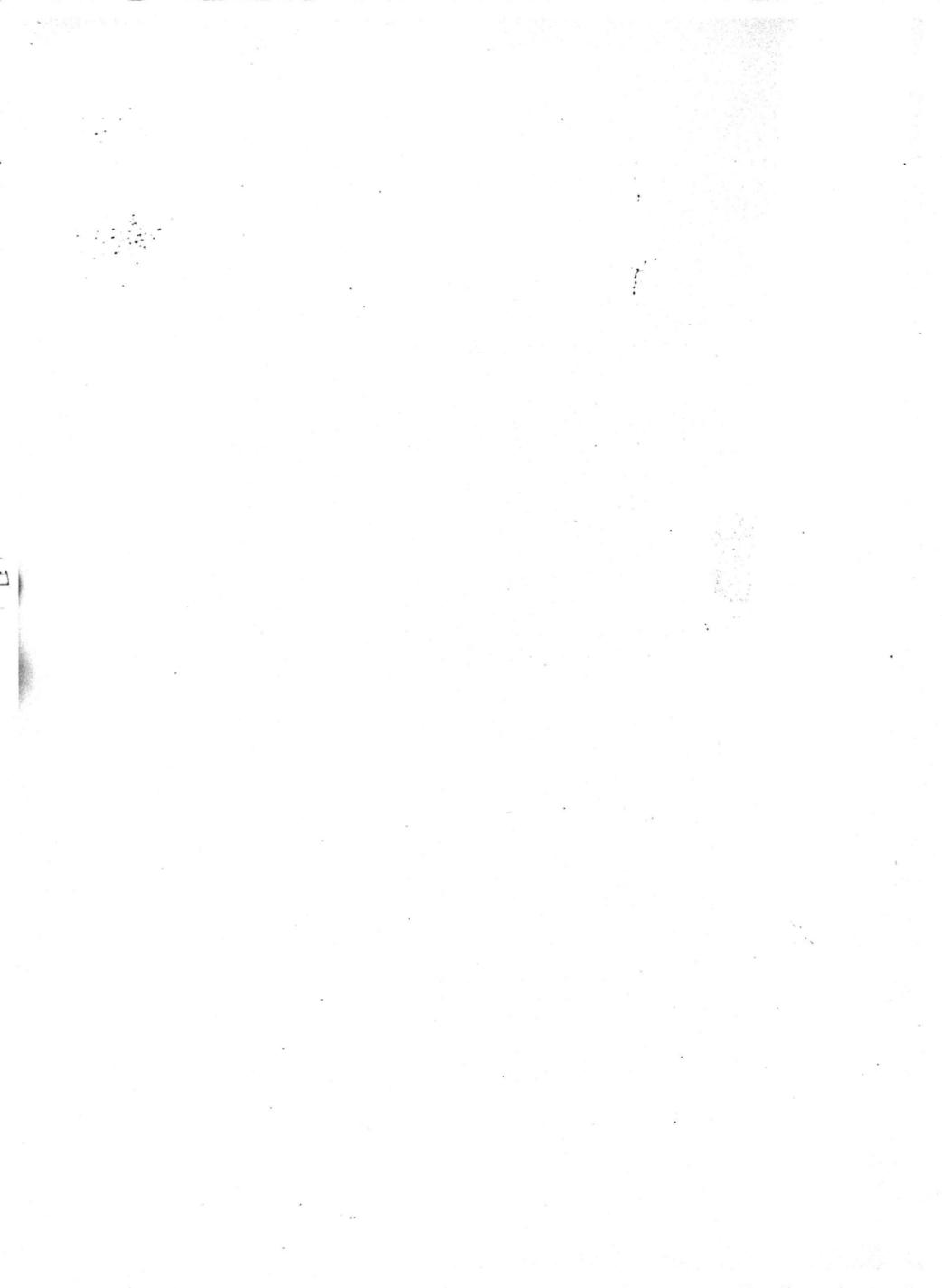
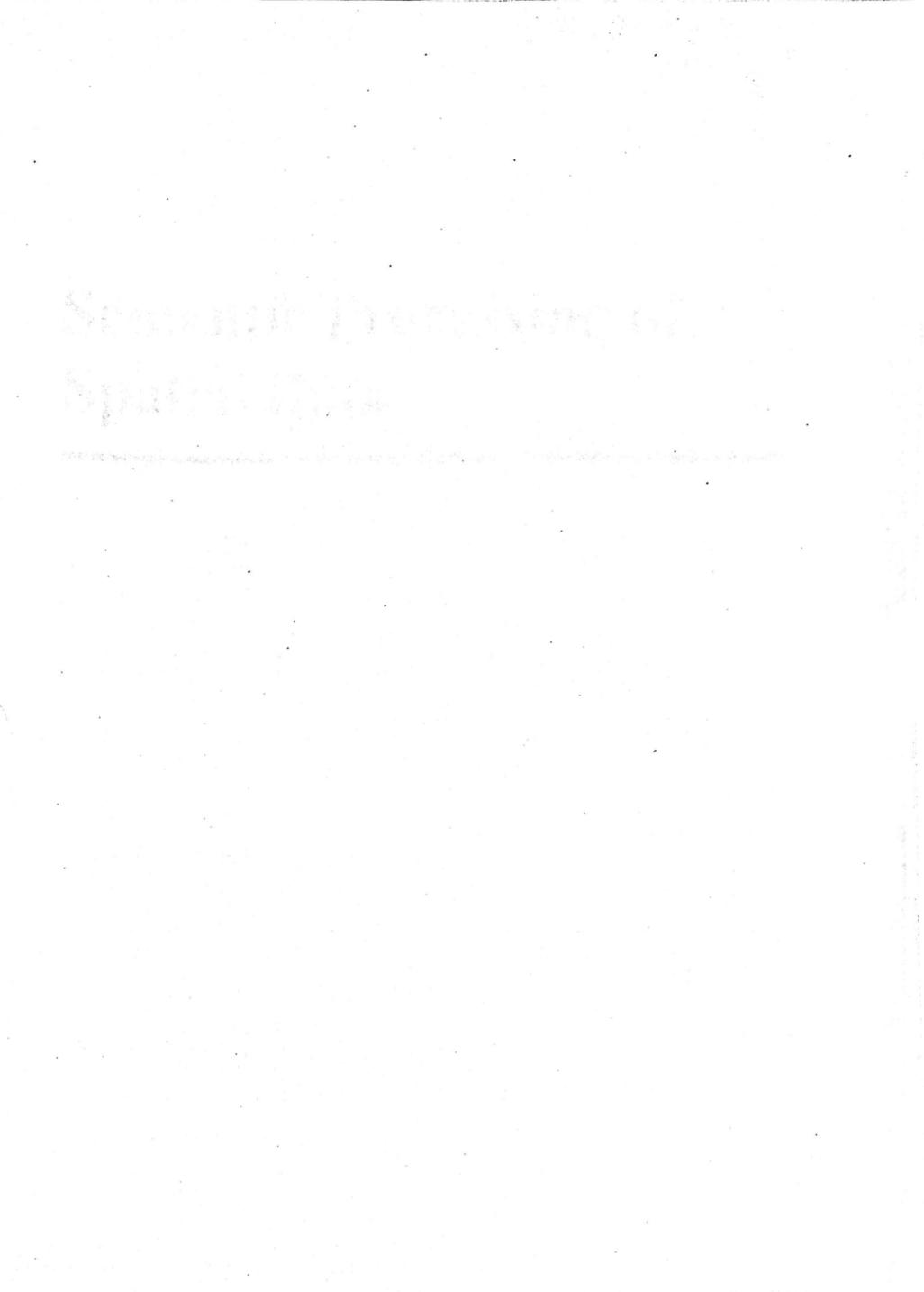
RESEARCH ON COMPUTING SCIENCE

Semantic Processing of Spatial Data

Serguei Levachkine Jean Serra Max Egenhofer

Instituto Politécnico Nacional Centro de Investigación en Computación





Semantic Processing of Spatial Data

Research on Computing Science

Editorial Board

Serguei Levachkine Jean Serra Max Egenhofer

Edition

Miguel Torres Rolando Quintero Marco Moreno

Cover Design

Miguel Torres Rolando Quintero Marco Moreno

Semantic Processing of Spatial Data

Serguei Levachkine Jean Serra Max Egenhofer (Editors)

Instituto Politécnico Nacional Centro de Investigación en Computación México, 2003

apagan garang duning-hour, barantup baranan da amada ma







Copyright © by Instituto Politécnico Nacional, 2003

Published by Instituto Politécnico Nacional Centro de Investigación en Computación Av. Juan de Dios Bátiz s/n esq. Miguel Othón de Mendizábal Unidad Profesional "Adolfo López Mateos", Zacatenco 07738, México D.F., MÉXICO

http://www.ipn.mx http://www.cic.ipn.mx http://geopro.cic.ipn.mx

The editors and publishers of this book have made their best effort in preparing this book, but make no warranty of any kind, expressed or applied, with regard to the information contained in this book.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by means electronic, mechanical, photocopying, recording or otherwise, without the prior permission of Instituto Politécnico Nacional. Except for personal or classroom use provided that copies bear the full citation notice provided on the first page of each paper.

Printing: 100 Printed in Mexico

TABLE OF CONTENT

		Page
PREF	ACE	i
1. RA	STER-TO-VECTOR CONVERSION	
	V. Alexandrov and S. Levachkine, "Cognitive promptings for semantic-mind analysis and object-oriented data integration of information flows"	2
	S. Levachkine and V. Alexandrov, "Semantic-mind analysis and object-oriented data integration of visual information: A primer"	11
	S. Levachkine, "System approach to R2V conversion for analytical GIS"	22
	A. Velázquez and S. Levachkine, "Text/Graphics separation in raster-scanned color cartographic maps"	34
	A. Gelbukh, SangYong Han, and S. Levachkine, "Using sources of evidence to resolve ambiguities in toponym recognition in cartographic maps"	. 42
	E. M. Felipe Riverón and M. E. Ramos García, "Enhancement of digital color halftone printed images"	52
	J. Angulo and J. Serra, "Mathematical morphology in color spaces applied to the analysis of cartographic images"	59
	E. Bodansky, "System approach to a R2V conversion: From research to commercial system"	67
2. AU	TOMATIC MAP GENERALIZATION	75
	A. Ruas, "Spatial analysis and agent principles to automate generalization process"	76
	D. Lee, "Generalization within a geoprocessing framework"	82
	T. Cheng and Zhilin Li, "Measures for semantic quality after polygon generalization"	92
	M. A. Moreno Ibarra and S. Levachkine, "Automated digital map description for generalization"	102

3. AUTOMATIC GEO-IMAGE INTERPRETATION			
U. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen, "Multi-resolution, object-oriented fuzzy analysis of remote sensing data to GIS ready information"	110		
O. Pogrebnyak and P. Manrique Ramirez, "Image filter for efficient impulse noise removal and fine detail preservation"	127		
J. Ren, Y. Shen, and L. Guo, "A novel image retrieval"	133		
SooHyun Cho, YunKoo Chung, and JaeYeon Lee, "Mosaicing system with strong of robustness in camera motion"	140		
R. Quintero Tellez, "Increasing resolution of Digital Elevation Models using bicubic parametric patches"	145		
4. INTELLIGENT GIS TOOLS	151		
S. Mustière, N. Gesbert, and D. Sheeren, "A formal model for the specifications of geographic databases"	152		
S. Balley, "Assisting users in selecting and restructuring data sets"	160		
M. J. Torres Ruiz and S. Levachkine, "Generating spatial ontologies based on the spatial semantics"	169		
C. Granell Canut, J. Poveda, and M. Gould, "Incremental weak composition and invocation of geographic web services"	179		
J. Poveda and M. Gould, "Multidimensional binary indexing for neighbourhood calculations in spatial partition trees"	188		
J. Feng, N. Mukai, and T. Watanabe, "Multi-level transportation network adaptable to traffic navigation"	197		
Lau Bee Theng and Wang Yin Chai, "Sketch spatial query representation model"	205		
M. Pech, D. Sol Martínez, and J. González, "Adaptation and use of spatial and non-spatial data mining"	215		
C. Zepeda Cortés and D. Sol Martinez, "An implementation of spatio-temporal annotated constraint logic programming (STACLP)"	221		
A. F. Razo Rodríguez and D. Sol Martínez, "GISELA: a web-based interface for risk management GIS using XML"	226		
A. Guevara, "The spatial enabling of information paradigm"	229		

5. GIS	SAPPLICATIONS	*	237
	Á. M. Felicísimo and J. Muñoz, "Dispersed by Wind Project: A Wind Fields"	A Motion Analysis over	238
	E. Ortiz, E. Ramirez, A. Rodríguez, E. Villaseñor, and J.M. I assess spatial distribution of Mexico City emissions inventory"	Herrera, "Using GIS to	245
AUTI	HOR INDEX		249
REV	EWERS		253

and the second of the second o

and the second state of the second of the acceptable of the accept

REMEDIAL.

2011 1 41 / 12

The state of the s

and the second of the second o

the state of the s

and a common state of the second state with

. The second of the second problem of the control of the second of the s

e per la la la cilabria. Se las estas e

graph the second of the second control of the second control of the second control of the second control of the

and the second of the second o

and the second of the second o

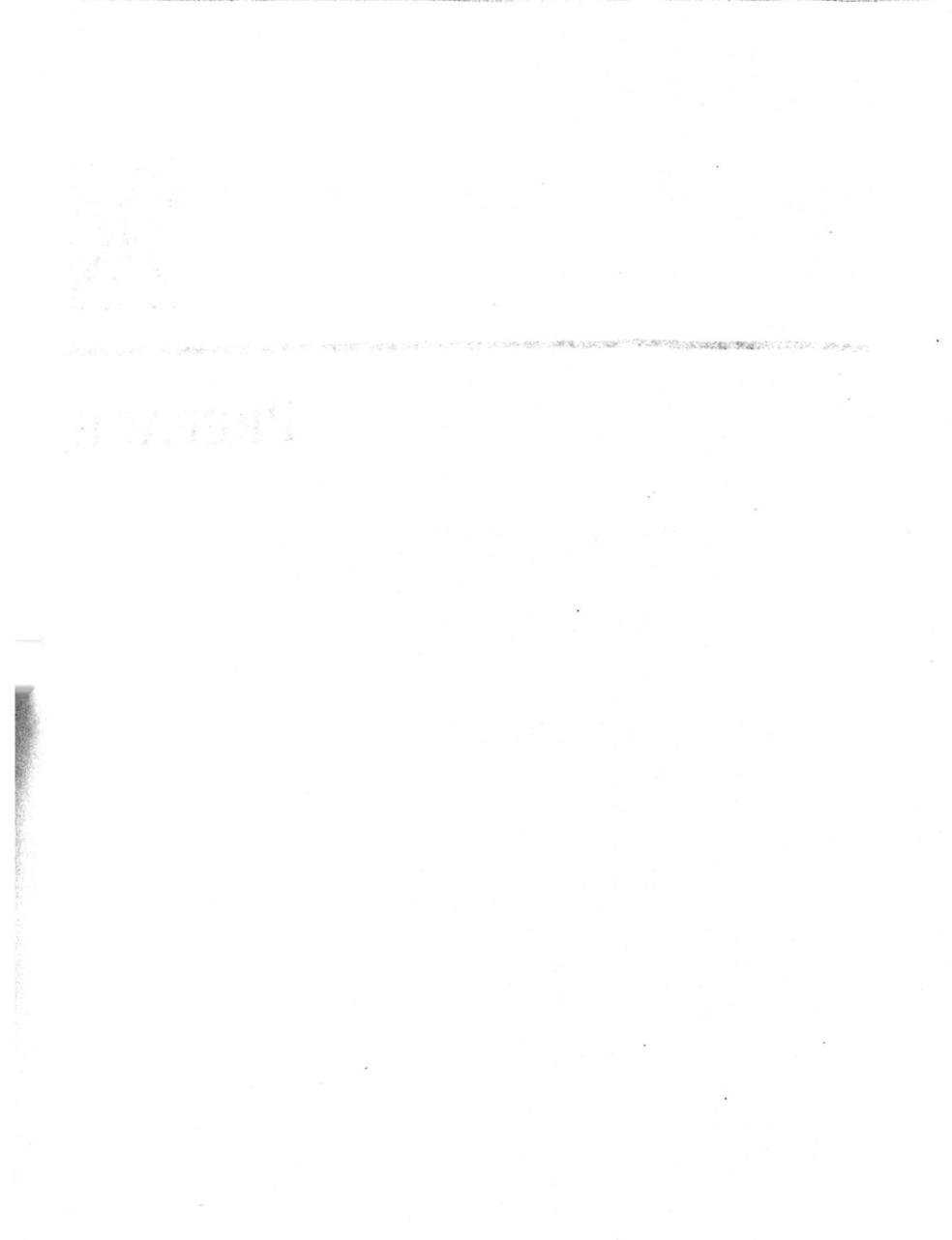
en de la companya de

en la companya de la La companya de la co



PREFACE

n de la companya de



This book contains refereed and revised papers presented in GEOPRO 2002 and GEOPRO 2003, the International Workshops on Semantic Processing of Spatial Data, which took place in Mexico City, Mexico in December 2002 and November 2003. Semantic processing of spatial data is an emerging area of Geomatics that focuses on representation and use of knowledge to support (semi) automatic Geographic Information Systems (GIS). Due to the growing demand for both interactive and automatic intelligent GIS, the field of semantic processing of spatial data has an exciting and promising future. The GEOPRO workshop is relatively small but unique Geomatics forum that exclusively concerns with this topic in the fields of raster to vector conversion, map generalization, image interpretation, spatial databases, and GIS applications. Excellent keynote speakers are another peculiarity of GEOPRO. We believe that the appearance of such event is extremely timely.

GEOPRO is addressed to the international community of researchers and practitioners at all levels of experience to share insights into semantic processing of spatial data. The workshops welcome strong participation from researchers in both industry and academia. GEOPRO was sponsored by National Polytechnic Institute (IPN), Centre for Computing Research (CIC), Mexican Petroleum Institute (IMP/PIMAYC), Mexico, and Environmental System Research Institute (ESRI), USA.

The response to the workshop was overwhelming that ensures its promising future (77 submitted papers). Every paper submitted to GEOPRO was evaluated by at least two reviewers (30 accepted papers). We are grateful to both authors and reviewers for their careful work during this review process. Many of the papers that appear in this book were thoroughly revised and improved in response to reviewers' suggestions.

This book is organized into five sections, reflecting the sessions in the GEOPRO workshops: Raster-to-vector conversion (8 papers), Map generalization (4 papers), Geo-image interpretation (5 papers), Intelligent GIS tools (11 papers), and GIS applications (2 papers). Each session included a keynote talk and/or tutorial given by recognized specialists.

We are gratefully acknowledge the support provided by the sponsors of GEOPRO: IPN (Instituto Politécnico Nacional, Mexico), CIC (Centro de Investigación en Computación, Mexico), IMP/PIMAYC (Instituto Mexicano del Petroleo/Programa en Matemáticas Aplicadas y Computación, Mexico), CONACYT (Consejo Nacional de Ciencia y Tecnología, Mexico), ESRI (Environmental System Research Institute, USA), IGN (l'Institut Géographique National, France), Ecole des Mines de Paris, Centre de Morphologie Mathématique, France. Our special thanks to Dr. Juan Luis Díaz de León - Director of CIC, Dr. Rodolfo Suárez - Director of IMP/PIMAYC, and the members of Geoprocessing Laboratory of CIC for their noble support.

See you at GEOPRO 2004!

November 3, 2003

Serguei Levachkine Miguel Torres Workshop Organizers

and the second of the second control of the second of the

to page for an experience of the second of the granded of the second of

per control of the control of the state of the state of the control of the contro

(AME OF YORK IS approved

atti i sagrani

RASTER-TO-VECTOR CONVERSION

COGNITIVE PROMPTINGS FOR SEMANTIC-MIND ANALYSIS AND OBJECT-ORIENTED DATA INTEGRATION OF INFORMATION FLOWS

Victor ALEXANDROV*

*Saint Petersburg Institute for Sciences 39 14th Line, Saint Petersburg, Informatics and Automation-Russian Academy of 199178 RUSSIA

E-mail: alexandr@mail.iias.spb.su

ABSTRACT

An emerging area of digital data processing is the computer-based intelligent analysis of information flows. In this paper, we discuss some cognitive promptings that can lead to intelligent data (audio, images, video, mixed) analysis and synthesis. A companion paper in this book describes how to extract semantic components from unordered data sets (Gestalt problem) in visual information data (Analysis) and an application of our approach is illustrated with a raster-scanned color cartographic map interpretation system—Analogical-to-Raster-to-Vector (A2R2V).

KEY WORDS

Semantics, Information flow, Visual data, Semantic-mind analysis, Object-oriented data integration.

Think about meaning...
The words are becoming themselves.

Lewis Carroll

1. INTRODUCTION

The use of computers for data processing, storage, and transmission in a historic perspective includes the following subsequent stages [17]:

- Data (temporal series, matrices, etc.) representation and processing using traditional mathematical models. The search for computer-oriented mathematical models in signal and image processing.
- -Generalization of the concept of data. Software development of storage, search, and reference for texts. Development of computer programs of translation from one language into another.
- Development of decision-making tools and expert systems.

Serguei LEVACHKINE⁺

[†]Department of Artificial Intelligence Centre for Computing Research-IPN UPALMZ, Ed. CIC, Mexico City, 07738 MEXICO

E-mail: palych@cic.ipn.mx

- Continuous data integration (text, audio-video) leads to the subsequent stage of software development intelligent agents such as Data Mining, Copernicus, etc.

This was being thought about among researchers since the early 1980s. Simon [18] wrote: "The arrival of a third generation of machines made it possible to experiment with information (data) provided by instruments in many fields of observation: in visual images, in spoken words, in the fields of physics, medicine, economics, linguistics, etc. A new field of observation and study was transferred from philosophy to experimental sciences. This was a general and crucial phenomenon in the history of Man's efforts to understand Nature: the telescope gave birth to astronomy and metallurgical, chemical, electrical and vacuum techniques gave birth to physics. There is a certain paradox in the fact that the computer, designed for commercial accounting and scientific computation, gave birth to pattern recognition and artificial intelligence".

For the pure mathematician, data recognition is a trivial problem that can be expressed formally as follows: Let X be a representation space, preferably a "nice" topological space, and let Ω , the interpretation space, be a finite set of names. Recognition or identification is a mapping $E: X \to \Omega$, to which certain properties are described and from which elegant theorems can be deduced.

This, however, is not where the problem lies: in practice, the question is one of constructing E, i.e., of providing operators or programs which given any x from X, enables us to decide automatically onto which ω from Ω the element is mapped.

An extended definition of E, to be held in full memory, is out of the question even for small-scale problems, because here we come up against the problem of computational complexity. In the search for usable operators, data recognition is continually confronted with problems of information complexity; so many data recognition problems are exponential that we are constantly obliged to adopt less than optimal, polynomial solutions. Data recognition is first and foremost a battle against complexity.

The other guiding thread appears to us to be the semantics of the general data recognition problem, which varies according to the question under consideration. Is there in fact a general, universally applicable method for constructing a data recognition operator?

We must therefore treat each problem in a specific manner and search for any items of information that will enable us to construct the required operators. Our view is that information is to be found in the properties of

- the representation space,
- the interpretation space, or spaces.

However, the solution of applied problems requires the bridge between the strong structure of mathematical concept of space and empirical properties of the data as input information that leads us to the following conclusion: *semantic information* is to be found in the properties of

- data representation;
- · data interpretation as knowledge of subject domain.

This means that we use the following model of a recognition process: find an element from a finite set, which is equivalent to the unknown object. When the element is found, the object is attributed with its all known properties. Essentially, this is the principle of "identification by indistinguishability", first formulated by Leibnitz [19].

Information flow (IF) is binary digital data stored and processed by the computer. This is the basic element of digital technologies, i.e., Turing's (Computer) world. On the other hand, IF is data that obtain from the outer (to computer) world and that have very different representations—environmental monitoring, text, music, speech, images, etc.; thus, IF is a part of the Human's world. At present, the problem to adapt computer systems to human perception and cognition is apparent (MPEG-7), but yet not carried out. In the present work, we would like to discuss some crucial aspects (Semantic-mind analysis (SMA) and Object-oriented data integration (OODI) of IF) of this problem, outline the difficulties encountered on the way to its solution, and guide the reader toward the bridge between computer and human worlds. All this can be observed within the general context of human-machine interaction as a semantic approach to computer data analysis and synthesis.

In [1], we considered one aspect of this approach. The goal of the present work is to sketch the general frameworks of SMA (analysis) and OODI (synthesis) in different types of information flows. We apply the proposed approach for the case of the visual information flows and in particular for cartographic raster data [20].

2. DATA STRUCTURE AND INFORMATION

The concepts of data, knowledge, and information have held the steadfast attention of scientists during centuries. Although, knowledge as a research topic is one of oldest in the history of science, we have only a general philosophical understanding without a strong formal definition. Such a definition is required for knowledge-based computer systems (e.g., on-line and Internet education).

The main difficulty in formal definition of these concepts probably lays in the fact that man as an information carrier and translator has very specific structure of their representation and processing and thus cannot easily abstract and detect these concepts from a structured shell.

Known concepts of data, knowledge, and information explicitly show their direct link to the form of representation. "A knowledge representation is a certain method in which the experience becomes structurally defined in wide-spread and common terms" (Marvin Minsky).

Research concerning knowledge structuring can be divided into two lines. The first is related to the study of individual intelligence, but it does not yet systemize it. Therefore, this research does not provide an entire description of knowledge representation in the human brain. On the other hand, collective intelligence is much more accessible for research because numerous forms of knowledge representation have been constructed over the centuries.

Several examples of the social experience of mankind show a high degree of influence on the final result of man's activity with regard to the choice of one form or another of knowledge representation. A typical example in science occurs when two scientists making the same discovery cannot understand this fact because they use different forms (notations) of knowledge representation. To avoid such situations, it is important to represent any information in compact graphic form that reflects its basic essence. The representation must provide a possibility to determine certain details of information flow using as tools that are as simple as possible. We would like to highlight here that this way provides good results not only in education or knowledge popularization but also directly in practice and science. This is probably why computer science specialists prefer computer programs that provide precise details of processing to long-winded ambiguous explanations or fogware.

Thus, research of the interrelation between information content and its structured representation is important from both theoretic and practical points of view.

¹Formal definitions of SMA and OODI will be given in §3.

2.1 INFORMATION

When investigating an object, anyone is stated (probably in an intuitive level) three questions:

- · What is this?
- How is this related to something that I know?
- · What are the characteristics of this object?

In other words, the human being attempts to identify the object, to order new information that corresponds to his early experience, and finally detects qualitative properties of the object. He partially obtains this information in an empirical manner by analogy with other objects in his eigensystem of representation. This is an example of the developing system of knowledge representation.

Problems of identification, ordering, and semantics appear in any research but not necessarily in this order. As a rule, these problems are repetitively applied to define (and modify) the solutions of each problem obtained in the previous step. We shall consider these three problems as follows and attempt to define their meaning within the context of structured representation of information.

2.2 IDENTIFICATION

Identification is the oldest human activity and at first glance represents sufficiently well a simple procedure of name assignment of objects, processes, and concepts. If $X^* = (x_1^*, x_2^*, ..., x_n^*)$ is a representation of original object, X is the representation space and $\Omega = (\omega_1, \omega_2, ..., \omega_p)$ the set of names, then identification is mapping ξ from the representation space to the set of names:

$$\xi: X \to \Omega$$
.

Of course, the procedure of mapping should be constructive. However, in attempting to apply such a procedure we encounter several problems:

- Should we assign to any object a name? (the problem of Leibnitz' "identification by indistinguishability");
- How novel and informative should a concept or process to assign it a name be? (the problem of label), and
- How should new objects, concepts and processes be named for further easy use (the problem of name structuration)?

These three problems graphically illustrate that identification is not a mechanical procedure of name assignment (although we should perform such mechanics for pointer construction, for example). Thus, in subsequent sections we shall also discuss these three problems in relation to the concept of structure.

2.2.1 LEIBNITZ' PRINCIPLE OF IDENTIFICATION BY INDISTINGUISHABILITY

Identification is related to classification, association, analysis, and synthesis of knowledge, i.e., it is a necessary link in cognition and probably is the only link: "A hard classification implies the hardness of words. A soft classification implies an inflexible look at things" (Charles Luttwidge Dodgson).

The first problem is essentially Leibnitz' principle of identification by indistinguishability: "Two objects are indistinguishable if all their properties are the same" (Gottfried Wilhelm Leibnitz).

The problem of identification was first strongly formulated by Leibnitz in his Ph.D. thesis and led him to the concept of congruence: $ABC \cong ABX$. Congruence or universal characteristics according to Leibnitz mean "... man can draw while not being an artist..." The method consists of the rule for identifying X as a set of points, objects, and things and their properties by the name and properties of known object C. At first glance, this method provides unlimited possibilities for learning and cognition. However, anyone who frequently uses the Internet for scientific information search constantly requires conceptual identification—comparison of things searched for with known information that can be represented in different forms and languages.

Moreover, if we use Leibnitz' principle we always must take into consideration the following two limitations:

Limitation 1. Absolutely identical objects do not exist. An infinite number of properties characterize any real object. Therefore, we use only a subset of properties when comparing objects or rejecting others. A given problem or subject domain defines the choice of properties, which are essential for object comparison; hence, the same objects can be identical with respect to one subject domain and different with respect to another. Thus, the first problem is choice of identification properties, and

Limitation 2. Values of quantitative and qualitative properties can only be defined approximately, e.g., the weight is approximately kg, the color is blue, etc. Thus, indistinguishable objects at the crude level of computing or measurement of parameter values can be distinguishable only after re-computing and re-measuring. Thus, the second problem is exactness of representation of object properties.

Sense of the principle of identification by indistinguishability is detection of object identity by known properties. Applications of this principle are different in different subject domains: in algebra—relations of equality and identity; in geometry—relations of congruence and similarity; in linguistics—synonyms; in statistics—coefficients of correlation, etc.

2.2.2 THE PROBLEM OF LABEL

This problem is well illustrated by Di Bono: "Having a name, the knowledge fragments become 'frozen' and 'untouchable' because the label can only be used for constant value. We must consider a world that is constructed from names such as house—from the bricks, which must be broken into parts and investigated separately to understand the whole". This is the point! The reader encounters difficulties in understanding the main idea of a scientific paper that contains numerous abbreviations or introduces new terms for unjustifiably large number of concepts. Thus, the reader must usually break down and digest unusual concepts.

G. Zipf [10] analyzed the naming of new concepts and proposed the following law: name is a function of the frequency of use of a new concept in some limited social group; thus, professional language, dialect, slang, and other language subsets are generated. Naming is the desire to increase efficiency (velocity) of interaction in human society.

The appearance of new names inside a certain group is not sufficient for their introduction into a larger group; otherwise, the tendency to optimize transmission of messages can lead to the inverse result. For example, if each number has its own hieroglyph (digit) then number memorization and use would be difficult despite the fact that each number must be written, for example, shorter than in decimal notation.

Therefore, the problem of label should be solved taking into consideration the following two conditions²:

- Sufficient stability of the concept to be labelled, and
- Sufficiently frequent use of the concept so that efforts for memorizing a new name are lesser than its perception through description.

2.2.3 PROBLEM OF NAME STRUCTURATION

This problem involves the assignment of names in such a manner that their associations to objects are simplified to the fullest possible extent. Let us consider a number. By thinking, we studied to perceive quantitative information. First, we perceived one stone and two stones as objects not linked to a unified scheme; thus, counting skills were developed step-by-step. At some time, it became possible to assign different and unsystemized names to various quantities, e.g., a dozen. More complex calculations that employed a wide range of numbers rendered such an archaic method loose. A positional number system with its strong structure of number description and generation of a description of any big or small number does not yet require such types of names.

Thus, we use names like a million with the same ease as when we write 10⁶ or 10 to the sixth.

2.3 ORDERING

In Mathematics, the problem of structuration of object names in ways useful for applications and the problem of object ordering are solved by means of unified methods. Probably this is one of the causes of the high descriptive power of Mathematics. Although mathematical solution is quite natural, there is no similar correspondence between name and ordering in other sciences. To illustrate this thesis, let us consider computer-based numerical analysis.

A peculiarity of numerical analysis is the numerical model: one-to-one mapping between input and output data. Nevertheless, great difficulties can appear for ill-posed problems. However, these difficulties are exceptional. In most cases, strong mathematical description of the original problem provides adequate solutions. As a rule, input and output data as well as the model's parameters are objects of the same nature and can be strongly defined. This explains the elegance of the algorithmic model in numerical analysis. Moreover, admitted operations on objects represented by numbers are known a priori. Only interpretation of the result (assuming that all definitions are correct) can result in uncertainty of the numerical model.

Information analysis does not possess a mathematicslike axiomatic base. In particular, the set of operations on objects is not defined by ordering of natural numbers as in numerical analysis.

The concept of ordering in Information Science is often used in the narrow sense as a synonym of certain temporal unfixed object relationships. For example, people in a queue can be ordered in different manners: by height; professional or sporting interests, etc. In contrast with this pragmatic approach, we understand ordering as something implied from basic object content. For example, the meaning of the word six implies that the object with the name three precedes such an object (in linear systems); the meaning of the word father implies its following location of the object, son (in hierarchical systems). Of course, other examples can show situations in which these orderings do not hold. In this case, we can use the words as temporal labels and not their usual meanings. Note that not only numbers admit such (partial) ordering. The main difficulty here is that non-numeric information expressed by words does not admit one-toone ordering due to the multiple meanings of each word. A possible solution is multi-dimensional ordering. The Mendeleev Periodical Table is an outstanding example of this type of ordering. The Table shows how to take into consideration different properties of chemical elements and how to order each element.

²In the history of natural language, the problem has been solved this way.

The concept of ordered number could be established independently of its quantitative substance as the result of abstraction from qualitative differences of the equally arranged sets. Understanding independent number substances became clear only in the 19th century (*Peano*).

To express solely the number's ordering substance, we could use letters without their quantitative meaning only by using alphabetical order. In this sense, the concept of natural numbers \aleph has been developed. In \aleph , each element is defined by its location in the series.

The French word *l'ordinateur* literally means *fixing* order, which is closer in meaning to modern computer use than the word in English, computer, which literally means calculator and mainly characterizes the prescription of computers in initial stages of the computer era.

V. Bush (Manhattan Project) was the first who noted this fact and who proposed a new approach to the order organization of concept indexing for search in the computer. In 1945, Bush published the semi-utopic work, "As We May Think", which remained during many years the most cited publication in the field of man-machine interaction. Amazingly, the work contains a description of browser —a system for text-graphics information search. This system, called Memex, included a large library of texts, photos, and drawings. Although Bush showed great forecasting talent, Memex was not a computer because it used microfilms and photo-elements. The main peculiarity of Memex system was the possibility to input relations among library elements. The corresponding mechanism was inevitably bulky but logical. If the user had two documents (each in single image) that he wanted to relate, he struck via a special relation a name button and this name appeared at the bottom of each image. To obtain the document related by a certain relationship to another, he simply entered its code via the same button. This we call hypertext at present.

Some time after D. Engelblatt (developer of the first computer mouse) developed a system that linked hypertext with newly invented devices and graphics (multimedia).

In contrast, the search engines of the modern Internet use the principle of search by keywords. This principle is not efficient because an engine displays a great number of vague references from different sources related to each keyword. Among methods of identification in the Internet, the most interesting principle of information search belongs to the server www.altavista.com, which shows the complete route to the site (page) where a keyword is found. However, this approach is not a semantic (or meaningful) search. ISBN (book systematization standard) uses the subject-alphabetical principle of ordering but does not provide detection and construction of semantic-meaning concepts as inter-disciplinary and inter-topic relationships (associative pointers). Construct associative pointers means the establishment of

relationships between a given and well defined, tree-like structure of unique pointers (*subject domain*) and a set of semantically meaningful concepts (*course domain*).

Problems of semantics and meaningful search for words in IF are considered in the subsequent section.

3. SEMANTIC-MIND ANALYSIS AND OBJECT ORIENTED DATA INTEGRATION OF INFORMATION

"Semantics—the study of the meanings of words" [2]. Therefore, semantics is the search for cognitive, associative object identification in IF. Keywords in text, sound-shapes in audio flows, segments in image flows (data invariants), etc., are the object-oriented data of IF.

"Mind—the part of human being that governs thought, perception, feeling, will, memory, and imagination" [2]. In other words, meaning is the capability to understand, to feel, to perceive, and to imagine, i.e., restoring an entire piece of knowledge by some segment or part. For example, the program Guess a Melody [3] identifies the musical piece by a few musical phrases; the-raster to-vector conversion system A2R2V [4–6] is oriented toward searching for names of the converted set of pixels, thus, object-oriented data integration to Geographic Information System (GIS).

On the other hand (See first paragraph of this section), semantics is the adequate, meaningful search for words in IF (e.g., Google search www.google.com as zero approximation), while the meaning is the extraction of the subject domain in IF (e.g., adaptation of a Physics textbook or articles in a specialized journal for the secondary school).

Subsequently, Semantic-mind analysis (SMA) of IF is the meaningful search for object names and definition of the subject domain as a set of found names (e.g., 32 letters require five bits of information by Shannon, while the same letters require fewer bits of information by Morse. The difference is that the Morse Code took into consideration frequency of use of letters in text and named letters by symbols, such as point, dash). Objectoriented data integration (OODI) is input into a particular computer-based application of the output of SMA. This input suggests special SMA-output data organization, compression, storage, processing, etc., which is dependent on that application. For example, vector object integration to GIS is straightforward. Moreover, it is more desirable to store that object in GIS under the corresponding name only. In the following sections, we will use the abbreviation SMA/OODI for newly introduced concepts.

Cartographic data (CD) (raster or vector) as IF are one of the most complex subjects for SMA/OODI, because they are simultaneously contain different types of information carriers, including graphics, images, texts, symbols, etc. [6] [20].

3.1 SMA/OODI: METADA AND MPEG-7

For the sake of simplicity with regard to the following explanations, let us now consider the concept of metada. Metada are aids (via HTTP, FTP, etc.) for network users to follow up information resources and optimize their primary and secondary uses (see, e.g., IEEE European Colloquiums' Multimedia Database and MPEG-7). SMA/OODI can be also seen in the context of MPEG-7 technology as a system potentially adapted to process IF as metada (user-oriented processing). It is motivated by the fact that at present, digital audio-visual information (AVI) can be accessed by anyone not only for consumption but also for yield. This converts us at least potentially into content makers. We can publish and transmit digital information yielded by us via the Internet. However, day-by-day more and more simple procedures of audio-visual content acquisition, processing, and transmitting constitute only one part of the problem. The other part is that access to existing data should be equally simple because of the huge amount of AVI yielded daily throughout the world. Unfortunately, identification of desired (useful) information by searching and filtering has become more and more complicated. Even if open-source resources are used for access in such specific area as GIS (open-GIS and similar), the problem remains due to data interoperability and homogenization. Thus, GIS-oriented people now refer to the Spatial Semantic Web [7].

Therefore, the problem of fast and efficient identification of audio-visual contents in IF is emerging. This problem has motivated the Multimedia Content Description Interface Project by MPEG, also known as MPEG-7. MPEG-7 attempts to define standards for description of different AVIs that include images, image sequences, speech, audio, graphics, 3D models, and synthesized audio independently of representation formats [3].

Our general considerations in this work are aimed to follow up this research line and hopefully lead to better understanding of this research.

3.2 THE LAW OF PROGRESSIVE SIMPLIFICATION

We visually perceive the outer world as an optical process of transmission of images to the retina and subsequently construct the scene model as the spatial-temporal structure of objects and their relationships by local analysis and synthesis. This allows for extracting image semantic relationships and advancing to the verbal level of representation of initial visual information by logical analysis [8].

Subsequent passage from visual perception to verbal (logical) cognition provides progress in many areas (Computer Science included), well illustrated in the following [9]: "... In the history of handwriting, we observe not only the correspondence between the

techniques of writing development and the form simplification but also these two tendencies are identical 'de facto' because the technical problem of writing as registrator and translator of human speech should have been also solved. This is a clear representation of the widest sphere of human language by means of maximal optimization of visual symbols, i.e., etherification is the law of progressive simplification..."

Human reasoning, memory, and cognition as self-substance do not exist. On the contrary, they are simply the names (labels), synonyms that associatively reflect the result of the human's brain functioning as continuous cognition, i.e., the tool (processor) of etherification. The human brain's self-sufficiency is based on the law of progressive simplification. A particular case of etherification is the principle of least effort by Zipf [10]³.

Basic sources of information exchange and their types, and evolution of perception and thinking in the human world are shown in Table 1. If we look at this table from the point of view of computer-based systems (Turing's world), some analogs are straightforward. Let us only highlight some of them, leaving to the reader to complete others as an exercise.

Suppose that you are a computer specialist and have never heard about cognitive science or have never learned any other science. Are you familiar with concepts such as genetic algorithm, environment, sampling, learning, selflearning, knowledge, etc., from your area? Of course you are. Then look again at the table.

Computer genetic algorithms and programming involve methodologic background similar to the natural processes and laws that govern the transmission of genetic information. Thus, the concept of inherent knowledge can be spread and accepted in both worlds (Stage I).

Environment and environmental models are now commonly accepted concepts of Turing's world. They are rather the computer media (Unix, Linux, Windows, WordNet, open GIS, Spatial Semantic Web, etc.) in which computer programs are run than physical media. At present, computer experiments are common and available to nearly everyone. Moreover, newly invented computer environments encourage experiments that we call by analogy research instinct (Stage II).

Sampling is widely used in image processing and pattern recognition to imitate the characteristics and behavior of something already known from "parents", "teachers" (sample set, seeds, etc.) to acquire new or identify old information from unordered data sets (Stage III).

Learning and self-learning are represented as powerful tools of new knowledge acquisition in both worlds.

³In other words, we adopt here a strongly nominalist point of view. It is rather our hypothesis and model than a fact that everyone should accept. The only aim of this "sound" affirmation is to guide the reader toward the computer-oriented model as quickly as possible (§3.3).

Moreover, learning and self-learning computer systems usually oriented to tune system parameters are most efficient in automatic modelling (Stage IV).

Knowledge-based computer systems applied to digital

development of self-consistence virtual reality similar to human's imagination (Stage V).

Stage	Source	Information process	Model	Transmission tool
I	Biological kind (Nature)	Transmission of genetic information	Inherent knowledge	Inherence, instincts of life and death
II	Environment	Self-learning	Environmental model	Research instinct
III	Parents, flock	Sampling	Behavioral model	Imitation
IV	Teachers, society	Transmission of exciting knowledge, learning	Model of life	Language
V	Cognition	Synthesis of new knowledge	Model of the world	Imagination

Table 1. Evolution of perception and thinking (levels of information exchange) ordered by rows. We aim for this to show some analogs between human and machine worlds in this evolution.

3.3 TOWARD A PROBLEM-ORIENTED SEMANTIC ENVIRONMENT

In taking up again previously mentioned promptings within the context of SMA/OODI, we should note the following.

Images or events of any nature involve semantics. Semantics is universal and context-interpretative in a certain finite space of meanings and events. This statement is based on the understanding (probably multivalued) of any image or event of the outer world even if they do not yet possess a semantic context.

At present, organization of system analysis of data flows that aim to detect the interpretative in different contexts structured elements and a fast search for commonly accepted data structures experience great difficulties. This means that further significant development of data flow processing is impossible without a knowledge-based semantic-oriented analysis of data structures. The latest developments in the Internet graphically illustrate our statement. Of course, the knowledge bases themselves are not functional without special methods of preliminary analysis.

Here, correspondence of the problem-oriented environment to the level of the problem under consideration and the capacity to process meaning arises once again. For example, signal digital representation is solely based on the abstract concept of band capacity and does not take into account semantics: what is this signal—text, music, or image?

To illustrate certain principles with regard to meaning processing, we temporarily disregard the concept of generalized context.

If we denote something, then we assume that the meaning is understood, explicit. We do not go from sounds to images and from images to meaning: from the beginning, we are embedded in meaning, and being able

to express it in sentences. Meaning prescribes possible denotations and conditions. Moreover, meaning is the object of the following sentence. If a certain name is assigned to a sentence, then clearly each name that itself denotes an object can be the object of the new name denoting its meaning: N1 addresses N2, which denotes the meaning of N1; N2 addresses N3, etc. The language for each of its names must contain some name for the meaning of this name. Such infinite multiplication of verbal essences is called a Freguier paradox. In the meaning-object relationship, Freguier paradox locates the place of search for the meaning and indicates the moment of appearance of so-called context-meaning dependence. We use Freguier paradox as a methodology of the dynamic structures of data relationships in the problemoriented environment. These relationships in difference with arbitrary environment are defined by means of associatively organized identification and allow eventual, i.e., meaningful structuration [20]. This provides a new look at problems of efficient organization of the computer environment for data processing and understanding.

The first problem—compact representation of information concerning complicated systems— forces the search for new structured forms of knowledge representation. From our point of view, self-similar recursive structures provide adequate description tools. Moreover, self-similar recursive structures not only handle information but also manage information in a similar manner as genomic programs. One more important fact is their simplicity and regularity, although the result of application does not provide such as impression [11].

Secondly, the problem of memorizing the data structuration must be solved in a flexible, manageable manner so that signs that represent information segments are ordered by values [1] [13].

The problem of development of computer knowledgebased systems is illustrated very well by Hofstadter [12]: "A computer does not have automatic sensitivity to the images it processes. Of course, we could not expect this. It only executes the program like an old saw. The computer does not tire by adding columns of numbers even if all numbers are equal. Men get tired. What is the difference? Obviously, the machine misses something that allows it to have unlimited patience for repeated operations. The missing detail can be described in a few words: this is the capacity for self-observation, contact with the outer world; this is the capacity to perceive an image of proper activity and carry it out at any level of abstraction. Metaknowledge and knowledge are completely mixed between themselves in a unique flow and are mutually enriched. This renders self-observation as an automatic implication of the memory's structuration. How is this amazing flow organized in the human brain?"

We cannot answer Hofstadter's question even approximately. Undoubtedly there is the need of hierarchical knowledge ordering, but the hierarchy should possess a special form that Hofstadter calls "... complete mixture in unique flow" [13].

Natural organization of memory demonstrates once again the efficiency of a system approach with flexible inclusion of different tools for the problem of structuration of information. For example, Arbib [14] wrote: "Many people have discussed the problem of whether the human brain is a sequential or a parallel computer? This is false opposition. Considering the eye's motion, we observe some sequence of operations but understand that strong parallel computing is required within each time segment".

If we can successfully designate to machines the capacity to make decisions on an experimental basis, fully using any given insufficient data, furthering precision and extending these decisions with newly arrived additional data, then we would have a computer that does not require an explanation of how-to-do, but only what-to-do. Of course, the language of human-machine interaction should possess less logical and arithmetical depth than the internal language of the computer in this case.

4. CONCLUSION

of tool images Psychographic visual communication appeared much earlier than language i.e., the abstract verbal form of the semantic-meaningful representation of the outer world. We attempted to highlight this fact in §3.2 by citing [9]. Development of traditional mathematical models has only led to numerical data processing. By inertia, similar approaches have been relegated to the computer data processing (e.g., a huge number of useful-less, pixel-oriented image processing approaches: the amount of data required to apply such approaches is often equivalent to the amount of the original data). There is no need to prove that recognition of the known painting requires less information than an attempt to recognize the unknown. Modern computer technologies (protocols, formats, etc) empirically demonstrate the necessity to identify different types of images, including cartographics, paintings, photos, chart-flows, etc. Morphologic classification could be useful here, but experiences great difficulties in computer analysis due to its weak formalization.

In the present work, we recall and show one possible approach to the problem: detection of semantic-meaningful components from information flows. These components can be different with regard to dependence on the problem under consideration.

In Table 1, we attempted to exhibit evolution of human processing of information flows and to put forward some analogs with machine processing. From our point of view, these analogs are quite correct because efficient human-machine interaction is the stumbling block of modern computer technologies.

The problem herein discussed is of great current interest (remember the now-popular image-processing slogan *Back to the intelligence!* www.ijcai-03.org). We conjecture that the most promising line of progress toward the solution of this problem lies in successively increasing automation of the separate links of the approach considered herein. We suggest as *the main principle* of such automation, the maximal use of data-semantic content. Indeed, semantic information can be optimally organized and effectively processed by a computer system.

The companion paper in this book [20] illustrates our approach by describing the *Analysis* system and the color cartographic map interpretation system (A2R2V) that encapsulate basic elements of semantic analysis and synthesis of visual data. Both systems constitute symbolic language descriptions of objects of information flows rather than the traditional programs of data treatment.

ACKNOWLEDGEMENTS

The authors of this paper wish to thank Prof. Jean Serra for very useful comments and especially for reminding us of Saint Augustine's words⁴.

REFERENCES

[1] V. Alexandrov, M. Kharinov, S. Levachkine, Conception of hierarchical dynamic structure in application to audio and video data recognition, In: Hamza, M. (ed.), Proceedings of IASTED International Conference on Artificial Intelligence and Soft Computing (ASC 2001), 21-24 May 2001, Cancun, Mexico, 348-353. ACTA Press,

⁴ Nec ipsa tamen intrant (in memoriam), sed rerum sensarum imagines, illic praesto sunt cogitationi reminiscenti eas. Quae, quomodo fabricatae sint? Quis dicit? conf. X-8

Anaheim Calgary Zurich (2001) (ISBN 0-88986-283-4; ISSN 1482-7913).

- [2] Webster II (Ed.: Berkley Books, N.Y., 1984).
- [3] H. Kosch, http://www.acm.org/sigmod/record/issues/0206/5.kosch2-new.pdf
- [4] S. Levachkine, A. Velizquez, V. Alexandrov, M. Kharinov, Semantic analysis and recognition of raster-scanned color cartographic images. Lecture Notes in Computer Science, Vol. 2390, 2002, 178-189.
- [5] S. Levachkine, System approach to R2V conversion for analytical GIS, In: Levachkine, S., Ruas, A., Bodansky, E. (eds.), e-Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2002), 3-4 December 2002, Mexico City, Mexico (2002) (ISBN: 970-18-8521-X).
- [6] S. Levachkine, Raster to vector conversion of color cartographic maps for analytical GIS, In: Llados, J. (ed.), Proceedings of Fifth IAPR International Workshop on Graphics Recognition (GREC 2003), 30-31 July 2003, Barcelona, Catalonia, Spain, 77-91 (2003) (ISBN: 84-932156-6-X).
- [7] M. Egenhofer, Spatial semantics on the Spatial Semantic Web. In: Levachkine, S., Serra, J., Egenhofer, M. (eds.), Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2003), 4-5 November 2003, Mexico City, Mexico (2003).
- [8] V. Alexandrov, The eye and visual perception, Journal of Optical Technology, 66(9), 1999, 809-816.
- [9] J. Toynbee, Comprehending history in Russian edition: Дж. Тойнби, Постижение истории (Ed.: Прогресс, Москва, 1996).
- [10] G.K. Zipf, Human behavior and the principle of least effort (Ed.: Addison Wesley Publishing Co., Inc., Cambridge, MA, USA, 1949).

- [11] V. Alexandrov, N. Gorsky, Image representation and processing: a recursive approach. (Ed.: Kluwer Academic Publishers, Dodrecht Boston, London, 1993).
- [12] D. R. Hofstadter, Gödel, Escher, Bach: an eternal golden braid (Ed.: Basic Books, N.Y., 1979).
- [13] S. Levachkine, A. Guzman, Hierarchies as a new data types for qualitative variables, Journal of Data Knowledge Engineering (DKE), Elsvier (2002).
- [14] M. Arbib, Brains, Machines and mathematics (Ed.: McGraw-Hill, N.Y., 1964).
- [15] V. Alexandrov, N. Gorsky, From humans to computers: cognition through visual perception. (Ed.: World Scientific Publishing Co. Pte., Ltd., Singapore, 1991).
- [16] V. Alexandrov, M. Kharinov, A. Velázquez, S. Levachkine, Object-oriented color image segmentation, In Hamza, M. (ed.), Proceedings of IASTED International Conference on Signal Processing, Pattern Recognition, and Applications (SPPRA 2002), 25-28 June 2002, Crete, Greece, 493-498, ACTA Press, Anaheim, Calgary, Zurich (2002) (ISBN: 0-88986-338-5; ISSN: 1482-7921).
- [17] V. Alexandrov, Developing systems in science, technique, society and culture. (Ed.: Nauka, Saint Petersburg, 2002).
- [18] J.-C. Simon, Patterns and operators. The foundations of data representation. (Ed.: McGraw-Hill, 1984).
- [19] V. Alexandrov and A. Arsentieva, Information and developing structures. (Ed.: Leningrad Institute for Informatics and Automation of the Academy of Sciences of the USSR, 1984).
- [20] S. Levachkine and V. Alexandrov, Semantic-mind analysis and object-oriented data integration of visual information: a primer. In: Levachkine, S., Serra, J., Egenhofer, M. (eds.), Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2003), 4-5 November 2003, Mexico City, Mexico (2003).

SEMANTIC-MIND ANALYSIS AND OBJECTED-ORIENTED DATA INTEGRATION OF VISUAL INFORMATION: A PRIMER

Serguei LEVACHKINE+

*Department of Artificial Intelligence Centre for Computing Research-IPN UPALMZ, Ed. CIC, Mexico City, 07738 MEXICO E-mail: palych@cic.ipn.mx

ABSTRACT

A companion paper in this book emphasizes the emergency of computer-based intelligent analysis of digital information flows, provides some prompting toward its solution, and sketches an approach to intelligent data (audio, images, video, mixed) analysis and synthesis. In the present paper, we describe how to extract semantic components from unordered data sets (Gestalt problem) in visual information data (Analysis). An application of our approach is illustrated by describing a raster-scanned color cartographic map interpretation system—Analogical-to-Raster-to-Vector (A2R2V).

KEY WORDS

Semantics, Information flow, Visual data, Semantic-mind analysis, Object-oriented data integration, Semantic analysis of cartographic data.

1. INTRODUCTION

For the sake of completeness, we reproduce here some definitions that have been introduced and discussed in the companion paper in this book [20].

Information flow (IF) is binary digital data stored and processed by the computer. This is the basic element of digital technologies. On the other hand, IF is data that derives come from the outer (to the computer) world and that have very different representations— environmental monitoring, text, music, speech, images, etc. At present, the problem in adapting computer systems to human perception and cognition is apparent (MPEG-7), but not yet carried out. In the present work, we would like to describe some applications of Semantic-mind analysis (SMA) and Object-oriented data integration (OODI) of IF by following the methodology proposed in [20]. These applications aim to show within the general context of the human-machine interaction the semantic approach to computer data analysis and synthesis.

"Semantics—the study of the meanings of words" [2]. Therefore, semantics is the search for cognitive, associative object identification in IF. Keywords in text,

Victor ALEXANDROV*

*Saint Petersburg Institute for Informatics and Automation-Russian Academy of Sciences 39 14th Line, Saint Petersburg, 199178 RUSSIA

E-mail: alexandr@mail.iias.spb.su

sound-shapes in audio flows, segments in image flows (data invariants), etc. are object-oriented data of IF.

"Mind—the part of the human being that governs thought, perception, feeling, will, memory, and imagination" [2]. In other words, meaning is the ability to understand, to feel, to perceive and to imagine, i.e., restoring a whole knowledge by some segment or part. For example, the program Guess a Melody [3] identifies the musical piece by a few musical phrases; the raster-to-vector conversion system A2R2V [4–6, 9] is oriented toward searching for names of converted set of pixels, thus, object-oriented data integration to Geographic Information System (GIS).

On the other hand (See above), semantics is the adequate, meaningful search for words in IF (e.g., Google search www.google.com as zero approximation), while meaning is the extraction of subject domain in IF (e.g., the adaptation of a Physics textbook or the articles of a specialized journal for secondary school).

Subsequently, Semantic-mind analysis (SMA) of IF is the meaningful search for object names and definition of the subject domain as a set of found names (e.g., 32 letters require five bits of information by Shannon, while the same letters require fewer bits of information by Morse. The difference is that the Morse Code takes into consideration frequency of use of letters in text and named the letters by symbols-dot, dash). Objectoriented data integration (OODI) is input into a particular computer-based application of the output of SMA. This input suggests special SMA-output data organization, compression, storage, processing, etc. dependent on that application. For example, vector object integration to GIS is straightforward. Moreover, it is far superior to store that object in GIS under the corresponding name only. In the following, we will use the abbreviation SMA/OODI for newly introduced concepts.

Cartographic data (CD) (raster or vector) as IF are one of the most complex subjects for SMA/OODI because they simultaneously contain different types of information carriers: graphics; images; texts; symbols, etc. [6].

Our method is object-oriented, while the majority of state-of-the art approaches are pixel-oriented [11-15].

Pixel-oriented methods are artificial and destroy most essential image characteristics, including general composition, object shapes, image dynamics, and contextual information. All this results in increasingly complex methods and algorithms for machine recognition of data objects based on pixel-oriented approaches that do not materially improve the results. The majority of effective image formats preserves the natural image structure and thus provides object-oriented data integration. This thesis is especially true for cartographic images that being intermediate between man-made drawings and natural images provide a nearly ideal model for SMA, and OODI is quite natural here within the context of GIS. More arguments are to be found in the [1, 4, 5, 6, 9, 20] (See [7] mathematical morphology approach and fuzzy approach [18] with similar ideas to GIS-ready data processing).

In [1], we considered one of the applications of our approach to audio information flows. In particular, we detected the semantic components from 'Für Elise' by Beethoven. We compared our object-fitting hierarchical compact decomposition with wavelets and fractal decomposition and in addition gained better-structured and nearly twice the amount of compressed signal representation. Moreover, in decomposing the signal by our method signal objects were partitioned into a natural hierarchy with no additional procedure.

The goal of the present work is to apply SMA (analysis) and OODI (synthesis) in the case of visual information flows (*Analysis*) and in particular for cartographic raster data (*A2R2V*).

2. ANALYSIS

In this section, we shall explain how visual data can be transformed into their semantic-mind representation by describing the system termed *Analysis*—first SMA/OODI application.

The majority of image processing methods such as wavelets and fractal decomposition or texture analysis usually miss the vital components—the semantic structure.

In our approach to visual information structuring, called the Analysis system, we apply adaptive dynamic data structures for object-fitting hierarchical analysis of video-data. This system is a tool to reveal interrelated networks of context independent semantics of the initial data structure.

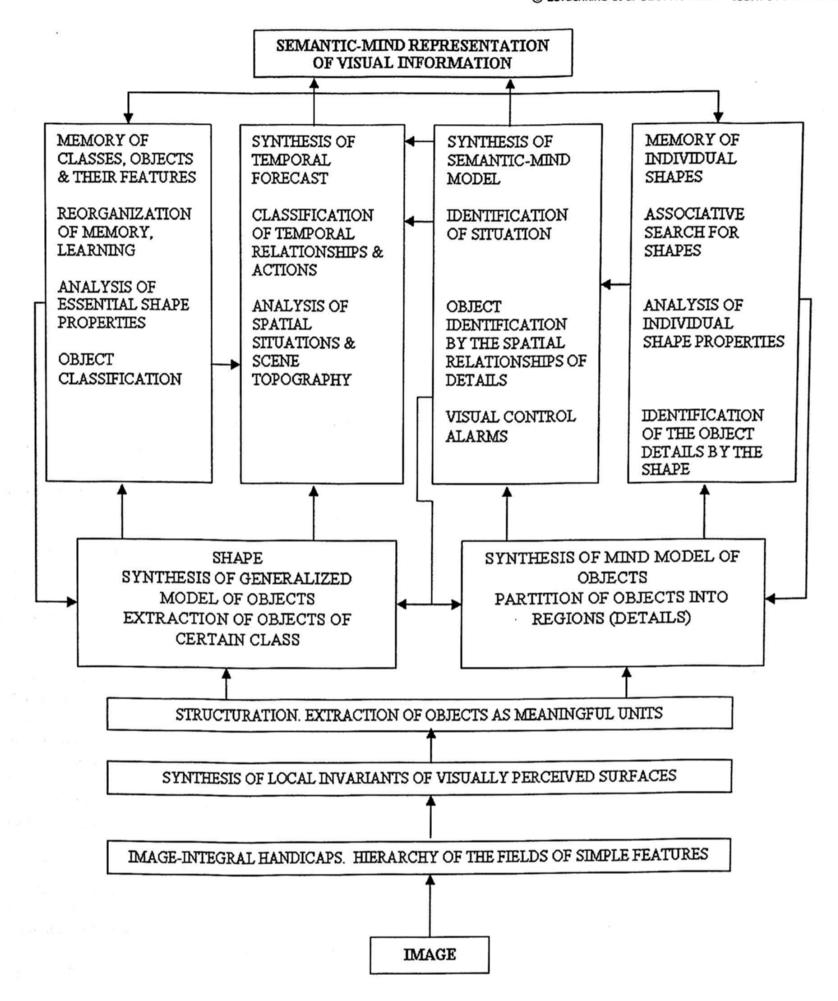
Our approach is based on location of semantically important image segments using a rating principle and subsequent iterative synthesis of hierarchical dynamic trees to build the coherent structure of selected segments. Our considerations have led us to the idea that distribution of segments (F_I) and levels (I) of a tree correlates with the growth law $F_I \approx I^{0.618}$ similar to the empirical Zipf's laws [10].

Adaptive dynamic data structure is the result of the structuring process and contains image segments, which are essentially important for subsequent image processing and understanding.

The most evident application of image semantic decomposition is preliminary structuration of video-data for subsequent object identification and target-oriented compression of images.

To transform visual information into its semanticmind representation, we used the previously method of associative dynamically adaptive data tree-structures. Flow chart 1 shows the computer-oriented scheme of transformation of visual data into the semantic-mind model [1, 20].

and the same of th



Flow chart 1: Transformation of visual information into semantic-mind representation.

2.1 MAIN COMPONENTS OF ANALYSIS

The majority of segment location methods require initial partition of the image into a set of segments. The main problem with this non-trivial task of image analysis is lack of general methods for reasonable segment selection. We used the semantic-mind approach to overcome this difficulty [1, 20].

Analysis is oriented toward revealing the independen semantic contents in source data. The task is to build context-independent hierarchical networks of flat anospatial structures that set off this system from others.

A next step is generalization of Analysis into image synthesis. The combination of image analysis and synthesis would allow application of new methods for image processing and development of semantic-oriented compression systems.

Application of Analysis focuses on the problem of effective and optimal data storage. We solve this problem by selection and storing only semantically important data.

We have been developing the iterative method for object approximation by means of multilevel image segmentation [1, 16]. The hierarchy of source image segments tends to form the original adaptive dynamic data structure, which can be used to preserve semantic relationships of independent image parts. Elements of this data structure should be meaningful and recognizable for the human operator or automatic programs processing the image.

To solve the problem, the following algorithm has been designed:

- Source image is presented as RGB color-divided matrix. Pixels of the same color are segments of the first-level image partition (seeds).
- At subsequent iterations, the program merges segments whose color components contain little divergence, i.e., the differences among them fit into the selected range. In alternative cases, independent structure can be built for each color or gray-level component of the converted image. Special recursive procedure controls connections between segments and homogenizes processing throughout the entire picture.
- The process results in an interrelated semantic network of image segments that is the pyramidal adaptive structure for dynamic image representation.
- Source image I' is transformed into special computer representation I—the set of a x b pixels.
- For effective image processing, it is necessary to change the image structure once again: image I is represented as a hierarchy of segments. m^{th} level of this hierarchy $L_m = \{F_p \ p \in D_m\}$ contains segments defined by the set of segment indexes D_m , where $I = \bigcup_{p \in D_m} F_p$. Links between segments correspond to edges of the indexed tree with special organization in which each segment is a tree node.
- An element of mth level is denoted F_p^m. Segments of the first level are pixels of image I: L₁ = {F_p¹/1 ≤ p ≤ a * b}.
- Segment construction procedure validates the following equations for segment number in different levels:

$L_i > \# L_{i+1}; \# L_{255} = I;$ $\forall p \in D_m \exists q \in D_{m+1}: F_p^m \subset F_q^{m+1}.$

Object location in terms of this structure is equivalent to location of segment networks with suitable node attributes (color, gray-level, spatial, or more complex segment characteristics). The essential problem arises when no coherent segment network corresponding to the required segment is found. This predicament is the main obstacle to full automation of image analysis in this approach.

To overcome this difficulty, Analysis allows the user compulsory manual restructuration of semantic structure. This restructuration can be considered a manner to apply user-oriented pragmatics to image analysis.

The main objective of our system is assistance in the search for practical solutions in general fields of image analysis and problem-oriented restructuration of videodata. This is reached by means of location of image-semantic information and subsequent processing of meaningful image segments.

Basic components of Analysis are the following [1]:

- Algorithm of detection of equibrightness (gray-level) and equicolor segments. Homogenization throughout the entire image of segment features, using dynamic global thresholding (thresholds that are required for detection of certain segments); thus, it is adaptive to different contexts.
- Mechanism of secondary indexing of detected segments, which are organized as dynamic, irregular trees of the hierarchical relationships of these segments.
- Procedure of restructuration of image segment relationships to satisfy to operator visual control requirements and respond to alarms.

Note that possible applications of Analysis are immediate development of the following methods:

- Semantic relationship analysis with contextual image processing as problem-oriented video compression.
- Tree-structured synthesis as replication of real-object structures.
- Morphologic image classification.

Based on analysis of different computer formats, we arrived at the conclusion that most effective formats preserve natural image structure and thus the conviction to use the SMA/OODI-based approach.

In contrast, pixel-oriented methods are artificial. They the destroy majority essential of image characteristics—general composition, 2D and 3D object shapes, image dynamics, and contextual information. image representation requires line-ordered description of pixel characteristics, while image compression, restoration, analysis, and object location utilities require different data organization and formats. Such a format group includes vector image representation, fractal compression, wavelet decomposition, and a number of special image formats that store only important information; the best known instance of the group is ontical character recognition (OCR) programs that transform source image into lines of text.

Due to the huge amount of information required in satellite and cartographic image processing, the thematic search for video databases, etc., effective automatic methods of image analysis to detect objects and set of image segments to carry out semantic upload emerged and have motivated this research.

2.2 IMAGE-SEMANTIC COMPONENTS

Segments were obtained as the result of the iterative procedure of successive increasing of admitted gray-level and color thresholds in segment-merging form, subsequently increasing pyramidal hierarchical structure of flat segment networks §2.1. Each segment of this structure can have the ancestor or descendant. Thus, structure obtained is called the adaptive-dynamic data structure.

The segment of image is a node of this spatial structure, whose attributes are primary numbers defined by averages of color/gray-level segment features and a set of pixels that represent the area and shape of the segment. This allows to organize the object-oriented identification of semantically meaningful image regions. Image semantics in this context corresponds to association of segments of different hierarchical levels identified with identifying conceptions from the subject domain (e.g., for cartographic data, detection of a segment identifying a coastline or highway becomes semantically meaningful. Further, this segment is renamed as coastline, highway, etc.).

As mentioned previously, Analysis involves the interactive procedure of *compulsory restructuration* of segment relationships as pragmatic tools of semantic analysis of visual data. In other words, the system's learning and self-learning with the prescribed set of associative identifiers are possible in the interactive or semi-automated regime.

2.3 CRITERIA OF OBJECT DETECTION

Of course, the problem of image semantic component detection is not trivial. However, some formal criteria for image characteristics that appear to be semantically meaningful for humans can be defined. One such criterion that we have used is the hyperbolic law of the form $1/F^a$, where a = 0.618, similar to Zipf's distribution [10].

Statistical distribution of image segment number on levels of segment hierarchy shows this hyperbolic regularity. Figure 1 displays the results of experiments with six test images and represents dependence of region number (n) on level number (N) in segment hierarchy

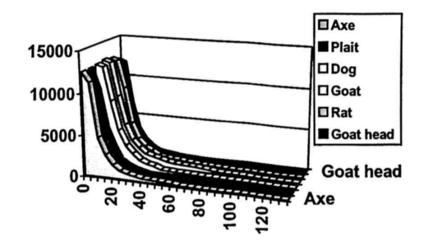


Figure 1. Rating distribution of amplitude and color regions.

The criterion of segment merging by close gray-level values was used (Figures 2a and 2b). This merging procedure is essentially different from traditional reduction of the color palette.

Similar hyperbolic regularities have been obtained not only with photos, raster images, etc. but also with the images of the realism and pseudo-realism paintings [20].



Figure 2a: Test image Axe at different semantic levels.



Figure 2b: Test image Plait at different semantic levels.

2.4 OBJECT LOCATION UTILITIES

Successive segment merging by such a criterion leads to segment structuring in a multilevel hierarchy. This hierarchy or multilevel image partition is an efficient method of semantic identification of the image's objects. Relationships between dynamic tree nodes indicate neighboring semantically meaningful regions. Because the image's regions are identified by corresponding tree nodes, the neighbor relationship among them can be completely defined by a table of adjacency. Modification (elimination of some edges, i.e., segment relationships) of the dynamic tree allows modifying the resulting region; thus, most exact object detection is reached. Each level of the tree of segments can be considered as alternative image interpretation in different semantics (Figures 2a and 2b; see also Figure 5, Section 4).

Adaptive dynamic tree structure regards the search for meaningful objects as fitting of the object features (color, location, geometry, topology, attributive information, e.g., the object's name from subject domain if known, etc. that on the whole defines the notion of image semantic object), to corresponding ranges and the following analysis of all admitted objects. This renders possible the use of automatic learning algorithms when the set of required objects is given and when it is necessary to define only corresponding feature ranges. The learning process can be organized as follows: User selects appropriate level of segment hierarchy and points out the set of suitable objects. These objects can be defined by merging object segments. Then, the program computes characteristics of located segments and relationships among them and establishes formal criteria for the search for similar objects. Figure 3 shows an example of associativedynamic data structure.

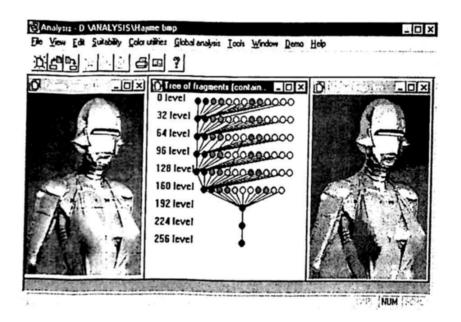


Figure 3: Example of associative-dynamic data structure.

2.5 EXAMPLE

To illustrate segment location, we use the following example: Source image is the segment of the earth surface (Figure 4a). The task is to segment land from sea.

First, we should construct the tree of segments.

Next, we select the appropriate level (64th) of tree structure, in which searched areas are approximated by segments of similar scale.

According to our task, small-scale segments that consist of one or two pixels are not essential. To visualize the different segments, we use contrast mapping of this level (Figure 4b).

From this mapping of semantically meaningful areas, we can point out any of the ground segments (it is shown in blue—the 1st node in the 64th level of the adaptive dynamic data structure).

To select all land surfaces, we need to edit the segment tree and join all land segments.

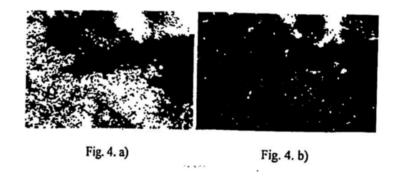


Figure 4. Example of segment location.

3. A2R2V

In this section, we will present another application of SMA/OODI by describing our raster-scanned color cartographic map interpretation system called A2R2V (Analogical-to-Raster-to-Vector) (see also [4-6, 9]).

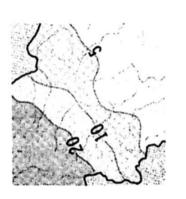
3.1 HOW DOES IT WORK?

A2R2V is based on SMA/OODI of color images. SMA/OODI of cartographic images is interpreted as a separate representation of cartographic patterns (alphanumeric, punctual, linear, and area). Our map interpretation system explores the idea of synthesis of invariant graphic images at low-level processing (vectorization and segmentation). This means that we ran vectorization-recognition and segmentation-interpretation systems simultaneously. Although these systems can generate some errors, they are much more useful for the following understanding algorithms and man-machine interaction because of its output in nearly recognized objects of interest.

We began map recognition from global binarization followed by classical OCR-identification with artificial neural networks supervised (ANN), clustering, knowledge-based recognition rules, and morphologybased vectorization. To overcome the problem of laborintensive training, we designed simplified images. For this purpose, we utilized linear combinations of color components or image representations (false color technique) and binary representation composing (composite image technique). These techniques are application-independent. However, within the framework of our approach map recognition may be treated as a common (application-dependent) task [17]. We followed the concept that important semantic information necessary to interpret an image is not represented in single pixels but in meaningful image segments and their mutual relationships [1, 4–6, 9, 20].

We set forth a conception of composite image representation and decomposition. The main goal of image decomposition consists of object linking by its associated names. We used image synthesis based on object-fitting compact hierarchical segmentation [16]. We performed composite representations (or simply composites) of the source image by means of a reduced number of color or tone components and segments. In this manner, visually perceived objects are not eliminated and the image's semantics is preserved (see Section 2).

Composite images form a book in which objects of interest can be found and recognized on appropriate page(s) [6]. Thus, a page number defines the method of thresholding and certain tuning parameters that can be learned interactively or automatically (Figure 5).



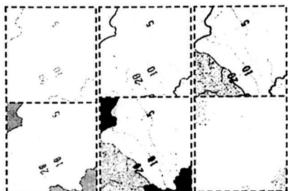


Figure 5: Source image (left) and some pages of a book of composite images (right).

3.2 A2R2V MODULES

Recognition systems of alphanumeric characters, punctual, linear, and area objects are subsequently described in [4-6]. They support segmentationrecognition-interpretation-vectorization A2R2V. Their names rather underline the desire to stratify the map by top-level vector thematic layers (punctual, linear, etc.) than by only recognition and vector points, arcs, etc. maintaining all of these together in one layer. In the following, we shall present the main ideas of the Fineto-Coarse Scale (F2CS) method. This method originated from the unsolved problem of vector description of raster objects and the fact that vector data in fine-scale maps are often available. In a certain sense, it represents a promising alternative. The method shows how to use feedback already acquired from semantic information to obtain new information.

Let us suppose that we have a vector cartographic image I₁ (or and already recognized raster image) in scale s₁ of given territory T and a raster cartographic image I₂ of T to be vectored (recognized) in scale s_2 and $s_1 > s_2$ (e.g., $s_1 = 1$: 100,000 and $s_2 = 1$: 50,000). Our goal is to use the information from I1 in vectorization of I2. Note that I_1 can be considered as generalization⁵ of I_2 : $I_1 = G$ (I_2) , i.e., if an object $O_2 \in I_2$, then there can exist $O_1 \in I_1$ such that $O_1 = G(O_2)$. We denote Ω —the set of all such objects O_2 from I_2 and Θ —the compliment of Ω in I_2 : I_2 = $\Omega \cup \Theta$. We also put $\omega = G(\Omega)$ and note that $\omega \subseteq I_1$. Obviously, to vector objects from Ω and Θ we need two different strategies. Objects from Ω can be vectored using features (position, color or colors, shape, etc.) of vector objects from ω . After objects of Ω have been vectored, we can vector objects from Θ by one of the A2R2V modules

⁵We do not discuss here what this generalization is

as new cartographic material. Results of vectorization obtained at each step of this process as well as processing by other A2R2V modules are feedback to reuse in the next step. To perform the F2CS method, we used conceptual clustering [8] based on the set of prescribed object features [4] and the concept of the associated to image function [9]. Note that the method also solves the

difficult problem of the search for object to be vectored in the entire raster map field. In this case, a vectored object can be found in the nearest neighborhood of its generalized analog and nowhere else. An example of application of the fine-to-coarse scale method is shown in Figure 6.

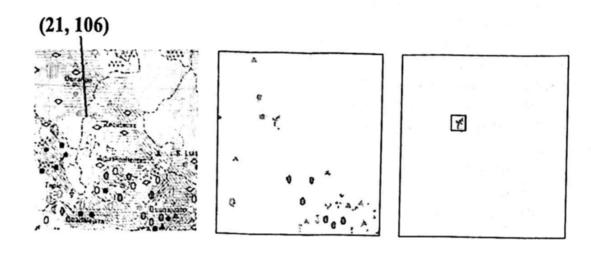
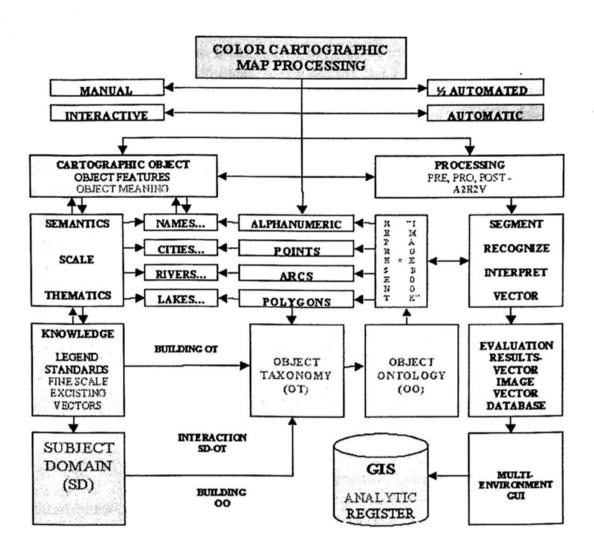


Figure 6: Recognition of punctual cartographic pattern. a) Original image in scale s_1 with input information: there is near geographic coordinates (21,106), a pattern denominated Palm. b) Image in scale s_2 ($s_2 < s_1$) obtained after application of punctual object recognition with conceptual clustering (criterion: β_0 -connected, β_0 = 0.9) [8], considering the associated to the image function restricted to one of the generated groups and adding pixels (region merging) [9]. Note that we searched for the pattern only near and for the already recognized *fine* pattern and nowhere else. c) Recognition of coarse pattern (function) Palm.

3.3 OBJECT-ORIENTED DATA INTEGRATION OF GIS-READY INFORMATION

We designated A2R2V as a multi-environment system that does not depend on particular data format and that is oriented to process color raster-scanned images. Its GUI is described as follows: At low level, it has modules of automatic recognition of alphanumeric, punctual, linear, and area objects in raster-scanned color cartographic images that have been programmed in C. Thus, a database with quantitative, qualitative, and nominal features is

associated with each recognized visual object. At the intermediate level, a Java module supports correct conversion of corresponding database to be processed at high level in one of the environments: Unix; Linux, or Windows. Therefore, the user of the system should not seek unified or particular representation of spatial databases in an environment, but only chooses the most convenient representation of the three provided. We are now developing a decision-making tool to support most adequately the particular GIS-application of the user's choice.



Flow chart 2: A2R2V functioning (components, modules, and interaction)

3.4 SUMMING UP

Now we shall explain the content of Flow chart 2 that summarizes our approach to color cartographic image raster-to-vector conversion. Our comments are as follows.

Any map interpretation system deals with processing of alphanumeric, punctual, linear, and area raster objects. The majority of state-of-the art systems based on image processing and graphic recognition [7, 11-15] consider color, shape and texture components of the concept of cartographic objects, usually omitting the meaning and topology. Of course, color, shape, etc., are important cues. However, they are only cues: there is no pure color in raster-scanned maps. What is the shape of the raster object; what is the arc centerline? In contrast, our definition of spatial semantic object (cartographic pattern) includes the following components:

Cartographic pattern = (color <space>; geometry <location, shape, etc.>; topology <point, arc, area, etc.>, and attribute <e.g., name of object; etc.>).

This definition allows full use of the peculiarities of very cartography for GIS-ready information

integration. Additional arguments may be found in [1, 4-6, 9, 16, 20].

Moreover, the majority of contemporary systems are pixel-oriented. They are artificial and destroy most essential image characteristics §1. Increasingly complex algorithms for machine recognition of maps that do not materially improve the results. We believe that the most effective image formats preserve natural image structure and thus provide OODI. This thesis is especially true for cartographic images that because they are intermediate between man-made drawings and natural images provide a nearly ideal model for semantic analysis; OODI is quite natural here within the context of GIS [6, 20]. This leads to natural taxonomic classification of cartographic patterns, the definition of cartographic subject domain (SD) (a set of names) dependent on map thematic, scale, legend, toponyms, and a priori knowledge concerning the map (existing/reworked vectors in fine scale). Different sources of evidence can be also put into SD to support efficient map processing. Interaction between SD and object taxonomy (OT) has led to conceptual structuration of cartographic data in hierarchical object ontology (OO) (nodes of OO are concepts-not words). Thus, we are looking for

- correct and adequate representation of raster objects as a thematic image book *prior to* processing/conversion. However, results of processing can be feedback and correct §3.2. In Flow chart 2, blocks in gray with text in red illustrate our system approach.
- 3) With regard to automatic interpretation of color cartographic images, this presents certain difficulties for state-of-the art image processing, pattern recognition, and artificial intelligence. The set of vector maps that one can expect as an output of such an interpretation is very useful for GIS, new map production, and old map updating. However, it appears unrealistic to obtain a fully automatic computer-based interpretation system free of errors [4-6, 9, 11-15]. Additionally, please note that high efficiency of interpretation is required for vector map production and updating first. It appears reasonable to obtain in both cases 90-95% successfully recognized objects. This is to avoid excessive work on corrections of errors produced by the computer system, which can sometimes be greater than manual raster-to-vector conversion [19]. Within framework of our approach, manual, interactive, semi-automated, and automatic processing may prove useful. Our point is: Apply them applicationdependently, use sources of evidence to respond alarms [4-6, 9].

In this section, we presented a system for automatic interpretation of raster-scanned color cartographic maps. The highlight of our system is two intelligent color image segmentation techniques. These allow searching for names of processed sets of pixels. Segmented and recognized objects are subsequently vectored to be finally included in GIS. These three stages (segmentation-recognition-vectorization) are Object-oriented data integration to GIS or GIS-ready information [18]. In our experiments, we used complex, full-size, raster-scanned color cartographic images with promising results [4–6, 9, 16].

4. CONCLUSION

We illustrated SMA/OODI by describing Analysis system and color cartographic map interpretation system (A2R2V) that encapsulate basic elements of SMA/OODI. Application of SMA/OODI for A2R2V allows increasing this system efficiency within the general context of raster-to-vector automatic conversion. Note that the latter problem is very complex and has attracted much attention during the last few decades. However, to date there is a lack of satisfactory solution of this problem. From our point of view, the proposed alternative is promising.

We attempted to demonstrate that a system, semanticbased approach to the raster-to-vector problem can be fruitful. Within the context of the A2R2V system this means, first, decomposition of source image by multiple semantic hierarchical networks. Second is segmentation with mutual recognition of appropriate primitives, while third is development of a unified, knowledge-based learning and self-learning system with optimal human-machine interaction. Our research is concerned with this approach.

ACKNOWLEDGMENTS

The authors of this paper wish to thank Prof. Jean Serra for very useful comments.

REFERENCES

- [1] V. Alexandrov, M. Kharinov, S. Levachkine, Conception of hierarchical dynamic structure in application to audio and video data recognition, In: Hamza, M. (ed.), Proceedings of IASTED International Conference on Artificial Intelligence and Soft Computing (ASC 2001), 21-24 May 2001, Cancun, Mexico, 348-353 ACTA Press, Anaheim, Calgary, Zurich (2001) (ISBN 0-88986-283-4; ISSN 1482-7913).
- [2] Webster II (Ed.: Berkley Books, N.Y., 1984).
- [3] H. Kosch, http://www.acm.org/sigmod/record/issues/0206/5.kosch2-new.pdf
- [4] S. Levachkine, A. Velázquez, V. Alexandrov, M. Kharinov, Semantic analysis and recognition of raster-scanned color cartographic images, Lecture Notes in Computer Science, Vol. 2390, 2002, 178-189.
- [5] S. Levachkine, System approach to R2V conversion for analytical GIS, In: Levachkine, S., Ruas, A., Bodansky, E. (eds.), e-Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2002), 3-4 December 2002, Mexico City, Mexico (2002) (ISBN: 970-18-8521-X).
- [6] S. Levachkine, Raster to vector conversion of color cartographic maps for analytical GIS, In: Llados, J. (ed.), Proceedings of Fifth IAPR International Workshop on Graphics Recognition (GREC 2003), 30-31 July 2003, Barcelona, Catalonia, Spain, 77-91 (2003) (ISBN: 84-932156-6-X).
- [7] J. Angulo and J. Serra, Mathematical morphology in color spaces applied to the analysis of cartographic images, In: Levachkine, S., Serra, J., Egenhofer, M. (eds.), Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2003), 4-5 November 2003, Mexico City, Mexico (2003).
- [8] J.F. Martínez-Trinidad and A. Guzmán-Arenas, The logical combinatorial approach to pattern recognition, an overview through selected works. Pattern Recognition, 34(1), 2001, 741-751
- [9] E. Gómez-González and S. Levachkine, Fine-to-coarse scale method for color cartographic images recognition. In: Levachkine, S., Serra, J., Egenhofer, M. (eds.), Proceedings International Workshop on Semantic Processing of Spatial Data (GEOPRO 2003), 4-5 November 2003, Mexico City, Mexico (2003)

- [10] G.K. Zipf, Human behavior and the principle of least effort. (Ed.: Addison Wesley Publishing Co., Inc., Cambridge, MA, USA 1949).
- [11] G.K. Meyers G.K and C.-H. Chen, Verification-based approach for automated text and feature extraction from raster-scanned maps, Lecture Notes in Computer Science, Vol. 1072, 1996, 190-203
- [12] M.-P. Deseilligny, R. Mariani and J. Labiche, Topographic maps automatic interpretation: some proposed strategies, Lecture Notes in Computer Science, Vol. 1389, 1998, 175-193
- [13] F. Dupont, M.-P. Deseilligny and M. Gonrad, Automatic interpretation of scanned maps: reconstruction of contour lines, Lecture Notes in Computer Science, Vol. 1389, 1998, 194-206
- [14] S. Frischknecht and E. Kanani, Automatic interpretation of scanned topographic maps: a raster-based approach, Lecture Notes in Computer Science, Vol. 1389, 1998, 207-220
- [15] G.M. Schavemaker and M.J.T. Reinders, Information fusion for conflict resolution in map interpretation, Lecture Notes in Computer Science, Vol. 1389, 1998, 231-242
- [16] V. Alexandrov, M. Kharinov, A. Velázquez, S. Levachkine, Object-oriented color image segmentation, In Hamza, M. (ed.), Proceedings of IASTED International Conference on Signal Processing, Pattern Recognition, and Applications (SPPRA 2002), 25-28 June 2002, Crete, Greece, 493-498, ACTA Press, Anaheim Calgary Zurich (2002) (ISBN: 0-88986-338-5; ISSN: 1482-7921).

and the control of th

- [17] D. Doermann, An introduction to vectorization and segmentation, Lecture Notes in Computer Science, Vol. 1389, 1998, 1-8.
- [18] U.C. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, M. Heynen, Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS ready information, In: Levachkine, S., Ruas, A., Bodansky, E. (eds.), e-Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2002), 3-4 December 2002, Mexico City, Mexico (2002) (ISBN: 970-18-8521-X).
- [19] Parcel mapping using GIS. A guide to digital parcel map development for Massachusetts Office of Geography Information and Analysis (MASSGIS) (August 1999) online: http://umass.edu/tei/ogia/parcelguide/
- [20] V. Alexandrov and S. Levachkine, Cognitive promptings for semantic-mind analysis and object-oriented data integration of information flows, In: Levachkine, S., Serra, J., Egenhofer, M. (eds.), Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2003), 4-5 November 2003, Mexico City, Mexico (2003).

SYSTEM APPROACH TO R2V CONVERSION FOR ANALYTICAL GIS

Serguei LEVACHKINE
Geoprocessing Laboratory
Centre for Computing Research-IPN
UPALM Zacatenco, CIC Building,
07738 Mexico City, MEXICO
palych@cic.ipn.mx

ABSTRACT

We present a system approach to automatic digitization of raster-scanned color cartographic maps. This approach contains all three steps involved in digitizing process: preprocessing, processing and post-processing. Up-to-date systems for vectorization raster maps using automatic or semi-automated programs essentially carry out only one kind of operation: they follow discrete points along a curve (tracing, snapping), or attempt to combine automatic and interactive tasks, if the tracing is met the ambiguities. In our proposal, the automation problem is approached from a unified point of view, leading to the development of software A2R2V (Analytical Raster-to-Vector) conversion that is capable to recognize and vector a maximum number of cartographic patterns into raster maps. We propose some strategies for solving the problem and illustrate it, describing single modules of A2R2V system. The place of the operator and knowledge in a R2V conversion system is considered as well.

KEY WORDS

Geographical Information System, raster-to-vector conversion, cartographic pattern recognition.

1 INTRODUCTION

The problem of automatic raster-to-vector (R2V) conversion is taken steadfast attention by researchers and software developers during last two decades. Numerous attempts to solve this problem have mainly originated from emerging area of automatic Geographical Information Systems (GIS). Unfortunately, completely automatic conversion system appears unrealistic and some authors suggested putting the operator into the center of a conversion system. Amazingly, this idea was independently expressed by the different authors with the same words from the Lukas' Gospel [1, 2]. Now the problem is which part of conversion process should be carried out by the operator and which part by the machine?

In the present work we analyze (keeping certain skepticism) the degree of possible automation of the single modules of a raster-to-vector conversion system. Mainly, we focus on the case of analytical GIS (the case of register GIS is considered in [2]). We believe that these two cases are representative enough and complement each other, constituting a complete description of both research and commercial (semi) automatic R2V conversion systems.

As we previously stated, the present work is clearly divided into two parts: 1) in Sections 1-4 we analyze the general concepts of automation of a R2V conversion system that are constituted the system approach; 2) in Section 5 we overview A2R2V conversion system in which we placed the ideas of our system approach. In Section 6 we emphasize the role of knowledge in automation of the conversion process.

The process of developing of a GIS can be represented by the following set of steps [1]:

- 1. State of the problem.
- 2. Definition of the methods for solution of stated problem.
- 3. Preparation of the input information required for solution of the stated problem by the defined methods.
- 4. Definition of the composition and logical structure of the cartographic data.
- 5. Analysis of existing cartographic information.
- 5.1 Recollection of the vector maps, which can be used in the GIS.
- 5.2 Definition of the vector map layers, which can be directly elaborated by the geodesic data.
- 5.3 Definition of the paper and raster maps for vectorization.
- 6. Designing the cartographic base for the GIS.
- 6.1 Inclusion of existent vector maps in the GIS.
- 6.2 Digitization of new cartographic materials.
- 6.3 Designing vector map layers by the geodesic data.
- 6.4 Unification of the vector map layers (conjunction of fragments, unification of the cartographic projection, elimination of logically contradicted data, etc.).
- Designing vector thematic maps.
- 8. Solution of the stated problem.
- Publication of the results.

Clearly the labor required for each of the steps listed above depends on the qualitative and quantitative characteristics of the cartographic materials used. These can be combined into a single index: *novelty*. Thus, we can classify a GIS into one of four novelty levels (in decreasing order of index):

A. A new vector cartographic material is required. However, for the territory under consideration only paper maps exist in the desired scale.

B. Basic vector cartographic materials exist for the territory under consideration, but there are not thematic maps of the desired type.

C. Thematic vector cartographic materials have been produced in a previous GIS for the territory under consideration. The new GIS are to be developed under the assumption of some fundamental changes.

D. Updating and development of an existing GIS without essential changes.

GIS with novelty level A and B require relatively more labor for the digitization of paper and raster maps than the others do [1]. Generally speaking, topographic and thematic vector-based maps have been produced for the territories in small-scale. Current GIS-development is directed towards problems in urban planning, emergency situation monitoring, election enumeration and other government activities that require the maps in big-scale. To carry out these projects, the volume of labor involved in the digitization of paper and raster maps is considerable, even if only the largest urban areas of a country are considered. Clearly the development of these problem-oriented GIS can only be completed if the task of vectorizing traditional cartographic materials is automated to the fullest extent possible.

Two main technologies⁶ are currently used for map vectorization [1]: (V1) Paper map digitization by electronic-mechanical digitizers, and (V2) Raster map (a map obtained after scanning of the paper original) digitization.

The process of digitizing paper maps cannot be automated; hence we suggest that the only practical approach to design GIS's is the development of new methods and software to automate vectorization of raster maps.

Raster map digitization technologies can be divided into four intersecting groups [1]: (D1) Manual; (D2) Interactive; (D3) Semi-automated, and (D4) Automatic.

In practice, manual raster map digitization methods (D1) coincide with paper map digitization methods (V1). A few examples will serve to illustrate this. In the case of punctual objects, the operator visually locates graphic

⁶ Or types of source data capturing from maps. Other data types commonly used in GIS are survey data, aero and satellite images, etc.

symbols and fixes their coordinates. In the case of linear and polygonal objects, the operator uses rectilinear segments to approximate curvilinear contours. The manual digitization rate is one to two objects per minute.

Interactive digitization uses special programs, which, once the operator indicates the starting point on a line segment, automatically follow the contours of the line (tracing). These programs are capable of tracing relatively simple lines. If the program cannot solve a graphical ambiguity on the raster map, it returns a message to the operator. Recently, vector editors capable of carrying out this digitization process have appeared, reducing the time consumption by a factor of 1.5 to 2. These can be called semi-automated systems⁷ [1, 2, 8, 9].

In theory, automatic digitization vector editors automatically digitize all objects of a given class, leaving the operator to correct errors in the resulting vector layers. The most popular vector editors use this approach. However, in practice, the high error level resulting from any small complication in the raster map means that alternative methods must be sought to reduce the high volume of manual corrections. It should be noted that the use of increasingly complex methods and algorithms for machine recognition of cartographic objects does not materially improve the results [1] (Cf. [4]). Why?

We suggested [1] that the answer is the development of a system approach to automatic map digitization, which uses the basic domination principle: "And he said unto them, 'Render therefore unto Caesar the things, which be Caesar's, and unto God the things, which be God's.' "8 (Cf. [2]). In the other words, methods and software should be developed which leave to the human operator only those tasks, which the computer cannot carry out. In the present case, this implies a detailed analysis of all existing map digitization processes, and the development of software to automate all technological vectorization operations, based on formal, heuristic, knowledge and interactive algorithms, which can effectively be used by the computer.

The above thesis is especially suitable for analytical GIS in contrast to register GIS, as the former do not require an extremely high level of geometric exactness in the cartographic materials, whereas they do require fast processing of a large number of vector layers. An example of analytical GIS is a GIS developed to solve territorial planning problems, while an example of a register GIS is GIS developed for a cadastral system. Our approach is effective for both types of GIS. However, the reader may think about the analytical GIS first.

Notice that the interactive and semi-automated modes can coincide if they are executed in real-time. Ot herwise we consider them as two different tasks

⁸ The Gospel from Lukas

An application of such system approach can greatly enhance the outcome not only of *processing*, the main stage of cartographic image recognition, itself, but also *pre-processing* (preparation of paper maps and their raster analogues) and *post-processing* (final processing of the results of automatic digitization). Thus, in the following three sections we shall consider all these processing as parts of the system approach.

2 RASTER MAP PRE-PROCESSING

The main goal of *pre-processing* is to prepare raster cartographic images in such a way as to simplify them and increase the reliability of their recognition in the automatic system.

Notation	Degree
Software for automation of the operation already exists	100
The operation can essentially be automated	75
The operation can be automated	50
The operation can partially be automated	25
The operation cannot be automated	0

Table 1. The degree of automation of raster-to-vector conversion operations.

Pre-processing operation	Degree
1. Preparation of the cartographic materials for	
scanning	
1.1 Restoration	0
1.2 Copying	0
1.3 Increasing the contrast of image objects	0
2. Scanning	
2.1 Test scanning	0
2.2 Definition of the optimal scanning parameters	50
2.3 Final scanning	0
2.4 Joining raster facets	100
2.5 Correction of the raster image geometry by the	100
reference points	
3. Preparation of the raster maps for recognition of	
the cartographic objects	
3.1 Edition of raster map	
3.1.1 Elimination of the map notations and map	25
legend	
3.1.2 Elimination of the artificial objects and	25
restoration covering by them images	
3.1.3 Restoration of the topology of cartographic	100
images in pixel level	
3.2 Separation of basic colors of the graphical	75
codification on a raster map	
3.3 Restoration of the color palette of a raster map	75
3.4 Stratification of raster map	
3.4.1 Stratification by reduced color palette	75
3.4.2 Logical stratification of the cartographic	25
objects	

Table 2. The degree of possible automation of pre-processing operations.

These general definitions of processing steps will be particularized in Section 5 Table 2 represents the proposed sequence of operations for the preparation of raster maps for automatic recognition and our very subjective opinion on their possible automation degree, using the notations given in Table 1¹⁰.

Restoration. Paper maps subject to storage and use show signs of wear on their surfaces; spots, scratches, and even notations. Even maps drawn on more stable media, such as cardboard, Mylar, organic glass, or aluminum are subject to the same defects. These signs of wear should be eliminated wherever possible; otherwise they impede the automatic recognition of cartographic features.

Copying. If the designer of a GIS project has copying equipment and the GIS is "analytical" not "register", then both manual and automatic vectorization are usually performed on copies of the map rather than on the original. The advantages include preservation of the original, parallel processing of the vectorization by several manual operators and registration of the graphical singularities of the objects of a given class for further processing of the copy.

Increasing the contrast of image objects. To simplify the process of vectorization, objects of the same class can be highlighted on a copy of the map using a contrasting color. Typically such marking involves enclosing linear objects such as rivers, roads or pipelines. In practice, outlines of polygonal objects, which do not have explicit borders (such as bogs, bushes, etc.), and are delineated only by dashed or patterned lines, must be drawn in. In particular, various polygonal objects may overlap, one represented by color, another outlined by a dashed line, and a third by patterned lines; in such cases, the objects must all be outlined explicitly.

Definition of the optimal scanning parameters. Both existing research and our experience in image recognition show that the efficiency of object identification depends on the choice of scanning parameters. Candidates are resolution, color palette, contrast, brightness, and visual effects. It is known that an optimal set of scanning parameters exists both for every scanner and for each map to be scanned [5]. The parameter set can only be tested by repeated test scans followed by visual or automatic analysis of the raster images obtained. An objective justification of the choice of optimal scanning parameters can be obtained by statistical calculation of the raster characteristics. Criteria for estimating the statistics can be developed from the point of view of effectiveness of the cartographic image recognition.

¹⁰ It would be interesting in the context of present work to compare not only the degree of automation of different procedures, but also their relative durations. This comparison will be a topic of our subsequent paper. In this work we may assume that the time is not essence for A2R2V conversion (however it may be so for a particular GIS-application)

Joining raster facets. The necessity of joining raster facets arises due to two main reasons; first, large-scale scanners are very expensive, and most users do not have access to them, and second, the cartographic materials required for the GIS-development often have different standards for different scales. In the present work, we may assume that software is available which can edge-match raster facets with minimal errors [4, 5].

Correction of the raster image geometry by the reference points. In practice, the raster image obtained after scanning is not uniformly deformed, due to printing errors, wear, scanning errors, and defects in edgematching. Raster transformation programs exist to eliminate or at least minimize these defects. This has the direct effect of increasing the accuracy of the vectorization of the final map and the indirect effect of ultimately improving image recognition. The general principle of raster map correction is that of plane transformation by reference points, i.e. displacement of certain points of the raster map those coordinates are known; followed by translation of the remaining elements of the raster correspondingly. Reliability of raster map correction is maximized when geodesic reference points are used as the control points. A satisfactory correction to the raster map can be provided by a coordinate grid on the original map. In this case, if the type and parameters of the cartographic projection are known, programs can be developed which generate a theoretically exact system of the reference points used for transformation of the raster map. If neither geodesic data nor a coordinate grid are available, larger-scale or same-scale vector maps which have already been corrected, or satellite images containing the reference points can successfully be used to correct the raster map. In this case, punctual features of the cartographic objects, such as river confluence's, road intersections, bridges, etc. can be chosen as control points.

Elimination of map notations and legend. After scanning, a raster map contains much graphical information that it is not necessary to vector, such as toponimics, annotations, explanations, copyright notice, etc. This information must be eliminated to simplify cartographic image recognition. The map legend must also be eliminated, and processed as a separate raster image.

Elimination of artifacts. Artifacts (traces of folds, stains, scratches, dust, etc.), which contain no information, are nearly always present on a raster map, and complicate subsequent cartographic object recognition. A difficulty in processing artifacts is the necessity of restoring the image previously covered by the artifact.

Restoration of the topology of cartographic images. As a result of printing and scanning errors, graphical images of the cartographic objects on a raster map frequently have topological defects. Typical defects are breaks in thin lines, which should be continuous (such as contour lines, rivers, etc.) and the fusion of lines that logically should

not intersect (e.g. contour lines). Topological errors in raster images complicate the recognition of cartographic objects, and gross errors, which can be noted by visual analysis of a raster map, must be corrected before the automatic recognition procedure is begun. The difficulty of solving this problem is increased due to the fact that such defects are often only visible, and can only be corrected, at the pixel level. Nevertheless, powerful programs for correction of raster maps at pixel level currently exist, providing hope for the solution of this problem [1, 4, 5].

Separation of basic colors. In graphical coding systems for cartographic information, a limited number of colors are used for a given map. However, after scanning, a raster map contains a large palette of colors, complicating image recognition. Analysis of the map legend colors and selection of the best quality segments in the raster field aid in reconstructing the basic graphical coding palette. Superfluous colors and their shades can then be substituted by the corresponding elements of the basic palette.

Stratification of raster map. A raster map, considered as a unified heterogeneous graphical image, is suitable for parallel human vision. In contrast, raster images, containing homogeneous graphical information, are suited to consecutive machine vision. Two approaches can be used for the stratification of the original raster map: 1) stratification by reduced color palette or 2) logical stratification of the cartographic objects.

In the first case, maps are derived from the original raster map which preserve only those pixels which have a strongly defined set of colors corresponding to the images of one class (for example, red and yellow, say, might correspond to the icons of populated places and the color of their outlines). In the second case, the map only preserves fragments of general raster images corresponding to the locations of cartographic objects of one class (for example, railways with adjacent thin zones).

The procedure of stratification by color is clear, therefore let us consider here the basic features of logical stratification of a raster map. The principle is that the presence of a cartographic object, which must be vectored in logically separated layers, is reliably known. Thus, the only task left for the recognition program is specification of the location of the object. This simplifies the problem and increases the reliability of the solution. Logical stratification of a raster map can be done manually. The operator moves a window of given size and shape, separating the fragments of the layer formed. The efficiency of logical stratification, even when performed manually, lies in that the main difficulty in vectoring raster maps is visual fixing of the coordinates of an object. In practice, this means that the operator must change the scale of the image and fix its coordinates using the cursor, perhaps two or three times, until a satisfactory result is obtained. On the other hand, defining the location of the object by specifying the size of a window is easier, especially when it is taken into consideration that the operator need not be concerned with any overlap of raster map facets, since the program can correct for this.

A typical situation that may arise is that a small-scale map already exists for a given territory. In this case, for logical stratification of a raster map, one must use methods for constructing buffer zones of the linear objects. These are already available in some popular vector editors [21].

3 RASTER MAP PROCESSING

The main goal of this principal stage of automatic vectorization of raster maps is the recognition of cartographic images; i.e. generation of vector layers and attribute information in electronic maps. From our point of view, the most promising line of software development is the development of methods, algorithms and programs that focus on locating and identifying specific cartographic objects. Each cartographic image has its own graphical representation parameters, which can be used for automatic object recognition on a raster map. The particular attributes depend on the topological class of the object. In traditional GIS, vector map objects are divided into three types: points, arcs and polygons, representing respectively punctual, linear and area objects. This classification can easily be extended to the analysis of cartographic images in raster maps. Objects are drawn on thematic maps in the form of graphical symbols, which are the same for all objects in a given group. Graphical images have geometric (location), topological and attribute (quantitative and qualitative parameters) information, which we can combine to form the concept of a cartographic image. The main geographical coding attributes of cartographic images of the three classes are shown in Table 3.

Object	Attribut	Attributes of the graphical representation			
Point	Form				
	Size				
	Painting	7			
Arc	Type				
	Color				
	Thickne	ess			
Polygon	Area			Contours	
	Filling	Dash	Crape	Туре	
	Color	Туре	Form	Color	
		Color	Size	Thickness	
		Thickness	Painting		
		Inclination	Density		
		Density			

Table 3. Main attributes used for graphical coding of cartographic images.

The classification of cartographic images is different when the vectorization of raster maps is considered. All objects on a raster map have area, and in this sense they are all polygons. It is not an easy problem to reduce the graphical coding elements of cartographic images to elements that correspond to the geometric categories "point" (a coordinate pair), "line" (a sequence of coordinate pairs) and "polygon" (a closed set of line segments which do not intersect and form the border of a geometrical figure). However, the classification of punctual, linear and polygonal objects must be preserved, because we can omit the relative stretch of the cartographic images in one or two directions (respectively lines or points) with respect to the stretch of the map field. A recognition program that recognizes, for example, punctual objects, does not have to distinguish between the punctual cartographic image itself or an element of a polygon fill pattern.

