Lecture Notes in Earth Sciences 79

Geoffrey H. Dutton

A Hierarchical Coordinate System for Geoprocessing and Cartography

With 85 Figures, 16 Tables and 2 Foldouts

Springer

Author

Geoffrey H. Dutton
Department of Geography, Spatial Data Handling Division
University of Zurich-Irchel
Winterthurerstrasse 190, CH-8057 Zürich/Switzerland
E-mail. dutton@geo.unizh.ch

"For all Lecture Notes in Earth Sciences published till now please see final pages of the book"

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Dutton, Geoffrey H.:
A hierarchical coordinate system for geoprocessing and cartography / Geoffrey H. Dutton. - Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 1999
(Lecture notes in earth sciences ; 79)
ISBN 3-540-64980-8

This monograph was a Ph.D. dissertation approved by the Philosophical Faculty II of the University of Zürich under the title "Working Through the Scales: A Hierarchical Coordinate System for Digital Map Data"

Die vorliegende Arbeit wurde von der Philosophischen Fakultät II der Universität Zurich im Wintersemester 1997/98 aufgrund der Gutachten von Prof. Dr. Robert Weibel und Prof. Dr. Kurt Brassel als Inaugural-Dissertation angenommen.

ISSN 0930-0317
ISBN 3-540-64980-8 Springer-Verlag Berlin Heidelberg New York

Typesetting: Camera ready by author
SPIN: 10685802 32/3142-543210 - Printed on acid-free paper

For *William Warntz*
Buckminster Fuller
Benoit Mandelbrot

Everything the same; everything distinct
— Zen proverb

Preface

Working on a Ph.D. can be lonely, frustrating, discouraging. In my own case, the work was both made easier and more difficult by having been involved in researching the subject matter of my dissertation for more than fifteen years: easier because I had worked out many of the basic concepts and techniques by the time I began the most recent phase; harder because there was by then a lot of material to integrate, and my main application area (map generalization) has so many complexities and vexing unsolved problems, some of which I attempted to tackle head-on. While I usually managed to budge them, the problems didn't always yield.

The main reason for my good fortune in being able to bring this project to some sort of closure is my connection to vital communities of extraordinary people, who if they did not actively encourage and support my efforts, gracefully tolerated my obscure and idiosyncratic visions with good humor, and possibly even respect. And during my time in Switzerland, I have been able to work steadily and uninterrupted, refreshed by periodic reunions with my family back in America. The unconditional support of my wife, Aygül Balcioglu Dutton and my mother, Sophie Pincus Dutton have nourished my spirit throughout my expatriation. I owe both of them so much.

There are so many others who have helped, in many ways. While this work has been a very personal odyssey, others have collaborated with me from time to time and supported the project. I am especially grateful to the National Center for Geographic Information and Analysis (NCGIA) for encouraging this work, by including me in specialist meetings for Research Initiatives 1, 3, 7 and 8, which prompted me to explore interesting avenues related to this project and resulted in several publications. In particular, I wish to thank Michael Goodchild (Santa Barbara), Barbara Buttenfield (Buffalo, now at the University of Colorado), and Kate Beard (Orono) for their efforts. It was Mike who encouraged Yang Shiren and others to develop the first working implementation of my model, brought us together to work on it, and reported on our progress in several venues. Dr. Shiren was a fountainhead of clever, practical ideas for implementing my data model, and I am highly appreciative for his collaboration. The resulting software prototypes demonstrated to me possibilities I had not considered, and convinced me of the essential value and practicality of these concepts. I am equally indebted to Barbara Buttenfield for inviting me to NCGIA Buffalo on several occasions to collaborate; there the central concepts that led to this project were distilled and the map generalization algorithms were first developed, some of which remain at the core of the current project.

All of the above would have come to naught had it not been for my advisors in Zürich, Kurt Brassel and Robert Weibel. Together they did managed the logistics to overcome various obstacles in my path, and have been extremely supportive and encouraging to me all the way. While in residence, my principal advisor has been Dr. Weibel, whose keen and detailed criticism and useful suggestions have helped to keep me on the right track. I am also grateful to the Swiss National Science Foundation for financial support for this research, through grants 2000-037649.93 and 2100-43502.95 to the Department of Geography. I am indebted to my colleague Frank Brazile as well; his cheerful skepticism prevented me from getting carried away too far with each new putative insight, and he also has brought to my attention tools that have been useful and could become more so as my project grows and matures. Frank's help in computing spatial orderings presented in appendix E is warmly appreciated.

