IMPLEMENTATION OF TRIANGULATED QUADTREE SEQUENCING FOR A GLOBAL RELIEF DATA STRUCTURE

Jaime A. Lugo
Department of Geography
Mankato State University
Box MSU-2 Mankato
MN 56001, USA

and

Keith C. Clarke
Department of Geology and Geography
Hunter College—CUNY
695 Park Avenue
New York, NY 10021, USA

ABSTRACT

A three-dimensional, multilateral data structure based on Morton sequencing, two-dimensional run-length encoding, and Dutton's Quaternary Triangular Mesh (QTM) was developed, implemented and assessed. This data structure, called the Triangulated Quadtree Sequence, or TQS, was developed to map a global raster dataset onto the surface of a three dimensional solid octahedron. TQS provides the means to translate from Morton Codes to QTM and to TQS structures. To implement the triangulated quadtree sequence, a modified interrupted Collignon map projection was developed to project the world into eight equilateral triangular facets. To assess the TQS, the regular latitude/longitude grid of the land portion of the ETOPO5 global relief database was translated and compressed into the TQS triangular lattice. The validity and usefulness of the model is assessed, and its potential uses discussed.

INTRODUCTION

Why do pixels always have to be squares, or why tessellate with squares? Since the first applications of computer science to automated mapping, cartographers have preferred to measure and model the Earth using square-based coordinate grids and to perform analyses on grids. As a result, most of today's research in hardware (and software) design is based on square structures such as quadtrees (Samet, 1990a). A square-based structure, although excellent for planar geometry and two-dimensional coordinate systems (such as the Cartesian Grid), is one of the least suitable geometric models available for developing a data structure to store, link, and aggregate global scale three-dimensional data. A multilateral data structure using a geometric building object other than the square (such as a hexagon, octahedron, or triangle) is almost always better suited for the modeling of three-dimensional data (Wuthrich and Stucki, 1991).

The main objective of this study was the implementation of a hierarchical indexing method for assigning a unique set of geocodes to a global scale database–given a set of coordinates and the level of resolution desired–and to propose a method by which the indexing system can be developed into a powerful analytical tool for rendering global datasets onto a three-dimensional Earth model. This geocoding method is based on the tessellation of equilateral triangular facets after an initial division of the earth into an octahedral structure, as proposed by Dutton (1984). The storage method is hierarchical, and therefore is able to reference different tessellation levels, with each level having a higher resolution (Tobler and Chen, 1986).

A quadtree structure was used for the proposed planetary indexing system, since the planetary relief model is referenced based on a hexagonal/triangular coordinate system. Various alternate quadtree structures were initially compared, including the B+-Tree (Abel, 1984), Morton Sequencing (Morton, 1966), the PM3' quadtree for vector data (Samet and Webber, 1985), the Quaternary Triangular Mesh (Dutton, 1989a), and many others (Samet, 1990a and 1990b). Each of these quadtrees were weighted using Tobler and Chen's (1986) requirements for a global coordinate system.

The Quaternary Triangular Mesh (QTM) is a spatial data model developed by Geoffrey Dutton for global data modeling and measurement of locational uncertainty (Dutton, 1988). Even though the QTM data structure is similar to the quadtree in that it is based on a fourfold branching hierarchical structure (representable by only two binary digits), similarities end there. Dutton's spatial data model is unique because it uses a triangulated, numerical scheme for geocode addressing within a spherical rather than a planar system (Dutton, 1984).

The QTM is especially important to this study because it matches the design goal of a nontraditional data structure, and has remained a model not yet fully implemented. Although other quadtree variations exist for storage of global data, such as the cubic quadtree structure discussed by Tobler and Chen (1986) and Mark and Lauzon (1985), and Lauzon et al.(1984) these models tend to ignore the modeling of global phenomena in favor of accessing and retrieving map and image data (Dutton, 1989a). Goodchild (1988) also pointed out the need for a spherical model for point pattern analysis based on Theissen polygons and for the tessellation of polyhedral skeletons.

In the current work, Dutton's QTM addressing was implemented, with some modifications, and extended into a quadtree sequence based on recursively divided triangles. The initial mapping was onto a global octahedron in a near equal area projection (the Modified Collignon), and the test implementation consisted of building software and testing it using the land portion of the ETOPO5 global terrain data set.