Notice that there may be other graphical objects involved in recognition of cartographic images, which are nearly always present in raster maps. Principally, these are letters and digits (toponimic, quantitative and qualitative characteristics of objects). Additionally, there may be other graphical elements in the map (footnotes to lines, insets, etc.) It is thus convenient to use the classification presented in Table 4.

Recognized object	Cartographic images and their		
	elements		
Point	Images of punctual objects		
	Polygon crape		
Arc	Images of linear objects		
	Explicit polygon contours		
	Dashed lines		
Polygon	Images of polygonal objects with		
	implicit contours given by:		
	FillingDash		
Text	 Crape Toponimics Altitude marks of relieve Road distances 		
	 Parameter values on the isolines 		
	 Passport characteristics of geodesic points and hydrometric posts, etc. 		
Additional graphics	 Text footnotes Guidelines Berg-dashes of relieve isolines 		
	 Circular diagrams, etc. 		

Table 4. Cartographic object classificati on from automatic vectorization point of view.

An important element of the automation of raster map vectorization is the development of an optimal sequence of steps for cartographic image recognition, successively eliminating elements already decoded from the raster map field and restoring images, which were hidden by the eliminated elements. The basic principle of this optimized ordering must be "from simple to complex". Nevertheless, the possibility of using information from objects already digitized (whether manually or by an automatic system) must be provided for in the development of a recognition strategy. For example, the punctual layer of hydrological monitoring posts can be successfully used for recognition of linear elements of the river network. Moreover, the symbols of these posts generally cover images of the river, complicating automatic identification of the rivers. Taking this into account, it becomes clear that hydrological monitoring posts must be vectored before the river network is digitized. Eliminating them from the raster map, one can use their locations and attribute data (mainly altitude marks) to aid in recognition of elements of the river network.

Further developing this approach, it is suggested to use already existing small-scale vector maps for recognition of corresponding cartographic images on large-scale maps. A small-scale map contains generalized (in a broad sense) information about a considerable proportion of the objects on the corresponding large-scale map. As a rule, the generalization involved in decreasing the map scale consists in the simplification of the geometric shape of the object and the elimination of a part of the object. For example, on a large-scale map, a river is represented by a polygon, but on the small-scale map, as a line. In general, a given object can be expected to change in topological type when the degree of generalization changes. Even if the topological type of an object is preserved after generalization, several objects on a large-scale map may correspond to a single object on a small-scale map. Examples of the correspondence between the objects in maps of different scales are presented in Table 5.

Scale		Examples		
Fine	Coarse	Fine	Coarse	
Point	Point	Altitude marks of relieve	Altitude marks of relieve	
	Line	Out of scale irrigation massif	Irrigation channels	
	Polygon	Out of scale habitual point	Territory of habitual point	
Line	Point	Linear rows of pores	Single pores	
	Line	Line of electric transmission	Line of electric transmission	
	Polygon	Line of riverbed	River water area	
Polygon	Point	Territory of water supply	Single pores	
	Line	Territory of water supply	Linear rows of pores	
	Polygon	Bog territory	Bog regions	

Table 5. Correspondence of cartographic objects in maps of different scales.

The use of small-scale maps solves a difficult problem in automatic digitization: the search for objects in the whole raster map field. In this case, a vectored object can be found in the nearest neighborhood of its generalized analogue, and nowhere else.

The search zone for paired punctual objects can be restricted to a circle with a radius defined by the correlation between the scales of the vectored maps and the maps used.

We suggest the use of the "caterpillar" algorithm (the name reflects the shape of the illustrated algorithm) for searching for paired linear objects. The caterpillar algorithm involves the construction of a system of line segments perpendicular to the contour of their small-scale analogue, divided in half by it. The length of each segment can be chosen by the correlation between the scales of the maps used and their density, i.e. by the curvature of the generalized line. Moreover, the search object is located along segments constructed in this way. The sequence of reference points of the search curve can thus be found. The reference points obtained can be joined by straight-line segments in an interactive digitization system without any intervention by the operator.

Automatic cartographic image recognition is simplified and its reliability increased by the use of corresponding vector layers of a small-scale map for digitization of isoline and other regular systems of linear objects (such as the coordinate grid, or urban blocks with linear or radial planning). But in this case not all lines of a large-scale map have small-scale analogues. For example, the contour lines on a vectored 1:50,000 map may have 10m density while on the corresponding 1:250,000 map they have 50m density. In such a case, the contour lines that have counterparts in the generalization (0, 50, 100, etc.) are vectored first by the caterpillar algorithm. Next, the "stairs" algorithm (the name reflects the shape of the illustrated algorithm) is applied for the recognition of the intermediate contour lines. The stairs algorithm constructs a system of curves between each adjacent pair of already vectored contour lines, which are perpendicular to each of these contour lines. The density of these curves is defined by the curvature of the basic lines, just as in the caterpillar algorithm. Moreover, points of the adjacent contour lines to be searched for are located along the curves constructed in this way. Between two index contour lines, the number of additional lines to be found is well defined (for example, between two contour lines of 100 and 150m four additional contour lines of 110, 120, 130 and 140m always exist and can be found). Once all the required reference points have been found, it is clear that they can be joined in succession using the program tools given by the caterpillar algorithm.

The sequence of reference points of a vectored liobject can be copied from the layer, which contain corresponding punctual objects. For example, shoreline structures (hydrometric monitoring posts, bridges, docks etc.) can be used as reference points to digitize the contours of rivers and lakes. The hydrometric monitoring posts are particularly useful here. Their coordinates and attribute data (name of the river or lake and altitude mark) can be used in automatic recognition algorithms for the elements of the hydrological network on the raster map. Notice that in this case automatic digitizing reverses the order of operations compared to traditional techniques. Traditionally, the operator first digitized the hydrological network manually, and then vectored the location points of the shoreline structures using vector-editing tools.

Operation	Degree
1. Development of the strategy of automatic	
digitization of raster maps	
2. Definition of sound matrices of raster maps	
2.1 Classification of recognized objects	0
2.2 Definition of the size, form and color filling of	75
basic sound matrices of raster maps	
2.3 Estimation of statistical weights of single	75
elements of cartographic images	
3. Recognition of cartographic images	
3.1 Digitization of objects which have vector	75
analogues	
3.2 Digitization of objects which have not	50
vector analogues	
3.3 Elimination of superfluous recognized objects	0
4. Recognition of attributive data of vectored	
objects	
4.1 Classification of attributive information	0
carriers	
4.2 Localization and identification of attributive	75
information	
4.3 Correction of errors of attributive data	75
recognition	
5. Elimination of recognized images from raster	
тар	
5.1 Restoration of image covered by recognized	75
object	
5.2 Correction of restored image	75

Table 6. The degree of possible automation of the processing operations.

In other words, maximal use of already existing information (directly or indirectly related to the vectored objects) employed as a general principle of automatic cartographic image recognition can increase efficiency and reliability. For example, algorithms that use digital models of a region, and that are based on small-scale maps can be produced for digitization of the hydrological network. If the layers are already vectored, this can be used to generate the sequence of reference points of the curves to be recognized; otherwise these points can be indicated manually as described above. This simplifies automatic digitization and increases its reliability.

Summarizing the processing of raster maps, we notice that the methods and algorithms used for this process must provide complete, even redundant cartographic image recognition to eliminate erroneous recognition of objects, since visual control and correction of the vector layers can be carried out more quickly than manual digitization of missed objects. To conclude the discussion in this section, we notice that the process of automatic cartographic image recognition (processing), from our point of view, should follow the scheme presented in the Table 6, where, as before, we assign scores indicating the degree of possible automation of the various steps involved in processing.

4 RASTER MAP POST-PROCESSING

The main goal of the *post-processing* of raster maps (after cartographic image recognition) is an automatic correction of vectorization errors.

For automatic correction of raster map digitization we suggest *two approaches*: using 1) the topological characteristics of objects in vector layers and 2) the spatial correlation (connectivity) of the corresponding vector layers.

The first approach is based on the fact that many cartographic objects in the system have well-defined topological characteristics, which can be used for the correction of vectorization errors. Let us give some obvious examples.

Isolines. The topological characteristics of isoline systems (for example, contour lines) are: a) isolines are continuous, b) they cannot intersect each other, c) each isoline is either closed or starts and finishes at a domain boundary, and d) polygons which cover the whole domain without intersections can be assembled from the arcs of the correct isoline system together with the domain However on a raster boundaries. map characteristics, as a rule, may be lost due to several reasons: a) the lines are broken where a parameter value for a given isoline is written, b) some sections of isolines are not well drawn in high density regions, and c) the raster images of some isolines merge due to defects of printing and scanning the paper maps. The "tick marks" (small segments of fixed length, which are perpendicular to the isolines and drawn in the direction of the decreasing of cartographic parameter) need special consideration. These elements of the map's graphical design, if not recognized as the parts of the isoline system, hinder the correct assembly of the polygons and either must be eliminated or (better) detached in a separate vector layer. They can be restored on the vector map and used for the automatic attribution of polygons assembled from the contour lines.

Hydrologic network. The relevant topological characteristics of the hydrological network (rivers, channels, brooks, etc.) are that its vector image, within the limits of each watershed, has the structure of a simple connected directed tree graph. However on raster maps, images of water bodies can be disrupted by images of other types: area (reservoir, cities, etc.) and punctual

(populated places, bridges, etc.) that break the structure of the hydrologic network.

Region maps. The topological characteristics of a region map (for example, the map of political division of a country) is that each arc, being a segment of the boundary of a disjoint region, is either closed or starts and finishes at another arc. Thus a correct system of non-intersecting polygons covering the whole surface of the map can be assembled from the boundaries of the regions, adding, if necessary, the domain boundaries. However on the vectored region map, the boundaries of some regions may be discontinuous, impeding the production of the correct map topology.

The second approach is newer and offers more potential. It consists in using connectivity, or spatial correlations among the various vector objects to correct the results of automatic digitization and the attributes of raster map cartographic images. Let us explain this by the following examples.

Hydrosphere elements. The vectored elements of the hydrosphere include objects of all three topological types: polygonal (e.g. seas, lakes, reservoirs, ponds, bogs), linear (e.g. rivers, channels, brooks) and punctual (e.g. springs, out-of-scale reservoirs). It is clear that nearly all hydrosphere elements have certain spatial relationships among one other. For example, springs or lakes can be the sources of rivers (main hydrosphere elements). Linear river sections can connect polygonal hydrosphere elements (extended sections of riverbeds, lakes or reservoirs). Rivers discharge into seas or lakes. These spatial relationships are not only necessary for a topologically correct vector image of the hydrosphere, but can be also used to correct automatic digitization results of the river network, where the digitized polygonal and punctual hydrosphere elements are available.

Relief. The relief of a region on the maps is represented by objects of two topological types: 1) punctual (trigonometric points, hydrometric monitoring posts, reference points with coordinates obtained by the Global Positioning System (GPS)), and 2) linear (contour lines, relief features—slopes, ravines, landslides, etc.). All these objects have correlations among each other due to their location and in particular to the altitudes given as attribute information.

We notice particularly the relationship between altitude and the hydrologic network; altitude decreases monotonically along a river in the direction of flow, thus each contour line intersects the river either not at all or exactly once. Due to this relationship, it is convenient to produce the vector map of the hydrological network before digitizing the contour lines and to subsequently use the digitized hydrological network for correction of the results of contour digitization.

Road network. The road network on the map represents roads, highways and railways of all sizes. Its elements form a graph, the majority of those nodes represent populated places and punctual objects of the transportation infrastructure (e.g. bridges, tunnels, railway crossings, docks, airports) and can be used for the correction of results of the automatic digitization of the road network. Moreover the construction of modern roadways leads to the transformation of the natural relief of the region that, as a rule, is represented by short linear objects, which describe the relief (e.g. embankments, excavations). If these objects are digitized before vectorization, then their geometrical and attribute information can be used in the algorithms to correct results of the automatic digitization of the road network.

Administrative and political divisions. The borders of the administrative and political division of a territory often follow rivers, roads and other linear elements of the cartographic structure. Prior digitization of these elements can be used for the correction of results of automatic digitization of the corresponding layers of raster maps.

The examples presented in this section show that the characteristics of internal structure and relationships between the vector objects can be used effectively in automatic correction of errors of the automatic vectorization of raster maps. In practice it means the development of more specific software for automatic cartographic image recognition.

Summarizing the discussion of this section, we propose that the process of automatic correction of results of automatic cartographic image recognition (post-processing) follows the scheme presented in the following Table 7, where we show, as before, our evaluation of the degree of possible automation of the corresponding procedures, expressed as the scores defined in Table 1.

Operation	Degree
Correction of vector layers based on peculiarities of their internal topology	75
Correction of vector layers based on their spatial and logical correlation with other vector objects	50
Final correction of vector layers in whole electronic map system	0

Table 7. The degree of possible automation of post-processing operations

5 A2R2V

5.1 INTRODUCTION

In this section, we overview single modules of our map interpretation system called A2R2V effected to date [3].

5.2 HOW DOES IT WORK?

A2R2V is based on semantic analysis of color images. Semantic analysis of cartographic images is interpreted as a separate representation of cartographic patterns (alphanumeric, punctual, linear, and area). Our map interpretation system explores the idea of synthesis of invariant graphic images at low level processing (vectorization and segmentation). This means that we ran "vectorization-recognition" and "segmentation-interpretation" systems simultaneously. Although these systems can generate some errors, they are much more useful for the following understanding algorithms because its output is nearly recognized objects of interest.

We begun map recognition from global binarization followed by classical OCR-identification with artificial neural networks (ANN), supervised clustering, knowledge-based recognition rules, and morphologybased vectorization. To overcome the problem of laborintensive training, we designed simplified images. For this purpose, we used the linear combinations of color components or image representations (false color technique) and binary representation composing (composite image technique). These techniques are application-independent. However, in the frameworks of our approach, map recognition may be treated as a common (application-dependent) task [6]. We followed the concept that the important semantic information necessary to interpret an image is not represented in single pixels but in meaningful image segments and their mutual relations [8].

We set forth a conception of composite image representation and decomposition. The main goal of image decomposition consists of the object linking by its associated names. We use an image synthesis based on object-fitting compact hierarchical segmentation [10]. We composite perform representations (or simply, composites) of the source image by means of a reduced number of color or tone components and segments. In this manner, visually perceived objects are not eliminated and the image's semantics is preserved. The image's semantics in this context corresponds to the association of segment fields of different hierarchical levels being identified with identifying conceptions from the subject domain. For example, detection of a segment set identifying a coastline or highway becomes semantically meaningful. Further, this set of segments is renamed as "coastline", "highway", etc.

The composite images form a "book" in which the objects of interest can be found and recognized on appropriate page(s). Thus, a "page number" defines the method of thresholding and some tuning parameters (Figure 1).

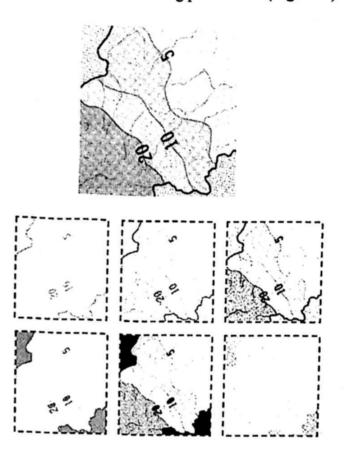


Figure 1. Source image (top) and some "pages" of a "book" of composite images

5.3 SACR

SACR is System of Alphanumeric Character Recognition. We use the false color technique of image segmentation [11], which we aim to separate alphanumeric characters (for subsequent recognition) into raster-scanned color cartographic maps, using model of color RGB. Note that this is a very difficult problem because there is either text embedded in graphic components, or text touching graphics. The processed images only maintain pixels that truly belong to objects we wish to segment, eliminating all pixels not of interest that are produced by noise or obtained due to erroneous selection of scan parameters (Figure 2).

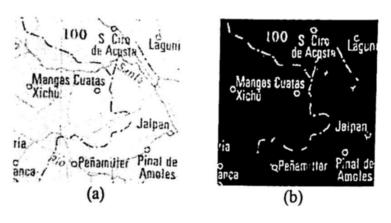


Figure 2. (a) - source image; (b) - segmented image

The following identification of segmented objects is particular, previously prepared application. We first build the strings of characters and use a set of ANN for

identification of single letters and gazetteers (toponimic dictionaries) for word identification.

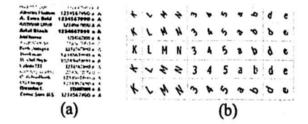


Table 8. (a) - 16 selected fonts for testing previously trained ANN; (b) - example of inclination of the characters Abadi MT Condensed Light font in the sizes of 13 ÷17 points

Moreover, ANN are previously trained on a set of artificially created characters of different fonts, sizes, inclinations, etc. (Table 8). However, we apply these ANN to identification of real characters into cartographic images using 1,025 alphanumeric characters (upper and low cases, different fonts, sizes and inclinations) extracted from real raster maps. 95.5% of which were successfully recognized by the system. A gazetteer is used to verify either an identified string of characters forms meaningful word, or not [3].

5.4 SPOR

SPOR is System of Punctual Object Recognition. This nearly full automatic system, which generates vector layer of punctual cartographic patterns (Figure 3), is described in detail in [1].

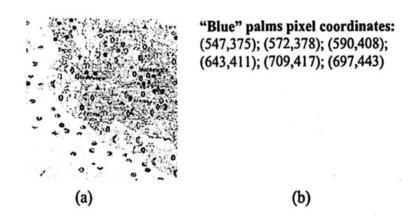


Figure 3. (a) – recognized "blue" palm re storation (in white) over source image; (b) – pixel coordi nates of recognized palms

5.5 SLOR

SLOR is System of Linear Object Recognition. This is mainly semi-automated system that provides full image treatment: pre-processing (image enhancement and segmentation), processing (a set of supervised identifiers and knowledge-based recognition rules), and post-processing (morphology-based vectorization). Thus, human-machine interaction is optimized [3]. A result of full processing of linear cartographic patterns is shown in Figure 4.



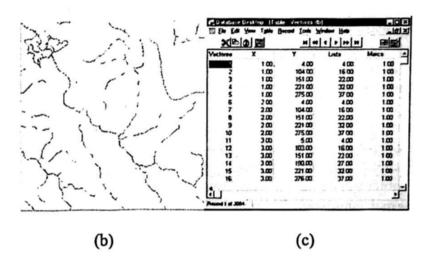


Figure 4. (a) - source image; (b) - vector image of rivers; (c) - associated to (b) vector database

5.6 SAOR

SAOR is System of Area Object Recognition. A result of application of composites to an administrative map is shown in Figure 5.

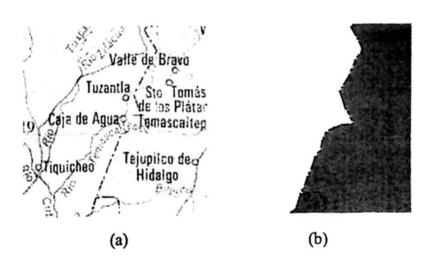


Figure 5. (a) - source image; (b) - recognized polygons

5.7 COMMENTS AND SECTION 5 SUMMING-UP

(a) We used OCR-identification with ANN in SACR and SPOR systems only, training ANN on sets of artificial characters, but applying them to real map characters; (b) We used perceptron multi layer-multi neuron feedforward to recognize alphanumeric and punctual characters. In the first segmentation, we selected and labeled the connected components to be analyzed by

ANN; (c) If characters touch graphics, we developed a method to separate and further recognize touching symbols [3, 11]; (d) Size of training set is 95,000 samples; (e) ANN have a special output to show shapes that do not appear to be characters of interest; (f) Linear character recognition does not use OCR; (g) Composites provide area character recognition because segmented area objects are labeled by the system in same, but are different for each type of objects, gray-level values (Section 5.6).

The processing of spatial data is a complicated interaction of the hardware, software, data, methods and people. Following to [7], GIS is a computer system that consists of a database in which the spatial and descriptive information of geographical environment as a part of the real world is stored. Moreover it permits the input, maintenance, analysis, transformation, manipulation and representation of spatial data of some geographical point. The basic element of any GIS is a set of vector cartographic maps. Note that some popular commercial systems are very expensive and often ugly to perform problem-oriented applications due to the absence of format standards, ill representations of spatial databases or impossibility to process color images.

We design A2R2V as a multi-environment system, which does not depend on particular data format and oriented to process color raster-scanned images. Its GUI is described as follows. At low level, it has the modules of automatic recognition of alphanumeric, punctual, linear and area objects into raster-scanned color cartographic images11, which have been programmed in C. Thus, a database with quantitative, qualitative and nominal features is associated to each recognized visual object! At the intermediate level, a JAVA module supports correct conversion of corresponding database to be processed at high level in one of the environments: UNIX, LINUX or Windows. Hence the user of our system should not seek for unified or particular representation of spatial databases in an environment, but only chooses most convenient one from given three. We are now developing a decision making tool to support most adequate to particular GISapplication user's choice.

In this section, we presented a system for automatic interpretation of raster-scanned color cartographic maps. The highlight of our system is two "intelligent" color image segmentation techniques. They are followed by a set of identifiers that includes classical OCR-identification, supervised clustering, knowledge-based recognition rules, and morphology-based vectorization. The identifiers and vectorization worked well because they receive as input from the segmentation step nearly recognized objects of interest (alphanumerical, punctual, linear, and area). In our experiments, we used complex raster-scanned color cartographic images, which, being

intermediate between drawings and natural images provide a nearly ideal model for testing because they have characteristics of both [3].

At the same time, automatic interpretation of color cartographic images presents certain difficulties for stateof-the art in image processing and also artificial intelligence. A set of vector maps that one can expect as an output of such interpretation is very useful for Geographical Information Systems, new map production, and old map actualization. However, it appears unrealistic to obtain a fully automatic computer-based interpretation system free of errors [1-2, 12-16]. Additionally, please note that high efficiency of interpretation is required for vector map production and actualization first. It seems reasonable to obtain in both cases 90÷95% successfully recognized objects. This is to avoid excessive work on corrections of errors produced by the computer system that can sometimes be greater than manual raster-tovector conversion [17].

We believe that only a system approach to the problem can be fruitful. In the context of the A2R2V system, this means first, decomposition of source image by multiple hierarchical components to achieve a stable, accurate representation in the presence of degraded images. Second is segmentation with mutual recognition of appropriate primitives. Third is the development of a unified knowledge-based learning and self-learning system with optimal human-machine interaction for color cartographic image treatment. Our future research will be concerned with this approach.

Finally, attempting to answer the question stated in Introduction, we define the degree of automation of each module of A2R2V system (in terms of system's errors based on our experimentation) as follows in Table 9¹².

A2R2V module	Degree of participation (%)		
	Human	Machine	
SACR	5÷15	85÷95	
SPOR	5÷10	90÷95	
SLOR	5÷30	70÷95	
SAOR	< 15	> 85	

Table 9. Which part of conversion process do the operator and which part by the machine carry out?

6.CONCLUSIONS

Summarizing the analysis of the process of automatic raster map digitization we notice that this problem is of great current interest and we conjecture that the most promising line of progress towards its solution lies in successively increasing automation of the separate links of the system approach considered here. We suggest as a

¹¹ In the other words, these modules provide raster-to-vector automatic conversion of scanned color cartographic images.

¹² In this work we keep away an important problem of quantitative estimation of the efficiency of the conversion system [2, 18-20]. This problem will be considered in a s ubsequent paper (Cf. footnote 5)

main principle of such automation, maximal use of already existing and reworked vector and attribute information in the preparation, digitization and correction algorithms of each succeeding vector layer. This information can be organized and effectively used as a knowledge base in the conversion system. Thus, the system approach presented here can be combined with knowledge-based approach to R2V conversion. We already included into A2R2V system some elements of "intelligence" to increase its effectiveness. The following examples illustrate this statement: 1) a simple idea to use the gazetteers in SACR increased significantly that system efficiency; 2) the knowledge-based recognition rules combined with morphology-based vectorization in SLOR led to automatic generation of vector databases associated to linear cartographic patterns; 3) the use of knowledge allows to run simultaneously "low-intermediate-high level processing" as in the false color and composite techniques. Now we develop the "fine-to-coarse scale technology" of the conversion in which the "knowledge" of cartographic patterns into small-scale map aids in recognizing the corresponding patterns into large-scale map of the same territory.

[2] stated that an efficient R2V conversion system is not just a union of its parts, but their complex interaction that defines the effectiveness of the system. From our point of view, the best way to design that interaction is the use of knowledge.

REFERENCES

- [1] S. Levachkine and E. Polchkov, Integrated technique for automated digitization of raster maps, Revista Digital Universitaria, 1(1), 2000, on-line: http://www.revista.unam.mx/vol.1/art4/ (ISSN: 1607-6079)
- [2] E. Bodansky, System approach to a raster-to-vector conversion: From research to commercial system, In: Levachkine, S. Bodansky, E., Ruas, A. (eds.), e-Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2002), 3-4 December 2002, Mexico City, Mexico (2002) (CD ISBN: 970-18-8521-X).
- [3] S. Levachkine, A. Velázquez, V. Alexandrov and M. Kharinov, Semantic analysis and recognition of raster-scanned color cartographic images, *Lecture Notes in Computer Science*, Vol. 2390, 2002, 178-189.
- [4] S. Ablameyko, T. Pridmore, Machine Interpretation of Line Drawing Images, Technical Drawings, Maps, and Diagrams (Ed.: Springer-Verlag, Berlin-Heidelberg-New York, 2000).
- [5] S. Levachkine, Digitalización automatizada de mapas raster, *Revista Digital Universitaria*, 1(0), 2000, on-line:
- http://www.revista.unam.mx/vol.1/art3/arti3.html/ (ISSN: 1607-6079)

- [6] D. Doermann, An introduction to vectorization and segmentation, *Lecture Notes in Computer Science*, Vol. 1389, 1998, 1-8.
- [7] M. F. Goodchild, Geographical Information Science, International Journal of Geographical Information Systems, 6(1), 1992, 31-45.
- [8] Definiens Imaging GmbH e-Cognition: Object Oriented Image Analysis. http://www.definiens-imaging.com/ecognition/
- [9] Able Software Co.: R2V http://www.ablesw.com/r2v/
- [10] V. Alexandrov, M. Kharinov, A. Velázquez and S. Levachkine, Object-oriented color image segmentation, Proc. IASTED International Conference on Signal Processing, Pattern Recognition, and Applications (SPPRA 2002), Crete, Greece, June 25-28, 2002, 493-498.
- [11] A. Velázquez, H. Sossa and S. Levachkine, On the segmentation of color cartographic images, *Lecture Notes in Computer Science*, Vol. 2396, 2002, 387-395.
- [12] G.K. Meyers and C.-H. Chen, Verification—based approach for automated text and feature extraction from raster-scanned maps, *Lecture Notes in Computer Science*, Vol. 1072, 1996, 190-203.
- [13] M.P. Deseilligny, R. Mariani and J. Labiche, Topographic maps automatic interpretation: Some proposed strategies, *Lecture Notes in Computer Science*, Vol. 1389, 1998, 175-193.
- [14] F. Dupont, M.P. Deseilligny and M. Gonrad, Automatic interpretation of scanned maps: Reconstruction of contour lines, *Lecture Notes in Computer Science, Vol.* 1389, 1998, 194-206.
- [15] S. Frischknecht and E. Kanani, Automatic interpretation of scanned topographic maps: A raster-based approach, *Lecture Notes in Computer Science, Vol.* 1389, 1998, 207-220.
- [16] J.G.M. Schavemaker and M.J.T. Reinders, Information fusion for conflict resolution in map interpretation, *Lecture Notes in Computer Science*, Vol. 1389, 1998, 231-242.
- [17] Parcel mapping using GIS. A guide to digital parcel map development for Massachusetts Office of Geography Information and Analysis for MASSGIS. August 1999: http://umass.edu/tei/ogia/parcelguide/
- [18] H. Osamu, D.S. Doermann, Quantitative measurement of the performance of raster-to-vector conversion algorithms, *Lecture Notes in Computer Science*, Vol. 1072, 1996, 57-68.
- [19] L. Wenyin, and D. Dori, Genericity in graphics recognition algorithms, *Lecture Notes in Computer Science Vol. 1389*, 1998, 9-20.
- [20] P. Ishan, Liang Jisheng, A.K. Chhabra, R. Haralick, A performance evaluation protocol for graphics recognition systems, *Lecture Notes in Computer Science Vol.* 1389, 1998, 372-389.
- [21] Resident Ltd.: MapEdit http://www.resident.ru/

TEXT/GRAPHICS SEPARATION IN RASTER-SCANNED COLOR CARTOGRAPHIC MAPS

Aurelio VELÁZQUEZ¹, Serguei LEVACHKINE²

¹Instituto Mexicano del Petróleo
Eje Central Lázaro Cárdenas 152, México D. F. 07730, MÉXICO

avelaz@imp.mx

²Centro de Investigación en Computación-Instituto Politécnico Nacional
U.P. Adolfo López Mateos, Edif. CIC, México D.F. 07738, MÉXICO

palych@cic.ipn.mx

ABSTRACT

In this paper, we propose a method to separate and recognize to the fullest extent possible those characters that are touching or even overlapping each other. The characters are processed in raster-scanned color cartographic maps. The map is segmented first to extract all text strings including those that are touched other symbols and strokes. Second, OCR-based recognition with Artificial Neural Networks (ANN) is applied to define the coordinates, size and orientation of alphanumeric character strings in each case presented in the map. Third, four straight lines or a number of curves computed in function of primarily recognized by ANN characters are extrapolated to separate those symbols that are attached. Finally, the separated characters input into ANN again to be finally identified. Experimental results showed 95÷97% of successfully recognized alphanumeric in raster-scanned color maps.

KEY WORDS

Cartographic Color Map, Text Segmentation, Character Recognition, Artificial Neural Networks.

1. INTRODUCTION

The maps such are electrical schematics, cartographic color maps, topographic maps, cadastral maps, engineering drawings, technical illustrations, public service maps and other thematic maps contain a lot of text.

Cartographic color maps are plenty of punctual, linear and area objects. With points, lines and areas a map is able to model the real world and the man-made boundaries, using a map scale, which is a ratio of distance on the map to distance on the Earth. To describe those objects there can be used symbols (portrays) and labels (alphanumeric characters) presenting a great variety of features, some of them in equal shape but in different color. Different colors are used to represent different objects, including a number of character's fonts. These can be colored and following different paths in all kind of angles.

The process of development of a Geographical Information System (GIS) includes the selection of the paper and raster maps for vectorization Levachkine et al. [1]. To be included into GIS, the paper maps should be changed to a computer readable format, normally a raster format. After that, the raster maps can be converted into vector format that is most adequate to GIS-applications. In the context of raster-to-vector conversion of graphical documents, the problem of text recognition is of special interest, because textual information can be used for verification of vectorization results (post-processing).

The retrieval of all presented elements in a map can be made manually or by means of a computer system. In the former case, the map is scanned in a raster format and then converts to vector. Before a raster-to-vector conversion a map segmentation and recognition are usually employed.

General frameworks. The text segmentation and its subsequent recognition in raster images are very difficult problems because, in general, there is either text embedded in graphic components, or text touching graphics Doermann [14]. These challenging problems have been received numerous responses from the graphic recognition community Nagy [15]. However, there have not been developed efficient programs that solve the task automatically. Thus, the main idea of the most works is to put the operator in the loop. As proposed, for example, by Ganesan [16], the operator can draw line through the text, marking it as text and revealing its orientation all in one step. Fletcher et al. [17] and Tan et al. [18] developed the algorithms to extract text strings from text/graphics image. Both methods however assume that the text does not touch or overlap with graphics. For maps, the problem is much more complex since the touching or overlapping as well as many other character configurations are commonly presented in maps. Cao et al. [20] proposed a specific method of detecting and extracting characters that are touching graphics in raster-scanned color maps. It is based on observation that the constituent strokes of characters are usually short segments in comparison with those of graphics. It combines line continuation with the feature line width to decompose and reconstruct segments

underlying the region of intersection. Experimental results showed that proposed method slightly improved the percentage of correctly detected text as well as the accuracy of character recognition with OCR.

Segmentation. Applying color and spatial attributes to segment thematic maps, Silva [2] used a 300-dpi resolution in a RGB color system to perform a Karhunen-Loeve transformation. Luo et al. [3] used the directional morphological operations. They coded images by runlength-encoded as an enchained list, deleting the text that is represented by lines, and finally subtracting the new image from the original one to obtain an image with text without lines. In [4], Li described the Comb algorithm based on the best common structure of local minima found at a moment to search for global minima. He used the concept of maximum a posteriori (MAP) and Markov random fields (MRF) as the frameworks. To segment text from engineering drawings Adam et al. [11] used Fourier-Mellin transform in a five-step process. Using a heuristics, they found broken chains. In [12], Hase et al. described a three-step algorithm of segmentation called "multi-stage relaxation". However, they do not recognize characters. In [5], Levachkine et al. used false colors in a RGB model. They applied different combinations of R, G and B basic colors to segment map objects, and then a neighborhood analysis to recover or eliminate pixels.

Extraction and recognition. Some proposals for character extraction and recognition to be mentioned are as follows. In [6], Myers et al. described the verification-based approach for automated text and feature extraction from raster-scanned maps. They used a gazetteer to propose a forecasting hypothesis, which characters are in labels and where is their position in the map, having the information from other map in a different scale. Character and text boxes are used in [7] by Wenyin et al. The authors considered only horizontal and vertical text in which a character box is a rectangle with rate sides are no larger than 10 pixels to join character boxes. Thus, they built the text box that can grow horizontally or vertically under a threshold to fit the letters. Using directional morphological operations Luo et al. [3] separated the text from lines but not from curves. Deseilligny et al. [8] proposed different knowledge levels to solve the task. They begun with an analysis of related components (semiologic), then built the character chains (syntactic), detected related characters (higher semiologic level) and, finally, following the natural language rules corrected the text (semantic). Using templates Friscknecht et al. [9] linked them with symbols and characters. The approach does not require the complete template. It is pondered and hierarchically built. To retrieve street names Nagy et al. [10] used one of the four black layers. Taking the hue component from a HSV model for segmentation, they subtracted the street layer from the black layer and then made a connected component analysis to distinguish text characters. An efficient system to recognize characters by means of adaptive ANN is described in [13] by Velázquez et al. To

train ANN, they used characters from a word processor in different fonts, sizes and inclinations by applying them to identify a great variety of characters in cartographic maps.

The rest of paper is organized as follows. In Section 2, we describe an alphanumeric segmentation-recognition system. In Section 3, we consider a case of touching and overlapping characters presented in raster-scanned color cartographic maps. A method (*V-lines* and *V-curves*) to separate and further recognize the touching and overlapping characters is described in this section as well. Section 4 contains paper's conclusion.

2. SEGMENTATION OF OVERALL MAP CHARACTERS

A raster map has to be segmented first. All its elements should be retrieved with their coordinates and features, and then sent to the corresponding thematic layers. These layers could be symbols, rivers, landmarks, roads, railroads, pipelines, isoclines, natural and artificial surroundings, words and numbers, lakes and other punctual, linear and polygonal bodies.

Cartographic maps are the most complex graphical documents due to the high density of information that they contain. A typical example is shown in Figure 1 (RGB image). There are labels of different types, sizes, orientations and colors, roads, rivers, symbols and artificial surroundings. Some strokes are touching each other's.



Figure 1. Example of color cartographic map with different types of characters.

To obtain a binary image from color image, the former must be changed to a gray-level. One way to make this change is to convert the RGB model to the YIQ model, where the luminance (Y) is a gray-level image. Another way is to average the R, G and B values. Figure 2 shows the gray-level image obtained from the image of Figure 1. In this work, we used both conversion procedures as well as their combination as described in [5].



Figure 8. V-lines to separate the letter 'S' from the symbol is bellow it.

A similar process to separate the letters 'n' and 'o' from the letter 'S' is below them is applied to the image shown in Figure 7d to obtain the whole word "Ochentaiuno" (a special process is required to separate the 'n' from the 'o'). Used lines for this task are shown in Figure 9.



Figure 9. V-lines to separate the letters 'n' and 'o' from the letter 'S' are below them.

If the upper case characters are not horizontal, we can trace a diagonal rectangle following the same angle that those letters have. In figure 10, the first 'S' is touching the river's line which name is 'BALSAS'.



Figure 10. Label of 'BALSAS' river is touched the line of the river by the first letter 'S'.

With the other five letters, we can build a rectangle that covers all the word letters. Using the upper and lower pixels from the first and last characters of the chain, we compute their mean value obtaining two points to trace a line with them as shown in Figure 11.

Then the left or upper (if the chain is vertical) point is moved one pixel up and one left so that we have two new points, the right or lower point is moved one pixel down and one right so that there are four points for the first dynamic rectangle.



Figure 11. Starting with a line following the label's angle, a dynamic rectangle is built.

Computing the position of each character in the largest lines, we can find if there are outside pixels. The shortest lines are used for the same purpose with the first and last characters. If there are outside pixels in a line, it is moved one pixel in corresponding direction. If more than one line has outside pixels, all those lines are moved. The process is continued until no more outside pixel are found.

Now, it is possible to identify the missing letter using only the pixels inside of the rectangle box, sending it to ANN and testing the word 'BALSAS' in the gazetteer.

On the other hand, for upper case letters, four lines should be computed. Figure 12 shows the word 'Chicayán' touching a line of the river labeled for it.

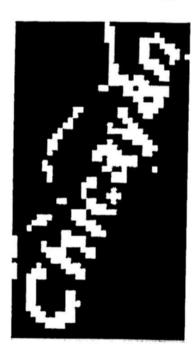


Figure 12. Label of 'Chicayán' river is touched the line of the river with the letter 'á'.

Using the procedure employed for capital letters, it is possible to construct a rectangle as shown in Figure 13. The largest lines will be used to find the four V-lines, if they are present. Normally, three of them are present.

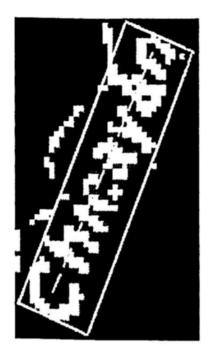


Figure 13. First, it is necessary to build a rectangle as it was made with capital letters.

Each line should be moved with pixel-by-pixel procedure until it is reached the next numbered level. An additional adjustment is made: each ending point is moved up three pixels and the line that better fits one of the inner lines is selected as the "leader". To the other three lines, if exist, we assign the same angle that the leader has. Figure 14 shows these lines.

Cutting the label with line one, the missing letter can be analyzed. It could be recognized as a letter 'd', but the word 'Chicaydn' does not exist in gazetteer. Thus, we attempt now with line two. Then, the letter can be interpreted as an 'a' and the word 'Chicayan' is already in the gazetteer.

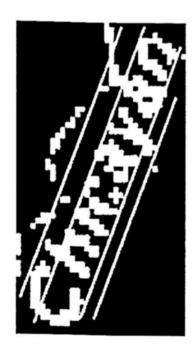


Figure 14. Four lines to unglue letter 'a' from the riverbed.

There are other manners in which objects can touch an alphanumeric character: at its left or right, at its top or bottom as in example shown in Figure 15, where letter 'M' is touched at its left by a state boundary.

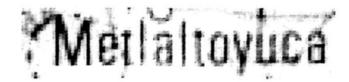


Figure 15. The word is touched at one extreme of the chain.

The V-lines are not helpful themselves to separate the characters. However, the V-lines are useful to build a "growing" rectangle that is fitted to the character's pixels, identifying the characters with the ANN and using the gazetteer until it matches with a correct word. The growing rectangle is shown in Figure 16.

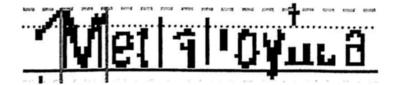


Figure 16. "Growing" rectangle to identify the letter 'M'.

To build the rectangle, we employ the following steps: 1) if the character is at left, we start at the beginning of the first letter. At this moment, we have lines two, three and four; there is another line, the second from the top to bottom. We use a tolerance of one third of the distance between lines two and three. A perpendicular line, that begins at line three plus tolerance and ends at line two minus tolerance, is moved left until an appropriated pixel is found, 2) the line moved is the first line of the rectangle. Other is formed by copying this line two pixels left. The other lines are two and three, unless there are pixels outside those lines, but inside of the tolerance. The first line found is the "anchor", all others can be removed, and 3) left line is moved pixel-by-pixel. There could be a motion of upper or lower lines always inside of the tolerance. Each time the line is moved, the pixels in the rectangle are analyzed and tested by ANN and the gazetteer. The process is continued until a correct word is found or the distance between the line and the anchor is more than one and half times of the distance between upper and lower lines.

Unfortunately, there are some man-made errors on the maps and, even though our method outputs the complete chain of characters, the word could be misspelled. Figure 17a shows a word where letters 'l' and 'n' are touched by two lines. After the processing with V-lines shown in Figure 17b is applied the chain "Tulacingo" is built, but it is not in the gazetteer, because the right word is "Tulancingo" in which the letter 'n' was missing. This error can be corrected by another procedure [22].

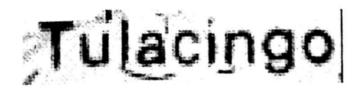


Figure 17a. Original chain touched two arcs of a boundary.



Figure 17b. Characters were unglued and the chain 'Tulacingo' was build.

On the other hand, some labels are nearly impossible to detect because the background features are too close to their own features. Figure 18a shows such a label and Figure 18b shows its binary image with the damaged characters hard to identify.

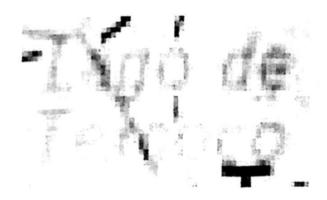


Figure 18a. Label overlapped by other objects with similar attributes.



Figure 18b. Binary image of Figure 18a, showing distorted characters.

Another example is shown in Figure 19. It is impossible to detect, in automatic way, the chain of characters because all of them are touched other elements.



Figure 19. Original image a) and its binary representation b).

In last two cases the operator intervention is certainly required as of the philosophy by Bodansky [23] or (better) by Gelbukh et al. [22].

V-curves. Take a look to Figure 20a. It displays a curvilinear text associated to a riverbed with the letter 'g' touching the line of the river (Figure 20b shows corresponding binary image).



Figure 20a. Color image with a curvilinear text "Tulancingo" (name of the river).



Figure 20b. Binary image of Figure 20a.

An application of V-lines method is difficult in such a case. However, we can use the following procedure that we call *V-curves* to solve the task. The text is divided into blocks of fixed (or nearly fixed) inclination. To each block the V-lines method is applied. Thus obtained lines are connected by a linear extrapolation, forming linear splines. These splines are V-curves. The following steps are similar to those that were used in V-lines method.

Because of the paper space limit we have to stop our explanations on V-curves method. Probably, it will be a topic of our subsequent paper.

4. CONCLUSION

In this paper, a method to separate and recognize touching and overlapping characters in raster-scanned color cartographic images has been developed. The method performs the process to segment the character layer and most valid (or "geographically meaningful") words are built. Though some words cannot be obtained complete, the system is able to suggest one word from a gazetteer to support the operator to resolve the ambiguous cases. OCRbased recognition procedure with ANN applied to the case of study possesses some peculiarities. ANN were tested first with synthetic characters. Among 18,432 synthetic samples, 23 were not recognized, giving 99.87% of successfully recognized characters. After that the same ANN were employed for the characters of cartographic maps on a set of 2,125 samples. The results gave 93.21% of success. These results were improved with the V-lines to some 96.73%.

REFERENCES

- [1] S. Levachkine and E. Polchkov, Integrated technique for automated digitization of raster maps, *Revista Digital Universitaria*, 1(1), 2000, on-line: http://www.revista.unam.mx/vol.1/art4/ (ISSN: 1607-6079)
- [2] C.J. Silva, Segmentation of thematic maps using color and spatial attributes", Lecture Notes in Computer Science Vol. 1389, 1998, 221-230.
- [3] H. Luo and K. Rangachar, Improved directional morphological operations for separation of characters from maps/graphics, *Lecture Notes in Computer Science, Vol.* 1389, 1998, 35-47.
- [4] S.Z. Li, Toward global solution to map image restoration and segmentation using common structure of local minima, *Pattern Recognition*, 33(1), 2000, 715-723.
- [5] S. Levachkine, A. Velázquez, V. Alexandrov and M. Kharinov, Semantic analysis and recognition of raster-scanned color cartographic images, *Lecture Notes in Computer Science*, Vol. 2390, 2002, 178-189.
- [6] G.K. Meyers and C.H. Chen, Verification-based approach for automated text and feature extraction from raster-scanned maps, *Lecture Notes in Computer Science*, *Vol. 1072*, 1996, 190-203.
- [7] L. Wenyin, and D. Dori, Genericity in graphics recognition algorithms, *Lecture Notes in Computer Science Vol. 1389*, 1998, 9-20.
- [8] M.P. Deseilligny, R. Mariani and J. Labiche, Topographic maps automatic interpretation: Some proposed strategies, *Lecture Notes in Computer Science*, Vol. 1389, 1998, 175-193.
- [9] S. Frischknecht, E. Kanani, Automatic interpretation of scanned topographic maps: a raster-based approach, Lecture Notes in Computer Science, Vol. 1389, 1998, 207-220.
- [10] G.A. Nagy, S. Seth Samal, T. Fisher, E. Guthmann, K. Kalafala, L. Li, P. Sarkar, S. Sivasubramanian and Y. Xu, A prototype for adaptive association of street names with streets maps, *Lecture Notes in Computer Science Vol. 1389*, 1998, 302-313.

- [11] S. Adam, J. M. Ogier, C. Cariou, R. Mullot, J. Labiche and J. Gardes, Symbol and character recognition: application to engineering drawings, *International Journal on Document Analysis and Recognition (IJDAR)*, 2000, 89-101.
- [12] H. Hase, T. Shinokawa, M. Yoneda and C. Y. Suen, Character String Extraction from Color Documents, *Pattern Recognition*, 34(1), 2001, 1349-1365.
- [13] A. Velázquez, J. H. Sossa and S. Levachkine, Reconocimiento eficiente de caracteres alfanuméricos provenientes de mapas ráster por medio de clasificadores neuronales, Computación y Sistemas. Revista Iberoamericana de Computación, 6(1), 2002, 38-50.
- [14] D. Doermann, An introduction to vectorization and segmentation, *Lecture Notes in Computer Science*, Vol. 1389, 1998, 1-8.
- [15] G. Nagy, Twenty years of document image analysis in PAMI, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 2000, 38-62.
- [16] A. Ganesan, Integration of surveying and cadastral GIS: From field-to-fabric & land records-to-fabric, Proceedings of 22nd ESRI User Conference, 2002, http://gis.esri.com/library/userconf/proc02/abstracts/a0868.html
- [17] L.A. Fletcher and R. Kasturi, A robust algorithm for text string separation from mixed text/graphics images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6), 1988, 910-918.
- [18] C.L. Tan and P.O. Ng, Text extraction using pyramid, *Pattern Recognition*, 31(1), 1998, 63-72.
- [19] A. Velázquez, Localización, recuperación e identificación de la capa de caracteres contenida en los planos cartográficos. *Ph.D. Thesis*. Centre for Computing Research-IPN. Mexico City, Mexico, 2002 (in Spanish).
- [20] R. Cao and C.L. Tam, Text/Graphics separation in maps, Lecture Notes in Computer Science, Vol. 2390, 2002, 168-177.
- [21] V. Alexandrov, M. Kharinov, A. Velázquez and S. Levachkine, Object-oriented color image segmentation, *Proc. IASTED International Conference on Signal Processing, Pattern Recognition, and Applications (SPPRA 2002)*, Crete, Greece, June 25-28, 2002, 493-498.
- [22] A. Gelbukh and S. Levachkine, Error detection and correction in toponym recognition in cartographic maps, In: Levachkine, S. Bodansky, E., Ruas, A. (eds.), e-Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2002), 3-4 December 2002, Mexico City, Mexico (2002) (CD ISBN: 970-18-8521-X).
- [23] E. Bodansky, System approach to a R2V conversion: From research to commercial system, In: Levachkine, S. Bodansky, E., Ruas, A. (eds.), e-Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2002), 3-4 December 2002, Mexico City, Mexico (2002) (CD ISBN: 970-18-8521-X).

COMBINING SOURCES OF EVIDENCE TO RESOLVE AMBIGUITIES IN TOPONYM RECOGNITION IN CARTOGRAPHIC MAPS*

Alexander GELBUKH *‡ SangYong HAN‡

* Natural Language Processing Lab, Center for Computing
Research CIC-IPN, UPALM Zacatenco,
Mexico City, 07738 MEXICO
E-mail: gelbukh@cic.ipn.mx; www.gelbukh.com
and

†Department of Computer Science and Engineering, Chung-Ang University, 221 Huksuk-Dong, DongJak-Ku, Seoul, 156-756, KOREA E-mail: hansy@cau.ac.kr Serguei LEVACHKINE †
† Image Processing and PR Lab, Centre for Computing
Research-IPN, UPALM Zacatenco, Ed. CIC,
Mexico City, 07738 MEXICO
E-mail: palych@cic.ipn.mx

ABSTRACT

Graphical documents such as cartographic maps contain a great variety of textual elements appearing in different spatial positions, in different fonts, sizes, and colors, touching and overlapping graphical symbols. This greatly complicates automatic optical recognition of such textual elements in the process of raster-to-vector conversion of graphical documents. In this work, we propose a method that combines OCR-based text recognition in raster-scanned maps with heuristics specially adapted for cartographic data to resolve the recognition ambiguities using various sources of evidence. Our goal is to form in the vector thematic layers geographically meaningful words correctly attached to the cartographic objects.

KEY WORDS

Raster-to-Vector Conversion, Geographic Information Systems, Optical Character Recognition, Spelling Correction

1. INTRODUCTION

Huge amount of geographic information collected in the last centuries is available in the form of maps printed or drawn on paper. To store, search, distribute, and view these maps in the electronic form they are to be converted in one of digital formats developed for this purpose. The simplest way of such conversion is scanning the paper map to obtain an image (a picture) stored in any of the raster graphical formats such as TIFF, GIF, etc. After that, a raster-to-vector conversion can be applied to include obtained vector maps into a Geographic Information System (GIS).

Though raster representation has important advantages in comparison with the hard copy form, it still does not allow semantic processing of the information shown in the map, for example:

- Search for objects: Where is Pittsburgh? What large river is there in Brazil?
- Answering questions on the spatial relations: Are Himalayas in China? Is Nepal in Himalayas? Is a part of Himalayas in China?
- Generation of specialized maps: Generate a map of railroads and highways of France.
- Scaling and zooming: Generate a 1:125 000 map of Colombia. Now, show more details at the point under cursor.
- Compression: Objects such as points, arcs, or areas can be stored much more efficiently than pixels

Note that these are semantic tasks rather than image manipulation. E.g., when zooming in or out, objects and, most importantly, their names should appear or disappear rather than become smaller or larger. Indeed, when zooming out the area of London, the name Greenwich should not become small to unreadable but should disappear (and appear in an appropriate font size when zooming in).

This suggests storing and handling of a map as a database of objects (points, arcs, areas, alphanumeric, etc.)—vector database—having certain properties, such as size, geographic coordinates, topology, and name. Specifically, the name of the object is to be stored as a letter string rather than a set of pixels as originally scanned from the hard copy. Thus, such vector representation can solve the listed above semantic tasks, but only to some extent [1].

^{*} Work done under partial support of Mexican Government (CONACYT, SNI, IPN), Korean Government (KIPA Professorship for Visiting Faculty Positions in Korea), and ITRI of the Chung-Ang University. The first author is currently on Sabbatical leave at Chung-Ang University.

However, automatic recognition of such strings (toponyms) in the raster image of the map presents some particular difficulties as compared with the optical character recognition (OCR) task applied to standard texts such as books:

- The strings are out of context, which prevents from using standard spelling correction techniques based on the linguistic properties of coherent text. Often such strings are even not words of a (modern) language, which further limits applicability of the standard linguistic-based spelling correction methods [12].
- The background of the string in the map is very noisy since it can contain elements of geographic notation such as shading or hatching, cartographic objects such as cities or rivers, and even parts of other strings, e.g., name of a city inside of the area covered by the name of the country; see Figure 1.
- In addition, the letters of the string are not properly aligned but instead are printed under different angles and along an arc; this happens with the names of linear and area objects, e.g., rivers or countries; see Figure 1.
- Unlike standard task, in toponym recognition it is not only required to recognize the string itself but also to associate it with a specific cartographic object, e.g., city, river, desert, etc.

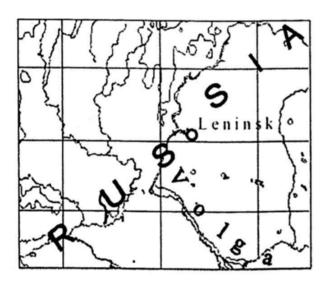


Figure 1. Intersection of string in a map

On the other hand, in many cases additional information is available that can give useful cues for ambiguity resolution. One of such information sources is existing databases (usually available from the country Government, postal service, etc.) providing spatial relationships between entities (e.g., a list of cities classified by administrative units) or even exact coordinates; see [14] for extensive discussion of this topic.

In this paper we discuss how such additional information can be used to workaround the problems arising in recognition of the inscriptions in the maps, associating them with specific cartographic objects, and importing information on these objects from available databases. First we describe the general scheme of our method. Then we discuss various sources of evidence taking into consideration, when available, in error detection and correction: information of the existing names and linguistic information, on the distribution of the letters of the string in the source raster image, and pre-existing geographic information such as coordinates of objects. Then global verification of consistency of the recognition results is described. Finally, conclusions are drawn and future work directions are outlined.

2. PREVIOUS WORK AND PRESENT PAPER OVERVIEW

The text segmentation and its subsequent recognition in raster images are very difficult problems due to the presence of the text embedded in graphic components and the text touching graphics [2]. These challenging problems have received numerous contributions from the graphic recognition community [3]. However, there have not been yet developed any efficient programs to solve the task automatically. Thus, in the most works human operator is involved. For example, [4] proposes that the operator draws a line through the text, marking it as text and revealing its orientation.

In [5] and [6], the algorithms are developed to extract text strings from text/graphics images. However, both methods assume that the text does not touch or overlap with graphics. For maps, the problem is much more complex, since the touching or overlapping as well as many other character configurations are commonly presented in maps. That is why [7], [8], and [9] developed the methods for text/graphics separation in raster-scanned (color) cartographic maps.

In [9] a specific method of detecting and extracting characters that are touching graphics in raster-scanned color maps is proposed. It is based on observation that the constituent strokes of characters are usually shot segments in comparison with those of graphics. It combines line continuation with the feature line width to decompose and reconstruct segments underlying the region of intersection. Experimental results showed that proposed method slightly improved the percentage of correctly detected text as well as the accuracy of character recognition with OCR.

In [7] and [8], the map is first segmented to extract all text strings including those that are touched other symbols and strokes. Then, OCR using Artificial Neural Networks (ANN) is applied to output the coordinates, size, and orientation of alphanumeric character strings present in the map. Then, four straight lines or a number of curves computed in function of primarily recognized by ANN characters are extrapolated to separate those symbols that are attached. Finally, the separated characters are input into ANN again for their final identification. Experimental results showed 95–97% of successfully

recognized alphanumeric symbols in raster-scanned color maps.

In the present work, we use the output obtained with this method in combination with pre-existing geographical information in semantic analysis of ambiguities for "geographically meaningful" word formation. We focus on text processing rather than image processing.

The proposed system is based both on the traditional techniques used in the general-purpose OCR programs and on the techniques we developed especially for cartographic maps. In particular, Section 5 deals with the problems and solutions common to any OCR task. However, even in these cases there are some differences with respect to the usual OCR situation. The algorithm described in Section 5.1 (check against a dictionary of existing words) in our case has to deal with much more noisy strings than usual OCR programs developed for clean black-on-white running text. The same can be said of Section 5.2 (non-uniform spatial letter distribution): in maps the letters are often placed at significant distances one from another, cf. Figure 1; as well of Section 5.3 (check against the general laws of a given language): maps have many foreign or indigenous words that do not conform to the main language of the given territory.

In contrast, Section 4 is specific for maps. In Section 4.3 (check against geographic information such as expected coordinates) the consistency with the available information about the location of an object is used, which is specific for cartographic maps. Also the information on the expected type of the object (river, mountain, etc.) is used. In Section 4.4 (global consistency check) it is verified that each object is recognized only once. These techniques do not have direct analogs in standard OCR research and thus are contributions of our paper.

Finally, we do not use many techniques standard for usual text OCR, which are applicable to running text but not to toponyms in maps, for example: morphological and syntactic analysis, semantic consistency verification [13]; paragraph layout determination, etc. In a way, the new techniques we introduce in the Section 4 play the same role of verification of contextual consistency, but in the manner very specific to cartographic maps.

3. MAIN PROCEDURE

We rely on a basic OCR procedure¹ (not discussed here; see [1], [7], and [8]) that recognizes in the map individual letters and groups together the letters of a similar font and color located next to each other, thus forming a hypothetical string. In this process, errors of various types

¹ Our method does not depend on how text strings have been extracted and recognized. Neither does it depend much on the type of graphical document being processed. It can be adapted to different subject domains.

can be introduced; our purpose is to detect and correct them.

The recognition algorithm for the whole map works iteratively. At each step, the basic OCR procedure selects for processing the longest and most clearly recognized string and returns it for error correction and subsequent adding to the database being constructed. Upon its processing, the string is removed from the raster image, and the next string is selected. The algorithm stops when no more letter strings can be found in the raster image.

This design allows for recognition of the names of large areas, which are usually represented by large letters scattered across the corresponding area, with many names of smaller objects between the letters of the area name. In the example shown in Figure 1, first the word Leninsk will be recognized and removed from the image, then the word Volga, and only then the letters of the word Russia can be grouped together in a string.

The basic OCR procedure returns, for each string it recognizes, the string itself, e.g., "RUSSIA," and the geographic coordinates in the map of the frame containing each individual letter, e.g., R, U, etc.

After this process, two major issues arise:

- How to associate the textual objects found in the map with the geographical objects found in the same map? In Figure 1, what are the type (city, river, mountain, etc.) and the coordinates of the object called Leninsk? What is the name of the city located near the center of the map?
- How to detect and correct possible recognition errors in the textual elements?

To solve these problems, various sources of evidence are to be taken into account. In the following, we will consider each such source of evidence, first for the association problem and then for the error detection and correction problem.

4. ASSOCIATION OF A NAME WITH AN OBJECT

As we have assumed, the basic OCR procedure returns two types of information:

- Geographical objects. These can be of three types: punctual, linear, and area objects. For them, the basic OCR procedure returns the corresponding coordinates in the map (in pixels or in the corresponding geographical units) as will be discussed in the following.
- Textual information. The basic OCR procedure returns a string along with the coordinates (again, in

pixels or in geographical units) of a box containing each of its letters.

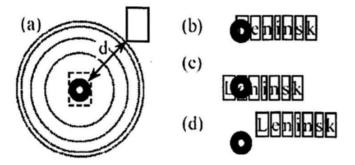


Figure 2. Simplified model for punctual objects

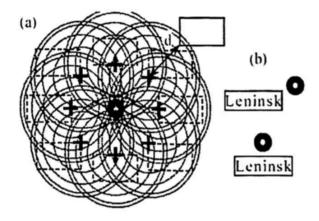


Figure 3. Improved model for punctual objects

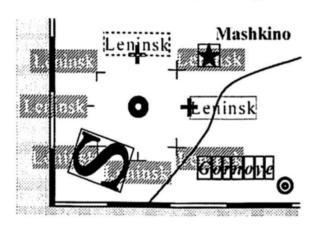


Figure 4. Constrained placement strategy

The next task for the map recognition system is to relate the strings (toponyms) with the objects found in the map. This is a non-trivial task due to several peculiarities.

First, it is highly heuristic since one needs to model the way in which the human cartographer assigned the labels to the objects. Second, not all objects have a corresponding label as well as not all labels correspond to objects detectable in the map.

The assignment procedure consists of two major parts: estimating of probability of a string to be related to an object, and final assignment of the strings to objects in a way that maximizes such probability. In the subsequent sections we shall consider these two tasks separately.

4.1 LIKELIHOOD OF RELATEDNESS BETWEEN A STRING AND AN OBJECT

Given a geographic object and a string, both along with their coordinates (in pixels or in geographical units), we can estimate the probability of that the string is related to the object. Using the Bayes formula, we can do it by modeling the process of placement of the names in the map by the cartographer.

Indeed, denote by R the event that the string is related to the object and by P the event that the cartographer placed the string to a specific position in the map (where we observed it). By Bayes formula, the desired probability is:

$$P(R \mid P) = P(P \mid R) \frac{P(R)}{P(P)}$$
 (1)

Since P(P) does not depend on a specific object and thus does not affect the disambiguation decisions, the desired probability is determined by the following two factors:

- P(P | R) reflects the strategy used by the cartographer to place the names of the objects next to the objects,
- P(R) reflects the relatedness of the name with the given object.

These values can be estimated heuristically taking into consideration various sources of evidence. We define these sources as mean proportional value:

$$P = \sqrt[n]{\prod_{i=0}^{k} P_i}, \tag{2}$$

where P is P(P | R) or P(R), correspondingly, and Pi are the probabilities contributed according to each source of evidence.

In what follows, we discuss various independent sources of evidence used in our method.

4.2 SPATIAL EVIDENCE

To define the probability $P = P(P \mid R)$ of placing the object's name in the observed specific position where it has supposedly been found in the map, we should model the strategy used by the cartographer for placing the name of this object. Then, we can assume that various (independent) random factors may cause the cartographer to deviate from the "optimal" position. The effect of various independent factors is approximated well by the normal distribution:

$$P = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{d^2}{2\sigma^2}},\tag{3}$$

where d is the distance from the actual inscription to the optimal position predicted by the model, and σ is a coefficient (dispersion) depending on the scale of the map and the fonts used; its selection is discussed in the following. If the model predicts several possible placements $x_1, ..., x_n$ with probabilities $p_1, ..., p_n$ and dispersions $\sigma_1, ..., \sigma_n$, then we assume:

$$P = \sum_{i=1}^{n} p_{i} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{d_{i}^{2}}{2\sigma_{i}^{2}}},$$
 (4)

where d_i are the distances from the observed placement to the corresponding coordinates x_i.

The placement strategies are different for punctual, linear, and are objects.

Punctual objects. For a punctual object (such as a city), which is represented by only one coordinate pair p, our previous work [15] suggested the following strategy of placing its name. We assumed that the inscription is expected to be next to the point p. Thus, we computed the distance d as minimum distance from the point p to any of the frames containing the individual letters of the string, as shown in Figure 2 (a). Though this simple model is a reasonably good approximation, it is not very precise. For example, both placements shown in Figure 2 (b) and (c) are predicted by the model to be optimal while (d) is not; this is contra-intuitive.

The model can be improved as shown in Figure 3 (a). First, we observe that the names of punctual objects (unlike the names of linear or area objects; see Figure 1) are aligned along a straight line; thus, instead of the frames of individual letters as in Figure 2, the frame containing the whole string can be considered. Second, we consider eight possible placement strategies shown in Figure 3 (a). The name is placed next to the object, not overlapping with the object, in some small distance from the object. This distance is approximately the size of one letter. We suggest that the dispersion σ_i from (4) should be also approximately of the size of a letter frame.

With this improvement, examples of (locally) optimal placements are shown in Figure 2 (d) and Figure 3 (b), while the placements shown in Figure 2 (b) and (c) are, in accordance to our intuition, not optimal.

The eight strategies have different probabilities pi in (4). For example, in the languages with left-to-right writing system, the placements to the right of the object are

preferred to those to the left. The procedure for determining these parameters is described here in Section 4.5

What is more, not all of the eight placement strategies can be possible in a specific environment. A specific placement strategy is not possible if the string would significantly overlap with any of other objects found in the map:

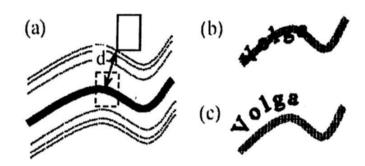


Figure 5. Simplified model for linear objects

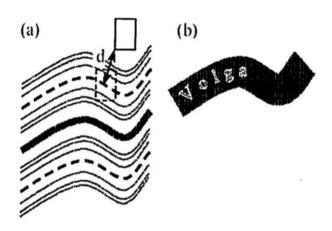


Figure 4. Improved model for linear objects

- Letters of other textual elements. The string may be placed between the letters of a string of a larger font, given that it does not overlap with individual letters.
- Other punctual objects. The string may, however, overlap with other linear or area objects.
- Borders of the map. The string cannot trespass beyond the area of the map.

An example is shown in Figure 4, where only two placement strategies are possible. In such cases, in the formula (4) the probabilities p_i of the impossible cases are set to zero and the other p_i are re-normalized. Alternatively, instead of setting the corresponding probabilities to zero, they can be significantly decreased (penalized).

Note that we assume that the operation of relating the names with the objects is performed after independent recognition of all objects and all strings in the map, so that the positions of other strings and objects are known at this moment.

Linear objects. For linear objects (such as rivers) represented by a sequence of coordinate pairs p_i , we suggest similar improvements over the procedure proposed in [15]. In the latter work, we indicated that the

¹ The model can be further improved taking into consideration that each individual distribution in (4) is not symmetrical: deviations that do not change the distance from the object are most probable, and toward the object are less probable that those away from the object. This can be done by a suitable deformation of the coordinate system; we omit here the details of this procedure.

slightly simplified way, to measure the distance between a letter and the broken line, two adjacent points x_i , x_{i+1} nearest to the letter are found and the distance from the letter to the straight line connecting the two points has been determined.

Similarly to the case of punctual objects, this would lead to a situation shown in Figure 5 (a), which incorrectly predicts the case (b) rather than (c) to be optimal. As in the case of punctual objects, we also suggest considering two placement strategies shown in Figure 6 (a), which correctly predict the case (c) and not (b) in Figure 5 to be optimal. The parameters (such as p_i and σ) and constraints (such as those shown in Figure 4) are treated much in the same way as in the case of punctual objects discussed above; we skip here the details.

An exception is the linear objects with the width significantly greater than the size of the letters in the string. In this case, the old model should be used, as shown in Figure 6 (b); namely, the string is expected to be found in the middle of the line. Note that our processing of such objects is different from that of area objects in that the string does not need to cover the whole length of the object.

Area objects. For an area object S (such as a province) represented by a sequence of coordinate pairs xi corresponding to its contour, our previous work [15] suggested the following approach. The inscription is expected to be in the middle of the area and the letters are expected to be distributed by the whole area. Thus, we can take $d = \iint_{\mathbb{R}^n} f(x, y) dx dy$ in (3), where f(x, y) is the minimum distance from the point (x,y) to any of the letters of the string. The integral is taken over the intersection S' of the area S and the whole area of the given map (in case a part of S proves to be out of the boundaries of the given map). Note that a similar integral along the contour would not give the desired effect. Since the area objects are much less numerous than other types of the objects in the map, we do not consider computational efficiency a major issue for our purposes. Neither precision is important for us. Thus, it is enough to compute the integral by, say, Monte-Carlo method.

Now we can re-interpret this procedure along the lines of the approach described in detail for the punctual and linear objects. Namely, the string that minimizes the integral above is the predicted "optimal" placement, with individual letters uniformly covering the surface of the area object. The observed placement can differ from the predicted one with the probability given by (3).

However, in this case we deal with a set of objects—individual letters—and not with one object, the whole string. Their distribution by the map can be considered independent. Thus, the probability of a specific configuration of n letters is:

$$P = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{d_i^2}{2\sigma^2}},$$
 (5)

where d_i are the distances between the predicted and actual location of individual letters.

This new interpretation allows for a meaningful choice of the parameter σ in (5), which in [15] was left undefined. The deviation in the placement of each letter can be of the order of about 1/3 of the distance between the letters in the predicted string, for the inscription not too look too misplaced.

Computationally, the task of finding of the optimal (predicted) string that minimizes the integral discussed above can be treated, for example, with 2n-dimensional gradient descent.

As in the case of linear objects, exceptions are to be considered. If the area object is too small in comparison with the font used in the string, it should be considered as punctual object. If the area object is similar to a line (very much longer in one dimension than in the other one), it can be treated as a thick or thin linear object, see Figure 6 (b).

4.3 APPROPRIATENESS OF A NAME FOR AN OBJECT

The previous section dealt with the component $P(P \mid R)$ of (1) reflecting the placement strategy used by the cartographer. In contrast, this and the following sections deal with the component P = P(R), which reflects the appropriateness of a particular name for a particular object, independently of the physical location of the string on the map.

Each of the following subsections discusses a specific contribution $P = P_i$ in the total probability. These contributions are combined by (2). In all cases, except for the next subsection, such probabilities are, though, binary: the combinations of a name and an object are classified into possible and impossible ones.

Typographic Evidence. As we have mentioned, the basic OCR procedure returns the coordinates of each letter. This can give us two characteristics of the recognized string:

- Whether the letters are aligned along a straight line,
- The distance between each adjacent pair of letters.

Only the names of linear and area objects (e.g., rivers or lakes), but not punctual objects (e.g., cities), can have non-linear letter alignment. Non-linear alignment is admitted for non-punctual objects but not required.

It is the responsibility of the basic OCR procedure to evaluate the probability P of that a string is linearly

aligned, which is to be used in case of a punctual object. Note that this condition is not applicable to linear and area objects.

Notational Evidence. Notation in the map gives additional information to filter out impossible combinations of names and objects. In some maps, rivers are explicitly marked as "river" or "r." and similarly mountains, peninsulas, etc. Specific font family, size, and color are usually associated with various types of objects (e.g., cities, and rivers). Though this information can provide very good filtering, it is not standard and is to be automatically learnt or manually specified for each individual map, which limits the usefulness of such filtering capability in a practical application.

Automatic learning of notation is discussed in Section 4.5. Alternatively, the system can provide the operator with the means to specify such notational elements, at least the prefixes such as "river". Similarly, font features for a specific type of objects can be automatically learnt from a large map or specified by the operator.

The importance of recognition of such notational information is two-fold. First, it helps filtering out impossible combinations: for example, the name of a punctual object cannot be specified as the name of a river. Another use of notational information is discussed in the next subsection.

Some precautions should be taken with such type of filters. For example, in Spanish rivers are marked as "rio" 'river'; however the string RÍO DE JANEIRO should not be filtered out as a possible name of the city (given that capital letters are not properly distinguished in the map).

Geographic Evidence. This is a very powerful source of evidence, though it relies on extensive databases not always available. Suppose the string is found in a dictionary (database) that provides at least two types of spatial information on the corresponding object:

- Its inclusion in a larger area, such as a province, state, etc. These areas form a hierarchy.
- Its geographic coordinates.

This information can be used to verify that the object in question recognized in the map satisfied the constraints specified by the database for the string in question.

Note that when only the hierarchical information is available (for example, "Jalapa city is in Oaxaca state"), this can be used to filter out undesirable variants only if the coordinates are available for one of larger areas, one or more steps up the hierarchy (but small enough to serve for disambiguation). Alternatively, it might happen that the corresponding larger area has been earlier recognized in the same map. Unfortunately, due to the order of recognition from smaller to larger objects (see the

beginning of Section 3), it is hardly probable. The corresponding check can be performed at the post-processing stage—global verification, see Section 4.4, when all areas have been already recognized.

In the best case, the full coordinate information is available in the dictionary for the object. Then the task of verification is greatly simplified, provided that the coordinate grid is reliably recognized for the given map.

The dictionary frequently contains several objects with same name, of the same or different type. When analyzing a map of *Canada*, the object corresponding to a recognized string *London* is to be a small Canadian city and not the large British city, so that the correct number of inhabitants for the object could be imported from the dictionary to the database being constructed. When analyzing an inscription *Moscow* in the coordinates (57°N, 35°E), its interpretation as a river rather than city is more probable.

Note that for correct identification of geographic information associated with a toponym, some information about notational conventions is important for addressing the dictionary. Indeed, for the string "river Thames" what is to be looked up in the dictionary is "Thames" and not "river Thames".

Linguistic Evidence. This is a substitute for the lack of knowledge on notation in a specific map. In some languages, the names of rivers, mountains, cities, etc., tend to follow some patterns that can be specified in the linguistic module of the recognition system. For example, in English a name ending in -town is more probable for a city than for a river. In Russian, a name ending in -ka is probable for a river or village, but not for a mountain. In Korean, a name ending in -do would probably indicate an island and -gan a river.

Obviously, these clues should be taken into account as factors in the total probability and not as rigid constraints (unless they are rigid constraints in the language in hand).

4.4 VERIFICATION OF GLOBAL CONSTRAINTS

After all inscriptions in the map have been recognized, some global constraints should be checked.

Uniqueness. To each object only one inscription should correspond. If two inscriptions have been associated with the same object, one or both of them is to be re-assigned. Even though the information on the probability of each of the two candidates is available at this point and could allow for automatic selection of one of the candidates, we believe that such conflicts should not be arbitrated automatically but the human intervention is to be requested instead. Of course, the probability information can be used to suggest most likely variant to the human operator.

An exception from this rule is linear objects such as long rivers. Several inscriptions can be assigned to such an object if their text is the same, the distance between them is much larger than their lengths, and their length are much smaller than the length of the object (river).

Inclusion. The hierarchical information available from the dictionary (see Section 4.3) can be applied at this point. Recall that our algorithm recognizes the names of, say, cities before recognition of the names of areas. So at the time of recognition of the string "Xalapa" the information "Xalapa City is in Veracruz State" could not be checked since we did not know yet where Veracruz State is in the map. Now that all strings have been recognized, this information can be checked (we already know where Veracruz is) and the error discussed in Section 5 (Xalapa mistaken for Jalapa recognized in Oaxaca State) can be detected.

4.5 MODEL CALIBRATION AND AUTOMATIC LEARNING OF PARAMETERS

The process described in the previous sections depends on a number of parameters, such as dispersion values or notational conventions. For their automatic learning iterative model calibration is used.

First, some approximate values are set as discussed in the previous sections. Then the automatic procedure of recognition of the map is executed. As a result, a (possibly incorrectly) recognized map is obtained.

Our hypothesis is that many of the elements in such a map will be recognized correctly from the first attempt. So statistics built for the results of this recognition—such as the average deviation of the strings from the predicted locations—is expected to be a good approximation of the real values.

With this new information, the parameters of the model (such as dispersion values) are adjusted, and the automatic recognition is performed again. The process is repeated iteratively a predefined number of times or until convergence. Since the results of the whole procedure are discrete values—associations between strings and objects—convergence can be indicated by repetition of exactly the same result.

In our previous work, we have successfully applied this procedure to learning the parameters of a syntactic parser for natural language sentences [10].

With this procedure, not only numerical parameters can be learnt, but also notational conventions such as the fonts and colors associated with specific types of objects (rivers, cities, mountains), typical prefixes or suffixes of their names (such as r. for river o mt. for mountain), etc.

An alternative way of automatic detection of such prefixes in a large map is the use of a dictionary. For each string consisting of several words, both the complete variant and the variants without the first (or last) word are to be tested. If for a specific type of objects (e.g., rivers) in most cases the string is found after taking off a specific word (e.g., "river"), then it is to be considered as notation for this type of objects.

5. SPELLING CORRECTION IN TOPONYM RECOGNITION

Due to a very complicated layout of objects and textual elements in cartographic maps, words can be recognized with errors, e.g., "RNSoSIA" for "RUSSIA" where U is erroneously recognized as N due to a nearby river, and the circle representing a city is erroneously taken for the letter o, see Figure 1. We suggest detecting and correcting such errors using the following algorithm.

- 1. Each string obtained from the basic OCR procedure is looked for in a list (dictionary) of expected toponyms, which (if the word is found) provides the semantic information associated with it, such as the type of object (e.g., city, river), its spatial relationships (e.g., administrative unit it is in), and its geographic coordinates if available. This information is verified using different sources of evidence, such as spatial distribution of the letters in the raster image, the coordinates of the letters, etc., as described in Section 4, and the probability of association of the string with the chosen geographic object is obtained.
- In addition, similar strings (e.g., RUSSIA, ASIA, Angola, etc. for RNSoSIA) are looked up in the dictionary and for them, the same information is retrieved and the same check is performed, an additional source of evidence being the probability of the corresponding changes in the letters of the string, as described below.
- 3. The variant with the best score (probability) S_I is considered.
- If this best variant is good enough (S₁≥α, where α is a user-defined threshold), then:
 - 4.1 If the score of the best variant significantly differs from the score of the second best one $(S_1 / S_2 > \beta, \beta)$ is a user-defined threshold) then this variant is accepted and is added to the database together with its associated information.
 - 4.2 Otherwise, human intervention is requested, and the variants are presented to the operator in the order of their scores.
- 5. Otherwise $(S_I < \alpha)$, no correction is applied to the recognized string. It is checked against the linguistic restrictions on the words of a given language, see Section 5.3.

- 5.1 If no anomalies are found, it is considered a new toponym absent in our dictionary. It is added to the database as is and is associated with a nearby object using the algorithm discussed in the previous section.
- 5.2 If an anomaly is found, the string is considered not recognized and human intervention is requested.
- 6. After all strings in the map are recognized, global check is performed, see Section 4.4. If this check fails, human intervention can be requested. Alternatively, the process of error correction can be repeated for this string, and then the global verification for the objects involved in the resulting changes.

As specified in Step 2, additional sources of evidence are taken into consideration when substituting a string for another similar string. Below we consider each of them.

Combination of different sources of information and not just finding the string or its spelling variant in the dictionary is important. For example, geographic information can be used to filter out the candidates that are very close to their spelling to the original string returned by the basic OCR procedure but are not located in the area in question. For instance, let the OCR procedure returned the string Xalapa in the area of Mexican State of Oaxaca. Such a string indeed exists in the list of Mexican cities, but the corresponding city is in the state of Veracruz. On the other hand, there is a city Jalapa precisely in the state of Oaxaca. Thus, it should be considered more probable that the string Xalapa was a result of a recognition error and that the correct string is a similar string Jalapa.

5.1 TEXTUAL INFORMATION

We suppose that there is available a list (dictionary) D of toponyms that can be found in a map. The list can contain much more toponyms than the map in hand—for example, all cities of the country, all seas of the world, etc. Such a list can be compiled as a combination of different sources such as governmental statistical databases, police databases, analysis of newspapers available in the Internet, etc.

For a given string s, e.g., RNSoSIA, a set of all strings similar to s in the dictionary D can be constructed [10]. A string s' is called similar to a string s if it differs from s in at most a certain number of the following disturbances:

- Substitution of a letter for another letter,
- Omission of a letter,
- Insertion of a letter.

With each such disturbance, a probability can be associated; in case of several disturbances, the

corresponding probabilities are multiplied to obtain the overall probability of that s (RNSoSIA) has been obtained from s' (say, RUSSIA) by this sequence of errors. For the string itself (s'=s if it is in D), the probability is 1.

The probabilities of the disturbances can depend on the specific letters involved, if this information is available. For instance, the probability of substitution of I for J is higher than W for L. Similarly the probability of omission of I is higher than that of M. In a cartographic map, the probability of insertion of o is high because of the notation for cities.

The iterative procedure described in Section 4.5 can be used to automatically adjust the model to the specific map. If the map is large or has some standard type and quality, the model can be trained by means of processing a part of the same map or another map of similar quality and manually verifying the results.

5.2 SPATIAL LETTER DISTRIBUTION INFORMATION

The distance between adjacent letters gives information on the probability of insertion or deletion type error. Deletion-type error (a letter is to be inserted to obtain a valid word) is highly probable if the distance between two neighboring letters is about twice larger than the average distance between the letters in the string (it can be the space between different words too). Similarly, insertion-type error (a letter is to be deleted from the string to obtain a valid word) is highly probable if the mean distance between the letter in question and its neighboring letters is about twice smaller than the average. Note that in these cases the corresponding correction of the word is not only acceptable but also required: the score of a string with this type of defects is decreased.

5.3 LINGUISTIC EVIDENCE

The checks described in this section are applied only to the strings not found in the dictionary for which the dictionary-based correction failed (no suitable similar string is found in the dictionary), see the Step 5 of the algorithm from Section 5. In this case, general properties of the given language can be used to detect (though not correct) a possible recognition error.

One of simple but efficient techniques of such verification is bigram (or trigram) control [11]. In many languages, not any pair (or triplet) of letters can appear in adjacent positions of a valid word. For example, in Spanish no consonant except r and l can be repeated; after q no other letter than u can appear, etc. The statistics of such bigrams (or trigrams) is easy to learn from a large corpus of texts. The multiplication of the bigram frequencies for each adjacent pair of letters in the word (and similarly for trigrams) gives a measure of its well-formedness, which can be compared with a user-defined threshold; if a

bigram not used at all in the given language appears, the word is immediately marked as probably incorrect.

Other properties of words specific to a given language can be verified; e.g., in Japanese all syllables are open. If a recognized string for which no variants of correction by the dictionary are found does not pass any of the linguistic filters, it is presented to the human operator for possible correction. Note that since toponyms are frequently words of another language or proper names of foreign origin, linguistic verification can produce a large number of false alarms.

6 CONCLUSION AND FUTURE WORK

We have shown that the problem of recognition of inscriptions in the map, assigning them as names to specific objects (e.g., cities), and importing—using these names as keys—properties of these objects (e.g., population) from existing databases involves both traditional techniques of image recognition and methods specific for cartographic map processing. Our algorithm combines various sources of evidence, including geographic coordinates and object inclusion hierarchy, to choose the best candidate for error detection and correction. (In this work we focused on maps with texts. There are many maps with numerical labels—elevations, geographical coordinates, and so on. See [1], [7] for discussion on this type of maps.)

One obvious line of future development is refining the heuristics used in the discussed sources of evidence and adding new sources of evidence. For example, the basic recognition procedure can return the probability (the degree of certainness) of each letter in the string, or even a list of possible letters at the given position in the string along with their respective probabilities. The idea is that if the basic recognition procedure is certain that the letter in question is exactly the one it recognized (as opposed to just looking like this), the letter should not be changed in error correction, and vice versa.

Another issue possibly to be addressed in the future is the computational complexity, especially that of the method used to compute the integral in Section 4.2.

However, the most important line of future research are improvements to the automatic training of the statistical models, automatic learning of the notational information, and automatic determination of the parameters used in various heuristics of our method.

na de Mercolo de Alango qui le la la contra le la contra l Mina de la calabata de Alango qui la contra la con

REFERENCES

- [1] S. Levachkine, A. Velázquez, V. Alexandrov and M. Kharinov, Semantic analysis and recognition of raster-scanned color cartographic images, *Lecture Notes in Computer Science*, *Vol. 2390*, 2002, 178-189.
- [2] D. Doermann, An introduction to vectorization and segmentation, *Lecture Notes in Computer Science*, 1389, 1998, 1-8.
- [3] G. Nagy, Twenty years of document image analysis in PAMI, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 2000, 38-62.
- [4] A. Gancsan, Integration of surveying and cadastral GIS: From field-to-fabric & land records-to-fabric, *Proc 22nd ESRI User Conference*, 7-12 July 2002, Redlands CA http://gis.esri.com/library/userconf/proc02/abstracts/a0868.html
- [5] L.A. Fletcher and R. Kasturi, A robust algorithm for text string separation from mixed text/graphics images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(6), 1988, 910-918.
- [6] C.L. Tan and P.O. Ng, Text extraction using pyramid, *Pattern Recognition*, 31(1), 1998, 63-72.
- [7] A. Velázquez, Localización, recuperación e identificación de la capa de caracteres contenida en los planos cartográficos. *Ph.D. Thesis*. CIC-IPN. Mexico City, 2002 (in Spanish).
- [8] A. Velázquez and S. Levachkine, Text/graphics separation and recognition in raster-scanned color cartographic maps, *Proc.* 5th IAPR International Workshop on Graphics Recognition (GREC 2003), 30-31 July 2003, Barcelona, Catalonia, Spain, 2003, 92-103
- [9] R. Cao and C.L. Tam, Text/Graphics separation in maps, Lecture Notes in Computer Science, 2390, 2002, 168-177.
- [10] A. Gelbukh, Syntactic disambiguation with weighted extended subcategorization frames, *Proc. Pacific Association for Computational Linguistics (PACLING 1999)*, 25-28 August 1999, Canada, 1999, 244–249.
- [11] A. Gelbukh. A data structure for prefix search under access locality requirements and its application to spelling correction, *J. Computación y Sistemas*, 2003.
- [12] R.C. Angell, G.E.Freund and P.Willett, Automatic spelling correction using a trigram similarity measure, *Information Processing & Management*, 19(4), 1983, 255-261.
- [13] G. Hirst, A. Budanitsky. Correcting real-word spelling errors by restoring lexical cohesion, *Computational Linguistics* (to appear).
- [14] S. Levachkine, Raster to vector conversion of color cartographic maps for analytical GIS, *Proc. 5th IAPR International Workshop on Graphics Recognition (GREC 2003)*, 30-31 July 2003, Barcelona, Catalonia, Spain, 77-91 (2003).
- [15] A. Gelbukh and S. Levachkine, Resolving ambiguities in toponym recognition in raster-scanned cartographic maps. *Proc.* 5th IAPR International Workshop on Graphics Recognition (GREC 2003), July 30-31, 2003, Barcelona, Catalonia, Spain (2003) 104-112.

ENHANCEMENT OF DIGITAL COLOR HALFTONING PRINTED IMAGES

* Laboratory of Image Processing and Pattern Recognition,
Computing Research Center, National Polytechnic Institute, MEXICO
Institute of Cybernetics, Mathematics and Physics, Havana, CUBA

** Mexican Petroleum Institute

* edgardo@cic.ipn.mx, ** mramos@imp.mx

ABSTRACT

This paper presents an approach for enhancing digital color images printed in halftoning. Due to the repetitive nature of the process for image digitizing, and the common structure of patterns used to create halftoning in color images, it has been observed that the alterations of pixel colors, in general, are also repetitive. Based on the frequency of occurrence of pixels with a given color, and on the values of their RGB components, we have defined eight heuristic rules. We consider 'artifacts' those pixels with particularly altered values of their components having lesser frequency of occurrence. These rules then define as valid pixels those having a high frequency of occurrence, and as 'artifacts' those with low frequency of occurrence. The color of less frequent artifacts is then changed to some color of any more frequent valid pixel. This approach, based also on user defined essential colors, generates images that are notably enhanced. The method, which we call Conditioned Selective Enhancement (CSE), has been applied with good results to the problem of enhancing several types of digital color images printed in halftoning.

KEY WORDS

Color Images Enhancement, Raster to Vector Conversion, Halftone Images, Color Halftone Image Digitizing, Artifacts Elimination.

1. INTRODUCTION

Color images are printed in halftone by using subtractive mixing CMY(K) of colors. To extract automatically the information of interest from these images, it is necessary to transform them to a digital format. When digitized, for example, with a conventional color flatbed scanner, numerous artifacts appear due to alterations to the RGB components of each pixel. The net result is that much of the original color of pixels is altered appreciably. If artifacts are not eliminated, it becomes difficult to analyze an image by computer, making it practically impossible to

use the color to recognize and interpret color features and patterns in the image.

When digitizing color maps, the problem acquires a more serious connotation, because of the great variability of graphic information commonly incorporated into it. The main characteristics of digital color maps obtained in such form are described in [1] [2] [3] [4].

The use of conventional filters to clean uniform and impulsive noise in digital images is not effective in this case because of the increase in the number of colors, changes in the original colors, and loss of color information.

The main purpose of this work is to generate a final image with a minimum of artifacts, while preserving the basic structure of graphic patterns and retaining final colors same that are as close as possible to the original ones.

The experimental focus of this paper is the enhancement of digital color images of maps printed in color halftone, thereby allowing the automatic extraction of the color information contained in it, for subsequent incorporation into a Geographical Information System (GIS). The successful solution of this problem promises significant savings of money and time because the graphical information extracted from maps prepared a long time ago can be fused in a straightforward manner with data fed into the design of a new GIS [4].

The image used for our examples is a common color map of Mexico digitized to 300 dpi in a conventional flatbed scanner [7] from a printed paper on scholar atlas book [8].

THE PROBLEM

Figure 1 shows 11 x 11-pixel subimages of coconut palm and lemon symbols extracted from a map legend, and a 13 x 13-pixel letter 'a' extracted from the map itself. The coconut palm symbol has 33 colors and the lemon 31 colors, including the white background. The letter 'a'

appears with 46 different colors. The dark control pixel on the right bottom corner does not affect in the analysis.

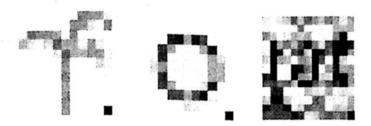


Figure 1. Symbols from a map of Mexico.

Figure 2 shows the same graphical symbols, but extracted now from the map itself (the coconut palm symbols are of size 15 x 15 pixels). The numbers of colors of each symbol are shown in Table 1. The presence of artifacts in all these images is evident.

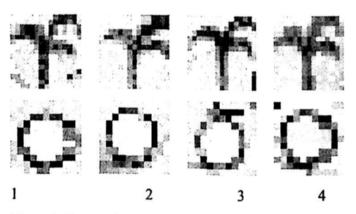


Figure 2. Four versions of the same symbols extracted from the map.

Image	Size (pixels)	1	2	3	4
Coconut Palm	15 x 15	74	73	79	70
Lemon	11 x 11	54	39	41	43

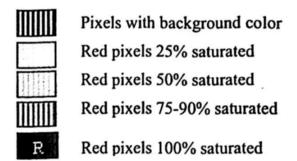
Table 1. Color number in images of Figure 2.

2. SELECTIVE CONDITIONED ENHANCEMENT

In the following analysis, where a halftone red dot R has been used, the following a priori considerations have been stated:

- Colors used for printing in subtractive mix CMY(K) are saturated. A red dot is created superimposing coincidently magenta and yellow dots.
- The register error is considered null
- Dot gain is null.
- The color of substrate is "white" and its distribution is uniform.

During the digitizing process, if sampling occurs exactly on the halftone dot, components of the pixel acquire correspondingly the hue of the dot, with brightness depending on the particular color (Figure 3). The remaining pixels acquire a color that depends on the percent of the samples over the halftone red dot and on the background color.



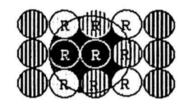


Figure 3. Sampling of a halftone red do t with arbitrary background color.

If the background color is white (Figure 4), which contains all colors, the corresponding color components (magenta and yellow) are relatively reduced, and the value of the other component (cyan) increases. As result, the dot color is less saturated, with the same hue, but brighter. If the pixel acquires less color from the red dot, its color appears whiter.

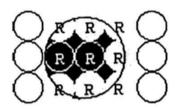


Figure 4. Sampling of a halftone red do t over a white background.

If the background color is black (Figure 5), which makes no color contribution, the corresponding color components (magenta and yellow) of the dot are also reduced, and the contribution to the other component is null. As result of this, the dot color acquires a shadowed color, with the same hue, but less saturated and less bright. If the pixel acquires less color from the red dot, its color appears blacker.

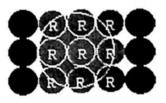


Figure 5. Sampling of a halftone red dot over a black background.

If the background color is other than white or black, and the color of a halftone dot is arbitrary with the area between dots is also being different, the variety of colors acquired by color components of pixels is higher. That is the reason why when digitizing a color halftone image pixels with many different colors appear, thus constituting 'artifact' that must be eliminated. When the frequency of sampling increases, the rate of 'artifact' pixels with respect to the total number of pixels in the image tends to diminish.

We consider as 'artifacts' those pixels with particular values of their components and lesser frequency of occurrence. A way to eliminate these pixels is to associate them with pixels of "legitimate" colors, whose components exhibit the characteristics of a set of eight heuristic rules detailed in Section 5. In this way, these pixels take the colors that initially should have been acquired by them when the image was digitized.

3. ESSENTIAL COLORS

In our approach, essential colors are those colors that will prevail in the final enhanced color image over the abnormal colors of 'artifact' pixels. They are those colors that appear in the legend and in backgrounds of the color map being enhanced. In spite of this, the user has the option to select the essential colors from the image he desires to enhance. Colors of 'artifacts' in the enhanced image are transformed to these essential colors, depending on the outcome of applying the 8 heuristic rules.

4. COLOR SAFE PALETTE

The Color Safe Palette (CSP) is a color palette that handles a limited number of colors. It has 216 colors, composed of only six possible R(ed), G(reen), and B(lue) values of each pixel in the image [6]. Possible values of each component are: 0, 51, 102, 153, 204 and 255. The resulting 216 colors are those obtained from the combination of these values.

When images are enhanced, the colors of pixel in the digital image are grouped into these 216 colors as a first alternative. In this task, three options are possible: Up, Down, and Middle. This means that the component values of each pixel are shifted to the next value in these three directions, that is, when Up (U), values are shifted to the immediately superior level, Down (D) to the immediately inferior level, and Middle (M), values are shifted Up and Down from the intermediate values of the CSP.

The Middle option is preferable in general because the alteration of transformed colors tends to be less than in the other two solutions, resulting in a higher subjective better quality. The grouping rules for the preceding approach are shown in Table 2.

5. BASIC PRINCIPLES TO DEVELOP THE HEURISTIC

The heuristic in the approach proposed was based on four basic principles:

- 1. Reduction of the final number of colors.
- 2. 'Artifacts' were considered those pixels having nonessential colors and lesser frequency of occurrence.
- Due to the repetitive nature of both digitizing and color halftone printing processes, it was assumed that the enhancing process could be governed by strict rules created from experience and observation.
- Color images do not have gray-level pixels, except white and black.

Color Group	Range of original component colors	Value of final component
1	$0 \le R, G, B \le 25$	0
2	$26 \le R, G, B \le 76$	51
3	$77 \le R, G, B \le 127$	102
4	$128 \le R, G, B \le 178$	153
5	$179 \le R, G, B \le 229$	204
6	$230 \le R, G, B \le 255$	255

Table 2. Grouping of image colors to the CSP colors (option M).

6. HEURISTIC RULES

Heuristic rules of the CSE approach are detailed in Table 3.

	Mutual		Numeric	examples
Rule number	Rule relations of Condition		Pixels with RGB components	will change their components to
1	3D	2D' = 2D 1D' = 1D ± 51	255, 102, 153	255, 51, 153
II	(2I, 1D)	2I' = 2I; 1D' = 1D ± 51	153, 0, 153 153, 102, 153	153, 51, 153 (Excluding gray levels)
III	3I; 3D; (2I, 1D); (1I, 2D)	I, D ≤ 153	102, 51, 51 0, 51, 102) 153, 153, 153	0, 0, 0
IV	(2I, 1D)	$2I_e' = 2I$ $1D_e' = 1D$ ± 51	153, 102, 102	204, 102, 102 (Essential color)
v	(2I, 1D)	2I' = 2I ± 5I 1D' = 1D	102, 204, 204	102, 153, 153
VI	(2I, 1D)	2I' = 2I ± 51 1D' = 1D ±51	51, 102, 102	102, 153, 153 (Always + or always -)
VII	3D	$1D' = 1D$ $1D_a' =$ $1D_b$ $1D_b' =$ $1D_a$	204, 51, 102	204, 102, 51
VIII	(2I, 1D)	2I' = 1D 1D' = 2I	153, 102, 102	102, 153, 153

Table 3. Heuristic Rules.

Conventions for RGB components

C: Component of non-essential color;

Ce: Component of essential color;

Ca: Component of non-essential color A;

C_b: Component of non-essential color B:

Initially. I: Equal; D: Different; Finally. I': Equal; D': Different.

Examples. 2I: two equal initial components;

1D_c': one different final component of essential color.

2I = 2I': it means that two equal initial components finally remain the same.

7. JUSTIFICATION OF THE ORDER OF RULES

The order given to heuristic rules is based on the following principles:

- In an ordered numeric list, elements with different components are more frequent than with any other combination.
- It is more likely that changes in component values occur in adjacent levels than in more distant levels.
- It is more likely that at the same time a single element varies than two or more elements.
- It is more likely that two (or more) elements change their component values in the same direction than in opposite direction.

8. ALGORITHM

```
Briefly, the algorithm is as follows:
Input: digital color halftone image, in BMP format.
Output: conditionally enhanced image, in BMP format.
 Input of digital image
 Determining image size
 Conversion of colors to the Color Safe Palette (CSP)
 Selection of essential colors from the image
       Sequential scanning of image
       Control of RGB component of each pixel
        Ordering in a list from most to less frequent, pixel
        colors according to their frequency of occurrence,
     For the whole image, up there is not new grouping,
     execute the X rule, for 1 \le X \le 8
         Grouping of pixels with rule X
         Reordering the list from more to less frequent
         Substitution in image color of pixels with new
         color defined by the rule X
```

9. RESULTS

The following results were obtained by using the algorithm described in the previous section. In all cases, only essential colors remained in the final image.

In order to compare results with the enhanced images in Figure 2, Figures 6-8 show the enhanced coconut palm image, but now including text. Similarly, Figures 9-11 show the enhanced lemon images.

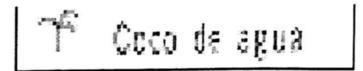


Figure 6. Original coconut palm extracted from the legend of map (63 colors).

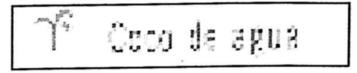


Figure 7. Image showing the coconut palm with CSP colors (11 colors).

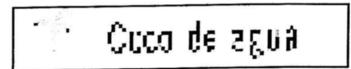


Figure 8. Enhanced coconut palm image (6 colors).

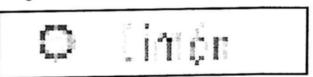


Figure 9. Original lemon extracted from the legend of map (58 colors.)

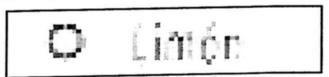


Figure 10. Image shown with CSP colors (21 colors).



Figure 11. Enhanced lemon image (7 colors).

Figure 12 shows an 11x11 pixel image of the third letter 'a' from the text segment "Guanajuato" in the map. (a) Original, with 44 colors (b) CSP image with 20 colors, and (c) Enhanced image with only 4 (essential) colors.

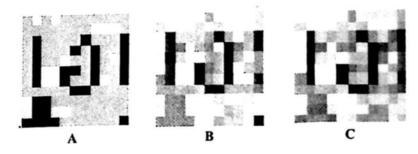


Figure 12. Image of third letter 'a' from the text Guanajuato in the map: A, original, B, with CSP colors, C, selectively enhanced.

Figure 13 shows square images from the map of Mexico (with 251 001 pixels). They have 8 bit/pixel (256 colors): A, original; B, with CSP colors, and C, image selectively enhanced.



Figure 13. A. Original; B. With CSP colors; C. Image selectively enhanced.

Table 4 shows the total number of colors, number of essential colors, number of non-essential colors, and the percent of essential and non-essential colors for each image in Figure 13.

Image →	A	В	C
Total number of colors	256	42	11
Number of essential colors	0	11	11
Number of non-essential colors	256	30	0
% Of non-essential colors	100	71.4	0
% Of essential colors	0	28.6	100

Table 4. Total number of colors in each image and percent of essential and non-essential colors related to that number.

Table 5 shows the total number of pixels in the enhanced image of Figure 13C, the number of ungrouped pixels, and the percent of error, calculated on the basis of the total number of pixel not grouped with respect to the total number of pixel in the image. In both Tables 4 and 5, results are shown in shaded cells.

Image	1
Total number of pixels	251001
Number of pixels not grouped	0
% Of error to group	0

Table 5. Total number of pixels in imag e of Figure 13C and percent of error when grouped.

10. ERROR ANALYSIS OF THE METHOD

The error incurred by our new approach is calculated on the basis of the number of pixels with non-essential colors that were not grouped, with respect to the total number of pixels in the image, for those essential colors predefined by the user in the operation. On this basis, Table 5 shows, that the error in the selectively-enhanced image of Figure 13 is zero. To compare in some measure the enhanced quality of images, a part of the map legend was compared with a system being used to segment alphanumeric characters [12].

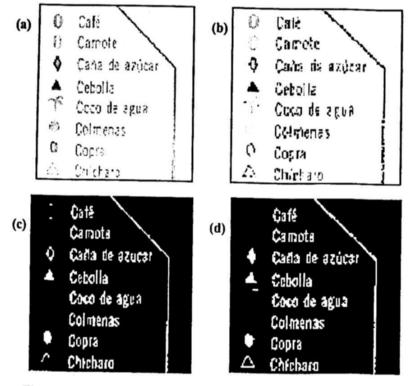


Figure 14. (a) Original Image (b) Enhan ced Image (c) Result after recognizing characters from the or iginal image (d) Result after recognizing characters from the enhanced image.

Table 6 details differences and Table 7 gives the final results. Total of characters properly segmented is $(46 + 11)/60 \times 100 = 95\%$; with the original image the result is (46 + 3)/60 = 81.67%. The general evaluation of results of the method can be seen more clearly in [9] [10].

Characters from the map legend					
Segmentation from the original image Image (c)	Segmentation from the enhanced image (Image (d)	(Image (d) with respect to (c)			
Caté Camote	Çafê Camote	Letter 't' is transformed to 'f'. Accent to letter 'e'. Letter 'a' loss quality			
Caña de azucar	Caña de azúcar	Accent to letter 'u'.			
Cetolla	Cebolla	Split letter 'C'; 'e' is worst. Letter 'b' is completed. Letters 'u' and			
Coco de agua	Coco de agua	second 'a' were completed.			
Colmenas Copra	Colmenas Copra	'm' is completed 'a' is improved.			
Chichard	Chicharo	First 'h' is completed and 'i' is with dot.			

Table 6. Comparing results from segmentation.

String	Character number	Enhanced	Poorer	Not change
Café	4	3 (a, f y é)	0	1 (C)
Camote	6	0	l (a)	5 (c, m, o, t, e)
Caña de azúcar	12	1 (ú)	0	11 (C, a, ñ, a, d, e, a, z, c, a, r)
Cebolla	7	l (b)	2 (C, e)	4 (o, l, l, a)
Coco de agua	10	2 (u, a)	0	8 (C, o, c, o, d, e, a, g)
Colmenas	8	1 (m)	0	7 (C, o, l, e, n, a, s)
Copra	5	l (a)	0	4 (C, o, p, r)
Chicharo	8	2 (h, i)	0	6 (C, c, h, a, r, o)
8	60	11	3	46

Table 7. Final results.

11. ADVANTAGES AND DISADVANTAGES

Advantages of the new approach are the following:

 The method does not require expensive drum scanners to digitize color maps printed in halftone, except when the size of image requires it.

- Maps of very high quality and special conditions of conservation are not required.
- 3. It maintains the 4- and 8 pixels connectivity in point, linear and alphanumeric character symbols.
- 4. It is not required to know a priori any information about the particular shape, size, position, or real colors of graphical patterns in the map, except the colors to be selected as essential by the user.
- The sampling frequency of the digitizing process, and the halftoning method used for printing are not important.
- It can be used in any color image, even if it was not printed in halftone or captured by a specific type of digitizer.
- 7. Because it is a pixel-grouping method, it can be used in the segmentation of color images [11].
- 8. It does not require previous processing of images with classical 'blind' enhancing methods to previously eliminate uniform and impulsive noise.
- Once the user has selected essential colors, the process requires only one run.

On the other hand, disadvantages of the method are the following:

- It is not applicable to gray-level halftone digital images.
- 2. It causes a weak loss of information when the first grouping is carried out to the CSP colors.
- 3. By the nature of color halftone printing, colors in narrow parts of point, linear, and alphanumeric characters, sometimes are associated with the background color.
- Some changes can take place in the original image colors, but generally these changes have little or no effect on the correct recognition of graphical patterns.

12. CONCLUSION

The algorithm of the conditional selective enhancement method fills a significant need for recognizing with higher reliability graphical patterns in digital color maps printed in halftone. The solution is based on the execution of eight heuristic rules that transform the altered color acquired by pixels when a color image printed in halftone is digitized. The transformation is based on essential colors defined by the user. Once the image has been enhanced, if the colors remaining in the image are only essentials, the total percent of error is zero. The number of pixels that do not undergo color transformation depends in general on the essential colors selected by the user.

ACKNOWLEDGMENTS

The authors of this paper wish to thank the Computing Research Center (CIC), Mexico; Institute of Cybernetics, Mathematics and Physics (ICIMAF), Cuba; General Coordination of Postgraduate Studies and Research (CGPI), Mexico, and National Polytechnic Institute (IPN), Mexico, for their support. Additionally, the authors wish to thank to Dr. Rafael C. Gonzalez, of the University of Tennessee, for his comments and suggestions.