Day in and day out my department proved to be a friendly atmosphere and close to an ideal environment in which to work: I was provided with a workstation, a computer account, access to libraries and software packages, and basically left alone to follow my research agenda. There has been no coursework or exams to distract me, just occasional interesting colloquia, a few administrative chores and the task at hand. Having seized this opportunity, I hope that this and other results of my efforts justify my advisors' and my family's faith in my capabilities.

Still, I have not managed to accomplish everything I had hoped for when this project commenced two years ago. My goals then — and now — included developing a more comprehensive set of tools to solve contextual map generalization problems, and while a good amount of conceptual progress was made toward this end, its implementation is far from complete, indeed barely begun. But this has been a work in progress for close to fifteen years now, and there is no reason for me to expect a sudden dénouement. It is my fond hope that others will find something of value in my work, and will take up the challenge of making it more operational and useful in an increasing variety of contexts. If I can help any of these efforts move forward, it would be my pleasure to collaborate.

Especially, I would like to thank Luisa Tonarelli for her patience, good humor and expertise in guiding my efforts to publish my dissertation in this series.

Geoffrey Dutton — June 1998
Zürich — QTM ID: 1133013130312011310

Abstract

The work reported in this thesis has sought to determine if a scale-specific notation for location can benefit GIS and cartography. After discussing some of the problems caused by relying on standard coordinate notations, the quaternary triangular mesh (QTM), a hierarchical coordinate system based on a planetary polyhedral tessellation is presented. This model encodes terrestrial locations as leaf nodes in a forest of eight triangular quadtrees, each of which can have up to 29 levels of detail (doublings of scale). Its multi-resolution properties are described in relation to common planar and spherical coordinate systems used for geographic location coding. One of QTM's advantages derives from its inherent ability to document the locational accuracy of every point in a spatial dataset. Another is its ability to retrieve such data across a range of spatial resolutions, enabling users to evaluate the fitness of map data for use at particular scales. In addition, the tessellation's spatial partitions can be exploited to identify regions where crowding of features occurs at certain scales. After describing this system's origins and comparing it to related data models, techniques for encoding and decoding locations, accuracy assessment, spatial conflict detection and other operations are presented. The focus then moves to the use of QTM to generalize digital map data for display at smaller scales. After a brief overview of current approaches to generalization, a strategy for using QTM for line simplification and conflict detection is outlined. A basic strategy for using QTM for line simplification was developed that is embodied in an algorithm having eight control parameters. This method is capable of using embedded spatial metadata to guide selection of retained points, and a way to encode such metadata is described; exploitation of point-specific metadata was seen as one of the more successful and promising aspects of the strategy. A series of empirical tests of this algorithm using several island shoreline datasets is described and illustrated. Outcomes of these experiments are compared with results produced by an existing, widely-used line simplification algorithm. Similarities and differences in results are discussed, attempting to identify reasons for failure or success of either method. The thesis concludes with an evaluation of QTM's potential contributions to this application area, along with a description of some of the current challenges and limitations to spatial data handling that QTM may be able to address. Additional overviews of contents may be found by consulting Scope of Work in chapter 1, as well as the summaries at the end of each chapter.