TRIANGULAR TESSELATION AND QUADTREES

Triangles, unlike squares, have corners that match important surface nodes, and in spherical coordinate systems, their edges touch the global great circles. Therefore, triangle edges can represent linear features that are easily translated into a topological vector structure (e.g., the Triangulated Irregular Network (Peucker, et al. 1978; Peucker, et. al., 1986).

Given the task of representing a global data base, the first step is to build an initial set of triangles. The simplest solid figure with equilateral triangles for sides and with nodes on the surface of a sphere is the octahedron. Each of the facets on the octahedron's surface was tessellated into four geometrically identical triangles. If the edge midpoints of a triangle are connected, the result will be four new triangles–with an identical geometric structure to their parent triangle–every time the original triangle is tessellated. Following the Morton Sequencing quadtree method, Dutton assigned values of 0 (center), 1 (apex corner), 2 and 3 (opposite corners) to each subset triangle (Figure 1). Labeling the subsets from 0 to 3 allows the model to store each tessellated triangle into a 2-bit binary code. Since a triangular tessellation always produces four sub-triangles, it is possible to follow any triangular quadtree address from a chosen global quadrant down to any level of resolution desired. Each facet on the octahedron (or octant) after the first tessellation always has its center triangle labeled 0 and its apex labeled 1, while the other vertices are assigned 2 and 3 (counterclockwise).

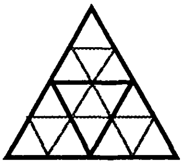


Figure 1a: A tessellated triangle produces more triangles

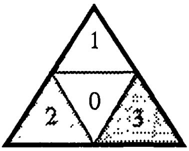


Figure 1b: Facet Breakdown Numbering. (Dutton, 1988)

APPLIED MAP PROJECTIONS

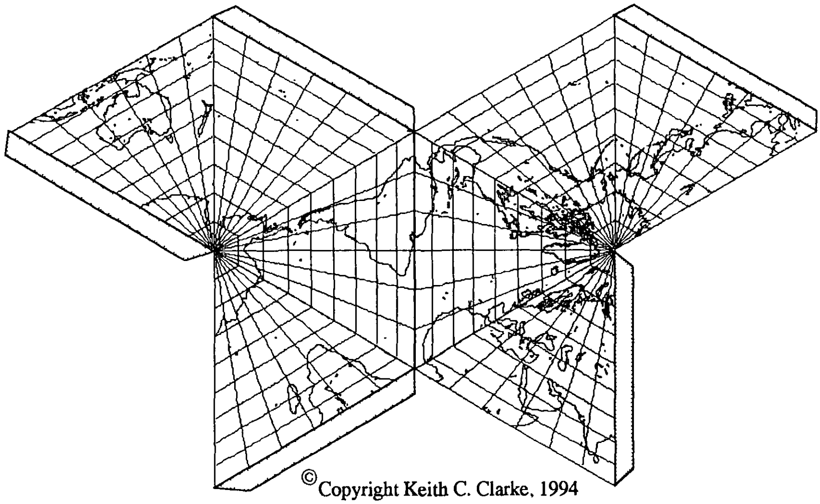
A major concern in developing a planetary relief model from an octahedron is the apparent error that may be introduced if the model does not account for the Earth being a geoid or ellipsoid rather than a perfect sphere. As discussed by Dutton (1984), every time an octahedron is sequentially tessellated it more closely resembles a multilateral globe. Dutton (1984) proposed to solve the geoid dilemma by first tessellating an octahedron to several thousand facets, then projecting it onto a geodesic projection. No specific projection was suggested in any of Dutton's papers (1984 to 1991); he claimed that more research was needed to determine a feasible solution.

To solve the riddle of what projection gives a feasible solution, over one hundred projections were examined, using *An Album of Map Projections* (Snyder and Voxland, 1989). It was determined that only a pseudo-cylindrical, triangular projection with straight parallels and meridians–such as the Collignon Projection–permits the world's outline to be projected onto a three-dimensional octahedron. It is especially important that latitude be represented as near linear steps and orthogonal at all points to the projection's central meridian, for the use of the triangle rotation algorithm in TQS. After several modifications, the Collignon projection was used to georeference both the World Data Bank and the ETOPO5 Global Relief Dataset into eight triangular lattice structures (octants).

