	(0,3,0)	(3,3,0)	(0,3,3)	(3,3,3)	011	
	0 28	(2,2,2)	(2,0,0)	(0,0,2)	(2,0,2)	(0,2,0)	(2,2,0)	(0,2,2)	(2,2,2)	010	
	0 14	(1,1,1)	(1,0,0)	(0,0,1)	(1,0,1)	(0,1,0)	(1,1,0)	(0,1,1)	(1,1,1)	001	
	0 00	(0,0,0)	(0,0,0)	(0,0,0)	(0,0,0)	(0,0,0)	(0,0,0)	(0,0,0)	(0,0,0)	000	
			000	001	010	011	100	101	110	111	
Upper Three Bit Planes											

Figure 1: Color gun values (in the order: red, green, blue) for the video look-up table used from combining thematic data with analytical hill shading.

have defined a special palette consisting of several series of color shades; each series has a particular hue, but with different values (see Figure 1). Using this color palette, we have devised and implemented a procedure which produces a hill shading image in many colors, with each hue corresponding to a different area type (soil, rock type, etc.). The eight gray levels of the hill-shading image can be drawn using colors 0 to 7. This is equivalent to drawing that image in the lower three bit-planes of the image buffer. Then, a thematic map with up to 7 classes (from a decoded 2DRE file) is drawn in the upper three bit planes without altering the contents of the lower bit planes. On the Terak, this is done by using negative color indices of the form  $-8*i$ ,  $i=1,7$ , where  $i$  represents the thematic class. It is then possible to overlay a series of thematic coverages, without any need to re-compute the hill-shaded 'base-map'. The six bit-planes of the Terak limit the display to at most 8 thematic classes; the 8 discrete levels for each color gun further limit the



display, so that only 7 thematic classes are practical. The method can be extended trivially on more powerful systems.

#### SUMMARY

As a DEM display system, QDEM is little more than an implementation of DEMGS (Kikuchi and others, 1982) in C. However, the change from row-and-column to Morton sequencing for the grid DEM means that DEM data can be efficiently interfaced with other geographic data within a GIS. The use of a full DEM matrix, referenced by Morton numbers, appears to be the most appropriate way to include topographic data in a quadtree-based GIS.

#### ACKNOWLEDGEMENTS

We wish to thank the Fulbright program for their assistance in allowing Juan A. Cebrian to conduct research at SUNY at Buffalo as a visiting scholar from Spain. Also, J.P. Lauzon provided useful comments on the work reported in this paper.

#### REFERENCES

- Abel, D.J., 1984, A  $B^+$ -tree structure for large quadtrees: Computer Vision, Graphics, and Image Processing, 27: 19-31.
- Allam, M.M., 1978, DTM's application in topographic mapping: Proceedings, Digital Terrain Models (DTM) Symposium, ASP-ACSM, St.Louis, Missouri, May 9-11, 1978, p. 1-15.
- Chen, Z.-T., 1984, Quad tree spatial spectrum: Its generation and application: Proceedings, International Symposium on Spatial Data Handling, Zurich, Switzerland, August 1984, vol. 1, pp. 218-237.
- Elassal, A.A., and Caruso, V.M., 1983, Digital elevation models: U.S. Geological Survey Circular 895-B, 40pp.
- Gargantini, I., 1982, An effective way to represent quadtrees: Communications of the ACM, 25:905-10.
- Kernighan, B.W., and Ritchie, D.M., 1978, The C Programming Language: Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 228pp.
- Kikuchi, L., Guevara, J.A., Mark, D.M., and Marble, D.F., 1982, Rapid display of digital elevation models in a minicomputer environment: Proceedings, ISPSR Commission IV Symposium (Auto-Carto V), pp. 297-307.
- Klinger, A., 1971, Patterns and search statistics: in Optimizing Methods in Statistics, J.S. Rustagi, ed., New York: Academic Press.
- Klinger, A. and Dyer, C.R., 1976, Experiments on picture representation using regular decomposition: Computer Graphics and Image Processing 5, 69-105.

Lauzon, J.P., 1983, Two-dimensional Run-encoding for Spatially Referenced Data: unpublished M.A. project, Department of Geography, State University of New York at Buffalo.

Lauzon, J.P., Mark, D.M., Kikuchi, L., and Guevara, J.A., in press, Two-dimensional run-encoding for quadtree representation: Computer Vision, Graphics, and Image Processing, forthcoming.

Mark, D.M., 1979, Phenomenon-based data-structuring and digital terrain modelling, Geo-Processing 1: 27-36.

Mark, D.M., and Abel, D.J., 1984, Linear Quadtrees from Vector Representations: Polygon to Quadtree Conversion: CSIRONET Technical Report #18, Canberra, Australia (also submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence, February 1984; revised October 1984).

Mark, D.M., and Lauzon, J.P., 1984, Linear quadtrees for geographic information systems: Proceedings, International Symposium on Spatial Data Handling, Zurich, Switzerland, August 1984, vol. 2, pp. 412-430.

Mark, D.M., and Lauzon, J.P., in preparation, The space efficiency of quadtrees: An empirical examination including the effects of two-dimensional run-encoding. Submitted to Geo-Processing, July 1984.

Morton, G., 1966, A computer oriented geodetic data base, and a new technique in file sequencing: IBM Canada Limited, unpublished report, March 1, 1966.

Noma, A.A., and Spencer, N.S., 1978, Development of a DMATC digital terrain data base system, Proceedings, Digital Terrain Models (DTM) Symposium, ASP-ACSM, St.Louis, Missouri, May 9-11, 1978, p. 493-505.

Pequet, D.J., 1984, Data structures for a knowledge-based geographic information system: Proceedings, International Symposium on Spatial Data Handling, Zurich, Switzerland, August 1984, vol. 2, pp. 372-391.

Peucker, T.K., Fowler, R.F., Little, J.J., and Mark, D.M., 1978, The triangulated irregular network, Proceedings, Digital Terrain Models (DTM) Symposium, ASP-ACSM, St.Louis, Missouri, May 9-11, 1978, p. 516-540.

Samet, H., Rosenfeld, A., Shaffer, C.A., and Webber, R.E., 1984, Use of hierarchical data structures in geographic information systems: Proceedings, International Symposium on Spatial Data Handling, Zurich, Switzerland, August 1984, vol. 2, pp. 392-411.

Swann, R., Thompson, J., and Daykin, S.E., 1978, Application of low cost dense digital terrain models, Proceedings, Digital Terrain Models (DTM) Symposium, ASP-ACSM, St.Louis, Missouri, May 9-11, 1978, p. 141-155.