REFERENCES

- [1] S. P. Levachkine E. A. Polchkov. Digitalización automatizada de mapas raster, *Proceedings of 3rd International Symposium CIC'99*, México, D. F., 1999, 257-265.
- [2] S. P. Levachkine and E. A. Polchkov. Cartographic pattern recognition. *Proceedings of International Symposium MICAI'00*. Acapulco, México, 2000, 209-218.
- [3] E. M. Felipe, M. E. Ramos, S. P. Levachkine, Problemas potenciales en la digitalización automática de los mapas cartográficos en colores, *Proceedings of International Congress on Computation CIC '2000*, Mexico D. F., 2000.
- [4] S. Levachkine, A. Velázquez, V. Alexandrov & M. Kharinov, Semantic analysis and recognition of raster-scanned color cartographic images, *Lecture Notes in Computer Science*, Vol. 2390, 2002, 178-189.
- [5] E. M. Felipe, M. E. Ramos, Mejoramiento de imágenes de semitonos en colores digitalizadas, 6th Iberoamerican Workshop on Pattern Recognition TIARP 2001, CIC-IPN, Mexico D. F., 2001, 13-24.

- [6] Lynda Weinman page: http://www.websitetips.com/color/colourchart.html
- [7] HP PrecisionScan LTX. User's Manual.
- [8] Atlas NUTESA (Ed.: Gráfica Internacional, S. A., España, 1979).
- [9] E. Felipe, M. Ramos, Mejoramiento Selectivo Condicionado: una nueva estrategia para el mejoramiento de imágenes digitales en colores impresas en semitonos, Informe Técnico, CIC-IPN, Septiembre 2002. (ISBN 970-18-8365-9).
- [10] E. M. Felipe, M. E. Ramos, Mejoramiento selectivo condicionado: una nueva estrategia para el mejoramiento de imágenes digitales en colores impresas en semitonos, *Proceedings of 7th Iberoamerican Workshop on Pattern Recognition CIARP 2002*, Mexico, D. F., 2002.
- [11] E. M. Felipe, M. E. Ramos, Segmentación de imágenes en colores mediante mejoramiento selectivo condicionado, *Proceedings of 11th International Congress on Computation CIC 2002*, CIC-IPN, Mexico D. F.
- [12] A. Velázquez, S. Levachkine, V. Alexandrov. 2001. Color Image Segmentation Using False Colors and Its Applications to Geo-Images Treatment: Alphanumeric Characters Recognition, IGARSS 2001 International Geoscience and Remote Sensing Symposium, Sydney, Australia July 9-13, 2001, 55. http://www.igarss2001.org/bin/IGARSS_Final_Program.pdf

MATHEMATICAL MORPHOLOGY IN COLOR SPACES APPLIED TOTHE ANALYSIS OF CARTOGRAPHIC IMAGES

Jesús ANGULO, Jean SERRA
Centre de Morphologie Mathematique - Ecole des Mines de Paris
35, rue Saint-Honoré, 77305 Fontainebleau, FRANCE
angulo@cmm.ensmp.fr, serra@cmm.ensmp.fr

ABSTRACT

Automatic analysis of cartographic images is an important task for the development of intelligent geographical information systems. Both geometrical features and color are powerful cues to extract spatial semantic objects. This contribution deals with the use of the various color pieces of information for partitioning color images and for extracting geometrical-color features with mathematical morphology operators.

KEY WORDS

Cartographic images, color image processing, mathematical morphology, connections, watershed segmentation, color top-hat, color gradient

1 INTRODUCTION

Automatic analysis of cartographic images is an important task for the development of intelligent geographical Information Systems [11].

From an image processing viewpoint, the contents of a cartographic color map is typically composed of color regions (each color is associated to a semantic label) and of small structures, such are text, symbols, lines, etc. Therefore, both geometrical features and color are powerful cues to extract spatial semantic objects.

The extraction of these semantic elements from an image can be made manually or supported by a computer system. Many efforts are currently carried out in order to propose satisfactory solutions for the automated interpretation of cartographic images [10, 12].

We identify two main steps: on the one hand, the segmentation of the image in order to define the color regions and in order to extract the text/graphic details, and on the other hand, the character recognition with OCR, symbol identification, color indexation, etc. [21]. In this contribution, we present a method for a full image analysis of cartographic maps based on mathematical morphology operators. The approach deals with the use of the various color pieces of information for hierarchical partitioning the image into homogeneous color regions

and for extracting a binary layer of the geometrical-color details. This powerful information can be the input to the subsequent pattern recognition algorithms which are not considered in this study.

The algorithms are illustrated by means of some examples of color cartographic images, Figure 1.

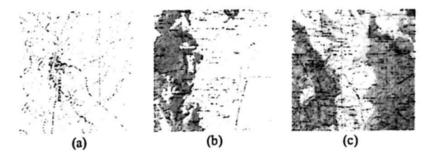


Figure 1. The color cartographic images f used in the examples.

The rest of the paper is structured as follows. First, the choice of a suitable color space for morphological image processing is discussed in Section 2. In Section 3 a reminder of the problem arising from the application of mathematical morphology to color images is included. We continue in Section 4 with a new extension of two morphological operators to color images. Then, in Section 5 are given the algorithms of our approach for the analysis of cartographic images. Finally, conclusions are included in Section 6.

2 COLOR SPACES FOR IMAGE PROCESSING

The choice of a suitable color space representation is still a challenging task in the processing and analysis of color images. The RGB color representation has some drawbacks: components are strongly correlated, lack of human interpretation, non uniformity, etc. A recent study [9] has shown that many color spaces (HLS, HSV,...) having been developed for computer graphic applications, are unsuited to image processing. A convenient representation must yield distances, or norms, and provide independence between chromatic and achromatic components [19]. In our works, we adopt an improved family of HLS systems that satisfy these prerequisites. This family of spaces is named: *Improved HLS* (IHLS).

There are three versions of IHLS: using the norm L_1 , the norm L_2 or the norm $max \rightarrow min$. The equations of transformation between RGB and the new HLS systems are given in [9] [19] and summarized in [2].

For the sake of simplicity, all the examples of the paper were obtained according to the equations: L = 0.212 R + 0.715 G + 0.072 B, $S = \max(R, G, B) - \min(R, G, B)$, $H' = \arccos\left[\frac{R-1/2G-1/2B}{(R^2+G^2+B^2-RG-RB-GB)^{1/2}}\right]$, $\Rightarrow H = 360^{\circ} - H'$ if B > G, H = H' otherwise.

We would like also to compare it with the L*a*b* color space. The principal advantage of the L*a*b* space is its perceptual uniformity. However, the transformation form the RGB to L*a*b* space is done by first transforming to the XYZ space, and then to the L*a*b* space [23]. The XYZ coordinates are depending on the device-specific RGB primaries and on the white point of illuminant. In most of situations, the illumination conditions are unknown and therefore a hypothesis must be made. We propose to choose the most common option: the CIE D₆₅ daylight illuminant.

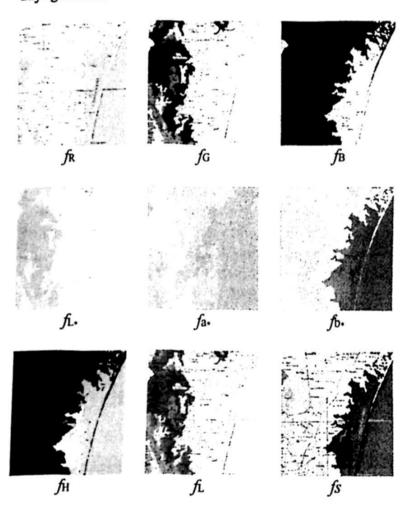


Figure 2. Color components of an image example (Figure 1(b)) in the RGB, the L*a*b* and the Improved HLS color spaces.

Let $f = (f_R, f_G, f_R)$ be a color image, its grey-level components in the improved HLS color space are (f_H, f_L, f_S) and in the L*a*b* color space are (f_L, f_A, f_B) . In Figure 2 are given the different color components of an image example.

3 MATHEMATICAL MORPHOLOGY AND COLOR IMAGES

Mathematical morphology is the application of lattice theory to spatial structures [16]. First introduced as a shape-based tool for binary images, mathematical morphology has become a very powerful non-linear image analysis technique with operators for the segmenting, filtering and feature extraction in grey-scale images. Formally, the definition of morphological operators needs a totally ordered complete lattice structure: there are no pair of points for which the order is uncertain. Therefore, the application of mathematical morphology to color images is difficult due to the vectorial nature of the color data.

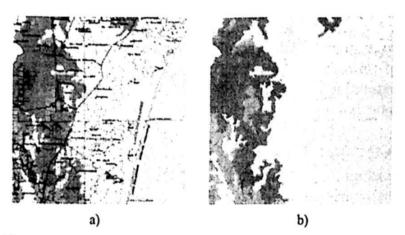


Figure 3. Example of morphological color filtering by vectorial connected operators (f is the image on Figure 1(b)): (a) Opening by reconstruction, $\varphi_{\Omega}^{rec}(f)$, and (b) closing by reconstruction $\varphi_{\Omega}^{rec}(f)$. The structuring element is a square of size _ and the lexicographical order is $\Omega = (L|_{\alpha=10} \rightarrow S \rightarrow H_{h_0} = o^a$.

We have proposed a flexible method for the implementation of morphological vector operators in complete totally ordered lattices by using lexicographical orders which are defined on the RGB color space and on the improved HLS system [3]. In Figure 3 are shown a color opening and a color closing of a cartographic image using one of the lexicographic orders presented in [3].

The inconvenient of the vector approach is the computational complexity of the algorithms which leads to slow implementations. Moreover, different choices must be made in the lexicographical orders: priority of components, degree of influence of the components, etc.

A second drawback, more specific, deals with the processing of the hue component, i.e. with data that are defined on the unit circle [7].

However, in practice, for many applications (e.g. segmentation and feature extraction) the total orders are not required as well as increment based operators (e.g. gradients and top-hats) may be used for the hue component. Hanbury and Serra [7] have developed the application of morphological operators on the unit circle and more precisely, they have defined the circular centered gradient and the circular centered top-hat.

Another recent study [18] proposes a theory where the segmentation of an image is defined as the maximal partition of its space of definition, according to a given criterion. See also [20]. The criterion cannot be arbitrary and permits to maximize the partition if and only if the obtained classes are connected components of some connection (connective criterion). Therefore, the choice of a connection induces specific segmentation. In this paper, we adopt this framework and we investigate different connections for segmenting the cartographic color images.

4 COLOR GRADIENTS AND COLOR TOPHATS

We introduce in this section the extension of the gradient and the top-hat notions to color images.

4.1 GRADIENT

The color gradient function of a color image f at the point x, denoted $\nabla f(x)$, is associated to a measure of color dissimilarity or distance between the point and the set of neighbors at distance one from x, K(x). For our purposes, three definitions of gradient have been used,

- Morphological gradient, ∇f(x): This is the standard morphological (Beucher algorithm [15]) gradient for grey level images (f: E → T, where E is an Euclidean or digital space and T is an ordered set of grey-levels), ∇f(x) = δ_k(f) ε_k(f).
- Circular centered gradient, $\nabla_c a(x)$: If a(x) is a function containing angular values (a: $E \to C$, where C is the unit circle), the circular gradient is calculated by the expression [7], $\nabla_c a(x) = \sqrt{[a(x) \div a(y).y \in K(x)]} \sqrt{[a(x) \div a(y).y \in K(x)]}$ where $a \div a' = |a a'|$ iff $|a a'| \le 90^\circ$ and $a + a' = 180^\circ |a a'|$ iff $|a a'| \le 90^\circ$.
- Euclidean gradient, $\nabla_E f(x)$: Very interesting for vectorial functions ($f(x) = (f_1(x)...f_n(x))$), it is based on computing the Euclidean distance d_E , $\nabla_E f(x) = \sqrt{d_E(x,y).y} \in K(x) \sqrt{d_E(x,y).y} \in K(x)$

We define a series of gradients for a color image:

- 1. Luminance gradient: $\nabla^L f(x) = \nabla f_L(x)$
- 2. Hue circular gradient: $\nabla^H f(x) = \nabla_c f_H(x)$
- 3. Saturation weighing-based color gradient: $\nabla^{S} f(x) = f_{S}(x) \times \nabla_{C} f_{H}(x) + f_{S}^{C}(x) \times \nabla f_{L}(x) \text{ (where } f_{S}^{C} \text{ is the negative of the saturation component);}$

- 4. Supremum-based color gradient; $\nabla^{\sup} f(x) = \sqrt{\left[\nabla_C f_S(x) \cdot \nabla_C f_L(x) \cdot \nabla f_H(x)\right]}$
- 5. Chromatic gradient: $\nabla^c f(x) = \nabla_E (f_{a+} f_b)(x)$
- 6. Perceptual gradient: $\nabla^P f(x) = \nabla_E (f_L f_{a+} f_b)(x)$.

In Figure 4 is depicted a comparative of the gradients of a color image. As we can see, the quality of the gradients is different. We show in Section 5 that the color gradient is a scalar function which can be used with the watershed transformation for segmenting the color images and we discuss how the different gradients perform on the segmentation results.

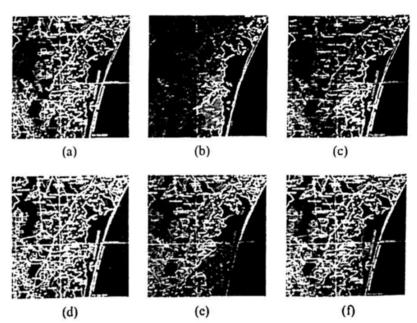


Figure 4. Examples of color gradients (f is the image on Figure 1(b)): (a) $\nabla^L f$, (b) $\nabla^H f$, (c) $\nabla^S f$, (d) $\nabla^{\sup} f$, (e) $\nabla^C f$ and (f) $\nabla^P f$.

4.2 Top-hat

The top-hat transformation is a powerful operator which permits the detection of contrasted objects on non-uniform backgrounds. In the sense of Meyer [13], there are two versions of the top-hat (the residue between a numerical function and an opening or a closing). It involves increments and hence can be defined to circular functions like the hue component. The three definitions of top-hat used in this work are:

- White top-hat, $\rho_B^+(f)$: The residue of the initial image f and an opening $\gamma_B(f)$, i.e. $\rho_B^+(f) = f \gamma_B(f)$ extracts bright structures.
- Black top-hat, $\rho_B^-(f)$: The residue of a closing $\gamma_B(f)$ and the initial image f, i.e. $\rho_B^-(f) = \gamma_B(f) f$, extracts dark structures.
- Circular centered top-hat, ρ_B^a(a): Fast variations of an angular function (defined in the unit circle) [7], i.e.
 ρ_B^a(a) = sup {inf [a(y) a(x).y ∈ B]}

It is possible to define a vectorial top-hat $\rho_B^a(f)$ (white and black) using the Euclidean distance, but this case is not considered in this paper, because we just deal with the separable top-hats in the HLS color representation.

In a similar way than the gradients, and starting from these scalar transformations, we propose three definitions for the top-hat of a color image f:

- 1. White-achromatic top-hat: $\rho_B^{A+}(f) = \rho_B^+(f_L) \vee \rho_B^-(f_S)$ It characterizes the fast variations of bright regions (i.e. positive peaks of luminance) and the fast variations of achromatic regions on saturated background (i.e. unsaturated peaks: black, white and grey on color regions).
- 2. Black-achromatic top-hat: $\rho_B^{A-}(f) = \rho_B^-(f_L) \vee \rho_B^-(f_S)$. Dually, it catches the fast variations of dark regions (i.e. negative peaks of luminance) and the fast variations of achromatic regions on saturated background (i.e. unsaturated peaks: black, white and grey on color regions).
- 3. Chromatic top-hat: $P_B^C(f) = [f_s \times \rho_B^0(f_H)] \vee \rho_B^+(f_S)$. This operator extracts the fast variations of color regions on saturated color background (i.e. saturated color peaks on uniform color regions) and the fast variations of saturated color regions on achromatic (unsaturated) background (i.e. saturated color peaks on achromatic regions).

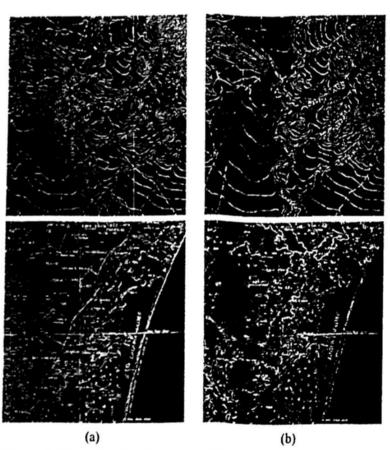


Figure 5. Examples of color top-hats (f are the images on Figure 1(a)-(b)): (a) Black-achromatic top-hat, $\rho_B^{A-}(f)$ and (b) chromatic top-hat, $\rho_B^{C-}(f)$. The structuring element B is a square of size 3.

Figure 5 shows the top-hat of two color images (the white-achromatic top-hat is not meaningful for these examples). We observe that the extracted objects can be different and on the other hand, a certain kind of structures are better defined on one top-hat than on the other. Their contributions are consequently complementary.

Usually, the top-hat is accompanied by a thesholding operation, in order to binarise the extracted structures. We present in Section 5 an interesting method for making the thresholding operation easier.

5 MORPHOLOGICAL ANALYSIS OF COLOR CARTOGRAPHIC IMAGES

The morphological approach for analyzing the color cartographic images is summarized in the overview of Figure 6. The further details of these steps are discussed below.

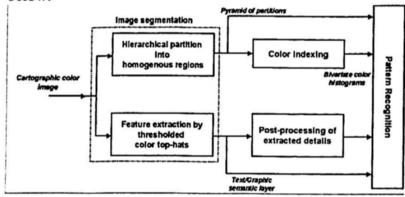


Figure 6. Overview to the proposed approach.

5.1 HIERARCHICAL PARTITION INTO HOMOGENEOUS REGIONS

The aim of the first step is the partitioning of the image into disjoint regions whose contents are homogeneous in color, texture, etc. In order to have a more flexible and rich approach, we propose to use a multiscale segmentation, that is, the partitioning is composed of a hierarchical pyramid with successive levels more and more simplified. In [4] we have introduced two algorithms for hierarchical color image segmentation. The first one is based on a non-parametric pyramid of watersheds, comparing different color gradients. The second segmentation algorithm relies on the merging of chromatic-achromatic partitions ordered by the saturation component. Several connections (jump connection, flat zones and quasi-flat zones) are used as connective criteria for the partitions. Both approaches involve a color space representation of type HLS, where the saturation component plays an important role in order to merge the chromatic and the achromatic information during the segmentation procedure. The presented methods are both good and fast. We would like now to evaluate the application of these algorithms to the cartographic images. We describe here the fundamentals of the algorithms and we comment on the preliminary results.

5.1.1 Waterfall algorithm for color images

The watershed transformation, a pathwise connection, is one of the most powerful tools for segmenting images. The watershed lines associate a catchment basin to each minimum of the function [5]. Typically, the function to flood is a gradient function which catches the transitions between the regions. The watershed method is meaningful only for grey tone images (is based on the existence of a total ordering relation in a complete lattice). However, it can be easily used for segmenting color images by defining a scalar gradient function corresponding to the color image. Using the watershed on a grey tone image without any preparation leads to a strong oversegmentation (large number of minima). A well-known method for avoid the over segmentation involves a nonparametric approach which is based on merging the catchment basins of the watershed image belonging to almost homogenous regions; this technique is known as waterfall algorithm [6].

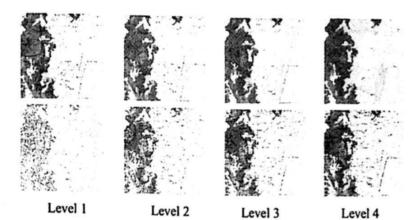


Figure 7. Pyramid of segmentation by waterfall algorithm (f is the image on Figure 1(b)). First row, mosaic images and second row, watershed lines.

Let g be a positive and bounded function $(0 \le g(x) \le M)$ and let W(g) be its watershed. An efficient algorithm for implementing the waterfalls is based on building a new function h: h(x) = g(x) iff $x \in W(g)$ and h(x)=M (h is obviously greater than g) and then, g is reconstructed by geodesic erosions from h [5], i.e. $g = R^{\bullet}(g,h)$. The minima of the resulting function g correspond to the significant markers of the original g, moreover, the watershed transform of g produces the catchment basins associated with these significant markers. In practice, the initial image g is the gradient of the mosaic image m (after a watershed transformation, m is obtained by calculating the average value of the function in each catchment basin). By iterating the procedure described above, a hierarchy of segmentations is obtained. Dealing with color images has the drawback of the method for obtaining the mosaic color image m_i of the level i. We propose to calculate the average values (associated to the catchment basins) in the RGB components, i.e. $m = (m_R, m_G, m_B)$. The gradient of level i+1 is obtained from m_i , i.e. $g_{i+1} = \nabla m_i$. In practice, all the presented gradient functions can be applied on m. It is possible to consider a contradiction the fact that, for the mosaic image, the values are averaged in the RGB

color components and then, the gradients (and consequently the watersheds) are computed using other color components. However, this procedure of data merging allows to obtain good results and on the other hand, the calculation of the mean of angular values (H, a*, b* components) is not trivial.

The example of Figure 7 illustrates the color waterfall technique (using $\nabla^{S} f$), with the different levels of the pyramid. The segmentation results corresponding to the different gradients are given in Figure 8. Other tests have been performed on a representative selection of color images and the results have been similar [4]. The use of only the brightness (∇^L) or only the color $(\nabla^H \text{ and } \nabla^C)$ information produces very poor results. We can observe in Figure 4 that the supremum-based color gradient is the most contrasted and obviously achieves to good results of segmentation. The perceptual gradient, which has very interesting properties for colorimetric measures in perceptually relevant units, leads to better results for the dark regions. However, the best partitions have been obtained with the proposed saturation weighing-based color gradient. The rationale behind this operator is the fact that the chromatic image regions correspond to high values of saturation and the achromatic regions (grey, black or white) have low values in f_S (or high values in f_s^c). According to the expression of ∇^s , for the chromatic regions the priority is given to the transitions of ∇^H and for the achromatic regions the contours of ∇^L are taken.

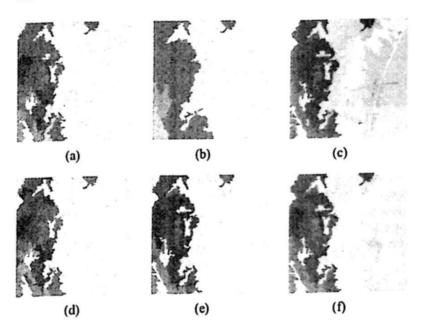


Figure 8. Examples of level 4 of waterfall pyramid using different color gradients (f is the image on Figure 1(b)): (a) $\nabla^L f$, (b) $\nabla^R f$, (c) $\nabla^S f$, (d) $\nabla^S f$, (e) $\nabla^C f$ and (f) $\nabla^P f$.

5.1.2 Ordered partition merging for color images

This segmentation approach is based on the application of other morphological connections to color images. For the sake of simplicity, we just use here the *jump connection* [17].

Let σ_i^{jc} be the jump connection of module k which segments the function obtaining a partition P, i.e. $P\sigma_k^{IC}(f)$. The jump connection is defined for functions $f: E \rightarrow T$ where T is a totally ordered lattice. As for the watershed, the application to color images involves special considerations. In the HLS color systems, the σ_i^{jc} could be applied to each grey level component, obtaining a partition for each component, i.e. $P\sigma_k^{JC}(f_L)$, $P\sigma_k^{JC}(f_S)$, $P\sigma_k^{JC}(f_H)$. Remark that we must fix a color origin for the hue component in order to have a totally ordered set which involves some% disadvantages [7]. Typically, nonsignificant small regions appear in the partitions (oversegmentation). The segmentation may be refined by the classical region growing algorithm, based on merging initial regions according to a similarity measure between them. For the region merging process, each region is defined by the mean of grey levels and the merging criterion is the area a of the region (regions with area smaller than a are merged to the most similar adjacent regions).

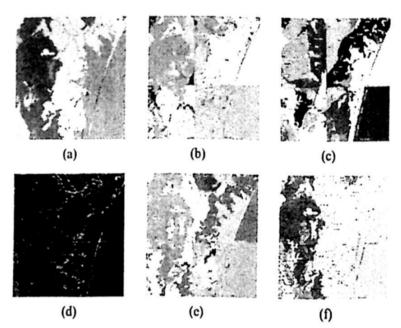


Figure 9. Examples of segmentation by jump connection k = 20 + 1 region merging a < 50, σ , (f is the image on Figure 1(b)): (a) Chromatic partition, $P_{\sigma}(f_L)$, (b) achromatic partition, $P_{\sigma}(f_L)$, (c) saturation partition, $P_{\sigma}(f_S)$, (d) saturation-based weighted contours of chromatic and achromatic partitions, g_S , (e) segmentation of weighted partition by watershed transformation, $P_{\sigma}(f)$, (f) contours of watershed lines on initial color image.

In Figure 9(a)-(c) are shown the partitions by jump connection improved by region merging. Now, the question is how the obtained partitions can be combined. As we can see in the example, the partition $P_{\sigma}(f_H)$ represents well the chromatic regions (very interesting for the cartographic images), as well as $P_{\sigma}(f_S)$ the achromatic ones. We propose the following strategy. Starting from the mosaic image associated to $P_{\sigma}(f_S)$, denoted m_S can use this function in order to weight the contours of the chromatic and achromatic partitions, in a similar way than for the saturation weighing-based color gradient, i.e. $g_S(f)(x) = m_S \times P_{\sigma}(f_H)(x) + m_S^c \times P_{\sigma}(f_H)(x)$, see Figure 9(d). The segmentation of the weighted partition g_S is obtained again by using the watershed transformation, $P_{\sigma}(f_H) = W(g_S)$, Figure 9(e)(f).

The increasing values of parameters k and a lead to a new pyramid of segmentation. The performance of this segmentation is relatively good. The main problem is the adequate choice of values for these parameters.

Therefore, we can conclude that for the aim of partitioning the cartographic image the most indicated method is the waterfall algorithm.

5.2 FEATURE EXTRACTION

The feature extraction involves basically obtaining the color top-hats which extract all the text/graphic details, as we have shown in the precedent discussion, followed by a threshold.

Now, the great difficulty arises from the binarisation of the top-hats in order to generate the semantic layer of the color image.

Besides the problem for finding the optimal threshold value, if we try the thresholding transformation directly on the top-hat image, it is probably that the result will be very noisy, Figure 10(a).

We propose to use an area opening operator. The area opening [22], γ_{λ}^{a} , is a connected filter that removes bright structures whose area is less that a give threshold λ but preserves the contours of the remaining objects.

Taking a size of $\lambda = 50$, it is possible to remove many noisy details. Moreover, the residue between this filtered image and an area opening of large size $(\gamma_{50}^a - \gamma_{large}^a)$ allows us to remove the background contribution, making the threshold easier, see the examples of Figure 10(b) (c).

In fact, this is also a step of segmentation (the threshold method is another connective criterion [18, 20]). We consider that the segmentation of the cartographic image is composed of the partitioning using the waterfall algorithm and the text/graphics detail extraction using the thresholded top-hats.

5.3 COLOR INDEXING AND POST-PROCESSING

Color is an important attribute for image retrieval: color is an intuitive feature for which it is possible to use an effective and compact representation. Color information in an image can be represented by a single 3-D histogram or three separate 1-D histograms. These color representations are invariant under rotation and translation of the image. A suitable normalization also provides scale/size invariance. The histograms are feature vectors which are used as image indices. A distance measure is used in the histogram space to measure the similarity of two images [1].

In [2], we defined two bivariate histograms: h_{HS}^{ρ} : (putting together the hue component and the saturation component) and h_{LS} (luminance and saturation components) associated to the HLS color representation. We presented an algorithm for the partitioning of these histograms using morphological tools which yields another interesting method for segmentation color images. In view of the compact representation of the chromatic-chromatic image information, the bivariate histograms are also very useful for the color distribution indexing of an image.

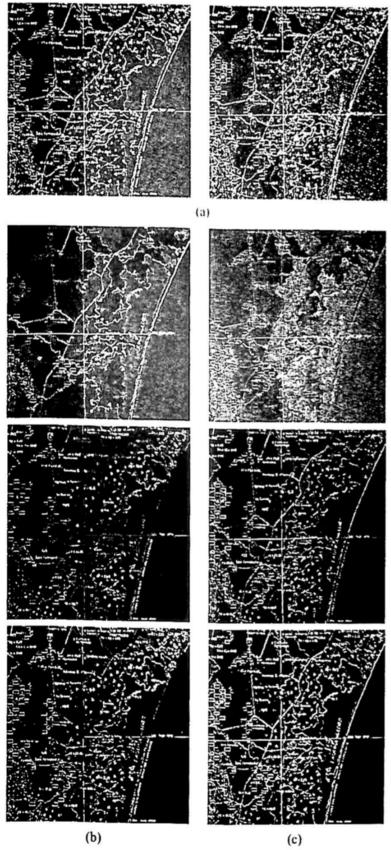


Figure 10. Examples of thresholding simplification by area Opening γ_{λ}^{a} . (a) Initial black-achromatic top-hat and binary image after thresholding at u = 50. (b) and (c) First row: area opening of size $\lambda = 500$ and $\lambda = 40000$, second row: residues between the area opening of size 50 and the corresponding large area opening, third row: binary images after thresholding residues images at u = 50.

In the case of cartographic color images, one of the mosaic color images of the segmentation pyramid (where many non significant color details have been removed) can be used for computing the bivariate histograms, see Figure 11. Then, the histograms could be considered the index for a retrieval system based on the color content.

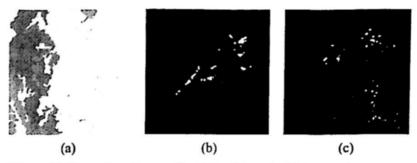


Figure 11. Examples of normalized bivariate color histogram images: (a) Initial image (mosaic of level 4 of the segmentation pyramid), (b) chromatic histogram f_{IIS} , (c) achromatic histogram f_{IIS} .

Finally, there are other morphological operators which can be applied to the binary semantic layer in order to simplify the subsequent steps of pattern recognition: for the separation between text and graphics and for the reading of the text (mainly using an OCR systems). For instance, it is possible to build the skeleton of the binary structures for better outlining the objects (characters, lines, symbols, etc). The most interesting transformation is the morphological thinning [16] which preserves the homotopy and yields robust skeletons, see examples in Figure 12.

The color of the text/graphics is also meaningful and therefore, the simple masked binary layer with the original color image constitutes an interesting information, see Figure 12.

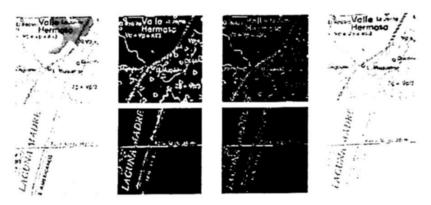


Figure 12. Examples of post-processing of semantic layers (first row, from the black-achromatic top-hat and second row, from the chromatic top-hat; of the image on Figure 1(b)): (a) Initial color image, (b) extracted text/graphic layer, (c) morphological thinning, (d) text/graphic layer masked with original color.

6 CONCLUSION

In this paper, we have proposed a new method for analysis of cartographic color images based on mathematical morphology operators. We have discussed the extension of the gradient and the top-hat notions to color images in the hue/luminance/saturation spaces. These morphological color operators can be used for hierarchical partitioning of images into homogeneous

regions and for details extraction. We also have described a technique for indexing the color distribution and for thresholding easily the extracted details.

We demonstrated on the preliminary image examples that the proposed approach is able to achieve to good segmentation results, providing robust and reproducible algorithms (very few parameters to set). We would like to evaluate deeply the performance of these techniques, comparing with other kinds of non-morphological techniques.

In summary, we believe that this approach can be used as a first step in an automated system for the management of cartographic maps in the framework of geographical information systems.

Acknowledgements

The authors would like to thank Miguel Torres and Serguei Levachkine of the Centre for Computing Research- IPN, Mexico City, for providing the images used in this study.

References

- [1] J. Angulo and J. Serra, Morphological color size distributions for image classification and retrieval, in *Proc. of Advanced Concepts for Intelligent Vision Systems.* (ACIVS'02), Ghent, Belgium, September 2002.
- [2] J. Angulo and J. Serra, Segmentación de imágenes color utilizando histogramas bi-variables en espacios color en coordenadas polares," submitted to *Computación* y Sistemas, May 2003, 10 p.
- [3] J. Angulo and J. Serra, Morphological coding of color images by vector connected filters, in *Proc. 7thIEEE Int. Symp. on Signal Processing and Its Applications (ISSPA'03)*, Paris, France, July 2003.
- [4] J. Angulo and J. Serra, Color segmentation by ordered mergings, in *Proc. IEEE Conf. on Image Processing (ICIP'03)*, Barcelona, Spain, September 2003.
- [5] S. Beucher and F. Meyer, The Morphological Approach to Segmentation: The Watershed Transformation, in (E. Dougherty Ed.), *Mathematical Morphology in Image Processing*, Marcel Dekker, 433–481, 1992.
- [6] S. Beucher, Watershed, hierarchical segmentation and waterfall algorithm, in *Mathematical Morphology and Its Applications to Image Processing, Proc.ISMM'94*, pp. 69-76, Kluwer, 1994.
- [7] A. Hanbury and J. Serra, Morphological operators on the unit circle, *IEEE Transactions on Image Processing*, 10(12): 1842–1850, 2001.
- [8] A. Hanbury, Mathematical morphology in the HLS colour space, in *Proc. 12th BMVC, British Machine Vision Conference*, Manchester, pp. II-451-460, 2001.
- [9] A. Hanbury and J. Serra, A 3D-polar coordinate colour representation suitable for image analysis,

- submitted to Computer Vision and Image Understanding, November 2002, 39 p.
- [10] S. Levachkine, A. Velazquez and V. Alexandrov, Color Image Segmentation Using False Colors and Its Applications to Geo-Images Treatment: Alphanumeric Characters Recognition, in *Proc. of the IEEE International Geoscience and Remote Sensing Symposium,IGARSS'01*, Sydney, Australia, 9-13 July 2001.
- [11] S. Levachkine, A. Velazquez, V. Alexandrov and M. Kharinov, Semantic Analysis and Recognition of Raster-Scanned Color Cartographic Images, in (Blostein and Young-Bin Kwon Eds.), *Lecture Notes in Computer Science*, Vol. 2390, 178–189, Springer-Verlag, 2002.
- [12] S. Levachkine, M. Torres and M. Moreno, Intelligent Segmentation of Color Geo-Images, in *Proc. of the IEEE International Geoscience and Remote Sensing Symposium*, (IGARSS'03), Toulouse, France, 2003.
- [13] F. Meyer, Contrast features extraction, in Cuantitative Analysis of microstructures in Materials Science, Biology and Medicine, Chermant, J.L. Ed., Riederer Verlag, 374–380, 1977.
- [14] F. Ortiz, F. Torres, J. Angulo and S. Puente, Comparative study of vectorial morphological operations in different color spaces, in *Proc. of Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision*, SPIE Vol. 4572, 259–268, Boston, November 2001.
- [15] J.F. Rivest, P. Soille and S. Beucher, Morphological gradients, *Journal of Electronic Imaging*, 2(4):326–336, 1993.
- [16] J. Serra, Image Analysis and Mathematical Morphology. Vol I, and Image Analysis and Mathematical Morphology. Vol II: Theoretical Advances, Academic Press, London, 1982 and 1988.
- [17] J. Serra, Connectivity for sets and functions, Fundamenta Informaticae, 41, 147–186, 2002.
- [18] J. Serra, Connection, Image Segmentation and Filtering, in *Proc. XI International Computing Conference (CIC'02)*, Mexico DF, November 2002.
- [19] J. Serra, Espaces couleur et traitement d'images, CMM-Ecole des Mines de Paris, Internal Note N-34/02/MM, October 2002, 13 p.
- [20] J. Serra, Image Segmentation, in *Proc. IEEE Conf.* on *Image Processing (ICIP'03)*, Barcelona, Spain, September 2003.
- [21] A. Velazquez and S. Levachkine, Text/Graphics separation and recognition in raster-scanned color cartographic maps, in *Proc. of the 5th IAPR International Workshop on Graphics Recognition, GREC'03)*, 92–103, Barcelona, Spain, 2003.
- [22] L. Vincent, Grayscale area openings and closings, their efficient implementation and applications, in (Serra and Salembier Eds.), Mathematical Morphology and its Applications to Signal Processing, Proc. ISMM'93, UPC Publications, 22–27, 1993.
- [23] G.Wysecki and W.S. Stiles, Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd edition. John Wiley & Sons, New-York, 1982.

SYSTEM APPROACH TO A RASTER-TO-VECTOR CONVERSION: From Research to Commercial System

Eugene BODANSKY
Environmental Systems Research Institute, Inc – ESRI 380 New York St., Redlands, CA 92373-810, USA

<u>ebodansky@esri.com</u>

ABSTRACT

This paper analyzes the main factors that influence the effectiveness of raster-to-vector conversion systems. Since complex systems such as these are more than simply the sum of their parts, we emphasize analysis of the whole system.

The paper shows that currently developing new automatic vectorization methods very often cannot increase the effectiveness of conversion systems. The effectiveness of conversion systems depends to a much greater extent on correct division of tasks between the operator and computer and on the capabilities of the raster and vector editors used for pre- and post-processing. Our analysis revealed some important problems that have escaped scientists' attention. Some of them are developing and analysis of semi-automatic vectorization.

KEY WORDS

Data Capturing, Raster-to-Vector Conversion, Vectorization, GIS.

1. INTRODUCTION

Data capturing from maps, engineering drawings, electrical schematics, and other paper line drawings is the combination of several tasks: manual digitizing, scanning of paper line drawings, heads-up digitizing, binarization and pre-processing of the raster images, automatic and semi-automated vectorization of black and white raster images, post-processing of vector data, text and symbol recognition, and interpretation of vector objects. After conversion, data could be loaded into vector databases where objects of line drawings have to be represented as points, lines, and polygons. This task is very labor-intensive and time-consuming.

Much research has been done in the field of data capturing and many articles are dedicated to text and symbol recognition, automatic vectorization, segmentation, compression, and smoothing the result vectors, geometric shapes recognition, measuring of effectiveness of vectorization, and so on.

A number of papers [1-6] have analyzed the data capture process and different raster-to-vector conversion systems.

[1] discusses the data capturing process and briefly mentions heads-up digitizing, automatic (batch mode), and semi-automated (tracing) vectorization, together with the analysis of the pro and contra arguments.

Papers [2] and [3] are dedicated to the specific conversion systems.

In [2] the authors describe the system that was intended for conversion and interpretation of land register maps, which satisfy the Italian Land Register Authority standards. The standards simplify the problems of vectorization and interpretation of raster objects because they define the guidelines for drawings and rules that form a graphic language and restrict the objects shown on these maps.

In this paper, the authors suggest a new algorithm of vectorization that is based on the processing of raster images in RLE format. To simplify vectorization raster object recognition is executed before vectorization.

They suggest that the efficiency of conversion systems or the conversion systems' performance be assessed on the basis of average elapsed time needed to process maps. This suggestion is in line with the opinion of Adrien Litton who writes in [1]: "Your goal is to produce the highest quality vectors in the shortest amount of time."

The general architecture of one commercially available CAD conversion system, GTX, is described in [3]. The authors discuss not only vectorization algorithms but also pre- and post-processing. They emphasize the importance of the operator by showing that automated CAD conversion is never a completely automatic process. At a minimum, a human must check the results of the conversion. In the worst case, the operator must spend hours correcting the revealed errors. In addition to automatic vectorization, the system suggests semi-automated vectorization (tracing).

There is an analysis of the conversion process in [4]. The author writes about the significance of semantic

interpretation of the result vector objects. The last step of the conversion of components is to use one's knowledge about the type of document to assign a semantic label. The result is a description of a document, as a human would give it. The semantic description makes easier for the operator to control the result and to correct revealed errors but usually itself requires the interference of the operator.

The most common opinion is that conversion systems can be made more efficient only by using additional information (gray scale raster images, libraries of used symbols, and so on) and by using more complex algorithms for automation of the vectorization process.

Only [6] emphasizes the problem of correct distribution of tasks between operator and computer. The authors believe that it can make conversion systems more efficient.

For more than a quarter of a century, there has been research and development of the problem of automatic vectorization all over the world. The main goal of these efforts is to accelerate, reduce the costs, and improve the quality of conversion. Nevertheless, manual digitizing from paper maps and scanning, with further heads-up digitizing, has been used up to the present day. In [7], the authors state that automatic vectorization is the easiest and quickest method of data conversion if a source document is in a very good condition. If not, it will require a lot of editing afterwards to reveal and correct all errors. Time spent on editing such a map could easily exceed time spent on document conversion done by some other method. The authors conclude that manual digitizing from analog maps "is the most basic method of digitizing traditional paper maps."

In this paper, we explain why we do not believe that the new methods of automatic vectorization could significantly increase the efficiency of conversion systems and show how to achieve a significant increase in efficiency.

2. VECTORIZATION

We begin our investigation with the problem of vectorization because it is a mandatory task of the conversion process.

Is it possible to strictly define what is the correct vector description of the raster image and what has to be the result of vectorization? Unfortunately the answer has to be "no," if the trivial cases (horizontal and vertical lines of constant thickness) are to be excluded. Strict and objective opinion about this matter does not exist, and only the user can judge whether the result of vectorization is correct. We illustrate this point with several examples.

Example 1. Figures 1a and 1b show the results of vectorizing the same fragment of a contour map. These vectors and width of the corresponding linear objects, which are usually calculated during vectorization, can be used to restore the source raster image. Which result is

better? Is it the result that allows one to restore the raster object? The answer is no.

Any cartographer will tell you that both solutions are incorrect. Contours cannot intersect and intersections in this image can be explained only by scanning error and the noise of the raster image.

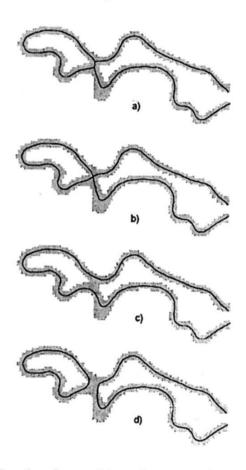


Figure 1. Results of vectorizing a fragment of a contour map.

Correct results are shown on Figure 1c or 1d. To select the correct result, it is necessary to analyze the source line drawing and to study the context. The opinion of the cartographer will be the most important.

Figure [2] shows three different results of vectorizing a fragment of a map. Which one is correct? If we do not know that these are roads, we might think that it's 2a. If we know that these are roads, then the correct result is 2b, if each line represents one side of the road, or 2c, if double lines were used as the line symbol to represent roads.

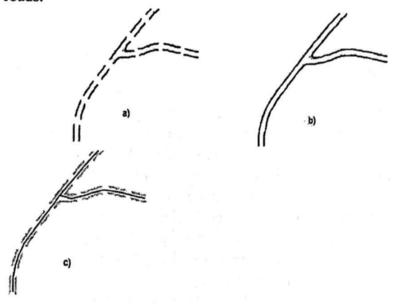


Figure 2. Results of vectorizing a fragment of a road map.

It is possible to find the correct result only if the entire context is taken into account.

Figures 3a and 3b show the results of the vectorizing a fragment of a city map. If the map shows a building, then the correct result of the vectorization is represented by Figure 3a, namely the centerline of the linear object and the border of the building interpreted as a solid. If the map shows two linear objects, the correct result is represented by Figure 3b.

Figure 4 shows a map consisting of several thematic layers. Each thematic layer may require a different vectorization. Texts must be recognized, roads vectorized, buildings outlined. If it is not possible to get hold of the separates, the map should be divided into several sub-maps, each containing one thematic layer, before using automatic vectorization. Dividing a map into thematic layers is complicated, and in the most complex cases, the algorithms that would provide a stable solution in the automatic mode have not yet been found. Sometimes in order to use automatic vectorization methods, the layers are manually copied onto Mylar.

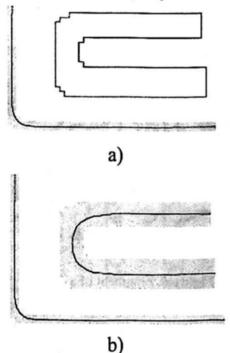


Figure 3. A fragment of a city plan

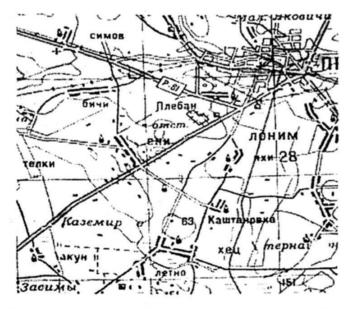


Figure 4. A fragment of a map with several thematic layers

In practice, it is extremely difficult to obtain the correct result by full automatic vectorization of documents. This difficulty stems, in large part, from the fact that the desired solution depends heavily on the following factors:

- noise level
- what linear symbols are used
- quantity and meaning of the thematic layers
- subject area

3. THE EFFICIENCY OF THE CONVERSION SYSTEM

As stated in article [2], efficiency should be measured by the amount of time it takes the operator to convert a document, provided the result satisfies the user's requirements. The cost of this time is equivalent to the cost of labor alone, because if the system is used intensively enough, the cost of the system itself and the cost of the operator's training can be disregarded.

The efficiency of a conversion system depends on what vectorization method is used. If head-up digitizing or manual digitizing is used, vectorization takes up a lot of time, the most of the time of the conversion process. This is why so much emphasis is put on the developing algorithms for automatic vectorization.

There were developed a lot of different methods of automatic vectorization [2, 8, 12, 16,17]. In the references we list only some of the articles dedicated to this problem. Many of them show good results while vectorizing black and white raster images of simple line drawings of good quality. Companies that develop commercial conversion systems also have implemented some good automatic vectorization algorithms.

In addition to errors caused by the imperfections of automatic vectorization algorithms, there are errors caused by noise and other factors listed above. That is why even with good methods of automatic vectorization it is practically impossible to completely avoid errors in automatic vectorization.

Many attempts have been made to compare different systems of automatic vectorization. Some of them offer a quantitative estimation of vectorization errors [8-11], such as an average deviation from the model lines, the number of extra vertices, and the length of incorrectly recognized straight segments, stroke lines, or arcs. It may be useful for the comparison of vectorization algorithms. This estimation is intended for evaluation of the amount of time the operator would have to spend on monitoring and correcting errors. However, this time is affected not only by vectorization errors, but also by the capabilities of the raster and vector editors, by tools that the operator has, as well as by what type of document is being operated on. Here are two simple examples.

Example 1. Because of the noises and errors of the raw vectorization, topological errors may appear. It is difficult to detect them by sight, and eliminating them manually is a difficult task. If a vector editor can

automatically detect and correct such errors, these errors will have little influence on the efficiency of the system.

Example 2. Engineering drawings and electrical schematics contain many straight lines and circle arcs, while contour or hydrographic maps consist mainly of free curves. Without special tools, correcting errors made during vectorization of free curves is more difficult than correcting straight lines.

The fact that manual and head-up digitizing are still being widely used, suggests that frequently the time spent on editing after automatic vectorization is commensurate with the time spent on digitizing.

Our experience with just some algorithms of automatic vectorization leads us to conclude that vectorization of linear objects of constant thickness is quite accurate and in most cases satisfies the user's requirements. The large number of post-vectorization errors and the considerable amount of time necessary for their correction result not from the low geometrical quality of the automatic vectorization but from the complexity of the converted documents and the poor quality of the raster images (see, e.g., figs. 4 and 5).

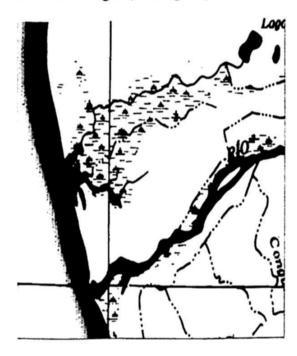


Figure 5. Fragment of a hydrological map

Consequently, in order to increase the efficiency of the conversion system, vectorization methods are needed occupy the middle ground between digitizing and automatic vectorization. Such methods would reduce the number of errors caused by the complexity of the converted documents and their bad quality, as well as mistakes by the operator. Semi-automatic vectorization methods would spare the operator from performing monotonous, labor-intensive routine work and would use the operator only when human judgment is necessary. In addition, these methods would allow one to implement selective vectorization (separate objects or layers) and, in especially complex cases, very easy to switch to heads-up digitizing.

4. INTERACTIVE VECTORIZATION METHODS

Currently, there are two interactive vectorization methods: tracing and raster snapping. Operators who use tracing have to only select a starting point and the direction of the tracing and then the vectorization software will trace the linear object until it comes to an intersection. That is much easier than digitizing each vertex. In addition, it yields computer-generated vectors along the lines, which are usually of higher quality than manual ones [1].

Tracing and raster snapping are indispensable tools for selective digitizing. But they are used for digitizing full documents also.

The fragment of a contour map, shown on Figure6, has more than ten intersections, one of them with six intersecting lines. But contours cannot intersect.

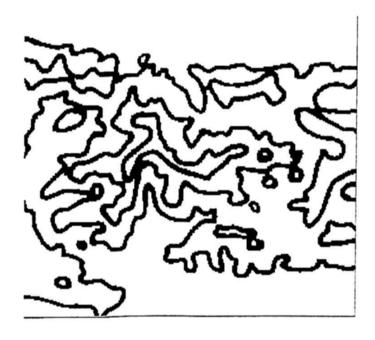


Figure 6. Fragment of a contour map.

What is simpler, to vectorize this fragment automatically and correct all the intersections afterwards, or to use tracing so that the operator could use head up digitizing at the intersections? The answer will depend on the extent of the automatization of the intersection correction procedure in a given system.

Amazingly, it is difficult to find research dedicated to interactive vectorization methods like tracing or raster snapping. If you look for the word "tracing" on the Internet you will receive tons of links to numerous commercial companies, and almost none scientific publications.

Is this so because tracing does not involve any major theoretical problems? No. Some automatic vectorization algorithms essentially perform tracing, that is, they build a vector description of each linear object by moving along this object. The algorithm Sparse Pixel Vectorization is an example of this kind [8]. Nevertheless, in order to achieve a full-fledged, truly interactive vectorization regime, it is first necessary to solve a number of problems that have not yet been adequately addressed. Among these problems are elimination of edge effects; on-line recognition of linear objects, intersections, solids, and

ends; the influence of the starting point on the final solution; smoothing and compression of simultaneously changing lines, and the influence of thresholds.

During tracing it is not important how long it takes vectorization of the full document, but tracing each linear object has to be done in real time. There is no difference if the operator waits for the result 0.1 sec or 0.001 sec. That's why some algorithms and methods that are too time-consuming for automatic vectorization of the whole document, can be successfully used for tracing. It is possible to accelerate tracing with localization of the segment of the image where the vector solution is being built. Segmentation can be performed in a variety of ways, by dividing the image into the tiles of fixed size, dynamic dividing into overlapping areas, selecting parts of the vectorized connected component, etc. However the parts of the image located beyond the borders of the segments can influent on the result vectors. The edge effects have to be suppressed.

Actions of the operator may be required too frequently because of noises. The horizontal line on the Figure 7 has more than ten intersections because of connections with digits and symbols. So tracing that ignores intersections may prove extremely useful. Maybe it is possible to develop the program that automatically builds a straight segment or an arc that starts at a given point and extends in a given direction?



Figure 7. Fragment of a parcel map

Tracing can be used with raster snapping that automatically places a cursor on the nearest point of either centerline, end, intersection, corner, or solid center. This possibility saves a lot of the operator time.

Raster snapping has an independent value when one has to perform vectorization of documents with many straight segments, that contain intensive noise or other layers, making automatic vectorization or tracing too difficult because of the large number of intersections.

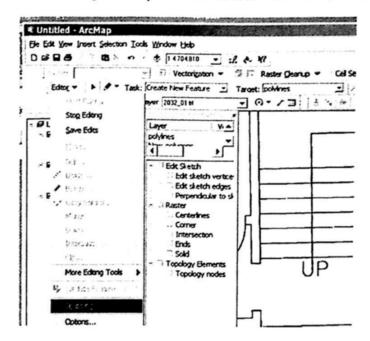


Figure 8. Interface for calling raster snapping in ArcScan

Figure 8 illustrates the procedure for raster snapping in the ArcScan conversion sub-system of ArcGIS. It allows raster snapping on a centerline, nodes intersections, corners (critical points), ends of lines, and centers of solids.

If raster snapping is performed on the ends and corners, then snapping just to two points can vectorize every straight line. At the same time, it is not necessary to zoom in the image to achieve high accuracy.

5. DIVISION OF LABOR BETWEEN OPERATOR AND MACHINE

The efficiency of a conversion system depends largely on how responsibilities are distributed between operator and machine. Humans are peculiarly adept at recognition. They easily identify various thematic layers, isolate and recognize texts and symbols, find critical points, locate the continuation of broken lines, situate the contours of areas filled with symbols, recognize shaded areas and much more, all things that are difficult for a computer program to do automatically.

Let us take the problem of closing gaps as an example [13]. Usually algorithms for solving this problem use two thresholds, search radius and fan angle. However, in some discontinuity points of lines several candidates for continuation can be found in spite of these two thresholds. Even if a more complex algorithm is used or additional parameters are introduced, there is no way to obtain the correct result in all cases. Is it worth it, then, making algorithms more and more complicated? Would not it be better to pass these problems on to the operator and to simplify and stabilize the algorithms instead?

The same issue arises with intersections. Figure 1 showed four possible versions of the solution of one intersection. An operator can easily decide which version is correct, but then it is necessary to redraw the intersection and this is a time-consuming and labor-intensive process if the operator will do it.

Better if the operator's decision can be communicated to the program via templates. Figure9 shows four templates that correspond to the four solutions of this intersection. The program, no longer required to provide a decision, could be simplified and would guarantee obtaining a

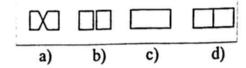


Figure 9. Templates for the intersection in Figure 1.

stable solution. The operator must simply choose the correct template, and thus be spared the lengthy and tedious task of drawing.

Therefore, let us take this into account when designing a conversion system. Let operator do what he can do easily, and let us use the computer and automatic methods to solve those problems that render themselves to formalization and have strict stable solutions. "Render unto Caesar the things that are Caesar's, and to God the things that are God's."

In order to use algorithms of automatic and semiautomatic vectorization, it is often necessary to know the maximum thickness of the linear elements. Frequently this value is not already known. To determine it, one has to measure the line thickness in several places, usually zooming in to an area and pointing the cursor at edge of the line and then the other with the distance tool. This has to be repeated several times.

Can this procedure be simplified? We developed a relatively simple algorithm that will measure line thickness in any point. All the operator has to do is to place cursor next to the line. Figure 10 illustrates how ArcScan performs this procedure. Local thickness of the line equals 12 pixels.

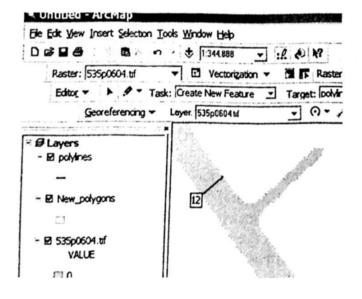


Figure 10. Measuring local thickness

Verification of vectorization results takes a long time. The authors of [14] proposed an algorithm that would identify how much the result of vectorization deviated from centerline. This algorithm can be used to automatically identify spots where deviations exceed

some given value and mark them (see Figure 11). The deviations can result from either the vectorization algorithm itself (a) or from post-processing of the obtained result, such as compression (b) and smoothing (c).

Automatic verification of the accuracy of vectorization result can significantly increase the efficiency of the conversion system.

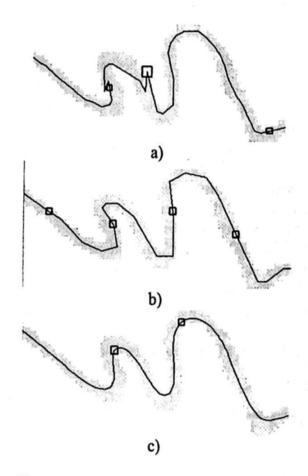


Figure 11. Automatic identification and marking of large local deviations of centerlines

While there are presently a lot of methods and programs for text recognition, there are no effective text recognition programs for graphic documents. Maps, engineering drawings, electrical schematics, and other graphic documents have lots of text on them, text that may touch linear objects or be otherwise difficult to distinguish, automatically, from graphic elements. To recognize text of graphical documents, it is necessary first to separate it from linear objects, solids, and symbols and to define its orientation.

There are no efficient programs that do it automatically. The operator can do it. As proposed in [15], the operator can draw line through the text, marking it as text and revealing its orientation all in one step.

Many conversion systems allow us to clean the raster image of noise before vectorization. To perform this task, it is necessary to find speckles and holes as connected components that meet requirements for size, area, and sometimes more sophisticated characteristics. But small graphic elements (dashes, dots, and others) can be identified as noise too. So it is more efficient to select speckles and holes automatically and highlight them. Then the operator can verify the result of selection and, if

necessary, to correct it by selecting or unselecting connected components before cleaning.

When processing the map of a city, it is often difficult to vectorize rectangular buildings. If it is a relatively small-scale map, contours of buildings can have quite a big noise. The program for an automatic recognition of rectangle buildings with big noise is complex and doesn't always give good and stable results. Manual drawing rectangle contours of arbitrary orientation is time-consuming. But an operator can recognize solids that are rectangular buildings easily and there are simple programs that approximate borders of the solids with rectangles. So it is possible to develop an effective interactive procedure of one click vectorization of rectangular buildings.

One way to increase the effectiveness of conversion systems is by using learning algorithms. The corners between straight-line segments and the boundary points of circle arcs are called critical points. The recognition of critical points is an important component of the conversion process, because critical points will help to correct recognition of geometrical objects. All the algorithms for this task use some thresholds. Often it is difficult to evaluate these thresholds because they depend on so many factors: maximum and minimum curvatures, noise, thickness of lines, and so on. But it is relatively easy to show critical points on the screen (Figure 12). The effectiveness of the conversion system will be increased if an algorithm can be developed that can automatically evaluate necessary thresholds and parameters using information about location of some of the critical points.

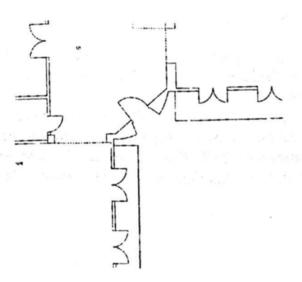


Figure 12. Fragment of architectural design

Sometimes, addition of a simple tool drastically expands a system's capabilities. ArcScan added a seemingly insignificant instrument that allows toggling colors of foreground and background pixels (see Figure 13).

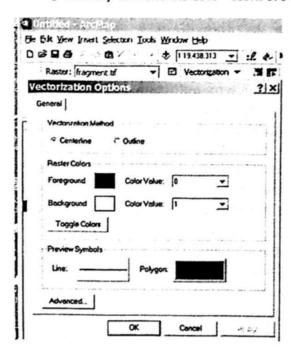


Figure 13. Tool for toggling colors of foreground and background pixels

When combined with semi-automatic vectorization (tracing), however, it allows one to vectorize a non-trivial image, in which linear objects are represented both by solid and double lines (Figure 14).

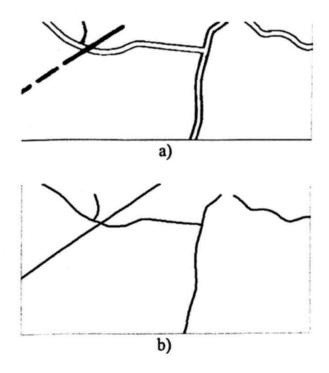


Figure 14. A source raster image (a) and the result of vectorization (b)

6. CONCLUSION

The article analyzes the task of data capturing and the principles of designing conversion systems. It demonstrates with examples how the efficiency of the conversion systems can be greatly increased if the division of labor between operator and machine will be done right, when an operator does what he can do easily, and the computer solves those problems that render

themselves to formalization and have strict stable solutions.

It further highlights the significance of the algorithms and methods of semi-automatic vectorization, which have been largely ignored by the scholarly community. Our approach involves a comprehensive analysis of the system as a whole, rather than looking at it as a mere collection of separate instruments.

The new version of the conversion system, ArcScan, which was developed by ESRI and is a part of ArcGIS, is a prototype of the new generation of conversion systems. In preparing the article, we used ArcScan to illustrate some of our statements, assumptions, and conclusions.

REFERENCES

- [1] GIS Data Conversion: Strategies, Techniques, Management (Ed.: Pat Hohl 1998).
- [2] L.Baotto, V.Consorti, M.Del Buono, S. Di Zenzo, V.Eramo, A.Esposito, F.Melcarne, M.Meucci, A.Morelli, M. Mosciatti, S.Scarci, M.Tucci, An Interpretation System for Land Register Maps, *IEEE Computer*, 25(7), 1992, 25-33.
- [3] A.J.Filipsky, R.Flandrena, Automated Conversion of Engineering Drawing to CAD Form, *Proceedings of the IEEE*, 80(7), 1992, 1195-1209.
- [4] Lawrence O'Gorman, Basic Techniques and Symbol-Level Recognition – An Overview, Lecture Notes in Computer Science Vol. 1072, 1996, 1-12.
- [5] Sergey Ablameyko, Tony Pridmore, Machine Interpretation of Line Drawing Images. Technical Drawings, Maps, and Diagrams (Ed.: Springer-Verlag, 2000).
- [6] Serguei Levachkine, Evgueni Polchkov, Integrated Technique for Automated Digitization of Raster Maps, On-line Journal: Revista Digital Universitaria, 1(1), Art. 4, 2000 (www.revista.unam.mx/vol.1/art4)
- [7] Parcel Mapping Using GIS. A guide to Digital Parcel Map Development for Massachusetts Local Governments. Prepared by University of Massachusetts Office of Geography Information and Analysis for MASSGIS. August 1999. http://umass.edu/tei/ogia/parcelguide

- [8] Dori Dov, Wenyin Liu, Sparse Pixel Vectorization: An Algorithm and Its Performance Evaluation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(3), 1999, 202-215.
- [9] Hori Osamu, Doermann David S, Quantitative Measurement of the Performance of Raster-to-Vector Conversion Algorithms, Lecture Notes in Computer Science, Vol. 1072, 1996, 57 68.
- [10] Jaisimha M.Y., Dori D., Haralick R., A methodology for the characterization of the performance of thinning algorithms, *Proc. of Int. Conf on Document Analysis and Recognition (ICDAR'93)*, Tsukuba, Japan, 1993, 282-286.
- [11] Phillips Ihsan, Liang Jisheng, Chhabra Atul K., Haralic Robert, A Performance Evaluation Protocol for Graphics Recognition Systems, *Lecture Notes in Computer Science Vol. 1389*, 1998, 372-389.
- [12] Dave Elliman, A Really Useful Vectorization Algorithm, Lecture Notes in Computer Science Vol. 1941, 1999, 19-27.
- [13] Eugene Bodansky, Alexander Gribov, Closing Gaps of Discontinuous Lines: A New Criterion for Choosing the Best Prolongation, Lecture Notes in Computer Science Vol. 2423, 2002, 119-122.
- [14] Bodansky Eugene, Pilouk Morakot, Using Local Deviations of Vectorization to Enhance the Performance of Raster-to-Vector Conversion Systems, *International Journal on Document Analysis and Recognition*, No. 3, 2000, 67-72.
- [15] Arvind Ganesan, Integration of Surveying and Cadastral GIS: From Field-to-Fabric & Land Records-to-Fabric, *Proc. ESRI User Conference 2002* (http://gis.esri.com/library/userconf/proc02/abstracts/a0868.html)
- [16] 16. Ogniewicz R.L., Kubler O, Hierarchic Voronoi Skeletons, *Pattern Recognition*, 28(3), 1995, 343-359.
- [17] Desseilligny M.P., Stamon G., Suen Ch.Y., Veinerization: A New Shape Description for Flexoble Sceletonization, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(5), 1998, 505-521.

AUTOMATIC MAP GENERALIZATION

a dirección este la sufficiencia de la co

SPATIAL ANALYSIS AND AGENT PRINCIPLES TO AUTOMATE GENERALISATION PROCESS

Anne RUAS
Laboratoire COGIT - IGN-France
2 avenue Pasteur
94 165 Saint Mandé FRANCE

ABSTRACT

To generalise digital data, some efficient algorithms are required. Recent researches have enriched the library of algorithms allowing interactive generalisation of GIS. The automation of the process requires that the generalisation system is able to find automatically 'where and how to generalise'. The 'where' depends on the capacity of analysis: the system should be able to identify which objects do not respect the specifications. The 'How' depends on the existence of procedural knowledge which are rules that connect conflicts with the algorithms which are supposed to solve such conflicts. The AGENT prototype has been developed to provide the geographical objects with the capacity to qualify themselves (i.e. to recognise automatically the conflicts) and to apply appropriate algorithms to solve the existing conflicts. Moreover, a mechanism of control has been developed in order to backtrack and try another algorithm whenever the results are not as good as required. This paper explains the principles of self-generalisation at different level of details as well as the importance of spatial analysis for generalisation.

KEY WORDS

Generalisation, Agent, Spatial analysis, Constraints.

1. INTRODUCTION

The process of generalisation aims at producing a new representation of geographical space from a more detailed one. It could be compared to the conception of an abstract from a full paper. The abstract should content less words, but the main ideas should be detected, preserved and enlighten. For geographical data some graphical rules must be satisfied to produce a representation that can be read and understood. The objective is to represent the same geographical space with less details while allowing the readers to build a mental representation reflecting the reality. Research on generalisation aims at designing a system able to build an abstract representation for any type of geographical area and any type of level of details. It requires a good understanding of the process of abstraction.

Early research in the eighties focused on the definition of the operations of data modification such as simplification, aggregation, displacement and object removal (see [1]). Then researchers designed algorithms for each of these operations, starting from the simplest ones to the more complex such as displacement [2]. The result of this research can be seen on some GIS packages of generalisation that allow a user to select an algorithm and apply it to one or several objects selected interactively. But the automation would require an orchestration of these algorithms as suggested by Mackaness [3]. Tests such as [4] clearly showed that 'how' to generalise (i.e. what to apply, where, and when) was not a trivial question as no predefined sequence could produce good results with the existing algorithms. The choice of an algorithm and its parameter values depends on the properties of each object or set objects to generalise.

From that point, three approaches have recently been proposed by the research community:

- To design new algorithms that would model generalisation as a constraint solving problem [5]. Constraints of proximity, of size are represented at the co-ordinates level. Classical mathematical methods such as least square one are used to displace in an iterative way the position of each point in order to respect constraints,
- 2. To try a set of combination of operators and to choose the solution that best satisfies a set of constraints. The theoretical number of trials depends on the set of objects and on the possible combination of operators. Simulated annealing technique is used by [6] to reduce the quantity of trials and to choose if not the best at least a good solution,
- 3. To try to replace the cartographer analytical capacity by spatial analysis tools that would guide step by step the process. This method is presented hereafter.

These three methods are all very interesting. The two first are certainly very valuable and efficient for small scale changes whereas the last one can easily include dramatic changes (such as object removal and collapse) but it requires more spatial analysis tools. A first remark is that all of these methods are based on constraints as it was proposed by K. Beard [7].

This paper presents an approach based on agents and constraints to automate the generalisation process. It

begins with a short explanation of the principles of the method before developing the important role of spatial analysis.

2. PRINCIPLES

The model we are proposing is based on constraints, levels of analysis and agents. It has been first proposed by [8] [9] and has been enriched during the European Esprit AGENT project [10]. The model is fully detailed in [11], we will just present hereafter the main principles before focusing on spatial analysis in section 3.

The main global principle of this model is that generalisation is viewed as a state change process. The initial (non-generalised) data are generalised step by step using algorithms that are chosen and triggered dynamically and automatically. The choice depends on two important states: the initial one -which is preserved as the reference- and the current one. The aim of the process is, at each step, to find a good next step. Control mechanisms are part of the process to check dynamically if a 'trial next step' can or not be validated to become the current state. This step by step process follows the human generalisation process.

2.1 THE CONSTRAINTS

Constraints are the representation of 'what should be respected' to obtain a good generalisation. They are formalised by means of functions such as:

 $size(building) > 350m^2$.

We distinguish two kinds of constraints:

- Constraints of generalisation represent the new rules that should be respected according to the final product specifications. They include graphical constraint -such as building size- and constraints related to information reduction. Usually, constraints of generalisation are not initially satisfied and algorithms are used to improve their state.
- Constraints of maintenance constrain the information to 'look like' the initial state in order to produce a final representation that allows the user to build a mental representation close to the reality. For example a generalised entity should have a shape if not equal at least that 'looks like' its non-generalised shape. These constraints are also used to avoid an homogenisation of the representation. For example, after generalisation, the centre of a city should be different from its suburb (if required in the specifications). Constraint of maintenance are initially satisfied but might be violated after applying the algorithms. They are necessary to avoid or to reduce the possible distortions.

Constraints help to know 1/ 'where to generalise' (e.g. this building should be dilated because it is too small: it does

not respect the size constraint on buildings) and 2/ if the proposed solution is acceptable or not (e.g. the solution of building dilation (i.e. size emphasising) should not be validated because the shape of the building changed too much: it does not respect the constraint of shape preservation on building). Of course a large set of constraints (on buildings, on streets, on rivers, etc., related to size, shape, density, etc.) are necessary to perform generalisation and often it is difficult to find one solution that strictly respects all the constraints. The objective is to find a good balance for each object (or set of objects), if possible the best. Whenever a constraint is not satisfied a conflict is created. Some conflicts are more sever than others do. The challenge is to reduce the severity of the conflicts to reach a state where most of the constraints are satisfied.

To facilitate the use of constraints during the process, it has been proposed to associate with each entity (each road, each building) the state of its property towards the function of constraint [8]. To apply an algorithm the system needs to know that this building has a size, which is significantly smaller than the minimum building size. To do so, each entity is described by its initial and current properties (such as size, position, shape, orientation) and each current property (e.g. the current size) is evaluated against the corresponding function of constraint. For example if the current building size of this_building is 280 m², and the building size function is 'size(building) > 350m²' then the state of the size of this_building is bad and requires to be improved by means of a generalisation algorithm such as size_emphasing.

Consequently each entity is described by the state of each of its properties. These states are used to trigger generalisation algorithms and to control that a proposed solution still respects the constraints of maintenance. A building that has been simplified in order to respect its granularity constraint should have a generalised shape that also respects the constraint of shape maintenance (a nearly squared shape should not be transformed into a round or a flat shape). Moreover, as an entity can have a set of non-satisfied constraints, priority of treatment are computed to allow the system to choose which constraint should be solved first. These priorities are computed dynamically according to the nature of the constraint and the severity of the constraint violation. As an example, if a set of buildings is too dense and buildings are too close, it is better to reduce the quantity first (by means of object removal) before trying to displace the buildings as there is not enough room for all of them.

2.2 THE AGENTS

As seen previously, each geographical entity is described by a set of properties. The global state of an entity is the synthesis (or the aggregation) of all its property states. In other words, the state of an entity is computed from the conflicts that exist on each of its properties. Whenever an other words, the state of an entity is computed from the conflicts that exist on each of its properties. Whenever an entity has a bad global state it should be generalised. The idea is to allow each entity to compute its own global state and to trigger itself its own generalisation. To do so, we used the concept of agent.

An agent is an entity that acts (and interacts) by itself. It means that it can give to itself the order to simplify its border to respect its granularity constraint or to emphasise itself to respect its size constraint. In terms of implementation, it means that each agent has an 'engine' that allows it 1/ to compute its state (the state of each of its property and the synthesis), 2/ to choose an algorithm to improve itself and to trigger it 3/ to control its new state, to validate the solution if it has improved or to backtrack and choose another algorithm if it has not. Phase 1 and 3 depend on the state of the properties towards the function of constraints whereas phase 2 depends on procedural knowledge. The schema of the process of an agent is given in figure 1. In our system, each geographic entity becomes an agent, which means that it has an engine that allows its self-generalisation. The term situation is used to describe a geographical entity whatever its level of detail.

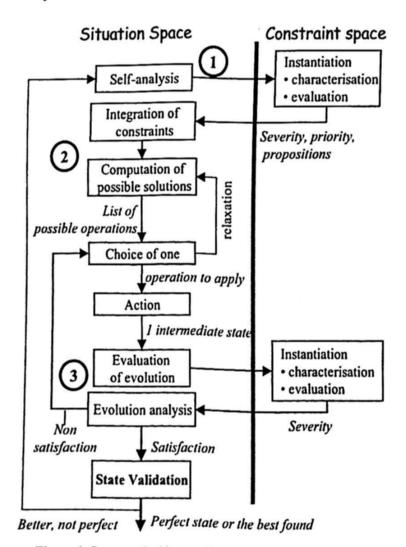


Figure 1: Process of self-generalisation for a single agent [8].

Amongst generalisation operations, some are individual (such as simplification, size emphasising) while others are contextual. Contextual operations are applied to a set of objects. Displacement, object removal, aggregation are contextual operations as the decision of 'where to displace', 'which object can be removed' or 'which objects can be aggregated together' can only be taken

with an overview on a set of objects. To do so, we introduced the concept of meso-agents [8]: A meso agent is a group of agents that act collectively as if it was a single entity. A meso agent is an agent (consequently it generalises itself) composed of agents. To distinguish them, we call micro agent an agent that is not composed of other agents. The difficulty is then to understand which kinds of meso agents should be created for generalisation purposes, and to design methods to detect them as they are not initially represented in the database. Different approaches exist: we can create a meso agent each time a conflict can not be solved at the micro level. We can also create meso agents that correspond to geographical meaning such as an urban block, a district, a town or a street network. Generally these two approaches are rather coherent one to the other. Research done at the COGIT laboratory for generalisation purposes led to the definition of a set of meso agents necessary for the generalisation process and is illustrated in figure 2.

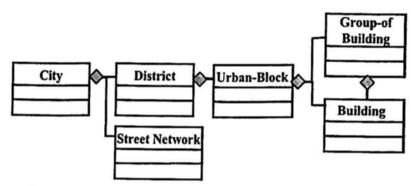


Figure 2: Mcso objects type within a town for generalisation [12].

Amongst meso agents, some, such as urban block are used to trigger generalisation algorithms, while others, like districts, are only used to represent some constraints of maintenance. An urban block agent triggers and controls building removal and building displacement, a city agent triggers and controls street removal, whereas a district agent controls the fact that its semantic property is preserved during the process (its diversity of buildings nature).

To choose the appropriate algorithm to reduce a conflict or a set of conflicts, procedural knowledge [7] is described at the constraint level. Each constraint holds this knowledge described by means of rules such as: if state is very bad then use such or such algorithm (with such parameter value), if state is bad, then use such or such algorithm, otherwise if state is medium then use such or such algorithm. Information related to procedural knowledge is given in [11].

This model of generalisation has been used during the European project AGENT. The team developed a GIS package named Agent on top of Laser-Scan Limited Lamps2 GIS. During the project, only methods related to roads and urban areas have been developed. The project proved that this model is appropriate for generalisation. The use of this package in NMAs already allowed to improve the productivity, reducing the quantity of interactive work. Certainly the best result concerns the

road network generalisation. More important, the system has proved flexible. It has been used for different databases, at different resolutions. On the other side, the evaluation we made also showed that research should be pushed forward not only to introduce new algorithms but also to design more and better spatial analysis tools.

3. THE ROLE OF SPATIAL ANALYSIS FOR GENERALISATION PURPOSE

The core of the generalisation process is based on our capacity to develop (and to master) measures able to describe geographical information as well as possible. These measures replace the analytical capacity of a cartographer who detects that a situation is not correct and who is able to change it while preserving its main properties. Measures are used to detect conflicts as well as the information and properties that should be preserved during the process: Spatial analysis tools are necessary to trigger and to control the process as suggested by [3]. Functionally, we can decompose the set of measures into two parts:

- The first part is used to recognise and create meso entities such as district, groups of building or ring road
- The second is used to describe the properties of the entities such as measures of distance, size, shape, orientation or density.

Besides the fact that spatial analysis tools are very important for a large set of geographic applications such as risk management, the library of measures is not enough identified as a goal in itself, even if GIS designers are proposing spatial analysis tool packages. At least for generalisation, each developer or researcher develops its own appropriate library from available tools on GIS or from a mathematical library such as CGAL. Designing a measure requires three steps:

1/ the identification of the mental concept coming from our perception,

2/ the transformation of this concept into a mathematical formula that can be implemented,

3/ the interpretation of the value of the measure (the measurement) in the context of generalisation purpose.

The last step is certainly the most complex. The interpretation of a value means that we should be able to give a meaning for each value according to a goal. For example, if a road has a fractal dimension equal to 1.32, what does it mean? If it is changed to 1.25, is it an important change? Does it mean that the shape changed dramatically? Is it acceptable? If some measures are trivial, others such as measures of shape are much more complex as there is no mathematical definition for shape. Example of this complexity can be found in the research done to describe the sinuosity of roads [13]. Every one can describe the sinuosity of a road in his own way, but how to describe it mathematically? Often, it is nearly

impossible to reach our human subtlety of description even though the measure is mathematically accurate : measures do rarely describe our mental concepts. However it is necessary to progress in the digital description of geographic structures to enrich our very poor data base - where objects are represented only by means of few attributes and their location - with the explicit representation of information necessary to automate a process (whose objective is to simplify the representation while preserving the geographical meaning). If a cartographer is able to visually detect what is important and what should be preserved and how, it is because he has 1/ knowledge on geographical space 2/ a drawing of the specific arrangements of objects that he can connect to its knowledge to reconstitute a mental image of the reality and 3/ experience. For example the generalisation of a river network requires that 1/ we know in theory what is a river network, what are its properties 2/ we see it globally to identify which parts can be removed without losing the global structure of this specific network, and this according to the spatial relationships that exist between the network and other objects. The automation of such a process requires 1/ to reconstitute the river network as a single entity 2/ to analyse its personal and contextual properties and 3/ to develop an algorithm that will reduce the quantity of rivers without losing the global property of a river network nor the main characteristics of this river network.

The complexity of such analysis is coming from the fact that most of the meso entities useful for the generalisation process are entities that can not be precisely defined. They are phenomena that can be named (such as a town) and that are useful to geographical reasoning. We are close to the Gestalt theory [14], which looks for pregnant forms that emerge from a set of information. Designing methods to detect such phenomena is not easy because no accurate or single definition exist. For example, there are thousands of possible definitions of a town.

To illustrate this difficulty but necessity of analysis let us take the example of building alignments. Building alignments are important on a map because they structure the space. They are the reflect of a certain kind of urban settlement, more specifically of middle class districts. Their maintenance is important during generalisation process because they hold implicitly this geographical (i.e. historical and economical) knowledge. Moreover, generalisation algorithms - such as building displacement, aggregation and removal - tend to distort them. Figure 3 shows an example of AGENT prototype generalisation of an urban block where two alignments are completely lost as they were not detected by the process. At least a 'good generalisation' should have preserved one of both alignments.

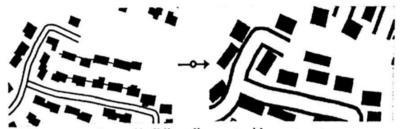


Figure 3: Lost of building alignment with agent prototype.

The detection of such structures would help not only to constrain the displacement of buildings but also to develop a new method of building removal (named typification) adapted to such meso-objects.

Recent research [15] presents a method to detect and characterise this kind of pattern. The alignment in itself is not so difficult to detect by means of projection and clustering techniques (see [15]). The point is to be able to distinguish regularity from irregularity, as, for different reasons, exact regularity never exists. An alignment is perceptually important if and only if it has *some* regularities of arrangement, size, inter-distance, shape and orientation (see figure 4).

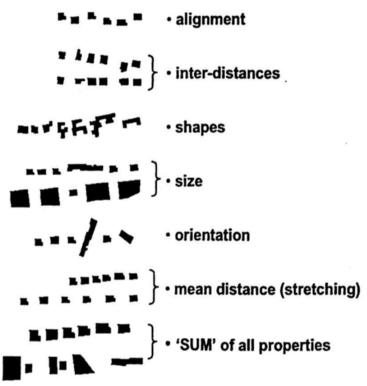


Figure 4: What makes a pattern.

If measures of alignment, distance, building shape, size or orientation, are not too complex, nor their regularity - as the standard deviation can be used to compute regularity-, the difficulty here is to go from digital values to an interpretation necessary to distinguish the main alignments that should be preserved. On going research based on expert knowledge acquisition will soon be published to present our proposed method and results. To sum up, some examples of alignments are proposed to cartographers who mark them. This information is used to find - if possible - the functions between the quantitative values computed by the measure with the qualitative values given by experts. These qualitative values are the result of the knowledge of the cartographers. In such a way their expertise can be introduced in the system. This method allows to compare heterogeneous properties such as alignment, shape and size - and to build an

aggregation function necessary to mark the global quality of each alignment. It then becomes possible to mark them and to select only regular alignments (the smallest mark, the best in figure 5).

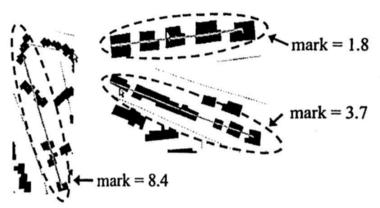


Figure 5: Evaluation of building alignments quality.

The best alignments are introduced as new meso-agents in the generalisation process with appropriate attributes and generalisation methods (Figure 6).

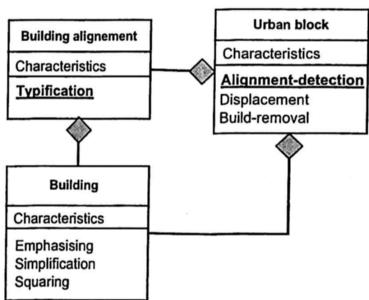


Figure 6: Adding Building alignments to the data model.

As a global result, the fact that we are now able to detect automatically these structures improves the Agent generalisation results: Each *urban block* meso agent has the task to detect and create these new meso agent that would be themselves responsible for their own generalisation: an *alignment* will apply to itself typification to reduce its number of houses while preserving the pattern, and the meso *urban block* will perform displacement while preserving the relative position of buildings (see figure 6).

Another aspect to enlighten is that the progress made in generalisation induced a real change in the data modelling as illustrated in figures 2 and 6. New classes of objects associated with their methods of creation and their characterisation (i.e. the computation of their properties) have been introduced in the initial modelling. This effort of formalisation tends to fill the gap between the abundance of the information of the real world and the weakness of the information contained in classical databases where only single object location and nature are described.

databases where only single object location and nature are described.

The detection and characterisation of building alignments illustrates the use of spatial analysis for generalisation purpose. Of course, lots of progress are still necessary as much more information should be explicitly represented to allow the computer to preserve the main information implicitly contained in the initial database.

4. CONCLUSION

In this paper we have presented the principles of one model of generalisation that has already been implemented by the AGENT team on the Lamps2 package Agent. This model, as opposed to others such as [5] and [6], has the advantage of pushing the explicit representation of the information that governs the process. This new information can be either a better description of the initial data - such as the explicit representation of distance, size, shape, density or orientation - by means of new attributes or the creation of new entities - such as towns, districts, urban blocks or building alignments - that are also characterised by new attributes. This effort towards the explicit representation of information exceeds the simple case of generalisation. For example, it could be used for multiple representation databases as these meso objects are often the real link between existing databases. Moreover it opens the way to new geographical analysis that require different levels of analysis, as the study of the influences of individual entities on more global ones. The phenomena of aggregation - which is the core of generalisation process - is recognised as being a core and complex domain in geography. The introduction of methods of analysis that allow the automated recognition of such entities should certainly improves research in the broader domain of geography.

ACKNOWLEDGEMENT

I want to thank Serguei Levachkine to invite me to the GEOPRO 2002 conference as well as Nicolas Regnauld for his helpful comments.

REFERENCES

- [1] K.S. Shea and R. McMaster, Cartographic generalisation in digital environment: When and How to generalize, *Proc. Autocarto 9*, Baltimore USA, 1989 56-67.
- [2] M. Bader, Energy minimization methods for feature diaplement in map generalization *PhD Thesis* University of Zurich, Suisse, 2001.

- [3] W. Mackaness, Issues in resolving visual spatial conflicts in automated map design, *Proc. Spatial Data Handing'94*, Edinburgh, UK, 1994, 325-340.
- [4] A. Ruas Automatic generalisation project: Learning process form interactive generalisation. *OEEPE Official Publication N°39*, 2001, 11-98.
- [5] L. Harrie and T. Sarjakoski, Simultaneous graphic generalization of vector data sets. *GeoInformatica* 6(3), 2002, 233-262.
- [6] J.M. Ware, C.B. Jones and N. Thomas, Automated cartographic map generalization with multiple operators: a simulated annealing approach, to appear in *The International Journal of Geographical Information Systems* 2003.
- [7] K. Beard, Constraints on rule formation, in Buttenfield and McMaster (Ed.) *Map Generalization* (Longman, 1991) 121-135.
- [8] A. Ruas, OO-constraint modelling to automate urban generalisation process, *Proc. Spatial Data Handling'98*, Vancouver, Canada, 1998, 225-235.
- [9] A. Ruas Modèle de généralisation de données géographiques à base de contraintes et d'autonomie. *PhD Thesis* University of Marne La Vallée, ftp://ftp.ign.fr/ign/COGIT/THESES/RUAS, 1999.
- [10] M. Barrault, N. Regnauld, C. Duchêne, K. Haire, C. Baeijs, Y. Demazeau, P. Hardy, W. Mackaness, A. Ruas, R. Weibel, Integrating multi-agent, object oriented and algorithmic techniques for improved automated map generalization, *Proc. of 20th International Cartographic Conference*, Beijing, China, 2001, 2110-2116.
- [11] A. Ruas, Data modelling for constraints-based generalisation, to appear in CUES, 2003,
- [12] A. Boffet, Méthode de création d'information multiniveaux pour la généralisation cartographique de l'urbain PhD Thesis Université de Marne la Vallée, ftp://ftp.ign.fr/ign/COGIT/THESES/BOFFET, 2001.
- [13] X. Barillot and C. Plazanet, Analyse des formes des routes, in A. Ruas (Ed.) Généralisation et représentation multiple (Paris: Lavoisier Hermes Sciences, 2002) 203-224.
- [14] M. Wertheimer, Über Gestalttheorie, Erlangen, 1925. In the translation by Willis D. Ellis published in his Source Book of Gestalt Psychology, New York: Harcourt, Brace and Co, 1938.
- [15] S. Christophe and A. Ruas, Detecting building alignments for generalisation purposes, *Proc. Spatial Data Handling*, Ottawa, Canada, 2002, 419-432.

GENERALIZATION WITHIN A GEOPROCESSING FRAMEWORK

Dan LEE
Software Development Department, ESRI, Inc.,
380 New York Street, Redlands, CA 92373
USA

ABSTRACT

Map generalization, the process of data transformation, reduction, and integration, requires a powerful and flexible environment in modern GIS. With a new architecture and user experience, ArcGIS, the objectoriented generation of ESRI's GIS product, provides a spatial framework to support GIS and mapping needs. Geoprocessing, combining its earlier command operation with a modern user interface, has become an integral part of the upcoming releases. Developing generalization tools within a geoprocessing framework has given us opportunities to explore new technology and data models and to make enhancements using better techniques. This paper briefly reviews the research and development in the past few years, introduces the geoprocessing concepts and environment, and discusses how map generalization tools have been enhanced and implemented in the geoprocessing framework, and outlines what remains ahead.

KEY WORDS

geoprocessing, generalization, data integrity

1. INTRODUCTION

Many data providers, map producers, especially national mapping agencies, and other GIS organizations have built high resolution, high accuracy databases to represent the geographic world. The potential of using these detailed master databases to serve multiple-purpose and multiple-scale applications can be greatly extended if the automated generalization becomes available in modern GIS systems. Tremendous efforts have been made within the research and development community from deriving numerical methods (as summarized in [1] and [2]) to implementing generalization functions into commercial GIS products and receiving evaluations from major national mapping agencies [3]. Better qualification of generalization solutions and full integration of generalization capability for deriving new datasets and compiling cartographic products has become inevitable.

1.1 Previous development

In pursuing GIS-based map generalization, the concepts and classifications of generalization operations were defined [4] and a set of the most requested generalization tools was created for the coverage data model in ESRI's Workstation ArcInfo in late 1990s [5]. These tools were implemented following generalization principles and produce less complexity and reduced detail in the output, as shown in Figure 1, while preserving close representations of the geographic objects and meeting data integrity requirements. Our main tasks included defining generalization rules, creating algorithms, setting up logical procedures, facilitating post-processes, and supporting user's requests and benchmarks. The above practice prepared us to meet new challenges.

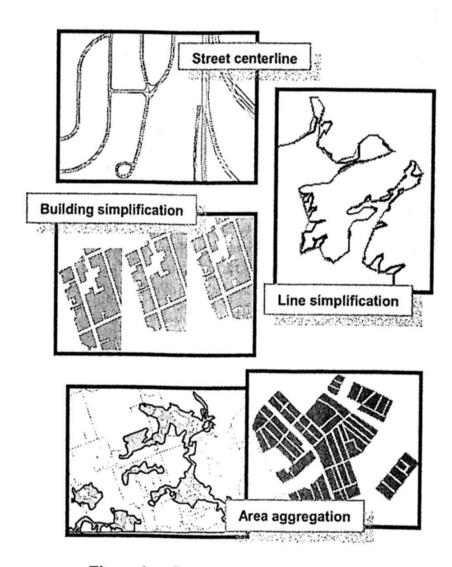


Figure 1 Coverage tools for generalization

1.2 Critical experiences

Successful automated generalization translates human knowledge in manual generalization into explicit rules and logics so that they can be coded in computer language. Very few manual generalization guidelines exist in textbooks, and they are usually too general and incomplete. For example, the instructions for area building simplification state: "The measured area of the simplified outline should remain roughly the same as the area of the original", "General form should be maintained", and "If possible, draw rectangles" [6]. When we created the Arc command BUILDINGSIMPLIFY, we had to analyze existing maps, use reverse engineering and common sense to extract explicit rules for the implementation. Here are some examples:

- A building must be simplified if it contains one or more sides shorter than a specified length.
- Building simplification should preserve and enhance orthogonality, that is, making near-90-degree corners exactly 90 degrees.
- A building can be simplified by filling up corners, cutting off or widening isolated small spaces (intrusions or extrusions), or by straightening or averaging a number of consecutive sides, while keeping the measured area roughly the same.
- A building that does not retain a minimum area will be excluded, if the user chooses to.
- Under relatively large reduction, a building can be turned into a rectangle taking the shape of the bounding box oriented along the longest side and an area close to the original.
- For connected buildings, only the outer boundaries are simplified and the buildings should remain connected after simplification.
- Five simplification statuses are to be recorded to support post editing; they are: properly simplified, partially simplified due to potential conflict, simplified but too small, partially simplified group, non-simplified group.

A previous paper presented more details on building simplification [7]. For every generalization tool we develop, a set of such explicit rules must be defined and enhanced over time.

Geographic databases usually store features in various for example as political boundaries, transportation, vegetation, and so on. Each can contain point, linear, and polygonal feature types and their attributes. Generalization may involve designing a new classification for a target output, selecting features for inclusion, reducing details through simplification, aggregation, typification, and so on, and resolving spatial and cartographic conflicts for the final products. These operations must be carried out in a logical order; sometimes decisions may depend on the status of the intermediate results; interactive inspection and editing may be needed. In response to some users' requests and major benchmark specifications, we delivered procedures that transform large-scale data into a small-scale space with fairly encouraging results [8]. However, without a complete integration of generalization capability, from enriched databases that support generalization analysis

and decisions to a powerful, flexible framework that facilitates automated processes and interactive compilation, the limitations of the early solutions were obvious.

2. GEOPROCESSING IN ARCGIS

With object-oriented technology and the new geodatabase for modeling the world, ArcGIS marks a new generation of ESRI software. The upcoming release of ArcGIS 9.0 will present a geoprocessing framework for carrying out GIS operations. The integration of generalization tools into ArcGIS has been in progress with the ultimate goals to support database generalization and cartographic generalization from geodatabases.

Taking a relatively broad definition, geoprocessing in ArcGIS 9.0 [1] refers to the application of core GIS operations that create new spatial data from existing or derived data. The basic GIS capabilities found under this umbrella include data conversion, spatial analysis, and data management. A typical geoprocessing operation takes input geodatabase features, performs an operation on them, and returns an output geodatabase feature class. For example, the Buffer tool takes point, line, or polygon features and creates a buffer polygon feature class based on user-specified parameters.

The two core components in ArcGIS Desktop are ArcCatalog (database creation and management software) and ArcMap (start-to-end mapping software). The geoprocessing tools have been made accessible from the dockable ArcToolbox window and the Geoprocessing window in both software environments. Figure 2 shows the access to the ArcToolbox window and the Geoprocessing window in ArcCatalog and the partially expanded view of the toolboxes and toolsets.

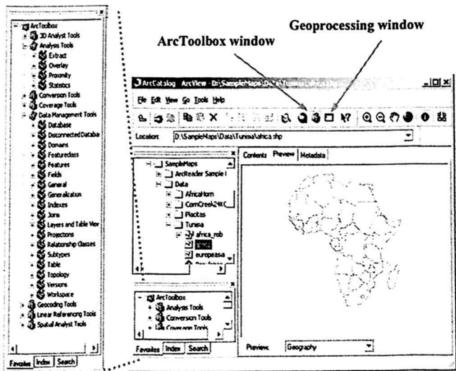


Figure 2. Access to geoprocessing in ArcCatalog

To perform geoprocessing tasks, you can choose one of the following four methods: tool dialogs, command lines, model tools, and scripts (Figure 3). A tool dialog can be invoked from the ArcToolbox window. The dialog gives an easy user interface for you to specify data and

parameters to perform a single operation. A command line, similar to Workstation ArcInfo command line, allows you to specify and execute a cammand in the Geoprocessing window. Once you type in a tool name, the Geoprocessing window will prompt you with the

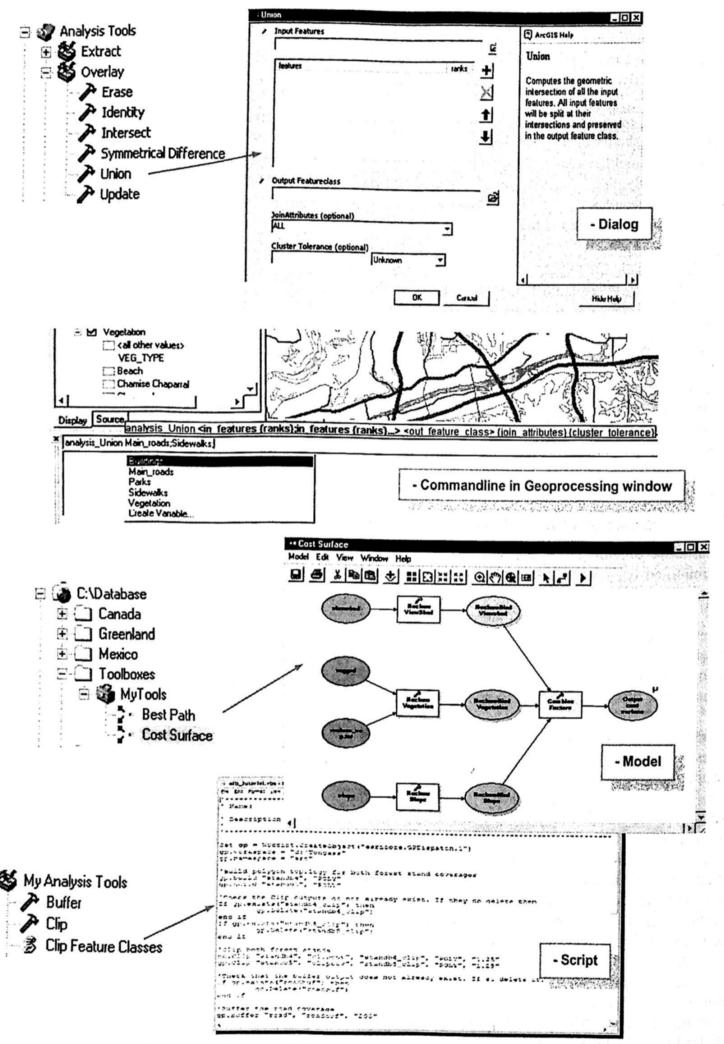


Figure 3. The four methods for performing geoprocessing tasks

usage of the command so that you can enter parameters and options accordingly, then execute the command. Both dialog and command line allow one tool execution at a time. A model can be created in Modelbuilder, which provides you with a graphical environment to construct a diagram of the steps-representing a model-to complete a geoprocessing task. A model executes processes in chained sequence. A script offers an efficient and effective way of managing geoprocessing tasks. especially those involving a large volume of data, repetitive work, and more complex decision-making. The development platform for ArcGIS Desktop applications is known as ArcObjects, a collection of the Component Object Model (COM)-based software components with GIS functionality and programmable interfaces [10]. The Geoprocessor is an ArcObject that supports the COM interface IDispatch, which enables interpretive and macro languages, such as VBScript, JScript, and Python, to access COM objects. This IDispatch object is called IGPDispatch and it exposes all geoprocessing tools to scripting clients.

A variety of environmental settings, such as default workspace location, output extents, cluster tolerance, and so on, can be set and applied at the application level, the model level, or a specific tool level. The geoprocessing framework sets the fundamental, flexible environment for users to manage geographic data operations.

3. DEVELOPING GENERALIZATION TOOLS IN GEOPROCESSING

The integration of generalization into ArcGIS will tremendously extend the power of using master databases for multiple purposes and multiple scales applications. Our development is underway with the ultimate goals to support both database (or model) generalization and cartographic generalization, as distinguished by researchers [11] from geodatabases.

For database generalization, a new dataset or database is derived from a master database with a reduced level of detail, usually for a smaller scale analysis or representation. Such a process can be a single operation on particular features, for example simplification of rivers or aggregation of buildings, or a logical sequence of generalization operations in conjunction with other necessary steps to reach the desired result among multiple feature classes. The geoprocessing framework in ArcGIS described above provides an ideal environment for managing database generalization and preparing data for cartographic finishing. Each generalization tool will be made in compliance with all other geoprocessing tools and can be executed via one of the four geoprocessing methods mentioned above. Certain parameters and options, for example minimum spacing, symbol sizes, and so on, can be specified as environmental settings (not in place yet) and used to guide generalization processes.

2.1 Coverage generalization tools

To continue supporting coverage model applications, the Coverage Tools toolbox in geoprocessing contains tools created from Workstation ArcInfo commands, each of which takes coverage input and produces coverage output. All of the generalization commands implemented in Workstation ArcInfo, as mentioned above, are included in the Generalization toolset in this toolbox. The conversion tools going between coverage and geodatabase features (see more about geodatabase features below) make it possible to process data interchangeably.

2.2 Geodatabase feature generalization tools

The geodatabase model is an object-oriented data model created with ArcGIS and stores geographic data in a commercial off-the-shelf DBMS [12]. The geodatabase extends the traditional coverage model with support for intelligent features and complex networks. Geographic features can now be defined and stored as objects with rules, behaviors, and relationships to other objects.

Based on our earlier research and understanding about generalization [13], a set of generalization functions have been defined to be built for geodatabase features; some are relatively simpler and more straightforward; others are more complex and involving features in context. We have begun to implement the simpler tools that each performs a unique generalization operation.

One of the geoprocessing toolboxes is named Data Management Tools, under which a Generalization toolset is created (see Figure 2) to contain geodatabase feature generalization tools. Like all the other geoprocessing tools, generalization tools will take geodatabase features as input and produce geodatabase feature classes as output. Currently, for line generalization, an enhanced Line Simplification tool and a new Line Smoothing tool have been added into this Generalization toolset (Figure 4). More generalization tools have been planned for future releases.

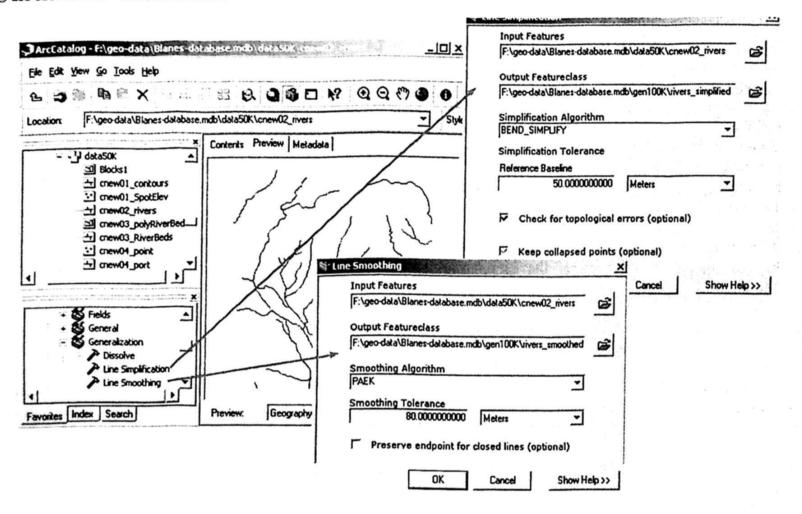


Figure 4 Line Simplification and Line Smoothing tool dialogs (Thanks to the Institut Cartografic de Catalunya for providing the testing data.)

2.3 General principles and requirements for developing generalization Tools

To ensure the quality of generalization and data integrity, some general principles and requirements are closely followed. The Line Simplification tool will be used as an example to explain and illustrate our development considerations through out this section.

Using efficient and effective techniques

Without going into long discussion, an efficient and effective generalization technique simply means it carries out the unavoidable reduction of feature complexity "in such a way that typical characteristics are least affected" [14] and the computation is elegant, easy to understand and control, and inexpensive.

To simplify digital lines, the two primary tasks are to compress over-digitized vertices and to remove small undulations so that "the course of the feature is to be perceived without ambiguity" [15] at the intended scale. There have been a wide range of published algorithms for line simplification, some are listed in [1] and [2]. We chose two distinctive algorithms: POINT_REMOVE, which is an enhanced version of the well-known Douglas-Peucker algorithm [16] and compresses lines or serves minor line simplification quite effectively, while creating more angularity; and BEND_SIMPLIFY, which is an inhouse designed algorithm that reduces extraneous bends

along lines and preserves the essential shapes of the lines and aesthetic quality quite well [17]. Figure 5 shows the nature and the difference between the two line simplification algorithms.

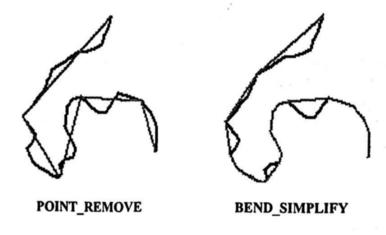


Figure 5 The two distinctive line simplification algorithms

Determining parameters

Usually generalization parameters determine the degree and quality of generalization. For research and learning, it might be nice to expose all involved parameters to the users so that they can analyze and understand the impact of each parameter by fixing other parameters. However, for easy usability and production work, too many parameters can be confusing and hard to control; simplicity in parameter design would be preferred.

For the Line Simplification tool, currently only one parameter needs to be set for either algorithm; other additional necessary parameters are internally derived, taking either empirical values or values that are logically related to the specified parameter.

The question always arises: how to come up with a reasonable value for a parameter? For geoprocessing tools, wherever possible a default value will be provided for a numerical parameter. However, for the line simplification tool or any generalization tools, unless the scale related specifications or some measurements about the data are available, no reasonable values can be suggested for the parameters. Based on our experience, setting the line simplification tolerance for both algorithms can begin with a value in a ground unit close to or a little greater than one converted from the minimum allowable spacing between lines on the map at the target scale; trials and errors may be expected in order to reach a suitable tolerance value.

Resolving topological errors

Many generalization operations more or less alter the geometric representation of the features. As soon as that happens, the spatial relationship among features might be destroyed or become incorrect. For example, when the shape of a building outline is simplified, it may overlap a neighboring building outline, which in reality is impossible. A generalization tool should avoid creating these types of topological errors, or if they are created during the process, try to resolve them.

Topological errors that might be created in line simplification are: line-crossing, coincident lines, and collapsed zero-length lines. The user has the option to have these errors detected and resolved. There could be many different ways to do so. Since these errors usually occur in congested areas, they indicate that the simplification tolerance may not be suitable for those areas, usually that it's too large. So here is how we resolve the problems:

The data will be simplified using the specified tolerance. In order to find out if errors are created, the Delaunay Triangulation will be constructed from which information about line-crossing and coincident lines can be extracted. Each pair of involved line segments will be located, a reduced tolerance (half of the original) will be applied to re-simplify these segments, and the resulting line segments will be used to update the triangulation. If errors still occur, a reduced tolerance (half of the last used) will be applied to the involved line segment. This iteration will repeat until no more errors are found. Figure 6 shows a comparison between an input line and its simplified form. The small bend the arrow points to is much smaller than those in the left circle, but could not be removed as were those in the right circle without causing line-crossing; so it was under-simplified and kept in the result.