The modifications involved on the Collignon require that the globe be first interrupted and that each quadrant be located and tested to determine whether it belongs to the northern or southern hemisphere. If it falls below the Equator, the quadrant is inversely projected (all points y values are multiplied by −1.0). The chosen central meridian is then used to determine the coordinates that define the longitudinal range of the quadrant's base, or equator. A C language program was written using the modified Collignon projection to project the world into eight equilateral triangular octants of an octahedral skeleton. As a result, the octahedral facets may be applied either collectively or individually. An algorithm was written that returned both a global octant number and a projected location within the octant.

The most important modification to this projection was the rescaling of the original projection to fit an equilateral triangle; the Interrupted Collignon is no longer completely equal-area. As a result, the Interrupted Collignon provides the specific Collignon location for any given latitude/longitude (x,y) coordinate in an equilateral triangle lattice. This feature not only locates all coordinate points on the globe, but serves as the method by which Morton, QTM and TQS addresses may be issued to every node on the octahedral skeleton. The quadrants can be seen in Clarke's Butterfly Projection (see Figure 2), where each quadrant is arranged by translation and rotation to resemble a butterfly. This projection can also be assembled into the shape of a three-dimensional paper octahedron.

Figure 2: The Interrupted Collignon(Clarke's Butterfly) Projection.



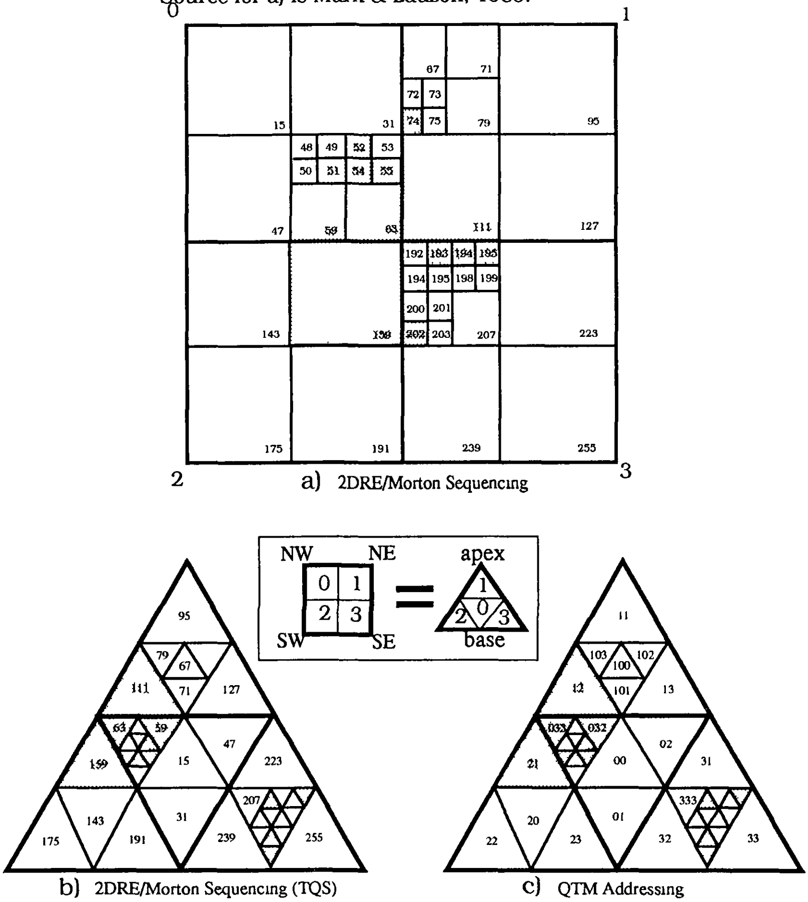
The Triangulated Quadtree Structure (TQS)

The proposed three-dimensional model requires a quadtree structure capable of georeferencing coordinates and elevation data to the triangle nodes within the octahedral structure, up to any level of tessellation. Such a quadtree must also be capable of linking large numbers of entities, attributes and relationships between the dataset and the octahedral facets. The Triangulated Quadtree Sequence is a combination of Morton sequencing, two-dimensional run-length encoding, and the QTM structure, modified to fit the previously defined specifications.

In order to store the outline of the world in the octahedral structure of the Interrupted Collignon Projection, the geocode given to each triangular quadrant must follow a static pattern similar to a rectangular Morton sequence (see Figure 3a). With a static pattern it is then possible to automatically assign Morton numbers and QTM addresses simultaneously. First, the center facet of each tessellated triangle was labeled zero [0] (Dutton, 1988); its north-south orientation is always opposite to the location of the central triangle's base. Second, the apex (labeled one [1]) shares the same base with the center facet (note that both triangles face in opposite directions).