Zusammenfassung

Die vorliegende Arbeit geht der Frage nach, ob eine neue, massstabsunabhängige Notation für Positionen (Koordinaten) in Geographischen Informationssystemen (GIS) und in der Kartographie von Vorteil sein könnte. Zunächst werden einige der Probleme herkömmlicher Korrdinatennotationen untersucht, um dann das "Quaternary Triangular Mesh (QTM)" als Alternative zur Beschreibung und Speicherung räumlicher Daten einzuführen. QTM ist ein hierarchisches Koordinatensystem, das auf einer planetaren polyedrischen Tessellation basiert. Ausgehend von einem Oktaeder, beschreibt es Positionen auf der Erdkugel als Blätter in einem Wald von acht Dreiecks-Quadtrees (entsprechend den acht Oktaederflächen). Jeder Quadtree kann bis zu 29 Auflösungsstufen umfassen (je Stufe verdoppelt sich der Massstab). Die besonderen Eigenschaften dieser hierarchischen Datenstruktur, insbesondere die Möglichkeit mehrstufige Auflösungen innerhalb einer Datenstruktur zu repräsentieren, werden in Relation zu herkömmlichen planaren und sphärischen Koordinatensystemen beschrieben. Ein wesentlicher Vorteil dieses Datenmodells ist, durch seine Struktur die Genauigkeit eines jeden Datenpunktes beschreiben zu können; ein weiterer, dass man Daten für verschiedene Auflösungen extrahieren und so auf einfache Weise die Tauglichkeit von Daten für verschiedene Kartenmassstäbe evaluieren kann. Ausserdem eignet sich die mehrstufige Raumaufteilung gut, um Gebiete mit zuviel Detail in einem bestimmten Massstabsbereich zu erkennen. Nach der Diskussion der Ursprünge des QTM-Modells und dem Vergleich mit verwandten Systemen beschreibt die Arbeit mehrere Algorithmen zur Kodierung und Dekodierung von Positionen, Genauigkeitsabschätzung, Ermittlung räumlicher Konflikte (für die kartographische Generalisierung), sowie weitere Basisoperationen. Danach wendet sich die Arbeit vornehmlich der Verwendung von QTM zur digitalen Generalisierung kartographischer Daten zu. Nach einer kurzen Diskussion bestehender Ansätze für die Generalisierung wird eine Strategie für die Verwendung von QTM für die Linienvereinfachung und für die Erkennung potentieller Konflikte zwischen kartographischen Objekten vorgestellt. Unter anderem wurde eine Methode für die QTM-basierte Linienvereinfachung entwickelt, die zu einem Algorithmus geführt hat, der von acht Parametern gesteuert wird. Dieser Generalisierungsalgorithmus kann durch Metadaten, die im QTM abgelegt werden, gezielt gesteuert werden und so sicherstellen, dass wesentliche Punkte einer Linie erhalten bleiben. Eine Methode für die Extraktion von Metadaten zur Beschreibung der Sinuosität kartographischer Linien und deren Kodierung im QTM wird ebenfalls beschrieben. Die Verwendung von punktbezogenen Metadaten zur

Steuerung der Generalisierungsalgorithmen hat sich als sehr vorteilhafter und vielversprechender Aspekt der vorliegenden Strategie erwiesen. Dies wird auch durch eine Reihe empirischer Tests erhellt, die anhand verschiedener Datensätze von Küstenlinien durchgeführt wurden. Die Ergebnisse dieser Untersuchungen werden einem herkömmlichen, weit verbreiteten Linienvereinfachungsalgorithmus gegenübergestellt. Ähnlichkeiten und Unterschiede in den Resultaten der zwei Ansätze werden diskutiert, um so die Gründe für Erfolg oder Misserfolg der jeweiligen Methode zu identifizieren. Die Arbeit schliesst mit einer Diskussion konkreter Anwendungsbereiche für QTM und seinem potentiellen Beitrag zur Lösung aktueller Probleme Geographischer Informationssysteme und in der Kartographie. Weitere Hinweise über den Inhalt der vorliegenden Arbeit können dem Kapitel 1 (Abschnitt Scope of Work) sowie den Kurzzusammenfassungen am Ende jedes Kapitels entnommen werden.