200

The collapsed zero-length lines could only happen to closed (circular) lines when the tolerance is relatively too large. This type of error can be detected easily without using triangulation. When an error is found, a reduced tolerance will be used to re-simplify the original closed line. Again, the iteration will repeat until the line won't collapse anymore.

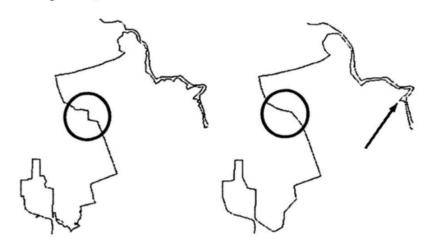


Figure 6 Before (left) and after (right) simplification: where the arrow points at is obviously less-simplified compared to the shape change in the circles, as the result of resolving line-crossing errors.

Although the result of line simplification won't contain the above three types of topological errors, other types of topological errors may still occur, for example, a severely simplified line might end up being on the opposite side of a point feature. The resolution to this problem has not been implemented.

Flagging status

It is very likely that the automated generalization process cannot produce a complete and satisfactory result. Therefore, the generalization process status and problems must be identified and flagged to support the evaluation of the result, the analysis of the parameters and option used, the post processing (interactive or semi-automatic), and the research for enhancements in the incomplete and problem areas. Quite often, the flagged areas may require a different type of generalization operation.

Again in the line simplification case, the iterative approach of resolving the topological errors described above may result in a line simplified by possibly the specified and reduced tolerances in different parts. To make the user aware of the situation and be able to review the under-generalized lines easily, two new attributes, MaxSimpTol and MinSimpTol (the maximum and minimum simplification tolerances used to simplify a line), are written for each line in the output feature class. The user knows immediately what range of tolerance is used for a particular line and whether the specified tolerance is suitable for the majority of the data. The partial attribute table of a simplified line feature class in Figure 7 illustrates the MaxSimpTol and MinSimpTol

values. In this case, the specified tolerance was 99 map units. Where 49.5, 24,75, 12.375, and 6.1875 are listed, topological errors were found and resolved by reducing tolerances for the problem line segments in one through four iterations.

hydShpShp_p80_Prk	geocomplD	Shape_Length	NaxSimpTol	MinSimpTol
hydshpshp_p80Er	275	41.1384907851158.	99	93
Test	276	638 673537399199	99	6.187
+ -31A	277	290.641681833964	99	6.1875
- Unimattydro	278	7528.15411176947	99	95
- hydCovFc_p80Er	279	110.048856035798	49.5	49.5
Hydrorc_p80Er	290	113.076568766616	49.5	49.5
hydshoFc_b80Er	281	120.830355327962	24.75	24.75
hydShpFc_p80Er	282	111.127947010314	24.75	24.75
hydCovShp_b80Er	283	94.5468366530967	99	99
hydcovshp_p80Er	284	103.163450556642	99	99
hydrcshp_b80Er	285	57.9773993830592	99	99
hydfcShp_p80Er	286	68.168316055854	49.5	49.5
hydshpshp_b80Er	287	564 77989318734	49.5	49.5
hydshpshp_p80Er	288	86.4770099908222	99	12375

Partial attribute table showing the Figure 7 MaxSimpTol and MinSimpTol values.

The under-simplified areas may imply the need for a different generalization operation, for example two closely located lines representing a narrow river may not be simplified without crossing each other. Perhaps they should be collapsed into a single line representation. A collapsed zero-length line may need to be excluded.

Insuring data integrity

There are many aspects that define data integrity. To generalization, data integrity means, at least, that the data is complete for the target output, consistently created and processed, properly linked, and timely updated. Some generalization operations reduce the level of details at individual feature level, for example, collapsing an area building to a point; therefore, the generalized features and their source features have one-to-one relationships. Others may reduce the number of features in a group or combine features in a group, for example, aggregating trees in close proximity into forest areas; therefore, the generalized features and their source features may have one-to-many relationships. It is important that the generalized features are properly linked to their source features so that the feature attributes can be calculated and transferred to the generalized features.

One of the data issues in line simplification is about how to properly handle shared geometry. Shared geometry is very common in route networks; for example, one part of a road may be shared by different routes. Theoretically the shared part of the features should be simplified the same for each route. With the availability of the topology engine in ArcGIS, we are able to identify such coincident line segments. For each group of coincident lines, the geometry of the lines will be simplified only once and the resulting geometry will be used to replace all coincident

segments in the group. In other words, shared geometries in route network are consistently processed with the coincident segments simplified and still coincident.

Also in line simplification, if the user did not choose to detect and resolve the potential topological errors, then there is a good chance that some closed lines will be collapsed into zero-length lines, as mentioned above. These features become invalid in the geodatabase and cannot be stored in a line feature class. To inform the users about these "lost" data and allow them to keep track of where the lines are, a point feature class will be generated to carry the endpoints of the collapsed or lost lines with their source line object IDs; the user can then decide to delete them, if they are indeed unimportant, retrieve the original lines through the linked object IDs, or do something else.

Although lines are simplified in shape, they still represent the same individual geographic objects, that is, they have one-to-one relationships. Therefore, normally all the source line attributes including the object IDs are copied over to the corresponding simplified lines.

2.4 Importance of generalization models

The generalization process may not be straightforward; to model the process is always a challenge. The Modelbuilder mentioned earlier helps us to experiment with different procedures, adjust the workflow according to different themes and target maps, and make the generalization processes easy to manage. You can create and edit a model diagram in Modelbuilder to put the generalization steps in a desired sequence. The diagram can be saved as a model in a user-specified toolbox and modified easily to repeat the same or similar processes for different datasets or for the same data with different parameters and options.

The model diagram in Figure 8 illustrates an experimental building generalization sequence. The goal is simple:

- Larger buildings (larger than 6000 sq ft) will be kept as areas and only their footprints will be simplified
- Medium-size buildings (4000 6000 sq ft), since they are too small to remain as areas but large enough to be included, will be represented by points
- Small buildings (smaller than 4000 sq ft) will simply be excluded.

The input clip_bldgA is a large-scale area building feature class. It is first converted to a coverage so that the coverage Building Simplification tool can be used to simplify the building footprints, keeping only the three large buildings. Then, in a parallel process, the Select tool selects the three medium-size buildings from clip_bldgA based on a query, that is, size greater than 4000 and smaller than 6000 sqft, followed by the Feature To Point tool that collapses the buildings into their corresponding centroid points. All other buildings are left out. It is possible that some of the small buildings are aggregated if

they are in close proximity; constraints can be added so that the aggregation won't cross major roads. But for the simplicity of illustration, aggregation was not considered. Three other simple models were also used to generalize the rivers (collapse, simplify, and smooth), roads (select, collapse to centerlines, and extend), and contours (select, simplify, and smooth). More details will be given at the presentation due to the limitation of the paper length. Figure 9 shows the result of the experimental generalization.

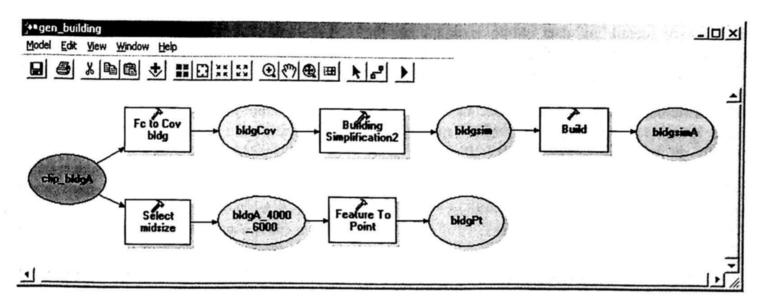


Figure 8 Example model for building generalization

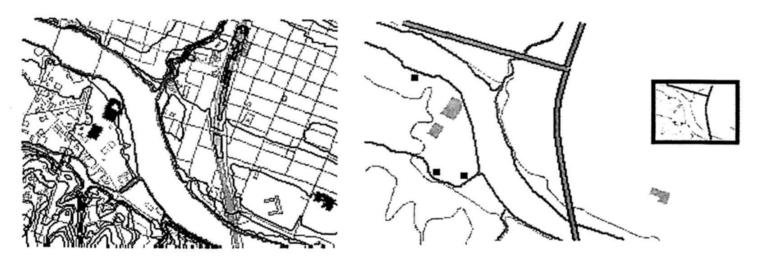


Figure 9 Before (left) and after (right) generalization; result at a smaller scale (in the black box)
Source data: the USGS 1:24000 mapsheet, Austin East, TX, USA

In reality a much more complex model or even a number of models might need to be set up for generalization. Features to be included for the target database or map must be selected from the master database; different generalization operations, parameters, and sequences are to be decided for various themes and features; some intermediate and final results need to be inspected and possibly edited. As shown in the generalized map at the reduced scale in Figure 9 (about four times reduction), some buildings look too close to rivers and may need to be displaced that the space between a building and a river is perceivable and no new conflicts are created. A complete generalization workflow could include automated processes done by models and any necessary

interactive or semi-automatic processes. Being able to record the whole workflow and easily alter it for repeated use and updating would be very helpful in production. Our research and development continue.

4. ONGOING AND FUTURE WORK

For line simplification in particular, there are still areas that need enhancement: bottle-neck areas may need to be widened, small consecutive bends, usually representing the switchbacks in mountain roads or rivers, may be combined into fewer and larger bends, multiple features need to be taken into account when detecting spatial conflicts, line symbol width will need to be considered for

cartographic quality. The recursive or iterative approach of resolving topological errors may not be the only or best way. Other techniques and choices are to be explored.

The Line Smoothing tool has been implemented following similar general principles as for line simplification. Although the topological errors occurring in smoothed results can be detected and flagged, the resolution of these errors has not been fully defined.

We will continue to derive techniques that satisfy the essential requirement of generalization, that is, reducing the level of detail in data while maintaining the characteristics of the presented geography. It is not the purpose of the this paper to discuss the details of the rules, algorithms, and steps for all the planned generalization tools, but tremendous research has been made to analyze feature properties and preserve their representation faithfully in reduced forms. Figure 10 shows some preliminary results of developing area aggregation tools and collapsing dual-lines to centerlines tool using triangulation.

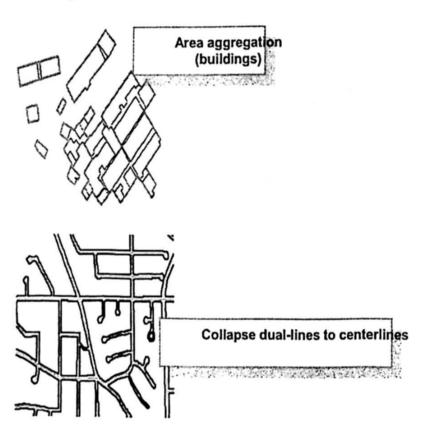


Figure 10. Development of area aggregation and collapse tools in progress

The full integration of generalization in ArcGIS must consider interactive operations in addition to automated batch processes. Interactive generalization means allowing interactive selecting of features, applying generalization to the selected features, dynamically viewing the changes while altering a parameter, being able to undo and redo the changes, and saving the results as needed. When a mapped area contains mixed levels of complexity, a uniform generalization process may not be the desired solution. The interactive generalization could

help to apply different parameters and choices to different areas based on visual judgment, and do it right.

Until automated generalization can produce perfectly satisfactory results, if that is ever possible, the completion of generalization tasks will still need to count on the necessary interactive editing and refinement of the automated results. Some of the unresolved areas or imperfect solutions can be flagged, as already discussed above. An efficient post-editing management environment should be part of the geoprocessing framework. Meanwhile, specialized post-editing tools will need to be developed; they can be semi-automatic, meaning that a human being makes the decisions and choices that the computation cannot make, and the computer does the work. For example, the coverage tool Create Centerlines flags unresolved intersections (too complicated to connect properly in the automated process). An interactive specialized post-editing tool could let the user pick which lines should be connected and which intersection style to use (assuming a set of pre-defined intersection styles, such as "T"-intersection, "Y"-intersection, and so on, is available) and let the program finish making the desired intersection. The current geoprocessing framework has not yet supported interactive generalization and postediting, but the design and planning are underway.

One very important and often neglected side of the integration of generalization in our GIS systems is database design. Building effective classification of features and enriching databases with necessary geometric, relational, and attributive information to support generalization is very critical. No matter how sophisticated the techniques used to create generalization tools are, their ability to understand and achieve what a human being can see and do is always limited [18]. The fact that computational methods simply cannot interpret geographic differences from purely feature geometry and may result in the loss or distortion of spatial integrity after generalization, such as disconnected road or stream networks, broken boundaries, or misplaced features, has led to the increasing interests and demands on multiplescale representations and database-driven cartography. Our research and prototype efforts have begun.

5. CONCLUSIONS

There is obviously a long journey ahead in pursuing generalization solutions. The development of generalization tools in the geoprocessing framework is just the first major step towards our goal of supporting database and cartographic generalization.

As geoprocessing development advances, more generalization tools will be added and more features in context will be taken into account in generalization operations. Our main focus will be on providing flexible and practical ways of managing generalization work

(interactive or automatic), formalizing generalization guidelines and models for further automation and enhancement, following up with post-editing management and tools, and exploring database potentials in balancing the computational limitation and expenses.

REFERENCES

- [1] R.B. McMaster and K.S. Shea, Vector-based Generalization in *Generalization in Digital Cartography* (Washington D.C.: Association of American Geographers, 1992) 71-98
- [2] Elsa Maria Joao, The algorithmic approach in *Causes and Consequences of Map Generalization* (London: Taylor & Francis Ltd, 1998) 10-13.
- [3] B. Baelia, M. Pla, and D. Lee, Digital Map Generalization in Practice, the working paper for ICA Workshop on Progress in Automated Map Generalization, Barcelona, 1995.
- [4] ESRI, Automation of Map Generalization in ArcInfo the cutting edge technology, search for map generalization at http://arconline.esri.com/arconline, 1996.
- [5] ESRI, Map Generalization in GIS: Practical Solutions with Workstation ArcInfo Software, search for map generalization at http://arconline.esri.com/arconline, 2000.
- [6] Swiss Society of Cartography, Cartographic Generalization Topographic Maps, Cartographic Publication Series, No.2 (Zurich: Swiss Society of Cartography, 1987) 22
- [7] Z. Wang and D. Lee, Building Simplification Based on Pattern Recognition and Shape Analysis, *Proceedings of the 9th International Symposium on Spatial Data Handling*, Beijing, China, 2000, 58-72.

rapid the state of the state of

A Solidar Tollandare in Control Control Control (1989)

- [8] D. Lee, Data Conversion and Generalization from 1:1000 Scale Database to 1:5000 Scale Database, benchmark report to Hong Kong Land Information Center, 1997.
- [9] ESRI, ArcGIS Overview, http://www.esri.com/software/arcgis/overview.html, 2003.
- [10] ESRI, ArcObjects Technology, http://www.esri.com/software/devsolutions/arcobjects/index.html, 2003.
- [11] R. Weibel, and C.B. Jones, Computational Perspectives on Map Generalization, *GeoInformatica*, 2(4), 1998, 307-314
- [12] ESRI, ArcGIS Geodatabase, http://www.esri.com/software/arcgis/geodatabase.html, 2003.
- [13] ESRI, Automation of Map Generalization in ArcInfo, search for map generalization at http://arconline.esri.com/arconline, 1996.
- [14] E. Imhof, Cartographic Relief Presentation (Berlin: Walter de Gruyter, 1982) 55.
- [15] R.W. Anson, Basic Cartography for Students and Technicians, Vol.2 (London and New York: On behalf of International Cartographic Association, Elsevier Applied Science Publishers, 1988) 37.
- [16] D. H. Douglas and T. K. Peucker, Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or Its Character, *The Canadian Cartographer*, 10(2), 1973, 112-123.
- [17] Z. Wang, Manual versus Automated Line Generalization, *Proceedings of the GIS/LIS '96 Annual Conference and Exposition*, 1996, 94-106.
- [18] D. Lee, Making Databases Support Map Generalization, *Proceedings of the GIS/LIS '96 Annual Conference and Exposition*, 1996, 467-478.

MEASURES FOR SEMANTIC QUALITY AFTER POLYGON GENERALISATION

Tao CHENG and Zhilin LI
Department of Land Surveying and Geo-Informatics
The Hong Kong Polytechnic University
Hung Hom, Kowloon,
Hong Kong
Email: {lstc; lszlli}@polyu.edu.hk

ABSTRACT

Map generalization changes the semantics and geometry of map objects according to the context defined by users. How to evaluate and ensure the quality of generalization has become a major issue in contemporary digital cartography. The semantic change after generalization has been seldom studied compared with the other two aspects, i.e. geometry and topology. This research investigates the effect of generalization operations on the semantics of maps objects. A set of quantitative measures for semantic change is put forward. A case study of a land use map is carried to illustrate the practical usefulness of these proposed measures, with merging as an example for polygon generalization. The results indicate that these measures are not only sound in theory but also meaningful in practice

KEY WORDS

semantics; generalization; quality assessment, measures

1 INTRODUCTION

Map Generalization is a process of selection and simplification of details of a map according to the scale and/or the purpose of the map [1]. Such a process is used to derive small-scale maps from large-scale maps. Many operations are applied to achieve the aim of generalization. An operation defines the geometric transformation that is to be achieved; and a generalization algorithm is then used to implement the particular transformation. There are many operations used in generalization and the basic ones are simplification, selection/elimination, merging aggregation, symbolization, collapse, exaggeration, displaceme, t, typification and smoothing [2, 3, 4].

In order to derive the simplified model (or representation) of reality at a smaller scale, these operations essentially modify the geometry, topology and/or semantics of the objects from a high resolution to a low resolution. For example, simplification and exaggeration only modify the

metric aspects of the objects; selection/elimination, merging, symbolization and aggregation essentially modify the topological aspects; and displacement is primarily metric, but can also cause topological changes in some cases [5]. The changes in topology and geometry may result in semantic changes, and vice versa. The topological changes indirectly imply metric changes as well, but not vice versa. For instance, where the real world distance between a lake and a village is too small to be displayed graphically at a given scale, it is certain that the railway and road between the lake and village also cannot be represented, resulting in a spatial conflict in representation. Displacement is a solution, but leads to a loss in positional accuracy [6]. Further, due to the removal of the railway/road, the semantics of the map is then changed, which leads to loss in the semantic accuracy.

How to evaluate and ensure the quality of generalization is becoming a major issue in contemporary digital cartography. The assessment of automated generalization results has so far largely relied on visual and qualitative methods that are specific for particular procedures [7,8,9,10], particular aspects of the spatial objects (attribute accuracy) [11,12], and particular shape changes (only for linear objects) [13]. Very few quantitative assessment methods are available when multiple objects are involved or the entire map needs to be characterized [4, p.150]. Recently a method has been proposed by Barl [14] for evaluating generalization quality by comparing the difference between the observed result (after generalization) and the theoretical result (reference α initial characterization of the object). In Bard's approach, three geometric properties are taken into consideration, i.e. position, concavity ratio and size. Galanda [15] modeled the constraints for polygon generalization, which could serve as a guideline for general quality control. But semantic issue is rarely touched in such work.

Indeed, the change in semantics after generalization has been seldom studied compared with the other two aspects, geometry and topology. This research is to investigate the effect of generalization operations on the semantics of polygon objects on thematic maps, as semantics plays a fundamental role in the modeling and representation of geographic objects [16, p.559]. A set of quantitative measures will be developed to measure such effects.

The next section defines a strategy for the evaluation of the quality of semantics after generalization. A set of quantitative measures, i.e. accuracy, consistency and completeness, is described in Section 3. The effect of generalization on semantics is analyzed in Section 4, where merging is taken as a case for deep analysis. An experimental test is reported in Section 5. A discussion on the practical usefulness of these measures is given in Section 6. Some concluding remarks are given in Section 7.

2 ASSESSMENT OF SEMANTIC QUALITY: A STRATEGY

2.1 MERGING AS A CASE FOR STUDY

For the generalization of polygons on thematic maps, the two basic operations are merging and aggregation. Here merging refers to a process to eliminate small areas or sub-polygons. This is also referred to as coarsen in some literature. After the merging process, the original objects cease to exist (see Figures 1 and 2). Aggregation refers to the process that deletes edges between similar objects and forms a composite object. The semantics of the original objects are then transferred to the new composite object, but the original objects do not cease to exist [17]. Merging is usually done with priority given to the neighbor that shares the longest border or the largest area. These two approaches to merging have been implemented in Arc/Info as elimination [18]. Indeed, Merging is actually a geometry-driven approach. On the other hand, aggregation is usually carried out because of common thematic characteristics with the neighbors, thus it is actually a class-driven approach. The semantic change in aggregation is not as obvious as in merging. Therefore, this paper will concentrate on the semantic change due to merging operation. Although merging can be applied to point, line and area objects, we mainly consider the area objects (i.e. polygons) in thematic maps.

2.2 3 ELEMENTS FOR DATA QUALITY

For evaluation of semantic quality, attribute accuracy, completeness, consistency and currency are the elements to be considered [19].

Accuracy is the probability of correctly assigning a value. It describes the stochastic errors of observations on attributes and is a measure that indicates the probability that an assigned value is acceptable.

Completeness refers to the symmetric difference between the perceived reality and the database at a given moment. Consistency is the validation of semantic constrains. It is the result of the validation of semantic constrains of the objects, attributes or relations.

Currency measure changes over time and describes the semantic quality at a given time, say T.

Here accuracy, completeness and consistency of the semantics after generalization are discussed as currency of the semantics involves the temporal change of the data.

2.3 GENERALIZATION AT 3 SPATIAL LEVELS

It has been pointed out that map generalization could be carried out at one the three spatial levels 15, 20], i.e.

- Map level,
- Class (group) level and
- individual feature level.

For polygon generalization, an individual feature means a polygon. Such a three-level decomposition is also referred to as macro, meso and micro levels by other researchers [21,22]. Quality evaluation of generalization at these three levels has also been conducted [14].

The elements of semantic quality for different spatial levels are shown in Table 1. It can be seen that the accuracy can be mapped to a polygon, a class and a map; consistency and completeness can only mapped to a class and a map since it don't make sense for a polygon.

Table 1. An evaluation matrix for semantic assessment

Data Quality		Spatial levels		
Component	Polygon	Class	Map	
Accuracy	+	+	+	
Completeness		+	+	
Consistency		+	+	

3 MEASURES FOR SEMANTIC QUALITY

In the following sections, a set of quantitative measures for the semantic quality will be developed. The discussion is mainly related to category maps such as a landuse map that has discrete (or nominal) value attributes.

3.1 ACCURACY

Accuracy can be assessed at three levels. At polygon (i.e. feature) level, the attribute accuracy can be described by the certainty index for the new area objects created after generalization. The accuracy of the new area objects created after generalization will be discussed in Section 4.4 and can be calculated by Equations 13–15.

At class level, the accuracy for Class K can be calculated as follows

$$\mu^{C_{k}} = \frac{\sum_{i=1}^{N} \mu_{i}^{C_{k}} * Area(i)_{C_{k}}}{\sum_{i=1}^{N} Area(i)_{C_{k}}}$$
(1)

where, i is the objects which belong to Class K after generalization; $\mu_i^{C_K}$ is the certainty of Area i belonging to Class K, which can be derived based upon Equations 13-15; N is the number of objects belonging to Class K after generalization; $Area(i)_{C_K}$ is the area of Polygon i which belongs to Class C_K .

At map level, the accuracy should be evaluated for the whole map as follows:

$$\mu^{M} = \frac{\sum_{j=1}^{M} \sum_{i=1}^{N_{j}} \mu_{i}^{C_{j}} * Area(i)_{C_{j}}}{\sum_{j=1}^{M} \sum_{i=1}^{N_{j}} Area(i)_{C_{j}}}$$
(2)

where M is the number of classes after generalization; N_j is the number of objects belonging to Class j after generalization; $\mu_i^{C_j}$ is the certainty of Area i belonging to Class C_j , which can be derived based upon Equations 13-15; $Area(i)_{C_j}$ is the area of Polygon i belonging to Class C_j .

It can be seen that valuations of accuracy at three-levels are mainly related to the object and classes after generalization. They all have a value within 0 to 1. The higher the value is, the better the accuracy is.

3.2 CONSISTENCY

Consistency for a given constrains is given by the ratio between the number of violations and the number of checks of constraint, which was defined by Salgé [19] as,

$$\rho = \frac{N^{\nu}}{N} \tag{3}$$

where N^{ν} is the number of violation and N is the number of checks of the constraint.

The obvious violation in semantics after generalization is the change of class types for the objects. Therefore, Equation 3 is modified to describe the attribute change at the class level and the map level.

At class level, the consistency can be defined as

$$\rho_o^{c_k} = \frac{N_{o_k}^{\ \ \nu}}{N_{c_k}} \tag{4}$$

where $N_{O_k}^{\ \nu}$ is the number of area objects which belonged to Class C_K (before generalization) and changed their class type after generalization; N_{C_k} is the number of area objects belonging to Class K (before generalization).

As for the map level, the consistency can be defined as

$$\rho_{o}^{M} = \frac{N_{o}^{V}}{N} = \frac{\sum_{j=1}^{M} N_{o_{j}}^{V}}{\sum_{j=1}^{M} N_{c_{j}}}$$
 (5)

where M is the number of classes (before generalization); $N_{O_j}^{\nu}$ is the number of area objects which belonged to Class C_K (before generalization) and changed their class type after generalization; N_{C_j} is the number of area objects belonging to Class j (before generalization).

In addition to number of objects that change their class types after generalization, the consistency of semantics can be represented by the attribute change in terms of areas of the objects. Bregt and Bulens [23] defined the attribute change index as follows:

attribute change index
$$= \frac{sum \ of \ absolute \ area \ differences \ per \ class}{total \ surface}$$
(6)

In Equation 6, the area difference is that of the original with generalization per class.

Equation 6 can be considered as another measure for the consistency at the map level. In order to have a same format, Equation 6 is re-written as Equation 7:

$$\rho_{Area}^{M} = \frac{Area_{o}^{V}}{Area} = \frac{\sum_{j=1}^{M} Area_{o_{j}}^{V}}{\sum_{i=1}^{M} \sum_{j=1}^{N} Area(i)_{C_{j}}}$$
(7)

where M is the number of classes (before generalization); $Area_{O_j}^{\ \ \nu}$ is the absolute area difference of objects belonging to Class j before generalization and after; $Area(i)_{C_j}$ is the area of Object i which belongs to Class j before generalization.

At class level, the consistency in area can be defined as Equation 8.

$$\rho_{Area}^{C_K} = \frac{Area_0^V}{Area_{C_K}} = \frac{\sum_{i=1}^N Area(i)_{C_K} - \sum_{i=1}^{N_b} Area(i)_{C_K}}{\sum_{i=1}^{N_b} Area(i)_{C_K}}$$
(8)

where N is the number of objects belonging to Class C_K after generalization; N_0 is the number of objects belonging to Class C_K before generalization, $Area(i)_{C_K}$ is the area of Polygon i which belongs to Class C_K before generalization.

3.3 COMPLETENESS

The completeness can be given by two figures: the rate of over-completeness (or commission) and the rate of missing data (omission). Completeness applies for the objects of a class, attributes and relations between objects [19, p.147]. Since objects and classes will be lost after

generalization, here we only discuss the situation of omission. Omission is defined by Salgé [19] as follows,

$$\tau^- = \frac{N^-}{\max(N, N^0)} \tag{9}$$

where N^0 is the number of occurrence in the perceived reality; N is the number of occurrence in the sample; N^- is the number of occurrence in the perceived reality which does not exist in the sample.

Here we assume the original map before the generalization to be the perceived reality, and the map after generalization to be the sample. In such a case, $N^0 \ge N$ and $\max(N, N^0) = N^0$. Therefore Equation 9 is modified as Equation 10 to evaluate semantic completeness at the class level in terms of object omission for each class.

$$\tau_{O}^{-c_{K}} = \frac{N_{O_{C_{K}}}^{-}}{N_{O_{C_{K}}}^{0}} \tag{10}$$

where $\tau_o^{-c_K}$ is the ratio of object lost in Class C_k ; $N_{o_{c_K}}^-$ is the number of objects lost for Class C_K (omitted) after generalization; $N_{o_{c_K}}^0$ is the number of objects belonging to Class C_K before generalization.

Equations 11 and 12 are the semantic completeness at map level in terms of object omission (the loss of objects for the whole map after generalization) and class omission (the loss of thematic classes after generalization), respectively.

$$\tau_o^{-u} = \frac{N_o^-}{N_o^0} \tag{11}$$

where τ_o^{-1} is the ratio of object lost for the whole map N_o^- is the number of objects lost for the whole map after generalization; N_o^0 is the number of objects for the whole map before generalization.

$$\tau_C^{-M} = \frac{N_C^-}{N_C^0} \tag{12}$$

where τ_C^{-M} is the ratio of class lost for the whole map N_C^- is the number of class lost (omitted) after generalization; N_C^0 is the number of class types before generalization.

3 EFFECT OF GENRALIZATION ON SEMANTCS: MERGING AS A CASE

Here an example is used to illustrate the effect of generalization operation on the semantic change, with merging as an example.

4.1 SITUATION 1: COARSEN

The first situation (see Figure 1) is that the small area D is contained in (or isolated by) a large area A (as

background). After generalization, D is merged into A. We use A' to represent the area after merging, although in the database it is still represented as A.

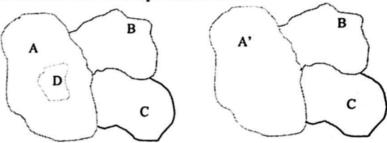


Figure 1. The case of merging (1) - Small area D is merged into background area A.

4.2 SITUATION 2: AREA- AND BORDER-BASED

The second situation involves a small area D adjacent to several larger areas (Figure 2A). The merging will depend upon some geometric criteria defined. In general, there are two ways to merge Area D. The first way is to merge D with Area A due to its biggest size (Figure 2B). This approaches of merging is referred to as area-based merging in this paper. The second way is to merge D with Area C due to its longest common boarder with D (Figure 2C). This approach is referred to as border-based merging in this paper. These two ways are geometry-based approaches and they are implemented GIS software such as ArcInfo (see [18]).

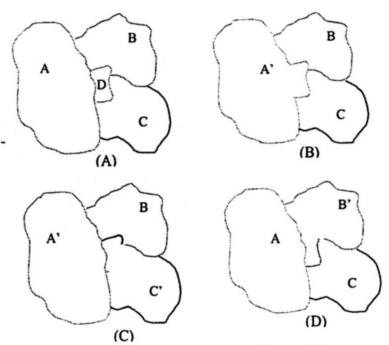


Figure 2. The case of merge (2) - Small area D is merged into one of its adjacent areas.

(A) Small area D and its adjacent areas; (B) Small area D is merged into A, (C) Small area D is merged into C; (D) Small area D is merged into B

4.3 SITUATION 3: SEMANTICS-BASED

The third way can be a semantics-based or knowledge-based approach, i.e., to merge D with Area B due to its closest semantic similarity with D (Figure 2D). Semantic similarity of two area objects has been discussed in several papers. For examples, the sharpness of a boundary

for a polygon was defined based upon the purity of polygons [24]; the sharpness of boundaries for regions in raster format was defined based upon the fuzzy membership value [25, p.64]; a semantic similarity evaluation matrix was proposed in [26].

Other semantics, such as priority given to "wood into forest" rather than "wood into lake", may also be considered as criteria for merging. Indeed, it can be considered as knowledge-based approach. The use of the prior-knowledge of thematic maps has been proved an effective approach to landuse map generalization [27].

4.4 THEMATIC ACCURACY (CERTAINTY) OF INDIVIDUAL POLYGON AFTER MERGING OPERATION

Although these three ways has different criteria for merging, the thematic certainties of the new created area A', B' or C' can be calculated as follow (although in the database they are still represented as A, B, C, respectively):

$$u'_{r} = \frac{Area(A) \cdot u'_{r} + Area(D) \cdot u'_{r}}{Area(A) + Area(D)}$$
(13)

$$u_F^{C_s} = \frac{Area(B) \cdot u_F^{C_s} + Area(D) \cdot u_D^{C_s}}{Area(B) + Area(D)}$$
(14)

$$u_C^{c_c} = \frac{Area(C) \cdot u_C^{c_c} + Area(D) \cdot u_D^{c_c}}{Area(C) + Area(D)}$$
 (15)

These values are used for computation of thematic accuracy at class and map levels.

Here we assume Areas A, B, C and D are fuzzy objects [28] and they belong to Classes A, B, C and D with membership function values as $\{u_A^{C_A}, u_A^{C_B}, u_A^{C_C}, u_A^{C_D}\}^T$, $\{u_B^{C_A}, u_B^{C_B}, u_B^{C_C}, u_B^{C_D}\}^T$, $\{u_C^{C_A}, u_C^{C_B}, u_C^{C_C}, u_C^{C_D}\}^T$ and $\{u_D^{C_A}, u_D^{C_B}, u_D^{C_C}, u_D^{C_D}\}^T$, respectively. In Equations 13-15, $u_A^{C_A}, u_B^{C_B}, u_C^{C_C}$ are the membership function values of Area A, B or C belonging to Class A, B or C, respectively; $u_D^{C_A}, u_D^{C_B}, u_D^{C_C}$ are the membership function values of Area D belonging to Class A, B or C, respectively.

If Areas A and D are crisp objects which belong to Class A and D, respectively and certainly, i.e. $u_A^{c_A} = 1$ and $u_B^{c_A} = 0$, Equation 13 becomes Equation 16.

$$u_{*}^{C} = \frac{Area(A)}{Area(A) + Area(D)}$$
 (16)

If Areas A and D are thematically similar, i.e. $u_A^{c_A} = 1$ and $u_D^{c_A} = 1$, then $u_{c_A}^{c_A} = 1$. It means there is no uncertainty created by merging if A and D fully belong to a same class.

If the value of Equation 13, 14 or 15 is close to 1, it indicates the change in semantics is small and the quality of generalization is good; on the other hand, if its value is

close to 0, it indicates the change in semantics is a lot and the quality of generalization is low.

3 CASE STUDY

A land use map of a sub-area of Wuhan, China (Figure 3) is used for test in this study. It is assumed that this land map is free of uncertainty and is used as source data for generalization.



Figure 3. A detailed land use map (Baoqin Area, Hankou, Wuhan, China)

At the polygon level, the merging operation is applied to eliminate the small areas with highest priority given to the metric constraint of minimal area. Here the two algorithms available in ArcInfo (i.e. longest shared boundary (border-based) and largest area (area-based)) are used to compare their difference in effect on semantic quality after generalization. In order to check the influence of metric constraint (i.e., size of the minimum area, or minimal mapping unit - MMU), three threshold values (1000, 2000 and 4000) are applied to both algorithms. The different values of the MMU correspond to different target map scale. Figures 4A and 4B show the merged results based upon largest area and border respectively, with the MMU= 2000.

5.1 ACCURACY

Figure 4C represents the semantic accuracy of Figure 4B (which are calculated based upon Equations 13-15). It represents the thematic accuracy at polygon level.

At class level, the semantic accuracy (μ^c) for the border-based algorithm with MMU=2000 is listed in Column 2 of Table 2, which are calculated based upon Equation 1.

Based upon the accuracy at the class level, the semantic accuracy at the map level is calculated based upon Equation 2 and the measure of $\mu^M = 0.979$.

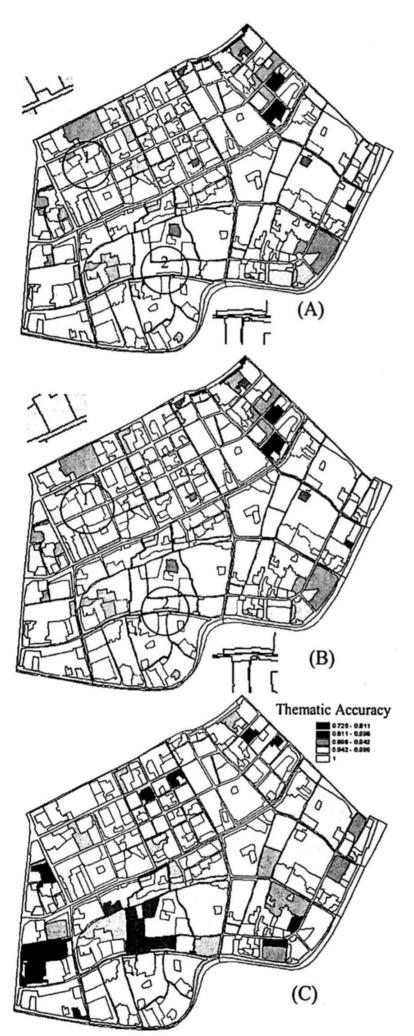


Figure 4. Generalization results by Merging.

- The merged results based upon the largest area.
- B. The merged results based upon the largest border.
- C. Thematic accuracy of Figure B. Circle 1 indicates an area in a doesn't exist in b; Circle 2 indicates an area in B doesn't exist in A.

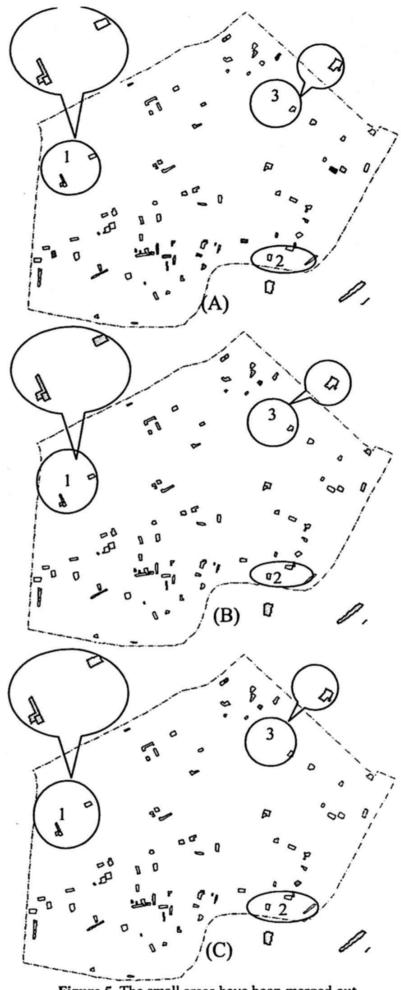


Figure 5. The small areas have been merged out.

A. represents the original class types

B. represents the class types after generalization based upon largest area;

C. represents the class types after generalization based upon largest border;

Three circles in each figure represent three typical situations. Circle 1 represents an area in C which changes its class type to different class types in A and B; Circle 2 represents an area in C which changes its class type to a same class type in A and B; and Circle 3 represents an area in C which doesn't change its class type in both A and B.

Table 2. Semantic quality at the class level

Table 2. Semantic quality at the class level										
Class	Accuracy (μ^{c})		(Consistenc		•		Comp	leteness (7	\overline{o})
Type	2B	1A	1B	2A	2B	3A	3B	1	2	3
99	0.983	0.001	0.003	-0.002	0.007	-0.025	0.003	0.029	0.100	0.200
C11	0.964	-0.028	-0.028	-0.070	-0.036	-0.149	-0.081	0.250	0.400	0.600
C12	1.000	0.000	0.000	0.000	0.000	-0.258	-0.193	0.000	0.000	0.571
C21	1.000	-0.009	-0.009	-0.024	-0.024	-0.139	-0.107	0.095	0.143	0.381
C22	1.000	0.000	0.000	-0.406	-0.406	-1.000	-1.000	0.000	0.500	1.000
C23	1.000	0.000	0.000	-0.189	-0.189	-0.344	-0.344	0.000	0.400	0.600
C25	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C26	1.000	0.000	0.000	-0.052	-0.052	-0.276	-0.276	0.000	0.167	0.583
C32	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C35	1.000	0.000	0.000	0.000	0.000	-1.000	-1.000	0.000	0.000	1.000
C36	1.000	0.000	0.000	-0.130	-0.130	-0.130	-0.001	0.000	0.667	0.667
C41	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
C51	0.982	-0.014	-0.014	0.004	0.004	-0.040	-0.008	0.286	0.286	0.429
C52		-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	1.000	1.000	1.000
C61	1.000	0.000	0.000	0.000	0.000	0.000	0.048	0.000	0.000	0.000
C62	1.000	0.000	0.000	-0.071	-0.071	-0.349	-0.349	0.000	0.167	0.500
C63	1.000	0.000	0.000	-0.239	-0.239	-1.000	-1.000	0.000	0.500	1.000
C65	1.000	0.000	0.000	0.000	0.000	-1.000	-1.000	0.000	0.000	1.000
C7		-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	1.000	1.000	1.000
D3	0.841	0.000	0.190	0.000	0.190	-1.000	-1.000	0.000	0.000	1.000
G11	0.975	0.000	0.026	0.000	0.026	0.000	0.438	0.000	0.000	0.000
MI	1.000	-0.193	-0.193	-0.472	-0.040	-1.000	-1.000	0.500	0.800	1.000
M13	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
M2	0.990	0.000	0.000	-0.049	-0.040	-0.107	-0.097	0.000	0.250	0.438
M21		0.000	0.000	-1.000	-1.000	-1.000	-1.000	0.000	1.000	1.000
М3	1.000	0.000	0.000	0.000	0.000	-0.098	0.143	0.000	0.000	0.250
R11	1.000	0.000	0.000	0.000	0.000	0.002	0.002	0.000	0.000	0.333
R12		0.000	0.000	-1.000	-1.000	-1.000	-1.000	0.000	1.000	1.000
R21	0.982	0.003	0.007	-0.021	-0.010	-0.106	-0.070	0.034	0.172	0.414
R22	1.000	0.008	0.000	-0.086	-0.095	-0.201	-0.225	0.000	0.333	0.667
R3	0.725	0.000	0.000	0.000	0.380	0.000	0.380	0.000	0.000	0.000
R31	0.981	0.000	0.003	0.002	0.011	0.008	0.042	0.087	0.159	0.275
R32	0.991	-0.012	-0.012	-0.100	-0.092	-0.429	-0.421	0.056	0.278	0.722
R41	0.962	0.007	0.005	0.042	0.038	0.057	0.057	0.043	0.043	0.217
R42	0.977	0.006	0.006	-0.012	-0.018	0.004	-0.011	0.250	0.450	0.650
SI	0.970	0.012	0.003	0.070	0.031	0.324	0.158	0.000	0.000	0.000
TI	1.000	0.000	0.000	0.000	0.000	-0.382	-0.382	0.000	0.000	0.500
T23		-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	1.000	1.000	1.000
T4	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
T42	1.000	0.000	0.000	-0.963	-0.963	0.000	0.000	0.000	0.750	0.000
U3	1.000	0.000	0.000	0.000	0.000	-1.000	-1.000	0.000	0.000	1.000
U9	1.000	-0.068	-0.068	-0.068	-0.068	-0.068	-0.068	0.500	0.500	0.500
WI	1.000	-0.038	-0.038	-0.067	-0.067	-0.144	-0.072	0.333	0.444	0.556
average	0.982	-0.077	-0.073	-0.184	-0.159	-0.345	-0.312	0.127	0.291	0.513
max	1.000	0.012	0.190	0.070	0.380	0.324	0.438	1.000	1.000	1.000
min	0.725	-1.000	-1.000	-1.000	-1.000	-1.000	-1.000	0.000	0.000	0.000
stdev	0.051	0.258	0.261	0.346	0.356	0.432	0.462	0.274	0.343	0.378

5.2 CONSISTENCY

Semantic consistency in area at the class level is summarized in Column 3 (ρ_{Area}^{c}) of Table 2, where negative value means the area for the class is reduced; and value "-1" means that there is no objects in this class (i.e. the class is omitted). It is calculated based upon Equation 8. It can be seen that the consistency index ρ_{Area}^{c} increases with the MMU at the class level. It implies that the inconsistency in semantics at the class level increase with an increase in MMU.

Semantic consistency at the map level is reported in Table 3. Both consistency indexes in terms of area (ρ_{Area}^{M}) and object (ρ_{o}^{M}) violation increases with an increase in MMU at the map level. The differences between the indices of the border-based and the area-based are not obvious at three spatial levels. This is also approved by the results that an average 98.2% of merged areas based upon the largest area and 97.2% of merged areas based upon the largest border change their class types after merging (see Figure 5). It implies that the inconsistency in semantics at the map level increases with an increase in MMU.

Table 3. Thematic consistency at the map level

MMU	ρ_{Area}^{M}		ρ_o^M	
	A	В	A	В
1000	0.006	0.006	0.089	0.087
2000	0.031	0.027	0.214	0.212
4000	0.123	0.087	0.416	0.409

(A and B represents the area-based and border-based merging, respectively)

In general, the consistency decreases with an increase in MMU at the class level and the map level. There is not much difference with the area-based and the border-based algorithm.

5.3 COMPLETENESS

Column 3 (τ_0^{-c}) of Table 2 summarizes the semantic completeness at the class level. Generally speaking, the completeness indicator increases with an increase in MMU at the class level. It implies that the completeness in semantics at the class level decreases with an increase in MMU. Since the number of objects lost in both algorithms are the same, the completeness for both algorithms are the same at the class level.

Table 4 summarizes the semantic completeness at the map level. It can be seen that the number of objects and classes after generalization decrease with an increase in MMU. In another word, the semantic completeness index at map level increases with the MMU. Since the number of objects lost and the number of classes lost in both algorithms are the same, the completeness for both algorithms are the same at the map level.

Table 4. Semantic completeness at the man level

MMU	No. of Objects	No. of $\tau_0^{-\mu}$		f Classes	τ_c^{-M} (A,B)	
	Merged		Α	В		
1000	39	0.068	4	4	0.092	
2000	92	0.114	5	5	0.216	
4000	179	0.274	12	12	0.421	

(Number of original objects is 425; Number of original classes is 44.)

It implies that completeness of semantics decreases with an increase in MMU. For the whole map, the reduction of object number is around 6.8% to 27%; the reduction of class number is around 10% to 40%.

In general, the completeness indicators increase with an increase in MMU at the class level and the map level. It implies that the quality of completeness of semantics decreases with an increase in MMU at the class level and the map level. There is no difference with the area-based and border-based algorithm.

6 PRACTICAL USEFULNESS OF THE QUALITY MEASURES: A DISCUSSION

The three quality measures of the semantics at three levels are discussed in above sections. What is their practical usefulness? Here we try to interpret these quality measures from the practical viewpoint. The accuracy of the semantics represents the purity of the polygon, the class or the map. At the polygon level, the purity of each polygon is calibrated so that any area has inclusion is clear and cautious can be taken for further analysis. At the class level, if we are especially interested in a class (such as a special landuse), we should consider if the algorithms could satisfy the requirement of the semantics for that class. At the map level, we may pay attention to the quality for the whole map to check if the map has sufficient semantic quality for further usage.

The consistency in area represents the mis-matches of areas for different classes after generalization. The area sizes of a class of polygons before and after generalization could be quite different. If this mis-match is too big to be accepted, the generalization operation should then be avoided. In other words, alternative generalization operations should be applied.

The completeness represents the omission of objects and classes at the class and map levels. Since the omission of objects might imply the omission of the owners of a parcel (a polygon in the landuse map), the indicator of completeness of object omission provides a further constraint for quality control of generalization. If a complete class of polygons is omitted after generalization, and if this class type is very important for the application, omission of that class should then be avoided. It also

implies that the particular generalization operation shouldn't be applied in such a case.

Generally speaking, the quality measures offer a means for quality control of generalization. It provides quality information of maps after generalization to map user and help them decide if the generalized map is qualified for further analysis. The quality information can also be used to facilitate the cartographers in choosing proper operations and algorithms for map generalization.

7 CONCLUSIONS

This paper picks up a particular type of maps, polygon maps, for discussion. The aim is to evaluate the effect of generalization on the semantic changes. Special attention has been given to the uncertainty in semantics created in the generalization operation, with merging operation as a case analysis. A set of quantitative measures, i.e. accuracy, consistency and completeness, has been developed for semantic quality assessment. A set of real-life data (landuse map) has been used to evaluate these measures. In such a test, different MMU values are used, which correspond to different target map scales.

It is found that the semantic quality in terms of accuracy, consistency and completeness decreases with an increase in the MMU (minimal mapping unit) value. It means that the semantic quality of the thematic map become poorer at a smaller scale. This is in accordance with our common sense. Such is also the case for the quality of topographic maps. Therefore, the set of quantitative measures developed in this project is theoretically sound and meaningful in practice.

It should be noted here that, although the set of quality measures for semantic change is only applied to merging operation in this study, it is applicable to other operations for polygon maps.

In this study, only the situation in which the original source data have no uncertainty has been analyzed. It is desirable to study the situation where there is uncertainty in the source data in further research. Also it would be of great interest to systematically analyze the changes of semantics after generalization with MMU and then to find out the correspondence between MMU and map scale.

REFERENCES

- ICA (International Cartographic Association), Multilingual dictionary of technical terms in cartography, Wiesbaden, Franz Steiner, 1973.
- [2] J. Keates, Cartographic Design and Production, (2nd edition). Longman Scientific, London, 1989, 261pp.

- [3] R. B. McMaster and K.S. Shea, Generalization in Digital Cartography, Association of American Geographers, 1992, 134p.
- [4] R. Weibel and G. H. Dutton, Generalising spatial data and dealing with multiple representations, in Longley, P.A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W. (eds.), Geographical Information Systems, Second Edition, Vol. 1, 1999, 125-155.
- [5] G. Dettori and E. Puppo, How generalization interacts with the topological and metric structure of maps, Proceedings of the 7th International Symposium on Spatial Data Handling, edited by M. J Kraak and M. Molenaar, 1996, P.559-570.
- [6] F. Harvey and F. Vauglin, Geometric match processing: applying multiple tolerances, Proceedings of the 7th International Symposium on Spatial Data Handling, edited by M. J. Kraak and M. Molenaar, 1996, P.155-171.
- [7] Z. Li and S. Openshaw, A comparative study of the performance of manual generalization and automated generalization of line features. Geographical Information 1992-1993 (AGI Year Book), 1992, 501-514.
- [8] M. De Berg, M. van Kreveld and S. Schirra, A new approach to sub-division simplification, ACSM/ASPRS Annual Convention and Exposition 4 (Proceedings AutoCarto 12): 1995, 79-88.
- [9] F. Brazile, A generalization machine design that incorporates quality assessment. In *Proceedings 8th International Symposium on Spatial Data Handling*, Vancouver, 1998, pp. 349-360.
- [10] R. Weibel and G. H. Dutton, Constraint-based automated map generalization, Proceedings of International Symposium on Spatial Data Handling, T. K. Poiker and N. Chrisman (eds.), 1998, P.214– 224.
- [11] E. M. João, The importance of quantifying the effects of generalization, In Müller, J. C., Lagrange, J. P. and Weibel, R. (eds), GIS and Generalization, Methodology and Practice, London: Taylor & Francis, 1995,183-193.
- [12] M. Painho, The effects of generalization on attribute accuracy in natural resource maps, In Müller, J. C., Lagrange, J. P. and Weibel, R. (eds.), GIS and Generalization, Methodology and Practice, London: Taylor & Francis, 1995, 194-206.
- [13] A. Skopeliti and L. Tsoulos, A methodology for the assessment of generalization quality, paper presented at Fourth Workshop on Progress in Automated Map Generalization, August 2-4, 2001, Beijing, China, http://www.geo.unizh.ch/ICA/docs/beijing2001/papers01.html
- [14] S. Bard, Evaluation of generalisation quality, paper presented in the Fifth Workshop on Progress in Automated Map Generalization, Paris, 2003, http://www.geo.unizh.ch/ICA/docs/paris2003/papers 03.html.

- [15] M. Galanda, Modelling constraints for polygon generalization, paper presented in the Fifth Workshop on Progress in Automated Map Generalization, Paris, France, 28-30 April 2003, http://www.geo.unizh.ch/ICA/docs/paris2003/papers
- [16] F. Harvey, F. Vauglin and A. Bel Hadj Ali, Geometric matching of areas comparison measures and associated Links, *Proceedings of the 8th International Symposium on Spatial Data Handling*, T. K. Poiker and N. Chrisman (eds.), 1998, P.557–568.
- [17] T. Cheng, Quality assessment of model-oriented generalisation. In Proceedings of the 4th Workshop ACI on Progress in Automated Map Generalisation. Beijing, China, 2001, http://www.geo.unizh.ch/ICA/docs/beijing2001/.
- [18] D. Lee, Recent Generalization Development and road ahead, paper presented in the Fifth Workshop on Progress in Automated Map Generalization, Paris, 2003,
 - http://www.geo.unizh.ch/ICA/docs/paris2003/papers 03.html
- [19] F. Salgé, Semantic accuracy, in Elements of Spatial Data Quality (edited by S.C. Guptill and J. L. Morrison, Exeter: Elsevier Science), 1995, P. 139-152.
- [20] Z. Li, Multi-scale representation of spatial data: An Overview. *Proceedings of Mapping Sciences* '98. 24-28 May 1998, Perth 409-416.
- [21] A. Ruas, OO-constraint modelling to automate urban generalization process, Proceedings of the International Symposium on Spatial Data Handling, Vancouver, Canada, 1998, 225–235.

entropy of the company of the section

- [22] S. Musti'ere and B. Moulin, What is spatial context in cartographic generalisation? *Proceedings of the Symposium on Geospatial Theorey, Processing and Applications*, Ottawa, Canada, CDROM, 2002.
- [23] A. Bregt and J. Bulens, Application-oriented generalization of area objects, in Molenaar, M. (editor), Methods for the Generalization of Geo-Databases, Delft: Netherlands Geodetic Commission New Series, Nr. 23, 1996, 57-64.
- [24] F. Wang and G. B. Hall, Fuzzy representation of geographical boundaries in GIS, *International Journal of Geographical Information Science*, 10(5), 1996, 573-590.
- [25] T. Cheng, A Process-oriented Data Model for Fuzzy Spatial Objects, Ph.D. thesis, Wageningen Agriculture University/ITC, Enschede, 1999, 163pp. (ISBN 90-5808-038-2).
- [26] Y.L. Liu, M. Molenaar and M. J. Kraak, Semantic similarity evaluation model in categorical database generalization, in *Proceeding ISPRS Commission IV* Symposium, Ottawa, Canada, CDROM, 2002.
- [27] W.X. Gao, J. Gong, and Z. Li, Knowledge-based generalization on land-use data, paper presented in the Fifth Workshop on Progress in Automated Map Generalization, Paris, 2003, http://www.geo.unizh.ch/ICA/docs/paris2003/papers 03.html.
- [28] T. Cheng, M. Molenaar and H. Lin, Formalization and application of fuzzy objects, *International Journal of Geographical Information Science*, 15(1), 2001, pp.27–42.

AUTOMATED DIGITAL MAP DESCRIPTION FOR GENERALIZATION

Marco MORENO, Serguei LEVACHKINE
Geoprocessing Laboratory
Centre for Computing Research-IPN
UPALM Zacatenco, CIC Building,
07738 Mexico City, MEXICO
marcomoreno@cic.ipn.mx, palych@cic.ipn.mx

ABSTRACT

The present article describes an approach to map generalization based on automatic map description. This approach is focused on the identification of the errors (inconsistencies) originated by the generalization. Map description is attempt to represent the relationships (invariants) that not change after the generalization by a set of description operators. The inconsistencies are identified by comparing the description of the map prior to and after generalization. The identified inconsistencies are corrected using local generalization parameters. This process can be iteratively repeated until all inconsistencies are corrected. The resulting map without inconsistencies indicates that the map's semantics is preserved after the generalization. Examples and results of the proposed approach are presented for the case of study of hydrological networks.

KEY WORDS

Automatic Generalization, Digital Map Description, Spatial Properties, Spatial Semantics.

1 INTRODUCTION

Over the recent years, digital cartography has been facing major changes in at least three ways: (a) the map making process has been supported by facilities of collecting and maintaining data and by supporting map design by means of computer graphics; (2) in the consequence of the former, Geographic Information Systems (GIS) have been developed to integrate the processes of data acquisition, manipulation, and representation; and (3) the process of description and understanding of maps is being explored to gain insight in the ways humans perceive and construct the surrounding geographic world through maps [1].

Cartographic map generalization is the process of designing maps overly detailed geographic data. Automation of cartographic generalization is a problem of primary importance in computer cartography. First, automatic cartographic tools allow us decreasing cost and time necessary to produce maps. Automation of generalization requires that the cartographic tools be able to find 'where' and 'how' to generalize. The 'where' depends on the capacity of analysis: the system should be able to identify which objects do not respect the

specifications. The 'how' depends on the existence of procedural knowledge and consists of rules that connect generalization conflicts with algorithms supposed to solve such conflicts [2].

Thus, this approach attempts to answer these questions. The 'where' is focused on identifying the inconsistencies and solving spatial conflicts, wich can be generated by the generalization. The 'how' is focused on modifying and establishing the parameters of the generalization functions to solve the inconsistencies.

The generalization problem is not completely solved universally due to the great number of existing case of studies. Erroneously, map generalization has in the past been seen as a set of geometric manipulations. It is true that generalization manifests itself as manipulation of geometry, but it is fundamentally driven by the need to convey specific meaning with respect to a particular map purpose [3].

Projects developed in [4] and [5] use a constraint based on approaches to map generalization, that is constraints control the map generalization process. A constraint denotes "a design specification to which solutions should adhere" [6], that allows considering both the user's needs and mapping principles such as minimal dimensions. In this approach the constraints are considered the invariants. These are used to specify the spatial object relationships that should be preserved in the generalized map.

In digital map generalization (in the interactive and automatic environment), certain spatial inconsistencies are generated. These inconsistencies represent topological and logical errors of spatial object systems; therefore, some relationships among objects that should be invariant are not completed.

The map description process can be used to reflect the map structure. Thus, the description of digital maps allows evaluating certain aspects in a simple manner prior to and after generalization, in other words, to evaluate a set of generalization constraints.

The basic difference between the alternative herein presented and other the state-of-the-art systems is as follows:

- Method to represent some map semantic invariants,
- Evaluation method to qualify generalized data, and
- Mechanism that makes possible to parameterize appropriately generalization functions.

The present article is organized as follows: in Section 2, properties of spatial objects (PSO) are described; in the next section, we present map description as tool of map generalization, showing the process employed to make the generalization. Section 4 presents the preliminary results, and finally, Section 5 contains our conclusions.

2 PROPERTIES OF SPATIAL OBJECTS (PSO)

All objects presented in maps show a land situation. This land situation is composed by thematics. The set of objects with the same primitives of representation (point, arc, and polygon) consist of a thematic (T_i) such as populations, roads, etc., wich represents a spatial object system (S_i).

$$\exists i = 1,...,n; T_i \cong S_i \text{ (where } \cong \text{ denotes equivalence)}$$
 (1)

An object system can be composed by a single layer, although spatial object system such as populations is composed of a few layers (possibly three, including punctual (O_p) , linear (O_l) , and area layers (O_a) that in this case are a group of three spatial object systems).

$$T_i = (SO_p \cup SO_1 \cup SO_a)_i; i = 1, ...,n$$
 (2)

The properties of spatial objects (P_s), are defined by relations (R) and attributes (A).

$$P_{Si} = (A \cup R)_i; i = 1, ...,n$$
 (3)

According to the nature of the property, it is possible to distinguish the following (Figure 1):

- Locational properties (λ). These properties are related to the position of spatial objects. The position is represented by numeric values of coordinates.
- Geometric properties (γ). These properties depend on some computation of distance (e.g. area, perimeter, and length). Their values are typically numeric.
- Topological properties (τ). These properties can be described without reference to their position, orientation, shape and size. A relationship can only be topologic when it is preserved under geometric transformations (translation, rotation, and scaling). It is generally represented by nominal values. For example, intersect, inside, adjacent.

- Directional properties (δ), when they concern the orientation (e.g., north, northeast, and so on). It is generally a directional feature that is represented by means of nominal values.
- Logical properties (I), when certain objects have a greater relative importance or combinations of properties that depends on the human expertise (e.g. faraway-west).

Each object system has a particular set of properties: {locational, directional, topological, geometrical and logical}. This set of properties is important for spatial data manipulations and semantic data representation.

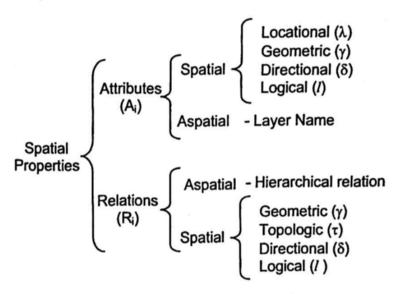


Figure 1. Taxonomy of spatial properties

The relations can be divided into two categories: intrinsic and extrinsic; intrinsic (R_I) represent the relationships among objects of the same spatial object system. On the other hand, extrinsic relations (R_E) describe the relationships among different spatial object systems.

$$R_{Si} = (R_I \cup R_E)_i$$
; $i = 1, ..., n$. (4)

Note that, properties denominated *semantic invariants* exist. These represent the essential properties of spatial patterns that do not change in scale. For instance:

- Altitude in which any hydrological network starts is greater than the altitude where it finishes.
- A population at the right side of the river.

The set of semantic invariants (I) is a subset of the set of properties.

$$I_{Si} \subset P_{Si}$$
; $i = 1, ..., n$ (5)

The map definition is "a conventionalized image of geographical reality, representing selected features or characteristics, resulting from the creative effort of its author's execution of choices, and thus designed for use when spatial relationships are of primary relevance" [7]. Considering this definition we can observe that characteristics and relations play a primordial role in map processing. Attributes and relations define the map

processing. Attributes and relations define the map structure. Any change, in the *characteristics* or *relations* modifies the map structure, this means that it is not the same map. While, the *content* of the set of spatial properties reflects the situation represented in a map, according to the map purpose.

The map semantics can be defined by the set of spatial properties, composed by *relationships* and *attributes*, which define the characteristics of the geographical information. That is the *spatial semantics* of a map is equivalent, in certain sense, to the content of set of PSO that has a map.

For instance, any topographic map has a particular group of spatial objects system with specific properties, these properties defines that it is topographic map, in other words, define its semantics. After any processing of map, it should preserve the essential properties.

3 MAP DESCRIPTION AS TOOL OF MAP GENERALIZATION

Description of spatial object systems represented in a map facilitates the task of finding some properties and relationships among objects. Certain properties are characteristics for certain spatial data layers and some properties are invariants in all scales. Additionally, it is possible to find relationships among different spatial data layers. Map Description represents a map using the relationships between the spatial objects and attempts to obtain the description as better as possible. The description should preserve the spatial semantics.

The map description allows representing properties of spatial data by means of operators. This proposal is focused on the identification of inconsistencies originated by generalization. The inconsistencies are relations that do not respect the semantic invariants. They represent an error in generalized map. These are identified by comparing map description prior to and after generalization.

3.1 MAP DESCRIPTION

The description process allows consideration of many spatial object systems. In nature, diverse configurations exist, wich contain objects of different types, as the case of a hydrological network. For this case, the objects (i.e., rivers) can be represented by lines or areas depending on scale. They are always connected with other elements of the network and related with other thematic.

To describe a topographic map, a set of operators can be used. Operators should involve each possible relationship (geometric, topological, directional, and logical). These operators should be consistent so that they can allow describing spatial object systems appropriately. In

addition, operators should be as basic as possible; at the same time, an operator should be applicable to the most of existing case of studies and use the invariants.

Certain spatial objects have more accurate than the others and they can be used as a means to initiate map objects description. In Table 1, we list some operators and their meaning and description.

Table 1. Map description operators

Operator	Meaning	Type	Domain		
			Турс	Values	
ArctoPnt (A,P)	Relation between arc A and point P	Geometrical	Numerical	[0CELL size]	
PlytoPnt (A,P)	Relation between Polygon A and point P	Topological	Boolean	{true, false}	
ArctoArc1 (B,A)	Relation between arc A and arc B	Topological	Nominal	{connected, not connected}	
ArctoArc2 (A,B)	Relation between arc A and arc B	Logical	Nominal	{Nearly perpendicular, Nearly parallel}	
ArctoPnt3 (A,B)	Relation between arc A and point B	Logical	Nominal	{N,NNW,NW, WNW,W,WSW, SW,SSW,S,SW, S,SSE,SE,ESE, E,NE,NE,NNE}	

Figure 2 shows a fragment of topographic map (1:50,000). The relationships between these objects are described by description operators. The operators applied for this map as follows (see Table 1 for the definition of operators):

```
Description Cell: 15
{
ArctoPnt1(LFOR, Mine2) = 800mu1,
ArctoPnt1(River8, Mine2) = 50mu,
ArctoArc1(LFOR, River 8) = Not connected,
ArctoArc2(LFOR, Highway1) = Nearly perpendicular,
ArctoArc2(River 8, Contour4) = Nearly perpendicular,
ArctoPnt3(River2, Population1) = N,
...
}
```

In this description, the group of operators is referred to an object that has the more exact coordinates. The Largest First Order River (LFOR) is obtained using the algorithm CLAJER [8]. The map description process is detailed explained in [9].

¹ Map units, e.g. meters.

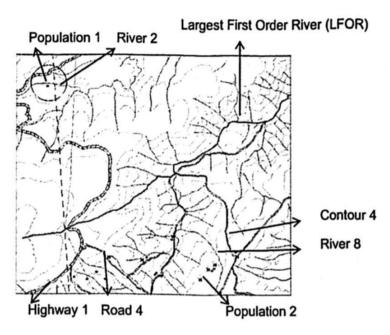


Figure 2. Fragment of topographic map

3.2 MAP GENERALIZATION

The core of the generalization process is based on our capacity to develop (and to master) measures able to describe geographical information as well as possible. These measures replace the analytical capacity of a cartographer, who detects that a situation is not correct and who is able to change it while preserving its main properties. Measures are used to detect conflicts as well as the information and properties that should be preserved during the process [2]. In this article the comparison of Map Description is used as a measure to detect and solve conflicts originated by the generalization.

During the generalization many properties of a map are manipulated principally by topological and geometrical operations, but some PSO should be preserved (invariants). The preserved properties should represent the particular aspects of the map structure (avoiding inconsistencies) in other words: the map semantics should be preserved.

3.3 GENERALIZATION BASED ON MAP DESCRIPTION

In this section, we describe a generalization method as tool of map description. This method consists of three functional stages: (1) Map Description, (2) Map Generalization and (3) Verification (Figure 3).

Map Description generates a description of relationships between the objects that compose a map by means of operators. (See Section 3.1)

Map Generalization is manipulated by Generalization Functions. These functions are composed by generalization operators [11].

Verification is the stage that evaluates the generalization process.

This method uses a source map (S_M) and a generalized map (G_M) . Description of the map source (D_S) and of generalized map (D_G) are generated. The Verification will be made comparing D_S and D_G (See Figure 3). Verification allows evaluating the structure of generalized map identifying inconsistencies. This will be used to feeding back parameters to Map Generalization. This process can be iteratively repeated until all inconsistencies are resolved. The target map (T_M) without inconsistencies indicates that the map's *semantics* is preserved after the generalization (See Figure 4).

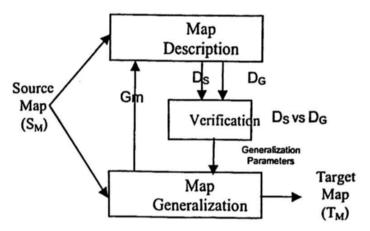


Figure 3. Generalization based on map description.

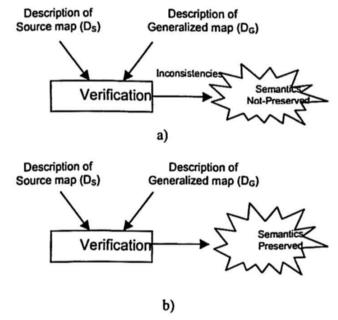


Figure 4. Verification to identify inconsistencies: a) Not satisfactory Generalization, b) Satisfactory Generalization

The results of verification that can be obtained are the following: Equal, Different and Equivalent (see Table 2).

Table 2. Results of Verification

Case	Verification (D _S vs D _G)	Meaning
Oa = Oa	Equal	Invariant respected
Oa ≠ Ob	Different	Inconsistency
Oa ≅ Ob	Equivalent	Invariant respected (in certain cases)

- Equal: This is obtained when the value of both operators is the same. It is representing that the invariant has been preserved (Figure 5).
- Different: This represent that the operator value is not the same in D_S and D_G. Then it is representing an inconsistency. The map structure is modified (See Figure 5b).
- Equivalent: This represent that the operator value is not the same in D_S and D_G, however in certain cases of study it is acceptable. The equivalent values not represent an inconistencies since they do not break the structure of the map. For instance; Figure 5c shows that the value ArctoPnt1 = NNE is accepted as equivalent. However, in Figure 5d, ArctoPnt1 = NW and it is not accepted as equivalent value. This situation represents an inconsistency. To define the equivalence, it is necessary to define threshold values. These values depend on the context and the subject domain.

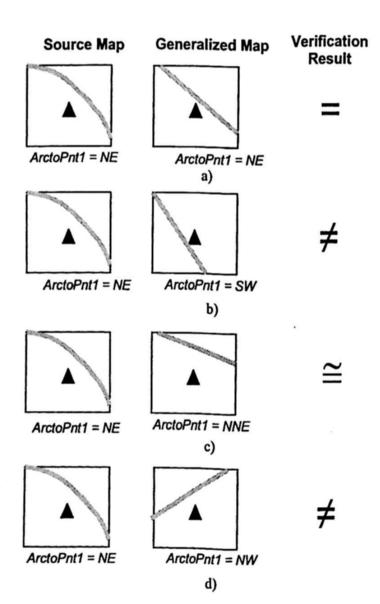


Figure 5. Results of verification: a) Equal, b) Different, c) Equivalent d) Different.

The identified inconsistencies are solved using local generalization parameters. These parameters will be matching to the specific situation. For instance; applying different tolerances in diverse parts of the arc, in other words, they are not necessarily the global parameters and the criteria of generalization, but are useful.

This method is useful because if parameters appropriately parameterize functions, inconsistencies generated by the generalization process can be resolved. For instance; in topographic data subject domain, one of possible set of parameters can be arc curvatures, number of vertices, arc intersections, etc.

4. PRELIMINARY RESULTS

Figures 6 and 7 show the simulations of generalization for hydrological network using as source the topographic map presented in Figure 2. In Figure 6, we use the generalization method developed in [8]. It can clearly be observed that in the original data, "Population 1" (surrounded by a circle) is at the north of "River 2" (Figure 2). In Figure 6 after the generalization this "Population" is at the south of the river, which is an inconsistency. The inconsistency is solved (Figure 7), because now the population is at the correct position with respect to the river.

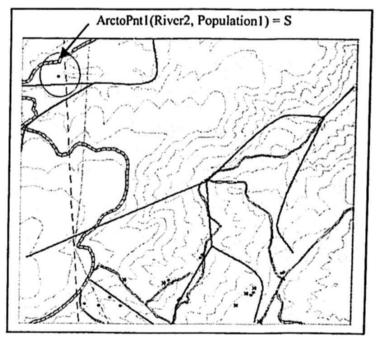


Figure 6. Generalization with inconsistencies

To correct the map inconsistency we used the position of the point (population) with respect to the river. On the other hand, a wide linear object was ill-parameterized and its shape is not appropriate to the scale (Figure 6). We corrected this inconsistency, using the curvature of the corresponding arcs (geometrical properties). The generalized object now seems more appropriate to the map scale (Figure 7). Thus, the data have more aesthetical quality.

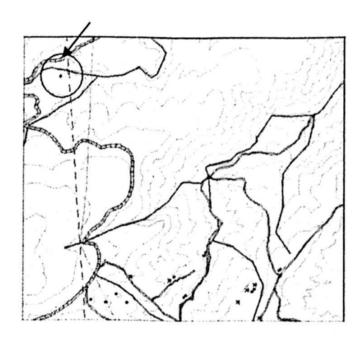


Figure 7. Generalization without inconsistencies

6. CONCLUSIONS

Automatic generalization is a Geomatics area that has attracted much attention during last few decades. A method herein presented provides an alternative to identify and solve certain conflicts in map generalization. Relationships between spatial objects allow solving inconsistencies generated by the generalization process. Map description enlarges the field of action of generalization operators because they can be parameterized in an appropriate manner, depending on the applications.

The main characteristics that provide the methodology herein presented are:

- 1. This proposal attempts to detect the map semantics by means of the elimination of inconsistencies.
- Data analysis on the whole, not isolated data layers, but generalizing spatial object system that interact with other spatial object system;
- 3. Use of techniques that allow evaluating data after generalization;
- Use of Properties of Spatial Object to describe spatial systems;
- Auxiliary to parameterize functions that manipulate spatial data, in particular generalization functions;
- Possibility to modify and incorporate new functions and properties.

The challenging tasks in generalization are to solve inconsistencies generated during the generalization process. Thus, it is necessary to carry out a methodology of general application to the great amount of cases of studies and to preserve inherent *spatial semantics* for each spatial object system. Developing our proposal, it is possible to solve partially some problems originated from the generalization.

ACKNOWLEDGMENTS

The authors of the present work wish to thank to the National Polytechnic Institute, the General Coordination of Studies of Postgraduate and Investigation (CGPI) and National Council of Science and Technology (CONACYT) for the support offered for development of this research.

REFERENCES

- [1] B. Berendt, T. Barkowsky, Freksa, C., Kelter, S. Spatial Representation with Aspect Maps. In: C. Freksa, C. Habel, K. F. Wender (Eds.), Spatial Cognition An Interdisciplinary Approach to Representing and Processing Spatial Knowledge. (Ed.: Springer-Verlag, 1998).
- [2] A. Ruas, Spatial Analysis and Agent Principles to Automate Generalisation Process. In: Levachkine, S.,Ruas, A., Bodansky, E. (eds.), eProceedings of the International Workshop on Semantic Processing of Spatial Data (GEOPRO 2002), 33-4 December 2002, Mexico City, Mexico (2002) (CD ISBN: 970-18-8521-X).
- [3] S. Lamy, A. Ruas, Y. Demazeu, M. Jackson, W. Mackaness, and R. Weibel, The Application of Agents in Automated Map Generalisation, *Proceedings of 19th International Cartographic Conference*, Ottawa, Canada (1999). http://agent.ign.fr/public/ica/paper.pdf
- [4] A. Ruas, Modele de generalization de donnees geographiques a base de contraints et d'autonomie, Ph.D. thesis, Universite de Marne-la-Vallee, 1999. (in French) ftp://ftp.ign.fr/ign/COGIT/THESES/
- [5] M. Galanda, Modelling Constraints for Polygon Generalization, International Cartographic Association (ICA), Map Generalization Workshop Paris 2003, Paris, France (2003).
- http://www.geo.unizh.ch/ICA/docs/paris2003/papers03.html
- [6] R. Weibel and G. Dutton, Constraint-based Automated Map Generalization, *Proceedings 8th International Symposium on Spatial Data Handling*, Vancouver, Canada (1998) 214–224.
- [7] T. Kanakubo, 'On the definition of "map", "cartography" and "cartographer" by Dr Christopher Board, chairman of the ICA working group on cartographic definitions', *Map-Journal of the Japan Cartographers Association*, (33) 3, 1993, 27-28.
- [8] M. Moreno, La Generalización Automática de la Información Geográfica Multiescala, M.S. Thesis, Centre for Computing Research, National Polytechnic Institute, Mexico City, Mexico, 2001. http://geo.cic.ipn.mx (in Spanish).
- [9] M. Moreno, Map Description as Intelligent Tool of Generalization. In: Levachkine, S.,Ruas, A., Bodansky, E. (eds.), eProceedings of the International Workshop on Semantic Processing of Spatial Data (GEOPRO 2002), 33-4 December 2002, Mexico City, Mexico (2002) (CD ISBN: 970-18-8521-X).
- [10] M. Bobzien and D. Morgenstern, Abstracting and Formalizing Model Generalization, International Cartographic Association (ICA), Map Generalization Workshop Paris 2003.

No.

after ald property of the legal of the second

Authorities of the relationship in the control of t

rotty borb, sprak look kinns in new con-

A the standard modernia is to be to be a first or and the standard modernia is to be a second or an analysis of the standard modernia is the stan

between afficially it consists in the con-

AUTOMATIC GEO-IMAGE INTERPRETATION

MULTI-RESOLUTION, OBJECT-ORIENTED FUZZY ANALYSIS OF REMOTE SENSING DATA FOR GIS READY INFORMATION

Ursula C. BENZ, Peter HOFMANN, Gregor WILLHAUCK, Iris LINGENFELDER, Markus HEYNEN
Definiens Imaging GmbH
Trappentreustr. 1, Munich, Germany

ABSTRACT

Remote sensing from airborne and space borne platforms provide valuable data for mapping, environmental monitoring, disaster management and civil and military intelligence. However, to explore the full value of these data, the appropriate information has to be extracted and presented in standard format to import it into geo-information systems and thus allow efficient decision processes.

Object-oriented analysis can contribute to powerful automatic and semi-automatic analysis for most remote sensing applications. Synergetic use to pixel based or statistical signal processing methods explores the rich information contents.

Here, we explain principle strategies of object-oriented analysis, discuss how the combination with fuzzy methods allows implementing expert knowledge and describe a representative example for the proposed workflow from remote sensing imagery to GIS.

The strategies are demonstrated using the first objectoriented image analysis software on the market, eCognition, which provides an appropriate link between remote sensing imagery and GIS.

KEY WORDS

Object-oriented image analysis, Remote sensing, multiresolution segmentation, fuzzy classification, GIS

1. INTRODUCTION

Remote sensing imagery of a large variety of spaceborne and airborne sensor provides a huge amount of data about our earth surface for global and detailed analysis, change detection and monitoring. Powerful signal processing methods are developed to explore the hidden information in advanced sensor data [12][21][47][50][53], e.g. for hyperspectral or high-resolution polarimetric SAR data [11][12].

However, all these algorithms suffer from the problem that no contextual information can be taken into account and aggregation of information from different resources is not very well supported. Additionally, results are mostly presented in raster format and are not well suited to incorporate into vector based geo-information-systems.

Thus there is a large gap between theoretically available information in remote sensing imagery and extracted and used information to support decision-making process in geomatics.

We propose a new strategy to bridge this gap.

Our approach focuses on

- the extension of the signal processing approach for image analysis by exploration of a hierarchical image object network to represent the strongly linked real world objects.
- usage of polygons for suitable interface to GIS (e.g. with .shp files to ArcView)
- fuzzy systems for improved and robust modeling of real world dependencies and detailed quality check of information product
- sensor and information fusion to use all available synergies

In the following we describe basic concepts of our approach, where some parts are taken from eCognition's UserGuide [56] explain some examples and discuss the enhanced possibilities due to the fuzzy classification.

1 Overview: Move from data analysis to image understanding using an hierarchical object network

In contrast to traditional image processing methods, the basic processing units of object oriented image analysis are image objects or segments, and not single pixels. Even the classification acts on image objects. One motivation for the object-oriented approach is the fact that the expected result of many image analysis tasks is the extraction of real world objects, proper in shape and proper in classification rather than single pixel

classification. This expectation cannot be fulfilled by common, pixel-based approaches.

Since at least two decades processing power of affordable computers allows image processing and image segmentation. First major advantages of object-oriented analysis were derived in studies for sea-ice analysis [13], object-oriented image matching [22] and certain approaches for data compaction [18].

The first general object-oriented image analysis software on the market was eCognition [56]. This software product was produced by DEFINIENS. Although eCognition is of course a specific combination of different contributing procedures, there are some basic characteristic aspects of the underlying object oriented approach, independent of the particular methods.

Directly connected to the representation of image information by means of objects is the networking of these image objects. Whereas the topological relation of single, adjacent pixels is given implicitly by the raster, the association of adjacent image objects must be explicitly worked out, in order to address neighbor objects. In consequence, the resulting topological network has a big advantage as it allows the efficient propagation of many different kinds of relational information.

Each classification task addresses a certain scale. Thus, it is important that the average resolution of image objects can be adapted to the scale of interest. Image information can be represented in different scales based on the average size of image objects. The same imagery can be segmented into smaller or larger objects, with considerable impact on practically all information, which can be derived from image objects. Thus, specific scale information is accessible.

Furthermore, it is possible to represent image information in different scales simultaneously by different object layers. Bringing different object layers in relation to each other can contribute to the extraction of further valuable information.

This can be achieved, for instance, by a hierarchical networking and representation of image objects. Besides its neighbors, each object also knows its sub-objects and super-objects in such a strict hierarchical structure. This allows precise analysis of the substructures of a specific region. Furthermore, based on sub-objects, the shape of super-objects can be changed.

Image analyses can be separated in two steps, sensor specific analysis of object primitives and scene specific analysis based on the detected and recognized object primitives.

This decoupling makes image analysis very flexible. Remote sensing experts develop sensor specific methods to extract certain kinds of objects primitives from sensor data, e.g. trees for buildings. This information is available in object features and attributes and can be combined in the subsequent scene dependent processing:

As soon as trees or buildings are identified, general knowledge can be applied, e.g. the expert knowledge of a forester, who does no more need to have specific remote sensing knowledge.

That means in principal everything, which is done from that moment, an object primitive is identified as part of a tree or tree, this object and its networked environment can be analyses with a *forest* logic. Instead of processing all areas of an image with the same algorithms, a differentiated procedure can be much more appropriate, very similar to the different experts in manual image interpretation: The special knowledge of urban planners or forest engineers is used for the dedicated analysis tasks to get appropriate results. To enable this *localized usage* of expert algorithms is a specific strength of object oriented image analysis.

Characteristic for the object oriented approach is, finally, a circular interplay between processing and classifying image objects. Based on segmentation, scale and shape of image objects, specific information is available for classification. In turn, based on classification, specific processing algorithms can be activated for subsequent segmentation and to refine classification and object creation. In many applications the desired geoinformation and objects of interest are extracted step by step, by iterative loops of classifying and processing. Thereby, image objects as processing units can continuously change their shape, classification and mutual relations.

Similar to human image understanding processes, this kind of circular processing results in a sequence of intermediate states, with an increasing differentiation of classification and an increasing abstraction of the original image information. On each step of abstraction, new information and new knowledge is generated and can be used beneficially for the next analysis step. Thereby, the abstraction not only concerns shape and size of image objects, but also their semantics. It is interesting that the result of such a circular process is by far not only a spatial aggregation of pixels to image regions, but also a spatial and semantic structuring of the image's content. Whereas the first steps are more data driven, more and more knowledge and semantic differentiation can be applied in later steps. The resulting network of classified image objects can be seen as a spatial, semantic network. The image analysis basing on a hierarchical object network leads from pure data analysis to scene understanding.

2 KNOWLEDGE BASED IMAGE INTERPRETATION

The design of successful image analysis methods requires knowledge about the underlying processes. The better the knowledge about the process and the better this knowledge can be implemented in the system, the more useful the extracted information will be.

Main characteristics of the process are

1) understanding of the sensor characteristics

- understanding of appropriate analysis scales and their combination
- identification of typical context and hierarchical dependencies
- considering inherent uncertainties of the whole information extracting systems, starting with the sensor, up to vague concepts for the requested information.

There a numerous publications on methods with regard to appropriate analysis of sensor data with signal processing or statistical methods (item 1 of the above list). Here, we focus on the possibilities of the object-oriented approach and fuzzy analysis of the object network to enable implementation of knowledge about items 2,3 and 4.

2.1 SELECTION AND COMBINATIONS OF SCALES

Scale is a crucial aspect of image understanding. Although in the domain of remote sensing a certain scale is always presumed by pixel resolution, the desired objects of interest often have their own inherent scale. Scale determines the occurrence or nonoccurrence of a certain object class. The same type of objects appears differently at different scales. Vice versa, the classification task and the respective objects of interest directly determine a particular scale of interest.

There is an important difference between scale and resolution: as resolution commonly expresses the average size of area a pixel covers on the ground, scale describes the magnitude or the level of abstraction on which a certain phenomenon can be described. Thus, studying an image from different levels of scale instead of an analysis approach based on different resolutions is an adequate approach to understand relations within an image and interpret the scene more easily.

The following describes the multi-scale concept for analysis of an image with depicts an urban area, e.g. by high-resolution satellite sensor as Ikonos.

Looking from a close distance on the image, one can detect and recognize single houses, buildings, roads and other urban objects. If one enlarges the viewing distance, one cannot discover single buildings, but rather different settlement types or even quarters. These areas typically can be distinguished by different textures and by different size and shape, too. The quarter's texture comprises its objects and structures on a smaller scale—houses, roads, gardens etc.—and it is especially determined by their tone, shape, and also by their topological relationships.

At a larger distance one might discover the city area as a one single entity and some surrounding agricultural areas and / or forests.

This example describes a 3 scale level approach:

- 1) trees, buildings and roads
- groups of trees and groups of buildings aggregated to different settlement types
- 3) forest and urban area

Between these scales there is a hierarchical dependency. abstracting houses, buildings, roads and other objects, one obtains settlement areas or even quarters. The aggregation of several settlement areas yields a town. Ecosystems show analogous patterns: combining several trees builds a group of trees and combining more trees or groups of trees builds a forest. Forests and towns have a similar abstraction level. Both are of comparable scale and both are of high semantic abstraction. The hierarchical scale dependencies between the affected object classes are obvious: quarters are substructures of cities, and houses are substructures of quarters.

These hierarchical scale dependencies are implicitly selfevident in each observation and description of real world structures. However, reflecting, and especially explicit representation of these patterns adds valuable information to automated image understanding methods.

Houses in an urban area can be treated in a different way than single houses in forests, for instance; different characteristics are of interest. Thus, in order to analyze an image successfully it is necessary to represent its content on several scales simultaneously and to explore the hierarchical scale dependencies among the resulting objects.

It is obvious that these relationships and dependencies cannot be analyzed by just changing the resolution of the imagery. This would, moreover, lead to the loss of a lot of useful information.

2.2 IMAGE SEMANTICS - MUTUAL RELATIONS BETWEEN IMAGE OBJECTS

One of the most important aspects of understanding imagery is information about context. There are two types of contextual information: global context, which describes the situation of the image—basically, time, sensor and location—and local context, which describes the mutual relationships or the mutual meaning of image regions. It is obvious that the processing of context information is always consciously or subconsciously present in human perception and contributes essentially to its great capabilities.

In order to receive meaningful context information, image regions of the right scale must be brought into relation. This scale is given by the combination of classification task and the resolution of the image data. Imagine for instance the classification task to identify parks in very high-resolution imagery. A park is always a large, contiguous vegetated area. This different scale distinguishes parks from gardens. Additionally, parks are distinguished from pastures, for example by their embedding in urban areas. Single neighboring buildings are not a sufficient condition to describe parks. However, their neighborhood to single buildings is a suitable criterion for distinguishing gardens from pasture.

This simple example already shows how much the available context information depends on the scale of the structures, which are brought into relation. This astonishing fact explains why it is so difficult or even impossible to describe meaningful context relations using pixel-based approaches. Only representing image information based on image objects of the appropriate scale enables one to handle image semantics. Additionally, in order to make image objects aware of their spatial context it is necessary to link them. Thus, a topological network is created.

This network becomes hierarchical when image objects of different scale at the same location are linked. Now each object knows its neighbors, its sub- and super-objects. This additionally allows a description of hierarchical scale dependencies. Together with classification and mutual dependencies between objects and classes, such a network can be seen as a spatial semantic network.

The fact that image understanding always means dealing with image semantics was until now not sufficiently covered by the capacity of digital image analysis, especially in the field of remote sensing.

2.3 INHERENT UNCERTAINTIES AND VAGUENESS

Various types of uncertainty influence information extraction from remote sensing data. First of all, there are many factors which influence the processes of data acquisition, data processing and image generation, and which differ from scene to scene, even if the data comes from the same sensor. A very basic, inherent problem of earth observation data is that land cover can look different, depending on the season, time of day, light conditions and weather. Furthermore, the same type of objects appears highly differently depending on the sensor type, sensing geometry and the resolution.

The dependency between features and land cover or land use is mostly only roughly modeled and vagueness is inherent even in the concepts of land cover and land use. Sensor measurements—the basic source for image pixels—have limited radiometric resolution even after careful calibration of the instrument. The geometric resolution in remote sensing—and in any data acquisition process—is limited as well. This effect leads to class mixture within one resolution cell: if a resolution cell covers water-land transition, the relevant pixel represents to some degree water and to some degree the land cover of the shore area.

The image generation process converts sensor measurements to image data. Additionally, these data have to be compressed to reduce requirements for archiving and data transmission. In most cases, these data processing steps cause artifacts and ambiguities, which lead to noise and therefore to additional uncertainty in the final image data.

Usually only vague concepts exist for land cover and land use. There is no exact threshold between densely and sparsely populated area, or between low and high vegetation. Whenever thresholds are defined in terms of numbers, they are mostly unsatisfactory idealizations of the real world and therefore subsequently lead to problems during classification and performance estimation of the classification.

Information retrieval from remote sensing databases is based to a large extent on vague knowledge. Especially important context information is typically only expressed in terms of vague linguistic rules. For example, if trees are "nearly completely" surrounded by urban area, they are assigned to the class park.

Furthermore, in many cases the desired information for a specific classification task is not, or not sufficiently, contained in the available image data. This can be caused by spatial or radiometric resolution, because the signal to noise ratio is too low, or simply because the sensor does not deliver unambiguous signal for the desired class separation.

If these uncertainties are not taken into account in information extraction, classification will not be robust and transferable. There are several approaches, so called soft classifiers, which take these uncertainties into account. One of the most powerful soft classifiers are fuzzy classification systems, which are able to incorporate in the analysis approach from the very beginning on inaccurate sensor measurements, vague class descriptions and unprecise modeling. Degree of uncertainty is part of the fuzzy classification result [3][5][6][37][56].

3 IMAGE OBJECTS

The basic elements of an object-oriented approach are of coarse image objects. Image objects describe certain contiguous regions in an image. We distinguish between image objects primitives and objects of interest. Only object of interest match real world object, e.g. the building footprints or whole agricultural parcels. Object primitives are usually the necessary intermediate step before objects of interest can be found by segmentation and classification process. The smallest image object is one pixel.

Image objects can be linked to a hierarchical network, where they carry a high-dimensional feature space:

Image object statistics and texture

Within an image object all kind of statistics based on single input layers or combinations within the input image layer stack can be computed, e.g. the mean value of the ratio of two input channels A and B.

$$f_{r_{-}AB} = \frac{\frac{1}{n} \sum_{n} p_{A}(x_{n})}{\frac{1}{n} \sum_{n} p_{B}(x_{n})}$$

with n number of pixels x within object; p(x) value of pixel at location x.

Using image objects to calculate this statistic instead of boxes of pixels allows reliable statistic without smearing edges, since objects do not exceed edges.

Image object shape

The closer objects primitives are to objects of interest, the more image object shape features can be used as additional uncorrelated object features. Since they are usually independent from sensor characteristics they are robust versus sensor calibration and illumination conditions.

Typical examples are area of objects, their length-towidth ratio, number of straight borders and many more. Advanced shape features can be derived from object polygons and object skeletons, the latter describing the inner structure of an object.

Image objects statistics, texture (e.g. Haralick features [20]), shape can be regarded as intrinsic features. They are available for each object.

Topological object features

Due to the object network context features are provided. Within one scale relations to neighbored objects can be evaluated, whereby the size of the neighborhood can be defined as parameter. Between image scales hierarchical relations can be explored, where the distance of scales can be adjusted using distance scale parameter.

This hierarchical object network enables addressing of context and semantic for subsequent image analysis with innovative techniques:

The hierarchical network provides additional object features:

- Characterization of an image object based on its sub-objects using
 - a. Texture analysis based on sub-objects, classifying attributes of all sub-objects of an image object on average. Attributes can for instance be contrast or shape.
 - b. Line analysis based on sub-objects.
 - c. Class-related features: relationships to classified sub-objects, such as the relative area of other image objects assigned to a certain class, e.g. if an urban area on higher level contains many sub-objects classified as houses, this urban area can be described as dense vs. other less dense areas.

Characterization of an image object based on its superobjects, e.g. houses belonging to a super object urban can be classified as urban houses, whereas houses in rural areas can be classified as cottages or special buildings.

Semantic features

These higher order features are available after a first classification of image objects. They allow describing a park as forested region within urban area or shore regions as adjacent land regions to water. These very features are important for advanced remote sensing image analysis. The reduce ambiguities, allow land use classification in addition to pure land cover classification and thus lead to a first step of scene understanding

In the following we describe image object creation process.

3.1 CREATION OF IMAGE OBJECTS

Objects are created by image segmentation, which is the subdivision of an image into separated regions. Image segmentation is a long year research topic in the area of image analysis [25][35][36][46]. In almost all cases, segmentation is an optimization process. Regions of minimum heterogeneity given certain constraints have to be found. Criteria for heterogeneity, definition of constraints and strategy for sequence of aggregation determine the final segmentation result.

Heterogeneity can refer to primary object features, such as standard deviation or gray tones, shape of object, or texture on objects or on higher order object features, such as class assignment of objects. Segmentation methods using heterogeneity definition relying only on primary object features can usually only deliver object primitives, without a direct relationship to real world objects. However, these object primitives can be assigned to classes during a first classification step and then the higher order object feature "class assignment" is available for classification based segmentation [45]. This advanced segmentation step is able to create objects of interest with unique relation to the depicted real world.

Segmentation in eCognition ([1][2]) allows as well segmentation based on primary features (gray tone and shape) and – after an initial classification - the more advanced classification based segmentation. The method leads for most data distributions to robust results and is applicable under many conditions. Constraints can be used to ensure exact reproducibility of segmentation.

As already mentioned in the chapter 2.1 image scale is very important for meaningful analysis. Therefore eCognition provides segmentation on several scales. Scale selection is performed using certain parameters.

Details to eCognition's object creation approach are provided in the following chapters.

3.1.1 OBJECT CREATION IN ECOGNITION

eCognition's multi-resolution segmentation is a bottom up region-merging technique starting with one-pixel objects. In numerous subsequent steps, smaller image objects are merged into bigger ones. Throughout this pairwise clustering process, the underlying optimization procedure minimizes the weighted heterogeneity $n \cdot h$ of resulting image objects, where n is the size of a segment and h an arbitrary definition of heterogeneity. In each step, that pair of adjacent image objects is merged which stands for the smallest growth of the defined heterogeneity. If the smallest growth exceeds the threshold defined by the scale parameter, the process stops. Doing so, multiresolution segmentation is a local optimization procedure.

To achieve adjacent image objects of similar size and thus of comparable quality, the procedure simulates the even and simultaneous growth of segments over a scene in each step and also for the final result. Thus, the procedure starts at any point in the image with one-pixel objects. A treatment sequence based on a binary counter guarantees a regular spatial distribution of treated objects. However, for obvious reasons, such a sequence contains a stochastic, historical element.

To ensure reproducibility on same image, constraints can be used to force an exact reproducible segmentation. Here a global optimization criterion is used.

3.1.1.1 DEFINITION OF HETEROGENEITY

Heterogeneity in eCognition considers as primary object features color and shape. For fusion decision not the heterogeneity itself is important, but the increase of heterogeneity. This increase has to be less than a certain threshold.

$$f = w_{color} \cdot \Delta h_{color} + w_{shape} \cdot \Delta h_{shape},$$

$$w_{color} \in [0,1,], w_{shape} \in [0,1,]$$

$$w_{color} + w_{shape} = 1$$

The weight parameters allow to adapt heterogeneity definition to the application.

The spectral heterogeneity allows multi-variate segmentation giving to the image channels c certain weight w_c . Difference spectral heterogeneity is defined as following.

$$\Delta h_{color} = \sum_{c} w_{c} \left(n_{merge} \cdot \sigma_{c,merge} - \left(n_{obj_1} \cdot \sigma_{c,obj_1} + n_{obj_2} \cdot \sigma_{c,obj_2} \right) \right)$$
 with n_{merge} number of pixels within merged object, n_{obj_1} number of pixels in object 1, n_{obj_2} number of pixels in object 2, σ_{c} standard deviation within object of channel c. Subscripts $merge$, obj_1 and obj_2 refer to the

merged object, object 1 and object 2 prior to merge, respectively.

The shape heterogeneity is a value that describes the

improvement of the shape with regard to smoothness and compactness of object's shape.

 $\Delta h_{shape} = w_{compct} \cdot \Delta h_{compt} + w_{smooth} \cdot \Delta h_{smooth}$

with

$$\Delta h_{smooth} = n_{merge} \cdot \frac{l_{merge}}{b_{merge}} - \left(n_{obj_1} \cdot \frac{l_{obj_1}}{b_{obj_1}} + n_{obj_2} \cdot \frac{l_{obj_2}}{b_{obj_2}}\right)$$

$$\Delta h_{cmpct} = n_{merge} \cdot \frac{l_{merge}}{\sqrt{n_{merge}}} - \left(n_{obj_1} \cdot \frac{l_{obj_1}}{\sqrt{n_{obj_1}}} + n_{obj_2} \cdot \frac{l_{obj_2}}{\sqrt{n_{obj_1}}}\right)$$

with n_{merge} number of pixels within merged object, n_{obj} number of pixels in object 1, n_{obj} number of pixels in object 2, l_{merge} perimeter of merged object, l_{obj} perimeter of object 1, l_{obj} perimeter of object 2, b_{merge} perimeter of the merged object's bounding box, b_{obj} perimeter of object 1 bounding box, b_{obj} perimeter of object 2 bounding box.

Thus, the smoothness heterogeneity equals the ratio of the de facto border length l and the shortest possible border length b given by the bounding box of an image object parallel to the raster.

The compactness heterogeneity equals the ratio of the de facto border length *l* and the square root of the number of pixels forming this image object.

The weights w_c , w_{color} , w_{shape} , w_{smooth} , w_{cmpt} are parameters, which can be selected in order to get for a certain image data stack and a considered application suitable segmentation results.

The scale parameter is the stop criterion for optimization process. Prior to the fusion of two adjacent objects, the resulting increase of heterogeneity f is calculated. If this resulting increase would exceed a threshold t determined by the scale parameter, $t = \Psi(scale)$ then no fusion takes place and segmentation stops.

 $f > \Psi(scale)$ then segment fusion is discarded $f \leq \Psi(scale)$ then segment fusion takes place

If no more objects can be fused according to this constraint, segmentation stops.

The larger the scale parameter, the more objects can be fused and the larger the objects grow. Details are to be found in [2].



Figure 1: Exemplary segmentation result in eCognition

3.1.1.2 ALTERNATIVE CREATION MODES IN ECOGNITION FOR OBJECT PRIMITIVES

Appropriate object creation with respect to different applications may require alternative approaches to the described standard implementation in eCognition. Therefore external segmentation results can be inserted into eCognition, e.g. using thematic layers on in the latest version of eCognition alternative methods are available.

Segmentation according to the spectral difference of objects

Using the mode "spectral difference" large homogeneous areas can be created regarding spectral difference. Areas, which have a lower spectral difference than a certain threshold are merged. The scale parameter determines the threshold of the spectral difference of neighboring objects, below which they are merged.

Segmentation of sub-objects for the purpose of line analysis

Object oriented line analysis of image objects can be performed using a special mode of segmentation, which is available due to the object hierarchy. This mode uses only heterogeneity compactness. The scale parameter—here ranging from 0.5 to 1—determines the maximum relative border length of sub-objects to neighbors, which are not sub-objects of the same superior object.

For the analysis of image objects such as in Figure 2 the specific image object level can be sub-segmented. The results are compact sub-objects, which guarantee a minimum and maximum border length to the outer environment. Operating from center point to center point of these sub-objects means that it is possible to easily analyze the length of a curved line, average thickness, curvature etc.

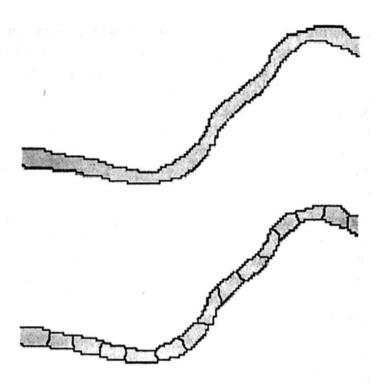


Figure 2: Linear structure subsegmented into compact objects

3.1.2 VALIDATION OF OBJECT CREATION PROCESS

Human interpretation and correction: Automatic segmentation replaces in the usual image interpretation workflow manual digitizing of polygons. Thus a strong and experienced source for the evaluation of segmentation techniques is of course the human expert. It can't be expected that automatic segmentation result will be fully convincing for human interpreter. Therefore, eCognition provides also possible manual interaction (manual section and manual fusion) to use automatic segmentation for the major segmentation part of the image and support subsequent correction by a human expert for small areas.

Automatic validation: There are several approaches to evaluate segmentation quality.

- a. Reference polygons (e.g. provided by manual digitizing) can be used to test the automatic segmentation. If the complete reference polygons is covered by automatically achieved segments, best scores are given,
 - if minimum number of segments are within the reference segment (lowest possible over segmentation)
 - if minimum area of segments outside of reference polygon is covered
- b. Strength of segment borders is analyzed. The higher the border between to segments, the less probable is their merging in the optimization process. Thus even certain variations of parameters or conditions will not change the segmentation result. Thus, the larger the number of strong borders relative to weak borders in an segmented image, the more stable and reproducible the segmentation will be for similar scenarios.

 reproducible the segmentation will be for similar scenarios.

3.2 HIERARCHICAL OBJECT NETWORK

Object creation is a complex and very often time-consuming process. However, it is in many cases of significant advantage for image analysis and geo-information production.

Objects carry much more information than single pixels or boxes of pixels do, and thus allow image analysis with higher accuracy. One example, in object-oriented analysis "rivers" can be distinguished from "lakes" by the different shape of these objects. This huge uncorrelated object feature space enables robust classification.

Linking objects to a hierarchical network provides further advantages.

3.2.1 HIERARCHICAL NETWORK CREATION IN ECOGNITION.

The different levels of image objects are generated by the described multi-resolution segmentation. In general, the higher the level and the larger the average size of image objects, the larger was the chosen scale parameter.

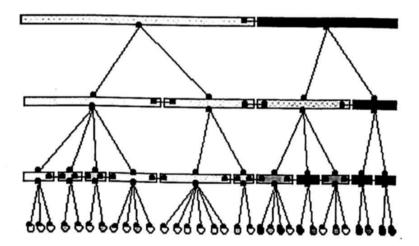


Figure 1: 4 level hierarchical network of image objects in abstract illustration

All segmentation procedures provided by eCognition operate on arbitrary levels in a strong hierarchical network. Since the level of pixels and the level of the whole image always exist by definition, each segmentation of a new level is a construction in between a lower and an upper level. To guarantee a definite hierarchy over the spatial shape of all objects the segmentation procedures follow two rules:

- Object borders must follow borders of objects on the next lower level.
- Segmentation is constrained by the border of the object on the next upper level.

Within cCognition

 Structures of different scales can be represented simultaneously and thus classified in relation to each other.

- Different hierarchical levels can be segmented based on different data; an upper layer, for instance can be built based on thematic land register information, whereas a lower layer is segmented using remote sensing data. Classifying the upper level, each land register object can be analyzed based on the composition of its classified sub-objects. By means of this technique different data types can be analyzed in relation to each other.
- Object shape correction based on regrouping of subobjects is possible

This network provides the bases for most powerful information extraction, because many successful strategies of human analyst can be approximated:Relations between scales and combinations of scales can be used, e.g. one could look based on the same image on trees, on groups of trees and on forest. Roads (extracted on low scale) leading through forest areas (extracted on high scale) can be classified as *forest roads*. Based on the width of the roads certain drivability can be assigned as feature in the output for a geo-information system. Furthermore, context information and semantic can be used to distinguish between trees within a forest or within an urban area.

3.3 CREATION OF VECTOR INFORMATION TO BRIDGE REMOTE SENSING AND GEO-INFORMATION SYSTEMS

Image objects not only enhance automatic classification of remote sensing imagery, they support also export of the extracted information to geo-information systems, since they can be easily converted to polygons using vectorization approaches. Within eCognition this vector structures are not only used for import and export, but also for advanced classification.

eCognition supports a simultaneous raster / vector representation of image objects. After segmentation, vectorization functionality allows production of polygons and skeletons for each image object. This vector information is produced in different resolutions for different purposes.

eCognition produces polygons along the pixel raster (Figure 4: polygons 1) or slightly abstracted polygons (Figure 4: polygons 2). The latter polygons are referred to in the following as base polygons. They are created with respect to the topological structure of image objects and are used for exporting vector information, too. More abstracted vector information represents the shape of image objects independently of the topological structure (Figure 4: polygons 3) and is used for the computation of shape features. Theses polygons are referred to as shape polygons.