Finally, the remaining facets are labeled by assigning the numbers two [2] and three [3] in a counterclockwise direction beginning at the apex. This step differs from Dutton's labeling scheme, where facets [2] and [3] may be arbitrarily labeled (Dutton, 1988). The result is a static addressing pattern that allows a triangular tessellation to be stored directly into a quadtree structure and vice versa, and which can also be stored using either Morton or QTM numbers. As seen in Figure 3, a raster image was originally stored using Morton sequences in a two-dimensional run-length encoding (2DRE) format. Note that the sequencing scheme used is logically identical to the Morton/2DRE addressing method in Figure 3a. The only difference is the applied geometric base. The next step is to apply the QTM addressing method. Once the raster image has been transferred from a square to a triangular lattice, it is a simple matter of substituting the Morton numbers for QTM addressing numbers. The results are shown in Figures 3b and 3c.

Figure 3: Although both a) and b) and c) are identical tessellations, a) is a 2DRE/Morton Quadtree, b) is a Triangular Quadtree Structure (TQS), and c) is a QTM triangular structure. The original image shown in 3a was directly imposed on 3b following the addressing pattern between Figures 3a and 3b. Each number shown in every one of these figures labels the highest value assigned to each particular quadrant, and based on the original image. Source for a) is Mark & Lauzon, 1985.



IMPLEMENTATION AND TESSELLATION OF THE TRIANGULATED QUADTREE STRUCTURE (TQS)

Recursive Tessellation for Determination of a TQS Address

First, each point is transformed from its original (x,y) coordinate (squared) lattice into a triangular lattice with the modified Collignon algorithm. The original raster image of the ETOPO5 dataset was transformed and stored into a set of eight separate octant quadtree files (one octant per quadtree). Each of the quadtrees in turn represents one of the facets of the global octahedron. After each latitude and longitude coordinate was computed from the original image (a 16-bit ERDAS "lan" format file, figure 4) and transformed into a triangular Collignon coordinate; the C program assigned a QTM address to each point within ETOPO5. The Collignon coordinate was then tested recursively to determine where within

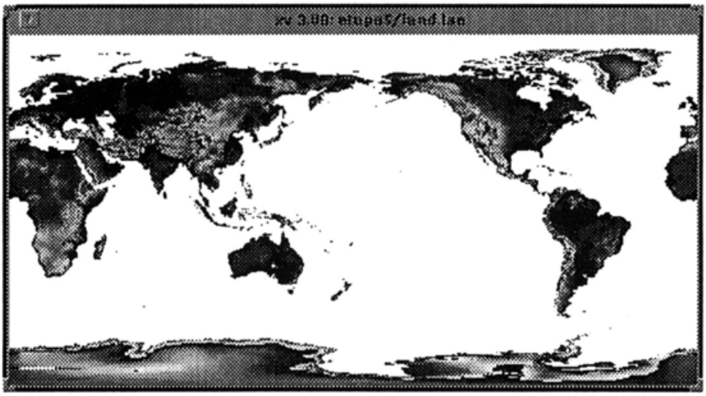
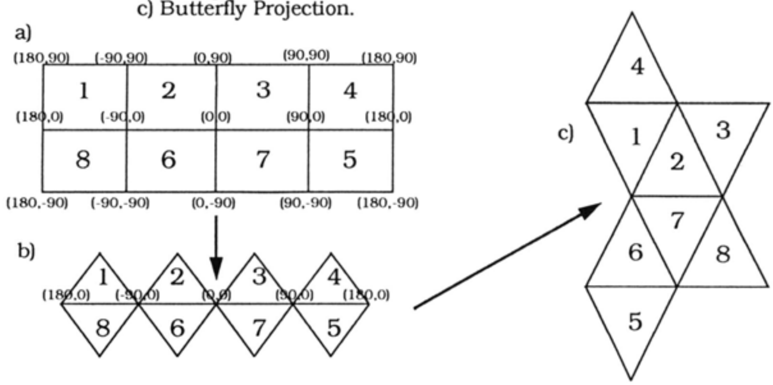