Contents

Preface VII

Abstract IX

Zusammenfassung X

1 Introduction and Problem Statement 1

1.1 Motivation 1

1.2 Limitations to Detail and Accuracy of Spatial Data 2

1.2.1 Scale Limitations 3

1.2.2 Precision Limitations 5

1.3 Documenting Spatial Data Quality 7

1.3.1 Data Quality Information and Map Generalization 10

1.3.2 Encoding Geospatial Data Quality 11

1.4 Challenges of Digital Map Generalization 15

1.5 Scope of Work 16

1.6 Chapter 1 Summary 18

2 Historical and Conceptual Background 19

2.1 A Typology of Spatial Data Models 20

2.2 A Global, Regular, Hierarchical Spatial Data Model 23

2.2.1 Local and Global Properties 26

2.3 Prior, Related and Derived Work 28

2.3.1 Polyhedral Map Projections 28

2.3.2 Global Environmental Monitoring 32

2.3.3 Global Terrain Modeling 33

2.3.4 Related Independent Research 36

2.3.5 Derivative Models and Applications 37

2.4 Chapter 2 Summary 39

3 Computational Aspects of a Quaternary Triangular Mesh 41

3.1 Geometric, Topological and Numerical Properties 42

3.1.1 Variance in Area of QTM Facets 42

3.1.2 QTM Quadrant and Node Numbering 46

3.1.3 QTM Coordinate Conversion 48

3.1.4 Structuring QTM Identifiers 52

3.1.5 Spatial Access Strategies 55

3.2 Attractors: Knitting Hierarchies Together 57

3.2.1 Attractor Formation and Naming 59

3.2.2 Attractors as Containers 61

3.3 Documentation and Discovery of Positional Accuracy 62

3.3.1 Method-induced Positional Error 63

3.3.2 Unearthing Positional Error 65

3.3.3 Revealing Fractal Properties 67

3.4 Chapter 3 Summary 69

4 Using QTM to Support Map Generalization 71

4.1 Map Generalization: Methods and Digital Implementations 72

4.2 Aspects of QTM-based Generalization 75
4.2.1 Evaluating and Conditioning Locational Data 75
4.2.2 Line Simplification 77
4.2.3 Spatial Conflict Identification 79
4.3 Parameterizing Line Generalization 88
4.3.1 Regulating the Amount of Detail 88
4.3.2 Controlling the Quality of Detail 89
4.4 Vector Line Generalization Using QTM 89
4.4.1 Data Modeling 90
4.4.2 Parametric Control of QTM Generalization 93
4.4.3 Employing Vertex-specific Metadata 96
4.5 Chapter 4 Summary 102

5 QTM Scale Filtering: Tests and Evaluations 105

5.1 Evaluating QTM Line Generalization 105
5.1.1 Software Testbed 106
5.1.2 Test Data 108
5.1.3 Methodology 110
5.2 Empirical Map Generalization Experiments 111
5.2.1 Results for Bainbridge Island, WA 115
5.2.2 Results for Nantucket Islands, MA 119
5.3 Evaluation and Discussion of Findings 122
5.4 Chapter 5 Summary 129

6 New Approaches to Cartographic Questions 131

6.1 On Cartographic Lines and Their Simplification 131
6.1.1 Imperatives of Scale 131
6.1.2 Weeding and Conditioning Data 132
6.1.3 Local versus Global Strategies 132
6.1.4 Choosing Tolerance Values 133
6.1.5 Optimizing Performance 133
6.1.6 Assessing Performance 134
6.2 Important Compared to What? 135
6.2.1. Generating Good Gestalts 135
6.2.2 Comparing Regional and Local Sinuosity Selection 138
6.3 Topics for Continuing Research 144
6.4 Looking Back, Looking Ahead 145

A Quaternary Triangular Mesh Data Processing 149

Encoding and Decoding Locations: *QTMeval* and related algorithms 149
Projective Geometry Algorithms 157
Simplifying Feature Data with QTM: Point Selection Algorithms 159

B Basic, Computational and Invented Geospatial Terms 165

C Generalization Experiments: Statistical Results 173

1.1 Bainbridge, WA: Summary Statistics for QTM Generalization Experiments
1.2 Bainbridge, WA: Distribution of Sinuosity Values after Generalization
1.3 Bainbridge, WA: Length Statistics for Coastline
2.1 Nantucket, MA: Summary Statistics for QTM Generalization Experiments
2.2 Nantucket, MA: Distribution of Sinuosity Values after Generalization
2.3 Nantucket, MA: Length Statistics for Coastline