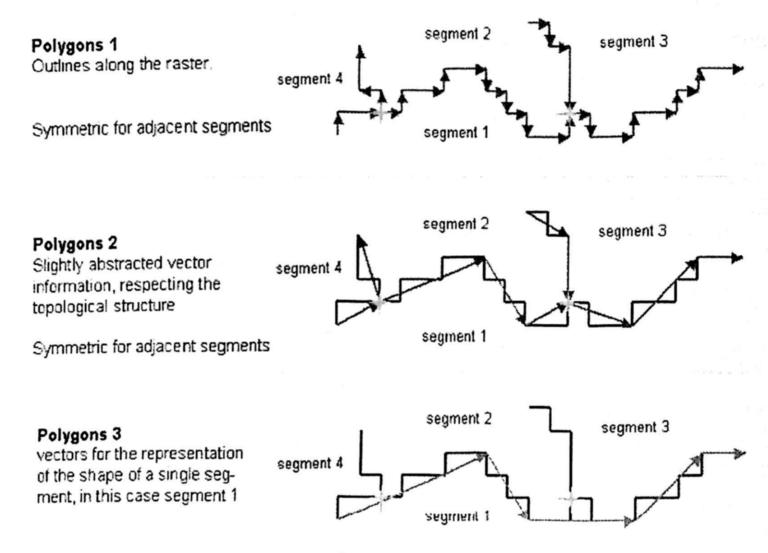


Figure 4: Different polygon types for the vectorization of segments /image objects.

The computation of base polygons is done by means of a Douglas Peucker [15] algorithm. The Douglas Peucker algorithm is one of the most common procedures for polygon extraction. It is a top-down approach, which starts with a given polygon line and divides it into smaller sections iteratively.

- Given the two end points of a polygon line—in eCognition typically these two starting points are topological points, see yellow marks in Figure—the algorithm detects this specific point on the polygon line with the largest vertical distance to a line connecting the two end points, see Figure.
- 2. At this detected point, the polygon line is cut into two shorter polygon lines, **Figure 5**b.
- 3. This procedure continues until the longest vertical distance is smaller than a given threshold, Figure 5c.

In other words: the threshold describes the strongest possible deviation of the polygon from the underlying raster. In eCognition, this threshold can be interactively defined to meet the needs of the application and is measured in pixel units.

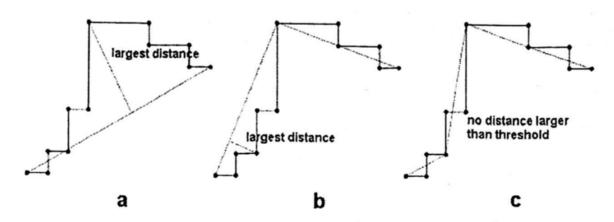


Figure 5: Douglas-Peucker algorithm: a) start configuration and detection of largest distance; b) new state after dividing into two sections; c) final result, no further division as no distance is larger than given threshold

The Douglas-Peucker algorithm in its pure application suffers in some cases in producing relatively acute angles. In order to improve the result, in eCognition angles smaller than 45 degrees are detected in a second run. From the two particular vectors at such an angle, that one is subdivided which will result in the largest angles. This procedure continues in iterative steps until there are no angles smaller than 45 degrees.

For high thresholds, which produce a strong abstraction from the original raster, slivers and intersections within and between base polygons can arise. This can be especially disturbing when these base polygons are used for export. In order to avoid this effect, an additional, optional algorithm detects intersections and fractionates the affected vectors.

The shape polygons are created by means of a derivative of multiresolution segmentation (3), in this case not applied to image regions but to single vectors. In contrast to the Douglas-Peucker algorithm this procedure is a bottom-up approach. Starting with base polygons, the single vectors are subsequently merged, optimizing a homogeneity criterion. It is important to understand that the heterogeneity of single shape vectors is defined as deviation of the underlying base vectors. Thus, a threshold of 0 will always produce shape polygons identical with the underlying base polygons. The resulting shape therefore depends also on the threshold of the base polygons. A threshold bigger then 0 will result in a abstraction than the base polygons. Concretely, the deviation is computed as the maximum of the difference of length between shape vector and underlying base vectors and the sum of the lengths of the vertical parts of the underlying base vectors to the shape vector. Iteratively, the two adjacent vectors of a polygon, which result in the smallest heterogeneity, are merged. This continues until the predefined threshold is reached.

3.3.1 OBJECT FEATURES BASED ON POLYGONS

The resulting shape polygons are independent of the topological structure and therefore specific for each single image object. A straight edge of a segment is represented as one vector, even if it contains a topological point

Thus, fractal parts of the boundary of an image object are represented in a characteristic way by a number of short vectors, whereas straight edges are represented by long edges. Based on these shape polygons meaningful shape features can be computed, e.g. the number of straight edges, average length of edges and maximum length of edges: Artificial targets are usually characterized by few long straight edges, whereas natural targets are more irregularly shaped.

Skeletons are advanced object features based on polygons. They describe the inner structure of an object, which provides new characteristics for an image object.

They provide centerline of objects and second and higher order branches. Thus instead of the street detected in remote sensing imagery just the center line of the street can be exported and later on by appropriate map production software assigned with the width of its importance and according to the defined map scale.

Skeletons then are created by identifying the mid points of the triangles in an image object and connecting them. Thereby the triangles' mid points are determined by connecting the midpoints of the triangles' sides.

To find skeleton branches, three types of triangles are created: branch-triangles (three-neighbor-triangle), connecting triangles (two-neighbor-triangles) and end-triangles (one-neighbor-triangles).

- Branch triangles indicate branch-points of the skeleton.
- two-neighbor-triangles indicate a connection point and
- end-triangles indicate end-points of the skeleton.

To obtain the skeletons, the generated points are connected. The longest possible connection of branch-points is defined as main line.

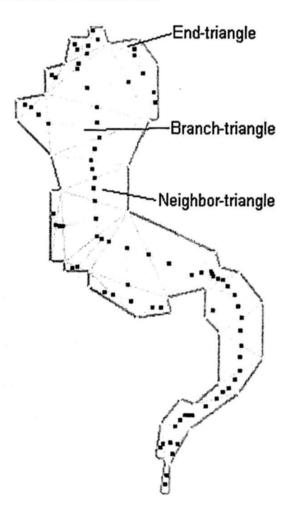


Figure 6: Creation of skeletons based upon a Delauney triangulation of image objects' shape polygons.

Skeletons describe the objects inner structure and shape in detail. Hence, they provide input for advanced automated shape correction. For example an automated object cut can be started to cut high order branches of a fractal object. Typical example is lower order streets connected to a main road. The automated object cut can be

can be started to cut high order branches of a fractal object. Typical example is lower order streets connected to a main road. The automated object cut can be understood as pruning the object's skeleton from outside to inside.

3.3.2 VECTOR FORMAT IMPORT AND EXPORT

Efficient import of extracted information into geoinformation systems is possible, because objects can be represented easily by polygons as shown in the previous chapter.

eCognition supports the import and export of thematic data in shape format. Since eCognition is a region based analysis system, only polygons are considered for import.

For internal usage this vector information is transformed to raster format. Export supports polygons, line and point information. While lines in vector format are based on the lines of the skeletons, points are equivalent to the centerpoint of the main line for each object to be exported.

4 FUZZY CLASSIFICATION

Fuzzy classification is beside neural networks [19] and probabilistic approaches [12] a very powerful soft classifier and typically represents and expert system for classification [53]. It takes into account

- uncertainty in sensor measurements
- parameter variations due to limited sensor calibration
- vague (linguistic) class descriptions
- class mixtures due to limited resolution.

Fuzzy classification consists of a n-dimensional tupel of membership degrees, which describes the degree of class assignment of the considered object *obj* to the *n* considered classes.

$$f_{class,obj} = \left[\mu_{class_1}(obj), \mu_{class_1}(obj), ... \mu_{class_n}(obj) \right]$$

Crisp classification would only give the information, which membership degree is the highest, whereas this tupel contains all information about the overall reliability, stability and class mixture.

Fuzzy classification requires a complete fuzzy system, consisting of fuzzification of input variables resulting in fuzzy sets, fuzzy logic combinations of these fuzzy sets and defuzzification of the fuzzy classification result to get the common crisp classification for map production.

Fuzzy logic is a multi-valued logic quantifying uncertain statements. The basic idea is to replace the two Boolean logical statements "true" and "false" by the continuous range of [0...1], where 0 means "false" and 1 means "true" and all values between 0 and 1 represent a transition between true and false. Avoiding arbitrary sharp thresholds, fuzzy logic is able to approximate real world

in its complexity much better than the simplifying Boolean systems do. They can take into account imprecise human thinking and can implement linguistic rules.

Hence, fuzzy classification systems are well suited to handle most sources of vagueness in remote sensing information extraction. The mentioned parameter and model uncertainties are considered by fuzzy sets, which are defined by membership functions.

Fuzzy systems consist of three main steps, fuzzification, fuzzy rule base and defuzzification, which are briefly described in the following.

4.1 FUZZIFICATION

Fuzzification describes the transition from a crisp system to a fuzzy system. It defines on an object feature certain sets, which assign for example all feature values in a certain range to a specific class, e.g. "low". The sets are defined using membership functions. They assign a membership degree between 0 and 1 to each (object) feature value with respect to the considered feature class. Depending on the shape of the function, the transition between "full member" and "no member" can be crisp (for a rectangular function) or fuzzy (see Figure 7, set M).

All feature values, which have a membership value higher than 0 belong to a fuzzy set. In general, the broader the membership function, the more vague the underlying concept; the lower the membership values, the more uncertain is the assignment of a certain value to the set.

Within the fuzzy system not feature values are combined but the fuzzy sets defined on these features values. Hence all calculations refer to membership degrees with the defined range between 0 and 1, independent of the dynamic of the originally crisp features. This simplifies working in a high-dimensional feature space with different dynamics and features of various types, e.g., backscatter from different sensors, geographic information, texture information and hierarchical relations.

For successful classification a deliberate choice of membership function is crucial. This is one of the most important steps to introduce expert knowledge into the system. The better the knowledge about the real system is modeled by the membership functions, the better the final classification result [9].

recovered in aging AM by Dastern and a residence in a per-

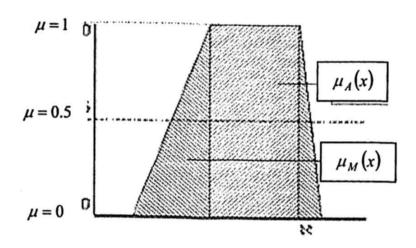


Figure 7: Rectangular and trapezoidal membership functions on feature x to define crisp set M (///), $\mu_M(x) \in \{0,1\}$ and fuzzy set A (\\\) $\mu_A(x) \in [0,1]$ over the feature range X;

It is possible to define more than one fuzzy set on one feature, e.g., to define the fuzzy sets *low*, *medium* and *high* for one object feature. The more the memberships overlap, the more objects are common in the fuzzy sets and the vaguer the final classification.

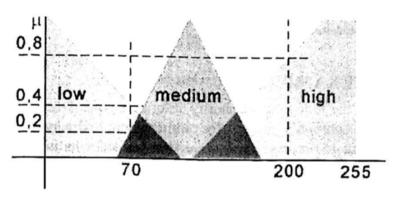


Figure 8: The membership functions on feature x define the fuzzy set *low*, *medium* and *high* for this feature.

Figure shows three fuzzy sets defined for the feature x: low, medium and high. They are characterized by overlapping triangular membership functions.

For an image object with a feature value of x = 70, the membership to class *low* is 0.4, to class *medium* is 0.2 and to class *high* is 0.0. If the feature value x equals 200, the membership to the classes is 0.0, 0.0, 0.8, respectively.

4.2 FUZZY RULE BASE

A fuzzy rule base is a combination of fuzzy rules, which combine different fuzzy sets. The simplest fuzzy rules are dependent on only one fuzzy set.

Fuzzy rules are "if – then" rules. If a condition is fulfilled, an action takes place. The following rule could be defined: "If" feature x is low, "then" the image object should be assigned to land cover W. In fuzzy terminology this would be written: If feature x is a member of fuzzy set low, then the image object is a member of land cover W. According to the definition in Figure 8, in case feature value x = 70, the membership to land cover W would be

0.4, in case x = 200, the membership to land cover W would be 0.

To create advanced fuzzy rules, fuzzy sets can be combined. An operator returns a fuzzy value that is derived from the combined fuzzy sets. How this value is derived depends on the operator. The logic operators are "and", "or" and "not". There are several possibilities to realize these operators. In most cases the simplest implementation is to use minimum operation to implement the fuzzy "and" and the maximum operation to implement fuzzy "or".

The results are very transparent and ensure independence of the sequence of logic combinations within the rule base (A "and" B gives the same result as B "and" A). In addition a hierarchic structure following common logic (e.g., A "or" (B "and" C) equals (A "or" B) "and" (A "or" C)) can be created easily.

A fuzzy rule base delivers a fuzzy classification, which consists of discrete return values for each of the considered output classes (see Figure 9). These values represent the degree of class assignment.

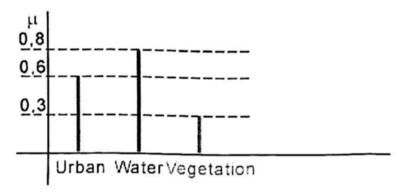


Figure 9: Fuzzy classification for the considered land cover classes urban, water and vegetation. The image object is a member of all classes to various degrees

$$\mu_{urban}(obj) = 0.6$$
, $\mu_{water}(obj) = 0.8$, $\mu_{vegetation}(obj) = 0.3$

Please consider that fuzzy classification gives a possibility for an object to belong to a class, while classification based on probability provides a probability to belong to a class. A possibility gives information on a distinct object. Probability relies on statistics and gives information on many objects. Whereas the probability of all possible events adds up to one, this is not necessarily true for possibilities.

The higher the membership degree for the most possible class, the more reliable is the assignment. In the example above, the membership to water $\mu_{water}(obj) = 0.8$ is rather high and in most applications this object would therefore be assigned to the class *Water*. The bigger the difference between highest and second highest membership value, the clearer and more stable the application. Classification stability and reliability can be calculated and visualized within eCognition as an advanced method for classification validation.

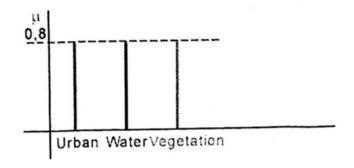


Figure 10: Fuzzy classification for the considered land cover classes *Urban*, *Water* and *Vegetation*.

In figure 10 the equal membership degrees indicate an unstable classification between these classes for the considered image object and can be a description of class mixture within the resolution cell or that class definition is not sufficient to distinguish between the classes. If the land cover classes can be distinguished on the data set for other objects, class mixture is very probable.

The high membership

$$\mu_{urban}(obj) = \mu_{water}(obj) = \mu_{vegetation}(obj) = 0.8$$

value shows that the assignment to this class mixture is reliable.

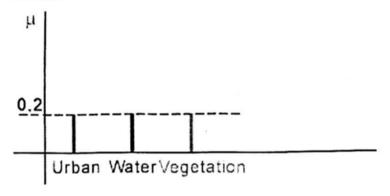


Figure 11: Fuzzy classification for the considered land cover classes *Urban*, *Water* and *Vegetation*.

The equal membership degrees in figure 11 again indicate an unstable classification between these classes for the considered image object, as in figure 10. However, the low membership value

$$\mu_{urban}(obj) = \mu_{water}(obj) = \mu_{vegetation}(obj) = 0.2$$

indicates a highly unreliable assignment. Assuming a threshold of a minimum membership degree of 0.3, no class assignment will be given in the final output.

This analysis of fuzzy classification provides an important input for classification validation but furthermore for information fusion in current and future remote sensing systems with multi-sensor sources and ancillary data. The reliability of class assignments for each sensor can be used to find the most possible and probable class assignment. A solution is possible, even if there are contradictory class assignments based on different sensor data, e.g., optical sensors are regarded as being less reliable than radar sensors if there is a heavy fog.

4.3 DEFUZZIFICATION

To produce results like maps for standard land cover and land use applications, the fuzzy results have to be translated back to a crisp value, which means that an object is either assigned to a class or not. For defuzzification of classification results the class with the highest membership degree is chosen.

Defuzzification is the reverse process of fuzzification. It delivers a crisp classification. If the membership degree of a class is below a certain value, no classification is performed to ensure minimum reliability.

As this output discards the rich measures of uncertainty of the fuzzy classification, this step should be the final step in the whole information extraction process.

$$f_{crisp} = \max \left\{ \mu_{class_1}(obj), \mu_{class_2}(obj), ..., \mu_{class_n}(obj), \right\}$$

with class assignment equal to the class i with the highest membership.

Further information on fuzzy systems in image analysis and remote sensing can be found in the following references [5][6][26][37].

5 EXAMPLE

In the following we shortly describe a typical example for eCognition's usage for information extraction from remote sensing imagery to update geo-information.

The goal of this example was to analyze a mosaic of highresolution aerial orthoimages of FMM, an Austrian company. Input files were the image mosaic and shape files showing building footprints and roads. This information should be updated and extended by polygons for impervious areas.



Figure 12: Provided polygons of streets and buildings



Figure 13: Image mosaic provided by FMM Forest Mapping Management, Austria

Based on image subset, segmentation and classification strategy is developed.

- roof
- not roof
 - o non impervious
 - o impervious
 - o probable impervious
 - could be impervious
 - not likely to be impervious
 - o shadow
 - shadow on vegetation
 - shadow on impervious area

Figure 14: Hierarchical rulebase structure in eCognition

On two scale levels classification takes place. The hierarchical rule base defines on scale 1 first the classes "roof" and "not roof". "not roof is further subdivided into "non impervious", "shadow", "impervious" and probable "impervious". "Shadow is classified as "shadow over vegetation", or "shadow over impervious area".

This hierarchy in rule base design allows a well-structured incorporation of knowledge with low mutual influence of object classes.

As the class names already show, linguistic and vague concepts are necessary to take uncertainty into account. The helpful concept to use shadow for further classification is only possible using the neighborhood concept that elevated targets as buildings throw shadow.

Using protocols this developed strategy is saved in a program routine and can be applied automatically on the whole image mosaic.

Operation	Level	Parameters
Segmentation	-> Level 1	(1, 0.00, 0.10)
Segmentation	-> Level 1	(15, 0.40, 0.60)
Load Class Hierarchy	All Levels	fmm_base_1.dkb
. Classification	Level 1	class-related, 5 cycles
Classification Based Fusion	Level 1	
Load Class Hierarchy	All Levels	fmm_base_2dkb
L Classification	Level 1	class-related, 5 cycles
T Classification Based New Level	Level 1	
Load Class Hierarchy	All Levels	fmm_base_3.dkb
Classification	Level 2	class-related, 5 cycles
Classification Based Fusion	Level 2	
L Classification	Level 2	class-related, 5 cycles
3 Vectorization	Level 2	(1.25, Permit Slivers, 1.00)
Export Objects	Level 2	ImageObjects.shp

Figure 15: Protocol for automatically analysis of mosaic

The protocol shows very clearly the interactive process of segmentation and classification.

Results are the

- a classification map and the reliability map.
 (Figure 16)
- and statistics with relation to certain classes and with respect to single objects (Figure 17)
- updated and extended shapefile (Figure 19)

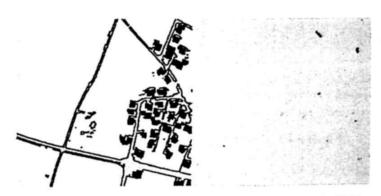


Figure 16: Classification map (left) of roads, buildings and impervious areas and reliability map (right); very bright and very dark objects in black/white presentation show less reliable classifications.

The reliability map (Figure 16, right) supports the semiautomatic workflow. Only those objects flagged as less reliable have to be manually assigned after inspection of the aerial images or – if no decision is possible base don the image in situ observations have to be performed.

Thus the methods do not replace all manual interactions, but reduces the amount significantly and increases objectivity to the large area of automatic classification.

Due to the final supported check by experts not only a time efficient process is possible, but also a product with high classification accuracy and reliability can be provided.

id	Best	Class ID B	est Member : M	ean 102451
5497	1	11	1	99.1171
	2	12	1	101.753
	3	14	1	24.5
	4	14	0.978938	204.209
	5	14	1	37.6667
	6	14	0.991359	211.292
	7	14	0.9828	221

Class	building	non impervious impervious			
Objects	218	91	883		
Sum Area	40818	341016	121684		
Mean Area	187.239	3747.42	137.807		
StdDev Area	168.376	13328.6	1857.64		
Min Area	1	0.25	0.25		
Max Area	1667	107179	54654.8		
Sum Length	3966.1	4887.51	8903.85		
Mean Length	18.1931	53.7089	10.0836		

Figure 17: Exemplary statistics for export.



Figure 19: Updated and extended shape file

Due to the object-oriented approach with

- the possibility to take context and semantic information into account and
- the ability for vector based input and output and
- due to the robust fuzzy classification with its advanced accuracy and reliability assessment

an operational system could be developed for analysis of aerial image mosaic and update of GIS- information.

The efficiency of the up-to-date time consuming and subjective analysis with many in-situ measurements can be improved and thus the quality of final geo-information can be increased while simultaneously reducing costs.

6 CONCLUSION

The main focus at Definiens is to produce software for the analysis of dependencies within complex systems. This can only be done if the high degree of mutual relationships and actions on different scales such as

context information, semantic and hierarchical structure is taken into account. With DEFINIENS Cognition Network Technology the basis is available to analyze images and texts from many different domains, and to combine the information from heterogeneous sources to support decision-makers in geomatics and biotechnology. The product eCognition, which is based on DEFINIENS Cognition Network Technology, is applied to geomatics

There are many things to improve for workflow and to increase the synergy with signal processing approaches. However, first two years time on market and many success stories from customers encourage further development.

ACKNOWLEDGEMENT

This document relies on many discussions within Definiens Geomatics unit and parts are taken from the user guide of eCognition created over the years with contribution of whole DEFINIENS geomatics unit.

REFERENCES

- [1] M. Baatz, M. Mimler, Bildobjekt-Primitive als Bausteine Extraktion von Objekten .of interest. bzw. anthropogenen Objekten basierend auf der expliziten Kanteninformation von Bildobjekt-Primitiven. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren Innovative Methoden Wichmann Verlag, Heidelberg, 2002, 179-188.
- [2] M. Baatz, A. Schäpe: Multiresolution segmentation an optimization approach for high quality multi-scale image segmentation. In Strobl, Blaschke & Greisebener (Edts): Angewandte Geographische Informationsverarbeitung XI. Beiträge zum AGIT-Symposium Salzburg 1999. Karlsruhe. Herbert Wichmann Verlag.
- [3] H. Bandemer, S. Gottwald, Fuzzy Sets, Fuzzy Logic, Fuzzy Methods with Applications (Wiley Press, New York, 1995)
- [4] U. Benz, M. Baatz, G Schreier, OSCAR-object oriented segmentation and classification of advanced radar allow automated information extraction Geoscience and Remote Sensing Symposium, 2001. IGARSS '01. IEEE 2001 International, Volume: 4, Syndey 2001, 1913 –1915
- [5] U. Benz, Supervised fuzzy analysis of single and multi-channel SAR data *Transactions on Geoscience and Remote Sensing*, 37(2), 1999, 1023-1037
- [6] J. Bezdek, S. Pal, Fuzzy Models for Pattern Recognition, Methods that Search for Structures in Data, IEEE Press, 1992
- [7] T. Blaschke, C. Glässer, S. Lang, Bildverarbeitung in einer integrierten GIS/Fernerkundungsumgebung. Trends und Konsequenzen. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg, 2002, 1-8.

- [8] B. Chaudhuri, N. Sarkar, Texture segmentation using fractal dimension. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (1), (1995), 72-77
- [9] R. Civanlar, H. Trussel, Constructing Membership Functions Using Statistical Data, *IEEE Fuzzy Sets and Systems* 18,1986, 1-13
- [10] D.L. Civco, J.D. Hurd, E.H. Wilson, M. Song, and Z. Zhang: A comparison of land use and land cover change detection methods. *Proc.* 2002 ASPRS Annual Convention, Washington, D.C. 12 p.
- [11] S.R. Coulde, E. Pottier, A review of target decomposition theorems in radar polarimetry *Transactions on Geoscience and Remote Sensing*, 34(2), 1996, 498-518
- [12] J. Curlander, W. Kober, Rule Based System for Thematic Classification in SAR Imagery, *Proc. IGARSS* 1992, 854-856
- [13] J. Daida, R. Samadani, J.F. Vesecky, Object-oriented Feature-tracking Algorithms for SAR Image of the Marginal Ice Zone *IEEE Transactions on Geoscience and Remote Sensing*, 28(4), 1990, 573-589
- [14]R. Donner, S. Schmidt, D. Walter, Ableitung von Nutzungsparametern landwirtschaftlicher Flächen aus multitemporalen ETM-Daten. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 189-198.
- [15] D. H. Douglas and T. K. Peucker. Algorithms for the reduction of the number of points required to represent a line or its caricature *The Canadian Cartographer*, 10(2), 1973,112--122
- [16]M. Ehlers, Fernerkundung für GIS-Anwender. Sensoren und Methoden zwischen Anspruch und Wirklichkeit. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 10-23.
- [17] M. Gähler, R. Janowsky, U. Schröder, Automatisierte Biotoptypenklassifizierung auf Basis höchstauflösender Flugzeugscannerdaten. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 233-240.
- [18] H. Ghassemian, D.A. Landgrebe, Object-oriented feature extraction method for image data compaction *IEEE Control Systems Magazine*, 8(3). 1988, 42–48
- [19]S. Gopal, C. Woodcock, Remote Sensing of Forest Change Using Artificial Neural Networks, *IEEE Trans. on Geoscience and Remote Sensing*, 34(2), 398-404
- [20] R.M. Haralick, K. Shanmugan, I. Dinstein, Textural features for image classification, *IEEE Trans. on Systems, Man, and Cybernetics* 3(6), 1973, 610-621
- [21] D. Haverkamp, C. Tsatsoulis, The Use of Expert Systems in Combination with Active and Passive Microwave Data to Classify Sea Ice, *NASA report*, 1992, 1625-1627
- [22] G. Heene, S. Gautama, Optimisation of a coastline extraction algorithm for object-oriented matching of multisensor satellite imagery Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. Volume: 6, Hawai 2000, 2632 –2634

- [23] M. Herold, A. Müller, S. Gunther, J. Scepan (2002): Object-oriented mapping and analysis of urban landuse/cover using IKONOS data *Poster at the 22nd EARSeL Symposium & General Assembly Geoinformation for European wide integration* Prague, Czech Republic, 4th 6th June 2002.
- [24] A. Hoffmann, J.W. van der Vegt, F. Lehmann, Towards automated map updating: detecting houses with new digital data-acquisition and processing techniques Geoscience and Remote Sensing Symposium, 2000, Proceedings, IGARSS 2000, Volume: 5, Hawai 2000, 2108-2110
- [25] T. Hofmann, J. Puzicha, J. Buhmann, Unsupervised texture segmentation in a deterministic annealing framework *In: IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(8), 1998, 803-818
- [26] G. Jaeger, U. Benz, Measures of classification accuracy based on fuzzy similarity *IEEE Transactions on Geoscience and Remote Sensing*, 38(2), 2000, 1462 –1467 [27] J. Janoth, M. Eisl, F. Klaushofer, W. Luckel, Segmentbasierte Verfahren zur Veränderungsanalyse und Waldklassifizierung mit hochauflösender Satellitendaten. *In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden.* Wichmann Verlag, Heidelberg, 98-107.
- [28] S. Kaya, T. J. Pultz, C. M. Mbogo, J. C. Beier, E. Mushinzimana, The Use of Radar Remote Sensing for Identifying Environmental Factors Associated with Malaria Risk in Coastal Kenya In: International Geoscience and Remote Sensing Symposium (IGARSS .02), Toronto, June 24-28, 2002.
- [29] T. Kellenberger, Extracting forest inventory parameters with IKONOS, DEM and DSM data. Website: http://www.geo.unizh.ch/rsl/research/LACOMM/ecognition/
- [30] K. Kim; Seung-Keol Choe; Jong-Hung Lee; Young-Kyu Yang, Managing scheme for 3-dimensional geofeatures using XML Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International, Volume: 7, Hawai 2000, 2899 –2901
- [31] R. de Kok, A. Buck, T. Schneider, U. Ammer, Modular project design in object oriented analysis. In Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren .Innovative Methoden. Wichmann Verlag, Heidelberg, 33-42.
- [32] S. Lang, Zur Anwendung des Holarchiekonzepts bei der Generierung regionalisierter Segmentierungsebenen in höchstauflösenden Bilddaten. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 24-33.
- [33] C. Leser, Operationelle Biotoptypenkartierung mit HRSC-Daten. Probleme und Lösungsansätze. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 88-97.

- [34] K. Leukert, Verwendung von GIS-Daten für die Objektextraktion. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 132-140.
- [35] B. Manjunath, R. Chellappa, Unsupervised texture segmentation using Markov random field models *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13, 1991 478-482.
- [36] J. Mao, A. Jain, Texture classification and segmentation using multiresolution simultaneous autoregressive models *In: Pattern Recognition 25*, 1992, pp. 173-188
- [37] F. Maselli, A. Rodolfi, C. Copnese, Fuzzy Classification of Spatially Degraded Thematic Mapper Data for the Estimation of Sub-Pixel Components, *Int. Journal of Remote Sensing* 17(3), 1996, 537-551
- [38] M. Möller, M. Rosenberg, M. Volk, Nutzung von Methoden der Fernerkundung und GIS zur Standortcharakterisierung als Beitrag zur Umsetzung der EG-Wasserrahmenlinie. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 150-158.
- [39] M. Möller, New remote sensing sensors and imaging products for the monitoring of urban dynamics. In: Proceedings of the 22nd EARSeL Symposium & General Assembly .Geoinformation for European wide integration, Prague, Czech Republic, 4th 6th June 2002. (in press)
- [40] M. Möller, Untersuchungen zur Modellierung und Analyse von höchstauflösenden Scannerdaten In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 208-221.
- [41] A. Müller, Object-oriented land use classification of polarimetric high resolution SAR and hyperspectral HyMapTM data Poster at the 22nd EARSeL Symposium & General Assembly .Geoinformation for European wide integration Prague, Czech Republic, 4th 6th June 2002.
- [42] M. Neubert, G. Meinel, Fortführung von Geobasisdaten durch die segmentbasierte Auswertung von IKONOS-Daten. erste Ergebnisse. In: J. Strobl, T. Blaschke, G. Griesebner (Hrsg.): Angewandte Geographische Informationsverarbeitung XIV. Beiträge zum AGIT-Symposium Salzburg 2002, H. Wichmann Verlag, Heidelberg, pp. 403-408.
- [43] M. Neubert, G. Meinel, Segmentbasierte Auswertung von IKONOS-Daten. Anwendung der Bildanalyse-Software eCognition auf unterschiedliche Testgebiete. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 108-117.

- [44] I. Niemeyer, M. J. Canty, Object-oriented change detection analysis with high-resolution satellite imagery. Poster at the 22nd EARSeL Symposium & General Assembly .Geoinformation for European wide integration, Prague, Czech Republic, 4th 6th June 2002. [45] I. Ng, et. al, Supervised segmentation, IEEE Trans on Signal Processing 31(2), 134-162
- [46] D. Panjwani, G. Healey, Markov random field models for unsupervised segmentation of textured color images *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(10), 1995, pp. 939-954.
- [47] E. Pierce, F. Ulaby, K. Sarabandi, M. Dobson, Knowledge-based Classification of Polarimetric SAR Images, *IEEE Trans. on Geoscience and Remote Sensing*, 30(4), 1994, 697-705
- [48] H. Pilz, J. Strobl, Versiegelungsanalyse auf Basis digitaler Farborthofotos mittels objektorientierter Bildverarbeitung. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren. Innovative Methoden. Wichmann Verlag, Heidelberg: 159-168.
- [49] Ch. Oliver, Sh. Quegan, *Understanding Synthetic Aperture Radar* (Artech House, Boston, 1998)
- [50] S. Serpico, F. Roli, Classification of Multissensor remote-sensing images by structured neural networks *IEEE Trans on Geoscience and Remote Sensing 33*(3), 1995, 562-577
- [51] P.C. Smits, A. Annoni, Cost-based feature selection for GIS-embedded data fusion Geoscience and Remote Sensing Symposium, 2000, Proceedings. IGARSS 2000 Volume: 6, Hawai 2000, 2614-2616
- [52] M. Song, D.L. Civco, A knowledge-based approach for reducing cloud and shadow. *Proc. ASPRS Annual Convention*, Washington, D.C., 2002, 7 p
- [53] C. Tsatsoulis, Expert Systems in Remote Sensing Applications, *IEEE Geoscience and Remote Sensing Newsletter*, 1993, 7-15
- [54] G. Willhauck, U. C. Benz, F. Siegert, Semiautomatic classification procedures for fire monitoring using multitemporal SAR images and NOAA-AVHRR hotspot data. *In: Proceedings of the 4th European Conference on Synthetic Aperture Radar*, Cologne, Germany, 4 6 June 2002
- [55] Y. Xia; Ho, 3D vector topology model in the Visualization System Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. Volume: 7, Hawai 2000, 3000 –3002
- [56] L. Zadeh, Fuzzy Sets IEEE Trans. Information and Control 8, 1965, 338-353
- [57] UserGuide eCognition, Definiens, website: www.definiens-imaging.com

IMAGE FILTER FOR EFFICIENT IMPULSE NOISE REMOVAL AND FINE DETAIL PRESERVATION

Oleksiy POGREBNYAK, Pablo MANRIQUE

Instituto Politécnico Nacional, Centro de Investigación en Computación, Av. Juan de Dios Bátiz s/n, Colonia Nueva Industrial Vallejo, C.P. 07738, México D.F. olek@pollux.cic.ipn.mx, pmanriq@pollux.cic.ipn.mx

ABSTRACT

A filtering algorithm applicable to image processing is presented. It was designed using rank-ordered mean (ROM) estimator to remove an outlier and robust local data activity estimators to detect the outliers. The proposed filter effectively removes impulse noise and preserve edge and fine details. The filter possesses good visual quality of the processed simulated images and good quantitative quality in comparison to the standard median filter. Recommendations to obtain best processing results by proper selection of the filter parameters are given. The designed filter is suitable for impulse noise removal in any image processing applications. One can use it at the first stage of image enhancement followed by any detail-preserving techniques such as the Sigma filter at the second stage.

KEY WORDS

image processing, nonlinear filters, detail-preserving filtering

1. INTRODUCTION

In practice, the quality of digital images often is not acceptable to provide reliable data interpretation due to It is highly desirable to get the random noise presence. image enhancement providing both effective noise cancellation/suppression and fine detail preservation. Linear filters ensure strong attenuation of Gaussian noise but they fail when data contain the impulse one. On the other hand, nonlinear filters have become very attractive in signal and image processing because of their ability to suppress noise of different nature, in particular, to remove impulse noise. Nonlinear filtering is also a well-known detail-preserving method. However, nonlinear filters are mainly designed to preserve edges of image objects only, but not fine details such as thin lines and small-scale objects. In particular, median, Wilcoxon^{1,2} and α-trimmed mean^{2,3} filters can remove small size objects considering them as outliers. As a result, they may to be unable to preserve fine details. A class of linear median hybrid

(LMH) filters was introduced by Heinonen and Neuvo⁴ to provide edge preservation with impulse noise reduction^{5,6}. Its subclass, FIR-median hybrid filters (FMH)⁷, provides preservation of thin lines as well^{5,6}. Impulse rejecting filters⁶ suppress impulse noise effectively and avoid unnecessary distortions of noise-free pixels. These filters use different impulse detectors to decide, if the current pixel is an outlier and should be filtered by some nonlinear filter or it can be unaltered otherwise. To provide simultaneous detail preservation, rather complicated impulse detectors must be used. For example, Mitra et al⁸ suggested rank-ordered mean (ROM) impulse rejecting filter with sophisticated fuzzy detector that can be optimized using image training data.

Besides, Sigma⁹ and KNN¹⁰ filters can preserve fine details well, but their robustness is insufficient to provide desired suppression of impulse noise^{2.5,6}. Since the standard Sigma filter does not have any robustness, it is not able to suppress impulse noise at all. The attempts to get more appropriate robust versions of two latter mentioned techniques known. First, the modified locally adaptive Sigma filter¹¹ possesses some robust features but it does not perform well for probability of spikes greater than 0.05. Second, the known adaptive KNN filter¹² also provides insufficient impulse noise suppression. Another attempt to get better detail preservation is a weighted median filter¹³ that has better detail preservation but lower noise suppression in comparison to the standard median one.

Recently, a robust KNN filter (RM-KNN filter) for efficient impulse noise suppression and good fine detail preservation was designed ¹⁴. The filter uses a robust RM estimator ^{15,16} derived from R-estimators following from the statistical rank theory and robust M-estimators. The limitation of K nearest neighbors is applied to the data within the filtering window to perform calculations in an recursive manner. The resulted RM-KNN estimator has an adaptive nature: the number of K neighbors is adjusted using some robust estimator of local data activity that acts as outlier detector. The variants of this filter that use the different local data activity estimators were presented ^{14,17} and investigated.

It was found that this kind of impulse cancellation filter is very sensitive to the quality of the local data activity estimator. Some variants of this filter substitute the RM-KNN estimate by the output of the standard 3x3 median filter in the case when the local data activity is high that results that the central pixel of the filtering window is outlier. In contrast, when the local data activity is too small, the decision may be made to take the central window pixel as the filter output, similar to known impulse rejecting filters⁶. If the activity has intermediate values, the filter performs recursive calculations according to the RM-KNN estimator algorithm¹⁵.

Unfortunately, the processing with the RM-KNN filter is slow because of the recursive nature of the robust estimator used. We found that the computation time may be significantly reduced eliminating the RM-KNN estimator when the data local activity estimator, or outlier detector performs better. Simultaneously, we found by simulations that the output of the ROM filter produces a more robust and more accurate estimate in comparison to the median estimator when the outlier is detected exactly or more or less reliably.

In this way, we designed a new adaptive impulse rejecting filter that uses an enhanced local data activity estimator and behave as an impulse rejecting filter: calculate the ROM estimate in the case when the data activity is high, and preserve the image pixels otherwise. In this paper, we present the designed filter that possesses the cancellation of impulse noise and preserves well image fine details.

2. PROPOSED IMAGE FILTER

Different impulse noise models were proposed and described in the literature^{2, 6}. We used the following image degradation model in the case of impulse noise presence^{15,17}:

$$u(x,y) = n_{im}(e(x,y)), \qquad (1)$$

where $e(x, y) = \vec{e}$ is the vector of an original image, $u(x, y) = \vec{u}$ is the vector of a distorted image, and $n_{im}(e(x, y))$ is the functional

$$n_{im}(e(x, y)) = \begin{cases} random \text{ valued spike with probabilty } P \\ e(x, y) \text{ otherwise} \end{cases}$$
 (2)

We assume that the spikes have uniformly distributed random values (0..255 for the byte-represented images). Such assumption makes the problem of impulse noise removal more complicated, because when spikes are represented by maximal and/or minimal values only, one can use some thresholding techniques for their detection and removal. Besides, the impulse noise described by the model (1) is more realistic.

With this model, the problem of impulse noise removal is to derive a robust filtering algorithm that can be able both to remove the outliers and to preserve the fine details well.

The output of the ROM filter can be represented as follows. Let the w(n) is a vector that represents the data within the 3x3 filtering window that is scanned on the entire image and is centered at i, j pixel excluding this pixel, u(i, j), itself:

$$\mathbf{w}(n) = \{ w(1), w(2), w(3), w(4), w(5), w(6), w(7), w(8) \}$$
 (3)

These samples can be ordered by rank, which defines the vector

$$\mathbf{r}(n) = \{r(1), r(2), r(3), r(4), r(5), r(6), r(7), r(8)\}$$
 (4)

where r(1), r(2), ..., r(8) are the elements of w(n) arranged in ascending order such that $r(1) \le r(2) \le ... \le r(8)$. In this case of even number of data the most robust estimate that corresponds to a Hoghes-Leman estimate by the rank sign test¹⁸ is

$$\hat{e}_{ROM}(i,j) = \frac{r(4) + r(5)}{2}$$
 (5)

According to the theory of rank tests, the estimate (5) in the case when the vector **w** has odd number of elements corresponds to the median of the data, which is more known and widely used in practice.

From the point of view of rank test/estimation theory, the estimator (5) has the same robust properties as the usual median estimator. However, taking into account that the image filter that is based on the estimator (5) excludes the corrupted central pixel of the filtering window, this estimator seems to be more robust.

One can expect that the ROM filter, which implements estimator (5) excluding the central window pixel u(i, j) at the stage of vector $\mathbf{w}(n)$ forming, will perform better than the standard median filter in the case when the pixel u(i, j) is detected as corrupted properly. Therefore, it is highly important to detect the outlier reliably. Besides, the performance of the image rejecting filter depends on the false outlier detection. Thus, we have to design an outlier detector, which has to be able both to detect an outlier well and to minimize the false detection, which is usually occur in the vicinity of small scale fine details of the image.

For the purpose of outlier detection, we modified the local data activity estimator, which was developed and presented previously¹⁹. This estimator is based on the robust estimate of data scale, the median of absolute

deviations from median (MADM) that is known from the theory of robust M-estimators²⁰. The calculation scheme of the previous estimator is described as

$$S(u(i,j)) = \frac{\operatorname{med}\{|u(i,j) - u(i+m,j+n)|\}}{\operatorname{MADM}\{u(i,j)\}}$$
(6)

where $\operatorname{med}\{|u(i,j)-u(i+m,j+n)|\}=\operatorname{MADCP}(i,j)$ is the median of the of absolute deviations from central filtering window pixel, k,l=-L..L, and $\operatorname{MADM}\{u(i,j)\}$ is calculated as

We found by simulations that the estimator given by Eq.(6) performs well at flat image regions, but in the vicinity of edges it produces the small values. This causes insufficient outlier removal near the object edges in the filtered image. To resolve the problem of insufficient sensitivity of the impulse detector (6) near the edges, it was modified to produce better results in the filtered images. The new version of the local data activity estimator is expressed as

$$S(u(i,j)) = \frac{[a \cdot MADCP\{u(i,j)\}]^2}{\ln(MADM\{u(i,j)\})}$$
(8)

where the coefficient a varies the sensitivity of the estimator. The estimator (8) performs well both in flat regions and near the image object edges. Unfortunately, it was found by simulation that the direct use of this estimator as impulse detector is insufficient and one has to take into account other features as well. The additional criterion that we used is the difference between the value of the central pixel of the filtering window and ROM estimate (5)

$$dif(i,j) = |u(i,j) - \hat{e}_{ROM}(i,j)| \tag{9}$$

Finally, the proposed impulse rejecting filter can be described as a sequence of the experimentally derived rules, which can be formulated as

$$\hat{e}(i,j) = u(i,j), \text{ if } dif(i,j) \le 5$$

$$|\hat{e}_{ROM}(i,j), \text{ if } dif(i,j) > \text{ the median of entire image} \\ \text{ or if MADM}(i,j) \le 1 \text{ and MADCP}(i,j) > 0 \\ \text{ or if } S(i,j) > 1 \\ \text{ or if } [dif(i,j) \cdot a]^2 / \ln(\text{MADM}(i,j)) > 1 \\ u(i,j) \text{ otherwise}$$

$$(10)$$

where $\hat{e}(i, j)$ denotes the output of the proposed filter and a is the noise sensitivity coefficient. To possess the fine detail preservation properties, the filter always perform calculation of the ROM estimate in 3x3 filtering window meanwhile the scanning window for local data activity estimator (8) generally can be of the different size. However, the estimator (8) that estimates the data within the window of the same 3x3 size produces the better results as it is shown in the next section.

3. SIMULATION RESULTS

We performed a number of different tests to study the properties of the proposed algorithm (10) and to compare it to the standard median filter. The criterion used for the comparison of the performance of the filters was peak signal to noise ratio (PSNR), which can be expressed as

PSNR =
$$10 \cdot \log_{10} \left(\frac{255^2}{\sum [e(x,y) - \hat{e}(x,y)]^2} \right)$$
, (11)

where e(x,y) denotes the pixels of the original (no corrupted) image and $\hat{e}(x,y)$ denotes the pixel of the filtered corrupted image (restored image).

In simulations, the parameters were varied: the percentage of the impulse noise, the noise sensitivity coefficient value, and the size of the sliding window L for determination of the local median in calculation of the parameters $MADM\{u(i,j)\}$, $MACP\{u(i,j)\}$ that are used in estimator (8) and in the filter output (10) forming as well.

To evaluate the deterministic properties of the designed filter, we performed filtering of the artificial test image shown in Figure 1 a) by the standard 3x3 median filter and the proposed one. One can see that in this case of the artificial image processing the proposed filter performs very similar to the median one.

To determine the noise suppression properties of the proposed adaptive ROM filter, the standard 512x512 test images ("Lena" and "Mandrill") shown in Figure 2 were corrupted by the random-valued impulse noise according to Eq.(1). The percentage of impulse noise was varied from 1% up to 15%. The filter parameter a from Eqs.(8),(10) was varied in a wide range as well as estimator (8) window size L was varied from 3 to 7.

Table presents the PSNR values, which were obtained according to (11) on images processed by the proposed filter with the optimal values of the coefficient a. The PSNR values of the standard median filter having different window size 3x3, 5x5, 7x7 are presented as well.

Analyzing these values, one can see that the designed filter performs better and provides significantly larger PSNR values in comparison to the median filter. It can be concluded from the analysis of this Table that the optimal size of the estimator (8) window is 3x3. The PSNR values of the processed "Mandrill" images is smaller than the ones of the "Lena" images that is caused by numerous small scale details contained in the original "Mandrill" image.

The PSNR criterion does not reflects well the quality of the filtered images in the sense of fine detail preservation. That is why it is necessary to check visually both the detail preservation and absence of outliers after filtering. Figure 2 illustrates the impulse noise removal by the designed filter and the median one. Figure 2 (a) shows the noisy test image "Lena", Figure 2 (b) presents the output of the standard 3x3 median filter and Figure 2 (c), (d) show images processed by the proposed filter having estimator (8) window size 3x3 and 7x7. Analyzing this Figure, one can see that the proposed filter possesses both good impulse noise removal and better detail preservation in comparison to the standard median filter.

4. CONCLUSION

We have presented an adaptive impulse rejecting filter for image processing applications. Its deterministic and statistical properties have been analyzed. The proposed filter possesses good impulse noise removal and preserves well edge and fine details in the processed images. The filter optimal parameters have been given. The presented results demonstrate obviously that the designed filter can remove impulse noise even from highly corrupted images. It was established that the optimal impulse detector size as 3x3, the same as the filtering window. This feature allows to simplify the filtering algorithm. The proposed filter can be used for impulse noise removal as well as for information abundance decreasing in image compression applications.

Because the proposed algorithm is strictly non-linear, it not changes the noise-free pixels and not introduces any "new" information. This feature means that the filtered image can be passed through a once trained classifier without the necessity of the new training.

REFERENCES

- [1] R.J.Crinon, "The Wilcoxon filter: a robust filtering scheme", in Proc. *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing* 1985, 668-671.
- [2] I.Pitas and A.N.Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications* (Ed.: Kluwer Academic Publisher, 1990).
- [3] J.B.Bednar and T.L.Watt, "Alfa-trimmed means and their relationship to median filters," *IEEE Trans.*

- Acoust., Speech, Signal Processing ASSP-32, 1984, 145-153.
- [4] P. Heinonen and Y.Neuvo, "Smoothed median filters with FIR substructures," in Proc 1985 *IEEE Int Conf. Acoust., Speech, and Signal Processing* Tampa, March 1985, 49-52.
- [5] Y.Fong, C.A.Pomalaza-Ráez, X.Wang. "Comparison study of nonlinear filters in image processing applications". *Opt. Eng. 28*(7), 1989, 749-760.
- [6] J.Astola and P. Kuosmanen. Fundamentals of Nonlinear Digital filtering (Ed.: Boca Raton-New York: CRC Press, 1997).
- [7] A.Neiminen, P. Heinonen and Y.Neuvo, "A new class of detail-preserving filters for image processing," *IEEE Trans. Pat. Anal. Mach. Intell.* PAMI-9(1), 1987, 74-90.
- [8] Eduardo Abreu, Michael Lightstone, Sanjit K. Mitra, Kaoru Arakawa. "A new Efficient Approach for the Removal of Impulse Noise from Highly Corrupted Images," *IEEE Trans. on Image Processing*, 5(6), 1996, 1012-1025.
- [9] J.-S.Lee, "Digital image smoothing and the sigma filter," Computer Vision, Graphics, and Image Processing, 24, 1983, 255-269.
- [10] L.S.Davis, A.Rosenfeld, "Noise cleaning by iterative local averaging," *IEEE Trans. Syst. Man Cybern* SMC-8(9), 1978, 705-710.
- [11] V.Lukin, N.N.Ponomarenko, P.Kuosmanen, and J. Astola, "Modified sigma filter for processing of images corrupted by multiplicative and impulse noises," *in Proc. of EUSIPCO-96, III*, Triesto, Italy, 1996, 1909-1912.
- [12] Atsushi Teramoto, Isao Horiba, Noburu Sugie, "Improvement of Image Quality in MR and CT Image Using Adaptive K-nearest Neighbor Averaging Filter," Proceedings of 8th International Conference on Signal Processing Applications & Technology, 1, San Diego, USA, 1997, 677-681.
- [13] D.R.K.Brownrigg, "The weighted median filter," Commun. ACM, 27, 1984, 807-818.
- [14] Ponomaryov V.I., Pogrebniak A.B. «Novel robust KNN filters with impulsive noise suppression for image processing» Mathematical Modelling and Estimation Techniques in Computer Vision, Proc. SPIE 3457, Franqoise Préteux, Jennifer L.Davidson, Edward R.Dougherty, Editors, Orlando, 1998, 19-30.
- [15] V.I.Ponomarev, A.B.Pogrebniak. Robust RM filters for radar image processing. Proc. SPIE 2955, Image and Signal Processing for Remote Sensing III, Jacky Desachy, Ed., 1996, 272-283.
- [16] Volodymyr V. Ponomaryov, Oleksiy B. Pogrebnyak. «Novel robust RM filters for radar image preliminary processing». *Journal of Electronic Imaging*, 8(4), 1999, 467-477.
- [17] Volodymyr I. Ponomaryov, Oleksiy B. Pogrebnyak, Felipe González Leon "Robust rank detail-preserving filters with noise suppression in image processing applications", *Visual Information Processing VIII*, Stefen K. Park, Richard D. Juday, Chairs/Editors, Proc. SPIE 3716, Orlando, 1999, 206-211.

[18] T.P.Hettmansperger. Statistical inference based on ranks (Ed.: New York Wiley, 1984).

[19] Volodymyr I. Ponomaryov, Luis Niño de Rivera Francisco J.Jallegos Funes, Oleksiy B. Pogrebnyak "Novel detail preserving robust filter for multiplicative and additive noise suppression in the image processing", Visual Information Processing IX, Proc. SPIE 4041, Stephen K.

Park, Zia-ur Rahman, Chairs/Editors, Orlando, USA, 2000, 172-179.

[20] F.R.Hampel, E.M.Ronchetti, P.J.Roussew, W.A.Stahel. *Robust Statistics. The Approach Based on Influence Function* (Ed.: New York Wiley, 1986).

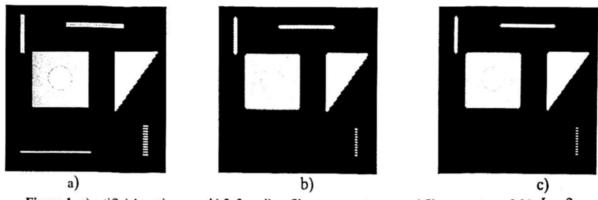


Figure 1. a) artificial test image b) 3x3 medi an filter output c) proposed filter output, $\alpha=0.05$ L=3



Figure 2. Noisy test and filtering images: (a) test image "Lena" corrupted by 10% impulse random-valued noise; (b) the output of the 3x 3 median filter; (c) the output of the proposed filter (10) with a = 0.044 and estimator (8) window size 3x3; (d) the output of the proposed filter (10) with a = 0.044 and estimator (8) window size 7x7

Table. Simulation results on impulse noise suppression

Window size		Lena		Mandrill			
of estimator (8) (median filter)	Impulse noise percentage	Filter sensitivity coefficient a	PSNR	PSNR for median filter	Filter sensitivity coefficient a	PSNR	PSNR for median
(1	0.034	45.1	36.4	0.019	34.1	23.64
	2	0.037	43.1	36.2	0.021	31.7	23.58
3	3	0.037	42.1	36.1	0.022	30.6	23.53
	4	0.038	41.0	35.9	0.023	29.7	23.48
	5	0.039	40.2	35.8	0.024	28.8	23.42
	6	0.04	39.6	35.6	0.024	28.3	23.38
	7	0.042	39.1	35.4	0.026	27.9	23.34
	8	0.044	38.6	35.3	0.026	27.4	23.27
	9	0.044	38.1	35.0	0.027	27.0	23.24
	10	0.044	37.6	34.8	0.027	26.6	23.18
	11	0.046	37.0	34.6	0.027	26.3	23.13
	12	0.046	36.5	34.3	0.027	26.0	23.06
	13	0.047	36.3	34.2	0.028	25.7	22.98
	14	0.047	35.9	33.9	0.028	25.4	22.94
	15	0.048	35.5	33.7	0.028	25.2	22.88
	I	0.031	43.7	32.4	0.019	34.1	21.27
1	2	0.031	41.7	32.3	0.02	31.8	21.26
5	3	0.035	40.9	32.3	0.022	30.1	21.26
	4	0.036	39.8	32.2	0.022	29.6	21.24
	5	0.038	39.2	32.2	0.022	28.8	21.23
	6	0.039	38.7	32.1	0.023	28.2	21.23
	7	0.04	38.3	32.09	0.024	27.8	21.23
	8	0.042	37.8	32.02	0.024	27.3	21.21
	9	0.041	37.4	31.97	0.025	26.9	21.22
	10	0.041	37.0	31.92	0.026	26.5	21.20
	11	0.041	36.6	31.81	0.026	26.2	21.19
].	12	0.042	36.2	31.75	0.026	25.9	21.18
	13	0.042	35.9	31.71	0.027	25.6	21.16
1.	14	0.042	35.6	31.61	0.027	25.4	21.16
	15	0.044	35.3	31.56	0.027	25.1	21.13
_	1	0.027	42.8	30.26	0.019	33.9	20.59
7	3	0.034	40.0	30.21	0.02	30.4	20.58
	5	0.039	38.5	30.16	0.022	28.6	20.58
	7	0.041	37.7	30.12	0.024	27.7	20.58
	8	0.041	37.3	30.11	0.024	27.2	20.57
<u> </u>	10	0.044	36.6	30.03	0.025	26.5	20.56
	11	0.047	36.2	29.96	0.025	26.2	20.55
	12	0.044	35.9	29.94	0.025	25.9	20.55
	13	0.046	35.6	29.95	0.026	25.5	20.54
<u> </u>	14	0.046	35.3	29.85	0.026	25.3	20.55
	15	0.044	35.1	29.84	0.027	25.1	20.53

A NOVEL IMAGE RETRIEVAL

Jianfeng REN¹, Yuli SHEN² and Lei GUO³

1,3</sup>Department of Automatic Control,

North Western Polytechnic University, P.R. CHINA

² Department of Electrical Engineering,

Zhan Jiang Ocean University, P.R.CHINA

rijff@163.com¹ lguo@nwpu.edu.cn³

ABSTRACT

In this paper, we present results of a project that seeks to transform the low level features to high level meaning. Firstly we extract the low level features called as the representative colors from the images, and then we present a new approach called WordNet to establish the link from the low level feature vectors to the semantics. In order to improve the retrieval efficiency, the relevance feedback is also applied into our system. Experimental results show that our method is promising.

KEY WORDS

Representative colors, wordnet, relevance feedback.

1 INTRODUCTION

Indexing diverse collections of multimedia data remains a challenging problem. Most current approaches to image retrieval mainly focused on two aspects: one is the visual features [1,2,3,4], such as the color, texture and shape, the other is the distance metrics [5,6,7].

However, in the visual feature approaches, the interfaces supplied to the user are non-intuitive and unnatural. The user searching for visual data usually has some idea of the image content and the image layout of the desired image, where the image content is described in terms of objects and global features rather than low-level features such as the color and texture. However, the image objects can be correctly and clearly expressed through keywords, which have powerful query abilities.

All the images from the database in our experiment are nature images, such as rose, blue sky and so on. Through human observation, these objects have obvious color characteristics. For example, rose appears red. This paper is conducted to explore such links between the colors and nature objects, which are expressed through keywords.

Even though significant progress has been made toward developing effective content-based descriptors, such as the standard descriptors by MEPG-7 [8], there is a difficulty to narrow the gap between the low-level features in image analysis and image understanding at the semantic level. Because people analyze, understand and classify the image content according to its semantic features.

Due to the importance of semantics, some approaches have been provided to bridge the gap between the low-level features and semantic level features. Shi-F Chang [9] proposed the novel idea of Semantic Visual Templates (SVT) to narrow this gap. Each template represents a personalized view of concepts (e.g. slalom, meetings, sunsets etc). The SVT is represented using a set of successful queries, which are generated by a two-way interaction between the user and the system. Aleksandra Mojsiovic and Bernice Rogowitz [10] proposed a method for semantic categorization and retrieval of photographic images based on low-level image descriptors. In this method, they first used multidimensional scaling and hierarchical cluster analysis (HCA) to model the semantic categories into human observers organizing images.

Through a series of psychophysical experiments and analyses, they refined the definition of these semantic categories, and used these results to discover a set of low-level image features. The main drawback of this method is that we need to do a lot of psychophysical experiments. Wiiam I.Grosky and Rong Zhao [11] presented the techniques, latent semantics indexing (SCI), to negotiate the gap. First, a corpus is formed of documents (in this case, images with a caption) from which features are computed. Then by singular value decomposition (SVD), the dictionary covering the captions is correlated with features derived from the pictures. However these methods seem to be inefficient due to the size of image collections.

In order to overcome the shortfalls mentioned above, in this paper, we propose a new approach based on the color semantics to narrow the gap between the low-level and semantic level, considering the intuitive characteristics of color. Firstly, the low-level features based on the color are presented. Then we use the WordNet in order to narrow the gap between the low-level feature and semantic level features through the training sample images and human interactions. In order to more efficiently retrieve images, relevance feedback has been adopted in our strategy. Experimental results demonstrate that this method could correctly retrieve images not only through the specified images but also keywords in a given domain.

This paper is organized as followed: we introduce the overview of our method in section 2. In section 3, the low-

level features based on the representative colors are developed. In section 4, we will present the WordNet to bridge the gap between low level features and high-level features. We will provide the different query strategies in section 5. In section 6, the relevance feedback is applied into our strategy. In section 7, a detailed analysis based on the experimental results will be presented. Finally, Conclusions and remarks are given in section 8

2 OVERVIEW

Our method mainly includes the following steps:

Step1: image decomposed into five regions14. Because there are no good methods to segment different objects in the images at present, we mainly apply the simple method provided in [14] to segment the images. After segmented, each region should include one object as possible as it could.

Step2: computing the frequency of eight domain color components in each region.

Step3: establishing the WordNet from the color features to semantics through the training samples and human interactions

Step4: re-define the query feature from the similar images using the relevance feedback technique until the user satisfy the results main goal of *pre-processing* is to prepare raster cartographic images in such a way as to simplify them and increase the reliability of their recognition in the automatic system.

3 COLOR FEATURE EXTRACTION

Color is perhaps the most expressive of all the visual features and has been extensively studied in image retrieval during the last decade. Some low-level features are developed by using the color histograms[2,12,13]. This type of representation tends to produce false positives. Color histogram give statistics of the image pixels but do not provide spatial, relational or content information in terms of objects in the scene. To improve performance, spatial information was incorporated into color-based image retrieval system by allowing multiple color histograms, representing different locations in the image [2,12,13].

In this paper, we provide eight representative colors in a region of interest. The reason that we employ this method is as following:

- It is provided mainly based on the observation that a small number of colors are usually sufficient to characterize the color information in image region.
- Representative color approach can overcome the drawback of the traditional histogram. The number of

- bins in a typical color histogram range from few tens to a few hundreds. The high dimensionality of the feature vectors result in high computational cost in distance calculation for similarity retrieval, and inefficiency in indexing and search.
- It also considers the spatial information and relationships between the different regions in the images.

The process of our method is as follows: Firstly the image can be decomposed into five fixed regions14, which is shown in figure 1, and then representative color histogram is computed in each region.

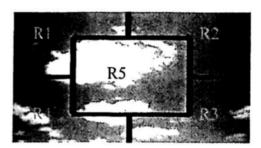


Figure 1. Image decomposition

3.1 COLOR SPACE

The different color spaces used in MPEG-7 include the familiar monochrome, RGB, HSV, YCrCb, and the new HMMD. Although many approaches are mainly based on the HSV and YCrCb rather than RGB due to the fact these methods can improve the perceptual uniformity [15], the RGB color format is the most common color format for digital images. The primary reason for this is because it retains compatibility with computer displays.

3.2 REPRESENTATIVE COLORS

According to human vision perception, a small number of image colors are sufficient to express the color information. Therefore, in our method, we regard eight colors as the representative colors, which are mapped into corresponding points in the RGB space, which are shown in table 1.

Red	(255,0,0)		
Green	(0,255,0)		
Blue	(0,0,255)		
Yellow	(255,255,0)		
Magenta	(255,0,255)		
Cyan	(0,255,255)		
Black	(0,0,0)		
White	(255,255,255)		

Table 1. Relationships between the representative color and its corresponding points

3.3 COLOR CLUSTERING

the image.

In order to compute the percentage of representative colors in the region, the colors in the image region should be clustered based on the nearest neighbor algorithm.

assume eight representative colors $\{C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8\}_{as}$ the clustering center, and # C, denotes the total number of pixels which are grouped into the clustering center Ci, and # C denotes the total number of pixels in the image, so $\frac{\#C_i}{}$ can be regarded as the percentage that representative color account for in the image. For example, in this paper, $\frac{\# C_1}{\# C_1}$ denotes that the percentage of the red component in

Because the image are decomposed into five fixed regions such as the left-up, right-up, left-down, right-down and center, we should express the spatial relationships between the different regions in the feature vectors. Here, let expresses the corresponding region in the image. For example, in a given region, after color clustering $\left[\frac{\#C_1}{\#C}, \frac{\#C_2}{\#C}, \frac{\#C_3}{\#C}, \frac{\#C_4}{\#C}, \frac{\#C_5}{\#C}, \frac{\#C_6}{\#C}, \frac{\#C_7}{\#C}, \frac{\#C_8}{\#C}, s5\right]$ is obtained

Now we have obtained the representative color features in the corresponding regions. However, these feature vectors could not be directly recognized by humans, just because people understand image at the semantic level. So a new approach called WordNet .is provided to solve this problem.

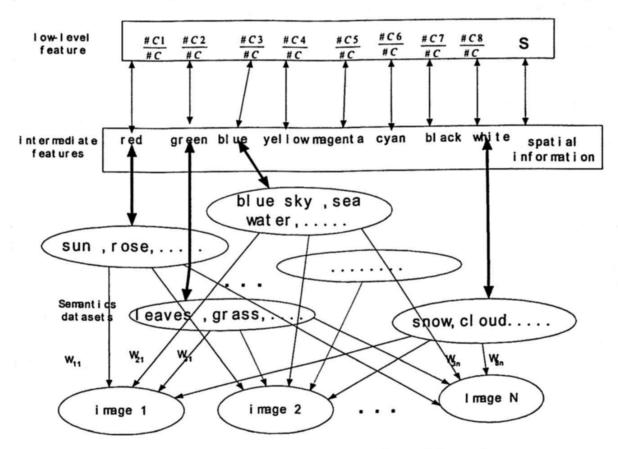


Figure 2. Link between representative colors and semantic keywords

4 ESTABLISH THE LINKS

In our paper, all the images are from the scene images, which have obvious color characters. We can easily find the link between the representative color and nature object. For example blue can express the meaning of blue sky or sea water. A detailed link between the color and the corresponding keywords is shown in figure 2

Through the above links, we can easily narrow the gaps between the low-level and semantic level features. And more, we also see that the intermediate features play an essential role in bridging the gap between them. Such links are stored in the database as the look-up table. In addition, the semantic datasets can be expanded through human interactions and relevance feedback, in other

words, the user can add and delete the keywords in the semantic datasets.

For a give region R5 located in the center of the image, the specified process from the low-level to semantic level is expressed as following:

If #C1 in the feature vector is maximum Then the object can be {rose, sun....} If #C3 in the feature vector is maximum Then the object can be {see water, blue sky....}

Through the above algorithm, each region can be indexed one representative color, which has a set of keywords associated. Accordingly, an image that is decomposed into five fixed regions will have not more than five sets of associated keywords, because the different regions may have the same representative color.

The weight associated on each link of the keywords with the image represents the degree of relevance in which the keyword describes the linked image's semantic content.

From above discussed, we see that each image object is associated with three different level features respectively at the low-level feature, intermediate level feature and semantic level feature.

5. QUERY STRATEGY

From Figure 2, an image can be expressed at three levels: low-level features, intermediate level and semantic level. Accordingly, there are three different modes of user interactions involved in typical retrieval systems.

Firstly, we can index image and retrieval image from the semantic level. In this case, the user types in a list of keywords representing the semantic contents of the desired image. For instance, for the image shown in the figure 1, we can list our keywords: we want to find cloud in the center and the blue sky in the surrounding. In this way, we not only describe the semantic content but also spatial information in the images.

Secondly, at the intermediate level, we can use the combinations of the representative colors to describe the image. For the image shown in figure1, we can describe it as following: white in the center and blue in the surrounding.

Thirdly, we also can retrieval the images based on the low-level features which are obtained by using the algorithm provided in section 3.

6. RELEVANCE FEEDBACK

Recently, relevance feedback based CBIR techniques [16,17] merged as a promising research direction. The central idea of the relevance feedback is that it does not require a user to provide accurate initial queries, but rather estimate the user's ideal initial query by using positive and negative examples feedback by user. The fundamental goal of this method is to estimate the ideal query parameters accurately and robustly.

All the approaches [16,17] perform relevance feedback at the low-level feature vector level, but failed to take account into the actual semantics for the image databases. The inherent problem with these approaches is that the low-level features are often not as powerful in representing complete semantic content of images as keywords in representing text documents.

In our paper, we provide relevance feedback not only at the low-level feature level but also at the semantic level. At the low-level feature level, we try to improve the estimate of the" ideal query point" by moving it towards good examples point and away from bad examples points. Here, we update the estimates by using the Rocchio's formula given below (1) for sets relevant documents D_{x} and non-relevant documents D_{x} given by the user [18].

$$Q' = \alpha Q + \beta \left(\frac{1}{N_{R'}} \sum_{i \in D'_{R}} D_{i} \right) - \gamma \left(\frac{1}{N_{N'}} \sum_{i \in D'_{N}} D_{i} \right)$$
(1)

However, semantic based relevance feedback can be performed relatively easily compared to low-level feature counterpart. The basic idea behind it is a simple scheme to update the weights W_{ij} associated with each link shown in Figure 2. And the weight updating process is described as below:

- 1. Initialize all weights W_{ν} to 1. That is, every keyword has the same importance.
- 2. Collect the user query and the positive and negative feedback examples
- For each keyword in the input query, check to see if any of them is not the keyword database. If so, add them into the database without creating any links
- 4. For each positive example, check to see if query keyword is not linked to it. If so, create a link with weight 1 from each missing keyword to this image. For all other keywords that already to this image, increment the weight by 1.
- 5. For each negative example, check to see if any query keyword is linked with it. If so, set the new weight

$$W'_{ij} = \frac{W_{ij}}{2}$$
. If the weight W_{ij} on any link is less than 1, delete that link.

7. EXPERIMENTAL RESULT

We have implemented the presented algorithms in the image retrieval system designed by us. In this image retrieval system, it can support three modes of query which have been previously discussed in Section5. All the experimental evaluations were implemented under Microsoft Windows 2000. The machine we used is a Pentium-3650M PC with 126MB DRAM main memory. The software is Matlab 6.0.

We download about 10000 scene images in the BMP format from the website. Every representative color may have more than 1000 scene images. In order to guarantee that nature objects have obvious colors, these images mainly are the combinations of the following nature

that nature objects have obvious colors, these images mainly are the combinations of the following nature objects such as flowers, mountain, blue sky and so on. According to its semantic content of the image and the corresponding representative colors, each representative color can be associated with a few sets of keywords. For example, the keywords such as blue sky and sea waters may be classified into the representative color "blue". Through the algorithms provided in section 3 and section 4, each image and associated keywords are put into the database. Figure 3 show the main interface of our image retrieval system.

The experimental research is concerned primarily with the retrieval time and the accuracy of the retrieved images.

The first experiment is conducted to evaluate the retrieval time. Table 2 shows the retrieval time at three different levels.

Query level	Retrieval time (s)		
Query by Example image	132.13		
Query by intermediate color	0.300		
Query by keywords	0.270		

Table 2 the retrieval time

The pair (the Cod In Version in

From the above table, we can see that query by keywords can improve the retrieval speed greatly compared with the query by example image.

The second experiment is implemented in order to evaluate the retrieval accuracy based on the relevance feedback. A retrieved image is considered as a relevant one if it has the similar representative color to the query image. The retrieval accuracy is defined as:

$$R = \frac{re \text{levant images retrieved in top N returns}}{N}$$
 (2)

In our experiment, N is usually set to 20. Thus, the retrieval accuracy is denoted by R20.

From the database, four random images are selected as querying the whole image database at the semantic level and low-level feature respectively. Figure 4 shows the accuracy based on the relevance feedback at the semantic level, and figure 5 shows the accuracy based on the relevance feedback at the low level feature.

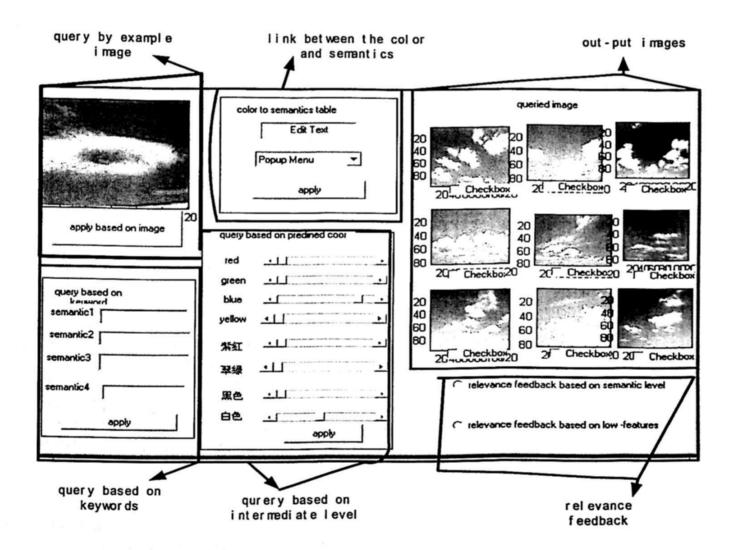


Figure 3: System Interface

As we can see from the results, our system achieves the high accuracy (>80%) after a few of relevance feedback for any given query. Unlike other methods where more user relevance feedback [17,18] may lead to lower the retrieval accuracy, experimental results demonstrate our method can be stable.

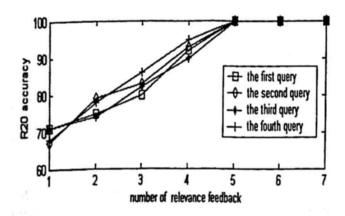


Figure 4. Accuracy based on the relevance feedback at the semantic level

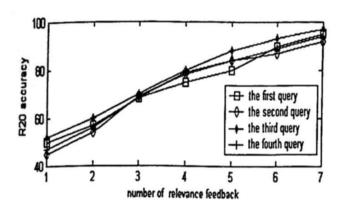


Figure 5. Accuracy based on relevance feedback at the low level

In addition to verifying the effectiveness of our proposed method, we also compared our method against the other techniques. We have chosen to compare our method with the retrieval technique used in [19]. The comparison is made through four random queries based on relevance feedback. Figure 6 shows the comparisons with the method in [19].

It is easily seen from results, our methods based on relevance feedback will improve the retrieval accuracy substantially.

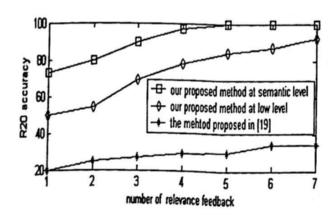


Figure 6. Performance comparison

8. CONCLUSIONS

In this paper, we provide the representative colors for indexing the images. Our method not only captures the color content of the images but also characterizes the spatial information of color in the image. In addition, we present a new approach called the WordNet to narrow the gap between the low-level features from the representative colors and semantic content of the images.

Finally, in order to improve the retrieval performance, we applied the relevance feedback to retrieve images both at the semantic level and at low-level feature. Experimental results show this method is stable. In the future work, we will further narrow the gap from the shape descriptor to semantics after we have obtained the representative colors.

REFERENCES

- [1] M. Flicker, H, Sawhney, and W. Niblack, "query bay image and video: The QBIC system," IEEE Computer, 28, 1995.
- [2] M, Swain and D. Ballard, "Color indexing" International Journal of Computer Vision, 7, pp 11-32, 1991.
- [3] C.C. Gotlieb and H.E. Kreyszig, "Texture descriptors based on co-occurrence martrices," Computer Vision, Graphics, and Image Processing, 51,1990.
- [4] J.Z. Wang et al, "Content-based image indexing and searching using Daubechies's wavelets," Int J. Digital libraries, 1, pp 311-328,1998.
- [5] S. Santini and R. Jain, "Similarity measures," IEEE Trans. On PAMI, 9,1999.
- [6] A, Natses, R. Rasto, "Walrus: a similarity retrieval algorithm for image database," SIGMOD Record, 28, 1999
- [7] D, Bimbo and P, Pala, "Visual image retrieval by elastic matching of user sketches," IEEE Trans on PAMI, 19, 1997.
- [8] Text of ISO/IEC 15938-3 Multimedia Content Description Interface-Part 3: Visual, Final Committee Draft, ISO/IEC/JTC1/SC29/WG11, Doc, N4062, Mar, 2001.
- [9] S.-F. Chang, W. Chen, and H. Sundaram, "Semantic Visual Templates: linking visual features to semantics Semantic Visual Templates: Linking Visual Features to Semantics," Proceedings 1998 International Conference on Image Processing (ICIP'98), Chicago, Illinois, October 4-7, 1998.
- [10] Mojsilovic, and B. Rogowitz, "Capturing image semantics with low-level descriptors", Proc. Int. Conference on Image Processing, ICIP 2001, Thessaloniki, Greece, September 2001.
- [11] Rong Zhao and William I.Grosky, "Narrowing the Semantic Gap---Improved Text-Based Web Document Retrieval Using Visual Features," IEEE Transaction on Multimedia, Vol,4, No.2, pp189-200, 2002

- [12] W. Niblack, R.Barber, W.Equitz, M.Flickner, E. Glasman, D. Petkovic, P. Yanker, C.Faloutsos, and G.Taubin. "The Q BIC project: Querying images by content, using color, texture, and shape.". In SPIE Conference on Storage and Retrieval for Image and Video Databases. Volume 1908, pp 173-187,1993.
- [13] T, Kato et al. "A sketch retrieval method for full color image database-query by visual example." In proceedings, 11th IAMPR International Conference on Pattern Recognition, pp530-533, 199
- [14] Markus Stricker, Alexander Dimai," Color Indexing with Weak Spatial Constraints" SPIE Conference, Feb, 96,San jose.

- [15] G.wyszecki and W. S. Stiles, Color Science: Concepts and Methods, John Wiley & Sons, 1982.
- 16] Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," in Proc. IEEE Int. Conf. On Image Proc., 1997.
- [17] Y. Ishikawa, R. Subramanya, and C.Faloutsos, "Mindread: Query Databases through multiple examples" in Proc. Of the 24th VIDB Conference, (New York), 1998. [18] RUI,Y, huang ,T.S, and Mehrotra, S "content-based image retrieval with relevance feedback in MARS," in proc. IEEE Int Conference. On image proceeding, 1997.
- [19] Rui, Y., Huang, T.S. "A Novel Relevance Feedback Techniques in Image Retrieval", ACM Multimedia, 1999.

MOSAICING SYSTEM WITH STRONG OF ROBUSTNESS IN CAMERA MOTION

SooHyun CHO, YunKoo CHUNG, JaeYeon LEE
Visual Information Research Team
Electronics and Telecommunications Research Institute. KOREA
jsh62860@etri.re.kr

ABSTRACT

Image mosaicing has been collecting considerable attention in the field of computer vision and photogrammetry. Unlike previous methods using a tripod, we have developed which can handle images taken with a hand-held camera to accurately construct a panoramic image. This paper proposes the automatic image mosaic system implementation, which it sees to use feature detection of the image, which is extracted a feature point adjustment from continuous two images hour automatically from pixel price of image and in order to accomplish. It will extract the feature point of each image and correspondence relationship of feature point in base matching point using perspective transform based Taylor series it will be cool, it uses and it initial matching. After initial matching the filtering it does the feature point, which goes wrong and removal after doing as a favor, it does a mosaic image creation.

KEY WORDS

image processing, biquadratic algorithm, affine transformation.

1. INTRODUCTION

In the last few years the interest in mosaicing has grown in the vision community because of its many applications. The automatic construction of large and high-resolution image mosaics is an active area of research in the fields of photogrammetry, computer vision, image processing, medical image, real rendering, robot -vision and computer graphics. Image mosaics involve aligning a sequence of image into a larger image and are an important issue in many virtual reality problems. Mosaicing is a common and popular method of effectively increasing the filed of view of a camera, by allowing several views of a scene to be combined into single view. The traditional approach, which uses correlation intensity based [1] [2] image registration, suffers from computation inefficiency and is sensitive to variations in image intensity. To improve the efficiency of image mosaics, we used a feature-based approach. Two images belonging to a planar scene are related by an affine transformation and perspective transformation using Taylor series. One of the images is used as the reference image, and the second image is aligned with the reference image. To find the coordinate

transformation between the two images, we first conduct corner detection to find the corners in these two images. Next, we perform a corner matching process to find the corresponding corner points corresponding point, which does a filtering processing from each image. We can estimate the transformation parameters using the Taylor series. This algorithm is implemented in the project, as well as image warping and blending between any overlapped images given Figure 1: Flowchart of mosaic create system, as well as image warping and blending between any overlapped images given by user. System flowchart is shown in Fig 1. In the next section, we give implementation and algorithm of the system and experimental and results. Finally, conclusions are given in section 4.

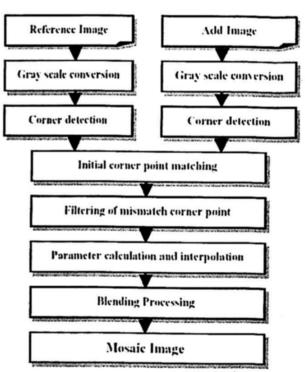


Figure 1. Flowchart of mosaic create system

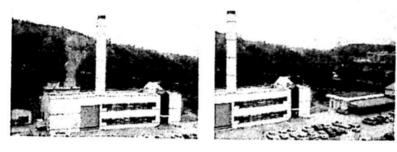
2. SYSTEM IMPLEMENTATION AND ALGORITHM

2.1 IMAGE ACQUISITION

In this paper, we used hand-held camera. The set may consist of two images taken of scene at different times, from different viewpoints. It is possible that movements of camera, (i.e. pan, tilt, rotation, translation, scale, shear). Outside scene the image which it requires from the of

course inside scene is possible. First, if image involves more than one band, say RGB, it will be converted to gray-scale image equation 1. That is I'() is formulate as follow. The reference image and add image of original images are shown in Figure 2.

$$I'(x,y) = 0.2999xI(x,y)_{red} + 0.587xI(x,y)_{green} + 0.114xI(x,y)_{blue}$$
(1)



a) Reference Image

b) Add image

Figure 2. Original Image

In this approach, we first detect corner points using the SUSAN (Smallest Unvalued Segment Assimilating Nucleus) principle in the each two images [3]. The response is processed to output the set of corners. The mask is places at each in the image and, for each point, the brightness of each pixel within the mask is compared with that of the nucleus, i.e. the center point. The comparison uses the Equation 2 here, \vec{r}_0 is the position of nucleus, r position of and other point, $I(\bar{r})$ the brightness of any pixel, t is the brightness difference threshold and c the output. Considering 7x7 pixel region window (w) that point in the center and optional point p(x, y) in the image, from inside the w different one point q(x, y) it does. I(p)and I(q) shows intensity values each point from p and q. Calculate the number of pixels within the circular mask which have similar brightness to the nucleus. Input image is gray scale value, T the brightness difference threshold and G the Gaussian value Equation 3. The results of corner detection image are shown in Figure 3.

$$C(\vec{r}, \vec{r_0}) = \begin{cases} 1 & \text{if } |I(\vec{r}) - I(\vec{r_0})| \le t \\ 0 & \text{if } |I(\vec{r}) - I(\vec{r_0})| > t \end{cases}$$
 (2)

$$S = \left(\sum_{i} e^{-\left(\frac{I(x,y) - I(x + dx, y + dy)}{I}\right)^{6}}\right)$$
 (3)





a) Reference image

b) Add imag

Figure 3. The results of corner detection image

2.3 INITIAL CORNER POINT MATCHING

Feature matching is a key component in many computer vision applications, for example stereovision, motion tracking, and identification. Of all possible features, "corner" is the most widely used: there two-dimensional structure providing the most information about image motion. A number of correlation-based algorithms attempt to find points of interest on which to perform the correlation. To match the correlation corner points between the two images, we used SSD (Sun of Squared Difference) method [4]. We measure the similarity between the two correlation windows using those detected corner points. Correlation scores are computed by comparing a fixed window in the second. SSD method is practical method, which produces reliable results with a minimum of computation time in comparison with the other method. Here I_1 is reference image, I_2 is add image, N is mask size. The result of Initial corner matching image is shown in Figure. 4

$$G = \sum_{i,j=N/2}^{N/2} (I_1(x+i,y+j) - I_2(x+i+dx,y+j+dy))$$
(4)

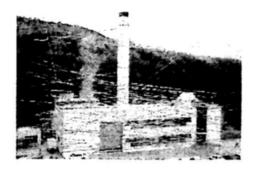


Figure 4. The result of initial corner matching image

2.4 FILTERING OF MISMATCH CORNER POINT

After initial corner point matching, in order to remove mismatches corresponding point, which does a filtering processing from each image General method is vector filtering [4], the median flow filtering, rotational cross-correlation filtering method in other to removes of mismatch point. This method is, which feature points definite local region of overlapping region, we used similarity values of rotation angle and length. The results of initial corner matching image and after filtering matching image for the line are shown in Figure. 5

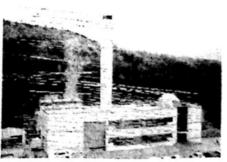


Figure 3. The result of filtering image

2.5 PARAMETER CALCULATIONS AND INTERPOLATION.

Using homogeneous coordinates, 2D planar projective transformation plus affine transformation method employs the following equations (5):

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$
(5)

(x,y) = image coordinates (X,Y,Z) = world coordinates

Geometric correspondence is achieved by determining the mapping function that governs the relationship of all points among a pair of images. There are several common mapping function models in image registration. The general form for the mapping function induced by the deformation is [x, y] = [X(u, v), Y(u, v)] where [u, v] and [x, y] denote corresponding pixels in I_1 and I_2 , respectively, and X and Y, are arbitrary mapping function that uniquely specify the spatial transformation. In registering I1 and I2, we shall be interested in recovering the inverse mapping function U and V that transform I2 back into $I_1[u, v]=[U(x, y), V(x, y)]$ In this section, we extend the results of affine registration and Biquadratic using a Taylor series. They include (1) 3-parameter rigid transformation (translation), (2) 6-parameter affine transformation (translation, scale, shear) using first order **Taylor** series. (3) 12-parameter **Biquadratic** transformation (translation, rotation, full motion) using second-order Taylor series. We leverage the robust affine registration algorithm to handle the more perspective registration problem. A local affine approximation is suggested by expanding the perspective transformation function about a point using a first - order Taylor series. In general, Taylor series equation (6):

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \dots$$

$$+ \sum \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$
(6)

The approximation holds in a small neighborhood about point (x_0, y_0) . As a result, affine transformation equations: (7)

$$u = Ax + By + C$$

$$v = Dx + Ey + F$$
(7)

The affine transformation with first polynomial, it is possibility of the image mosaic create according to translation motion. But in the case of rotation, pan and tilt motion, it is difficult to expect good results. It is possible that all movements of camera, (i.e. pan, tilt, rotation,

translation, scale, shear), in order to mosaic image creation. The traditional method uses perspective transformation using 8-parameters. A weak point of the method, the calculation process being complicated and the error scope is big. In order to overcome the method's defects, this paper uses second – order Taylor series possible full motion mosaic image create. The method is Biquadratic using 12-parameter. The Equations 8 and 9 are Biquadratic (second polynomial) using second-order Taylor series.

$$u=U(x,y)$$

$$=U(x_{0},y_{0}) + \frac{\partial U(x_{0},y_{0})}{\partial x}(x-x_{0}) + \frac{\partial U(x_{0},y_{0})}{\partial y}(y-y_{0})$$

$$+ \frac{\partial U^{2}(x_{0},y_{0})}{\partial (x^{2})}(x-x_{0}) + \frac{\partial U^{2}(x_{0},y_{0})}{\partial (y^{2})}(y-y_{0}) + \frac{\partial U^{2}(x_{0},y_{0})}{\partial x^{2}}(x-x_{0})(y-x_{0})$$

$$= Ax + By + C + Dx^{2} + Ey^{2} + Fxy$$

$$v=V(x,y)$$

$$=V(x_{0},y_{0}) + \frac{\partial V(x_{0},y_{0})}{\partial x}(x-x_{0}) + \frac{\partial V(x_{0},y_{0})}{\partial y}(y-y_{0})$$

$$+ \frac{\partial V^{2}(x_{0},y_{0})}{\partial (x^{2})}(x-x_{0}) + \frac{\partial V^{2}(x_{0},y_{0})}{\partial (y^{2})}(y-y_{0}) + \frac{\partial V\partial V}{\partial x\partial y}(x-x_{0})(y-x_{0})$$

$$= Lx + Hy + I + Jx^{2} + Ky^{2} + Lxy$$
(8)

Compared with Affine transformation and Biquadratic, Affine transformation is possible slow moving camera movement and translation motion be unchanged viewpoint of users. But optical rolling motion of the camera to free motion it cannot obtain good result. Biquadratic algorithm can obtain good result free motions of camera, rotation, rolling, zoom in and zoom out. For example affine transformation, given the four corners of a tile in observed image I_2 and their correspondences on reference image I_1 , we may solve for the best affine fit by using the least squares approach. We may relate these correspondence in the form U=WA (See Equation 10).

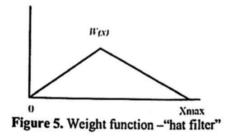
The pseudoinverse solution $A = (WTW)^{-1}W^{T}U$ is computed to solve for the six affine coefficients. After finding the transformation function for each successive image pair, we compute the transformation function of each image relative to the base image based on the associative of the matrix multiplication.

$$\begin{bmatrix} u_1 \\ \vdots \\ u_4 \\ v_1 \\ \vdots \\ u_4 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_4 & y_4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & y_4 & y_4 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \\ E \\ F \end{bmatrix}$$
(10)

Interpolation process can be classified into four methods (i.e. Nearest Neighbor interpolation, interpolation, Cubic Convolution, B Spline interpolation). In this approach, we used bilinear interpolation. A critical portion of warping images is interpolation for the resulting pixel values. Without a descent facility for interpolation, precise movement in the error minimization technique would not be possible. The process of bilinear interpolation requires four neighboring values to the coordinate at which we need to interpolation. The interpolation is performed in a separable manner as illustrated by the diagram below. First, the interpolation occurs in the X direction, followed by an interpolation in the Y direction.

2.6 BLENDING PROCESSING

The objective of blending is to provide a smooth transition between images and eliminate artefacts of minor misalignments resulting from parallax or imperfect pair wise registration. Blending is a process to reduce the discontinuities in intensity and color between images being composited. A simple feathering algorithm is used to blend all of the warped images. Generally blending algorithm is based on weighted average. There are two commonly used weight functions available weight with Euclidean distance and weight with triangle function. The paper used simple averaging and weighted averaging. This weighting function looks like a pyramid because it has unit weighting at the center of the image and falls off to zero at the edge. The result of using this function is the removal of virtually all traces of an edge between images, provided the registration was moderately successful. In this paper, we used sample weight function using "hat filter". "hat filter" are shown in Figure 6 and Equation 11.



$$W(X) = 1 - \frac{\left| x - \frac{X_{\text{max}}}{2} \right|}{\frac{X_{\text{max}}}{2}}$$
 (11)

3 EXPERIMENTAL RESULTS

To evaluate the performance of our proposed algorithm, we implemented in visual C++ language. The test images we used were obtained in the outside scene and inside using hand held camera for minimum time interval. Since only 2D motion parameters were estimated, the test images sere constrained to or close to planar image sequences. Each color image has 640.480, 320.240 pixels,

and image format has JEPG, BMP. In general, feature based correlation is sensitive to rotation motion. In order to feature matching, we need feature matching point more than 40 point. Using our proposed approach, part of matching took 60% of total processing time. Processing time results of two images for different processing are shown in Table.1. Total processing time is different each image, (i.e. image size, natural scene, inside scene, difficult scene...). When compared Figure 7.8 with Figure 9, the boundary two images loses becomes the natural image. Figure7 shows the result image using the affine transformation (6-parameter) and Figure 8 shows the result image using Biquadratic (12- parameter). Figure 9 shows the result image of blending processing. Figure 10 shows the results of test image. The display is shown in Figure 11.

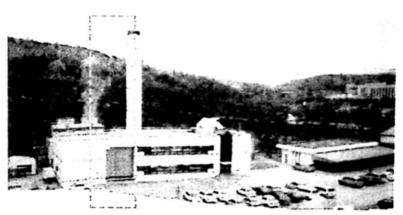


Figure 7. The result image using the Affine transformation (6-parameter)

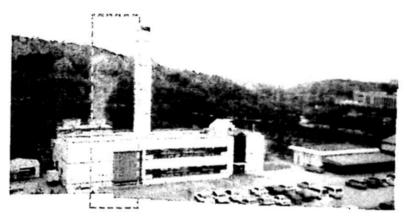


Figure 8. The result image using Biquadratic (12-parameter)

Image Size	Feature extraction	Initial matching	Filtering	Total Time
320*240	1	1.30	1	4
640*480	2	7	4	16

Table 1. Processing time

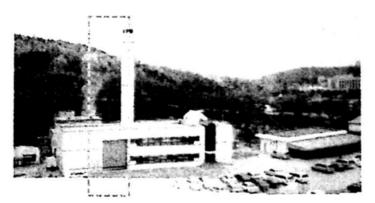
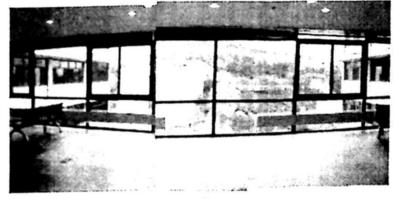


Figure 9. The result of blending image

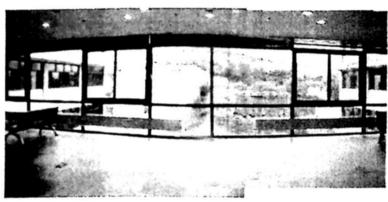


(a-1) Reference image

(a-2) Add image



(b) The result of mosaic image



(c) The result of blending image



(d-1) reference image



(d-2) add image



(c) The result of mosaic image Figure 10. The test image

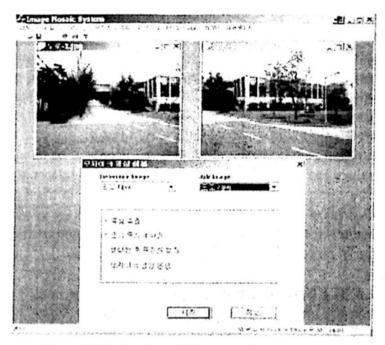


Figure 11. The interface of image mosaic create system

4 CONCLUSIONS

We have presented a distributed image mosaic system that can quickly, hand held camera full motion and automatically align a sequence of images to create a larger image. We see from Table 1 that the long time demanded for initial matching processing. A further direction of this research will be effective algorithm development for shorten the corner matching process time. Better automatic coarse registration techniques needed.

REFERENCES

- [1] R. Szeliski, Image Mosaicing for tele Reality applications, Cambridge Research Lab Technical Report, May 1984.
- [2] R. Szeliski, H. Yeung Shum, Creating Full View Panoramic image Mosaic and Environment maps, SIGRAPH 87, 1987, 251-258.
- [3] S.M Smith and J,M Brady, Susan a new Approach to how level image processing, Technical Report TR95SMSIc, 1995.
- [4] P.Smithm, D Sinclaif, R. Cipolla and K. Wood, Effective Corner matching, *British machine vision conference*, 2, ,1998, 545~556.
- [5] R. Szeliski, Video mosaics for virtual Environments, *IEEE Computer Graphics and Application*, 1966, 22-30.
- [6] N. Chiba, H. Kano, M. Minoh and M. Yasuda, Feature based image mosaicng, *IEICE*, Japan D-II, J82(10), 1589~1999 1999.

INCREASING RESOLUTION OF DIGITAL ELEVATION MODELS USING BICUBIC PARAMETRIC PATCHES

Rolando QUINTERO

Geoprocessing Laboratory, Centre for Computing Research, National Polytechnic Institute
Av. Juan de Dios Bátiz s/n Unidad Profesional "Adolfo López Mateos"
C.P. 07738, México, D.F., MEXICO
quintero@cic.ipn.mx

ABSTRACT

In this work, we propose an algorithm based on bicubic parametric patches to change the scale of DEM for real-time simulation processes preserving the semantics of the elevation data. We merge this algorithm with the discrimination algorithm presented in [1], which manages optimally the huge quantity of data contained in DEM. These algorithms allow navigating along the huge quantity of information within the elevation data and, at the same time, increase the level of detail for areas of interest. An application of this method is image processing. It may be used to zoom-in images with a non-linear re-sampling.

KEYWORDS

Digital Elevation Model, Parametric Patches, Image Resolution.

1. INTRODUCTION

Nowadays, Digital Elevation Models (DEM) have gained the terrain in applications for simulating natural disasters. The decision making tools for disaster prevention become important worldwide because the human lives could have been saved, if we have preventive information. Nevertheless, applications for the simulation of natural disasters require a huge amount of data. In many cases, the available data do not have enough quality for simulation processes. The National Institute of Statistics, Geography and Informatics of Mexico (INEGI) produces DEM with 50 meters of resolution [2][3], but some simulation processes require a better level of detail.

In this work, we propose an algorithm¹ to increase resolution of DEM for real-time simulation processes without change the semantics of the elevation data. Also, the application of this algorithm in image processing to

¹ A preliminary result of the development of the PhD thesis "Adaptive Methods for Generation of Digital Elevation Models" supervised by Dr. Serguei Levachkine

enlarge (zoom-in) images with a non-linear re-sampling is presented.

In our case, it is very important that the algorithm does not change the semantics of the DEM, thus it contains all relevant information about the model. We are working about the formal definition of DEM semantics. In Section 5 we describe the early results of our work about this formal definition. For the aims of this work, we will consider the semantics like the set of all the well-known values of elevation. The algorithm described in this work allows obtaining new data on the basis of these values.

In [1], we presented an application to manage the huge quantity of data contained in DEM for real-time rendering. In that application, we discriminated the less significant elevation data, without alter the semantics of these data. However, we cannot improve the quality of more relevant data to obtain additional information.

Using the algorithm presented in [1], we can solve the problem of 3D data representation and build virtual scenes, which are ready to navigate, either by simulations or by defined trajectories [6].

In the next section, we present some frameworks of the underlying theory of bicubic parametric patch representation. In Section 3, we give the pertinent considerations for the application of parametric patches and outline the proposed approach. In Section 4, some results and tests are presented and analyzed. Finally, the conclusions are outlined in Section 5.

2. BICUBIC PARAMETRIC PATCHES

In [1] we mentioned that elevation data can be represented as a polygon mesh. The step from polygon meshes to patch meshes is straightforward. If we consider a mesh of four-sided polygons approximating a curved surface, then a parametric patch mesh can be defined as a set of curvilinear polygons, which actually lie in the surface.

The definition given in [5] for a parametric surface (either B-spline or Bezier surfaces) Q(u,v) is in terms of two parameters, u and v, where $0 \le u \le 1$ and $0 \le v \le 1$, and the function Q is a cubic polynomial. The accurate values of the coefficients in the cubic determine the curve. A special and convenient way of defining these is to use 16 three-dimensional points known as control points. The shape of the patch is fully determined by the position of these points.

A bicubic surface is defined by

$$Q(u,v) = \sum_{i=0}^{n} \sum_{j=0}^{m} P_{ij} B_{i,j}(u,v) , \qquad (1)$$

where P_{ij} is an array of control points an $B_{i,j}(u,v)$ is a bivariate basis function. We can generate $B_{i,j}(u,v)$ in the following form

$$B_{i,i}(u,v) = B_i(u)B_i(v)$$
, (2)

where $B_i(u)$ and $B_j(v)$ are the univariate cubic basis function. The definition of these basis functions describes the type of surface to be generated. Next we give the definition of basis functions for the surfaces used in this work: Bezier and B-spline surfaces.

2.1. BEZIER SURFACE PATCHES

Bezier surfaces properties are extended from the Bezier curves formulation in [5][7], we can find out these properties. A bicubic surface is defined by its basis function. In the case of Bezier surfaces the basis function is defined by the Bernstein polynomials (for a set of n+1 control points):

$$B_k^n(u) = C(n,k)u^k(1-u)^{n-k}, (3)$$

where C(n,k) are the binomial coefficients:

$$C(n,k) = \frac{n!}{k!(n-k)!}$$
, (4)

then, the description of the Bezier surface is given by

$$Q(u,v) = \sum_{i=0}^{n} \sum_{j=0}^{m} P_{ij} B_{i}^{n}(u) B_{j}^{m}(v) .$$
 (5)

2.2. B-SPLINE SURFACE PATCHES

Similar to Bezier surfaces, the B-spline surfaces are characterized by its basis function. In B-spline surface the value of the patch depends exclusively on the values of a 4×4 array taken from the control point set. For a set of $n\times m$ control points we will have $(n-3)\times(m-3)$ patches.

Considering that, the B-spline basis function is defined as

$$B_0(u) = \frac{(1+u)^3}{6} ,$$

$$B_1(u) = \frac{3u^3 - 6u^2 + 4}{6} ,$$

$$B_2(u) = \frac{-3u^3 + 3u^2 + 3u + 1}{6} ,$$

$$B_3(u) = \frac{u^3}{6} ,$$
(6)

then, the patch Q_{ij} is defined for i=0,1,...,n-3 and j=0,1,...,m-3 by

$$Q_{ij}(u,v) = \sum_{k=0}^{3} \sum_{l=0}^{3} P_{i+k,j+l} B_k(u) B_l(v) .$$
 (7)

3. APPLYING BICUBIC PARAMETRIC PATCHES TO DEM

In [1] we presented the steps required to obtain and set up all data and parameters for real-time DEM rendering. In that work we defined a grid G(i,j) that contains the elevation data. To describe the rendering algorithm, we defined some parameters (their meanings are illustrated in Figure 1) first.

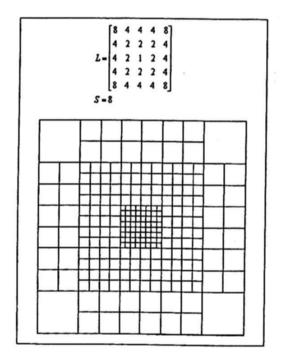


Figure 1. Visual representation of the parameters L and S of the algorithm. The grid at the bottom of the Figure illustrates how these parameters manage the number of polygons to be rendered.

- Matrix L of $H \times H$ defines the discrete level of detail to use. H is an odd number greater than 1, and $L[i,j]\neq 0$ for $1\leq i\leq H$ and $1\leq j\leq H$.
- Number S defines the optimization unit size; it means that it is necessary to optimize regions of S×S polygons.
- Vector o represents the observer position.

Using the parameters defined, we outline the algorithm.

```
RENDER (o)

1 (o_x, o_y) \leftarrow \text{RELATIVE-POSITION}(o, G)

2 \text{for } i = -\frac{1}{2}H \text{ to } \frac{1}{2}H

3 x \leftarrow o_x + S(i - \frac{1}{2})

4 \text{for } j = -\frac{1}{2}H \text{ to } \frac{1}{2}H

5 y \leftarrow o_y + S(j - \frac{1}{2})

RENDER-BLOCK(x, y, L[i + \frac{1}{2}H, j + \frac{1}{2}H])
```

```
RENDER-BLOCK(x, y, lod)

1 if lod > 0

2 for i = x to x+S step lod

3 for j = y to y+S step lod

4 RENDER-QUAD(i,j,lod)
```

RENDER-QUAD(i,j,lod)

- 1 RENDER-VERTEX(G[i,j])
- 2 RENDER-VERTEX(G[i+lod,j])
- 3 RENDER-VERTEX(G[i+lod,j+lod])
- 4 RENDER-VERTEX(G[i,j+lod])

The algorithm is based on taking bigger blocks of data as far as they are, and represent them by sampling G(i,j) in non regular steps along the grid. We can handle the number of polygons to be rendered by manipulating L-matrix and S- optimization size. In fact, the number of polygons to be rendered is given by,

$$N_{P} = \sum_{i=1}^{H} \sum_{j=1}^{H} \left(\frac{S}{L[i,j]} \right)^{2}.$$
 (8)

3.1. PARAMETRIC CONSIDERATIONS

Thus, we can discriminate data from DEM. However, in the case when we need more detailed data than those that are in G(i,j) we can apply parametric patches to obtain these intermediate data.

We will use the bicubic surfaces defined in Section 2. As a first approximation, we apply Bezier patches, which are commonly used in computer graphics, because they enable an efficient patch-splitting algorithm for rendering. But the main problem with Bezier patches is that the generated surface does not fit with the given points (control points). This is not desirable for the applications, thus we wish to increase the data resolution. We cannot fit all control points (the ones we have information) with the resultant Bezier patches. Nevertheless, by applying Bezier surfaces, we can compute how the whole data set behavior affects to a single point. See Equation 5.

On the other hand, we can apply B-spline patches. Thus, we can compute the inner points between the known ones (control points), using only local information. As we have mentioned before, a B-spline patch is always defined by a 4×4 control point array. So, with this type of surface we can find the new data without affecting the behavior of the whole data set.

3.2. THE ALGORITHM INTEGRATING PARAMETRIC SURFACES

To integrate the increasing resolution using parametric patches, we should modify the original algorithm in the following way. First, we should allow values less than one in L-matrix. Such values mean that we wish to obtain higher resolution for the block that is being rendered. On the other hand, we need to compute the parametric curve. This curve is stored in an alternate grid called Q. The values of Q will be defined by the control points and by the transformation matrix B (see [5] for matrix representation of Bezier and B-spline basis functions).

$$B = \frac{1}{6} \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix}. \tag{9}$$

Finally, the changes in the algorithm are the following:

```
RENDER-BLOCK(x, y, lod)
         if lod \ge 1
    2
             for i = x to x+S step lod
    3
                for j = y to y+S step lod
    4
                   RENDER-QUAD(i,j,lod)
    5
         else
    6
             Q \leftarrow \text{TO-SPLINE}(G, x, y, lod)
    7
            for i = x to x+S step lod
    8
               for j = y to y+S step lod
    9
                   RENDER-QUAD-SPLINE(i,j)
TO-SPLINE(G,x,y,lod)
         P \leftarrow \text{CONTROL-POINTS}(x,y)
         for u=0 to 1 step lod
    3
            U \leftarrow [u^3 u^2 u 1]
    4
            for v=0 to 1 step lod
               V \leftarrow [v^3 v^2 v 1]
    5
               O[u,v] \leftarrow U \times B \times P \times B^T \times V
```

RENDER-QUAD-SPLINE(i,j)

- 1 RENDER-VERTEX(Q[i,j])
 2 RENDER-VERTEX(Q[i+1,j])
 3 RENDER-VERTEX(Q[i+1,j+1])
- 4 RENDER-VERTEX(Q[i,j+1])

We only present the changes for integrating B-spline surfaces. Similar considerations must be taken for applying Bezier surfaces.