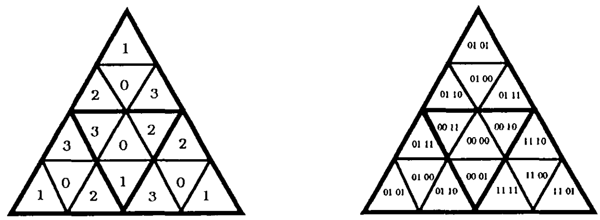
Figure 4: The original ETOPO5 raster image, global topography, land portion.

each triangle (facets 0, 1, 2, or 3) did the point belong. To save computing time, the maximum number of tessellations needed was determined before a TQS address was assigned for each point. Once the desired resolution was known, each point was transformed from a square-based lattice to the Collignon triangular lattice. Figure 5 shows the boundaries of an octant after a point is fitted through the Collignon algorithm. The most prominent feature is the grouping of points as one moves from the triangle base up to the apex (see Figure 2).

Figure 5: Transformation from a) square-based raster image, to
b) triangle-based lattice structure, to
c) Butterfly Projection.



In Figure 5a, the ETOPO5 data points are located in a rectangular, latitude-longitude model. Each rectangular octant was then projected into the model shown in Figure 5b, which is later arranged into its proper position within Figure 5c through eight separate rotation algorithms. The implementation of the triangulated quadtree structure began with the assignment of QTM addresses to each ETOPO5 point. The addressing scheme used differs from Dutton's model in that triangle addresses are always assigned in a counterclockwise order–Dutton's schema varied from triangle to triangle (Dutton, 1984; 1991). The order of facets through each tessellation is shown in Figure 6. This method was selected for two reasons: 1) a regulated order is essential if TQS is to be linked to any type of hierarchical structure, and 2) the computer algorithm method used to detect triangle location is based on this ordering scheme.



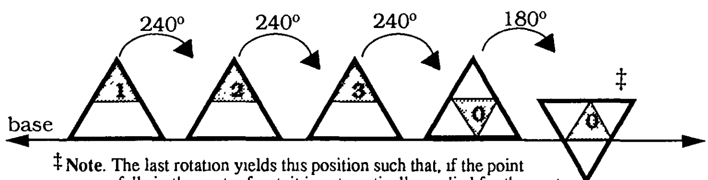
a) QTM: Address Ordering

b) QTM: Bit-Addressing

Figure 6: TQS Addressing Scheme.

The function assigns TQS addresses to ETOPO5, point by point, by determining whether the y-value falls between the midpoint's y-range and the maximum y- value. If so, this point is 1) assigned a QTM number representing its triangle location, and then 2) its x- and y-values are divided in half, along with the x- and y-ranges. This coordinate division allows for the subsequent tessellation of this point's location until the full QTM address is determined up to the level of resolution desired, or maximum depth is achieved; see Figure 7. If the point fails to fall above the midpoint range, then the point is continually rotated 240^{o} counterclockwise until the correct corner triangle location is determined. And, if it again fails to fall in a corner triangle, then by default it must fall in the center triangle. Therefore, by continually rotating and tessellating each x- and y-value, this method assigns a quaternary hierarchical address to every point in the dataset.

Figure 7: Triangle Rotations and Translations
Triangle rotation and translation for a point falling in the center facet. Note that the last translation inverts the point, then reduces its value by half (literally moving the point to the next tessellation level).



‡ Note. The last rotation yields this position such that, if the point falls in the center facet, it is automatically readied for the next tessellation.

Each time a point gets translated to determine whether it belongs to the facet that is currently above the midpoint range, it must be rotated counterclockwise 240^{o}, its x- coordinate gets subtracted one fourth of the xrange value, and its y-coordinate gets added one half the y-range value. These additions and subtractions translate the octant back to its original position, except that the next triangle in the tessellation order is now at the apex (above the midpoint range). Note that while it seems that a full triangle gets translated, only one point gets translated at a time. Triangles are used to help visualize the area being processed. Since geometric translation and rotation of points requires that all calculations be performed in radians, each point is first translated into radians, and then rotated using standard affine rotation formulae. Each point went through this process twelve times at increasing resolution to determine the full TQS address (12 tessellation levels).

Discussion of the TQS Addressing Method

Dutton's Quaternary Triangular Mesh provided the means to develop and assign a TQS address for every coordinate point on the ETOPO5 global relief dataset. Each TQS address can be stored into a maximum of three bytes (24 bits), for a global resolution of approximately 4,883 meters, or 5 minutes. This is accomplished by assigning a 2-bit value (from 0 to 3) each time a point gets tessellated. This 2-bit value gets bit-shifted to the left (twice) and then appended to the previous value until the TQS address consists of twenty four binary digits (0 or 1). Therefore, TQS addressing eliminates the need for a pointer to each tessellation level since each pair of binary digits represent each level–with the leftmost pair indicating the root node. As a result, this method allows the user to recursively select any level of resolution desired.