D Generalization Experiments: Cartographic Results 183
1.1 Bainbridge WA occupies portions of QTM facets 31202322 and 312102320
1.2 Bainbridge Island, WA: Non-hierarchical QTM Generalization Results Using QTM Facets without Point Lookahead
1.3 Bainbridge Island, WA: Hierarchical QTM Generalization Results Using QTM Facets without Point Lookahead
1.4 Bainbridge Island, WA: Non-hierarchical QTM Generalization Results Using Attractors without Point Lookahead
1.5 Bainbridge Island, WA: Hierarchical QTM Generalization Results Using Attractors without Point Lookahead
1.6 Bainbridge Island, WA: Non-hierarchical QTM Generalization Results Using Attractors with Point Lookahead
1.7 Bainbridge Island, WA: Hierarchical QTM Generalization Results Using Attractors with Point Lookahead
1.8 Bainbridge Island, WA: QTM Facet, Attractor and Bandwidth-tolerance Generalization Methods Compared
1.9 Bainbridge Island, WA: QTM Attractor Generalization Compared with Source Data at Derived Scales
1.10 Bainbridge Island, WA: Effects of QTM Lookahead and Sinuosity Parameters on World Vector Shoreline Data
1.11 Bainbridge Island, WA: Effects of Specifying Preferred Sinuosity for Non-hierarchical and Hierarchical QTM Line Simplification Using Three Equalized Classes and Smoothing
2.1 Nantucket, MA Occupies portions of QTM Facets 4022320 and 4022321
2.2 Nantucket, MA: Non-hierarchical QTM Generalization Results Using QTM Facets without Point Lookahead
2.3 Nantucket, MA: Hierarchical QTM Generalization Results Using QTM Facets without Point Lookahead
2.4 Nantucket, MA: Non-hierarchical QTM Generalization Results Using Attractors without Point Lookahead
2.5 Nantucket, MA: Hierarchical QTM Generalization Results Using Attractors without Point Lookahead
2.6 Nantucket, MA: Non-hierarchical QTM Generalization Results Using Attractors with Point Lookahead
2.7 Nantucket, MA: Hierarchical QTM Generalization Results Using Attractors with Point Lookahead
2.8 Nantucket, MA: QTM Facet, Attractor and Bandwidth-tolerance Generalization Methods Compared
2.9 Nantucket, MA: QTM Attractor Generalization Compared with Source Data at Derived Scales
2.10 Nantucket, MA: Effects of QTM Lookahead and Sinuosity Parameters on World Vector Shoreline Data
2.11 Nantucket, MA: Effects of Specifying Preferred Sinuosity for Non-hierarchical and Hierarchical QTM Line Simplification
2.12 Nantucket, MA: Sinuosity Analysis of Shoreline Using Three Equalized Classes and Smoothing

E Efficiency of QTM as a Spatial Index for Point Data 211
1 Spatial Ordering Statistics for 60 Cities and Towns, ordered by City Name
2 Spatial Ordering Statistics for 60 Cities and Towns, ordered by Latitude
3 Spatial Ordering Statistics for 60 Cities and Towns, ordered by Longitude
4 Spatial Ordering Statistics for 60 Cities and Towns, ordered by QTM ID
5 Spatial Ordering Statistics for 60 Cities and Towns, ordered by Shortest Path
6 Spatial Ordering Statistics for 60 Cities and Towns, ordered by Longest Path
7 Distance from Prior City on List, by Type of Ordering
8 Cumulative Distance for City Tours and Efficiency Scores

References 221

Index 229

List of Figures

1.1 Two types of discretization errors, along with human line tracing error 3
1.2 Scale-limited representational problems 5
1.3 Scaling moves point P outside triangle 6
1.4 Overlay of data with differing positional accuracy 12
1.5 Alternative ways to encode positional metadata 14
2.1 An iconic typology of spatial data models 21
2.2 Quaternary Triangular Mesh to Three Levels of Detail 12
2.3 Form and orientation of the quaternary triangular mesh 24
2.4 Da Vinci's *mappemonde* from c. 1514 28
2.5 Hexahedral world projection, vertices at poles 29
2.6 Clarke's Modified Collignon ("Butterfly") Projection 30
2.7 Zenithial OrthoTriangular World Projection 31
2.8 EMAP Polyhedral Global Sampling Model 32
2.9 Structure of the Geodesic Elevation Model 34
2.10 Relations between GEM's dual triangular hierarchies 35
2.11 Comparison of SQT and HSDS tessellations 37
2.12 The Semi-Quadcode QTM variant 38
3.1 Geometric Development of QTM Levels 1 and 2 from an Octahedron 42
3.2 Areas of QTM Facets for Levels 2-5 on a Unit Sphere 45
3.3 QTM node and facet numbering to level 2, ZOT projection 47
3.4 Two stages of encoding a point location 50
3.5 Five Options for locating coordinates decoded from a QTM ID 51
3.6 Enriched QTM ID 64-bit binary word structure 53
3.7 Quadtree global decomposition for HHCODE spatial indexing 57
3.8 Four levels of attractors occupied by a QTM facet 59
3.9 Four levels of vertex attractors for a polygon with 74 points 61
3.10 Systematic Error from QTM Point Decoding at Level 20 65
3.11 Total number of vertices for Swiss canton boundaries by QTM level 66
3.12 Swiss Canton and lake map data 67
3.13 Fractal analysis of Swiss canton boundaries ("Richardson diagram") 68
4.1 Bainbridge WA and Nantucket MA: No. of vertices by QTM resolution 77