4. TESTS AND RESULTS

There have been done some performance tests with different data sets. As we have mentioned, the performance of the algorithm is constant, no matter the volume of elevation data involved. The results of the application are shown in Figure 2, using the proposed algorithm. The Figure 2a shows the result of rendering all elevation data. Figure 2b presents the result using our algorithm with the following parameters:

$$L = \begin{bmatrix} 64 & 32 & 16 & 8 & 16 & 32 & 64 \\ 32 & 16 & 8 & 4 & 8 & 16 & 32 \\ 16 & 8 & 4 & 2 & 4 & 8 & 16 \\ 8 & 4 & 2 & 1/2 & 2 & 4 & 8 \\ 16 & 8 & 4 & 2 & 4 & 8 & 16 \\ 32 & 16 & 8 & 4 & 8 & 16 & 32 \\ 64 & 32 & 16 & 8 & 16 & 32 & 64 \end{bmatrix},$$

$$S = 64$$

$$\Rightarrow N_P = 23524$$
(9)

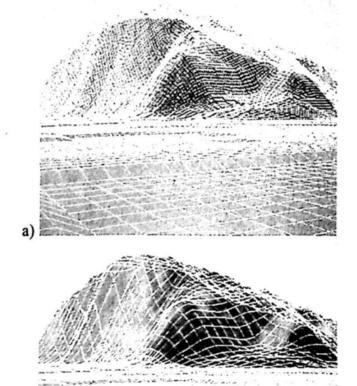


Figure 2. a) Result of rendering all elevation data. b) Result with the proposed algorithm.

Other data sets have been taken to test the zoom-in features of the algorithm. The results of this are shown in Figure 3. Figure 3a shows the data source used. Figures

3b and 3c present the results using our algorithm with B-spline and Bezier patches respectively.

As we can see, the B-spline patches allow us to interpolate the inner values of the image depending on their neighbors (because the B-spline basis function depends on 16 control points). The result contains the information about the local changes. The Bezier curve (Figure 3c) retrieves the information about all data, which allows applying smoothing to the image that not only depends on local values but also on the behavior of the whole data (memorizing of the data behavior).

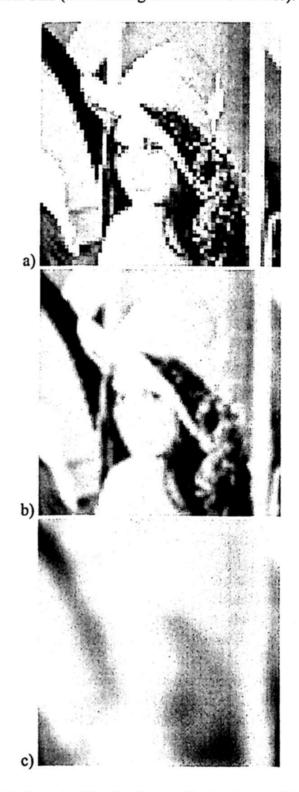


Figure 3. Example of the algorithm application (image of Leena on 64×64 resolution). a) Data source used. b) Image with B-spline 2x zoom. c) With Bezier 2x zoom.

In figure 4, we present the results of applying to a 32×32 image the most used algorithms to zoom images¹:

¹ The zoom factor applied was 16x.

Nearest neighbor gray-level interpolation (Figure 4a), Bilinear Interpolation [8] (Figure 4b) and Bicubic Interpolation [8] (algorithm proposed in this paper) (Figure 4c). Also, Figure 4 shows the image gradient to make notice the smoothness of the borders obtained with Bicubic Interpolation1.

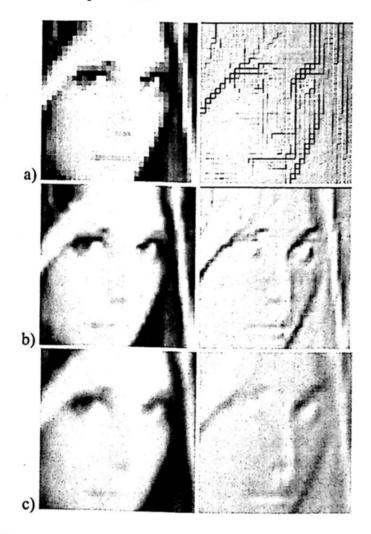


Figure 4. Results of applying different zooming algorithms to a 32×32 image. a) Nearest neighbor gray-level interpolation. b) Bilinear Interpolation. c) Bicubic Interpolation. The gradient obtained is shown as well.

5. CONCLUSIONS AND FUTURE WORK

In this work, the application has been designed to process spatial data in raster format. The implemented algorithm requires a low time to process the data, which are stored in DEM. The only restriction is that the images must correspond to the same scene. Two new digital images that are generated can be easily accessed in a faster way by the proposed rendering method.

The developed algorithm does not overload the processor, because it is very simple. Also, the rendering algorithm warranties a maximum number of elements to be rendered. In this way, we can manipulate this number (N_P) by varying the parameters of the matrix L. Additionally the algorithm does not modify the semantics of spatial data, it only discriminates or refines existing data without alter any characteristic of the elevation data.

This characteristic is essential to make the spatial semantic analysis.

An application of this algorithm in image processing is as follows: it may be used to zoom-in images with a non-linear re-sampling. Using parametric patches, it is possible to obtain the new values for the pixels between the known ones, containing local information (by means of B-spline patches) and global behavior (by means of Bezier patches) that improve the appearance of the enlarged image.

We are studying the application of recursive functions for the description of DEM, in order to obtain the semantics within the models. This can provide some significant advantages over the conventional ways of DEM description. First, it is possible to reach a very small granularity in the description and at the same time (another advantage) to obtain a very small (in amount of data) description. Thus, it can be possible to compress the huge amount of data in DEM. The application of the recursive functions can be seen as a model of adaptive approaches, in which an operation signature and a classification state (groups or classes) are registered in each recursion level. Under this scheme, the approaches can be refined every time adapting to the context of the land pattern. From this point of view, it is possible to say that the curves generated by the recursive functions adapt to the land shape.

DEM signatures as well as the classification states are obtained by applying recursively the classification function and calculating a measurement of the results2. The recursion finalizes when the measurement of the current result and the previous one are equivalent (the semantic has been preserved). With this information (the classification state and the operation signatures) it is possible to rebuild all the original DEM data.

As a future research; we must define the classification and measurement functions. As an early approach, we propose the classification function to be made of a set of basis functions. In the same manner as points, lines and polygons build a cartographic model, the basis functions will be used to describe DEM.

The selection of the basis function to be applied in each recursion level is an important problem to solve. We will have many basis functions then there will be many solutions in each level. To select the basis function and get the optimal result, we will apply different heuristics for the search of this optimal solution within the universe of possible solutions. In addition to the properties described in [9], the application of heuristic search is another reason that justifies the use of recursive functions.

¹ Note the smoothness of the border in the right cheek and eyebrow

² This measure is not defined yet, but it could be described as a classification quality rate.

ACKNOWLEDGMENTS

The authors of this paper wish to thank the Centre for Computing Research (CIC), General Coordination of Postgraduate Study and Research (CGEPI), National Polytechnic Institute (IPN) and CONACyT for their support. Additionally the authors wish to thank the reviewers for their pertinent comments.

REFERENCES

- R. Quintero, G. Guzman, M. Torres and M. Moreno, Spatial Analysis Based on 3D Rendering Techniques, Proc. of 6th Association of Geographic Information Laboratories Europe conference on Geographic Information Science (AGILE 2003), Lyon, France, April 24 26, 2003, 401-408.
- INEGI, Modelos Digitales de Elevación, Generalidades y Especificaciones, Instituto Nacional de Estadística, Geografía e Informática de México, Aguascalientes, México, 1999 (ISBN: 970-13-2511-7).

- INEGI, Normas Técnicas para la Elaboración de Ortofotos Digitales, Instituto Nacional de Estadística, Geografía e Informática de México, Aguascalientes, México, 1999 (ISBN: 970-13-2510-9).
- A. Watt and M. Watt, Advanced Animation and Rendering Techniques, theory and practice. Addison-Wesley, New York, 1995.
- A. Watt, 3D Computer Graphics, 2nd edn.
 Addiscon-Wesley, New York, 1993.
- M. Reddy and Y. Leclerc, Enabling Geographic Support in Virtual Reality Modeling with Geo-VRML, Journal of Cartography and Geographic Information Science, 26(3), 1999.
- D. Hearn and P. Baker, Computer Graphics, 2nd edn. Prentice Hall, New York, 1994.
- C. Gonzalez and R. Woods, Digital Image Processing, 2nd edition. Prentice Hall, New York, 2002.
- V. Alexandrov and N. Gorsky, Image Representation and Processing: A Recursive Approach, Kluwer Academic Publishers, London, 1985.

INTELLIGENT GIS TOOLS

A FORMAL MODEL FOR THE SPECIFICATIONS OF GEOGRAPHIC DATABASES

Sébastien MUSTIÈRE, Nils GESBERT, David SHEEREN
COGIT Laboratory - Institut Géographique National
2-4 av. Pasteur - 94165 Saint-Mandé Cedex
France
{sebastien.mustiere, nils.gesbert, david.sheeren}@ign.fr

ABSTRACT

In order to capture and manage their geographic databases, database producers use relatively huge and complex documents: the specifications. Actually, these are the only sources of information describing precisely what is the content of a database, i.e. what is the meaning of each part of the data schema, which object are captured and how they are represented. As such, they are an important tool to describe, manage, distribute, transform and federate databases.

Unfortunately, even if they can be precisely organised, these specifications are usually expressed in natural language and thus in a poorly formalised language. Their manipulation is then a hard task, and their automatic manipulation is almost impossible. In this paper we claim that formalising existing specifications is an important task and we propose an object-oriented model in order to do that.

This model contains two main parts: a part formalising which objects of the real world should appear in the database, and a part describing how these objects are represented in the database. These model rely on a set of typical criteria encountered in database specifications: geometric criteria, relationship criteria, and nature criteria.

KEY WORDS

Database, specification, formalisation, model.

1 INTRODUCTION

As for any database, the content of geographic databases is described by the data schema and sometimes by metadata. Common geographic words are used to name the different part of the schema, like names of classes, attributes and relations in an object-oriented model.

But, for each database, there exist an exact meaning beyond these words. For example, if a class is named River, it may actually designate only permanent river in a database, or only natural river excluding human-made canals in another database. More, the name river may designate only rivers that are wider than 10 meters. Indeed, a database is associated to a certain level of detail, and only relevant objects are captured.

How a real world object is represented in the database is also a precise thing for each database. For example, the geometric line describing the river may be the axis of the river or one of its border.

All these definitions are stored in particular documents: the database specifications. These specifications are first used by database producers to develop precise guidelines for data capture. They can also be distributed to data users, in order to help them to understand in detail what the database contains.

Sometime, different specifications can exist for one database: the content specifications describe which objects of the real world should be represented in the database, and the specifications of data acquisition which detail the process for data capture into the system. Furthermore, subsets of specifications are usually provided to users along with the delivery of the spatial databases.

Actually, these documents are the only sources of information describing precisely what is the content of a database. Metadata may also exist, but in practice they are mainly used to describe global criteria like the reference system used, the geographic extent, the date of the capture, the source for capture, some evaluations of precision, and so on. They do not describe the exact semantic of each concept used in the database.

Unfortunately, even if they can be precisely organised, specifications are usually expressed in natural language and thus in a poorly formalised language. Their manipulation is then a hard task, and their automatic manipulation is almost impossible.

In the next section, we explain some possible use of more formalised specifications. Then, we propose a generic model of specifications for geographical databases. In section 3, we explain how to link data schemas to our model of formal specifications. In section 4, we detail basic elements of the model. This model has been defined from the study of existing specifications used in IGN, the French National Mapping Agency.

Before that, let us make a remark about the notion of ontology that receive more and more attention in the GIS community. This work may be related to the work on spatial ontologies. But, due to the numerous meanings of this term, we prefer to avoid it. If an ontology is thought of as a way to represent the intended meaning of a vocabulary [[1],[2]], this work is about the development of an ontology from database specifications. But if an ontology is thought of as a way to express the links between different vocabularies in order to facilitate the understanding between systems (in other words, a kind of thesaurus) [[3]], this work is not about the direct development of ontologies, even if it may take advantage of them. Indeed, our purpose is really a formalisation of specific definitions encountered in each database specifications.

2 ON THE UTILITY OF FORMALISING SPECIFICATIONS

In term of quality, it is recommended that the specifications present some characteristics. Among these, [[4]] mention that the specifications must respect all expressed needs, avoid useless details and must be clear, precise and complete, without ambiguities. But specifications are generally represented in natural language and described in a relatively informal way, dependent on the considered database. Moreover, specifications associated with different databases may be not organised in the same way. This lack of formalisation raises some important issues:

- it makes sometimes long to browse specifications in order to find a special piece of information;
- it leads to different interpretations due to the imprecise descriptions;
- it complicates their comparison due to the heterogeneity of their description;
- it induces some difficulties for automatic exploitation of specifications.

This issues have consequences on different tasks in the GIS area.

2.1 DATA ACCESS

The first task data users are confronted with is data access. Beyond technical and commercial issues, users must first be able to understand the content of a database in order to know if it is adapted to answer to their needs: this the notion of "fitness for use". They must be able to answer to questions such as "Does this database contains the information I am looking for?", "Are these objects described with an adapted level of detail, or is it over or

under-detailed ?", "Between these two databases, which one is the more adapted to my needs?", "What should I filter in this database in order to make it simpler and more adapted to my needs", and so on.

All this tasks necessitates to understand clearly what is the meaning of the elements of the database. But present metadata can not fulfil these requirements as they usually describe the databases in a too global and limited manner.

In this context, the more the specifications are clearly organised, and the more they are homogenised, the easier it will be for users to understand and compare databases.

More, in the long run, we may be able to develop decision support system able to advise users with suitable data for their needs. Such a system can only be developed if, on the one hand, we are able to formalise user needs and, on the other hand, we are able to formalise what a database contains and for what it can be useful. Formalisation of specifications is thus one of the bricks necessary to build such a system.

2.2 DEVELOPING AND MANAGING DATABASES

From the data producer point of view, a more formal model of specifications could also help database designers to better define their specifications. As any model, this model is of course constraining and may be a limitation to the possible specifications that can be defined. But these limitations provide also a framework to help the definition of specifications more organised, better understood and less ambiguous.

A common model for all specifications produced by one data producer (which is actually not the case) has also the advantage of facilitating the exchange of knowledge in the company.

Another advantage of such a model is to facilitate the maintenance of specifications. For example, if specifications are updated, a model will allow to point directly at the part of the specifications that have changed when one must inform database users and producers.

More globally, we believe that working on a model of the specifications framework is good way to analyse deficiencies of current databases to better define them.

2.3 UNIFYING DATABASES

Many geographical databases exist to represent a same part of the world, seen at different levels of details and with different points of view. Unfortunately these databases are relatively independent, and there exists a growing need for the unification of different databases into a single one making explicit the relations between them [[5],[6],[7],[8]]. This unification can in particular

lead to a database managing multiple representations [[9],[10]]. Such unification would help to:

- Maintain the databases and propagate updates [[11]]
- Perform some quality analysis, by using one database to control another one or by identifying inconsistencies [[12],[13]].
- Increase the potentiality of applications development from these databases. Some applications can take advantage of using databases with multiple representations [[14],[15]].

An important issue for the unification is the ability of understanding what are the differences between the databases, which first requires to understand what exactly contains each database. These informations are actually described in the specifications. Their formalisation should allow, as we explained for the issue of data access, to better understand separated databases. Additionally, it should facilitate their comparison, as the specifications will be defined with a common model.

A relatively automatic processing of these specifications appears also very useful. These specifications, once presented in a formal model, can be used to guide an integration process of several spatial databases. In such a process, their automatic exploitation could simplify the research of the correspondences between them. First, their description in a formal language could reduce their lack of precision and therefore the inconsistencies of representations between the databases. Second, their formalisation allows displaying them to the users in a more organised way. Then, their formalisation should allow an automatic comparison of them. For example, one could automatically determine that the specifications of the class Stream in one database are more restrictive than the specifications of the class River in another database. Thus, we know that each object of the class Stream should have a corresponding object in the class River, but the reverse is not true. More precisely, we can determine how the classes correspond. For example we could determine that River and Stream correspond to each other only when the length of the river or stream is greater than 200 m, or when the river or stream is permanent.

3 LINKING A FORMAL MODEL OF SPECIFICATIONS TO A DATA SCHEMA

In the preceding section we claimed for a formalisation of specifications. In this section and the next one we propose a model for this formalisation. This model has been developed from the study of actual specifications taken out of two databases from the IGN (French National Mapping Agency): BDTopo (a topographic database with a metric resolution) and BDCarto (a road network oriented database with a decametric resolution). This model is expressed in UML. Let us notice that, for testing the feasibility of the model, we instantiated in the XML language from the river and road network of the databases.

3.1 A METAMODEL

Our specification model is linked to the metamodel of the geographical database, represented in a simplistic but quite generic way by three classes: "Class," "Attribute" and "Association" (see figure 1). We do not, for the sake of compatibility, use more details about it since we need to be able to handle databases in any format.

Let us notice that this model is a metamodel where the classes of the database (road, river, etc.) are thought of as instances of the metaclass *Class*. Similarly the attributes and associations are thought of as instances of the metaclasses *Attribute* and *Association*. The specifications about these classes, attributes and associations are expressed through the other elements of the model, described hereafter.

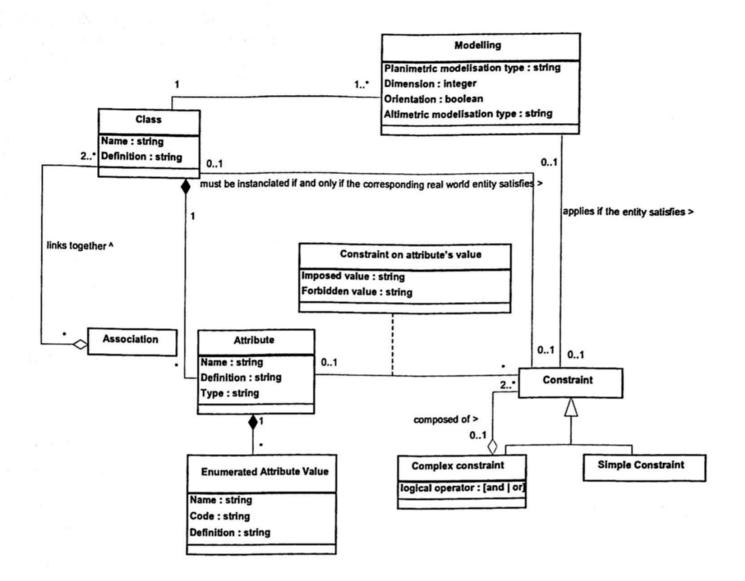


Figure 1. An extract of the general specifications model: the meta-model of the spatial database with the existence constraint, the modelling constraint and the constraint on attribute's value.

3.2 CONSTRAINTS OVER REAL WORLD OBJECTS

The main part of the model is a class hierarchy that allows to express a condition concerning a real world entity, such as "being greater than 100 meters" or "being inside a urban area". It consists of a base class "Constraint" and its subclasses.

It is important to notice that these "constraints" are not classical integrity constraints for the database. They are not constraints on the objects of the database (like "this attribute must be within such range of values, or an object of this class can not intersect an object of this class...). These constraints are used to express conditions over real world object, like "a real world river is represented in the database, if it is more than 10 meters wide".

To represent every possible condition that a real world entity could be asked to satisfy, we specialise the class "Constraint" into "Simple constraint" and a recursive "Complex constraint" which is composed of constraints

linked together by a logical operator: AND or an inclusive OR. This way, we get a tree structure whose nodes are complex constraints, containing logical operators and linked to their children through the "is composed of" relation, and whose leaves are simple constraints. This allows to express complex constraints such as "being inside a urban area and either longer than 100 meters or wider than 30 meters," which is a complex constraint composed of a simple constraint and another complex constraint, in turn composed of two complex constraints. Simple constraints specialise in turn into three subclasses "Geometric constraints," "Relationship constraints" and "Nature constraints". Before describing these constraints more in detail in section 4, the following section describes how the constraints are used to express specifications on the geographic database.

3.3 WHAT AND HOW TO REPRESENT IN THE DATABASE: EXISTENCE AND MODELING

The class of constraints over real world object is the pivot class to express specifications. These constraints are linked to elements of the metamodel by several ways, either to express existence constraints, modelling constraints, or attribute constraints.

Constraints are first used to define the necessary and sufficient condition under which a real world entity should be represented in the base, like "forests are taken into account in the database if and only if their area is greater than 8 ha." This condition is called the "existence constraint" of the class, and is linked to it through the "must be instantiated if and only if..." association (figure 1).

Another purpose of the constraints is to indicate which modelling applies in which case when a class can accept several ones, as in "a building is represented by its ground surface if it is more than 50 m² large, else its centre point is used." This kind of constraint is named a "modelling constraint" and is linked to the "Modelling" class (figure 1). To indicate the way the real world entity's geometry is represented in the database, we use the "Modelling" class, whose attributes allow to specify, between others, the spatial dimension of the database object and the way its geometry is deduced from the real entity's one (for example "outline of the roof" or "ground surface"). A class may have several possible modellings, in which case each one should be linked to a modelling constraint in order to indicate how is chosen among them. For example, let the following specification for the "River" class be: "a river is represented by its axis if it is less than 20 m wide, else it is entered as a surface" We get two instances of "Modelling," one with dimension

"line" and planimetric modelling "axis," linked to the geometrical constraint "width less than 20 m," and the other one with dimension "surface" and planimetric modelling "border," linked to the opposite constraint "width more than 20 m."

Constraints can also be used to express constraints on attribute values. Attributes are given a name, definition and type, which is generally sufficient. For example, "number of lanes: integer" does not a priori need to be further detailed. But there are cases however where it is not sufficient. Sometimes it is necessary to define how the attributes are instantiated. For examples constraints can be used to express specifications such as "attribute 'Accessibility' of the class 'Road' is set to 'Impossible' f the width is less than 2 meters or if the road is unpaved

4 BASIC ELEMENTS OF THE FORMAL MODEL

We defined a constraint model, in order to formalise the content of the constraints, and facilitate their comparison (see figure 2). The elements of this part of the model are described in the following sections.

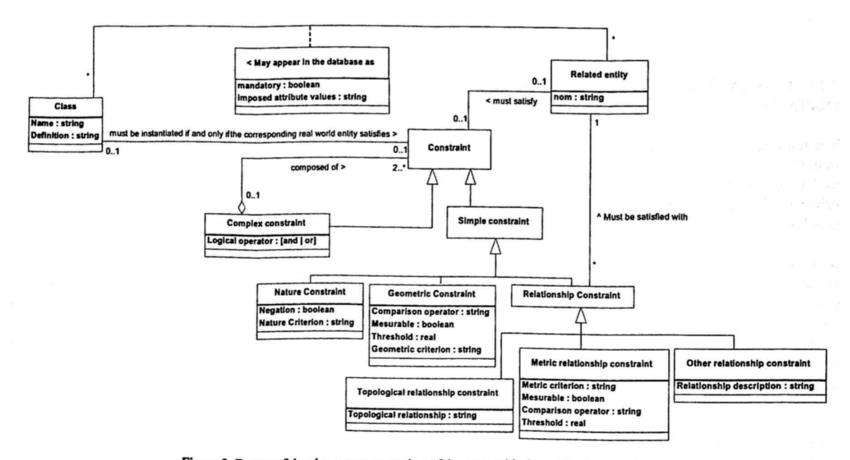


Figure 2. Extract of the elementary constraints of the geographical specifications model

4.1 GEOMETRIC CONSTRAINTS

"Geometric constraints" allow to specify a geometric criterion (length, area, etc.) and to demand that its value be above (or below) a given threshold. These constraints are commonly met, either as existence constraints or as modelling constraints.

These constraints are the easiest to formalise, as they are all relatively similar and clearly defined. The set of metric criteria used is restricted and a predefined set of possible values for the "geometric criterion" can thus be found to cover most cases: length, width, area, height. These constraints are also easy to compare, as no further knowledge is needed to know that "the width must be more than 20 meters" is more restrictive than "the width must be more than 10 meters".

4.2 NATURE CONSTRAINTS

"Nature constraints" are constraints specifying the nature of the entity. For example, we can describe there a condition like "the building must be a house" or "the forest must be composed of conifers".

It is difficult to formalise these constraints more than by just expressing the fact that the entity is or is not something. Thus, without additional knowledge, it is difficult to compare these constraints, except that some constraints are the opposite of some others. In order to go further, one would have to use some thesaurus of geographical terms [[3],[16]]. Those thesaurus can express that the words "unpaved road" and "track" are partly similar, or that "house" is a kind of "building," or that "lac" is the French translation of "lake." These knowledge can be of great help to compare nature constraints, even if some ambiguities can appear, as certain identical terms can be understood differently in different databases.

4.3 RELATIONSHIP CONSTRAINTS

"Relationship constraints" allow to handle conditions that do not bear on the entity itself but rather on the surrounding ones. For example we could have a constraint like: "a path is taken into account only if it leads to a house." Among these relationship constraints, we further distinguish metrical relationships, like "to be more than 50 m away from a house," and topological relationships, like "to be inside a forest," which are common. To specify with what the relationship is supposed to occur, we use the dedicated class "Related entity" that designated real world entities (and not database entities). If giving the name of the related entity does not suffice, for example if we do not want it to be any kind of road but to be a surfaced one, we link it once again to another instance of "Constraint" expressing "to be a surfaced road."

One frequently encountered special case is when the specifications' text directly refers to the database rather than to the real world, as in: "culs-de-sac are taken into account if they lead to an object to the class of of theme A (particular buildings)." In this case, we still create an instance of "Related entity" (named particular building) and we link it to the mentioned database classes (those of theme A) by the "May appear in the database as" association class with "mandatory" attribute set to true. In other cases, in order to facilitate future automatic processing of the data, we also link the entity to the database classes that may represent it (if there are some), but we then set the "mandatory" attribute to "false" to express that this link is only informative and not really part of the specifications. In some other cases, the related entity may not be linked at all to the database classes, if no class is used to represent this entity, as in the constraint "being outside the town" if the concept of town is not represented in the database.

These kind of constraints may be very precisely formalised if we use some models of spatial relations to describe them, like models to describe topological relations [[17],[18]] or distance and orientation relations [[19]]. These constraints may also be compared by using some models of conceptual neighbourhood between relations [[20],[21]].

4.4 OTHER CONSTRAINTS TO BE INTEGRATED IN THE MODEL

This classification of constraints was not established of course immediately, due to the above-mentioned lack of formalism in the specifications. In order to validate the model, we then instantiated it on some excerpts of our two sets of specifications. Instantiation of the model brought to light two constraint types which had not been taken into account: representativeness constraints, and constraints relative to attribute value changes.

Representativeness constraints play almost the same role as existence constraints, except that they apply when the database object does not correspond to a single real world entity but is rather meant to be representative of a group. At the moment, we distinguish two subcases: there is one single object which represents the group (aggregation), or there are several of them (representative elements), in which case there is generally no one-to-one or even oneto-several possible mapping between database objects and real world entities, only a group-to-group relation. An example of aggregation is this excerpt of specification regarding the class "Reservoir": "if there are many reservoirs close together whose diameter is less than 10 m, then they are merged." An example of representative elements could be buildings in a city in a generalised database.

Constraints regarding attribute value changes generally appear as geometrical ones: "attribute changes for the

road section class are not allowed for a length less than 100 meters." These constraints are neither existence constraints explaining which objects should be represented, nor attribute constraints explaining how to set the value of the attributes. They describe when to take into account the changes of attribute to determine the number of database objects to be created along the road.

These new constraints should now be integrated into the model, whose validation is going on. We did not either study in depth specifications and constraints concerning database associations.

5 CONCLUSION

In this paper we presented a model for formalising specifications of geographic databases. This model has two main parts: a part to define which object of the real world are represented in the database, and a part to define "how they are represented. This model rely on a set of predefined constraints on real world objects, like geometric, relationship and nature constraint.

This model is still under development and is part of an ongoing work. We intend to use it in a process of unification of databases. First, the model is used in the context of matching data schemas. Indeed, it is necessary to compare specifications in order to match precisely schemas. Second, the model is used in the context of analysing differences between databases [[22]]. Indeed, it is important to differentiate between "normal" differences justified by the differences in the specifications, and "wrong" differences due to errors in the databases.

More globally, we believe that database specifications are rich documents insufficiently used. We hope that going in the direction of their formalisation is a good avenue to increase their utilisation.

ACKNOWLEDGEMENT

We thank Anne Ruas who suggested the exciting idea of working on the formalisation of geographic databases specifications.

REFERENCES

- Guarino N. 1998. Formal Ontology and Information Systems. IOS Press, Amsterdam.
- [2] Fonsesca F.T. and Egnehofer M.J. 1999. Ontologydriven geographic information systems. In Proceedings of ACM GIS'99.
- [3] Smith B. and Mark D.M. 2001. Geographical categories: an ontological investigation. In International Journal of Geographical Information Science, vol.15, n.7, pp.591-612.

- [4] Fougères A.-J. and Trigano P. 1999. Construction de spécifications formelles à partir des spécifications rédigées en langage naturel, In *Document numérique*, 3(3-4), pp. 215-239
- [5] Devogèle T., Parent C. & Spaccapietra S. 1998. On spatial database integration. In *Int. Journal of Geographical Information Science*, vol.12, n.4, pp.335-352.
- [6] Weibel R. & Dutton G., 1999. Generalising spatial data and dealing with multiple representations. In Geographical Information Systems Principles, Techniques, Application and Management, 2nd Edition, Longley P.A., Goodchild M.F., Maguire D.J. & Rhind D.W. (eds), Vol.1, pp.125-155
- [7] Vangenot C., 2001. Multi-représentation dans les bases de données géographiques. PhD. Thesis, Ecole Polytechnique Fédérale de Lausanne.
- [8] Ruas A. 2002. Pourquoi associer les représentations des données géographiques?. In Généralisation et Représentation Multiple, Ruas (ed), Hermes-Lavoisier, pp.45-54.
- [9] Bertolotto M., De Floriani L., Marzano P., 1995. A Unyfying Framework for Multilevel Description of Spatial Data, Spatial Information Theory - A Theoretical Basis for GIS. In Lecture Notes in Computer Science, n.988, A.U. Frank, W. Kuhn (Editors), Springer-Verlag, 1995.
- [10] Rigaux P. and Scholl M. 1995. Multi-scale partitions: Applications to spatial and statistical databases. 4th International Symposium on Advances in Spatial Data, SSD'95, pp.170-183.
- [11] Lemarié C. and Badard T. 2001. Cartographic database updating. In *Proc of* International Cartographic Conference (ICA 2001), vol. 2, pp.1376-1385.
- [12] Egenhofer M.J., Clementini E. & Di Felice P. 1994. Evaluating inconsistencies among multiple representations, In Proceedings of the Sixth International Symposium on Spatial Data Handling (SDH'94), Edinburgh, Scotland, pp. 901-920.
- [13] Bel Hadj Ali A., 2001. Positional and Shape Quality of Areal Entities in Geographic Databases Quality information aggregation versus measures classification. In Proc of ECSQARU Workshop on Spatio-Temporal Reasoning and Geographic Information Systems, Toulouse 2001.
- [14] Car A. and Frank A. 1994. Modelling a Hierarchy of Space Applied to Large Road Networks. In Proc. of Int. Workshop on Advanced Research in Geographic Information Systems, Springer-Verlag, pp.15-24
- [15] Yuan M. and Albrecht J. 1995. Structural Analysis of Geographic Information and GIS Operation from a User's Perspective. In Proc. of COSIT: Spatial Information Theory, pp.107-122.
- [16] Kavouras M. and Kokla M. 2002. A method for the formalization and integration of geographical categorizations. In *Int. Journal of Geographical Information Science*, vol.16,no.5,pp.439-453.

- [17]. A Topological Data Model for Spatial Databases. In *Proc. of 1st Int. Symp. on the Design and Implementation of Large Spatial Databases*, LNCS 409, pp.271-286. Springer-Verlag.
- [18] Worboys M.F., 1996. Metrics and topologies for geographic space. In Advances in GIS research II: proc. of 7th International Symposium on Spatial Data Handling, Kraak and Molenaar (eds), Taylor and Francis, p.365-375.
- [19] Hong J.-H. 1994. Qualitative Distance and Direction Reasoning in Geographic Space. PhD. Dissertation, University of Maine.
- [20] Bruns H.T. and Egenhofer M.J. 1996. Similarity of Spatial Scenes. In Proc. of 7th Int. Symposium on Spatial Data Handling (SDH), Taylor and Francis, pp.173-184.
- [21] Paiva J.A. 1998. Topological equivalence and similarity in multi-representation geographic databases. PhD Thesis in Spatial Information and Engineering, University of Maine.
- [22] Sheeren D. 2002. L'appariement pour la constitution de bases de données géographiques multirésolutions vers une interprétation des différences de représentation. In Revue Internationale de Géomatique, Vol.12, n.2/2002, pp.151-168.

ASSISTING USERS IN SELECTING AND RESTRUCTURING DATA SETS

Sandrine BALLEY
Laboratoire COGIT - IGN

2, avenue Pasteur, 94 165 St Mandé Cedex
France

sandrine.balley@ign.fr

ABSTRACT

This paper addresses the issue of geographic data fitness for use. We want to assist users in finding data precisely adapted to their application's requirements. An approach is proposed to provide customised data sets, thanks to a system helping users to choose and to restructure existing data. This system mainly relies on a data set description model. This paper describes the issue and chosen approach for this beginning research project, but no result is presented yet. The core elements of the data set description model are presented at the end of the paper.

KEY WORDS

Geographic Data Access, Fitness for Use, Data Set Customisation, Data Set Description Model.

1. INTRODUCTION: USER ACCESS TO FITTED-FOR-USE GEOGRAPHIC DATA SETS

1.1. GENERAL IDEAS

Users of geographic data are not always experts in the geographic information domain: they may not be aware of the multiple slight differences making each geographic product more or less adapted to their application.

Some research work has been carried out to customise geographic processing or GIS tools [1] [2]. On the other hand, with regard to data sets, users up to now have had to do with rigid products which are proposed exactly as they were designed. However, user requirements are various, even concerning one and the same data set, as can be noticed in a pricing experiment described in [3].

This paper introduces a starting research project that aims at proposing data sets customised for user need, thanks to an interactive specification of data extraction and restructuring. Important research has been carried out concerning user access to data set description, especially through metadata [4]. The SDI Cookbook [5] defines three levels of metadata for spatial data infrastructures:

- The discovery level provides global information introducing the geographic product. It should enable the user to know which products exist.
- The exploration level gives details about data sets composing the product, and should enable to know whether the data will meet general requirements of a given problem.
- The exploitation level provides information required to load and use the data in the final application. It includes a data dictionary, the data schema, reference system and geometric characteristics, etc.

As shown in Figure 1, the metadata level we are focusing on is located on the exploration level and on part of the exploitation level. The user we consider is aware of existing data products (the discovery step is over). He wants to select among one of those the data containing the information specifically required by his application. He also needs to adapt the selected data, so that the information has the adequate modelling and can be used for the planned application as soon as the data are distributed.

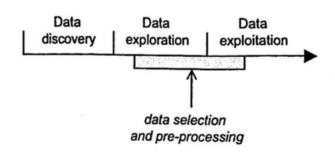


Figure 1. Levels of metadata in user access

The following example illustrates the process of user access and points out encountered difficulties.

A risk manager is looking for geographic data supporting flood simulations. For this purpose, he needs to have a representation of bridges, rivers and their crossing points. He has heard about a vector database called GeoBase and wants to check if it suits his requirements. Several questions must be answered.

- Does the information content correspond to the required content? To answer this question he checks the list of GeoBase feature types found on a metadata server: GeoBase does not provide any "bridge" feature type, nor any attribute domain including "bridge" or any synonym. However, had he been able to read the product specifications, he would realise that bridges are represented in GeoBase: a Road-Section feature whose attribute level has value 2 corresponds to a road bridging another feature. To know if it bridges a river or something else, the geometric intersection must be tested with River-Section features. So, the information is represented in an implicit way. It can be made explicit through a simple operation available in most GIS software. Moreover, the user cannot find any "road" feature type. Information is geometrically divided into "road sections", and represented by a large number of feature types distinguishing main roads, motorways, cycle tracks, etc.: GeoBase contains the required information, but its level of detail exceeds the user need.
- Is the data schema adapted to the application? In our example, it is not. However, the schema can be modified by some schema transformation operations: the Road-Section feature class can be split up by filtering the values for level attribute, so that the Bridge feature class appears. The River-Section features can be aggregated to generate River features. Moreover, the risk manager has some classical applications to perform: the River feature selection and join (to add some water level data), and the path processing on rivers. As GeoBase is a vector database, feature selection and join are enabled. On the other hand, rivers in GeoBase are surface features and do not constitute a network. The application is not feasible on GeoBase in its present state.

This example shows that describing a data set content and fitness for use is far more complex than listing feature types. The next two subsections further describe the problems of data selection and pre-processing.

1.2. ISSUES INVOLVED IN SELECTING ADEQUATE DATA SETS

The selection of a data set implies first of all the selection of an information content: a user specialised in water applications is likely to select the data set providing the most detailed representation of water bodies, whatever structure the data set has. The possibility to extract only needed information from a data set contributes to its fitness for use.

A first difficulty hampering selection is the lack of detailed and available descriptions of data sets content. Sections 1.2.1 and 1.2.2 expose this problem by

distinguishing content descriptions provided by metadata and by data sets technical documentation.

1.2.1. Using metadata to select adequate data sets

Metadata are the main way for users to explore available data sets, but the description provided by metadata standards do not particularly focus on the product content [6]. Metadata above all describe the data set in its globality (e.g. its global quality and spatial extension), and do not insist on the individual description of data set elements. The ISO 19115 standard [7] already provides three interesting metadata entities in this context: a subset of feature types occurring in the data catalog (MD FeatureCatalogDescription), an application schema in a graphic file (MD ApplicationSchemaInformation) and an image file illustrating a sample of the data set (MD BrowseGraphic). However, all of them are not mandatory: most of the time, data providers only fill the features type list. Moreover, these metadata entities provide limited information: geometric and semantic representations chosen for the data set features are not detailed.

1.2.2. Using description provided by technical documentation to select adequate data sets

The complete technical documentation written by the data producer (data schemas and product specifications) could provide potential users with more information, but, as we explain in this section, this technical documentation is not distributed and anyway would be difficult to interpret.

1.2.2.1. The technical documentation form

Data schemas.

Database terminology distinguishes three types of schemas for data description, from the most abstract to the most concrete point of view [8].

The conceptual data schema defines information represented by the data and the logical relationships organising this information. It can be used to communicate database content to users, independently of the way it is modelled or stored in a computer: it is platform-independent. In this paper, the feature types, relationships and their attributes defined in the conceptual data schema will be globally called "representation elements".

The logical data schema describes the data structure following the model of a particular database management system (DBMS). It lists tables and key attributes implementing the representation elements of the conceptual data schema.

The physical data schema describes the system of files adopted for data storage on the computer.

Figure 2 represents the three types of data schemas as the successive steps of an abstraction process translating real world features into computer representation.

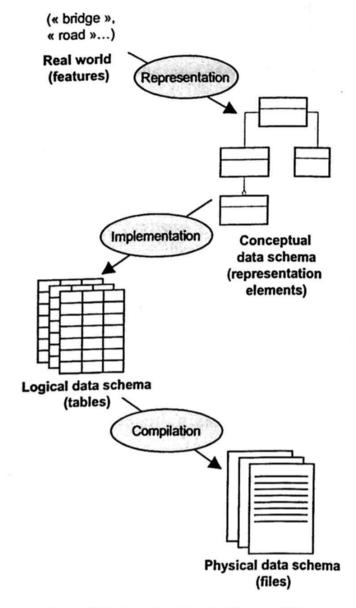


Figure 2. Real world and levels of data modelling

The conceptual data schema is the most relevant to assess the information content of data. However, it does not precise exactly what real world features are represented by representation elements, and how (e.g. what real world buildings can be included in the representation element called "building"). Indeed, real world features have different meanings for every user and representation elements have different definitions in every conceptual data schema. No universal ontology is able to link these two levels [9]. More precision about what information content is provided by a class of the conceptual data schema can be found in the specifications of the data set.

Data specifications.

The specifications provide conditions on real world feature characteristics to select those that should be represented by each representation element of the data set (e.g. buildings whose ground surface is more than 5m² and lower than 15m² are captured as "huts"). They also explain how the features must be represented (e.g. "huts" are represented as points indicating their centre). These textual conditions and rules are numerous and hardly readable by an end-user. Several research approaches are trying to make this documentation accessible to humans or machines. Gesbert [9] and Rüther [11] formalise specifications, which are usually complex text documents,

into conceptual data schemas. Their goal is to make data sets unification possible. Goder [12] has proposed a cartographic representation of data sets specifications, insisting on differences between products. As shown in Figure 3, tables (at the top) compare what concepts are represented and distinguished in the data sets. Representation rules are graphically displayed (at the bottom, for rivers of different width).

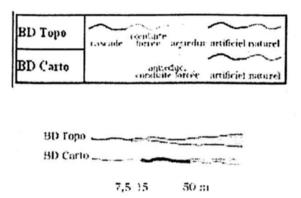


Figure 3. A user-friendly display of content and representation specifications for two IGN databases (BD Topo and BD Carto).

1.2.2.2. Conceptual data schema complexity

A second difficulty hinders data selection in addition to the difficulty involved in reading the technical documentation: it is the intrinsic complexity of the conceptual data schema. The conceptual data schema does not only represent the data set information content independently of other concerns: some representation and implementation choices are also implicated. The main choice is to model information in terms of homogeneous entities, even if it entails decomposing real world features. This is done in order to simplify data capture, storage, and management for the data producer.

For example, in some IGN data conceptual schemas, roads are composed of road sections, which are the longest sections without any intersection and with homogeneous attributes value. This choice is very useful: it renders the road network explicit, and it enables changes in attribute values along the road. However it does not necessarily make sense for users just checking if roads are represented in the data base.

To give a second example, in some IGN conceptual data schemas, distinct representation elements (such as "linear construction" and "surface construction") may represent the same real world feature (such as the concept of river embankment). It is a modelling choice made with data storage in mind, since point, line and area objects must be separated.

These modelling choices suit very well to data management, but they are not the most adapted to the data set content discovery and selection, which most of the time requires a lower level of detail. In the conceptual data schema, the representation elements are very numerous, and their names and hierarchical organisation do not directly reflect the basic real world features required by the user.

1.2.2.3. Conceptual data schema insufficiencies

A third problem makes conceptual data schema insufficient for data selection: not only does the conceptual data schema express the data set information content in a complex form, but also it does not express the whole information content. This is due to the importance of implicit information. As shown in the example of section 1.1 concerning the "bridge" feature, the conceptual data schema (and specifications) does not tell everything: it only describes features that are explicitly stored in tables. In fact, in some cases, simple needed concepts are provided somehow by the data set, but they are not explicitly specified in the conceptual data schema:

- In some cases, the concept has been aggregated or split to satisfy some implementation constraint,
- In other cases, the concept has not been taken into account during data set design, but it can be retrieved by simple operations on represented concepts.

As a conclusion, users who only have metadata description at their disposal, or who are unable to decode the complex technical description, cannot assess whether the information content they need is provided by the data set. That makes the selection of an appropriate data set arduous. The need for data pre-processing is described in the next section

1.3. ISSUES INVOLVED IN PRE-PROCESSING DATA SETS

Data are generally produced by organisations such as mapping agencies that have a limited number of immediate application purposes. So, data sets cannot be customised for every specific need, and they seldom exactly suit the user application requirements. Even if a data set provides the right information content, preprocessing may be a necessary step to adapt data to the application. There can be a problem of data structure (e.g. feature types must be split or aggregated), format, spatial representation (e.g. topologic relationships must be calculated so that roads form a network), spatial referencing (e.g. geographic coordinates must be converted to another system), etc.

Data transformation makes it possible to change data format or geometry, to modify a schema structure, to filter information, etc. But these transformations may be time and effort-consuming for non-expert users. Moreover, these home-made transformations are not referenced by the data provider, which does not guarantee any compliance with differential data sets when the data must be updated [13]. That is why a user catalogue query should be composed not only of an extraction query, but also of a pre-processing order. This customisation step should be supported and referenced by the data provider. Such a customisation of existing data sets for specific needs has not been widely studied. The UAPE system presented by De Oliveira [14] tries to adapt data and

applications to the user platform: it allows the user to design his own database and application, which then should be automatically implemented by a driver on the chosen GIS. Hubert [15] adapts data representation, through a cartographic generalisation process, to the user requirements of visual data display.

To summarise the issues involved here, crucial steps in user access to geographic data include the specification of a data set containing the required information (i.e. the choice of an information content based on metadata and data set technical documentation) and the restructuring of this data set to make it application compliant (i.e. the choice of a representation for the selected information content).

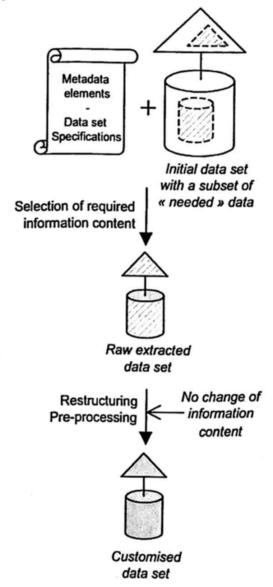


Figure 4. How to process customised data set specification

These steps are represented in Figure 4. They can hardly be achieved by end users because of the lack of available and explicit information describing data sets, and because of the complexity of the information structure. However, technical documentation of geographic databases and simple operations could help assess and improve data sets fitness for use. In our research, we intend to use these elements to build a system for customised data sets specification.

2. APPROACH

We aim at designing a system enabling users to order customised data sets, i.e. to assess existing data set content, to select within existing representation elements the information content the most adapted to their need, to restructure it, and to specify some additional preprocessing. The outputs of the system are orders to be placed to the data provider:

- An order for data extraction corresponding to the needed information content,
- An order specifying some pre-processing to be carried out on extracted data so that the data schema and format are consistent with the user application.

The two main elements of our approach are:

- A data set description model which links the user needed geographic features (e.g. roads) to implemented tables (e.g. road_section), possibly with data structure transformations (e.g. aggregation of road section objects).
- An interface to support interaction with the user. It is an essential part of the system. A lot of work must be completed to display the rich and complex information of the data set description model in a simple form. Since the specification of customised data sets is an interactive process that requires usersystem negotiation, a dialogue process controlled by a dialogue manager must be defined. Among other things, a dictionary will be required to match user key-words and system known concepts.

The steps of the specification process are presented in section 2.1. Section 2.2 focuses on the data set description model.

1.3. THE CUSTOMISED DATA SET SPECIFICATION PROCESSING

Figure 5 represents the successive steps that will be proposed by our system for interactive data set specification. These steps are then briefly described. Some of them rely on user-system interaction, others rely on the system only.

To illustrate sections 2.1 and 2.2, we take the example of a user who needs the road network, the rivers and their crossing.

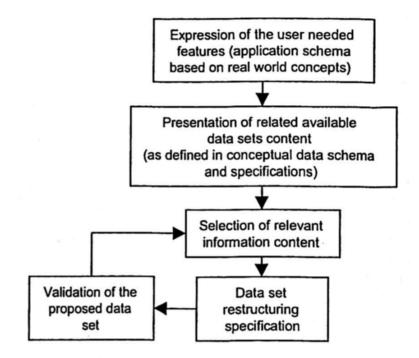


Figure 5. Main steps in the customised data set specification process

Expression of user needed features.

The first step is the expression of user needed features. The information required by the user for his application can be represented in an application schema. ISO provides guidelines for application schema definition [16]. Since we suppose that non-expert users need a simple way to express their need, we situate the user application schema at a less complex and detailed level: we will use semantic networks to propose features and define an application schema. An example of user application schema is shown in Figure 6.

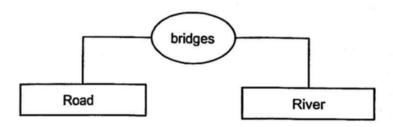


Figure 6. The user application schema

Presentation of data set content.

The second step is the presentation of available data sets information content. The system identifies existing representation elements of the conceptual data schema corresponding to the needed features, and presents them. Real data samples are also displayed, and all useful information available from metadata and data set specifications is provided so that the user can precisely assess the information content of the proposed representation elements. For the specific GeoBase data set, the proposed representation elements are those shown in Figure 7. If the system "knows" several data products likely to provide the adequate information content, it must present them distinctly and allow comparisons.

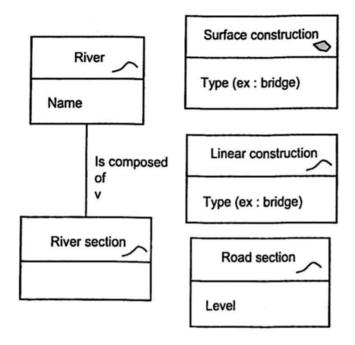


Figure 7. Representation elements of the GeoBase conceptual data schema, likely to correspond to the user application schema elements.

Selection of relevant information content.

During the third step, the user selects the most relevant representation elements. For example, after a look at the GeoBase content specifications and data samples, he selects the elements "river", "road section" and "linear construction" from the displayed representation elements (Figure 7).

This selection step should not be seen as a query definition: the goal is not to pick some objects from existing representation elements, but rather to define new representation elements. It can be seen as a view definition process, with this difference that the defined view is going to be restructured and extracted and will result in a new independent data set.

The representation elements must be chosen within a single data schema, even if several products have been presented during the previous step. Indeed, the system is not meant to integrate heterogeneous information, but to adapt and customise existing products.

Data set restructuring and pre-processing.

The fourth step is the restructuring and pre-processing of the selected data set. The system has to propose a customised data set as close as possible to the application schema expressed by the user, e.g. that explicitly represents the concepts that appear in the application schema. This can imply the need to restructure the data set to derive new components. For this step, the system can appeal to specific operations (i.e. class fusion, filtering, aggregation, etc.). This step relies on the data set description model and is further detailed in the next section.

Some simpler pre-processing operations such as changes of data format or coordinate system are necessary to fully adapt the data set to user requirements. They cannot be carried out on feature types, but directly on data. They can be ordered by the user and registered by the system, so as to be executed later, after the data extraction phase. On the other hand, treatments requiring complex tuning such as cartographic symbolisation or generalisation are not taken into account.

Validation: checking application compliance.

The fifth step is a validation step. It checks that the specified customised data set is accepted by the user and consistent with his planned applications. Indeed, as explained in section 1, each application requires specific data properties, which will be defined in a rule base. As every specific application requirement cannot be described, we limit this validation to common operations concerning measurement and selection, locating and addressing. Most complex applications are composed of such basic operations.

If the constraints induced by the user's planned activities are not fully respected by the specified custom-made data set, the system has to point this out. Previous steps of selection, restructuring and pre-processing have to be corrected or renewed.

1.1. THE DATA SET DESCRIPTION MODEL

In section 1, the gap between the real world features required by the user and the available descriptions of data set content was pointed out. Several modelling levels were distinguished. In this section, we present our draft data set description model that bridges the levels of needed features (real world), representation elements (conceptual data schema) and stored tables (logical data schema) together.

The problem analysis points out the need for an enriched data schema, as shown in Figure 8, with an intermediary conceptual level between the user application schema and the initial conceptual data schema.

Indeed, the initial conceptual data schema can hardly be connected to user needed features for several reasons that can be recalled here:

- It provides dense information. Depending on the level of detail, one real world feature may be related to several representation elements whose names are not always explicit. It is all the more dense in that it provides detailed information related to acquisition, representation and storage constraints.
- It does not describe the whole information content.
 The features that are implicitly represented in the data set are not taken into account.

To be easily exploited for user needs, this initial conceptual data schema should be simplified when too dense information is given, and enriched with concepts implicitly contained in data. This is the role of the enriched conceptual data schema.

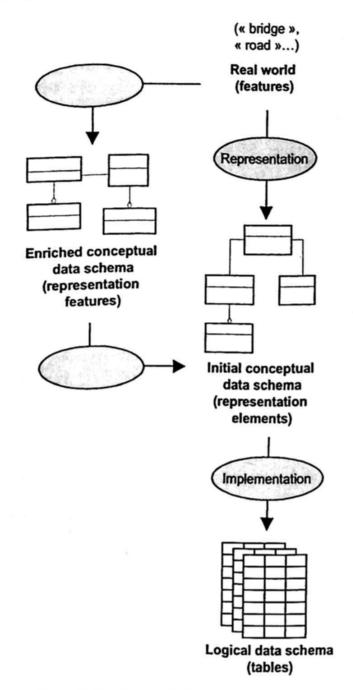


Figure 8. Use of an enriched conceptual data schema

In our data set description model, the relationship linking representation features of the enriched conceptual data schema to existing representation elements of the initial conceptual data schema is a "yield" relationship. This relationship must be supported by simple schema transformation operations, as shown in Figure 9.

Real world features mentioned in the user application schema are related to representation features of the enriched conceptual data schema through relationships of "possible representation".

To distinguish abstraction levels, we borrow the concept of representation stamp [17] from the MADS formalism for multiple representation data:

- User needed features carry the "usr" stamp,
- Representation elements of the initial conceptual data schema have the stamp "ini"
- Representation features derived from the initial conceptual schema by transformation operations carry the stamp "implicit" or "imp".

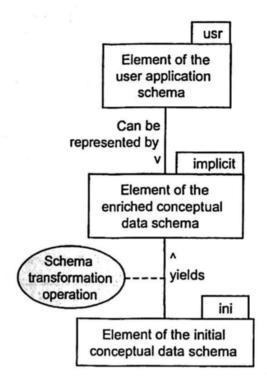


Figure 9. The "implicit representation" and "possible representation" relations

Some of the transformation operations can be automatically proposed by the system, such as the aggregation of the existing "Road section" element to create the desired "Road" feature. Others must be requested by the user.

In our example, the "Road" feature and the "bridging" relationship have to be built through operations shown in figure 9.

The features of the proposed customised data set maintain some properties of the initial conceptual data schema elements, which enable them to support user applications or not. The proposed data set in Figure 9 can be used to select roads and rivers concerned by a bridging relation. But location of bridging points or measure of the bridge length are not possible, whereas the initial data set contains the required information. If such operations are required by the user application, another solution must be found that derives a "Bridge" feature with a geometry.

3. CONCLUSION

This paper intends to propose an approach to deal with the problem of user access to geographic products suiting their need. We aim at building a system designed to specify a customised data set. The core elements of our data set description model have been presented. This model mainly relies on an enrichment of available data sets: it must describe precisely the properties of available representations elements. It finally must include the simple data manipulation operations that can be used to derive a data set explicitly representing the user needed features, in a way adapted to its intended use.

This approach makes the data provider in charge of the process of data restructuring and representation adaptation. By the way, the data provider keeps a trace of the data set modification and is able to assure future data set maintenance.

In this project, the source data sets will be limited to existing IGN data sets, and the described content will

probably be limited to a single theme like water bodies. Even if it dramatically reduces the area of possible custom-made data sets, it will enable us to explore a new type of access to geographic data. This is likely to improve usability of existing products without modifying their structure or their content.

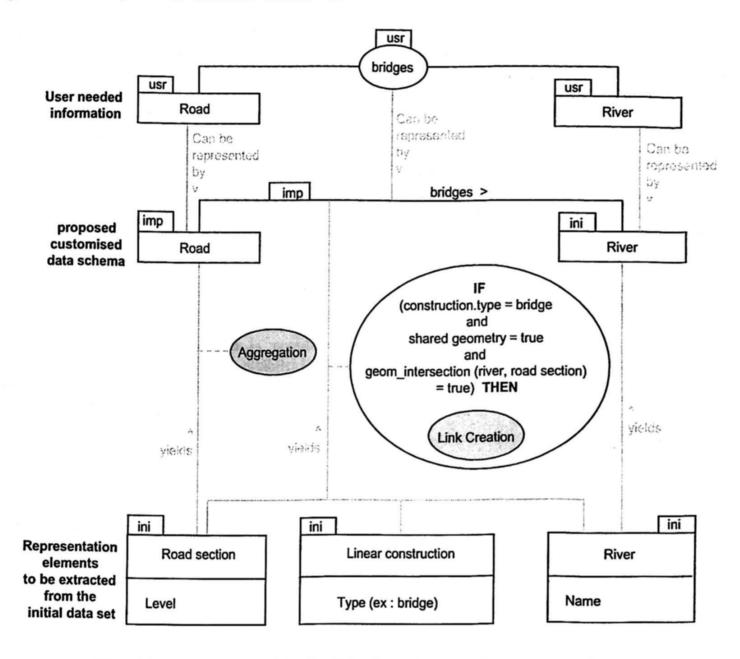


Figure 9. Data set restructuring. Links of implicit and possible representation are shown by pale arrows.

REFERENCES

- [1] A.U. Frank & W. Kuhn, Specification Languages for Open GIS, Revue Internationale de Géomatique, 9(2), 1999, 135-152.
- [2] B. Bucher, A Model to Store and Reuse Geographic Application Patterns, *Proc 4th Agile Conf on Geographic Information Science*, Brno, Czech Republic, 2001, 289-295
- [3] A.U. Frank & M. Jahn, How to Sell the Same Data to Different Users at Different Prices, *Proc 6th Agile Conf on Geographic Information Science*, Lyon, France, 2003, 357-360.
- [4] B.Vasseur, R. Devillers & R. Jeansoulin, Ontological approach of the fitness for use of geospatial datasets, *Proc 6th Agile Conf on Geographic Information Science*, Lyon, France, 2003, 497-504.
- [5] GSDI Technical Working Group, Developing Spatial Data Infrastructures: the SDI CookBook, v1.1, Douglas Nebert (Ed), 2001, on line: http://www.gsdi.org/pubs/cookbook/cookbook0515.pdf
- [6] P. Ahonen-Rainio, Description of the Content of geographic datasets, *Proc 8th SCANGIS Conf on Geographic Information Science*, As, Norway, 2001.
- [7] ISO TC211, ISO 19115 Geographic Information metadata, international standard, 2003

- [8] G. Gardarin, Bases de données. (Ed Eyrolles, 2000).
- [9] F.T. Fonsesca & M.J. Egenhofer, Ontology Driven Geographic Information Systems, Proc 7th ACM Symposium on Advances in Geographic Information Systems, Kansas City, USA, 1999, 14-19.
- [10] S. Mustière, N. Gesbert & D. Sheeren, Formal Model for the Specifications of Geographic Databases, Proc GEOPRO International Workshop on Semantic Processing of Spatial Data, Mexico City, Mexico, 2003.
- [11] C. Rüther., W. Kuhn.& Y. Bishr., An Algebraic Description of a Common Ontology for ATKIS and GDF, Proc 3rd AGILE Conference on geographic information science, Helsinki, Finland, 2000.
- [12] G. Goder, Représentation comparée de schémas et spécifications de contenu, Master thesis, 2003.
- [13] T. Badard, On the Automatic Retrieval of Updates in Geographic Databases Based on Geographic Data Matching Tools, Proc 19th International Cartographic Conference ICA/ACI, Ottawa, Canada, 1999, 47-56.

- [14] J.L. De Oliveira, F. Pires & C.B Medeiros, An Environment for Modeling and Design of Geographic Applications. *GeoInformatica*, 1, 1997, 29-58.
- [15] F. Hubert, B. Bucher, S. Balley & A. Ruas, A User Interface to Specify the Parameterisation of GIS Function: The Example of Building Generalisation Parameterisation, Proc EuroSDR Workshop on Visualisation and Rendering, Enshede, Netherlands, 2003.
- [16] ISO TC211, ISO/DIS 19109 Geographic Information
 Rules for application schema, draft international standard, 2003
- [17] C. Vangenot, C. Parent & S. Spaccapietra, Modeling and Manipulating Multiple Representations of Spatial Data, Proc SDH Conf on Advances in Spatial Data Handling, Ottawa, Canada, 2002, 81-93.

GENERATING SPATIAL ONTOLOGIES BASED ON THE SPATIAL SEMANTICS

Miguel TORRES and Serguei LEVACHKINE
Geoprocessing Laboratory (GEOLAB)
Centre for Computing Research (CIC)
National Polytechnic Institute (IPN)
Av. Juan de Dios Bátiz s/n Unidad Profesional "Adolfo López Mateos"
C.P. 07738, México, D.F., MEXICO
mtorres@cic.ipn.mx, palych@cic.ipn.mx

ABSTRACT

Spatial ontologies (SO) are used to improve the geographical objects representation in spatial databases. We generate SO by means of the interaction between a spatial subject domain (SSD) and spatial taxonomy (ST). SSD is composed by a priori spatial knowledge, which is related to the "essential properties" (spatial data semantics) of the geographical objects. On the other hand, ST describes the classification among the spatial data, according to the primitives of spatial objects representation. Using such method, it is possible to improve the spatial data representation establishing the geo-information conceptualization that results useful for subsequent spatial data processing and interpretation. Thus, this method of spatial data semantic processing can be considered as a knowledge discovery in the spatial databases (SDB).

KEY WORDS

Spatial Semantics, Spatial Ontology, Spatial Taxonomy, Spatial Subject Domain, Geographic Information System

1. INTRODUCTION

Nowadays, the spatial databases are very useful and powerful tools to handle, display, and process the geographical information. These databases integrate Geographic Information Systems (GIS), which are composed to store and process spatial data. To solve some ambiguities in the spatial data processing and interpretation, the geo-information should have good quality from the input to the representation. The "adequate" representation of spatial data is crucial for improving the decision making in different environments [1].

In this paper, we generate spatial ontologies based on the spatial semantics, which can be used to represent geographical objects by means of spatial concepts ("not words"). Such spatial data conceptualization aim to compress the data and facilitate the knowledge discovery into SDB.

Up-to-date GIS do not extensively explore the spatial data semantics. To develop a spatial semantic theory is a great challenge in the new trends of Geocomputation field. Thus, the spatial analysis can use alternative methods to represent spatial data: this data representation jointly with the semantic rules - both based on data semantics - can be stored in a knowledge-base to generate new concepts that form the spatial ontologies. These concepts are defined by the properties and behavior of geographical objects and explored the by human experience. In general, we seek to correctly represent spatial objects for their subsequent processing [2].

Several works related to semantic interoperability have been published. In particular [3] presents an approach to semantic similarity assessment combining two different strategies: feature-matching process and semantic distance computation.

An approach proposed in [4] has been achieved to improve the spatial data handling. As described in this work, LOBSTER system combines the artificial intelligence techniques to provide a query language more flexible and powerful than standard SQL. This system has been developed in Prolog for object-oriented modeling, geomorphology, and query optimization.

In [5] and [6] an ontology-driven GIS as a system integrator has been proposed. In these works, a special model to conceptualize the geographical information and solve problems related to the integration and interoperability in GIS of different types at different levels of detail has been described.

In [7] the Naive Geography is introduced as a body of knowledge that captures the way people reason about geographic space and time. Probably, future generations of GIS will incorporate formal models of naive geography.

Other works relate to the semantic approach to the spatial data processing based on the concept of geographic entities. [8] enables the seamless integration of several types of information through the use of flexible spatial object classes. These classes are composed by the

combination of other classes that represent the richness of the geographic world.

Nowadays, the interoperability in GIS is approached by using the spatial semantics representation. This interoperability is based on the integration of spatial schemas, query languages and sets of semantic rules, which can provide knowledge of data and geographical representation interfaces [9].

A peculiarity of our proposal is the definition of spatial semantics, which is based on intrinsic properties of the geographical objects. These properties integrate a knowledge-base, which can be represented by concepts that form the spatial ontologies. In other words to generate spatial ontologies, we define a data description that is composed by semantic properties of the spatial data to compose the spatial subject domain. Moreover, the spatial data are classified by spatial taxonomies. The latter are interacted with the subject domain to generate ontologies. In particular an improvement of representation of the spatial databases can be achieved by using our approach.

This proposal consists of generating ontology levels, which can be defined by top or down ontologies according to the particularization of the concepts. For this purpose, we consider 2D partition of a map that is composed by a set of primitives of representation. Such maps are stored in spatial databases. However, the approach is focused on describing the semantics of a particular map in two dimensions, which is formed by geographical objects. The approach attempts to find out the spatial semantics throughout spatial ontologies. Using this approach, we seek to enrich the concepts by means of the "interaction" and to improve the spatial data representation.

The rest of the paper is organized as follows. In Section 2 we describe a spatial semantics definition, which is proposed to use in spatial databases. The spatial subject domain and the spatial taxonomy are described in Section 3. Section 4 presents the spatial ontology definition and the interaction approach to generate it. Section 5 contains some results for topographic case of study. Our conclusions are outlined in Section 6.

2. SPATIAL SEMANTICS DEFINITION

To define spatial semantics, we use the essential characteristics that involve the spatial data. Our definition is based on providing a set of rules. This set is composed by relationships, properties, functions and behaviors, which define the characteristics of the geographical objects [2].

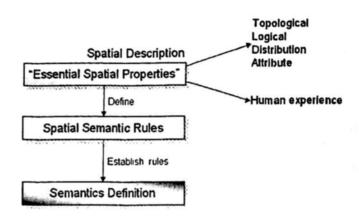


Figure 1. Spatial semantics definition

Figure 1 shows that the spatial semantics definition starts with a spatial description of the geographical objects. This description consists of topological, logical, attribute and spatial distribution properties, which define conceptually spatial semantic rules. To generate the description, it is important to take into account the human experience for establishing the constraints that compose the set of rules. Based on these considerations, we proposed the *spatial semantics definition* [2]. In which 2D spatial objects system is defined by a set of primitives of representation, which form a partition of a map.

A map is composed by tree basic primitives of representation, which are the following:

Point objects. The only geometric information given for a point object is the location, therefore it is represented by a node.

Line objects. The geometric information given for a line object is the location, shape and the length. This can be done by representing it as a polyline. The nodes contain the position information and shape information can be derived from the angles between the edges and their length. The edges can be linked to the line object through the polyline, we can express this link by saying that "the edge is part of the line object". Because of their definition, segments can take the role of edges in the description of the line object.

Polygons objects. The geometry of area objects is given by their boundaries. We assume that the definition of the area objects is such that they cover the total area of interest completely and that they do not overlap. In that case, they form a geometric partition of the area. This assumption implies that any part of an object boundary is always the boundary between two objects. If the geometry of these boundaries is described by polygons consisting of edges, then any edge will have an area object at its right hand side and an area object at its left hand side. The possibility of the distinction between "left" and "right" only exists if the edges are directed. Here two segments can take the role of edges. We can obtain the length of the perimeter and the area is the size measure.

Let P be the spatial partition as above and U_i the discourse domain (universe of geographical objects), which consists of a set of primitives of representation (lines, points or polygons). P is the set of partitions of the primitives of representation that can exist in that partition P and present in the same partition, such as shown in (1):

$$P = U_i \cap \left\{ Rp_i \vee Rp_p \vee Rp_a \right\}, i=1,...,n, \tag{1}$$

where:

 Rp_l is the primitive of representation "lines". Rp_p is the primitive of representation "points". Rp_a is the primitive of representation "polygons". i represents the thematic number that involves the spatial partition.

In all cases, these sets are associated with the geographical objects, which are denoted by f_G .

We also denote:

 O_T represents a set of operations related to the topological properties such as: meet, contain, cover, overlap, etc.

 O_L represents a set of operations based on the logical constraints, for example: the behavior of the land is considered to build a road in a specific zone.

 ${\cal O}_{\cal D}$ represents a set of functions that correspond to spatial distribution properties.

 $O_{\mathcal{A}}$ represents a set of the characteristics associated to the geographical objects such as: elevation, names, area, direction, etc.

The properties considered for this definition can be described as follows:

A. Topological properties (P_T) . These compose the spatial object layers. They represent the topologic and geometric shapes and the spatial relationships, keeping the consistency as well as the congruency of the geographical structures. Therefore P_T properties can be defined:

$$P_T = \{ \forall f_G \mid \exists p_t \in T, p_t \subset O_T \}, \tag{2}$$

where:

T represents the set of topological properties and, p_t indicates a particular topological property that can be considered.

B. Logical properties (P_L) . They provide rules that define the behavior, relationships and combined properties of the spatial structures. These characteristics are focused on the description of the phenomena, and can be used to simulate and model the geographical phenomena behavior, based on the human experience. Therefore P_L properties can be defined as:

$$P_{L} = \{ \forall f_{G} \mid \exists p_{I} \in L, p_{I} \subset O_{L} \}, \tag{3}$$

where:

L represents the set of logical properties and, p_l indicates a particular logical property that can be considered.

C. Spatial distribution properties (P_D) . They are used to locate any geographical object in the space. These properties consider the spatial reference of the geographical systems. Therefore P_D properties can be defined as:

$$P_D = \{ \forall f_G \mid \exists p_d \in D, p_d \subset O_D \}, \tag{4}$$

where:

D represents the set of spatial distribution properties and, p_d indicates a particular spatial distribution property that can be considered.

D. Attribute properties (P_A) . They describe the characteristics about a phenomenon. It is necessary to consider qualitative and quantitative attributes, because these specific attributes support the knowledge of the similarities between the geographical objects. Therefore P_A properties can be defined as:

$$P_{A} = \{ \forall f_{G} \mid \exists p_{a} \in A, p_{a} \subset O_{A} \}, \tag{5}$$

where:

A represents the set of attribute properties and, p_a indicates a particular attribute property that can be considered.

Therefore, the initial exploration to define a priori and intuitive spatial semantics (S_E) is the description of the content of all these properties and relations that are involved them, in any spatial partition P.

$$S_{E} = content \left(\bigcup_{i=1}^{n} U_{i} \cap \left\{ P_{T} \wedge P_{L} \wedge P_{D} \wedge P_{A} \right\} \right)^{1}$$
 (6)

All the properties are considered in the object systems of a finite space of geographical objects. Therefore, the content of the set of spatial rules reflect the topological, logical, spatial distribution and attribute properties of the spatial data. In general, in this definition, we are looking for improving the spatial data representation for their subsequent processing [2].

However, considering the *content* of all these properties and relations, the spatial semantics is defined by the *interaction* (between the spatial taxonomy and the spatial

¹ The concept of "content" will be particularized in the following sections in dependence on application.

subject domain) iterative process with the possibility to feedback this process to generate the spatial ontologies again in the following interaction. Each new defined concept is put into spatial subject domain. Then the interaction process is repeated up to no new concepts found. The resulting spatial ontology represents the semantics of the spatial object system.

In this sense, the spatial semantics relates the content and the data representation of entities or frameworks of the real world. The main purpose of spatial semantics is to identify the characteristics of geographical objects, which belong to different spatial databases and adapt them to the resolution levels of the spatial information [10].

The idea to use spatial semantics for representing geographical objects is focused on generating a knowledge-base. This depository is based on the human experience and the "essential" properties of spatial data. The knowledge-base can be transformed into a Spatial Expert System (SES), which stores a set of semantically defined rules [11]. Based on defined rules, a spatial taxonomy is generated. This spatial taxonomy is defined according to the primitives of representation of the spatial data. Spatial subject domain embedded into the knowledge-base interacts with the spatial taxonomy; that is Spatial Semantic Analysis (SSA). Figure 2 shows the scheme of the Spatial Semantic Analysis.

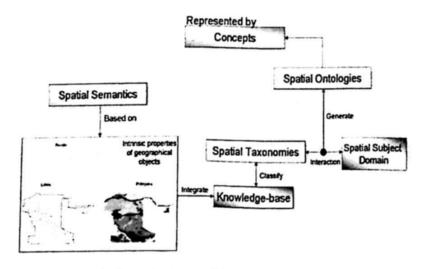


Figure 2. General scheme of the spatial semantic analysis

3. SPATIAL SUBJECT DOMAIN AND SPATIAL TAXONOMY

In our approach, the interaction between the subject domain and the taxonomy generates the "ontologies", according to the spatial semantic rules. These rules are composed by the "essential characteristics" of the geographical objects.

We define the rules to classify and represent the spatial entities. In this context, our approach leads to the "correct" representation of the spatial data. We have to look for the most adequate representation in each case of study.

3.1 SPATIAL SUBJECT DOMAIN

Spatial Subject Domain is defined as a set of "names" that describe the primitives of spatial representation. Thus, we can start with a priori knowledge of the geographical objects that appear, e.g. in the map legend. For instance, "blue" lines are united under the concept (name) "river" and "black" lines are united under the concept "fracture", etc. In reverse, the different concepts are united under the same description of the spatial representation that is "line". The interaction between the subject domain and the taxonomy is used to locate concepts into the spatial subject domain that correspond to a case of study, and to process these concepts to generate spatial ontologies [2]. Figure 3 shows the definition of spatial subject domain and its interactions.

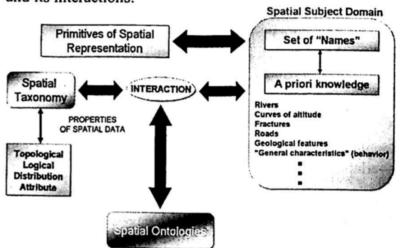


Figure 3. Interactions of Spatial Subject Domain

In our approach, it is important to count with a priori knowledge that is defined by the user depending on the case of study. However, the prompt is that this knowledge should be composed by the relationships, properties, functions and behaviors, which describe the characteristics of the geographical objects and these should be put into the spatial subject domain.

For this purpose, it is convenient to consider a particular case of study to define the knowledge, because the spatial subject domain interacts with the spatial taxonomy according to the spatial thematic.

To define the spatial subject domain, it is necessary to elaborate a description of the thematic to analyze, considering the main elements that compose this theme, such as the data model and the resolution levels of the spatial information. For instance, suppose that we have to generate "concepts" related to topographic data.

All the characteristics, which are considered in spatial description of the spatial subject domain, should represent relationships between themselves too. Spatial subject domain should recognize the different semantic levels of a priori knowledge that is stored in this domain. A framework of this description for the topographic data is presented in Figure 4.

```
<Spatial Data Description>
  <Layer Name: Topo>
  <Thematic:Topography>
  <Type:Line>
  <Geographical Properties>
    <Projection:UTM 14>
    <Datum:NAR D>
    <Units:METERS>
    <Spheroid:GRS1980>
    <Scale:1:50000>
 <Geographical Features>
    <Aqueduct relationship with the land >
      <Values domain:
           Underground
           Superficial>
    <Condition of Aqueducts >
      <Values domain:
           Construction
           Operation
           Not use>
   <Type of airports>
      <Values domain:
           International
          National
           Local>
   <Type of streets>
      <Values domain:
          First order
          Second order
           Third order
           Fourth order>
   <Population>
      <Values domain:
          State:
                  Low inhabitants
                  Medium inhabitants
                  Large inhabitants
          City:
                  Low inhabitants
                  Medium inhabitants
                  Large inhabitants
          County:
                  Low inhabitants
                  Medium inhabitants
                  Large inhabitants>
  <Type of roads>
     <Values domain:
          Pavement:
                  One rail
                  Two rails
                  Three rails
                  Four rails
                 More than four rails
          Unpavement
          Breach
          Path>
  <Hydrology>
     <Values domain:
         Intermittent
         Always presented>
  <Contour>
     <Values domain:
         Depression
         Altitude>
```

Figure 4. Partial description of spatial subject domain for topographic thematic

Using this description, it is possible to obtain "concepts" related to the topographic thematic, besides that, it is required that the spatial subject domain interacts with the spatial taxonomy to generate the spatial ontologies.

The description used to obtain the concepts is not general. Every case of study is different, because it presents characteristics that cannot be defined in a "general way". Thus, a priori knowledge is defined by every particular environment. In other words, it is not possible to develop a "general" spatial subject domain, not only for the complex structures that provides the shapes of the primitives of representation but also by the behavior, which is defined by the properties of every geographical object (line, point or polygon).

3.2 SPATIAL TAXONOMY

We define a spatial taxonomy as a classification method to describe every spatial entity (primitive of representation) [2]. The use of spatial taxonomies is served to classify and organize a priori knowledge in a hierarchical way. It can be communicated throughout spatial subject domain to make an interaction for generating the spatial ontologies. In this classification, the spatial entities can be generalized according to some spatial semantic rules.

We consider a subset of spatial data, which is generated by the spatial taxonomy for every spatial thematic layer that is stored into the spatial database.

Using this schema, we can define the spatial semantics of the geo-information. This method allows processing the geographical information according to the semantics of the geographical objects.

Figure 5 shows a general scheme of the spatial data conceptualization and representation as an interaction between the taxonomy and the subject domain.

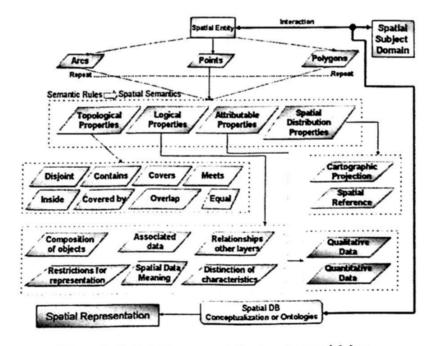


Figure 5. Spatial data representation based on spatial data conceptualization

In general, our approach uses the object properties and generates a set of spatial semantic rules. All generated rules are stored into a knowledge-base, which relates and classifies the geographical information according to the characteristics that correspond to each geographical object. The spatial data are classified in spatial digital layers, which are defined by basic primitives of representation of the spatial data such as arcs, points and polygons. Thus, a spatial taxonomy is generated. The taxonomic classification is based on the properties of the spatial data (Figure 5; the union of dotted rectangles).

To generate the spatial taxonomy, we establish spatial semantic rules. They are defined by the properties of the spatial data (Section 2). Therefore, to maintain these rules, we use a set of artificial intelligence techniques, which support the retrieval and acquired knowledge processes. The acquired knowledge represents other characteristics of the spatial data, which are used to partially automate the spatial representation. This representation sometimes calls for the human intervention to solve problems related to the ambiguity between the different contexts that the spatial data can take in a case of study.

4. SPATIAL ONTOLOGY GENERATION

Most widely accepted common conceptualization of the geographical data is based on the description of geographical objects and fields [8]. These objects are not necessarily related to a specific geographic phenomenon, because human-built features are typically modeled as objects. The spatial semantics definition (Section 2) is proposed and aimed to correctly represent spatial data in an alternative and universal way to generate spatial ontologies.

For this purpose, we consider a spatial ontology as an explicit and structured specification of conceptualization, that is, a description of the concepts and relationships that can exist between the geographical objects. In this context, the spatial ontologies represent definitions, functions, attributes, relationships, etc., of the geographical objects by means of "concepts" [2]. Besides that, ontologies can be considered as "languages", which use a specific vocabulary to describe entities, classes, properties and functions related to a certain view of the geographical world [10].

These concepts are considered as a knowledge, which can be put into the knowledge-based system. This knowledge-base can be used to share concepts, according to the spatial semantic rules that correspond to the spatial data by using an agent interaction language. [12].

In particular, our approach is designed to solve the ambiguities that can exist with single characteristics of the geographical objects, because the spatial ontology is defined by concepts (not by words) according to the geographical objects.

It is important to make a distinction between ontology and conceptualization [8]. According to Guarino, an ontology is a logical theory accounting for the intended meaning of a formal vocabulary (i.e., its ontological commitment to a particular conceptualization of the world), whereas a conceptualization is the formal structure of reality as perceived and organized by an agent, independently of the vocabulary used or the actual occurrence of a specific situation.

To mimic the ontologies in spatial data modeling, we consider three stages to model the spatial data representation. Real world stage, which is composed by the geographical objects and phenomena that will be modeled in the computer. Logical stage consists of formal definition of these objects considering their properties and behavior. Representation stage is a description of the geographical objects and phenomena based on the definition proposed in the logical stage. These stages are shown in Figure 6.

The schema shown in Figure 6 can be used to catch what human-mind perceives about the real world. The sense of the model has several concepts as rivers, contours, roads, soils, etc. The logical stage is composed by the formal description of geographical objects. Later, the representation stage represents the ontologies that can be generated from the formal description.

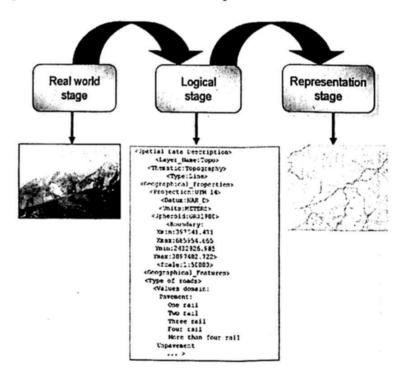


Figure 6. Stages to model the spatial data representation to generate ontologies

Therefore, we assume that spatial ontologies should be essential components of the logical stage for geographic data modeling [8].

Moreover, the geographical environment uses different levels of ontologies to guide processes for the extraction-representation (E-R) of more detailed geo-information and to allow the E-R of geographical objects in different stages of classification.

The spatial ontologies can be classified by levels according to their dependence on a specific task or point of view. These levels are generated for a specific spatial ontology (top-ontology) and it can be particularized to define a particular ontology (down-ontology). There are also different levels of information detail. Low-level ontologies correspond to very detailed information and high-level ontologies correspond to more general information.

In this situation, the generation of more detailed ontologies should be based on the high-level ontologies, such that each new ontology level incorporates the knowledge present in the higher level. These new ontologies are more detailed, because they refine general descriptions of the level from which they have been generated [10].

The levels of ontologies can be used to guide processes for the extraction of more general detailed information. The use of multiple ontologies allows the extraction of information in different stages of classification.

The use of explicit spatial ontologies contribute to better correct spatial representation, because every geographical object description is based on an implicit ontology, making it explicit avoids conflicts between the ontological concepts and the implementation.

On the other hand, spatial ontologies play an essential role in the conceptualization of spatial databases, allowing the establishments of correspondences and interrelations among the different domains of geographical objects and relations.

For instance, the concept "Limit" can be represented in different contexts in diverse spatial databases. "Limit" in some cases represents "coast boundary", separation between the "ground" and the "sea", "contour of value zero", "boundary" among two regions (states, countries, etc.), and so on (Figure 7).

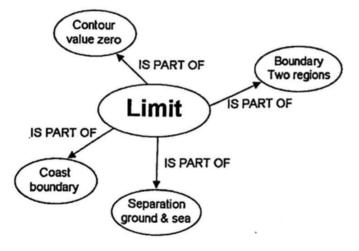


Figure 7. Semantic network related to the concept "Limit" interpreting the different concepts related to "limit" for subsequent spatial representation

For the concept "Limit", we can use a semantic network to interpret and organize the different context that it can take in the spatial ontology as presented in Figure 7.

Using this approach, we can generate specific spatial ontologies after defining the top-ontology to particularize the conceptualization in other specific ontologies (down-ontologies).

We propose a schema to generate the spatial ontologies. This schema consists of a set of spatial semantic rules that are stored in a knowledge-base. The rules are defined by the characteristics and behavior of the geographical objects. Spatial subject domain is used to work with a priori knowledge and interacted it with the spatial taxonomy to find out the concepts related to the spatial thematic aimed to generate the spatial ontologies. Figure 8 shows the schema to generate the spatial ontologies.

The use of ontologies in spatial databases enables knowledge sharing and information integration. The proposed approach provides dynamic and flexible information exchange and allows partial integration of spatial data when completeness is impossible.

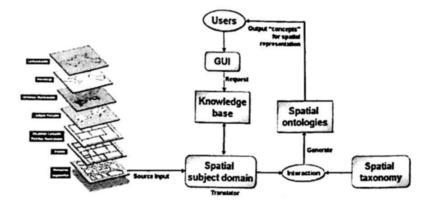


Figure 8. General schema to generate spatial ontologies

This can help to the next generation of spatial databases to solve semantic ambiguities in the available geo-information, because the context of the spatial data can change, according to the case of study or for the representation state by means of concepts of the geo-information.

5. PRELIMINARY RESULTS

The following results illustrate the interaction between spatial subject domain and spatial taxonomy, to obtain spatial ontologies by using the spatial semantics definition. Some tests have been made to prove our approach.

We design a spatial taxonomy related to topographic thematic, it is convenient to consider a particular case of study. Figure 9 shows the taxonomy of topographic data.

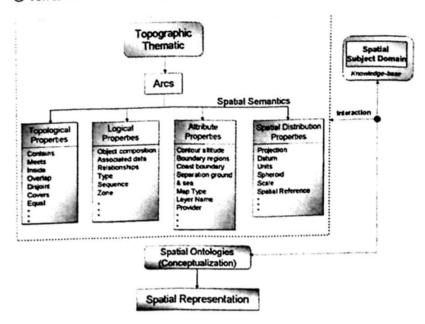


Figure 9. Spatial taxonomy for topographic thematic and its interaction

To generate the spatial ontology, it is required that the spatial subject domain interacts with the spatial taxonomy. As a result of this interaction, we can obtain

"concepts" related to the topographic data in different semantic contexts. For instance, in Figure 10 the concept "Limit" of spatial ontology is generated. This spatial ontology presents several "sub-concepts" related to "Limit". The users make a query to the geographical data according to the criteria or interested concept that they need.

The query functionality is the following: the user makes a request by means of a GIS-application, which sends it to the administration query module to process the request for obtaining the spatial ontology. Inside of GIS-application, a priori knowledge that is stored in the spatial subject domain interacts with the spatial taxonomy, considering in this case, the "arcs" as primitive of representation. Figure 10 shows the mechanism to obtain the spatial ontology by means of GIS-application. The architecture of the GIS-application is described in [2, 13].

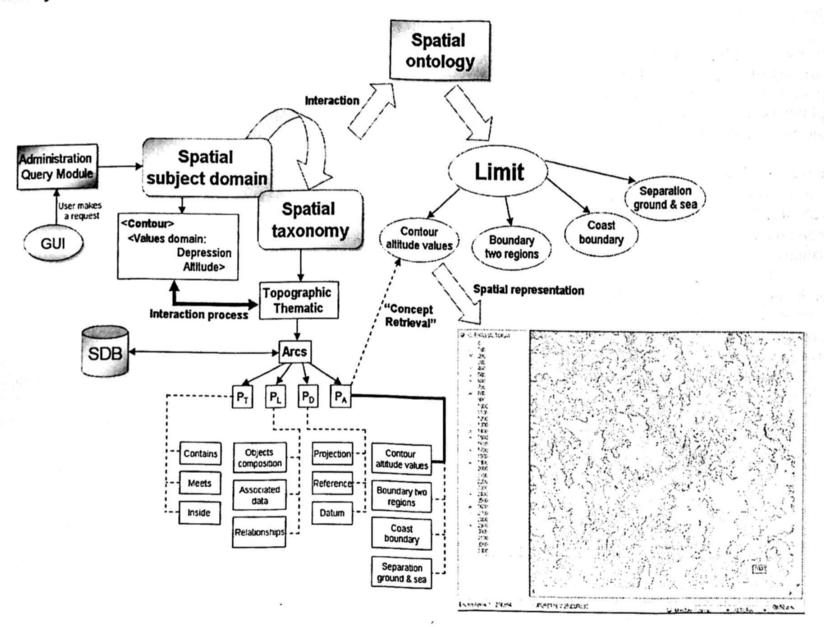


Figure 10. The concept "Limit" in spatial ontology that represents the contours

Figure 11 shows the query mechanism to describe the concept "Roads" into the spatial ontology. In this case, the ontology is composed by several "sub-concepts", which are ordered in a hierarchical way. In Figure 11, we see the different levels of the concepts, starting with a top-level

(Roads) and finishing with down-levels (One rail, two rails, etc.). When the ontologies present more levels of concepts, it is possible to particularize these in subconcepts, while the level is less (down-level), the concept is more particular.

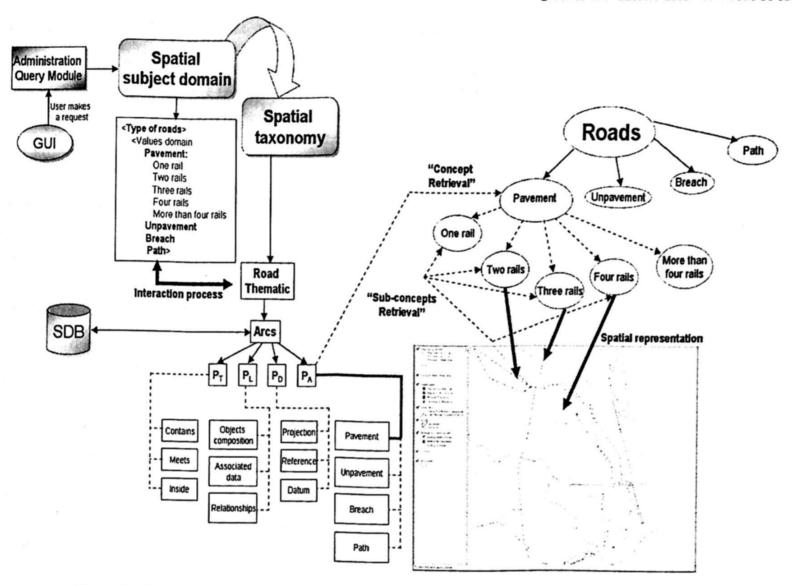


Figure 11. The concept "Roads" in spatial ontology that represents pavement roads with one, two, three and four rails

6. CONCLUSION

In the present work, the spatial semantics definition to represent geographical objects has been sketched. This definition is developed considering the intrinsic properties of spatial data, which are defined by the basic spatial representation primitives. The properties involved in this definition are: topological, logical, spatial distribution, and attribute.

By using our method, it is possible to classify geographical objects, generating a spatial taxonomy. The proposed definition of spatial subject domain is oriented towards an interaction with spatial taxonomy to conceptualize the spatial databases. Also, we presented a spatial data description that is based on a set of spatial semantic rules, obtained from the spatial object analysis. Using this set of rules, the request of the user evaluates and filters the geo-information according to the semantic criteria.

In essence, the spatial subject domain is defined as a set of "names" that describe the primitives of spatial representation. Thus, we can start with a priori knowledge of the geographical objects to examine the spatial data, which interact with the spatial taxonomy to generate spatial ontologies.

We attempt to show an alternative approach to represent spatial data considering their spatial semantic properties, besides that, it can act as a system to integrate geoinformation at different levels of detail.

The proposed spatial ontologies represent real world entities using a hierarchical structure, which is composed by "concept" (not words). These concepts can be formed by definitions, functions, spatial properties, rules, relationships that constitute the *spatial semantic analysis*.

Moreover, the spatial ontologies catch the semantics of geo-data to provide additional information related to the concepts. These ontologies can be used to establish agreements about diverse views of the world and consequently carry out the "meaning" of the geo-information. In many situations this geo-information embedded in the spatial representation of geographical phenomena in the human-mind.

Use of ontologies in spatial databases enables knowledge sharing and information integration. The proposed approach provides dynamic and flexible information exchange and allows partial integration of spatial data when completeness is impossible. This approach can aid to solve semantic ambiguities between the available geo-information, because the context of the spatial data can change, according to the case of study or for the representation state by means of concepts of the spatial data.

ACKNOWLEDGMENTS

The author of this paper wishes to thank the Centre for Computing Research (CIC), General Coordination of Postgraduate Study and Research (CGPI), National Polytechnic Institute (IPN) and the Mexican National Council for Science and Technology (CONACYT) for their support.

REFERENCES

- [1] M. Goodchild, Geographical Data Modeling, Computers and Geoscience, Vol. 18, 1992, 401-408.
- [2] M. Torres and S. Levachkine, Semantics Definition to Represent Spatial Data, In: Levachkine S., Bodanksky E. and Ruas A., (cds.), e-Proceedings of International Workshop on Semantic Processing of Spatial Data (GEOPRO 2003), Mexico City, Mexico, 2002.
- [3] M.A. Rodríguez, M. Egenhofer & R.D. Rugg, Assessing Semantic Similarities Among Geospatial Feature Class Definitions, *Lecture Notes in Computer Science Vol.* 1580, 1999, 189-202.
- [4] M. Egenhofer & A.U. Frank, LOBSTER: Combining AI and Database Techniques for GIS, *International Journal of Photogrammetric Engineering and Remote Sensing*, 56(6), 1997, 919-926.

- [5] F. Fonscea & M. Egenhofer, Ontology-Driven Geographic Information Systems, 7th ACM Symposium on Advances in Geographic Information Systems, Kansas City, MO, 1999, 14-19.
- [6] F. Fonseca, M. Egenhofer & P. Agouris, Using Ontologies for Integrated Geographic Information Systems, *Transactions in GIS*, 6(3), 2002, 25-40.
- [7] M. Egenhofer & A. Frank, Naive Geography, in Frank A. and Kuhn W., (Eds.) Spatial Information Theory, A Theoretical Basis for GIS, *International Conference COSIT '95*, Vol. 988, Lecture Notes in Computer Science, Springer-Verlag, Berlin, 1995, 1-15.
- [8] N. Guarino, Formal Ontology and Information Systems, in Guarino (Ed.), Formal Ontology in Information Systems, (Ed.: IOS Press, 1998), 3-15.
- [9] M. Gooldchild, M. Egenhofer, R. Fegeas & C. Kottman, *Interoperating Geographic Information Systems* (Ed.: Kluwer Academic Publishers, 1999).
- [10] F. Fonseca, M. Egenhofer and P. Agouris, "Using Ontologies for Integrated Geographic Information Systems", *Transactions in GIS*, vol. 6, no. 3, 2002, pp. 25-40.
- [11] N. Guarino, Formal Ontology, Conceptual Analysis and Knowledge Representation, *International Journal of Human and Computer Studies*, 43(5/6), 1999, 625-640.
- [12] A. Guzmán, C. Domínguez and J. Olivares, Reacting to unexpected events and communication inspite of mixed ontologies, In: C. Coello, A. Albornoz, L. Sucar, O. Cair and G. Kemper (cds.), Advances in Artificial Intelligence, Proceedings of Mexican International Conference on Artificial Intelligence (MICAI 2002), Vol. 2313, Lecture Notes in Computer Science, Springer-Verlag Merida, Yucatan, Mexico, 2002, pp. 377-386.
- [13] M. Torres, M. Moreno, R. Menchaca and S. Levachkine, Making Spatial Analysis with a Distributed Geographical Information System, *Proceedings of 21st IASTED International Conference on Databases and Applications (DBA' 2003)*, Innsbruck, Austria, February 10-13, 2003, pp. 1245-1250.

a measure frings are to care in significant

INCREMENTAL WEAK COMPOSITION AND INVOCATION OF GEOGRAPHIC WEB SERVICES

Carlos GRANELL, José POVEDA, and Michael GOULD Departamento de Lenguajes y Sistemas Informáticos Universitat Jaume I E-12071, Castellón, Spain {canut, albalade, gould}@uji.es

ABSTRACT

Geographic web services will soon become subsumed in the e-commerce world, a world where the composition of simple or atomic services to build compound services is a key characteristic. Since currently no detailed model of composite geographic web services exists, and within the scope of work on a European Union-funded project, we define such basic composition as part of an effort to test for geographic web service interoperability. Rather than adopting current static methods, we build on the concept of weak composition and provide a model for defining, composing and invoking compositions in a flexible manner. We demonstrate a prototype application for this purpose, and illustrate its utility through a simple arithmetic function composer scenario in which complex arithmetic expressions could be easily evaluated in base of the composition of atomic services. Benefits of this composition weak model over process-oriented alternatives are mentioned. Extension to geographic web services within an emergency management use case is proposed and necessary semantic and other extensions are outlined.

KEY WORDS

Web services composition, weak composition, semantic interoperability, interoperability testing.

1. INTRODUCTION

E-commerce has become so important over the past few years that it promises to become the global business paradigm of excellence in the near future. E-commerce services will naturally subsume the niche market of geographic information services, which are the ultimate focus of the work reported here. One of the principle characteristic for a mature and consolidated e-commerce will be to offer not only atomic services to be used ad hoc but also compound services based on the chaining of other services (atomic or also compound) which are then invocated using web technologies. The study of workflow (and corresponding services) has over the past few years

provided techniques for modelling the flow of activities, normally known a priori, that are executed procedurally within a controlled, homogeneous intra-organizational environment, and where a single participant controls the flow. Within the web context, however, it is not always possible to encounter services which are homogeneous with respect to their composition. In the heterogeneous and dynamic context that is the web, diverse services may appear and disappear without warning. For electronic commerce to reach its expectations, it would seem unwise to attempt migration of traditional process modelling techniques to these new open environments [1]. It would seem more logical to adapt the process models to inherent positive characteristics of the web, which implies weak composition (akin to late binding). For this reason, and within the realm of the European Union-funded project ACE-GIS [2] we propose a novel approximation defining a declarative composition model which is consistent with the open and dynamic characteristics of web architecture and behaviour. ACE-GIS proposes to build a developers platform for model-driven design and invocation of compound geographic information services (in addition to basic e-commerce services such as authentication and eventually payment). While judging the conformance of a novel web service to the implementation specifications within Open GIS Consortium is quite straightforward¹³, the leap from one-to-one conformance to true interoperability among diverse web services has yet to be realised. In fact, measurable interoperability has yet to be defined satisfactorily, which is why the authors undertook the present study.

The weak composition model described in [3] outlines a declarative composition model for web services, which is centered around descriptive aspects, on interoperability and on scalability. Basically, a web service composition is defined as a set of atomic or compound web services which interact according to certain logical rules. These logical rules specify the patterns of composition (serial, parallel, etc.) which describe the execution order of the services involved in the composition, and the connection flows established for the data flow between services.

¹³ See www.opengis.org and follow the link to Implementation Specifications.

In comparison with other composition models which are process oriented [4, 5], the approximation presented is essentially oriented to abstract interfaces and uses a backtracking-mode algorithm for the invocation. The resulting composition graph [3], as logical relations between services, is defined implicitly by the very structure of the composition. This graph may be termed a tree structure, whose leaf nodes are linked to atomic services, and whose internal nodes represent virtual intermediate compositions, where the root node represents and encapsulates the entire composition. During invocation this data structure is evaluated by the Interpretation Handler (see section 2) for the orchestrated invocation of the composition's services via the Invocation Handler (see also section 2). Only at execution time are the necessary connections identified and established in order to invoke all relevant services, hence one characteristic that defines the term weak composition. Internal nodes with compound descriptions provide both composition patterns and connection flow descriptions.

Weak composition differs from the architecture of traditional workflow management systems because the latter are essentially process oriented in the construction and execution of the composition graph. In existing approximations the developer normally must work with the entire graph, and so complexity increases noticeably as the composition grows. However, the approach described here holds the advantage that in the majority of process models, as complex as they may be, the logical composition graph may be decomposed into basic patterns, serial or parallel, of just two services at a time. The final composition encapsulates the complexity of the model in a manner similar to that of class hierarchy in object oriented languages; the main difference being that here we refer to abstract interfaces and not instantiations of classes.

The reader may be wondering why we have not exploited one of the several existing languages for the composition of web services, such as BPEL4WS, WSCI, BPML, etc. [6]. Within our perceived use case, of chaining 2 or 3 well-known geographic web services—say a Web Map Server to a Web Feature Server—to known e-commerce services such as authentication, is not necessary to overload the system with notions of conditional iteration of web services or feedback among them. Therefore, we consider web services composition without a procedural language, only based on appropriate pattern definitions and the inherent structure of the service description. In addition, there currently exists no international consensus on standard languages for composition; most are merely commercially-driven proposals [7].

Similarly, within traditional flow context, we recognise efforts toward defining standards for workflow interoperability (such as Wf-XML or SWAP) between workflow management systems [8], but this also has not shown a great deal of consensus.

For this reason, we have proposed a new approach different from current languages, and in the context of the goals of the ACE-GIS project.

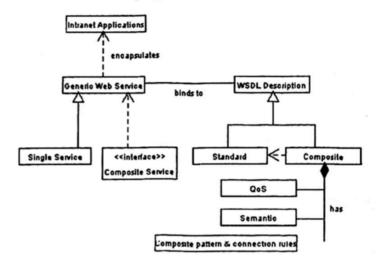


Figure 1. Relations in weak composition (after [9])

A fundamental requisite of organisations (businesses) in the near future [10], will be rapid and dynamic integration of applications and processes to adapt to the web environment. As a first step they often "wrap" applications accessible within the organisation as web services, to facilitate their access by all participants involved in a certain collaborative effort. Current web services technology permits the establishment of loosely coupled connections, which improves interoperability (syntactic at least) among heterogeneous services. At the moment of creating compositions of different available services, new relations appear among the basic elements which form the composition. These relations are shown in the class diagrams in figure 1.

From our viewpoint a general web service can be considered an atomic service or a composition of such general services. In either case the resulting composition should appear to the user as a single entity with which to interact. This affords several benefits including encapsulation of the underlying service complexity and straightforward reuse of a composition in future compositions. The standard manner of describing external behaviour of a web service is via Web Service Description Language, WSDL [11]. Each web service carries with it a WSDL description of the functionalities it offers. In treating a composition as an atomic web service (a so-called opaque service chain) this composition also carries a WSDL describing the entire functionality of the composition. However, in these composition descriptions only the abstract description is present, without details on concrete implementation: things such as access protocols or location points. Service compositions at the abstract level fit comfortably within the weak composition concept, unlike concepts of tightly coupled components within distributed computing models such as CORBA, RMI, etc.

In addition to the abstract WSDL description a composition incorporates the specification of composition patterns and connection flow established for the services

within a composition. These are encoded declaratively in an XML WSDL extension (see Annex for a complete example). Additional aspects such as the quality of service (QoS) desired in a composition, semantic attributes or security considerations, may also be embedded declaratively in the description. This together, abstract WSDL plus XML extension, forms an abstract interoperability interface of the composite service, a basic unit, which is interpreted by the Interpreter and Invocation Handlers (see figure 2) in the model presented. This composite WSDL description (which for the time being we will call WSDL++) is fundamental for the weak composition process, as the approximation follows an abstract component-based model, which facilitates interoperability and connectivity between services.

The remainder of the paper describes in some detail the process of composition and invocation, fundamental building blocks in the conceptual architecture which exploits the weak composition concept. Then we provide a simple example of a prototype implementation of composition and invocation. Finally, we provide conclusions and ideas for future research.

2. CONCEPTUAL ARCHITECTURE

To be able to define the function views that are established in the proposed architecture for composition and invocation of web services, we describe first a global vision of this conceptual architecture detailed in [3].

Basically, the architecture supporting the weak composition concept is composed of 5 layers providing a sort of middleware between the users and the external components such as web services, registries and catalogs:

- (i) User layer, formed by applications and interfaces directly accessible by the end user;
- (ii) Composition layer, responsible for generating the declarative description of the composition;
- (iii) Additional components layer, which contains other components of the composition, such as security, quality of service, semantics or system monitoring;
- (iv) Invocation layer, which interprets the composition and executes it;
- (v) Local storage layer.

From this conceptual architecture of middleware layers, we establish essentially two functional views of the system, to aid in composition and invocation of services:

- (i) Composition view;
- (ii) Invocation view.

The composition of services involves primarily components of the user and composition layers. At the user layer are found the components Searcher, Browser and Registry, while the components Composition Handler and Description Handler together form the composition layer of the conceptual architecture. Similar to the composition view, the invocation view contains components pertaining to the user layer, which are assigned to locate, select and navigate among the desired web services. Logically, this view also incorporates the invocation layer, which is necessary for the execution of web services compositions; this layer is comprised of the Interpreter Handler and Invocation Handler components.

The remainder of this section describes in detail the process of composition and invocation of web services, under the guidance of the weak composition conceptual architecture.

2.1 COMPOSITION VIEW

In figure 2 we illustrate a UML sequence diagram outlining the mode in which the search and composition of web services is carried out.

The first 4 steps define the search for available web services, atomic or already compound, supposing the presence of a service catalog accessible locally or remotely (ultimately via web). Once the services are selected, the composition process is initiated (steps 5-14). This composition is supervised by the user with assistance from the Coordinator component. During the composition process, first we establish a composition pattern and connection flow for the messages involved in linking a pair¹⁴ of web services (step 5). As mentioned earlier, component-oriented composition consists of decomposing complex compositions in basic patterns of pairs (avoiding the need for a flow language). In this manner the desired composition is constructed in multiple iterations of composing two at a time: for example, simple to simple forming complex, then complex plus another simple, etc., hence the qualifier incremental in the title of the paper. Current composition languages allow construction of several services simultaneously [6], however in our geographic web service context it is not realistic that a user would combine, for example, 10 services at the same time. Therefore, in principle, we prefer a simple and consistent manner to create compositions instead of increasing the cardinality of services composition each time. (This paired process is akin to software debugging, where it makes sense to correct and test one bug at a time. Here we are assured that each paired composition works before we proceed to link another to the composition.) In the prototype application presented the user manually establishes the connections between the two parts (services). One of the immediate goals is to eliminate this

¹⁴ There is no problem in defining patt erns of higher order or even some kind of model language.