The TQS address is now explained. Beginning from the left (root node) of the TQS number, each pair of binary digits represents the image's tessellation level up to the maximum pixel resolution; the two rightmost digits indicate the tessellation depth with the highest resolution. The TQS number is compressed even further by storing it in hexadecimal, rather than binary notation. This method introduces another source of compression into the geocoding process, though. When the binary number gets translated into hexadecimal notation, all leading zeroes get truncated from the number.

The result is a much smaller number than originally anticipated. For example, say that a point falls in TQS binary number `00 00 00 01 01 11 10 01 00 00 00`. Translated into hexadecimal notation, this number becomes [`5E40`]. If this number is again translated into a binary location, it becomes `1 01 11 10 01 00 00 00`. Note that, instead of the original 24 bits, the number now consists of only 15 bits. The missing levels are computed by adding the number of zeroes missing to the left of the binary address, until the number of binary pairs is equal to the maximum depth desired by the user (24 bits in this example). Sorting the TQS strings by length then allows maximal compression and multiresolution resampling.

Results From the Application of the TQS Addressing Method

The raster image of ETOPO5 as an ERDAS 16 bit lan file is 18,662,913 bytes. The small size of this data set comes from the fact that coordinates are implicit by order from the grid, not explicit. As such, reordering or sorting for resolution becomes impossible. With 4320 columns and 2160 rows, this file would expand to 214,617,600 bytes in ASCII with explicit coordinates before all points below sea level were excluded. Using just the land segment of the file, designed to test the multi-resolution aspect of TQS, left an explicit reference coordinate file of 52,254,720 bytes. The TQS, by storing an exact locational code for each point in hexadecimal notation, reduced the amount of storage needed to 48,540,279 bytes at twelve levels of recursion, equivalent to the same approximate resolution, stored as eight individual files (one for each octant).

Since the original coordinates were excluded from the octant files, the resulting files contain implicit location, instead of explicit location (octant number, elevation, and TQS address; no coordinates), and use binary form, not ASCII. Instead of storing the floating point elevation value (plus its link) required for representing each point in ASCII format at a 5-minute resolution, the TQS method stored the same point into a 24-bit binary value (three bytes instead of sixteen bytes). At a 5-minute resolution, the original coordinate and elevation values (in ASCII format) were stored as `+/-DD.MM +/- DDD.MM EEEE[EOL]` Degrees, Minutes, Elevation or the equivalent of twenty bytes. The same value in TQS format ranges from one to nine bytes, depending on resolution. ETOPO5 was further simplified for TQS by eliminating all those points with identical x- and y-coordinates within the ERDAS image file.

Elimination of data redundancy within a raster image was crucial because the original raster image adds identical x, y and elevation values as the image reaches the poles to account for the stretch caused by transposing the Earth's surface onto a flat grid. For example, the x, y and elevation values are identical for every pixel along the latitudes for the north and south poles, while every point is unique along the Equator. The Interrupted Collignon projection, unlike the original image, transposed the (x,y) coordinates closer and closer together as the points reached the poles, thus compressing the high latitudes into the apex of each triangle. Although it may seem as if accuracy is lost by concentrating more points in smaller and smaller areas, actually there is no more space at the poles than anywhere else on Earth. The error of reduction in areal resolution can be solved by implementing a three-dimensional planetary relief model (see Dutton, 1988; Lugo, 1994).

CONCLUSION

In conclusion, the Triangulated Quadtree Sequence (TQS) structure successfully linked the land portion of ETOPO5 to the Interrupted Modified Collignon Projection. The accuracy and precision of the TQS addressing numbers increased with every tessellation, until each address represented each elevation point up to the resolution of the original dataset. Once the georeferenced data is stored in the triangulated quadtree structure, areas with little relief can be generalized, while areas with high relief can be further tessellated up to the highest level of resolution.