4.2 QTM generalization via facets and attractors 78
4.3 Associating attractors to an arc using a minimum bounding triangle 82
4.4 A Directory for Associating Attractors to Primitives 86
4.5 Structure of Dual Attractor/Primitive Directory Used in the Study 87
4.6 A Feature-based Spatial Data Model for Map Generalization 91
4.7 Computing QTM Hierarchical Coordinates for vertices of a primitive 92
4.8 Estimating local sinuosity at specific polyline vertices 98
4.9 Non-linear transformations of sinuosity compared 99
4.10 Filtering Classified sinuosity Values to identify homogenous regimes 101
5.1 Bainbridge and Nantucket Islands: 1:80K source data at 1:295K 109
5.2 Bainbridge Island WA: Three simplification methods compared 117
5.3 Bainbridge Island: Relative distribution of Vertex Sinuosity 118
5.4 Nantucket MA: Three simplification methods compared 119
5.5 Nantucket Island: Relative distribution of Vertex Sinuosity 120
5.6 Mean segment size on ground and in map units at standard scales 124
5.7 Schaffhausen Canton, Switzerland: QTM and RDP generalizations 126
5.8 Effects of Varying QTM Retrieval and Point Decoding Resolutions 128
6.1 Bainbridge Island, 1:80K source generalized to QTM level 13 140
6.2 Bainbridge Island, 1:250K source generalized to QTM lvl 13 141
6.3 Nantucket Island, 1:80K source generalized to QTM level 13 142
6.4 Nantucket Island, 1:250K source generalized to QTM level 13 143

List of Tables

2.1 Quaternary triangular mesh statistics for 30 levels of geometric detail 25
2.2 Orientation of the QTM basis octahedron to the earth 27
3.1 Octant Area Statistics for QTM Levels 1-5 on a Unit Sphere 44
3.2 Attractor Identifiers for a pair of nearby locations in Switzerland 60
3.3 QTM Level 20 Positional Error Statistics for Transects 64
3.4 Mean Level 20 Positional Error by QTM Level 1 Quadrants 65
4.1 Generalization operators applied to types of spatial data. 72
5.1 Statistics for Nantucket MA and Bainbridge WA Test Datasets 110
5.2 Characteristics of Simplification Algorithms Compared 125

Abbreviations and Acronyms

AID	attractor identifier
CSL	classified sinuosity, local
CSR	classified sinuosity, regional
CSV	classified sinuosity value
DBMS	database management system
DEM	digital elevation model
DEPTH	delta-encoded polynomial terrain hierarchy

DLG	digital line graph
DTM	digital terrain model
EMAP	Environmental Monitoring and Assessment Program
EPA	(U.S.) Environmental Protection Agency
FGDC	(U.S.) Federal Geographic Data Committee
FID	Feature Identifier
GEM	geodesic elevation model
GIS	geographic information system
HSDS	hierarchical spatial data structure
MBT	minimum bounding triangle
MEL	mesh element
NMAS	national map accuracy standard
NOAA	(U.S.) National Ocean and Atmospheric Agency
PID	primitive identifier
PSV	preferred sinuosity value
QTM	quaternary triangular mesh
RDP	Ramer-Douglas-Peucker (algorithm)
SDTS	(U.S.) Spatial Data Transfer Standard
SQT	sphere quad tree
TIN	triangulated irregular network
TQS	triangulated quadtree surface
USGS	United States Geological Survey
UTM	universal transverse Mercator
WWW	world wide web
ZOT	zenithial ortho triangular (projection)