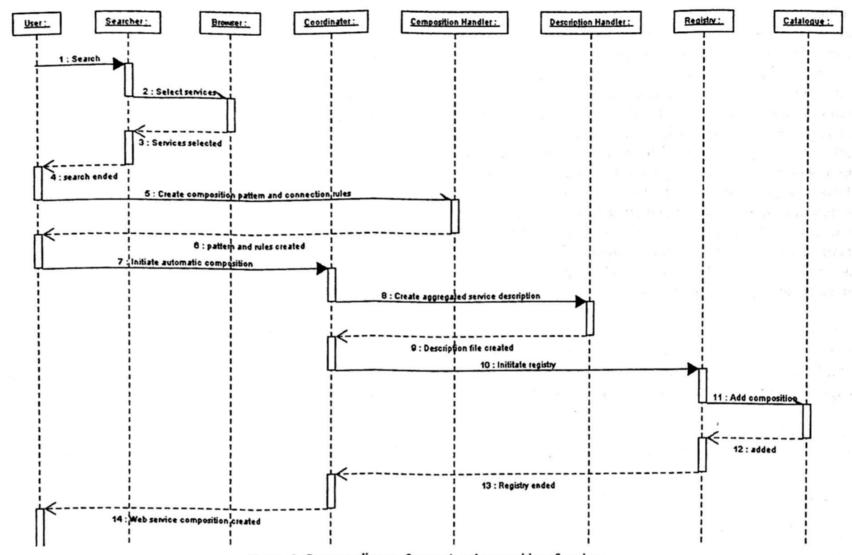


Figure. 2. Sequence diagram for search and composition of services

needs for human intervention to reach increasingly pure levels of automated composition, perhaps initially monitored by the user. This can only be accomplished through the addition of semantic translators, ontology parsers and addition of semantic tags [12], in line with the grand proposal of the semantic web [13]. These extensions are already contemplated within the presented conceptual architecture. For example, a semantic component describing web services and their data types will be in charge of enriching these descriptions from domain ontology (such as for example, transport experts' vocabulary) in order to allow an inference engine to be able to perform "matchmaking", to dynamically detect valid pairs of web services.

The composition process starting at step 7 is automated and controlled by the Coordinator component. During step 8 the Description Handler component encodes the composition patterns and flows in a declarative format based on XML [3]. Later the components from the Invocation layer interpret this encoding at the moment of invocation of the compound service. This format describes two fundamental aspects of the composition: how to connect and in which order to combine the two services, and what should be the external behaviour of the completed composition. The first aspect defines the type of composition, normally serial, and the connection flow for representing the data flow between the two services. The second aspect describes the abstract interface of the two services viewed as a single, independent web service.

On this occasion the abstract part defined in the WSDL specification is used. Finally the resulting WSDL++ description of the composition is added to the service catalog, thus exposing the new composition for future use (steps 10-13).

The WSDL++ description of the resulting composition is compatible with the WSDL standard. Any WSDL viewer should be capable of interpreting correctly WSDL++ descriptions produced by our prototype, ignoring the new tags added by our model.

2.2 INVOCATION VIEW

Figure 3 includes a UML sequence diagram which illustrates the mode of search and invocation of both atomic and compound web services.

In general terms the invocation of a composition requires an analysis of the associated description to encode the composition graph as a tree structure. Then, once this structure is created it is traversed in backtracking-mode, invoking the component web services and chaining the results obtained according to the connection flow defined in the WSDL++ description.

As in the composition view, first off the user selects the service or composition he or she wishes to invoke. The invocation process is totally automated and controlled by the Coordinator component. The user need only facilitate

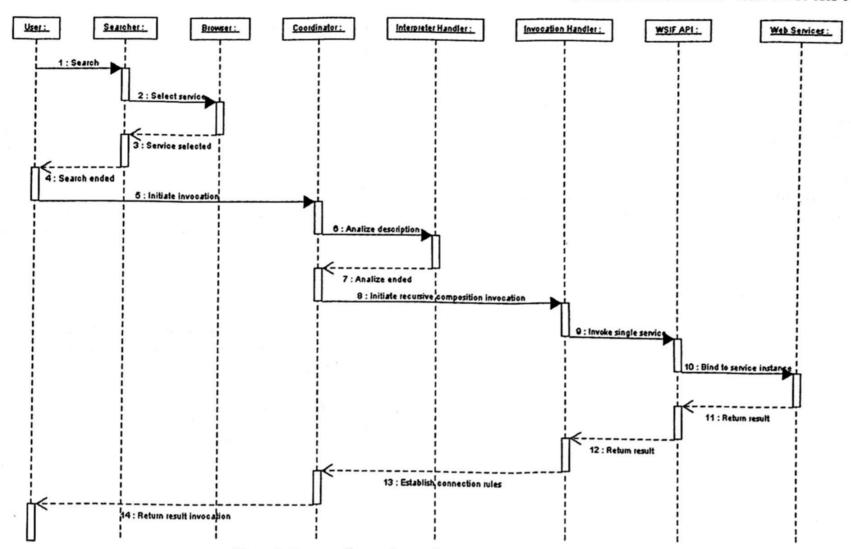


Figure 3. Sequence diagram for search and invocation of compositions

the initial parameters of the composition and await the composition's response (steps 5 and 14). The Interpreter Handler (step 6) realises two basic functions. The first, recursively interpret the WSDL++ of the composition. As all compositions are comprised of services, the Interpreter Handler analyzes the composition description in search of the services which compose it. If it encounters atomic services then the search halts and no further analysis is needed. If, however, it encounters compound services, the Interpreter Handler continues analyzing the new composition until it reaches a pair of atomic services.

This analysis algorithm extracts the composition graph from the very structure as defined by the composition description. The second function of the Interpreter Handler is the creation of the tree structure as an implementation of the composition graph. In this interpretation process we can appreciate how the composition graph is implicitly defined by the composition structure. As a consequence, at no moment does the user need to construct the complete graph at design time in order to create the desired composition.

Once the tree is created, the Invocation Handler component takes control to initiate the invocation process (steps 8-13). Invocation is a backtracking process where the leaf nodes are directly invoked as atomic web services (step 10) via the Web Services Invocation Framework (WSIF) [14], and open-source API provided by the Apache Software Foundation (www.apache.org) which

permits the invocation of web services in manner which is independent of protocol or access method (SOAP, HTTP GET/POST, etc.). Internal nodes on the tree are not directly executable because they describe virtual compositions. Their mission is to interpret the composition patterns and connection flows of the children nodes, that is, decide whether to execute them in series, parallel etc. and how to establish the data flow between these children nodes.

The current implementation of our model executes the service composition via the invocation of each instance and connects the information between services as a function of the data flow established by the user. This vision of static composition has availability problems if one of the services constituting the flow is not operational or available at the moment of execution. Future investigations are aimed at dynamic services composition at execution time, whereby the composition itself is able to detect functional anomalies and can substitute defective services with alternates in a transparent manner.

Aside from the mentioned limitations, backtracking-mode invocation through the weak composition model achieves a dynamic binding with concrete web service instances at execution time. The access mechanism to a service is not known until the moment of invocation. In this way, the composition using the abstract interfaces provides maximum flexibility compared to binding at design time of the composition.

3. AN EXAMPLE: ARITHMETIC FUNCTIONS COMPOSITION

In this section we present a web based prototype application which has been developed for the rapid integration, composition and invocation of web services, in an incremental fashion. Utilization of this prototype permits us to test and demonstrate a first approximation of the component based composition described here, its pros and cons, as well as the backtracking process in the composition invocation.

Following an interesting example of human arithmetic composition during World War II, described by Feynman¹⁵ [15], we compose functions based on a chain of basic arithmetic operations, each exposed as a web service.

This simple example is illustrative in order to demonstrate the capabilities of the prototype application. Future work, already under way, is to apply the prototype to geographic web services, beginning with well-known services such as those defined by Open GIS Consortium, such as a composition of a Web Map Service and a Web Feature Service. While these tests will be far more realistic and useful to an end user, they pose a degree of complexity which does not help us to explain the basic concepts in a didactic manner here.

Our simple example is based on the four basic arithmetic operations (addition subtraction, multiplication, and division) each of which is exposed as a separate web service. Each web service implements an operation through a single public method of the type:

Add (op1, op2): AddReturn
Subtract (op1, op2): SubtractReturn
Multiply (op1, op2): MultiplyReturn
Divide (op1, op2): DivideReturn

Imagine that we wish to solve compositions of arithmetic functions. For example if f y g are combinations of basic arithmetic functions, then $f \otimes g$ would be a valid composition of functions. The function f might be defined as the combination of a sum and a division, whereas the function g might be another combination of the operations subtraction and multiplication:

$$f(x, y, z) = \frac{x + y}{z}$$
$$g(u, v, w) = (u - v) * w$$

¹⁵ Nobel prize-winning physicist Fe ynman describes how, during the Manhattan Project in Los Alamos, he formed 'chains' of dozens of female data processors, each armed with a mechanical calculating machine and assigned to receive a punch card from the previous person, perform a single mathematical operation on it --addition, subtraction, cube root, etc.--and then pass it to the next person.

In this manner the composition $f \otimes g$ would take the following form:

$$f\otimes g=f\otimes g(x,y,z,v,w)=g(f(x,y,z),v,w)=((\frac{x+y}{z})-v)*w$$

If we analyze the resulting composition, it would be formed by a combination of simple arithmetic services. Evidently, it would be totally inefficient to attempt to create a separate and permanent web service for each possible arithmetic composition. In this case no service exists to resolve $f \otimes g$, however exploiting the combination of existing simple services it becomes possible to meet the special needs of the desired combination.

The remainder of this section describes in some detail how we realize the composition and invocation of arithmetic functions using the prototype which implements the weak composition model.

3.1 COMPOSITION OF ARITHMETIC FUNCTIONS

The construction of compositions in the weak composition model follows an incremental, top-down design, that is to say it is initiated with the union of atomic services to form intermediate compositions, until arriving at the compound compositions desired by the user. In order to create the composition $f \otimes g$, we previously would need to have created the intermediate compositions f y g. Consequentially, there are three new compositions awaiting in a registry, to be reutilized at another moment or by other users: f, g and $f \otimes g$.

Figure 4 shows the component Composition Handler for the creation of f at the moment of specifying the composition pattern and connection flow. This component is divided in two zones. The upper part shows information relevant for the user of two services to be combined, as for example the Add and Divide operations. In the lower part, the more interesting part, the composition pattern and connection flow for the two services are specified. Specifically, the output parameter AddReturn of the Add operation with the input parameter op1 of the Divide operation. The other input parameter of Divide, op2, is defined as a parameter external to the composition that is, introduced by the user at the moment of invocation.

Additionally, in this prototype it is possible to define input parameters as constants, whose value remains embedded in the WSDL description of the composition. Once the parameters defining the composition are configured, the Description Handler component creates the associated WSDL++ description and adds it to the registry of available services (see Annex for the complete WSDL++ description of composition f).

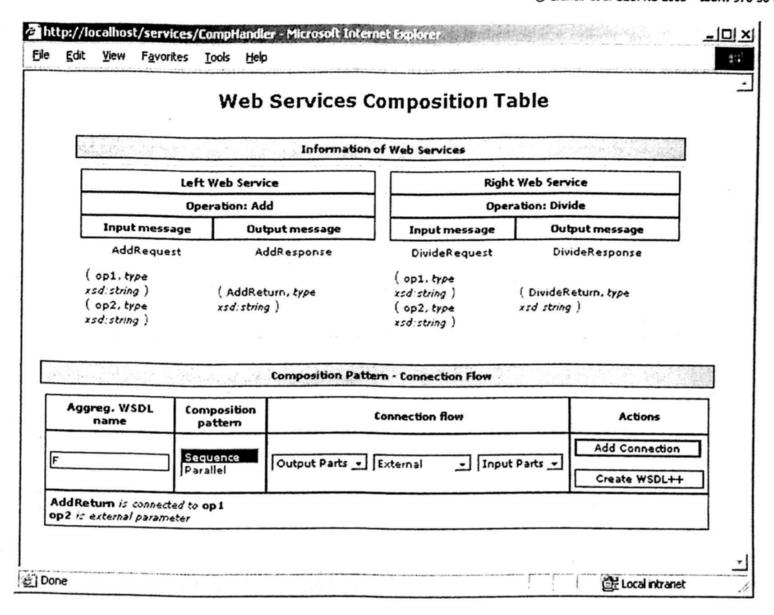


Figure 4. Specification of a composition pattern and the connection flow for the creation of composition f.

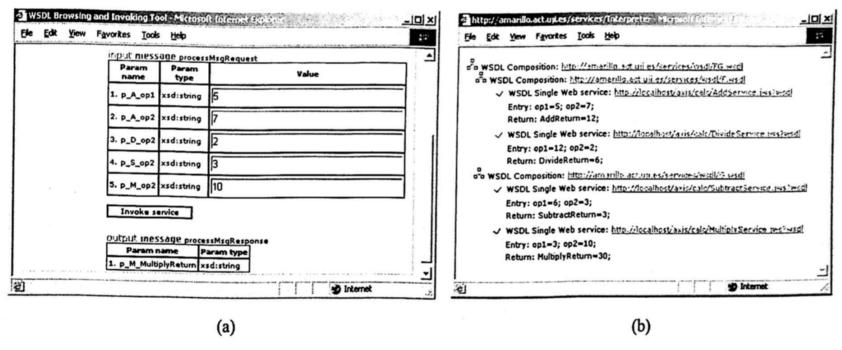


Figure 5. (a) - User sets initial parameters for $f \otimes g$ composition invocation; (b) - Result for $f \otimes g$ composition invocation.

In the same way, the user creates a composition g and then the final composition $f \otimes g$, from the intermediate compositions f and g. This iterative method of gradual composition encapsulates the complexity of the composition at design time. It is far easier for the user to compose $f \otimes g$ from f and g than from the basic arithmetic operations themselves.

Also, the incremental composition model alongside the possibility to independently invoke each intermediate composition created, facilitates debugging a priori. Utilizing this prototype it is possible to create each intermediate composition, execute it, and validate it correct operation. This characteristic helps save effort in later debugging steps in the final composition.

3.2 INVOCATION OF ARITHMETIC FUNCTIONS

The invocation process is centred on tree structure for the composition graph. The general invocation process of any composition is comprised of the following steps. The user introduces the initial composition parameters. Then the components of the Invocation layer are charged with creating the tree based on the description of the composition, with invoking the composition while establishing the data flows between connected services. Finally the component Coordinator returns the resulting data to the user.

Returning to the example of invocation of the composition $f \otimes g$, the user introduces the initial parameters of the composition (see figure 5a). Next, the Interpreter Handler is responsible for creating the tree structure. During the process of analysis the Interpreter Handler creates a node for each composition or atomic web service encountered, establishing double links between a father node and its children nodes. The initial invoked composition is identified by the root node. Each of the atomic services or intermediate compositions forming a given composition is converted into a child node of the node identifying the main composition. Specifically, the composition $f \otimes g$ is converted in root node and the compositions f y g as children nodes on the left and right respectively. Continuing the analysis recursively, the node associated with the composition f contains as children the atomic web services Add and Divide, while the node associated with the composition g contains the services Subtract and Multiply.

Once the tree is constructed, we need only to invoke it to obtain the composition's result. It is not possible in all cases to establish a priori the concrete value of the data flow between the services comprising the composition. Normally dependencies at execution time will impede the specification of concrete values for all parameters of web services. It is necessary to utilize a backtracking mode algorithm to traverse the tree, in order to realize

simultaneously the downward search and adjustment of inter-service data flow while returning. For example in invoking composition f, the services Add and Divide are defined in sequence. In this case it is not possible to begin execution of the service Divide without first obtaining the results of Add. In figure 5b we find the result of invoking the composition $f \otimes g$. In it we see the hierarchical structure of the invocation of the tree which represents the logical composition graph.

4. CONCLUSION AND FUTURE WORK

In this paper has been presented a model for weak composition of web services within the context of the ACE-GIS project [2], on-going research on geographic web service interoperability.

The model proposes a definition of web service compositions which are weakly or loosely coupled based only on abstract syntactic descriptions in WSDL. This weak composition is specified declaratively in XML, WSDL++, as an aggregation of external descriptions of the services as well as the composition pattern and connection flow.

From the inherent structure of this weak composition the Interpreter Handler generates a tree structure which the Invocation Handler then traverses in order to invoke the compound service.

With this model we attend to several aspects of the service composition problem:

- (i) Compositions are specified based on abstract descriptions of the services, linked to WSDL interfaces. This permits great flexibility in the integration of services which are a priori incompatible.
- (ii) Dynamic coupling of service instances is accomplished at execution time. The mechanism for accessing services is not known until the moment of invocation. This results in flexibility compared to design-time coupling.
- (iii) The oriented abstract interfaces composition and the backtracking-mode invocation of services via the logical composition graph are features of the proposed model. It permits that both compositions and simple web services may form a part of future compositions, which take on the appearance of atomic services.
- (iv) The construction of compositions follows an incremental design which hides composition complexity from the user (encapsulation), facilitates independent invocation of intermediate combinations, and facilitates the debugging and error detection in the final composition.
- (v) The combination of patterns and the inherent structure of the composition eliminate the need

- for a flow language. This main characteristic differentiates the weak composition concept from current languages for the composition of web services.
- (vi) This model defines a framework in which semantic aspects; quality of service, service cost, etc. could be included.

Thus far we have developed a prototype capable of guiding the user in the creation of compositions formed by an arbitrary number of existing web services. This prototype demonstrates the characteristics defining weak composition that is, permitting the composition of web services based on principles of interoperability and scalability.

As far as future work, we are labouring within the ACE-GIS project to compose a simple emergency management system combining a Web Feature Service (gas release plume model) and a weather service (reporting wind conditions at a given location); this is the logical next step toward supporting diverse geographic Additionally we are working with ACE-GIS partners (University of Münster) to incorporate semantic aspects (extensions of DAML-S) of the emergency management situation, with the goal of improving semantic interoperability within the composition. Finally, to help assure that the resulting service composition is somewhat fault tolerant with regard to service availability and quality of service, we are also investigating dynamic composition at execution time [16], so that the composition will be able to detect service faults and replace faulty services with substitutes meeting user requirements.

5. ACKNOWLEDGEMENTS

The work described here has been partly supported by the projects TIC-2000-1568-C03 (Spain Ministry of Science and Technology), IST-2001-37724 (European Union) and CTIDIB/2002/336 (Valencia government: Generalitat Valenciana).

REFERENCES

- [1] M.P. Singh, Physics of Service Composition, *IEEE Internet Computing*, 5(3), 2001, 6-7.
- [2] M. Gould, J. Poveda, & J. Huerta, ACE-GIS: Integración y Composición de Servicios Web Geográficos y de E-comercio. In: Toro, M. (Ed.): Proc. II Jornadas de Sistemas de Información Geográficos, El Escorial, Madrid, November 2002, 7-17. Project described in English at http://www.acegis.org

- [3] C. Granell, J. Poveda, & M. Gould, Composición Débil de Servicios Web, Submitted to professional journal. Also available on-line at: http://www3.uji.es/~canut/weakcomposition.pdf
- [4] D. Fensel & C. Bussler, The Web Service Modeling Framework WSMF, *Electronic Commerce Research and Applications Journal*, 1(2), 2002, 113-137.
- [5] B. Benatallah, Q.Z. Sheng, & M. Dumas, The Self-Serv Environment for Web Services Composition, *IEEE Internet Computing*, 7(1), 2003, 40-48.
- [6] S. Aissi, P. Malu, & K. Srinivasan, E-Business Process Modeling: The Next Big Step, *IEEE Computer*, 35(5), 2002, 55-62.
- [7] W.M.P. Van Der Aalst, Don't Go with the Flow: Web Services Compositions Standards Exposed, *IEEE Intelligent Systems*, 18(1), 2003, 72-76.
- [8] J.G. Hayes, E. Peyrovian, S. Sarin, M.-T. Schmidt, K.D. Swenson, & R. Weber, Workflow Interoperability Standards for the Internet, *IEEE Internet Computing*, 4(3), 2000, 37-45.
- [9] B. Benatallah, M. Dumas, M.C. Fauvet, & F.A. Rabhi, Towards Patterns of Web Services Composition. In S. Gorlatch and F. Rabhi (Eds): *Patterns and skeletons for parallel and distributed computing* (UK: Springer Verlag, 2002).
- [10] IDC analysts, Web Services Adoption Timeline and Related Business Opportunities, 2002. http://www.idc.com/en_US/st/extras/webServices.pdf
- [11] R. Chinnici, M. Gudgin, J.-J. Moreau, & S. Weerawarana, Web Services Description Language (WSDL) Version 1.2: Core Language. W3C Working Draft 11 June 2003. http://www.w3.org/TR/wsdl12/.
- [12] E. Sirin, J. Hendler, & B. Parsia, Semi-automatic Composition of Web Services using Semantic Descriptions, *Proc. of Web Services: Modeling, Architecture and Infrastructure Workshop at ICEIS 2003* Angers, France, 2003.
- [13] T. Berners-Lee, J. Hendler, & O. Lassila, The Semantic Web, *Scientific American*, 284(5), 2001, 34-43.
- [14] Web Services Invocation Framework, Apache Software Foundation, 2003. http://ws.apache.org/wsif/.
- [15] R.P. Feynman, The pleasure of finding things out: The best short works of Richard P. Feynman (Cambridge, Massachusetts: Perseus Publishing, 1999).
- [16] S. Ghandeharizadeh, G.A. Knoblock, C. Papadopoulos, C. Shahabi, E. Alwagait, J.L. Ambite, M. Cai, C.-C. Chen, P. Pol, R. Schmidt, S. Song, S. Thakkar, & R. Zhou, Proteus: A System for Dynamically Composing and Intelligently Executing Web Services. In the First International Conference on Web Services (ICWS), Las Vegas, Nevada, June 23-26, 2003.

MULTIDIMENSIONAL BINARY INDEXING FOR NEIGHBOURHOOD CALCULATIONS IN SPATIAL PARTITION TREES

José POVEDA, Michael GOULD

Departamento de Lenguajes y Sistemas Informáticos

Universitat Jaume I

E-12071, Castellón, Spain
{albalade, gould}@uji.es

ABSTRACT

We present a binary array encoding (location arrays) of the nodes in a spatial partition tree representing spaces of dimension k. This framework facilitates tree traversal for optimising access speed, and also supports simplified calculation of the neighbourhood of a subinterval of a particular partition. After defining the encoding we present a neighbour determination algorithm which extends work carried out by Samet [1] [2] and others on quadtrees. The primary extension is that the encoding and the neighbour determination algorithm extend to arbitrary dimensions beyond the 2-d quadtree case.

KEY WORDS

Spatial partition tree, binary location array, multidimensional indexing, multiresolution, neighbour calculation, terrain visualization

1. INTRODUCTION1

Quadtrees represent the primordial hierarchical spatial data structure, whose common salient feature is that of recursive decomposition of 2-d space. Quadtree types can be differentiated by the following characteristics:

- 1. The type of data used for its representation
- 2. The principle guiding the process of partitioning
- 3. The resolution (variable or not).

This data structure may be used for representing and handling point, curve, region and volume data. Decomposition of space can be regular at all the levels of the quadtree or non-regular, guided by the entry of the data to represent. The resolution of the decomposition (the number of times that subdivision is applied) can be set beforehand or can be guided for the properties of the input data. Depending on the specific application also a distinction can be made whether or not the structure is to define the border of a region, in the case of curves or

surfaces, or if it is used to define its interior in the case of areas and volumes.

One of the principal variants of the quadtree data structure is the region quadtree, so much so that Samet [1] uses the terms synonymously. As an example we present the region in figure 1 that initially we have represented like a structure of a binary array of 8 x 8 where the value 1 represents a pixel inside the region to represent and 0 if the pixel is found outside the region.

In the corresponding tree structure in figure 1, the root node corresponds to the complete array. Each son node represents a quadrant (labelled NW, NE, SW, SE) of the region represented by the node. The terminal nodes correspond to the blocks of the array in which it is no longer necessary to continue subdividing. A terminal node is represented in white or in black depending on whether its corresponding block is completely located inside the area to represent (all the elements of the array in this block contain the label "1") or that block is found completely outside the region of interest (all the elements of the array belonging to that block are labelled with "0").

These same spatial partitioning concepts extend to the third dimension in the representation of solids with octrees [3], although the encoding of blocks instead of quadrants becomes more complicated. In this paper we refer to spatial partition trees, in general, even though our initial examples treat the common 2-dimensional case using quadtrees. This is because the method and algorithm proposed here does not restrict spatial partitioning to 2-d but rather is extensible to spatial partitioning in arbitrary dimensions

2. MULTIRESOLUTION VISUALIZATION

Let us look at an application case where neighborhood calculation using a quadtree structure is a crucial element in interactive terrain visualization. For this application we propose to optimize the visualizatin of an extensive Digital Terrain Model (DTM) utilizing management of levels of detail (LOD), that is, multirresolution.

¹ The majority of this introductory section was adapted from Samet [1] [2].

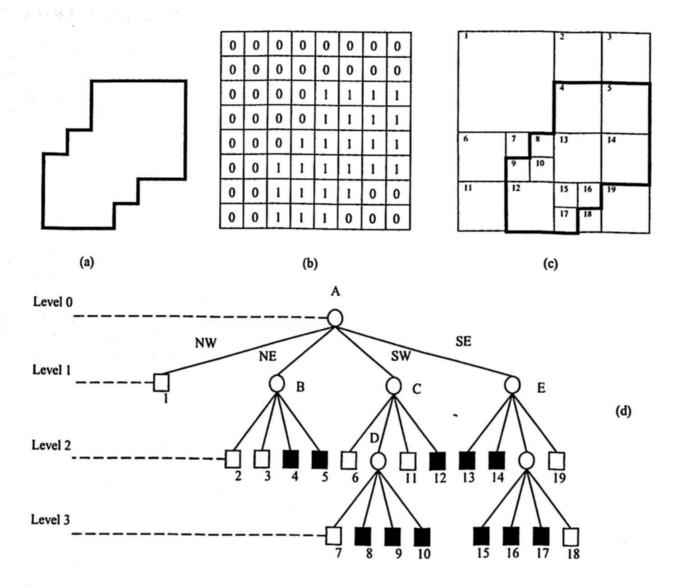


Figure 1. Example of quadtree (a) region, (b) structure of binary array, (c) maximum area blocks with constant value, (d) quadtree repres entation.

With this goal we define a process which renders to the screen relevant information, that is, only the information which might be appreciated from the viewpoint of the user. The data structure considered to manage the MDT with multiple levels of detail is the quadtree. In the quadtree nodes we store the elevation values which constitute the MDT as a regular grid with different levels of resolution, where each level of the tree corresponds to a level of resolution of the terrain.

In order to visualize the terrain, given a particular viewpoint, we traverse the quadtree and select the different levels of resolution represented by different levels in the tree. For the selection of nodes in the tree we use as main criterion the distance to the viewpoint, along with terrain roughness criteria. To insure against union of intervals with levels of resolution greater than 1, we run a balancing procedure over the pruned tree. Obtaining smaller regions after the pruning of larger adjacent regions, causes discontinuities (gaps) in the graphic representation of the DTM which cannot easily be corrected unless the difference in levels of resolution is no greater than one, see figure 2, in which case we make the correction interpolating the elevation of point b, taken from the region with higher resolution, which provoked the discontinuity, to the segment ac of the neighbouring region with lower resolution. For this process, therefore, it

is necessary to calculate the intervals which are neighbouring a given interval. Analogous to the multiresolution management, mentioned for the simplification of terrain geometry, it is also necessary to manage the simplification of the texture which overlays that geometry.

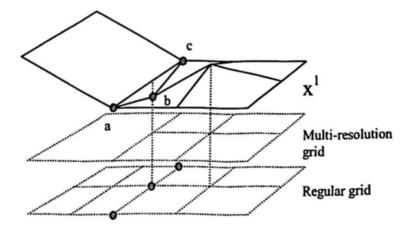


Figure 2. The union of differing levels of detail (LOD) here provokes a discontinuity in b: perspective view of the different layers utilized for generating the terrain.

3. BINARY ENCODING OF PARTITIONS

Frequently GIS related operations will require rapid access to a specific node in the tree. This normally requires traversal from root node downward through the tree, passing from node to node depending on the area representing the quadrant in question; this in turn supposes that position information is carried for each node. Normally quadtree quadrants are indexed according to a hierarchical base-4 ordering such as that proposed by Morton (the Z order in figure 1d). While this ordering simplifies operations in 2-d space and also facilitates human description in terms of four cardinal directions, a simpler referencing method is desirable as dimension increases. Here we describe such a method.

To increase access simplicity it is possible to apply binary encoding [4] to nodes so that these codes determine the traversal path down the tree. Rather than suppose four cardinal directions we encode each dimension with a 1-bit coordinate, describing the direction as positive or negative with respect to a given reference system. In particular, in two dimensions the positive directions --up and to the right (or East) -- are represented by a 1, while the negative directions --down and left-- are represented by 0. Consider the example in figure 3, where the "northwest" 16 quadrant at level 1 is assigned a 0 (negative direction) in the x1 dimension and a 1 (positive direction) in the x2 dimension. In this 2-d case the assumed reference system has its origin in the central point of the area where the quadrants meet. The example in figure 4 expands on this binary encoding, showing the notation to three levels.

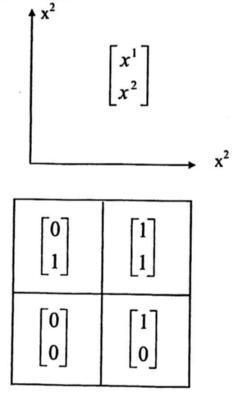


Figure 3. Binary array encoding in the 2-d case.

4. BINARY LOCATION ARRAYS

Taking into consideration the binary encoding of each dimension we define, as shown in figure 4, an array representation of each subset region which communicates both the location in the tree structure and, implicitly, the path to reach the node associated with that region.

The location array associated with a given node (here quadrant) is defined as the array of the father node, then adding on the right a new column vector with its own encoding according to the system in figure 3. The 3rd level node highlighted in figure 4 is thus represented by the 1,1 of the grandfather, the 1,0 of the father and then adding the 1,1 for its own level. Figure 5 shows additional examples to help clarify the encoding proposed. The location array representation of the shaded regions in figure 5 (four levels in 2-d space) would be the following:

$$A = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}; H = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$I = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$E = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}; F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$G = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Encoding higher dimensions follows the same principle described above. For example the following array P shows 5 levels of nodes (columns) in 4-d space (rows).

$$P = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Because the location of the quadrant¹⁷ (node) is defined by this simple binary location array, in section 6 we describe a sort of bit shifting on these arrays for simple neighbour calculation in any arbitrary dimension.

While it is all too tempting to use cardinal references, we again underscore that the positive/negative direction notation applies beyond the 2-d case equally to all dimensions.

¹⁷ Here the use of the term quadrant re fers to any interval of the spatial partition, at any dimension, and associated in the tree as a node.

5. MULTIDIMENSIONAL NEIGHBOUR CALCULATION

To achieve maximum data access efficiency and simplicity, necessary for applications such as interactive terrain visualization, it is useful to be able to reason about neighbour locations in the tree, identifying all possible neighbours of each node. These neighbourhood references are useful for tasks such as database paging or for adjusting resolution in a geometric model during an interactive flight over terrain [5] [6] [7]. If the terrain model includes multiple levels of resolution as a function of attributes such as distance to the viewer, then for each frame it would be necessary to recalculate the spatial partition tree and the neighbour relations of each node because quadrants of distinct levels of decomposition can cause discontinuities (gaps) which must be corrected by adjustment to the lower resolution. Because these applications will need to continuously recalculate these neighbour relations, it is essential that the algorithm for doing so deal efficiently with neighbours at multiple resolutions.

As a solution to clarify and optimise neighbour identification, we utilize the location array described above, in a simple algorithm of linear computational cost with respect to the depth of the tree.

To calculate the neighbour of E (in figure 5) in the positive direction at dimension 1 and at the same level of resolution (that is to say, quadrant of equal size or node at the same level), we begin with the location array of E and we visit each element, indicated with a dot, in the top row from right to left. Alternating the bits as we go along the first 1 becomes a 0, the second 1 a 0, the third also a 0, then the 0 a 1. When we negate a 0 and it switches to 1, at any point in the row, the process stops and the resulting location array describes the neighbour found, in this case F.

$$E = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = F$$

Now, if we continue applying the same algorithm we may calculate at the same level of resolution the neighbour of E in the negative direction in the second dimension ("down"), producing the location array which represents region G (see figure 5). This time we have alternated the bits in the second row (representing the second dimension), until a 0 is obtained, at the first move in this case.

$$E = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} = G$$

In general terms, then, we can state that in order to search for a neighbour in any positive direction, the stop condition in the bit shifting is the negation of any 0 to a 1. Equally then, to search for a neighbour in any negative direction the stop condition is the negation of any 1 to a 0.

Algorithm 1a (Neighbour identification in direction +xi)

- First element processed: last element of the row i, A(i,j) in the location array.
- 2.- If the processed element is a 0, we negate this element and the process stops having encountered a 1. The resultant array is the location associated with the neighbouring node in direction xⁱ.
- 3.- ELSE the processed element negates to a 0 and repeat step 2 with the next element to the left of that row.

For the case of a neighbouring node in the opposite (negative) direction, the algorithm would be modified as follows.

Algorithm 1b (Neighbour identification in direction -xi)

- 1.- First element processed: last element of the row i, A(i,j) in the location array.
- 2.- If the processed element is a 1, we negate this element and the process stops having encountered a 0. The resultant array is the location associated with the neighbouring node in direction xⁱ.
- 3.- ELSE the processed element negates to a 1 and repeat step 2 with the next element to the left of that row.

The principal advantages of this neighbour calculation algorithm are its generality with respect to spatial dimensions and its simplicity. An alternate representation of the algorithm (for either direction) is the given Calculate Neighbour algorithm.

6. THE MULTIRESOLUTION CASE

As stated earlier, applications such as real-time terrain generation and visualization depend on the optimal determination of multiple levels of resolution for any particular view [6], [7], [8], [9], [10]. Let us now demonstrate a method for exploiting the location array notation described above, to determine the neighbours of X (in figure 6) at multiple levels of resolution.

Algorithm CalculateNeighbour (A[i][j], ± e k)

Input. Location array A[i][j] of the region quadtree of interest, in direction e_k , where r is the level of resolution. Output. The array of the neighbour encountered in direction e_k

The array representation of the shaded regions in figure 6 would be the following:

$$X = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}; B = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}; D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}; F = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ; E = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} ; I = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

We first calculate for the first dimension the sameresolution neighbour of region X, in the positive direction, yielding D as this "eastern" neighbour.

$$X = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & i & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \dot{0} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = D$$

In the next step we identify all sub-arrays which represent neighbouring regions at lower dimensions. This is accomplished merely by eliminating the final column on the right side. The number of columns remaining informs of the level of resolution of each neighbouring region. In this case we obtain the four arrays associated with the neighbouring regions of X in the positive direction at the first dimension: D at the same resolution (4th level), G at the 3rd level, H at the 2nd level and finally I at the 1st level (or entire quadrant).

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = D; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = G$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = H; \begin{bmatrix} 1 \\ 0 \end{bmatrix} = I$$

Applying again the algorithm in the first dimension we calculate the neighbours of region X in the negative direction:

$$X = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = C$$

We obtain in the first step, the array characterizing the only neighbour of G in the negative direction. Repeating the process for the second dimension (the bottom row), in the positive direction, we obtain:

$$X = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} = A$$

and for the neighbours in the negative direction:

$$X = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & \dot{0} \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 1 & \dot{0} & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & i & 1 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = B$$

Therefore, we have determined in three steps that B is the neighbour of X at the same resolution and then, independent of the resolution, that X has three neighbouring regions which we deduced from the original array, as follows: B at the 4th level, F at the 3rd level and E at the 2nd level. The objective region X has no neighbour at the 1st level of resolution in this negative direction.

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = B$$

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = F; \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = E$$

Although beyond the third dimension we cannot easily visualize geometrically the indexing concepts described here, from a logical point of view we can demonstrate how the method applies equally well in k-dimensions. This in contrast to traditional quadrant notation following cardinal directions which are restricted to the 2-d plane.

To illustrate this point we will calculate, in a 4-d space (using the array P from section 4), the neighbours in the positive direction and in the 4th dimension, and additionally at all levels of resolution.

$$P = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

We apply the neighbour algorithm described earlier:

$$P = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & i \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & i & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

As is illustrated above, the process stops when a 0 to 1 bit shift occurs. We then determine using the method described earlier, that there exist three regions neighbouring P at 3 levels of resolution, described by the following arrays:

$$P_{\mathbf{I}} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

$$P_2 = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \text{ and } P_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

7. MULTIPLE NEIGHBORHOODS

Based on the neighbourhood algorithm and the location matrix notation of an interval I, we have seen how to calculate the location matrix of the neighbours of I at the same level of resolution. Based on that matrix we have also seen how to determine all other possible neighbours at lower resolution than I for a given quadtree partition. We now complete this section with the determination of the neighbours of I at higher resolution, which implies necessarily multiple vecinity in the same dimension. To illustrate this point we begin with the following quadtree partition, in figure 7.

We propose that the calculus of the neighbours of interval X in the positive direction and in the second spatial dimension, as seen in figure 7, results in the intervals A, B and C, whose associated location matrices are:

$$X = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}; C = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

On the basis of the location matrix defining the interval X, we calculate the neighbouring interval in the positive

direction of the second dimension and at the same level of resolution as I.

Once the location matrix of the neighbour interval at same resolution is obtained, we verify its existence in the quadtree to determine if it is a terminal node or, on the contrary, possess son nodes. For the terminal node case, the process concludes obtaining the location matrix of the

neighbour of X in the direction and dimension desired. For the case of son nodes, we repeat the process for only those sons in the opposite direction of the dimension of interest. Below we show how to calculate neighbour location matrices, using the tree structure navigation described.

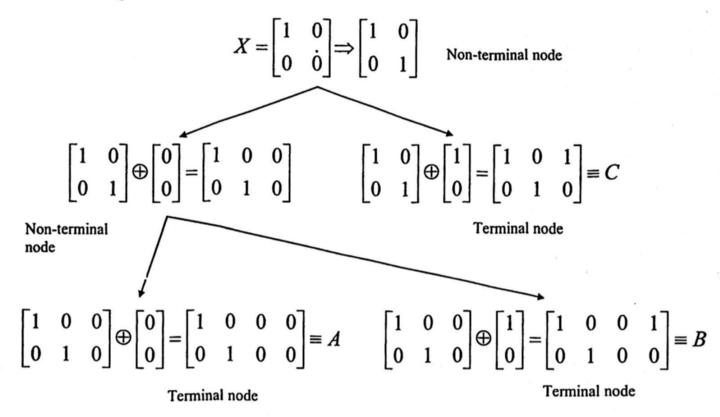


Figure 8. Calculus of the neighbours of X at higher resolution in the second dimension and in the positive direction

For the case of neighbours of X in the negative direction and the first dimension, we would derive an analogue process:

Therefore the terminal nodes represent the neighbours of interval X at equal or higher resolution. For the case of the other two directions which for the moment we have not considered, we show how we would arrive at singular cases. In the first place let us calculate the neighbours of X in the negative direction of the second dimension.

$$X = \begin{bmatrix} 1 & 0 \\ 0 & \dot{0} \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 \\ \dot{0} & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

As we did not encounter any unitary (one) value, we deduce that the interval X is on the border for the negative direction at the second dimension, and therefore has no neighbours in this partition. For the case of neighbours in the positive direction of the first dimension:

$$X = \begin{bmatrix} 1 & \dot{0} \\ 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

in which being a terminal node, concludes the defined process and thus we deduce that it is the only neighbour of X in the positive direction of the first dimension of equal or greater resolution. We conclude the calculation of neighbours of X at level of detail equal or less than X, identifying as neighbours all those terminal nodes obtained in the trees created. We are left with the location matrices obtained as neighbours of X in the positive direction of the second dimension, those corresponding to the matrices:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \equiv C \; ; \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \equiv A$$

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \equiv B$$

and for the negative direction of the first dimension

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \equiv D; \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \equiv E$$

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \equiv F$$

as may be verified in the figure 7, where we have labelled the intervals which are neighbours of X.

8. COMPLEXITY EVALUATION

Finally, we provide a proposition in order to define the complexity of the neighbour calculation algorithm described in this paper.

Proposition

Given that T is a quadtree of depth h, the neighbour of a given node n in a particular direction can be encountered at a maximum cost of O(h).

Demonstration

In the best case, in which a neighbour in a given direction has the same father in the tree, the computational cost will be of one comparison, such as in the case of E, G in figure 5.

In the worst case, that of E, F in figure 5, according to the algorithm presented here the number of operations (comparison/assignment) will be at maximum that of the level of resolution of the region considered, and therefore corresponding at maximum to the depth of the tree.

9. CONCLUSIONS AND FUTURE WORK

We have presented a binary encoding notation for representing spatial partitions, and therefore have introduced the concept binary location arrays. Secondly, we presented an algorithm for exploiting this notation for the simplified calculation of neighbours of any region, in arbitrary dimensions. This notation provides substantial generality over current quadtree notations which do not extend well into higher dimensions. Thirdly, we demonstrated how the neighbour calculation algorithm extends to determine neighbours at any level of resolution.

Future work may include application of the algorithm to interactive terrain visualization and other potentially interesting applications, and extensions for other data structures.

ACKNOWLEDGEMENTS

This work was partially supported by European Union projects Esprit 25029 (GeoIndex) and IST-2001-37724 (ACE-GIS).

REFERENCES

- [1] Samet, Hanan, "The Quadtree and Related Hierarchical Data Structures", ACM Computing Surveys 16, 1984, pp. 187-200.
- [2] Samet, H. "Applications of Spatial Data Structures", Addison-Wesley, 1995.
- [3] M. de Berg, M. van Kreveld, M. Overmars and O. Schwarzkopf, "Computational Geometry, Algorithms and Applications", Springer, Second Edition, 2000, pp. 291-303.
- [4] Goodchild, M.F. (Ed.), 1990. "Quadtree algorithms and spatial indexes". Technical issues in GIS. NCGIA Core Curriculum. Unit 37.
- [5] De Berg, M., Dobrindt, K. T. G., "On levels of Detail in Terrains", 11th ACM Symposium on Computational Geometry", 1995, pp. 1-15.
- [6] Falby, J. S., Zyda, M. J., Pratt, D. R., and Mackey, R. L., "Hierarchical Data Structures for Real-Time Three-Dimensional Visual Simulation", Computer and Graphics 17, 1993, pp. 144-151.
- [7] De Floriani, L. and Puppo E., "Hierarchical Triangulation for Multiresolution Surface Description", ACM Transactions on Graphics 14, 1995, pp. 363-411.
- [8] Poveda, J., Gould M., Sevilla, J. "GeoIndex Project: A virtual window to Geographic Informations Systems". In Rault, J.-C. (Ed.): La Lettre de l'Intelligence Artificielle, Proceedings of the International Conference on Complex Systems, Intelligent Systems & Interfaces (NIMES'98), Nimes, France. Nimes 1998.
- [9] Schee, L. H., Jense, G. L., "Interacting with Geographical Information in a Virtual Environment", Proc. Joint European Conference on Geographical Information. The Hague, NL, 1995, pp. 151-156.
- [10] Schroder, F. and Rossbach, P., "Managing the Complexity of Digital Terrain Models", Computer and Graphics 18, 1994, pp. 775-783.

12. ANNEX OF FIGURES

	10 11		11 11		
0		10 11		110 101	111 101
				110 100	111 100
00 01	01 01	100 011	101 011	11 01	
		100 010	101 010		
00 00	01 00	10 01		11 00	

Figure 4. Binary location arrays of a quadtree to the 3rd level

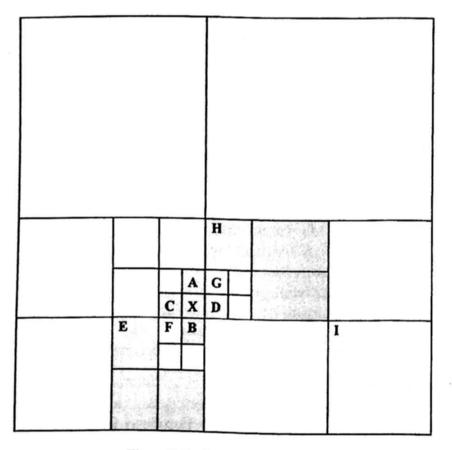


Figure 5. Region quadtree example

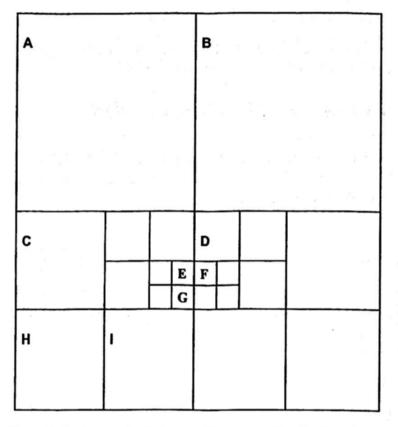


Figure 6. Study case of neighbours of X at several levels of resolution

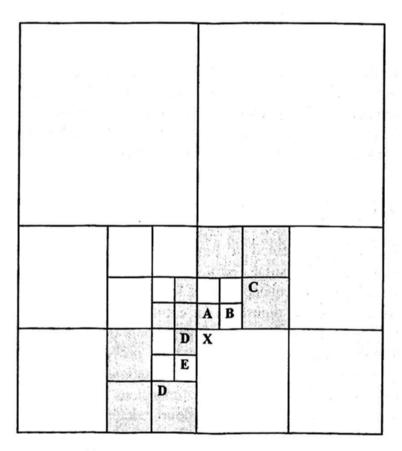


Figure 7. Partition used for calculus of ne ighbours of interval X in the positive direction of the second spatia I dimension. The result should be A, B and C, each a different (lower) levels of resolution.

MULTI-LEVEL TRANSPORTATION NETWORK ADAPTABLE TO TRAFFIC NAVIGATION

Jun FENG
Graduate School of Engineering,
Nagoya University
Furo-cho, Chikusa-ku,
464-8603 Nagoya, Aichi, Japan
feng@watanabe.nuie.nagoya-u.ac.jp

Naoto MUKAI
Graduate School of Information Science,
Nagoya University
Furo-cho, Chikusa-ku,
464-8603 Nagoya, Aichi, Japan
naoto@watanabe.nuie.nagoya-u.ac.ip

Toyohide WATANABE
Graduate School of Information Science,
Nagoya University
Furo-cho, Chikusa-ku,
464-8603 Nagoya, Aichi, Japan
watanabe@is.nagoya-u.ac.jp

ABSTRACT

In a country wide Intelligent Transportation System (ITS), there is a need for processing road map and transportation information in many levels of details. In this paper, an integrated method for representing transportation information with multi-level road map is proposed. This method adopts a spatial index structure for managing map objects in multiple levels and uses an integrated method for representing the travel junctions (or traffic constraints), travel cost on road segments and turn corners. Based on the datasets generated by this method, queries in ITS applications can be responded efficiently in different levels of details.

KEY WORDS

Application development, transportation network, road map, spatial index, traffic constraint.

1 INTRODUCTION

geographic database and the management facility/environment is one of the important subjects concerned with Intelligent Transportation System (ITS). A geographic database may contain static map data (including data of road network and other map objects), public transportation routes, and current travel cost (e.g., travel time) on segments of transportation network [1]. In a country-wide system, there is a need for processing the map and transportation information in many levels of details. This is due primarily to the desire of user to use (/see) relevant information only; too many details may hinder rather than help [2]. For example, a query of path search between two city halls belonging to two separated prefectures is solved more efficiently by dividing into three sub-queries: one sub-query for finding path between two cities based on the country-level and prefecture-level roads (i.e., highways and crossing-prefecture roads), and two sub-queries based on city-level roads (all the streets inside the city) for finding the path between the city halls and the upper level roads. Since the generalization of map information cannot be realized automatically [3], various methods have been proposed for maintenance of scaleless maps or multi-scales of maps [2, 3, 4]. However, transportation network is different from road maps. It is

important to identify one-way roads with attributes d links, traffic constraints, information about turns between links, or access conditions from one link to another[5]. Moreover, for some important route planning problems, the turn cost, which is encountered when we make a turn on a cross-point, are also taken into consideration [6]. The traditional representation method represents transportation network with nodes and arcs. The traffic constraints and turn costs are managed by adding new nodes and arcs. The total number of nodes and arcs in the dataset is multiplied [6, 7]. Furthermore, as the neighbouring nodes in different level of details do not keep the same, the transportation information sets should be generated for every level separately. And there are two problems for the generated datasets: one is the complex maintenance process; and another is that it is difficult to support a path search based on the transportation networks in different level of details.

To solve these problems, in this paper, we propose an integrated representation method for the multi-level transportation network. Our method adopts a spatial index structure for managing map objects in multiple levels and uses an integrated method for representing the travel junctions (or traffic constraints), travel cost on road segments and turn corners. Based on the datasets created by this method, queries in ITS applications can be processed efficiently by using proper search methods.

This paper is organized as follows. The related works for representation of multi-level road network and transportation network are presented in Section 2. The representation method for integrated management of multi-levels traffic conditions and spatial information about road network is proposed in Section 3 and Section 4 Section 5 analyses our method and Section 6 makes a conclusion on our work.

2 RELATED AND PREVIOUS WORK

The issue of this paper refers to the management of multilevels of map and the integrated representation of transportation information and spatial information about road network.

2.1 INTEGRATED REPRESENTATION OF TRAFFIC INFORMATION AND ROAD NETWORK

To represent the traffic information on road network, a typical method [7] represents the road network using a directed graph. In the graph, each arc depicts a one-way road and each node corresponds to a junction. Two-ways roads can be presented as a pair of arcs: one in each direction. However, extra nodes should be added to the graph when there are any access limitations (constraints of specific traffic controls). In other words, one node on the road network may be represented with several vertices corresponding to the junctions, and they are independent with each other. Figure 1 gives the representation of different types of roads and junctions: one-way road, twoways road without any access limitations, and T-junction with some access limitations (the center point on this Tjunction is represented by two vertices in this directed graph).

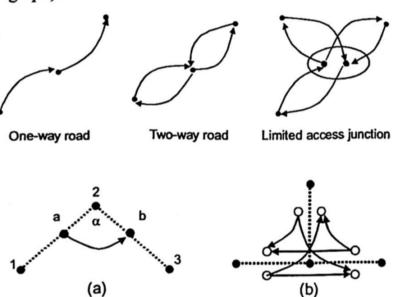


Figure 2. A primary graph (a, dotted lines) and its pseudo-dual graph (a,

black), and a T-road primary graph (b, dotted lines) and its restricted pseudo-dual graph (b, black) [6]

Figure 1. Different types of roads and junctions [7]

Since this representation method ignores the spatial attributes of map objects, only the routing queries are applicable well on this model. An architecture was proposed in [8] for keeping traffic information on nodes of road network. However, the information of traffic constraints and turn costs on the nodes is omitted in their discussion. To represent the traffic cost and the cost of turns, a method in [6] was proposed. The turn cost is represented by a pseudo-dual graph. Figure 2 gives the pseudo-dual graph for representation of turn costs on a primary graph (Figure 2(a)) and a T-junction with some access limitations (Figure 2(b)). The turn cost is represented with the additional nodes and arcs of the pseudo-dual graph: in (a), two nodes a and b are added to the primary graph, and an arc α from a to b represents the turn cost of leaving the traffic arc from node 1 to node 2 and entering the traffic arc from node 2 to node 3; and in (b), there are six nodes and six arcs which are added for representing turn costs between possible pairs of traffic arcs. The cost of search algorithms (e.g. Dijkstra's

algorithm [9]) becomes high. Moreover, the pseudo-dual graph is insufficient (and needs reference to the primary graph) for route drawing.

In this paper, we propose a flexible method for representing transportation networks, including basic road map, traffic constraints, travel cost on road segments and (or) turn cost from a traffic link to other "turnable" links. Moreover, our method is proper for representing multi-level transportation network.

2.2 METHODS OF MANAGING MULTI-LEVELS OF TRANSPORTATION NETWORK

To achieve an efficient navigation process on transportation network, many methods were proposed to pre-compute hierarchical road networks [10, 11, 12, 13]. Proper partitions are done to the transportation network, and some parts of the path length are pre-computed. In their methods, all the road segments are regarded as the same level of detail - basically, all the navigation process is based on the most detailed map. They do not support the navigation based on transportation networks on different level of details.

Contrary to those methods based on the most detailed map, multi-resolution data models provide ways of describing the world at various levels of detail. The importance of such data models in the context of spatial information is widely acknowledged, and there are several studies of their formal foundations [14, 15, 16]. Under these models, to realize the sharability among multi-levels or rapid access to multi-levels of maps, map objects are arranged in multi-levels based on a compromise between storage and computation. Although the sizes of storage, the speed of map zooming or effective spatial process on multi-levels of maps were studied in these works, they share the same complex process to keep the consistency among multi-levels of map information. To solve this problem, an access method, on Multi-levels Object-Relation (MOR-tree) structure, which possesses the ability of the followings, is proposed [4]:

- The ability of handling spatial data efficiently;
- The ability to provide integrated access to multiscale maps;
- 3. The ability of arranging the relations among multilevels of maps.

MOR-tree was proposed based on M^2 (Multi-level/Multi-theme) map information model [17] for managing multi-levels (of scale) of maps. M2 model is powerful to deal with various levels of maps uniformly, especially for managing multi-levels of road networks in comparison with other models [15, 16]. However, MOR-tree cannot be applied to transportation information directly, as the transportation information on road network is more complex than the original map. In this paper, an integrated representation method for multi-level transportation network is proposed.

3 REPRESENTATION OF TRANSPORTATION INFORMATION

Not only the kinds of information but also the management method of transportation information affects the processing efficiency of queries in ITS applications. In this section, we propose a representation method for integrating traffic information and spatial information about road network by considering the followings terms:

- The traffic conditions change continuously, and the snapshot of conditions is recorded as traffic information. In comparison with the traffic information, the map of road network is seldom updated, and can be regarded as static information. Therefore, if the static information is managed by an efficient structure, the changes of traffic information associated with the road map should not disturb the stability of the structure.
- The integrated representation should not only support the spatial query on road network and the temporal query on traffic information, but also support the interaction between these two kinds of queries.

A road network with nodes and links representing respectively the crosses and road segments can be regarded as an un-directed graph G, G = (V, L), where V is a set of vertices $\{v1, v2, ...vn\}$, and L is a collection of lines $\{l1, l2, ...lm\}$. Traffic information on the road network is regarded as a directed graph G', G' = (V, A), where V is a set of vertices $\{v1, v2, ...vn\}$, and A is a collection of arcs $\{a1, a2, ...ap\}$.

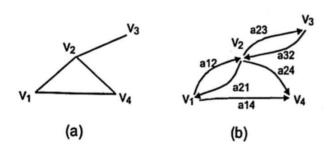


Figure 3. Road segment and traffic arc

Figure 3 depicts these two kinds of graphs. In the undirected graph of Figure 3 (a) road segments are represented by lines, while in the directed graph of Figure 3 (b) junctions are represented by arcs. One line for road segment in Figure 3 (a) may be corresponded to two arcs in Figure 3 (b) with two-direction traffic information. In addition to the directions of traffic, there are usually traffic controls (constraints) on road network: for example, the right-turn and U-turn are forbidden on some cross-points, which constrain the action of traffic. The typical road junctions with (or without) constraints are given in Figure 4 and Figure 5. Road junctions are represented by using [7]'s model in Figure 4(a)(1), Figure 4(b)(1) and Figure 5(1). Considering the shortcomings of this simple model, we propose super-node representation

method for integrating junctions (including traffic cost and traffic constraints) and road network.

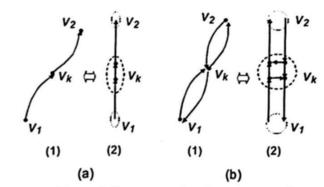


Figure 4. One-way road and two-ways road

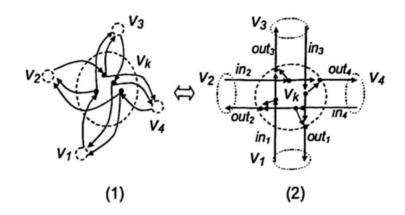


Figure 5. Cross node with constraint

A super-node can be defined as a node in road network: for example, vk in Figure 4 (a) (2), Figure 4 (b) (2) and Figure 5 (2). The information on the super-node contains the following parts (for simplicity of explanation, road junctions in Figure 5 (2) are used as an example):

1. Cost-arc: The arcs which have vk as their final vertex are called in-arcs, denoted as ini, and similarly the arcs which have vk as their initial vertex are called out-arcs, denoted as outj. The number of those arcs is called as in-degree (e.g., 4) and out-degree (e.g., 4), respectively. Every outi is defined as a Cost-arc, which consists of the final vertex of this arc and the traffic cost of this arc. Cost-arcs of vk in Figure 5(2) are:

$$\begin{bmatrix} out_1(v_1, \cos t_{k1}) \\ out_2(v_2, \cos t_{k2}) \\ out_3(v_3, \cos t_{k3}) \\ out_4(v_4, \cos t_{k4}) \end{bmatrix}$$

2. Constraint-matrix: The constraints on the super-node can be represented with an $n \times m$ matrix CM:

$$CM(v_k) = in_2 \begin{cases} out_1 & out_2 & \dots & out_m \\ C_{11} & C_{12} & \dots & C_{1m} \\ C_{21} & C_{22} & \dots & C_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ in_n & C_{n1} & C_{n2} & \dots & C_{nm} \end{cases}$$

And C_{ij} equals to 1 when there is restriction from *ini* to *outj*; C_{ij} equals to 0 when there is a junction from *ini* to *outj*. Constraint-matrix for vk in Figure 5 (2) is:

$$CM(v_1) = in_2 \begin{cases} in_1 & out_2 & out_3 & out_m \\ in_2 & 1 & 0 & 0 & 1 \\ in_3 & 0 & 1 & 1 & 0 \\ in_4 & 0 & 0 & 1 & 1 \end{cases}$$

Where there are restrictions on going from in1 to out1 and out4, from in2 to out1 and out2, from in3 to out2 and out3, and from in4 to out3 and out4. If there is no restriction for any ini of the super-node vk, Constraint-matrix of vk is filled with 0, and is regarded as NULL.

Moreover, our method is able to process the turn cost by extending Constraint-matrix to a Turn-Cost/ Constraint-matrix. The CM can be modified to a Turn-Cost/Constraint-matrix. For example, the Turn-Cost/Constraint-matrix for vk in Figure 5 (2) may be like this:

$$T_{-}CM(v_{1}) = in_{2} in_{3} in_{n} \begin{pmatrix} MAX & 10 & 40 & MAX \\ MAX & MAX & 10 & 30 \\ 10 & MAX & MAX & 30 \\ 10 & 40 & MAX & MAX \end{pmatrix}$$

Where MAX is defined as a large constant value. The element TCij in this matrix with a value of MAX represents a restriction from ini to outj: e.g., U-turn and right-turn are forbidden in this example. So, MAX is assigned to TCii (i = 1,..., 4), TC_{14} , TC_{21} , TC_{32} and TC_{43} . The value of TC_{12} represents the cost 10 (e.g., 10 seconds) of making a left-turn on the cross-point (from in1 to out2), while the cost of crossing the point vk from in1 to out3 is 40.

This method decreases the redundancies in the database and it is easy to integrate the traffic information and the basic road network. For the basic road network, the additional information for traffic information is managed on every node. When the number of nodes and traffic arcs is unchanged, the modification to any of the traffic information does not injure the stability of the spatial index structure (e.g., R-tree [18]) for road network. Therefore, queries in ITS application, which refer to the spatial information, can be solved by taking advantages of the spatial index. Another kind of queries, which refer to traffic information, can also be solved effectively. In the next section, we center on solving the second kind of queries by integrating the transportation information with the multi-level road map.

4 REPRESENTATION METHOD OF MANAGING MULTI-LEVEL TRANSPORTATION NETWORK

Multi-level transportation network is an extended part of the multi-level road map model, and is managed by an extended MOR-tree.

4.1 M² MAP INFORMATION MODEL

M² map information model is designed for integrated management of multi-level/multi-theme maps. M² model can be regarded as a forest consisting of two kinds of hierarchies: one is a directory tree, which is obtained by recursively decomposing map regions into a sequence of increasingly finer tessellations with regard to the granularities of administrative units; and another is a theme tree, which is obtained by uniquely dividing spatial objects into different themes. In every theme, map objects are divided into different levels of scales in regard to the display needs of map [19]. The information model is powerful to integrate various scales of maps uniformly and possesses advanced extensibility [17].

	(a) Amalgamation	(b) Selection		
Map data for upper and lower levels	Upper Lower	Upper Lower		
Refined datasets for upper and lower levels	A B C C Upper Lower	Upper Lower		

Figure 6. Refinement

Under M' model, road objects are assigned to multi-level datasets without repetition: the dataset in the lower level is a supplement of the dataset in the upper level. Typical situations of this assignment in two levels are given in Figure 6. In Figure 6(a), the relation between the map objects in upper and lower levels is amalgamation: the map object AD is represented as a line (AD) in the upperlevel map and as lines (AB, BC, CD) in the lower level. An object in the upper level can be regarded as the combination of adjacent objects in the lower level. The information about nodes A and D is the same in both levels but the relation between A and D is different. So, under our model the information in the lower-level map can be split into two parts: one part is the same information as that in the upper level; and another part is the information, which belongs to only the lower level. These two parts are called refined datasets, and are managed in the upper and lower levels, respectively. Considering that the relations between two datasets are needed only in generating maps of the lower level, these relations are managed in the refined dataset of the lower level. In (b), there is a selection relation between maps in two levels: the map object AB is displayed in two levels, but AC is not needed in the upper level. The information in the upper level is a sub-set of that in the lower level.

Therefore, there are an object AB and an object AC in the refined datasets on upper and lower levels, respectively.

4.2 MOR-TREE FOR MULTI-LEVEL ROAD NETWORK

Besides the ability of accessing spatial objects in one level (in one dataset) efficiently just like other spatial index structures (e.g., R-tree [18]), MOR-tree is designed so as to be able to differentiate the levels of objects and manage the relations among objects of multiple levels by adopting two kinds of hierarchies. A main hierarchy is proposed to differentiate the levels of road objects by assigning logical importance values to every object. The logical importance value is a natural number in agreement with the map level: e.g., 1 for objects in country-level datasets, 2 for those in prefecture-level datasets, and so on. The objects with higher importance (with a smaller value) are stored in the higher levels of the hierarchy. Another kind of hierarchy, called a composition hierarchy, is introduced to keep the relations among levels, which are pointed by the leaf nodes of the main hierarchy. The main hierarchy is based on R-tree, to achieve the outstanding spatial access performance. Each node in MOR-tree contains a number of entries. There are three kinds of entries: tree-entries, object-entries and composition-entries. The internal nodes may contain the first two kinds of entries, in contrast to Rtree. The leaf nodes contain object-entries. The objectentry points to a composition hierarchy, which consists of composition-entries.

Three kinds of entries have the following forms:

- Object-entry has the form (MBR, flag, comp-ptr), where MBR is the minimal bounding rectangle of composition hierarchy; flag is a natural number that indicates the importance level; and comp-ptr contains a reference to a composition hierarchy;
- Tree-entry has the form (MBR, flag, child-ptr), where child-ptr contains a reference to a subtree. In this case MBR is the minimal bounding rectangle of the whole subtree and flag is the smallest importance value of the child-nodes.
- 3. Composition-entry has the form (comp-id, n-ptr, nl-ptr), where comp-id is the identifier of object's composition; n-ptr contains a reference to the next composition of the parent node object: e.g., the object is a road segment, its first composition is one of its end points, n-ptr points to the next point on the same object; and nl-ptr contains a reference to the composition of the parent node object in the lower level: e.g., the intersection between the upper-level road and lower-level road.

Here, we give object-entry and composition-entry in Figure 7 as an example. In Figure 7(a), there is a link object al with compositions nal and na2 in level 1, so the entry in the index structure for al is an object-entry, which consists of MBR for al (For simplicity, we use al to replace MBR in this figure), flag of level 1 (here it is 1)

and comp-ptr which points to one composition of al (here it is nal). The compositions of al (nal, na2) are managed by composition-entries. n-ptr of entry for nal points to the entry for na2, and the other pointers of two composition-entries point to NULL. A more complex example is given in Figure 7(b), in which there are lower level compositions (nbl and nb3) of al. The objects bl and b2 belong to level 2. So, nl-ptr of nal points to the composition-entry nbl of level 2 and n-ptr of nbl points to nb3.

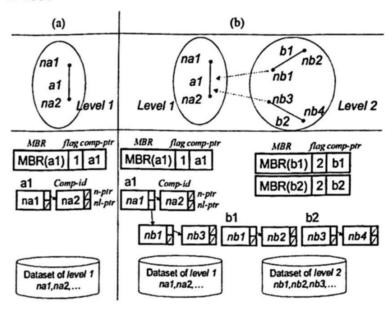


Figure 7. Example of MOR-tree for tw o-levels of road network

4.3 INTEGRATED REPRESENTATION OF MULTI-LEVEL TRANSPORTATION NETWORK

Basically, the transportation information is managed on every node of the map. However, the node, which connects to road segments in different levels, should be multi-level specially. the processed Consider transportation network given in Figure 8: a1, a1 and b1 to b3 are basic road segments in level 1 and level 2, respectively. The arcs (or dotted arcs) represent the transportation arcs in level 1(or level 2). For the simplicity of the figure, the traffic constraints are not drawn in this figure, but the constraint is set as U-turn forbidden, here. The MOR-tree for the basic road network in Figure 8 is given in Figure 9. The information about nodes is managed in level 1 and level 2 according to the level of basic map. The road segments between nodes of different levels can be recorded in the pointer of nl-ptr in the composition hierarchy. However, such pointers cannot assure that there are transportation arcs between these nodes: e.g., nal points to nbl and nb2, but there is no transportation arc from nal to nb2 (there is only an arc from nb2 to nal as b3 is a one-way road). Consider that the path search in level 1 is based on the arcs in level 1 and the search in level 2 is based on a different arc set.

For the search in *level 1*, *Cost-arcs* of node *a1* are:

$$\begin{bmatrix} out_{a1}(na_2,\cos t_1) \\ out_{a2}(na_3,\cos t_2) \end{bmatrix}.$$

While for that in level 2 the Cost-arcs of node al are:

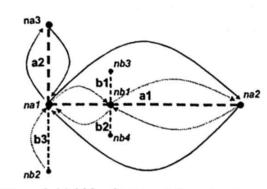


Figure 8. Multi-level transportation network

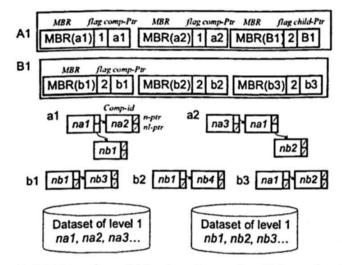


Figure 9. MOR-tree for multi-level road network: Ai: internal nodes of main hierarchy; Bi: leaf nodes of main hierarchy; a i, bi: composition hierarchy; nai, nbi: Composition-entries; i (1 or 2): flag of Object-entries or Tree-entries

$$\begin{bmatrix} out_{a11}(na_2,\cos t_{11}) \\ out_{a2}(na_3,\cos t_2) \end{bmatrix}$$

The Constraint-matrix of node al in level 1 and level 2 are:

$$CM(na_1)_1 = in_{a1} \quad \begin{pmatrix} out_{a1} & out_{a2} \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$CM(na_1)_2 = in_{a11} & out_{a2} \\ in_{a2} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$in_{b2} & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Because 1) the information in the lower level (level 2) is not to be used by the operation in level 1; and 2) the original MOR-tree is not consistent with the transportation situation, we split one node in the upper level (level 1) into two nodes belonging to two levels: one is the original node in the upper level, and another is a transportation-node in the lower level, which contains transportation information of the lower level. In MOR-tree for road network, nl-ptr contains a reference to the composition of the parent node object in the lower level. Here, we redefine nl-ptr as a pointer, which contains a reference to a transportation-node of the parent node object in the lower level. The new tree is called as

TMOR-tree, and its composition hierarchy is given in Figure 10.

The nodes na1, na2 in level 1 possess the transportation information referring to level 2, so new transportation-nodes for them are created in level 2. For the composition hierarchies of level 1, nl-ptr's refer to the corresponding transportation-nodes in level 2. And, the search operations referring to only the upper level can be done just like by using the original MOR-tree, and the operations referring to the two levels can be done by using the information managed in all the levels. The nodes, which connect only with transportation arcs on the same level, can be represented with a super-node directly.

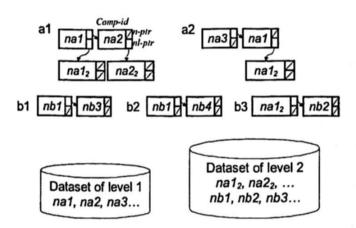


Figure 10. Composition nodes of TMOR-tree for transportation information

4.4 SPATIAL SEARCH AND PATH SEARCH

TMOR-tree supports not only the spatial search on multi-level maps but also the path search on multi-level transportation information. Spatial search to multi-level maps can be realized by accessing objects until a specific level via TMOR-tree. Since TMOR-tree takes the advantages of spatial index structure, spatial queries, such as region query, can be realized by accessing to the internal nodes and composition hierarchies until a specific level of TMOR-tree. The zoom-in/out operations can be realized by accessing the nodes under a specific internal node of TMOR-tree.

Path search on the transportation information on different levels can be realized by using an extended "ink-blot" search method. The ordinary "ink-blot" search method solves a path search from vi to vt likes this: it begins from expanding vi's connecting nodes in the sequence of the cost on the Cost-arc; if the target node has not been expanded, the search goes on by expanding the nodes connected by those expanded nodes. For the search based on the multi-level transportation network, we extend the method. For example, a path search from city hall A to city hall B in separated prefectures can be realized by searching based on transportation networks on different level of details. The search can be done in three steps:

1. At first, the search is based on the city-level transportation network inside city A using "ink-blot"

- search method. This step is terminated when a node belonged to the upper level is expanded.
- Then, the search goes on in the upper level until a node inside the region city B is expanded.
- At last, the search returns to city level and expands nodes until the target is found.

The correctness of this algorithm is assured by the "inkblot" method. Here, we leave the proof out. By using TMOR-tree, the changeover points between multi-levels are those nodes on the road network, which possess original node and transportation-node on multiple levels.

5 ANALYSIS

In this section, we introduce our prototype system and compare our *super-node* representation method with the node-arc methods used by [6, 7].

Our prototype system was developed in Java on an SGI O2 R5000 SC 180 entry-level desktop workstation. The system manages basic road maps in a part of Aichi Prefecture, Japan. The map is divided into map pages with the same size of $2000m \times 1500m$ (which is defined by Geographical Survey Institute of Japan [20]. As it has no effectiveness on the proposed methods in this paper, we ignore this concept in the followings). The total number of nodes *Nnum* is 42,062 and the number of links *Lnum* is 60,349 in the basic road map. These road segments are assigned to the country level (country-wide highways and national roads), the prefecture level (prefectural roads and main local roads) and the city level (city roads).

In city level transportation network, the average traffic arcs connecting to a node is about 2.87 (=2 Lnum / Nnum). When there is no traffic constraint for the basic road map, in node-arc method [7], there are 120,798 records (two times of the link numbers in road maps). As in our super-node method, the amount of information is related to the number of arcs of every node: here, the nodes with four, three, two and one out-arcs are about 24: 51: 13: 12. The total arcs managed in city level by SN method are 120,798. When there are traffic constraints, Right-turn and U-turn are forbidden in about half of the cross and T-junction points. Then, in NL method there are about 142,423 nodes and 135,293 arcs; while in SN method the amount of information keeps the same whether there are constraints (turn costs) or not. The number of arcs and nodes in city level (country and prefecture level) managed by super-node method (denoted as SN) and those by node-arc method (denoted as NA) are given in Figure 11 (Figure 12). In the figures, there are different values for different conditions of datasets in NA method. "Constraint" means there are traffic constraints and "Turn" means there are turn costs in the dataset. Because the number of nodes and arcs keep the same without depending on there are constraints (turn costs) or not, the datasets generated by SN method shares

the same value in this figure, which is denoted simply as SN method.

We have done experiments of finding nearest target object based on these datasets. The test results are given in Figure 13. In the figure, x-axis represents the density of targets on the road network, which is the ratio of the targets' number (Tnum) to the nodes' number (Nnum) in basic road map; and y-axis represents the CPU time for the search. D represents that adopting Dijkstra's algorithm does the search; and R represents the extended "ink-blot" search method proposed in this paper. D algorithm is executed based on the datasets generated by NA method with/without constraint (the number of objects is depicted in Figure 11); and R algorithm is done on the dataset generated by SN method with/without constraint. In Figure 15, we can observe that our SN R methods outperform those of NA D at any situations: especially, when there are traffic constraints on the networks.

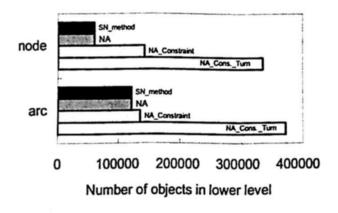


Figure 11. Numbers of arcs and nodes in city level managed by supernode and node-arc methods

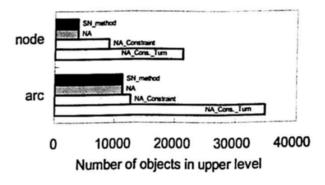


Figure 12. Numbers of arcs and nodes in country and prefecture level managed by super-node and node-arc methods

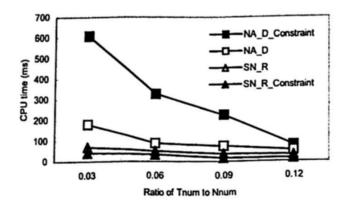


Figure 13. CPU time of finding nearest target object

6. CONCLUSIONS

In this paper, we proposed a representation method for multi-level transportation networks. Our method adopts a spatial index structure for managing map objects in multiple levels and uses an integrated method for representing the travel junctions (or traffic constraints), travel cost on road segments and turn corners. Based on the datasets generated by this method, queries in ITS applications can be responded efficiently in different levels of details. In our future work, the performance of the creation, modification and processing of the datasets created by our method will be evaluated, deeply.

REFERENCES

- [1] S. Shekhar and D.R. Liu, CCAM: A connectivity-clustered access method for aggregate queries on transportation networks: a summary of results, *Proc. of ICDE'95*, 1995, 410-419.
- [2] B. Becker, H.W. Six and P. Widmayer, Spatial priority search: an access technique for scaleless maps, *Proc. of ACM SIGMOD'91*, 1991, 128-137.
- [3] P.V. Oosterom, The Reactive-tree: a storage structure for a seamless geographic database, *Proc. of Auto-Carto*, Vol. 10, 1991, 393-407.
- [4] J. Feng and T. Watanabe, MOR-tree: An access structure for multi-levels of road networks on distributed environment, *Journal of the ISCIE*, *Vol.16*, *No.12*, 2003, (to appear).
- [5] M. F. Goodchild, GIS and transportation: status and challenges, *GeoInformatica*, *Vol.4*, *No.2*, 2000, 127-139.
- [6] S. Winter, Modelling costs of turns in route planning, GeoInformatica, No. 4, 2002, 345-361.
- [7] J. Fawcett and P. Robinson, Adaptive routing for road traffic, *IEEE Computer Graphics and Applications*, Vol. 20, No.3, 2000, 46-53.

- [8] D. Papadias, J. Zhang, N. Mamoulis and Y.F. Tao, Query processing in spatial network databases, *Proc. of VLDB 2003*, 2003, 802-813.
- [9] N. Christofides, graph theory: an algorithmic approach, Academic Press Inc.(London) Ltd. 1975.
- [10] N. Jing, Y.W. Huang and E.A. Rundenstteiner, Hierarchical encoded path views for path query processing: an optimal model and its performance evaluation, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 10, No. 3, 1998, 409-431.
- [11] Y.W. Huang, N. Jing and E.A. Rundensteiner, effective graph clustering for path queries in digital map, *Proc. of CIKM'96*, 1996, 215-222.
- [12] Y.W. Huang, N. Jing and E.A. Rundensteiner, Path queries for transportation networks: dynamic recordering and sliding window paging techniques, *Proc.* of GIS'96, 1996, 9-16.
- [13] E.P.F. Chan and N. Zhang, Finding shortest paths in large network systems, *Proc. of GIS'01*, 2001, 160-166.
- [14] J. G. Stell, Granulation for graphs, *Lecture Notes* in Computer Science, Springer-Verlag, No. 1661, 1999, 417-432.
- [15] Y. Leung, K.S. Leung and J.Z. He, A generic concept-based object-oriented geographical information system, *Int'l J. of Geographical Information Science*, Vol.13, No. 5, 1999, 475-498.
- [16] S. Timpf, Hierarchical structures in map series, 1998, on-line: http://www.geoinfo.tuwien.ac.at
- [17] J. Feng and T. Watanabe, Effective representation of road network on concept of object orientation, *Trans. IEE of Japan, Vol. 122-C, No. 12,2002*, 2100-2108.
- [18] A. Guttman, R-Trees: A dynamic index structure for spatial searching, *Proc. of ACM SIGMOD'84*, 1984, 47-57.
- [19] T. Yamanashi and T. Watanabe, Multi-layers/Multi-phases model adaptable to integrated GIS, *Proc. of VSMM'00*, 2000, 668-676.

[20] http://www.gsi.go.jp/MAP/CDROM2500/t2500.htm

SKETCH SPATIAL QUERY REPRESENTATION MODEL

Lau Bee THENG and Wang Yin CHAI Faculty of Information Technology University of Malaysia Sarawak Malaysia

ABSTRACT

Sketch is the most natural way of expressing a person's spatial query. Sketch-based spatial query requires essential modeling on sketch query to retrieve meaningful spatial configurations from the spatial databases. This research has proposed a spatial query model that retrieves spatial objects from spatial database by using sketch as query. This paper discusses on the conceptual modeling of spatial query representation.

KEY WORDS

Sketch, spatial databases, spatial query and spatial retrieval.

1. INTRODUCTION

In general, sketch consists of sketched objects with ambiguous shapes, scale and direction. The relations between all objects in a sketch are called object relations. Objects' relations are considered as one of the most important source of information for spatial query where object's geometry is considered only for metric refinements as in [1] and [7][8]. In their works, sketch is modeled into an association graph made up of sketched objects' relations alone. The retrieval is based on the matching of object relations where a sketch is translated into a reduced association graph for relative directions, metric refinements and topological relations. The reduced association graph is very complex to compute and requires long processing time even for matching a simple sketch with a database consist of 100 files with less than 10 objects in individual data record. Furthermore, the existing model is lacking of ability to support the real world spatial databases. Consequently an improved model to interpret and formulate spatial query for a sketch is essential.

There are a few areas of improvement for spatial query by sketch for real world spatial database system such as GIS [1]. The areas of improvements are stated below.

- How to produce a more efficient digital sketch representation for spatial query?
- 2. How to index spatial objects and relations in a spatial database for spatial query by sketch?

3. How to efficiently query a set of large continuous spatial database by sketch query?

The above three areas of improvement are indeed interconnected. A model that can meaningfully represent a digital sketch with most relevant details using less computations is required, hence the digital sketch representation can be used to perform indexing on the fly for spatial objects and relations in the spatial database, finally the query of large continuous spatial database can be conducted efficiently.

This research focuses only on the above areas to produce a sketch spatial query representation model that is able to extract and represent a sketch effectively and can also help to improve the retrieval from large continuous spatial database. As a sketch provides meaningful information such as relative size, relative distance, shape, relative direction and topology, the proposed model produces a spiral web structure that is able to represent spatial objects and relations in a sketch taking into account all the abovementioned information. By using the proposed model, a sketch is able to query the large continuous database object by assessing the similarity through a spiral web with the spatial objects from database. Hence the proposed model can support the retrieval from a set of large continuous spatial database such as a street map.

2. THE MODEL

This section discusses on the strengths of the proposed model namely the model over other existing model.

First, existing model uses MBR or Tilted MBR to approximate the sketch objects, hence objects A, B, C and D are treated as similar because they have exactly similar bounding rectangles though B, C and D have been rotated from A. The proposed model is able to preserve the originality of the sketch and represent its objects under various rotations differently.

Secondly, the model is sensitive to relative sizes that exist among objects in the sketch. Hence objects in Sketch A are represented with different values then objects in Sketch B. This is crucial as large continuous databases have plenty of possible matches, the sensitivity to scale can filter out the dissimilarities of objects.

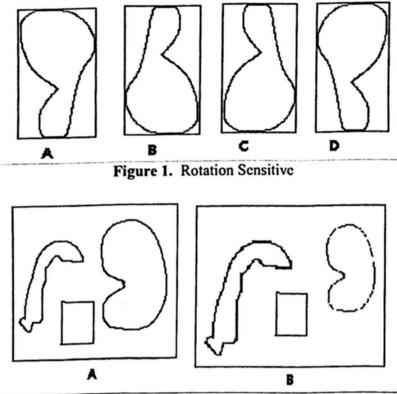


Figure 2. Relative Size Sensitive

Thirdly, the model is sensitive to neighborhood relations that exist between sketched objects within a sketch. It is able to preserve them under any spatial operations like rotation, scaling and movement. For example, C is contained by D, E touches F, G overlaps H, A is covered by B, A disjoin G, D disjoin F etc.

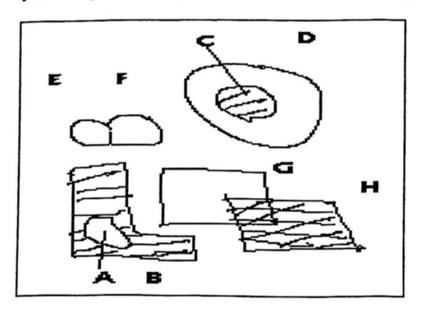


Figure 3. Object Neighborhood Sensitive

Currently, the existing model estimates the reduced object relations by ([n*(n-1)]/2) which still consists of redundancies though it is better than complete association relation (n*n). Fourthly, the model simplifies the number of spatial relations that need to be assessed to further reduce redundancy by using the formula: [2n - 3]. In Figure 5, the object relations computation is [8*(8-1)]/2 = 28 whereby, it is [(2*8)-3]=13 only using the proposed model for the eight objects in the sketch. The algorithm for computing the relation is shown below. Figure 4 shows the number of relations using the above algorithm where there are 5 relations for the 4 objects.

```
FOR i = 1 to TotalObjectCount

FOR j = 2 to TotalObjectCount

IF i ⇔ j THEN

BuildRelation (Object(i), Object(j)

END IF

NEXT

NEXT
```

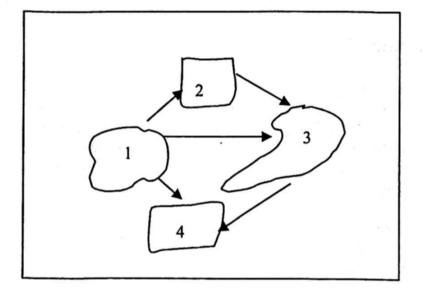


Figure 4. Reduced Relations for 4 Objects

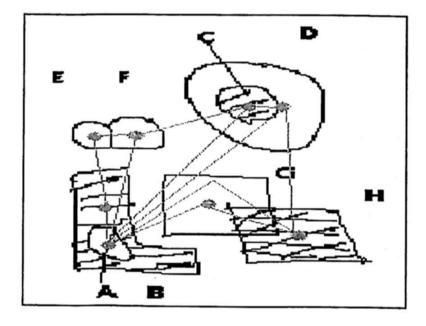


Figure 5. Reduced Object Relations

Basically, the spatial query by sketch model uses approximation to represent the sketch. However the shapes of each object bear a significant meaning when user sketches it. Fifthly, the model represents a sketch with its original object shapes without approximating them using any bounding rectangle or point. The sample in Figure 6 shows a representation of sketch A in Spiral Web B.

Sixthly, the model is inferential so that it can be used to determine the loose or unknown sketch properties from the known representation. For instance, the relative distance among the sketch objects can be viewed and estimated by looking into the ring and zone in the spiral web shown in Figure 6. Furthermore the relative directions among sketch objects can be easily recognized

from the location on the spiral web. Other property like relative size can be easily inferred from the objects in the spiral web as well. This inferential attribute has not existed in the existing spatial query by sketch model.

Currently, there is no formal model being developed specifically to meet the requirements of sketch-based spatial query and retrieval processing. Seventhly, the model provides a formal basis of sketch representation for query and retrieval processing.

Lastly, the model provides a sketch representation that each object is quantifiable and distinguishable from one to another. Hence it also provides quantitative assessment to similarity between sketch and objects in the spatial database for spatial objects retrieval.

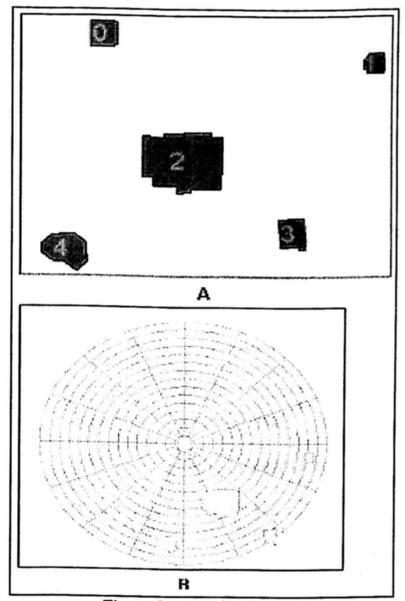


Figure 6. Approximation Free

3. HOW DOES THE MODEL WORK

The graphical illustration on how the model work is shown in Figure 21. Firstly, when a sketch is received, then representation process starts working to create a spiral web for the sketch. During the creation of spiral web, the values of all sketched objects also computed. Then the spiral web is used for retrieval from spatial database that consists of more than a thematic layer. The

created spiral web is used as an index on the fly for the spatial database. Then systematic search is applied to retrieve all the selected objects based on the index. After the systematic search with the spiral web index, each set of selected objects is stored as individual spatial scene. Each retrieved spatial scene is assessed for similarity values. For instance in Figure 21, one of the retrieved spatial scene produces the similarity of "1" as compared to each object in the sketch.

4. HOW TO REPRESENT SKETCH

After the brief discussions about the strengths of the model and how it works, this section describes how the query representation in the model works.

A sketch is extracted and represented as a spiral web structure that is made up of rings, zones and cells. The number of cells is a product of numbers of zone and ring They are determined by factors like relative distance of reference object to the furthest object, diameter of the reference object and relative direction in a sketch's orientation. For example, in Figure 9, there are 4 rings and 64 cells. The number of cell is not fixed as it grows or shrinks according to the abovementioned factors. The number of zone represents the relative directions. Figure 6 illustrates a sketch consist of 5 objects is transformed into the designed spiral web. There are 15 rings, 16 zones and 240 cells in the spiral web. All the sketched objects are also created as part of the spiral web. The representation of the sketched objects on the spiral web is determined by which ring and zone it falls on.

There are different computations for determining various positions of sketched objects on the spiral web as discussed in the following sections. The input for query representation is a digital sketch stored as a spatial data file that consists of sketched objects.

A. SPIRAL WEB STRUCTURE

The spiral web is the main component of query representation model. A spiral web is actually a spatial data structure built as a result of processing a sketch in which it contains the data from the sketch and also a set of cells. Spiral web is defined as a collection of zones and rings that is built to relate all sketched objects in a sketch. Each sketch has a spiral web that is tailored made for it. The number of zones and rings depend solely on the number of sketched objects in a sketch. A reference object is chosen to be the center of each spiral web. The reference object is the first sketched object drawn by user. The spiral web is highly sensitive to the reference object.

A spiral web is stored as a spatial object with these attributes: {SpiralWebObjectID, SpiralWebObjectShape}. SpiralWebObjectID is the index of spiral web; SpiralWebObjectShape is the geometry type of a spiral

web object in generic format that is polygon or line. The spiral web is in generic format that spatial data file like ESRI Shapefiles can provide. This gives the convenience of applying this model into any generic spatial data management system like GIS. With this simple spiral web structure, it is able to simplify the processing task in spatial query and retrieval by sketch.

B. ZONE COMPUTATION

People manipulate concrete relations rather continuous quantitative direction i.e. angles to express and reason about directions. Most previous work defines qualitative directions using either object projections or centroids. Each approach has its advantages and shortcomings. In this research, a centroid-based method is applied where the direction between two objects is determined by the angle between their centroids. The set of relative direction relations used is shown in Figure 8. These directions are used to represent the zones in the spiral web as shown in Figure 9. The 16 zones represent the 16 qualitative direction model. There is no fixed numbers of zone for a spiral web but it must follow the direction model such as 4 zones (north, east, south and west), 8 zones (north, northeast, east, south, southeast, northwest, southwest and west), 16 zones (north, northeast, north northeast, east northeast, south, southeast, south southeast, east southeast, east, northwest, north northwest, west northwest, south southwest, southwest, west southwest and west) or more by 4ⁿ. The derivation of zones depends on the number of objects in a sketch, the relative distances of each object pair formulated as shown in the algorithm in Figure 7. The sufficient condition is determined by the number of objects fall in each zone, if there is more than one object in a zone, then more zones need to be drawn. There are 8 zones in Figure 9 and there are 4 zones in Figure 9. The rationale behind these various zones is the less zone, the less details can be provided; the more zones, the higher the processing power is needed. However if there are sixteen zones and only three objects exist in the sketch, then the number of zones may be simplified accordingly such as reduces it to 8 or 4 zones only. For instance, in

Figure 10, there are four sketched objects in the sketch, therefore the spiral web only consists of 8 zones as it is sufficient to describe the sketch where Object B falls in Zone 8, Object C falls in Zone 1 and Object D falls in Zone 2 and 3 and Object E falls in Zone 4.

Algorithm to Determine Number of Zones

GET Sketch
FOR each object to total number of objects in sketch
Check location
Build 4-Zone
Estimate Value of object
NEXT object

IF 4-Zone is not sufficient to describe the objects THEN
FOR each object in sketch
Check location
Build 8-Zone
Estimate Value of object
NEXT object

IF 8-Zone is not sufficient for the objects THEN
FOR each object in sketch
Check location
Build 16-Zone
Estimate Value of object
NEXT object

END IF

END IF

Figure 7. An Algorithm to Determine Number of Zones for A Spiral Web

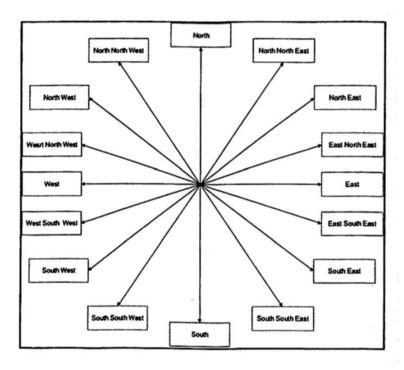


Figure 8. Directional Zones

C. RING COMPUTATION

In this research, a ring is defined as a circular ring derived specifically for the spiral web. Figure 11 shows the ring structure used for spiral web in which the rings are separated to illustrate the ring. In fact, they are joined together to make up the spiral web as shown in Figure 9. The computation of ring depends on the total distance of the reference object from the furthest object in the spatial scene. There is a formula derived specifically to build the spiral web as shown in Equation 1. In Figure 10, there are 5 objects where the furthest object is Object E, therefore the spiral is built with 4 rings based on the diameter of Object A that is the reference object in the sketch. Whereby in Figure 10, there are 5 objects in a sketch, and there are 3 rings created for the spiral web as the furthest object is Object E that falls into the 3rd ring.

L = EucDist / Diam

where:

L: Number of ring.

EucDist: Euclidean distance of reference object to the

furthest object in the sketch.

Diam: Diameter of the reference object.

Equation 1. Estimate number of rings in a spiral web

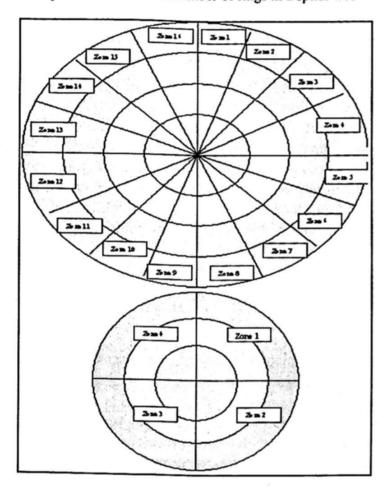


Figure 9. (a) A Sample of Spiral Web Structure Consists of 16 Zones and 4 Rings (b) Spiral Web Structure Consists of 4 Zones and 3 Rings

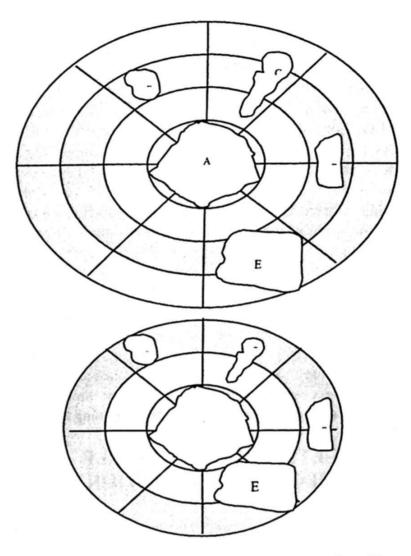


Figure 10. (a) A Sample of Spiral Web Structure Consists of 8 Zones and 3 Rings for A Sketch (b) A Sample of Positioning of Sketch Objects on the Spiral web Structure

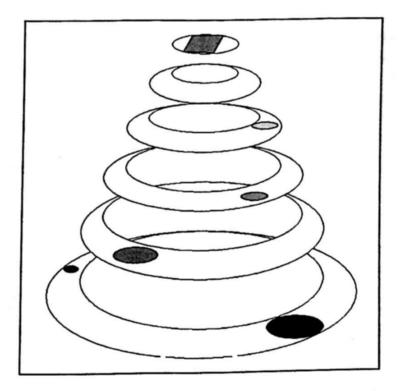


Figure 11. A Sample of Ring Structure Consists of 6 Rings for A Sketch with 6 Objects

D. SKETCHED OBJECT POSITIONING

Once a sketched object is being placed on the spiral web, it has a possibility to fall on single cell or multi cells. For cases of multi cells, there are Multi Zones Single Ring, Multi Zones Multi Rings, and Single Zone Multi Rings. As for single cell, there is Single Zones Single Ring. As for various positioning, the estimation of sketch object value also varies. In Figure 9, there are 16 zones and 4 rings. Object A is the reference object which is the center of the spiral web, the diameter of each ring is determined by Object A's diameter. Object B is a sample of Single Zones Single Ring object. Whereby, Object C is Single Zone Multi Rings object. Object D is Multi Zones Single Ring object. Lastly, Object E is Multi Zones Multi Rings. Due to this different sketch object positioning on the spiral web, there are 4 different computations used to estimate the object values of the sketched object on the spiral web. The detailed computation of the objects in these 4 categories is discussed in the following sections.

MULTI ZONES SINGLE RING OBJECT REPRESENTATION

The multi zones single ring is a category given to describe object that falls into more than one zone but only one ring in the spiral web. Object D shows the multi zones single ring sketched object representation where Object D is in the zone 4 and 5 but in ring 3 of the spiral web. Figure 9 shows zone categorization and Figure 11 shows ring categorization. Figure 10 shows different positioning of sketched object on the spiral web. By using the estimation in this model, the object value of Object D that has 0.4 of its area falls in Zone 4 and 0.6 area falls in Zone 5 and all area fall in Ring 3.

MULTI ZONES MULTI RINGS OBJECT REPRESENTATION

The multi zones multi rings is a category given to describe object that falls into more than one zone but only one ring in the spiral web. Object E shows the multi zones multi ring sketched object representation where Object E is in the zone 7, 8 and 9 but in ring 2, 3 and 4 of the spiral web. Refer to Figure 9 for the zone and ring categorization is shown in Figure 11.

Figure 10 shows different positioning of sketched object on the spiral web. Object E has 0.3 of area falls in Zone 10, 0.5 of area falls in Zone 9 and 0.2 of area falls in Zone 8. 30% of area of Object E falls in Ring 2, 45% area falls in Ring 3 and 0.25 of area falls in Ring 4.

SINGLE ZONE SINGLE RING OBJECT REPRESENTATION

The single zone single ring is a category given to describe object that falls single zone and single ring in the spiral web. Object B shows the single zone single ring sketched object representation where Object B is in zone 15 and ring 3 of the spiral web. Refer to Figure 9 for the zone and ring categorization is shown in Figure 11.

Figure 10 shows different positioning of sketched object on the spiral web. Object B has 100% of area falls in Zone 15, 100% area falls in Ring 3 alone.

SINGLE ZONE MULTI RINGS OBJECT REPRESENTATION

The single zone multi rings is a category given to describe object that falls single zone and multi rings in the spiral web. Object C shows the single zone single ring sketched object representation where Object C is in zone 15 and in ring 2, 3 and 4 of the spiral web. Refer to Figure 9 for the zone and ring categorization and refer to Figure 10 for the different positioning of sketched object on the spiral web. Object C shows the single zones multi rings sketched object representation.

5. HOW TO FORMULATE SPATIAL QUERY

After the completion of the query representation stage, the input to the spatial query formulation stage is a spatial data file containing a spiral web. The spatial query formulation is formed based on spiral web created in the Section 4. In short, for each processed sketch, there are more than a spatial query is formulated. The number of spatial query formulated is determined by the number of zones exists in the spiral web created for a sketch. If there are 4, then four spatial queries shall be formulated. The algorithm to formulate the spatial query is extracted and shown below.

Spatial Query Formulation

NEXT

LET PI = 3.142 LET a = Zone FOR angle = PI /Zone to (2*PI) STEP PI / a FORMULATE Query STORE Query

Figure 12. An Algorithm for Spatial Query Formulation

For instance, for the sketch shown in Figure 6, the spatial query formulation produces a list of spatial queries in spatial data file format as shown from Figure 13 to Figure 20 Query. Each spiral web represents a spatial query for the sketch. Query 2 is formulated when the original sketched objects are rotated for 45 degrees its original position in the spiral web.

As a result, Query 3 is a resulted query by moving the original sketched objects by 90 degrees from the original position. Consequently, Query 4, Query 5, Query 6, Query 7, and Query 8 are formulated by an increment of 45 degrees from the previous query. The retrieval of spatial objects for a sketch is done based on the queries made in the form of spiral web.

5. CONCLUSION

Though sketch as a query to spatial database management systems like GIS suffers plenty of deficiency such as in deficiency in relative size, shape, relative direction and relative distance similarity assessment, we are not supposed to ignore all these sketch behavior in processing spatial query. In fact, neighborhood relations and objects' geometry are equally important in sketching spatial query.

To model a good sketch representation model, this research has proposed a spiral web to represent a sketch's configuration setting. This paper details the conceptual modeling of the proposed model. It explains the model that can utilize the unique behavior of sketch in formulating a good spatial query representation model. It explores a new way of modeling sketch as spatial query in a real world spatial database system. With the spiral web representation, it shall be able to enhance the spatial retrieval by sketch. Currently, testing and evaluation of the model is underway. Hence the results discussions shall be published in the next paper.

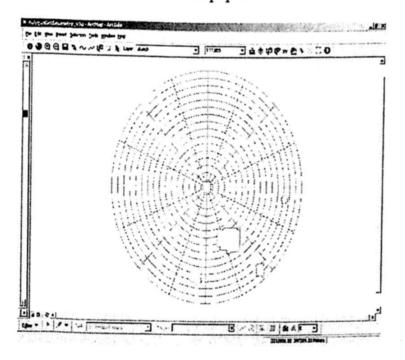


Figure 13. Query 1

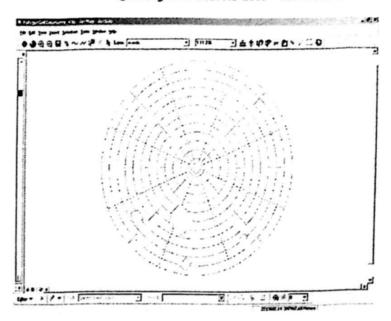


Figure 14. Query 2

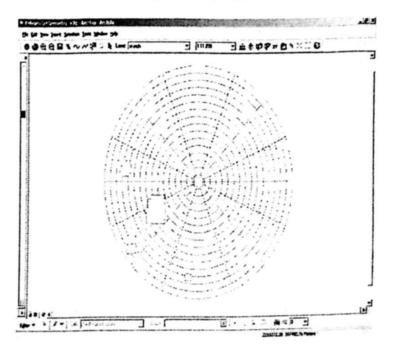


Figure 15. Query 3

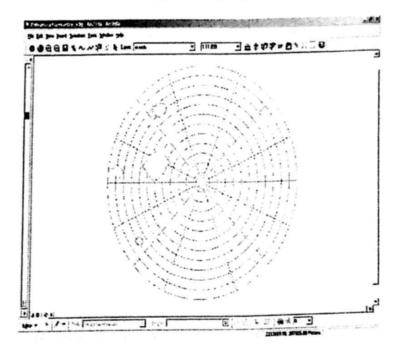


Figure 16. Query 4

© Theng et al GEOPRO 2003 - ISBN: 970-36-0103-0

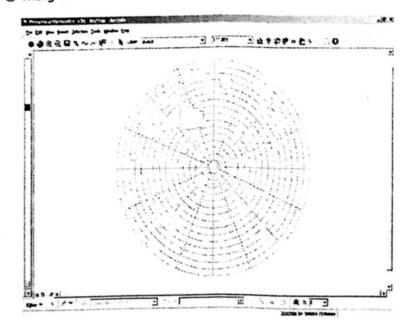


Figure 17. Query 5

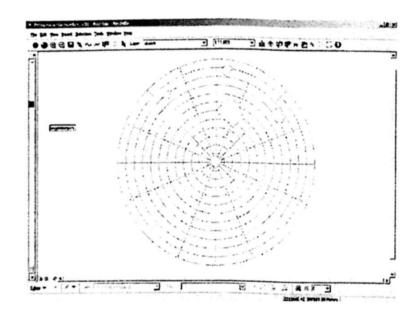


Figure 18. Query 6

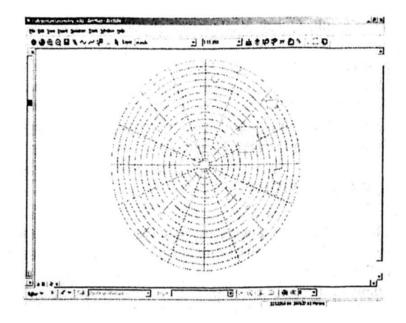


Figure 19. Query 7

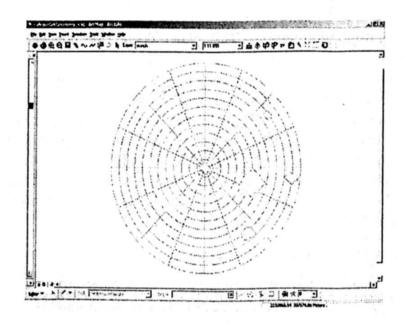
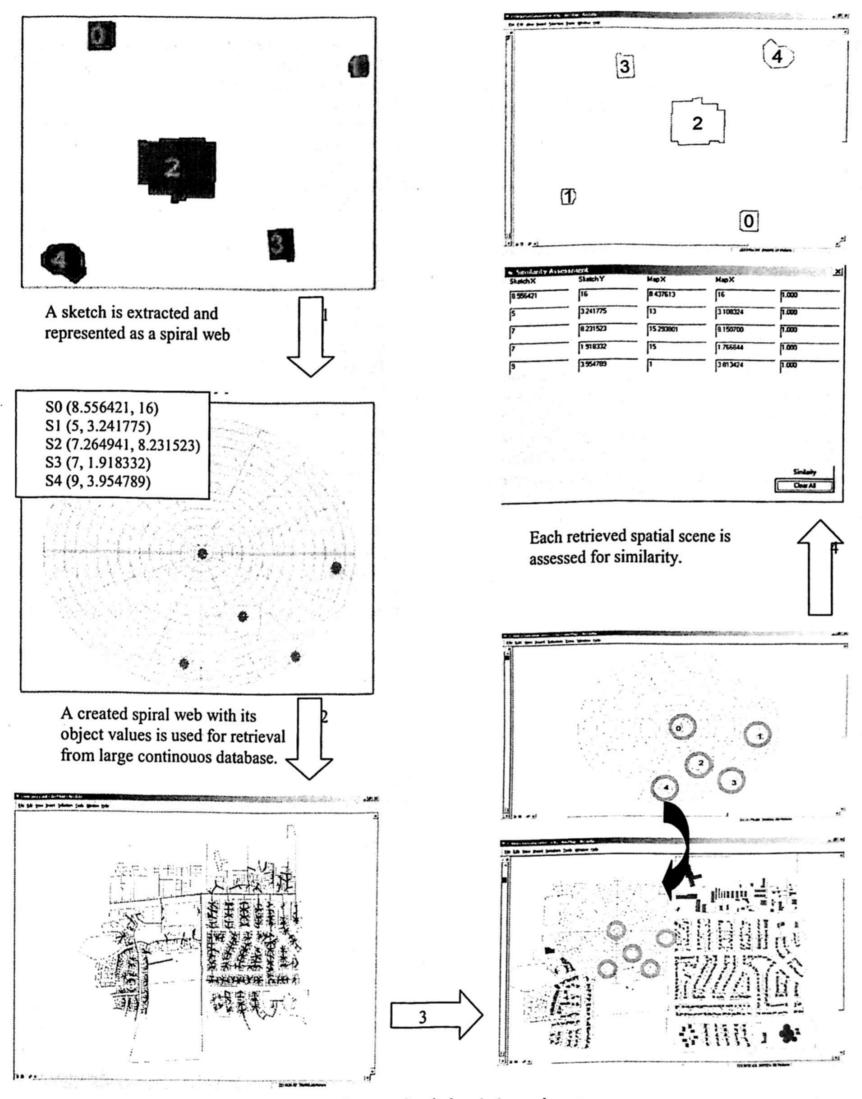


Figure 20. Query 8



A created spiral web is used as an index on the fly for spatial database. Then systematic search is applied to retrieve all the selected objects.