The result has been a model that fitted the Earth's surface onto the surface of a map projection that was easily transformed into an octahedron. This accomplishment will eventually lead to a three-dimensional model capable of providing an image of the globe for any number of attributes, including population density, income, meteorological data, spread of disease, and others. In other words, this model may be viewed as a 2.5-dimensional choropleth map of the world. On the other hand, while this method offers simplicity and geolocational precision, it lacks the ability to measure scale throughout the octahedron's surface. As the levels of tessellation change, so does the length of every triangular segment, meaning that scale continues to change throughout the globe's surface and becomes impossible to measure (Dutton, 1989b). In other words, while a coordinate may tell the user where an object is, it fails to reveal the actual size of the object.

Despite this weakness in scale measurement, the achievement of the stated goals makes the proposed application a new step toward the development of a fully functional multilateral data structure for three-dimensional study. Although people today take for granted that a coordinate system is a always a discrete rather than continuous set of points, there will always be feasible alternatives to be researched, developed and implemented.

REFERENCES

Abel, D. J. (1984). "A B+-Tree Structure for Large Quadtrees." *Computer Vision, Graphics and Image Processing*, v. 27, pp. 1-18.

Dutton, G. (1984). "Geodesic Modelling of Planetary Relief." *Cartographica*, v21, Numbers. 2 & 3, pp. 188-207.

Dutton, G. (1988). "Modeling Locational Uncertainty via Hierarchical Tessellation." ed. Goodchild Michael and Sucharita Gopal. *Accuracy of Spatial Databases*. New York: Taylor & Francis, pp. 123-140.

Dutton, G. (1989a). "Computational Aspects of Quaternary Triangular Meshes." Prime Park, MS: Prime Computer, Inc. (unpublished).

Dutton, G.(1989b). "The Fallacy of Coordinates." Prime Park, MS: Prime Computer, Inc. (unpublished).

Dutton, G. (1991). "Improving Spatial Analysis in GIS Environments." *Proceedings Auto-Carto 10*, 1991, pp. 168-185.

Goodchild, M.(1988). "The Issue Of Accuracy In Spatial Databases." *Building Databases for Global Science*, Mounsey, H., and Tomlinson, R. (editors). London: Taylor & Francis, pp. 31-48.

Lauzon, J., D. Mark, L. Kikuchi, and A. Guevara (1985). "Two-Dimensional Run-Encoding for Quadtree Representation." *Computer Vision, Graphics and Image Processing*, v30, pp. 56-69.

Lugo, J. A. (1994). "Implementation of a Triangulated Quadtree Structure for a Three-Dimensional Planetary Relief Model." M.A. Thesis. New York: Hunter College of the City University of New York.

Mark, D. M. and J. P. Lauzon (1985). "Approaches for Quadtree-based Geographic Information Systems at Continental and Global Scales." *Proceedings Auto-Carto 7*, Falls Church, VA: ASPRS/ACSM, pp. 355-364.

Morton, G. M. (1966). "A Computer Oriented Geodetic Data Base, and a New Technique in File Sequencing." IBM Canada Ltd.

Peucker, T. K. Fowler, R. J., Little, J.J., and Mark, D. M. (1986). Digital Representation of Three-dimensional Surfaces by Triangulated Irregular Networks (TIN). Technical Report Number 10, United States Office of Naval Research, Geography Programs.

Peucker, T. K., et al (1978). "The Triangulated Irregular Network." *Proceedings of the Digital Terrain Models Symposium of the American Society for Photogrammetry*/American Congress on Surveying and Mapping. St. Louis, Missouri, USA, pp. 516-540.

Samet, H. (1990a). *Design and Analysis of Spatial Data Structures*. New York: Addison-Wesley Publishing Company.

Samet, H. (1990b). *Applications of Spatial Data Structures*. New York: Addison-Wesley Publishing Company.

Samet, H. and R. Webber (1985). "Storing a Collection of Polygons Using Quadtrees." *ACM Transactions on Graphics*, Vol. 4, pp. 182-222.

Snyder, J. P. and P. M. Voxland (1989). *An Album of Map Projections*. Reston, VA: Department of the Interior; U.S. Geological Survey, Denver, CO.

Tobler, W. and Z-T Chen (1986). "A Quadtree for Global Information Storage." *Geographical Analysis*, v18(4), pp. 360-371.

Wuthrich C. A., and P. Stucki (1991). "An Algorithmic Comparison Between Square and Hexagonal-Based Grids." *Computer Vision, Graphics and Image Processing: Graphical Models and Image Processing*. v53(4), pp. 324-339.