Figure 21. How Does the Model Work

ACKNOWLEDGEMENT

A bundle of thanks to my advisor, Assoc. Prof Dr. Wang Yin Chai for his passion and patience throughout the research.

REFERENCES

- [1] Blaser, A. (2000). Sketching Spatial Query. Dissertation of PhD. University of Maine.
- [2] Blaser, A. and Egenhofer, M. (2000). Visual Tool for Querying Geographic Databases AVI2000 – Advanced Visual Databases, Salerno, Italy in V. Di Gesù, S. Levialdi, and L. Tarantini (eds.), pp. 211-216, May 2000.
- [3] Blaser, A., Sester, M., and Egenhofer, M. (2000). Visualization in an Early Stage of the Problem Solving Processing GIS Computers and Geosciences, 26 (1): 57-66.
- [4] Egenhofer, M. (1991)). Reasoning about Binary Topological Relations. In Gunther, O. an Schek, H.J. (eds). Advances in Spatial Databases, Springer Verlag LNCS.
- [5] Egenhofer, M. (1994). Pre-Processing Queries with Spatial Constraints. Photogrammetric Engineering & Remote Sensing 60 (6): 783-790.
- [6] Egenhofer, M. (1994). Spatial SQL: A Query and Presentation Language. IEEE Transactions on Knowledge and Data Engineering 6 (1): 86-95.
- [7] Egenhofer, M. (1996). Spatial-Query-By-Sketch. VL'96: IEEE Symposium on Visual Languages, Boulder, CO in M. Burnett and W. Citrin (eds.), pp. 60-67, September 1996.

- [8] Egenhofer, M.(1997). Query Processing in Spatial-Query-By-Sketch. Journal of Visual Languages and Computing 8 (4): 403-424.
- [9] Papadias, D. (2000). Hill Climbing Algorithms for Content Based Retrieval. Proceedings of the ACM Conference on Information Retrieval (SIGIR), Athens, July 24-28, pp. 240-247.
- [10] Papadias, D. and Mamoulis, N.(1999). Improving Search Using Indexing: A Study with Temporal CSPs. Proceedings of International Joint Conference of Artificial Intelligence, Stockholm, Sweden, 436-441.
- [11] Papadias, D. Kalnis, P, Mamoulis, N. (1999). Hierarchical Constraint Satisfaction in Spatial Databases. Proceedings of AAAI Conference, Orlando, FL, 142-147, AAAI/MIT Press.
- [12] Papadias, D. Mamoulis, N., Theodoridis (2001). Contraint-Based Processing of Multiway Spatial Joins. Algorithmica, Vol. 30(2), 188-215.
- [13] Papadias, D., Mantzourogiannis, M., Kalnis, P., Mamoulis, N., Ahmad, I. (1999). Content Based Retrieval using Heuristic Search. Proceedings of the ACM Conference on Information Retrieval (SIGIR), Berkeley, CA, 168-175.
- [14] Papadias, D., Momoulis N., Delis, V. (1998). Algorithms for Querying by Spatial Structure. Proceedings of Very Large Databases Conference (VLDB). New York.
- [15] Papadias, D., Karacapilidis, N. and Arkoumanis, D. (1999c). Processing Fuzzy Spatial Queries: A Configuration Similarity Approach. International Journal of Geographic Information Science Vol. 13(2), 93-128.

ADAPTATION AND USE OF SPATIAL AND NON-SPATIAL DATA MINING

Manuel PECH, David SOL, Jesús GONZÁLEZ
Universidad de las Américas-Puebla (UDLA)
San Andrés Cholula, Puebla, México, MEXICO
{sp205175, sol}@mail.udlap.mx, jagonzalez@inaoep.mx

ABSTRACT

The investigation described in this paper states the analysis and application of spatial and non-spatial data mining technology. Spatial data mining can be defined as the search of patterns that could exist in spatial databases. The test context is a database of the Popocatépetl volcano developed by the laboratory of Technologies of GeoInformation at UDLA-P. Spatial data mining was centered in clustering geometric objects by the implementation of the PAM (Partitioning Around Medoids) algorithm. Non-spatial data mining was made through the use of the SUBDUE system, which searches representative substructures in the data. The obtained results helped us to identified relevant characteristics which need to be improved as part of a contingency plan for the population living around the volcano risk zones (i.e. evacuation roads).

KEY WORDS

GIS, spatial data mining, non-spatial data mining, Open Gis.

1. INTRODUCTION

The Laboratory of Technologies of GeoInformation (Xaltal) [9] of the Universidad de las Americas-Puebla is developing the Popocatépetl Project [5, 11, 15, 17], which main objective is to offer a Geographic Information System (GIS) that incorporates all the information related to the Popocatépetl volcano, and allow a remote way to be accessed by different kinds of users.

A GIS is defined as a tool for geographic data manipulation [2]. It implements a great diversity of functions; some of these are the compilation, verification, storage, recovery, manipulation, updating and presentation of geographic data. One of its more important potentialities is the inclusion of modules for data analysis.

Our research involves the use of the data mining technology, applied to the volcano database, as a tool for knowledge discovery. The general objective is the adaptation, implementation and use of algorithms for spatial and non-spatial data mining to be applied to the Popocatépetl volcano database. Our intention is to provide a tool that allows the analysis, evaluation and optimization of the volcano information, as well as the advice that knowledge discovery can give at the time of decision making.

This work is divided into 7 sections. Section 1 shows the introduction to the investigation. In section 2 we explain what spatial data mining is. In section 3 we describe data mining and the SUBDUE system. In section 4 we introduce the Popocatépetl volcano database and the architecture of the developed system. Section 5 presents the implementation of the system. In section 6 some results obtained are presented and finally in section 7 the conclusions and the future work are given.

2. SPATIAL DATA MINING

Diverse studies on the methods for the spatial database knowledge discovery have been made. For example, Junas Adhikary [1] presents a classification of these methods and divides them in five groups:

- Methods using generalization are based on knowledge discovery and require the implementation of concepts hierarchies. In the case of a spatial database, these hierarchies can be thematic or spatial. A thematic hierarchy can be exemplified when generalizing specific concepts like apples and pears to fruits; and spatial hierarchies when generalizing a group of points in a map as a region and a group of regions as a country.
- A method using pattern recognition can be used to make automatic recognition and categorization of photographs, images, and text, among others.
- Methods using clustering are used to create groupings or data associations, when there is some knowledge similarity among the elements of the group (i.e., similarity by Euclidian distance).
- 4. Methods exploring spatial associations allow discovering spatial association rules, that is, rules that associate one or more spatial objects with another α other spatial objects (X → Y, where X and Y are a set of spatial or non-spatial predicates).
- Methods using approximation and aggregation allow knowledge discovery on the basis of the representative characteristics of the data set.

Spatial Data

Spatial data consists of information that describes spaces. Data is continuously obtained by diverse types of applications such as GIS and computerized cartography. Consequently, data analysis by means of manual techniques is sometimes a complicated task, due the large volume of data. In order to solve this problem, different methods have been proposed and applied to discover knowledge in spatial data. Most of these methods use machine learning techniques, database technology and statistics.

Spatial data mining is defined as the discovery of implicit and previously unknown knowledge in spatial databases [4]. Knowledge discovered from spatial data can be classified into several types, like representative characteristics, structures or clusters and spatial associations, just to mention a few.

Methods for knowledge discovery in spatial data

Geographic data in general has thematic and spatial data [1]. Thematic data is alphanumeric and related to the spatial objects. Spatial data, on the other hand, is described using two different properties: geometry and topology. According to [1], spatial location and size are considered geometric properties, whereas adjacency (the object A is right of object B) and inclusion (the object A is included in object B) are considered topological properties. In this way the methods discovering knowledge can be focused either on the thematic or in the spatial properties of spatial objects of a spatial database or both.

Methods using generalization

One of the most effective methods for discovering knowledge has been the learning from examples technique (with generalization). This method requires concept hierarchies.

Methods using clustering

Cluster analysis is a branch of statistics; the main advantage of using this technique is the feasibility to directly find groups (clusters) in the data without using any background knowledge, similar to an unsupervised learning approach used in machine learning. Diverse algorithms have been developed like PAM (Partitioning Around Medoids) [7], CLARA (Clustering LARge Applications) [7] and CLARANS (Clustering Large Applications based upon RANdomized Search) [12]. Next we describe the PAM algorithm.

PAM

In order to find k clusters (groups), PAM determines a representative object for each cluster. This representative object, called medoid, is the one that is located toward the center within the cluster. Once medoids have been selected, each non-selected object is grouped with the medoid to which is the most similar. More precisely, if O_j is a non-selected object, and O_i is a medoid (selected

object), we say that O_j belongs to the cluster represented by O_i , if $d(O_j, O_i) = min_{oc}d(O_j, O_c)$, where the notation min_{oc} denotes the minimum over all medoids O_c , and the notation $d(O_a, O_b)$ denotes the dissimilarity or distance between O_a and O_b objects. All the dissimilarity values are given as inputs to PAM.

To find the k medoids, PAM begins with an arbitrary selection of k objects. Then in each step, a swap between a selected object O_i and a non-selected object O_h is made, as long as such a swap would result in an improvement of the quality of the clustering. In particular, to calculate the effect of such a swap between O_i and O_h , PAM computes cost C_{jih} cost for all non-selected objects O_j . Depending on which of the following cases O_j is in, C_{jih} is defined by one of the following equations:

Case 1

Supposing that O_j currently belongs to the cluster represented by O_i , Furthermore, letting O_j be more similar to $O_{j,2}$ than to O_h , $(d(O_j, O_h) \ge d(O_j, O_{j,2})$, where $O_{j,2}$ is the second most similar medoid to O_j . Thus, if O_i is replaced by O_h as a medoid, O_j would belong to the cluster represented by $O_{j,2}$. Hence the cost of the swap is given by $C_{jih} = d(O_j, O_{j,2}) - d(O_j, O_i)$. This equation always gives a non-negative C_{jih} value, indicating that there is non-negative cost incurred in replacing O_i by O_h . An example is shown in Figure 2.1. Let us suppose that $d(O_j, O_{j,2}) = 6$ and $d(O_j, O_i) = 2$, then the C_{jih} value = 6 - 2 = 4.

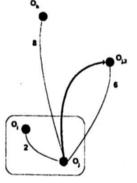


Figure 2.1. Case 1.

Case 2

 O_j currently belongs to the cluster represented by O_i . But this time, O_j is less similar to $O_{j,2}$ than O_h , $(d(O_j, O_h) < d(O_j, O_{j,2})$. Then, if O_i is replaced by O_h , O_j would belong to the cluster represented by O_h . This way the cost is given by $C_{jih} = d(O_j, O_h) - d(O_j, O_i)$. Unlike the equation in case 1, C_{jih} can be positive or negative, depending on whether O_i is more similar to O_i or to O_h .

Case 3

Supposing that O_j currently belongs to a cluster other than the one represented by O_i . In addition, assuming $O_{j,2}$ is the representative object of that cluster and O_j is more similar to $O_{j,2}$ that O_h . Then, even if O_i is replaced by O_h , O_j would stay in the cluster represented by $O_{j,2}$. Thus, the cost is given by $C_{jih} = 0$.

Case 4

 O_j currently belongs to the cluster represented by $O_{j,2}$. But O_j is less similar to $O_{j,2}$ than O_h . Replacing O_i with O_h would cause O_j jump to the cluster represented by O_h from cluster $O_{j,2}$. This way the cost is given by $C_{jih} = d(O_j, O_h) - d(O_j, O_{j,2})$. This cost always is negative.

Combining the four cases, the total cost of replacing O_i with O_h is given by:

$$TC_{ih} = \sum_{j} C_{jih}$$

Algorithm PAM:

1. Select k representative objects arbitrarily.

2. Compute TC_{ih} for all the pairs of objects O_i , O_h where O_i is currently selected, and O_h is not.

3. Select the pair O_i , O_h which corresponds to $minO_i$, O_h , TC_{ih} . If the minimum TC_{ih} is negative, replace O_i with O_h , and go back to step 2.

 Otherwise, for each non-selected object, find the most similar representative object.

5. Halt.

Spatial Data Mining based on Clustering Algorithms

In this section, we present two spatial data mining algorithms developed by Kaufman and Rousseeuw [7]: Spatial Dominant Approach (SD) and Non-Spatial Dominant Approach (NSD).

Spatial Dominant Approach: SD

There are different types of approaches to spatial data mining. A spatial database consists of spatial and non-spatial attributes. The non-spatial attributes are stored in relations. The general approach here is to use clustering algorithms to work with the spatial attributes, and use other learning tools to take care of non-spatial data over the spatial findings.

Non-Spatial Dominant Approach: NSD

The spatial dominant algorithms, such as SD, can be viewed as focusing asymmetrically on discovering non-spatial characterizations of spatial clusters. Non-spatial dominant algorithms, on the other hand, focus on discovering spatial clusters existing in the result of data mining in non-spatial data.

3. NON-SPATIAL DATA MINING

Data mining in general can be seen as the search for hidden patterns that may exist in databases [12]. The explosive growth in data and databases has generated a need for techniques and tools that can transform the data into useful information and knowledge.

Knowledge discovery in databases refers to the task of finding interesting knowledge, regularities, or high-level information from data sets, which can then be analyzed from different angles. Researchers in many different fields including database system, knowledge-base system, artificial intelligence, machine learning and statistic have shown great interest in data mining.

Some of the data mining techniques apply on structural data and others on non-structural data. A structural data is defined as data that describes the relationships among the objects described in the data. We can see the data objects as variables in the attribute-value representation, but now we also have relations among those variables.

In this work we used a data mining system that uses a graph-based learning technique. This technique has the potential to be competitive in the learning task, because it provides a powerful and flexible knowledge representation that can be used for relational domains.

SUBDUE.

The SUBDUE system [18] (developed at the University of Texas in Arlington) is a general tool that can be applied to any domain that can be represented as a graph. It discovers substructures that compress the original database and represents interesting structural concepts in previously-discovered data. By replacing substructures in the data, multiple passes of SUBDUE produce a hierarchical description of the structural regularities in the data. SUBDUE has the capability to use a constrained inexact graph match that can consider similar, but not identical, instances of a substructure as a pattern. SUBDUE uses the minimum description length principle to guide the search towards more appropriate substructures.

The SUBDUE system uses a graph representation. Objects in the data (concepts) become vertices or small subgraphs in the graph, and relationships between objects become directed or undirected edges in the graph. A substructure is a connected subgraph within the graph. This graph representation serves as input to the SUBDUE system. Figure 3.1 shows an example of an input database and its graph representation. The example is presented in terms of the house domain, where a house is defined as triangle on a square. T represents a triangle, C a square, E a star and R a rectangle. The objects in the figure (T1, C1, R1) become labeled vertices in the graph, and the relationship (on, shape) become labeled edges. The graph representation of the substructure discovered by SUBDUE from this data is shown in figure 3.2.

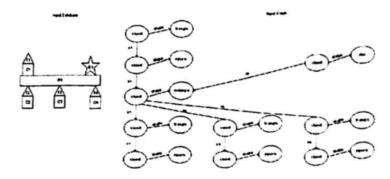


Figure 3.1. Graph Representation of the House Domain.

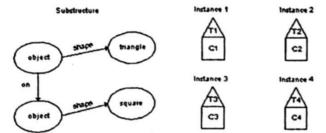


Figure 3.2. Substructure and Instances Discovered from the House Domain by Subduc.

An instance of a substructure in an input graph is a set of vertices and edges from the input graph that match the graph definition of the substructure. A neighboring edge of a substructure instance is an edge in the input graph that is not contained in the instance, but is connected to at least one vertex in the instance. An external connection of an instance of a substructure is a neighboring edge of the instance that is connected to at least one vertex not contained in the instance.

4. SPATIAL AND NONSPATIAL DATA MINING INTEGRATION

Database architecture

The database scheme complies with the Open GIS SQL92 specification [13]. The storage type used for geometries is the binary geometry schema (Well-know Binary for Representation Geometry). The database is made up of 150 entities containing descriptive and geometric information; in addition there are two more metadata catalogues.

System Architecture

The system was developed using the Java programming language and consists of 10 class packages (kdd, baseKDD, bd, shp, formats, factory, opengis, oracle, graficacion2D, paqueteDeGeometrias). Oracle was used as the database management system.

The Data Mining Process

The data mining process implemented in this project is based on the Spatial Dominant Approach (SD).

- Selection of the data layers that will integrate our initial data set.
- Data transformation for the application of the PAM algorithm.
- 3. Spatial data mining process (PAM).
- 4. Selection of a data subset generated by PAM (clusters).
- 5. Data transformation for the application of the SUBDUE system (graph creation).
- 6. Non-spatial data mining process (SUBDUE).
- 7. Pattern evaluation and interpretation.
- Knowledge application.

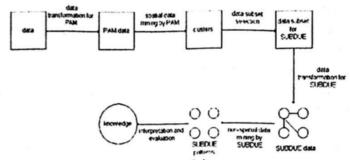


Figure 4.1. Data mining process.

5. SYSTEM IMPLEMENTATION

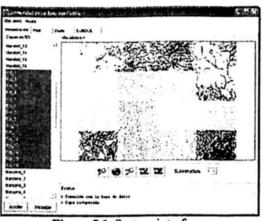


Figure 5.1. System interface.

The system is made up of the following four modules:

Visualization

This module allows the user to select and visualize the existing data layers in the database (figure 5.1). The user has the possibility to select one or more layers. Once the layers have been selected, they are displayed in the visualization area. In order to help the user to analysis them, there are options for applying colors to the lawyers for easier identification and the possibility to select a single spatial object and display its descriptive data. Additionally, we have the option to visualize the results of the data mining processes in a graphical way.

PAM

In this module the user applies the PAM data mining algorithm. Clusters of spatial objects (point, line, and polygon) are found on the basis of its closeness with other objects (using Euclidian distance). An example of an application for this technique is our wish to learn about the most important characteristics of the rivers located in the north zone of the Popocatépetl volcano. In order to delimit our data set, we can find clusters of the rivers and select the clusters within the zone of interest. Once we have identified the elements (clusters) we can apply SUBDUE to their descriptive data to find important characteristics about those rivers.

GRAPH

The Graph module is used to transform the descriptive information of the database (stored in relational form) into a graph representation. Each attribute value becomes a vertex and each attribute name becomes an arc. Once the graph has been created, it is stored in a text file. At this

point, data transformation is necessary since SUBDUE requires its input data set to be in graph form.

SUBDUE

This module is used to implement non-spatial data mining. It is invoked through a call to the operating system and its results directed to a text file. Later, the text file content must be loaded into the system to be able to visualize the results graphically. In order to mining the Popocatépetl database, we use the Spatial Dominant Approach.

6. RESULTS

In this section we present an example of the results generated by the system. For the application of the PAM algorithm we used three clusters. The representation of each cluster is pointed out by the following colors: cluster 1 red, cluster 2 blue and cluster 3 green. First, we apply the PAM algorithm to find clusters from the input data, in this case the Vialidad_1 and Vialidad_2 layers. Once we found the clusters we apply SUBDUE to the associated descriptive data. The representation of the results generated by SUBDUE is pointed out by the following colors: substructure 1 red, substructure 2 blue and substructure 3 green.

Figure 6.1 shows the results generated by PAM finding three clusters from the Vialidad_1 and Vialidad_2 entities that belong to the roads layer.

Processing the descriptive information from the Vialidad_1 and Vialidad_2 entities by SUBDUE (figure 6.2), we conclude that 60.07% of all roads are of the dirtroad type (pattern). The best substructure discovered by Subdue, which is shown below, supports this conclusion (figure 6.3).



Figure 6.1. Clusters from the Vialidad_1 and Vialidad_2 entities found by PAM.

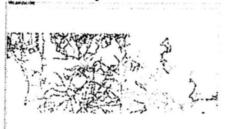


Figure 6.2. Patterns found by SUBDUE from the Vialidad_1 and Vialidad_2 layers.

Subgraph vertices 19 EVENT 24 7112.000000 25 DIRT-ROAD 26 0.000000 Subgraph edges [19 -> 24] CODE

[19 -> 24] CODE [19 -> 25] DESCRIPTION

[19 -> 26] ROUTE_NUMBER

Number of instances = 1345

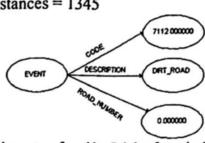


Figure 6.3. Substructure found by Subdue from the Vialidad_I and Vialidad_2 layers.

This substructure is very important because is telling us that there are some areas that need more material roads in order to implement a contingency plan to vacate the zone. In cases like this, we can use the PAM algorithm to find smaller clusters (from areas closer to the volcano) and find interesting knowledge with the Subdue system. Now would be trying to identify clusters with high population and with no material roads so that the existing dirty-roads could become a bottle neck in case of an emergency.

7. CONCLUSIONS AND FUTURE WORK

In this project we presented a system for applying spatial and non-spatial data mining techniques to the Popocatépetl volcano database. In the first case the clustering algorithm PAM was implemented, and in the second case the substructure discovery system SUBDUE was used. The results showed an efficient yield of PAM when working with small data sets. The complexity for one iteration is O((n-k)2). The response times of SUBDUE are based on the size of the input graph and the parameters established for their operation. The system architecture is made up of modules. Each class package is designed in such away that it is possible to add, modify or replace any of its elements. This is an important feature since new data mining technologies can be implemented in order to count with a lager number of tools to allow the discovery of more useful and beneficial knowledge.

This research showed that data mining technologies, developed to be used in other research fields, are feasible to be adapted according the context of geographic data. The PAM adaptation to work with data stored in the volcano database is a good example. The results generated by SUBDUE were transformed with the purpose of showing them graphically by using geometric objects.

The system is very valuable for its use by several kinds of users (in general, for decision making) in areas like contingency planning in case of a catastrophe. In order to enlarge the capabilities of the current system we propose for further work the issues below: Research on new clustering methodology techniques.

Handling of the SUBDUE data compression process. One of the most important features of this system is the replacement in the initial data set with the discovered substructures. This process is not supported in the current system; therefore, if in the result from SUBDUE there is a substitution, this one is not reflected in the results shown in the Visualization module. Building a Data Warehouse. The next step in the Popocatépetl volcano project evolution would be the implementation of a geographic warehouse in order to work with historical data and enhance the data mining processes.

ACKNOWLEDGEMENT

National Science and Technology Council (CONACYT). Project W35804-A. Access to High Quality Digital Services and Information for Large Communities of Users.

REFERENCES

- [1] Adhikary, Junas. Knowledge Discovery in Spatial Databases, Progress and Challenges. School of Computing Science, Simon Fraser University. 1996.
- [2] Bernhardsen, Tor. Geographic Information Systems, An Introduction, Second Edition. John Wiley & Son Inc. 1999.
- [3] Bunke, H., and G. Allermann. *Inexact graph matching for structural pattern recognition*. Pattern and Recognition Letters, 1983.
- [4] Frawley, W. J., G. Piatetsky-Shapiro, and C. J. Matheus. Knowledge Discovery in databases: An overview.
- [5] Gómez H. 2001. Urban Analysis Tool in a GIS. Bachelor Thesis. Computer System Engineering. Department of Computer Science, Engineering School, Universidad de las Américas-Puebla. May 2001.
- [6] Holder, L. B., D. J. Cook and S. Djoko. Substructure Discovery in the SUBDUE System. In Proceedings of the AAAI Workshop on Knowledge Discovery in Databases. 1994.
- [7] Kaufman, L., and P. J. Rousseeuw. Finding Groups in Data: an Introduction to Cluster Analysis. John Wiley & Son Inc. 1990.

- [8] Kolatch, Erica. Clustering Algorithms for Spatial Databases: An Survey. Department of Computer Science, University of Maryland, Collage Park. 2001.
- [9] Geo-information Laboratory. Universidad de las Américas-Puebla. Internet site, visit last time October 2002. http://mailweb.udlap.mx/~gisudla/.
- [10] Laurini R., and D. Thompson. Fundamentals of Spatial Information Systems, Academic Press. 1992.
- [11] Morales A. 2001. Construction and Modeling of a Geographic Database. Bachelor Thesis. Computer System Engineering. Department of Computer Science, Engineering School, Universidad de las Américas-Puebla. May 2001.
- [12] Ng, Raymond T. and Jiawei Han. 1994. Efficient and Effective Clustering Methods for Spatial Data Mining. Proceedings of the 20th Very Large Databases Conference (VLDB 98) Santiago, Chile.
- [13] Open Gis Consortium, Inc. *OpenGIS Simple Features Specification for SQL*, Revision 1.1. OpenGis Project Document 99-049. 1999. http://www.opengis.org.
- [14] Piatetsky-Shapiro, G., and W. J. Frawley, editors. Knowledge Discovery in Databases. AAAI/MIT Press, Menlo Park, CA, 1991.
- [15] Posada N. 2001. Implementation of a CBR for a Volcano Context. Master Thesis. Sciences in Computer System. Department of Computer Science, Engineering School, Universidad de las Américas-Puebla. December 2001.
- [16] Quinlan, J. R., and R. L. Rivest. *Inferring decision trees using the minimum description length principle*. Information and Computation, 1989.
- [17] Razo A. 2001. GISELA X3: Standard Modeling of Three-Dimensional Geometric Data with XML and Their Application in a Geographical Information System for Civil Protection. Master Thesis. Sciences in Computer System. Department of Computer Science, Engineering School, Universidad de las Américas-Puebla. December 2001.
- [18] SUBDUE. University of Texas at Arlington. Internet site, visit last time October 2002. http://cygnus.uta.edu/subdue/.
- [19] Witten, Ian H., and Eibe Frank. Data Mining, Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann Publishers.

AN IMPLEMENTATION OF SPATIO-TEMPORAL ANNOTATED CONSTRAINT LOGIC PROGRAMMING (STACLP)

Claudia ZEPEDA, David SOL {sc098382, sol}@mail.udlap.mx Universidad de las Américas-Puebla San Andrés Cholula, Puebla, México, MEXICO

ABSTRACT

In this paper we present important aspects throughout the implementation of Spatio-Temporal Annotated Constraint Logic Programming (STACLP), in which temporal and spatial data are represented by means of annotations that label atomic first order formulae. The reason to perform the implementation is due to that there is not software available for STACLP. At the same time, we explore the implementation of STACLP to obtain experience and to learn from it. This implementation uses XSB system to verify it as a good alternative of knowledge representation.

KEY WORDS

Constraint Logic Programming, Spatio-Temporal Reasoning, Knowledge Representation.

1. INTRODUCTION

One of the most promising directions of the current research in GIS field focuses on the development of reasoning formalisms that merge contributions from both Artificial Intelligence (AI) and mathematical research areas in order to express spatial qualitative reasoning. In recent literature [1] [2] [3] [4], there has been a great interest in studying spatial concepts from a cognitive point of view, giving rise to Qualitative Spatial Reasoning as a new research area between AI and GIS. Qualitative Spatial Representation addresses different aspects of space including topology, orientation, shape, size, and distance. Qualitative Spatial Reasoning also has been recognized as a major point in the future developments of GIS [1] [2] [3] [4]. It has been claimed that today GIS technology is capable of efficiently storing terabytes of data, but the key point is to abstract away from the huge amount of numerical data and to define formalisms that allow the user to specify qualitative queries. In [1] [2] [3] [4] is showed how the approach of expressing Spatio-Temporal Reasoning on geographical data is useful and they introduce STACLP Spatio-Temporal Annotated Constraint Logic Programming (STACLP) where temporal and spatial data are represented by means of annotations that label atomic first order formulae. For example, if we assume that a transport is described by its

name, the kind of things it transports and its spatial position(s) in certain time interval, using STACLP we can model information like Bus1 transports people, it is parked from 10am to 12am, and it is working from 12am to 5pm in some places of Calpan town but it is loading people between 1pm to 2pm.

In this paper are presented some important aspects throughout the implementation, using XSB system, of STACLP. The reason to perform the implementation is because up to this time there is not a software available for STACLP and using XSB looks to be a very interesting and useful way to explore STACLP and to learn from it. And at the same time, we can verify XSB system as an alternative of knowledge representation.

The paper is organized as follows. In Section 2, it is presented STACLP which combines spatial and temporal annotations and its semantics, and in Section 3, it is presented some features of XSB considered as necessary for the implementation of the meta-interpreter of STACLP. In Section 4 are showed some important aspects throughout the implementation, using XSB, of the meta-interpreter for STACLP and in this same section some examples of queries involving spatial and temporal knowledge are formulated in our implemented meta-interpreter. Finally, in Section 5, the conclusions are discussed and possible future work.

2. SPATIO-TEMPORAL ANNOTATED CONSTRAINT LOGIC PROGRAMMING

In [4] was introduced an extension to Temporal Annotated Constraint Logic Programming (TACLP) where both temporal and spatial information can be dealt with and reasoned about. The resulting framework was called Spatio-Temporal Annotated Constraint Logic Programming (STACLP), where temporal and spatial data are represented by means of annotations that label atomic first order formulae.

Time can be discrete or dense. Time points are totally ordered by the relation \leq . We denote by τ the set of time points and we suppose to have a set of operations to manage such points. A *time period* is an interval [r,s] with r,s τ and 0 $r \leq s \leq \infty$ which represents the convex and non-empty set of time points $\{t \mid r \leq t \leq s\}$. Analogously, space can be discrete or dense and we consider as spatial regions rectangles represented as [(x1,x2), (y1,y2)] which are intended to model the region $\{(x,y) \mid x1 \leq x \leq x2\}$, $\{y1 \leq y \leq y2\}$ [4].

In [1] is defined an annotated formula as a formula with the form $A\alpha$, where A is an atomic formula and α is an annotation. Also are defined three kinds of temporal and spatial annotations:

at T and atp (X,Y) are used to express that a formula holds in a time point or in a spatial point.

th I, thr R are used to express that a formula holds throughout, i.e. at every point, in the temporal interval or the spatial region, respectively.

in I, inr R are used to express that a formula holds at some point(s), in the temporal interval or the spatial region, respectively.

On the other hand, in [1] is showed that the set of annotations is endowed with a partial order relation which turns it into a lattice. Given two annotations α and β , the intuition is that α β if α is "less informative" than β in the sense that for all formulae A, $A\beta \Rightarrow A\alpha$. This partial order is used in the definition of new inference rules. In addition to Modus Ponens, STACLP has the two inference rules below:

$$\frac{A\alpha \qquad \gamma \sqsubseteq \alpha}{A\gamma} \qquad rule(\sqsubseteq)$$

$$\frac{A\alpha \qquad A\beta \qquad \gamma = \alpha \sqcup \beta}{A\gamma} \qquad rule(\sqcup)$$

The rule \subseteq states that if a formula holds with some annotation, then it also holds with all annotations that are smaller according to the lattice ordering. The rule says that if a formula holds with some annotation and the same formula holds with another annotation β then it holds with the least upper bound α β of the two annotations. We can find the constraint theory for temporal and spatial annotations in [1].

The class of annotations which combines spatial and temporal annotations was introduced in [4].

Definition (Spatio-Temporal Annotations). The class of Spatio-Temporal annotations is the pairing of the spatial annotations Spat built from atp, thr and inr and of the temporal annotations Temp, built from at, th and in, i.e. Spat*Temp.

For technical reasons related to the properties of annotations, in [4] is restricted the rule—only to least upper bounds that produce regions which are rectangles and the temporal components are time periods. Thus there are six cases considered. Here, only present the first case where last upper bound is considered and the other five cases can be found in [4].

$$\begin{array}{l} \operatorname{thr}[(x_1, x_2), (y_1, y_2)] \operatorname{th}[t_1, t_2] \sqcup \\ \operatorname{thr}[(x_1, x_2), (z_1, z_2)] \operatorname{th}[t_1, t_2] = \\ \operatorname{thr}[(x_1, x_2), (y_1, z_2)] \operatorname{th}[t_1, t_2] \Leftrightarrow \\ y_1 \leq z_1, z_1 \leq y_2, y_2 \leq z_2 \end{array}$$

This axiom allows one to enlarge the region in which a property holds in a certain interval. If a property A holds both throughout a region R1 and throughout a region R2 in every point of the time period I then it holds throughout the region which is the union of R1 and R2, throughout I. Notice that the constraints on the spatial variables ensure that the resulting region is still a rectangle.

The clausal fragment of STACLP, which can be used as an efficient Spatio-Temporal Programming Language, consists of clauses of the following form [1]:

$$A\alpha\beta \leftarrow C_1, \ldots, C_n, B_1\alpha_1\beta_1, \ldots, B_m\alpha_m\beta_m$$

where n, $m \ge 0$, A is an atom (not a constraint), α , α i, β , β i are (optional) temporal and spatial annotations, the Cj's are constraints and the Bi's are atomic formulae. Constraints Cj cannot be annotated. A STACLP program is a finite set of STACLP clauses.

Semantics of STACLP

The definition of the semantics for STACLP is as follows [1]:

- It is assumed that all atoms are annotated with th, in, thr or inr labels. In fact, at t and atp(x,y) annotations can be replaced with th[t,t] and thr [(x,x),(y,y)] respectively by exploiting the (at th) and (atp thr) axioms.
- Each atom in the object level program which is not two-annotated, i.e., which is labeled by at most one kind of annotation, is intended to be true throughout the whole lacking dimension(s). For instance an atom A thr R is transformed into the two-annotated atom A thr R th [0,∞].
- Constraints remain unchanged.
- The meta-interpreter for STACLP is defined by the following clauses:

$$demo(empty) \\ demo((B_1, B_2)) \leftarrow demo(B_1), demo(B_2) \\ demo(A\alpha\beta) \leftarrow \alpha \sqsubseteq \delta, \beta \sqsubseteq \gamma, \\ clause(A \delta\gamma, B), demo(B) \\ demo(A\alpha'\beta') \leftarrow \alpha_1\beta_1 \sqcup \alpha_2\beta_2 = \alpha\beta, \\ \alpha' \sqsubseteq \alpha, \beta' \sqsubseteq \beta, clause(A \alpha_1\beta_1, B), \\ demo(B), demo(A\alpha_2\beta_2) \\ demo((C)) \leftarrow constraint(C), C$$

 A clause Aαβ←B of a STACLP program is represented at the meta-level by

$$classe(A\alpha\beta, B) \leftarrow valid(\alpha), valid(\beta)$$

where valid is a predicate that checks whether the interval or the region in the annotation is not empty.

The resolution rule, third clause, implements both the Modus Ponens rule and the rule \square . It contains two relational constraints on annotations, which are processed by the constraint solver using the constraint theory for temporal and spatial annotations mentioned in Section 2.2. Fourth clause implements the rule \square combined with Modus Ponens and rule \square . The constraint $\alpha_1\beta_1 \sqcup \alpha_2\beta_2 = \alpha\beta$ in such a clause is solved by means of the axioms defining the least upper bound mentioned in Section 2.3. Fifth clause manages constraints by passing them directly to the constraint solver.

3. SYSTEM XSB

We used XSB system to implement the meta-interpreter for STACLP defined in Section 2.5. In this Section we present some features of XSB considered as necessary for the implementation of the meta-interpreter.

XSB is a Logic Programming and Deductive Database system for Unix and Windows. In addition to providing all the functionality of Prolog, XSB contains several features not usually found in Logic Programming systems, including [6]:

- Constraint handling for tabled programs based on an engine-level implementation of annotated variables and a package, clpqr, for handling real constraints.
- A number of interfaces to other software systems, such as C, Java, Perl, ODBC, SModels, and Oracle.
- Preprocessors and Interpreters so that XSB can be used to evaluate programs that are based on advanced formalisms, such as extended logic programs (according to the Well-Founded Semantics [9]); Generalized Annotated Programs [7]; and F-Logic.

Prolog is based on a depth-first search through trees that are built using program clause resolution (SLD) [5]. As such, Prolog is susceptible to getting lost in an infinite

branch of a search tree, where it may loop infinitely. SLG evaluation, available in XSB, can correctly evaluate many such logic programs if they are compiled as a **tabled** predicate. The user can declare that SLG resolution is to be used for a predicate by using table declarations. Alternately, an **auto_table** compiler directive can be used to direct the system to invoke a simple static analysis to decide what predicates to table.

4. IMPLEMENTATION OF THE META-INTERPRETER

In this Section we present some important aspects throughout the implementation using XSB of the meta-interpreter for STACLP, defined in Section 2.5, and some examples of queries involving spatial and temporal knowledge.

Implementation using XSB

Below we show the main code of the meta-interpreter for STACLP, the total number of code lines is about 300.

```
:- auto_table.

demo(true).

demo(((B1,Alfa1,Beta1),(B2,Alfa2,Beta2))):-

demo((B1,Alfa1,Beta1)),demo((B2,Alfa2,Beta2)).

demo((A,Alfa,Beta)):-

contR(Alfa,Delta),contT(Beta,Gama),

clause1((A,Delta,Gama),B),demo(B).

demo((A,Alfap,Betap)):-

une(Alfa1,Beta1,Alfa2,Beta2,Alfa,Beta),

contR(Alfap,Alfa),contT(Betap,Beta),

clause1((A,Alfa1,Beta1),B),demo(B),

demo((A,Alfa2,Beta2)).
```

clause1((A,Alfa,Beta),B) :claus((A,Alfa,Beta),B), validaR(Alfa), validaT(Beta).

The clauses *contR* and *contT* correspond to the rule (to Space and Time respectively mentioned in Section 2.2.

Clause *une* implements the rule (L) of Section 2.3. In Section 2.5 we described that at t and atp(x,y) annotations can be replaced with th [t,t] and thr [(x,x),(y,y)] respectively by exploiting the $(at\ th)$ and $(atp\ thr)$ axioms. However, it is necessary highlight that a consequence of this replacement is that $(at\ th)$, $(atp\ thr)$, $(at\ in)$ and $(atp\ inr)$ axioms can be rewritten as the two next axioms

```
(th in) th [t,t] = in [t,t]
(thr inr) thr[(x,x),(y,y)] = in[(x,x),(y,y)]
```

because are the *link* between (th) annotation and (in) annotation or between (thr) annotation and (inr)

respectively. So, these two axioms were implemented instead of the other four axioms.

On the other hand, we can see as the first clause auto_table, because without it there would be some cases where XSB may loop infinitely.

Examples

In this section we present some examples which illustrate how spatial data are modeled by annotations and integrated with temporal information using the implemented meta-interpreter.

Volcano Zone

This example describes spatial and temporal information in the context of Popocatépetl volcano. Decision making is a very important activity in this context. About 200,000 people distributed in 50 towns are in danger when the volcano starts its activity. "Plan Operativo Popocatépetl" office in Puebla has the responsibility of coordinating the actions to keep the integrity of the people. This office uses printed maps and printed reports to decide the best sequence of actions in case of danger. Usually it is difficult to justify the decisions because they do not have enough information. Our example describes the activity of buses. We assume that a transport is described by its name, the kind of things it transports and its spatial position(s) in certain time interval. For instance, Bus1 transports people and it is parked from 10am to 12am, and it is working from 12am to 5pm in some places of Calpan town but it is loading people between 1pm to 2pm. This can be expressed by means of the following clauses in our implemented meta-interpreter.

claus(transport(bus1,people),true). claus((does(bus1,parking), inr(3,3,4,4),th(10,12),true). claus((does(bus1,working), inr(1,2,2,3),in(12,17)),true). claus((does(bus1,loading),thr(1,1,2,2),th(13,14)),true).

claus(transport(bus2,people),true). claus((does(bus2,working), inr(3,4,4,7),th(10,12)),true). claus((does(bus2,loading), thr(1,1,2,2),th(14,15)),true).

Furthermore, a *town* can be described by its name and its area represented by thr annotation. The temporal information for a town is represented by th(0,24) annotation because, as we know, a town is all the time in the same spatial position.

claus((town(huejotzingo), thr(3,4,4,7),th(0,24)),true). claus((town(calpan), thr(1,2,2,3),th(0,24)),true).

Now we show how some queries, involving the spatial and temporal knowledge, can be formulated:

• Which buses did load at Calpan between 12am and 3pm?

demo(((does(Y loading) inr(P) in(12.15)))(sum(calpan))

demo(((does(X, loading), inr(R), in(12, 15))town(calpan), thr(R), th()))

```
X = bus1
X = bus2
```

The answer to this query consists of all the buses that were loading at Calpan during that time period. The region of Calpan is assigned to the variable **R** and then is solved which buses were loading in some place of that region. We use in annotation because we want to know all the different positions of every bus between 12am and 3pm while the inr annotation allows one to know the region every bus is in during that time period, even if its exact position is unknown.

If we asked for

```
demo(((does(X,loading), atp(R), th(12,15)), (town(calpan), thr(R), th()))).
```

then we would have constrained buses to stay in only one place, in this example Calpan town, for the whole time period.

```
The query demo(( (does(X,loading),atp(R), in(12,15)), (town(calpan), thr(R), th(_)) )).
```

asks buses which are in definite positions sometime in (12, 15).

 Which buses were working or parked in Huejotzingo before 12am?

```
demo(( (does(X,working), inr(R), in(0,12)),
  (does(X,parking),inr(R),in(0,12)),
  (town(huejotzingo),thr(R), th(_)) )).
  X = bus1
  X = bus2
```

The result are all the buses that were working or parked in the region, assigned to variable R, that corresponds to Huejotzingo town, before 12am.

In Section 4.1, we saw that in the implementation of the meta-interpreter the first clause is *auto_table*, and it was necessary because without it there would be some cases where XSB may loop infinitely. An example of this situation occurs if the query to resolve is about some inexistent information. If we use the same information of the last example and we formulate the query *Is Bus2 parked at Atlixco?* in this case there is not information about where is Atlixco town so it leads to an infinity loop. The reason of the infinity loop, in this example, is because the meta-interpreter tries to resolve the clauses **contR** or **contT** corresponding to the rule to Space or Time mentioned in Section 2.2., and when it fails, it tries to resolve the query once and again.

5. CONCLUSION

Thanks to the implementation of STACLP we could realize that the axioms about (at th), (atp thr), (at in) and (tp inr) can be replaced for the next two: (th in), (thr inr). We used XSB system, because we consider that it has the necessary features for the implementation of the meta-interpreter. Finally, it would be interesting to cope with an interface using Natural Language to specify queries. More ideas about the evolution of STACLP are in [1] [2] [3] [4].

4. ACKNOWLEDGEMENT

This work has been supported by the Mexican Council of Science and Technology (CONACYT) W 35804- A.

REFERENCES

[1] A. Raffaetá, C. Renso, F. Turini, Qualitative Reasoning in a Spatio-Temporal Language, CRGD Workshop 2001, http://www-kdd.cnuce.cnr.it/crgd/

- [2] C. Renso, A. Raffaetá. Temporal Reasoning in Geographical Information Systems. *DEXA Workshop*, 2000, 899-905.
- [3] P. Mancarella, G. Nerbini, A.Raffaetá, F.Turini, MuTACLP: A Language for Declarative GIS Analysis, Computational Logic, 2000, 1002-1016
- [4] A.Raffaetá and T. Frühwirth, Spatio-Temporal Annotated Constraint Logic Programming, *PADL 2001*, 2001, 1990, 259-273.
- [5] W. Chen and D. S. Warren, Tabled Evaluation with Delaying for General Logic Programs, *Journal of the ACM*, 43(1), 1996, 20-74
- [6]D. S. Warren, Applications of logic databases. Programming the PTQ Grammar in XSB (Ed.: Raghu Ramakrishnan, Kluwer Academic Publishers, USA. 1995).
- [7] M. Kifer, V. S. Subrahmanian, Theory of generalized annotated logic programming and its applications, *J. Logic Programming*, 12(4), 1992, 335-368.
- [9] J. Alferes, C. Damasio, and L. Pereira, SLX: a top-down derivation procedure for programs with explicit negation, *International Logic Programming Symp.*, 1994, 424-439.

GISELA: A WEB-BASED INTERFACE FOR A RISK MANAGEMENT GIS USING XML

Antonio RAZO, David SOL

Center for Research on Automation and Information Technologies CENTIA,

Universidad de las Américas Puebla

Sta. Catarina Martir s/n Cholula, Puebla 72820

MEXICO

{anrazo, sol}@mail.udlap.mx

ABSTRACT

This paper presents the design and implementation of a web-based interface for a risk management GIS using XML and open standards. The analysis and visualization of geographic information in a distributed environment like Internet is essential nowadays. In this paper we will present how we are using XML to transfer and display geographic data. Several efforts have been done to translate the Geographic Markup Language (GML) [1] (an XML specification proposed by OpenGis [2]) to a 2D representation using SVG [3]. In this project we are also looking to translate GML documents into a 3D representation using X3D and VRML [4]. This representation is ideal because of the three-dimensional nature of geographic information. Another advantage of using VRML is the GeoVRML [4] extension, because we can generate geo-referenced three-dimensional models. We use XSLT stylesheets to transform the GML document into VRML. XSLT is the technology proposed by the W3C to translate XML documents. The work presented in this paper will be implemented on a proposal for a risk management GIS for the Popocatépetl volcano.

KEY WORDS

GIS, Visualization, GML, VRML, XSLT, Risk Management

1. INTRODUCTION

Authorities responsible of risk management identify and evaluate potential disasters to reduce loss and damage of people and properties [5]. Our proposal will support the decision making process for risk management and will describe a geo-spatial data model for this context. The work described in this paper will be implemented with the Popocatépetl volcano data where risk reduction is an essential activity. When the Popocatépetl volcano starts its activity, about 200,000 people distributed in 50 towns are in danger [6]. On Figure 1 we can see the risk areas associated to the volcano and the population affected. Visualization of geographic data allows the experts to

analyze and evaluate different risk situations. Because risk management needs several authorities to have an opinion about the same phenomenon we are looking to support the analysis and visualization of this geographic information in a distributed environment.

The visualization in any Geographic Information System is a very important component, this task can be done by using several formats and services but in a distributed environment like Internet the use of a standard format is important. In this paper we will present the design and implementation of a web-based interface for a risk management GIS using XML and open standards.

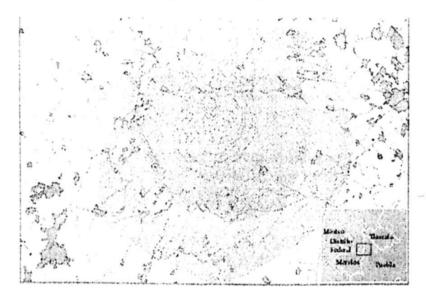


Figure 1. Risk Areas associated to the Popocatépetl volcano.

2. RELATED WORKS

The use of XML for distributed geographic information was presented on [7]. The prototype presented in this article describes how distributed databases, a mediator server and clients have been integrated using XML. The client, a Java applet received XML documents from geographic information integrated from diverse sources (OODB, Files, and RDB). The XML document contained geo-referenced information modeled after Oracle's GeoXML. This article mention several advantages of using XML, the independence of platform, the ease of use, the integration of descriptive and geographic

information and the validation of information using a DTD.

On [8] shows how to use XSLT stylesheets to transform a GML document into a 2D map with SVG. SVG is the Scalable Vector Graphic format, an XML proposal for vector data in two dimensions of the W3C [3]. The main advantage of using XSLT is that it is a standard to manipulate XML documents, and can be implemented by a client or by a server. The advantage of using SVG instead of an image format is that information from a SVG document is vector (instead of raster), so we can have multiple zooms, interaction and animation.

3. GISELA DESIGN

In a distributed environment, like Internet, information servers and user needs can have different configurations. For example, we can process data in one server and visualize it in another; we may also need information from several servers and process it on a client machine. XML is the ideal solution for these needs. Using GML we can represent geographic information, including geometric and descriptive data with XML. GML let us exchange and share data between different sources and servers. For data visualization we are working with XSLT to translate a GML document into HTML, SVG and X3D. We use HTML to present descriptive data, SVG for a 2D representation and X3D for 3D representation. Figure 2 shows this process.

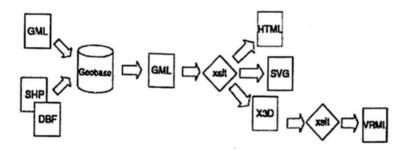


Figure 2. Giscla Design.

4. IMPLEMENTATION

Several tests have been made in the context of the Popocatépetl volcano. We acquire digital cartography of the zone consisting of roads, cities and towns, rivers, and terrain scale 1:20,000. We have been working with local authorities and with the Mexican National Center for Disaster Prevention (CENAPRED) to model the different risk associated to the volcano (lava, mudflows, and lahars). For the 3D representation we generated digital terrain models shown in Figure 3.

This information is stored on a relational database modeled after the OpenGis specification for SQL [9], extended to store the third dimension. For this process we have implemented classes in Java to export digital cartography in Shapefile format and GML documents to a

database server. Servlet connections using JDBC provide the information retrieval. For the generation and translation of XML, we use JDOM [10]. We have been working with several XSLT stylesheets to generate reports and tables in HTML, 2D maps using SVG and 3D maps using X3D (translated then to VRML to visualize it).

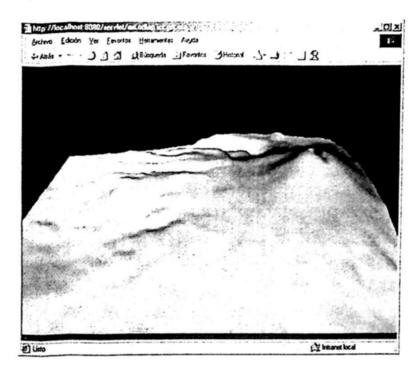


Figure 3. Popocatépetl digital model.

4.1 VRML AND GEOVRML

X3D is the XML equivalent of VRML, the translation of GML to X3D using an XSTL stylesheet is done in the following way, first we identify the geometric properties of every feature on the GML document. If the geometric property is a point, we generate a sphere, if it is a line or polygon we generate a sequence of lines in 3d, if it is terrain we generate a grid. When using VRML we need to transform the geographic coordinates into world coordinates but if we use the GeoVRML extension we just need to specify the spatial reference system and we can use the same coordinates [11]. After applying the stylesheet we have an X3D document, this specification is recent, so we need to translate this document into VRML to visualize it using a browser plug-in. In this case we use the stylesheet provided on [4].

5. CONCLUSION

We are using standards to share and integrate data from different sources. With a web-based interface users can access the data from different geographic locations and in different formats according to their needs. The use of VRML and its extension GeoVRML is an ideal representation of geographic information. We are looking to extend our services providing Metadata and dynamic map generation.

ACKNOWLEDGEMENT

The authors thank all the people working on the Plan Operativo Popocatépetl office at Puebla (Mexico) and the Mexican National Center for Disaster Prevention (CENAPRED) for their help and support. This project is supported by the Mexican National Council of Science and Technology, CONACYT project 35804-A.

REFERENCES

- [1] OpenGis 2000. Geography Markup Language (GML) v1.0. OGC Document Number: 00-029. Date: April 25, 2000. http://www.opengis.org
- [2] OpenGis 2002. OpenGis Consortium. Available: http://www.opengis.org, Date: February 2, 2002.
- [3] W3C 2001. World Wide Web Consortium. Available: http://www.w3.org, Date: November, 2001.
- [4] VRML 2002. Virtual Reality Modelling Language group. Available: http://www.vrml.org, Date: March 3, 2002.

- [5] K. Smith, Environmental Hazards Assessing risk and reducing disaster (Ed.: Routledge, New York: 3rd edition Routledge, 2001).
- [6] Gobierno del Estado de Puebla, *Plan de preparativos* para la emergencia del volcán Popocatépetl (Documento del Sistema Estatal de Protección Civil, 1997).
- [7] J. Zhang, M. Javed, A. Shaheen & L. Gruenwald, Prototype for Wrapping and Visualizing Geo-Referenced Data in a Distributed Environment Using XML Technology, Proceedings of the 8th ACM Symposium on Advances in Geographic Information Systems, New York, USA, 2000, 27-32.
- [8] W. Stephan, GML to SVG. Available: http://www.geoinfo.tuwien.ac.at/, Date: November 18, 2001.
- [9] OpenGis, Simple Features Specification For SQL. Revision 1.1. OpenGIS Project Document 99-049, Date: May 5, 1999, http://www.opengis.org.
- [10] JDOM Library JDOM. Available: http://www.jdom.org, Date: November, 2001.

THE SPATIAL ENABLING OF INFORMATION PARADIGM

J. Armando GUEVARA
President and Chief Executive Officer
GTT NetCorp, Inc.
Phone (813) 854-4968, Fax (813) 818-9804
jaguevara@gttnetcorp.com
www.gttnetcorp.com

ABSTRACT

Information Technology with the Internet and increasingly high-bandwidth wireless communications has become the backbone of the new world economy. Software-hardware-communications. Data-information-decisions. They all have become interlinked to create a new world order and as such, new work-flow paradigms and tools are required to effectively integrate and transform data into useful information to make decisions.

With the maturing of geospatial technologies such as GIS, GPS, Digital Photogrammetry, and Remote Sensing, locational facts have now become a key component of the decision-making process. Beyond system to system integration, the embedding of locational facts requires the design, development, and implementation of new solutions that integrate geospatial, information and communications technology aimed at optimizing process work flows, improve the decision making process and reduce operating costs in an innovative way. This integration process is defined as the "spatial enabling of information".

Coupled with these mature geospatial technologies, the advent of high-resolution imagery and new -more "intelligent"- technologies for the automation of pattern recognition and feature extraction, based on expert systems and object oriented technology, paves the way for a work-flow paradigm shift in the automated process of transforming data into information.

This paper presents the most relevant paradigm shifts in information and geospatial technology, it illustrates the three cardinal principles of postmodern technology: miniaturization, digitization and synthesis, and how they relate to innovative information-centric solutions. As part of the digital convergence paradigm, this paper elaborates on the power of the concept of "spatial enabling" and how it applies to new work-flow processes that encompass integrating evolving technologies to more effectively transform spatially enabled data to information.

KEY WORDS

Geospatial technologies, Automation, Software-hardware-communications, Data-information-decisions.

1. INTRODUCTION

With all the wealth of information and with the speed that it is reaching us, for those of us in the Information Technology field in general, and GIS in specific, it is not difficult to see what is coming. One thing for certain is that "it" will keep getting smaller, faster, and comparatively cheaper.

A true visionary will see beyond that sea of information, a Jules Verne or Leonardo Da Vinci type (flying before any one had conceived airplanes, going to the moon before the existence of rockets). Man is a true wonder when it comes to the creation of artifacts and pushing the threshold of the discovery.

In the early days of the radio a survey was made to ask users how they felt that the radio could be improved. Many came with suggestions like bigger knobs, better lid display, and others. Yet, nobody requested to see the person that was talking. This is part of envisioning where technology is going. But eventually someone did come that not only requested to see the person talking but made it happen.

GIS is a part of this scientific discovery journey. GIS is a result of the Geographer wanting to go beyond paper maps and into more complex and productive spatial reasoning. Leave the number crunching behind, let the machine do it, and take the new available time to more reasoning and making better decisions. But this is not unique to GIS. This is also part of what information systems in general are all about: organized data, organized information.

The evolution of certain technologies is a cascading effect. Take for example the airplane. The basic principles and fundamentals have not changed. Advances in scientific knowledge, methods and tools in aerodynamic tunnels, numerical and CAD systems, atmospheric and

space physics, propulsion systems, and so forth, have produced a more advanced, but very similar looking to the early concept, airplane. (Note: we will not need the airplane once teletransportation systems "a la Stark Trek" are invented.)

The reason I take the airplane example is because that is how I see GIS evolving. Cascading technologies that will improve and affect the evolution of GIS. There are two areas of domain that we can talk about GIS. On the one hand there is the R&D side of it and on the other, there is the use, application, of the technology. Again, this duality is not unique to GIS, but a result of being a technology.

1.1. PEOPLE AND TECHNOLOGICAL ADVANCES.

Of all the information technology advances, the one that marked the transition from a potential to an information revolution was the discovery of the microprocessor and with it personal computing devices. The microprocessor revolution resulted to be the enabler: hardware companies pushing the threshold of what could fit on a silicon chip as well as to what could be done with it. This path of miniaturization continues today.

PCs gave way to the Work Station market, and yet more even powerful PCs to the point where from the practical point of view it is hard today to distinguish one from the other. Xerox's research center, PARC, became the center of what was going to revolutionize the way how people interacted with computers (Kearns and Nadler, 1992). PARC forged the future for today's windows graphical user interfaces (GUI), the mouse, the laser printer, the bit map display, Ethernet, multimedia, the work station and concurrently with IBM the reduced instruction set computer (RISC). In those days Xerox was able to develop more new technology than it could commercially launch. Companies such as APPLE, INTEL and MICROSOFT thrived under this revolution. SUN gave way to the era of the workstations and the likes of Hewlett Packard and Silicon Graphics gave it presence and permanence in today's competing markets.

GIS in the midst of this revolution (70s, early 80s) had to tackle problems of primary algorithmic nature: solving the polygon overlay problem, line generalization, vector intersection, efficient spatial indexing and searching, and others. Like PARC for hardware and software, places like the Harvard Computer Graphics Laboratory and the GIS Laboratory at the State University of New York at Buffalo were the crib to what was to become GIS.

During the mid 70's and early 80's GIS was looking for its own identity attempting this by borrowing concepts and methodologies from Computer Science, Management Information Systems and other related disciplines. Today there is still no holistic body of scientific knowledge that can be termed as "a GIS discipline". Part of the problem

is that we tend to confuse GIS as a tool from GIS as a discipline. The tool is the enabler. The discipline is not one but multiple disciplines. GIS is a multidisciplinary tool, hence the body that contributes to the R&D of GIS is also of the same multidisciplinary nature.

Advances in computational geometry, data base management systems and more powerful central processing units (CPUs), more random access memory (RAM) and more disk space allowed, GIS in the early 80's to finally come to life as a practical technology. So where is GIS going in the twenty first century and the years to come?

2. TECHNOLOGICAL CONVERGENCE

GIS technology is part of a larger picture of technologies called Information Technologies (IT). The one particular thing about GIS is that its data is of spatial nature: a spatial datum is locatable in a given space under a predefined coordinate system, it is relate able to other spatial datum and the spatial datum can be described. Other than this, GIS faces the same kind of problems than any other IT system, more so with the growing presence of multimedia systems, where the spatial component is one more of the data sources that it must handle.

GIS has borrowed from mathematics (topology), geometry (computational geometry), computer graphics, computer aided design, data base management, geography and other related disciplines, to cast its technological purpose. This is at the core of what GIS is all about: an integrating technology. A technology unto which many other technologies converge.

Today's GIS have reached a level of functional technological convergence that makes it hard for the potential user to distinguish one commercial product from another. This is a sign of maturity. This maturity has been reached via various iterations of research and development. The first concern was that the technology actually worked, that it was reliable and robust. Later users began to demand a level of increasing functionality as their level of sophistication in using the technology also increased. The drawback of this increase of functionality was added complexity in learning and using the system. This created the pressure to simplify usage and make the systems more amicable and easier to use. GIS in this respect still remains largely in the domain of the specialist. This problem remains at the crux of making GIS available to the masses.

However, as we will see, certain technological shifts are providing the momentum to allow for GIS to migrate to the masses under a form that it is not any longer necessary for a common person to know how to use or what it is GIS. The underlying technology in the product being used just "tells them", where they are, how to get there, or describe the surroundings, among other "spatial things".

just "tells them", where they are, how to get there, or describe the surroundings, among other "spatial things".

3. HISTORIC TECHNOLOGICAL SHIFTS

The advent of the PC product, beginning with the APPLE MAC, gave way to a new computing paradigm that empowered the masses to know about computers and to undertake and solve more intricate problems with the aid of PCs. Early PCs however did not have the power necessary to run GIS software. The introduction of the INTEL 286 processor and later on the family of Intel processors in the MHz speed domain, made the potential for GIS to become a ubiquitous technology, a reality.

The benefits that the MAC and later on the WINDOWS software products brought is that they made us, the GIS R&D people, realize that unless there was an adequate metaphor to handle the real world in the computer, that users would have a hard time, hence a higher level of resistance, to using computers. The des top world, a finite and well-defined set of tools and activities were castled to graphical user interfaces that became intuitive with time and accessible to just about all computer neophytes.

It was not until the mid to late 80s that GIS companies began to adopt GUIs as their front end. However, most of these interfaces were just a "translation" of the command version of the system, to a graphical point and click version. These interfaces lacked the metaphor that has made Desktop computing so popular. Part of the reason is that handling the real world "real world" (not just the desk top real world) is a much more complex endeavor and representation. Part of the problem also is that GIS until now has only automated processes, and until recently, not really changed how we do things.

For instance, maps are still being done in many ways in the traditional sense. The difference is that instead of using drafting tools, we use digital tools. How a map is produced and used has not in essence changed. Hence our way of thinking in this respect has not changed very much either. There is however a wave of new technologies that may change forever how we produce and deal with maps. There is also a paradigm shift in how information technology is reinventing the enterprise. These shifts are affecting the evolution of GIS.

3.1. PARADIGM SHIFT

In his book "Paradigm Shift, the New Promise of Information Technology" Don Tapscott with Art Carson (1993) points out that many business leaders are working hard to reinvent their enterprises. The organization of old is considered in trouble. The command and control hierarchy established centuries ago is often bureaucratic, inwardly focused, unresponsive, unproductive, and stifling.

It is said that to meet the coming global challenges that the new enterprise has to be dynamic and be able to respond quickly to changing market conditions. Its structure is radically different by being flat, eliminating bureaucratic hierarchy. It is based on commitment rather than control. Business processes are reengineered for productivity and quality. It is open-focused outward-and networked redefining the traditional boundaries of the company.

The challenge to change the way the enterprises work entails a holistic view at five key components which mark the success or failure of any company: 1) talent, 2) time, 3) technology 4) tools and 5) tasks. GIS is concerned with these five issues and it will play an important role in the new technology paradigm, the new business environment, the new geopolitical order (relevant because of the strong geographic component) and the new enterprise. Figure 1 summarizes what these new situations mean (after Tapscott 1993).

These new situations challenge the current order of 1) technology, 2) organizations, 3) people and 4) methodologies and processes.

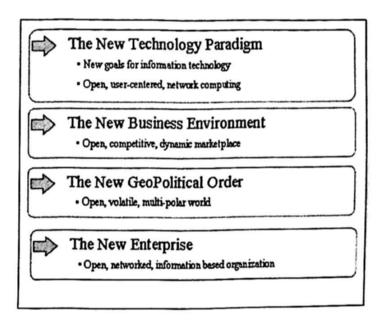


Figure 1. New Shifts.

The importance of looking at GIS under this light resides on the fact that, as previously mentioned, that GIS belongs to a larger array of technologies termed Information Technologies (IT), hence how the overall scope of IT affects the way companies become more efficient impacts the way GIS is evaluated. Also, GIS because of its spatial nature, is a tool that can readily support decision-making at many levels of this spatial domain.

For many years and until recently, IT has been primarily used to automate existing ways of working and existing business process, hence it is of no surprise that GIS has followed a similar path. Confronted with overgrown and inefficient corporate structures, the latter part of the 90s saw with great disappointment what was though to be a failure on the part of information technology.

saw with great disappointment what was though to be a failure on the part of information technology.

As corporations worldwide began an intimate inner look at how to improve themselves, pioneers such as IBM began to develop a series of measures that coupled with a more adequate use of IT has yielded tremendous results. This pioneering work gave way to a new paradigm based on three critical shifts in the application of information technology.

The new IT enables enterprises to have 1) highperformance team structure, 2) to function as integrated business despite high business unit autonomy, and 3) to reach out and develop new relationships with external organizations.

All in all, these new trends are redefining the personcomputer relationship into a paradigm that basically lays the principles and foundation for what this relationship is all about: the human-centered computing paradigm.

The human centered computing paradigm has as a center theme that computers learn about their users and adapt to their needs. It is a shift from GUIs to VUIs (voice-driven user interface). VUIs will go beyond "light on/light off" scenarios into complex communication protocols. Figure 2 depicts the impact of VUIs in the human-centered paradigm and GIS.

Human-Centered Computing:

Make computers and software easier to use by having them conform to how you communicate, and not the other way around. The goals is to have users concentrate on the task at hand as opposed to operating the hardware/software.

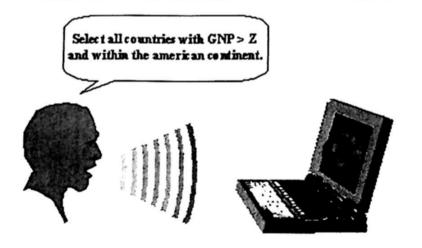


Figure 2. Human-Centered Computing.

GUIs from PARC and VUIs from concurrent and convergent technological shifts. Let us now examine the key technological shifts that will impact the evolution of IT in general and GIS in particular (shifts summarized from BusinessWeek, July, 2001):

3.2. SHIFT 1 - FROM TRADITIONAL SEMICONDUCTORS TO MICROPROCESSORS BASED SYSTEMS

The microprocessor, a computer on a chip, is at the center of this human centered computing paradigm. Traditional semiconductor technology, which twenty years and as recent as ten years ago, filled the cabinets of the mainframes and minicomputers in corporate data centers is being replaced by intertwined sets of microchips. Microprocessors are continually dominating leading edge computers (and "gizmos") of every size.

The microprocessor established the precondition for the new way of computing; it has changed everything we know about computing, the applications that are possible, computer architectures, and technology policy, not to mention the IT industry itself. Most important, the microprocessor enables empowered, distributed computing architectures to support the empowered, networked organization.

3.3. SHIFT 2 - FROM HOST BASED TO CLIENT/SERVER SYSTEMS

During the "first era of computing" devices interacted in a master/slave framework. Mainframes or minicomputers had slave terminals attached to them, typically dumb terminals with cryptic alphanumeric user interfaces. The enterprise also became populated with islands of personal computers, some attached to local area networks that enabled the simple sharing of files and devices like printers.

Now, because of the spectacular power of the microprocessor and the maturity of networking technology and standards, a fundamentally different style of computing has emerged, generally known as client/server computing, the approach coupled with Internet-based technology provides users to the access of a wide range of information, applications, and computing resources without worrying where they are or where they are connected.

Most importantly, software is processed not only on a host, but wherever it makes most sense. Software is not limited to one machine, but can be processed "cooperatively" on various computers across the network.

This client-server approach allows to integrate information across the organization. Client/server computing, like organizational empowerment, moves intelligence out into the enterprise. It enables powerful new multimedia applications that integrate data, text, voice, image, and video. This enables to have architectures that exploit the power of concurrent microprocessor technology.

For the years to come, the computing world will be a hybrid of very large servers (some kind of new generation of mainframes) and a new generation of mobile computing devices (more than just a PC as it is known today). Because mobile computing devices (specially wireless) offer clear advantages, such as speed, flexibility and enormous utility as a creative tool, these devices will allow to create "more intelligent computing" as they become part of a complex futuristic network.

3.4. SHIFT 3 - FROM VENDOR PROPRIETARY SOFTWARE TO OPEN SOFTWARE STANDARDS

One of the cornerstones for the success of the new paradigm in computing has been Open Systems, systems that are open to interconnectivity, interoperability and scalability. This shift has occurred to try to put an end to the "computing Tower of Babel". The open systems approach has laid the foundation for clear conceptual work and an understanding of the objectives pursued. Lets briefly define the various terms embraced by open systems of interconnectivity, interoperability and scalability.

Interconnectivity refers to hardware/software platform connectivity, i.e. all hardware/software connects and communicates to all hardware/software available in a user "transparent mode". In the client/server model, processing is distributed rather than centralized. Centrality if necessary can be created via "virtual" centers, centers that are distributed but perceived by the user as centralized.

Application *interoperability* refers to applications and computers from different vendors to work together on a network. This environment nurtures the portability of applications, data and people.

Application scalability allows the use of same application and system software on all classes of computers (from mobile devices to super computers).

The most important aspect of open systems is that it has the potential of shifting paradigms by actually changing in a more effective way how people do their work.

The move toward open systems can redefine workflows and changes people's work patterns through immediate and universal access to the information needed and processes to complete a given task. Consequently, decisions about migrating toward open systems require an organizational perspective in addition to the view of a technologist ("The Government Executive Open Systems Handbook", GT Publications Inc., 1993).

Open systems, based on industry standards not controlled by any one vendor, are transforming the computer industry and presenting a monumental challenge to commercial organizations alike. Standards are arising in all areas of computing, including communications, databases, user interfaces, operating systems, and software development tools.

Open systems result in information and software being portable, i.e. they can run on any hardware platform regardless of size or make. More important, standards also enable different systems to interoperate or communicate with each other.

Openness is not black or white. Standards and products have various degrees of openness, based on criteria such as vendor neutrality, platform availability, compliance with formal standards, and market penetration.

It is true that there are many competing standards, an oxymoron, indeed. But leading companies have concluded that adequate clarity exits to embrace the concept.

Research has showed that open systems have far-reaching advantages over the traditional approach. Standards are necessary to be effective in the new competitive business environment. Standards are required to link customers', suppliers', and competitors' systems. And standards are required to achieve integration of information. Standards in general and open systems in particular do not simply provide benefits. They are becoming imperative to create the kind of modular, flexible, powerful, networked computing architecture required by the new business environment (Tapscott 1993)

Standards are also transforming the computer industry by preventing the reinvention of wheels, as software suppliers include standard components and interfaces in their products.

3.5. SHIFT 4 - FROM SEPARATE DATA, TEXT, VOICE AND IMAGE TO MULTIMEDIA

In the early years of computing and even until recently, the immaturity of the various technologies involved and the absence of open standards meant that these four forms of information, data, text, voice and image, were separate and that separate technologies were needed to manage them.

Data processing systems handled numerical data. Word processing systems handled text. Telephone and dictation systems handled voice data. Photocopiers and microfilm systems handled image data. As the information contained in these systems becomes digitized, and as standards grow, the opportunity to integrate them is unfolding.

Multimedia computing is natural computing. These systems work the way people do, by integrating these forms of information into business processes and daily life.

3.6. SHIFT 5 - SOFTWARE DEVELOPMENT, FROM CRAFT TO FACTORY

Software has gone through a fundamental transformation. It has become an engineered profession using factory-of-the-future production techniques. Developers have moved towards an approach that enables the efficient reuse of software components (object oriented). Object-oriented computing allows programmers to create blocks of software called objects instead of large, complex, monolithic, tightly intertwined software programs. Objects are developed in standard ways and have standard behaviors and interfaces. These Lego-like pieces enable the rapid assembly of software rather than its laborious crafting.

With the advent of visual programming and computer aided software engineering, the software industry has become a parts industry, and through standards, software vendors create standard parts that enable customers to rapidly assemble computing environments.

This object-oriented paradigm has three components: object oriented analysis and design, object oriented programming and object-oriented databases. Of these three one and two are the best understood and applied. Object oriented databases are still conceptually and practically evolving as they attempt to encompass the relational model as well.

4. IMPACT OF TECHNOLOGICAL SHIFTS ON GIS

It has been over two decades since the first successful introductions of commercially available Geographic Information Systems (GIS). There has been many a contributor to make it happen, but none more so than people like Roger Tomlinson (the person who first coined the term "GIS") and Jack Dangermond (the person who led GIS to become a practical reality).

In conjunction with the Information Technology business, there have been significant advances in Geographic Information Systems since the term was first coined over three decades ago.

While GIS is certainly a subset of IT, the treatment of spatial data from a technology and organizational perspective has always been distinctly separate. For example, it is still quite common to find an Information Services Department and a GIS Department within any given organization. Since these early systems emphasized the importance of Geographic data with lesser importance given to the underlying information technology, we apply the designation <u>Gis</u>. Due to Gis' unique language, skills, and data processing, organizations have traditionally created user communities outside of the main IT domain. Fortunately, with the advent of newer technology and the acceptance of spatial information in the decision-making

process, this Gis and IT trend is merging. In more recent years, we've entered the GIS era where processing of spatial information has become less dependent on proprietary software and data structures and now embraces common mainstream IT architecture to the point that completes incorporation of spatial data processing into the IT realm is at hand, hence the designation "gIS".

From a historical perspective, the changes from Gis to GIS to gIS have occurred in roughly ten-year cycles. During the 1975 to the 1985 period, Gis was dominated by unique, costly, computer systems that offered little more that automated mapping capabilities. In fact, a significant achievement in this era was the simple capability of linking attribute data to line geometry was considered "state-of-the-art". People associated with the systems of this era were traditionally geographers in search of tools to perform their mapping task, which resulted in a tenuous relationship between the geoprocessing and traditional data processing groups.

During the 1985 to 1995 period, significant geoprocessing advances such as spatial analysis, attribute posting from relational databases, and high quality cartographic output from the systems evolved. However, these systems were still dominated by unique, proprietary software code and underlying databases, not to mention the high cost of obtaining and maintaining this technology. For these reasons, GIS was usually found only in organizations where the financial and human resources could support the system operation. Because of this esoteric nature of GIS technology in this period, full of spatial information incorporation into organization's decision-support process was never fully utilized.

With the dawn of major IT advances in the 1990's, such as powerful relational databases, miniaturization of the hardware platforms, and easy-to-use operating system, the business of geographic data processing began to evolve at a more rapid pace. The cost of ownership dropped dramatically and the number of system users rose in proportion. Incorporation of spatial information into the decision-support process became commonplace to the point that maps delineating the various geographic features are common communication tools. Because of this incorporation of mainstream IT tools into the GIS platform, this era is referred to as gIS.

All indications are the trend will continue to the point we enter an era referred to as gIs where the emphasis on *Information* is dominant and the underlying IT and spatial tools are a of secondary importance. There are a number of technical and social factors that support this assumption, with perhaps the most important being industry statistics.

The following graphic can summarize this evolution:



Figure 3. GIS evolution.

Gis depicts the era when geographic methods (in principle mainly cartography) were attempting to move to the computer, automated domain. During that period the computer industry, and in particular database technology (the relational model) as maturing and preparing for a paradigm shift.

GIS depicts the convergence of mature information systems technology with geography. Hence geographic information systems start to become a practical technology (a tool if you will). During this period GIS' systems approach is geocentric —everything revolves around the GIS.

gIS defines an era where it is recognized that geography as important as it may be, is part of a larger whole. The focus is on the information system and this establishes a GIS implementation approach that is more system-centric.

gIs defines the beginning of an era that is information-centric. The essence of having a GIS or any information system is to have information easily and readily accessible. The effort is focused in data integration, data warehousing, data mining —the conversion of data into information. I call this "the power of I". The following diagram illustrates the concept.

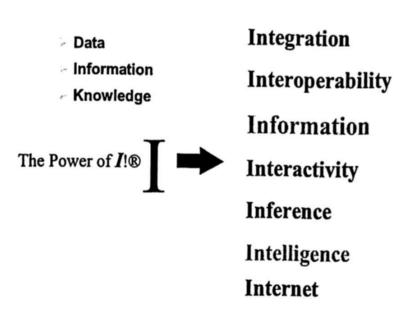


Figure 4. Different elements to integrate "I".

4.1. THE SPATIAL ENABLING OF INFORMATION®

We can summarize that the three forces that have driven recent information technological shifts are:

 Miniaturization – electronics getting smaller and more powerful

- 2. Digitization all data sources are now digital
- 3. Convergence force 1 and 2 coming together

These three forces have allowed to move from a data-centric world, to an information-centric world. As such, it is also the case of GIS, moving from a geocentric systems approach, to an information centric systems approach. Information Technology with the Internet and increasingly high-bandwidth wireless communications has become the backbone of the new world economy. Software-hardware-communications. Data-information-decisions. They all have become interlinked to create a new world order and as such, new work-flow paradigms and tools are required to effectively integrate and transform data into useful information to make decisions.

With the maturing of geospatial technologies such as GIS, GPS, Digital Photogrammetry, and Remote Sensing, locational facts have now become a key component of the decision-making process.

Beyond system to system integration, the embedding of locational facts requires the design, development, and implementation of new solutions that integrate geospatial, information and communications technology aimed at optimizing process work flows, improve the decision making process and reduce operating costs in an innovative way. This integration process is defined as the "spatial enabling of information®".

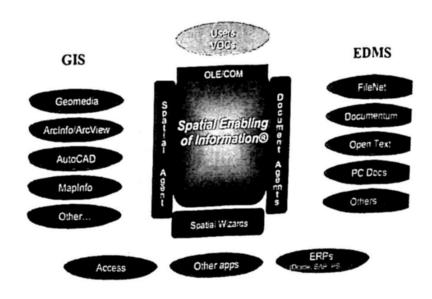


Figure 5. Spatial Enabling of Information.

Coupled with these mature geospatial technologies, the advent of high-resolution imagery and new -more "intelligent"- technologies for the automation of pattern recognition and feature extraction, based on expert systems and object oriented technology, paves the way for a work-flow paradigm shift in the automated process of transforming data into information.

The future GIS tools will become increasingly transparent to the user much like the IT business in general. Furthermore, it will have the capability to mold around the user's workflow instead of placing rigorous restraints on operations.

5. CONCLUSIONS

Visionaries, those that always ask "why not?".

In the quest for making a better world, indeed GIS has played and will continue to play a major role as a data integration tool. But beyond the tool, within an information centric view, GIS has provided the underlying bedrock to support a new emerging paradigm known as the "spatial enabling of information". This paradigm leverages more than anything does the force of convergence, the power of I, the power of transforming data into information readily and rapidly accessible to make decisions with an underlying spatial context.

In this paper we have outlined several technological shifts that have acted as forces in shaping the GIS of the future. GIS has matured in the far speeding highway of information technology. Its maturity lies in making this powerful integrating spatial technology "transparent" to the end user. The technological shifts presented have paved the way for GIS to enter into the next level of acceptance, as part of spatially oriented commodity products (the spatial enabling of information paradigm in action), in for example, a scenario like this:

"The airplane is about to land into an unknown city to me and it is early afternoon. I grab my handheld WCC (wireless computing and communication device) and "tell it" that I am hungry and want to have some sushi. That afterwards I want to see a movie near by, an action type, and then I want to go to my hotel. I sign off from the device. In the meantime my WCC has automatically "connected itself" to the local restaurant and cinema database and found out what I requested. Minutes later as I get my rental car I plug in my WCC into the car's WCC port.

The car gets driving directions from my WCC. The car hops into the electronic highway and takes me to have some really good sushi, "it" has bought tickets and drives my car to a cinema to see an action packed movie. Afterwards it instructs my car how to take me to my hotel for a good night sleep. No need to check-in, the WCC has already made the arrangements."

Underneath all of this is a technology called GIS. For all I know it could stand for great and Imaginative stuff.

REFERENCES

D. Tapscott, Paradigm Shift: How Information Technology is Reinventing the Enterprise, *Business Week*, 1993.

Datamation, Are You Ready for the DeskTop of the Future? (Ed.: Cahners Publishing Co., 1994).

- G. Gilder, Life After Television: The Coming Transformation of Media and American Life (Ed.: W.W. Norton & Co., New York, 1993).
- M. Hammer, & J. Champy, James, Reengineering the corporation (Ed.: Harper Business, 1993).
- D. Kearns & D. Nadler, Prophets in the Dark How Xerox Reinvented itself and beat back the Japanese (Ed.: Harper Business, 1992).
- D. Tapscott & A. Carson, *Paradigm Shift, the New Promise of Information Technology* (Ed.: DMR Group, McGraw Hill, 1993).

GIS APPLICATIONS

DISPERSED BY WIND PROJECT: A MOTION ANALYSIS OVER WIND FIELDS

Ángel M. FELICÍSIMO
Ingeniería Cartográfica, Geodesia y Fotogrametría
Escuela Politécnica
10071 Cáceres SPAIN
amfeli@unex.es

Jesús MUÑOZ Real Jardín Botánico Plaza de Murillo, 2 28014 Madrid SPAIN jmunoz@ma-rjb.csic.es

ABSTRACT

Remote sensing data in which each pixel stores vectorial information represent a challenge for most image processing and GIS applications. For an ongoing project which aim is to test if wind is the dispersal vehicle of several plant groups in the Southern Hemisphere, we have used scatterometer wind data (QuikSCAT) to generate anisotropic cost-distance surfaces. Because of the vectorial nature of the QuikSCAT raw data, we found several problems in the data processing and implementation within GIS. The aim of this paper is to present the problems we found and to propose some solutions. Main problems were related to (1) lack of complete daily coverage, (2) lack of circular statistical tools within GIS packages, (3) reprojection of vectorial data, (4) maintenance of spatial continuity, and (5) restoration of direction values after reprojection of vectorial data. The emphasis of this study is to present the solutions we took for each of the previous problems, which in the end allowed generating the costdistance surfaces with commercial GIS after programming some of the operations.

KEY WORDS

GIS, anisotropic cost-distance, vectorial data management, long-distance dispersal

1 INTRODUCTION

Hypothesis of difficult experimental demonstration are raised in some, if not many, scientific fields. Biological sciences are not an exception, and there are theories which foundations are more intuitions o beliefs than facts empirically checkable. The descriptive nature of such theories derives in many cases of the lack of instruments or technology to check the hypothesis behind. Fortunately, the oncoming technology occasionally allows testing such hypothesis to accept, reject or revise them.

As an example, the wind as been invoked to explain the strong floristic affinities shared among landmasses in the Southern Hemisphere. Long-distance dispersal by wind is favored by some authors (e.g. [1] for bryophytes, [2] for pteridophytes) and rejected by others who consider that present distributions in the Southern Hemisphere rather reflect old Gondwanan connections back to the Cretaceous (vicariance hypothesis, e.g. [3] for a general view, [4] for bryophytes).

Consequently, we can see that although wind is periodically presented as a dispersion vehicle, its role is not widely accepted perhaps because experimental confirmation was difficult. Certainly, the lack of reliable data on wind azimuth and speed at Earth scale hampered the empirical validation of the wind hypothesis until recently, when satellite scatterometers started to provide daily near-global coverage on wind properties with high statistical accuracy.

In this context, the general aim of Dispersed by Wind Project (DWP) is to test the hypothesis of long-distance dispersal by wind of several cryptogamic plant groups in the Southern Hemisphere (mosses, pteridophytes, lichens and liverworts). To this end, we have used data from the SeaWinds scatterometer (on board the QuikSCAT satellite) to model wind connectivity among landmasses and to test how well these "wind highways" explain the observed floristic similarities. In synthesis, we interpreted DWP as an example of multidisciplinary studies joining remote sensing, GIS and statistical analyses to model and check a biological hypothesis, a new and promising way in biological studies.

2 OBJECTIVES

The general aim of DWP is to test the hypothesis of longdistance dispersal by wind of several cryptogamic plant groups in the Southern Hemisphere. In order to reach this objective, we checked the degree of association between the floristic similarity index and wind connectivity 'similarity' matrices.

Consequently, the core of the DWP consists in verifying if the "winds paths" can explain similarities in flora composition. In the present work we will take care of the

problems related to GIS and remote sensing processing, a set of questions that must be outstanding because they reflects pitfalls which solution would highly improve GIS operability.

3 MATERIALS AND DATA

3.1 THE QUIKSCAT DATA

Most data currently used to monitor the biosphere derive from Earth-orbiting satellites. A large part of those data are commercial products managed by private or

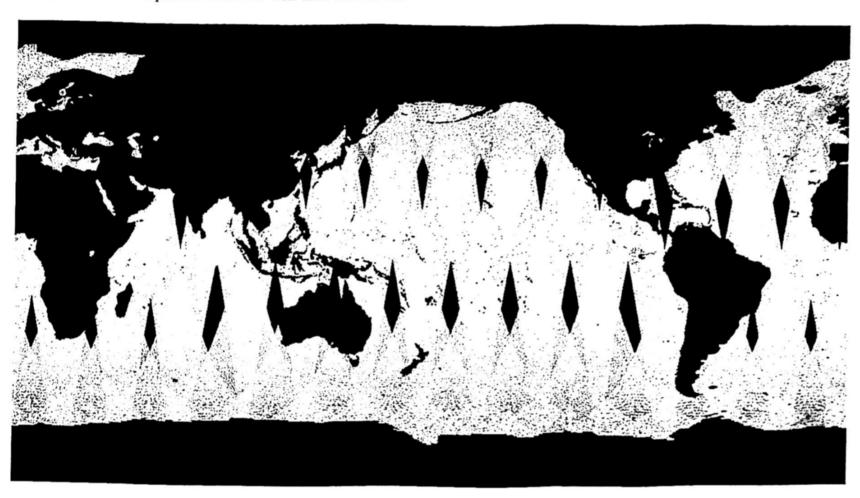


Figure 1. The image shows a typical coverage for one day of SeaWINDS on QuikSCAT Level 3 data.

governmental agencies (SPOT Image, European Space Agency, NASA, etc.). Others are freely available, and are an invaluable source of information for research at continental or global scale, as the budget needed for such studies using commercial products would be beyond most research projects budgets. Examples of studies benefited from such free datasets are those listed in [5], or in the World Climate Research Program, which long-term data sets are generated from observations of operational satellites [6]. Further studies using such sensors are [7] with SeaWiFS¹ or [8] with MODIS².

The SeaWinds on QuikSCAT scatterometer was developed by NASA JPL to measure the speed and direction of ocean surface winds. The QuikSCAT satellite

was launched into a sun-synchronous, 803-kilometer, circular orbit on 19 June 1999. The SeaWinds instrument is a specialized microwave radar (scatterometer) that measures near-surface wind speed and direction under all weather and cloud conditions over Earth's oceans [9].

All QuikSCAT standard products are in Hierarchical Data Format (HDF). HDF is a data format developed by the National Center for Supercomputing Applications (NCSA) at the University of Illinois³. The QuikSCAT HDF files contain a set of 16 Scientific Data Sets (SDS). An HDF SDS is a fixed dimensional array. The QuikSCAT SDS are a simple 0.25° rectangular grid that contains 1440 pixels from east to west and 720 pixels from south to north (JPL, 2001b) in a standard "plate carrée" projection (Figure 1).

http://seawifs.gsfc.nasa.gov/SEAWIFS.html

² http://modis.gsfc.nasa.gov/

http://hdf.ncsa.uiuc.edu

The data used in our study include daily data from June 1999 to March 2003, and were obtained from PO-DAAC (Physical Oceanography Distributed Active Archive Center).

3.2 BIOLOGICAL DATA

For DWP we have recorded distributions of mosses, liverworts, lichens, and pteridophytes for 27 localities in the Southern Hemisphere, including oceanic islands and continental areas (Figure 2).

The recorded data are four data sets: mosses (601 species), liverworts (461), lichens (597), and pteridophytes (192), totaling 1851 species. Only species shared by at least two localities were included.

3.3 SOFTWARE

Software includes IDL 5.4 (Research Systems, Inc., Boulder, CO, USA), ArcInfo 7.1.2 and ArcView 3.2 (Environmental Systems Research Institute, Redlands, CA, USA). We must say that there exist many software applications to work with remote sensing and, more generally, spatial data. At present, the limits between remote image processing applications and geographic information systems are more and more blurring. The

types of software applications, very different in their early days, have suffered a "convergent evolution" and presently they share many of the previously exclusive tools.

All current applications allow working with conventional remote sensing images, in which each pixel stores a reflectance value in a fixed wavelength interval. Data from active sensors (radar) also can be treated with a more restricted set of applications. All are scalar data, in which each pixel has assigned a single integer (digital value).

In the last few years we have assisted to the bloom of a new generation of data essentially different form the previous. In them each pixel stores information about a vector, not on a scalar as in the previous, general case. We will see how this type of data generates problems which solutions are not easy through current image processing applications and geographic information systems. It is not our aim to pillory GIS packages, but rather to highlight some pitfalls which solution would highly improve GIS operability, especially by incorporating tools for vectorial data analysis. To this end we will point out some problems we experienced and the solutions we adopted in the progress of the DWP, a project that uses wind vectorial data.



Figure 2. Study areas included in the Dispersed by Wind Project.

need to work simultaneously with remote sensing images and digital cartographic information has caused that both

4 METHODOLOGY

Statistical methods are based in the comparison of similarity/distance matrices. Our aim is to check the degree of association between floristic similarity and wind connectivity matrices.

4.1 DISTANCES O SIMILARITY MATRICES

4.1.1 Floristic similarity

Floristic similarity between each locality pair was calculated using the Ochiai index [10]. This index is based in concordance of species presence/absence and varies from 0 (no similarity) to 1 (same floristic composition). The results are four 27x27 floristic similarity matrices, one for taxonomic group.

4.1.2 Wind distances or connectivity

Anisotropic Cost Analysis (ACA) is a procedure to calculate the minimum accumulative-travel cost from a source to each cell location on a grid or raster model. The analysis calculates the accumulative cost over the field of wind vectors, considering such field as a friction surface. ACA is based in the measured wind-vectors fields. Connectivity extent between two points depends upon two factors: dispersion of the vector set, and the value of their modules (wind speed). Two well connected points will be joined by a vector chain without sinuosities. If those vectors have large modules the cost-path will be low.

ACA tools allow computing cost values from georreferenced data including the coordinates of the origin and end points, and the direction and module of the intervening vectors. We also should define the angular tolerance allowed as deviation from the main vector direction. This function assigns a minimum cost for the movement in the exact vector azimuth, and incrementally penalizes angular deviations, generating less strict models. In order to perform the anisotropic cost analysis, some conditions must be fulfilled, namely spatial topologic continuity, as well as distance and azimuth correctness. The result will be a surface in which the value of each pixel equals the cost of connecting it with the pixel origin. In our study, the above conditions were not fulfilled by the original data, and the GIS process were not negligible, as it is exposed next (Section 0).

Anisotropic cost analysis is not available in all GIS packages. Among common ones, only ArcInfo 7 (ESRI Inc., Redlans, CA, USA), and Idrisi 32 (ClarkLabs, Worcester, MA, USA) have this tool. Surprisingly, the tool is not implemented in ArcInfo 8 (formerly ArcGIS). ArcInfo tool (costpath command) is considerably more complete that its counterpart in Idrisi (varcost command), and this accounts for its use in the DWP.

To run the cost-path analysis requires (1) to determine the origin, namely, in DWP, each of the 27 sample points, (2) to define the wind field based in wind azimuth and speed, where azimuth defines minimum-cost direction and speed is considered the inverse of resistance, and (3) to define the angular tolerance, or the increment of cost corresponding to incremental deviations from the minimum-cost azimuth. As an example of the cost-path analysis, Figure 3 shows the results of the analysis for the April 1-10, 2000 period for Bouvet Island (54°26'S, 3°24'E).



Figure 3. Anisotropic cost surface from Bouvet Island (arrow) for period April 1-10, 2000 (increasing cost from white to dark in arbitrary units).

4.2 WIND CONNECTIVITY PROCESSING

Scalar data are simple data measured in a linear scale (e.g. reflectance values in a remote sensing image), and can be computed by conventional statistical methods. However, QuikSCAT pixels store vectorial information. A vector is a structure that combines both metric and topologic properties. Metric properties are the azimuth or angle respect a reference system, and the module or vector length. The topologic property is the vector direction (from-to). Under the convention adopted for QuikSCAT data, a wind direction of 0° implies a flow toward (not from) the North. The QuikSCAT vectors are stored as their u and v components in separate SDS of the HDF file.

While processing the original wind data to generate the cost surfaces needed to compare with the floristic similarities we found several problems:

4.2.1 First problem: Lack of daily data and circular statistics

QuikSCAT daily cover is not complete. Figure 1 shows empty areas result of the orbital trajectory. These areas must be removed because they break the spatial continuity necessary for anisotropic-path analysis. As those empty areas change daily, the adopted solution was to generate 10-day synthesis. This solution guarantees that each pixel is covered by several measurements, commonly between 5 and 7.

Most GIS software packages implement map-algebra calculations. However, wind-vectors cannot be analyzed throughout conventional statistics because they are distributed in a circular scale. A simple vectorial mean cannot be calculated straightforward, but the process must be decomposed into several elementary steps to use the appropriate circular data statistics [11, 12].

For example, it is not possible to calculate the average wind azimuth directly, but operating with u and v components, which are themselves linear variables. Decomposition of a vector with module S and azimuth γ is simple: $u = S \cdot \sin \gamma$ and $v = S \cdot \cos \gamma$.

Although such decomposition is a common procedure in circular statistics, it is not included in any GIS package, and it must be programmed. The average azimuth for a given dataset is then calculated from those u and v components. The first step involves calculating the arithmetic mean of the u and v components separately for each pixel using the 10-day SDS syntheses. The azimuth is calculated later with the expression $\gamma = \arctan(\bar{v}/\bar{u})$. Special care deserve pixels where u=0 and u=v=0 (no wind). Average wind speed follows the expression: $S = sqrt(u \cdot u + v \cdot v)$.



Figure 4. Three-day sequence of satellite data over Australasia showing the displacement of empty areas. 10-day syntheses guarantee complete area coverage.

4.2.2 Second problem: Projection properties and spatial continuity

QuikSCAT data could not be used directly in our project because the original simple cylindrical projection ('plate carrée') breaks the spatial continuity at Greenwich meridian. Under such conditions, two points next to each other but separated by Greenwich meridian would be, in practice, unconnected. The best solution would be also the simplest: to work directly with unprojected data in spherical coordinates (latitude/longitude). However, GIS packages only allow spatial analysis of data projected in a planar coordinate system. This is a severe pitfall of GIS packages that should be addressed.

The adopted solution was to project the data to Universal Polar Stereographic (UPS), a planar perspective projection where one pole is viewed from the other pole. The tangent point was the South Pole. Under such conditions this is a conformal projection with true-direction values from the South Pole. Figure represents average wind speed in the study area after transformation to UPS projection, which guarantees continuity over Greenwich meridian. A clear but thin belt of strong winds (pale tones) surrounds Antarctica, after which a wide calm belt (dark) is evident.



Figure 5. Example of SeaWinds on QuikSCAT data in UPS projection. Variable is average wind speed (increasing speed from dark to pale: 0-30 m s⁻¹).

4.2.3 Third problem: azimuth values restoration under UPS projection

Transformation from "plate carrée" to UPS projection guarantee spatial continuity, but it does not transform azimuth data to the new projection. This is another pitfall of GIS packages: raster data are projected without considering that some data are projection-dependent. For example, a pixel with azimuth value of 45 (wind towards the North-East) will keep this value after projection. Once in UPS projection this value will be correct only if the pixel location is exactly in the upper half of the

Greenwich meridian of Figure 5, otherwise information will be wrong.

In DWP we have developed a procedure that solves the above problems in the particular case of UPS projection. It takes two steps (Figure 6).

1. We define a new global Cartesian system in which the origin of coordinates (0,0) is the South Pole. Next we calculate the azimuth of the vector pointing towards the true North in the new system. The azimuth of a pixel i with coordinates [x(i),y(i)] in the new system follows the expression: true North azimuth = $\arctan [y(i)/x(i)]$. This true North azimuth is the deviation generated in the projection process, and will be employed to calculate the azimuth in UPS projection. Depending on the place of the pixel on the projected space, its true North azimuth will be distributed in the 0-360° range.

2. Denoting "true azimuth" to the azimuth values in the original "plate carrée" projection, calculation of azimuth in the UPS projection will follow the expression:

UPS azimuth = (true North azimuth + true azimuth) · modulus 360

This solution is not general but specific for UPS projection. However, it shows a problem that only can be solved after an important modification of projection routines, considering the vectorial or scalar nature of the raster data.

Table 1 (see end of paper) shows the summary of operations performed from SeaWinds HDF data reading to calculation of cost-path. Although the programming language is indicated, the processes are independent of software applications; they are shown here just as an example.

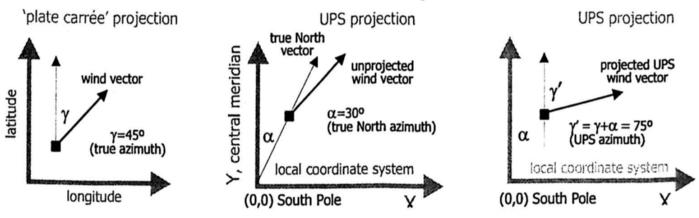


Figure 6. Calculation of azimuth vector in UPS projection.

5 CONCLUSION

In this paper we raise some questions arisen during DWP progress that are common to any work with vectorial data. Encountered problems are not negligible, because vectors appear often as spatial datasets (e.g.: aspect, dip in geological maps, animal dispersal movements, or sea currents, besides wind analysis as in the present study). Analysis with such data is the field of GIS. However, it seems evident that management and analysis of vectorial data with the available tools can cause troubles.

A downside is the impossibility of working with bidimensional (latitude, longitude) or three-dimensional (latitude, longitude, distance from origin) spherical coordinates. Occasionally, a GIS can represent the data from spherical coordinates, but the analyses must be done on projected data. This is not a severe problem when the study area is small. However, in global scale studies distortion in metric properties or rupture of topologic properties can prevent some analyses.

Further shortcomings come from the nature of the data employed in the DWP. Generally, GIS are not designed to manage vectorial data, and it will be necessary programming of operations that, for scalar data, would be straightforward. For example, an arithmetic mean or standard deviation of a dataset would be calculated as a mean direction and a circular standard deviation [12]. Their formulation is simple, but such calculation tools are not included in any GIS package.

Finally, in current GIS to project data means to change the location of spatial elements while maintains constant theirs attribute values. This is reasonable with scalar data (temperature, reflectance, etc.), but it produces severe errors when treating azimuths or other directional values, because the reference system changes.

How to explain the lack of literature on this type of problems? There is a wealth of publications on wind data, but they mainly deal with data capture algorithms [5, 15], sensor calibration and influential factors [16], quality data analysis [17], or the foreseeable impact of the new data in previous climatic models [18]. Perhaps the limited use of such data outside those fields resides [13] in that they are available only recently and in the prevalence of local over global studies. Even some of such global studies does not experience the problems found in the DWP because they are not affected by metric or topological constraints. For example, Chen [19] studied the results of models generated from QuikSCAT and TOPEX/POSEIDON

(altimetry) data over a large Pacific Ocean extension. In this case the only limitation was that the data were superimposable, and consequently he used the original "plate carrée" projection. The same happens in [20] studying sea surface temperature or [21] comparing temperature and QuikSCAT data over the tropics.

In the Dispersion by Wind Project we have performed analysis of vectorial data through a commercial GIS. Although the adopted solutions are appropriate for our project, we would like to attract developers' attention to improve GIS tools. Otherwise, management as well as analyses and operations with vectorial data, increasingly common with the launch of new satellites, can result prevented.

ACKNOWLEDGEMENT

J. Muñoz thanks M. T. Tellería, Director of the Royal Botanic Garden at Madrid, for unconditional support.

REFERENCES

- [1] B.O. van Zanten & T. Pócs, Distribution and dispersal of Bryophytes. *Advances Bryol.*, 1, 1981, 479-562.
- [2] P.J. Brownsey, New Zealand's pteridophyte flora plants of ancient lineage but recent arrival? *Brittonia*, 53, 2001, 284-303.
- [3] C.J. Humphries & L.R. Parenti, Cladistic biogeography: Interpreting patterns of plant and animal distributions (Oxford: Oxford University Press. 1999).
- [4] W. Frey, M. Stech & K. Meissner, Chloroplast DNA-relationship in palaeoaustral Lopidium concinnun (Hypopterigiaceae, Musci). An example of stenoevolution in mosses. Studies in austral temperate rain forest bryophytes 2. *Plant Systematics and Evolution*, 218, 1999, 67-75.
- [5] F.J. Wentz & M. Schabel, Precise climate monitoring using complementary satellite data sets. *Nature*, 403, 2000, 414-416.
- [6] G. Ohring & A. Gruber, Climate monitoring from operational satellites: accomplishments, problems, and prospects. *Advances in Spatial Research*, 28, 2001, 207-219.

- [7] M.J. Behrenfeld, et al., Biospheric Primery Production during an ENSO transition. *Science*, 291, 2001, 2594-2597.
- [8] C.O. Justice, et al., An overview of MODIS Land data processing and product status. Remote Sensing of Environment, in press, 2002.
- [9] J.P. Laboratory, QuikSCAT Science Data Product. User's Manual. Version 2.2, JPL Document 18053, California Institute of Technology, CA.
- [10] P. Legendre & L. Legendre, *Numerical Ecology* (Amsterdam: Elsevier, 1998).
- [11] K.V. Mardia, Statistics of directional data (New York, USA: Academic Press. 1972).
- [12] N.I. Fisher, Statistical analysis of circular data (Cambridge, UK: Cambridge University Press. 1993).
- [13] J.F. Commandeur & W.J. Heiser, Mathematical dérivations in the proximity scaling (PROXSCAL) of symmetric data matrices, Technical Report RR-93-03, Leiden University. Leiden, NL.
- [14] D.A. Jackson, PROTEST: a PROcrustean Randomization TEST of community environment concordance. *Ecoscience*, 2(3), 1995, 297-303.
- [15] D.-J. Kim & W.M. Moon, Estimation of sea surface wind vector using RADARSAT data. Remote Sensing of Environment, 80, 2002, 55-64.
- [16] W.T. Liu, Progress in scatterometer application. Journal of Oceanography, 58, 2002, 121-136.
- [17] A. Stoffelen, M. Portabella & A. Voorrips. Towards the assimilation of QuikSCAT winds. *Fifth International Winds Workshop Proceedings*. Lorne, Australia. 2000.
- [18] Y. Tahara. The preliminary study of the impact of QuikSCAT/SeaWinds ocean surface wind data to the JMA global model. *Fifth International Winds Workshop Proceedings*. Lorne, Australia. 2000.
- [19] D. Chen, Applying satellite remote sensing to predicting 1999-2000 La Niña. Remote Sensing of Environment, 77, 2001, 275-278.
- [20] F.J. Wentz, et al., Satellite measurements of sea surface temperature through clouds. *Science*, 288, 2001, 847-850.
- [21] C. Day, Two satellites observe details of ocean-atmosphere coupling. *Physics Today*, 56, 2000, 23-24.

USING GIS TO ASSESS THE SPATIAL DISTRIBUTION OF MEXICO CITY EMISSIONS INVENTORY

Elba ORTIZ, Eliot RAMÍREZ, Arturo RODRÍGUEZ, Eduardo VILLASEÑOR, Roxana TOLEDO Mexican Petroleum Institute, PIMAYC (IMP)

Eje Central Lázaro Cárdenas #152, Col. San Bartola Atepehuacan

C.P. 07738, México, D.F., MEXICO

Tel. (5255)3003-6000; Fax (5255)3003-8000

eortiz@imp.mx

ABSTRACT

The major contribution to the atmospheric contamination is due to the combustion of the oil fuels. The problem consists of locating the emission sources and the type of contaminants produced at the Mexico City. We propose an emissions inventory (EI) instrument to know the emissions by means of a spatial distribution. In the present work, the spatial distribution of the EI 1994 is performed, using a Geographical Information System (GIS). Also, we include the simulation grid for the area and industrial sources at the municipal and sub-municipal levels. The information for the EI has been provided by the National Ecology Institute (INE). The database contains 5019 registers of the several companies. The information consists of the mobile sources, traffic administration data and the total consumption values of gasoline and gas lp (liquefied from petroleum) from PEMEX-GAS project data. For the near 30 000 area sources involving commerce, service of business, population and housing the EI evaluation memoirs for Mexico City and the National Institute for Geography and Informatics (INEGI) data has been considered. The Information processing lead to define the zones of the Mexico City, which present the largest population, the major emission sources, the higher energy consumption, and the bigger contaminants produced in Metropolitan area in 1994.

KEY WORDS

Emissions Inventory, GIS, MCMA, INE, INEGI

1. INTRODUCTION

The relationship between the political, economical, and social processes involved in the Mexico City Metropolitan Area (MCMA) has growth with the time in different human activities in this metropolis. Figure 1 shows the zone to be studied (MCMA), which covers 2,396 km². In this area live more than 18 million people, with a fleet close to 3.5 million vehicles, around 35,000 industries and service establishments.

To support these activities, more than 44 million liters of fuel are burned daily. The MCMA is settled in a basin at 2,240 m (7340 ft) ASL, with predominantly East to West wind direction.



Figure 1. Location of study area: Mexico City Metropolitan Area.

As a result from that fuel consumption, a great number of pollutants are emitted to the atmosphere, and the particular topography of the basin. Mountains up to 1500m above ground level in its southern and western borders cause the accumulation of them. The photochemical reactions driven by ultraviolet (UV) radiation from the sun allows to generate the smog.

It is very important to identify where the emission sources are located and how kinds of pollutants are the most emitted in MCMA. A strategic tool to assess that is the EI, published by the National Ecology Institute (INE) and the Environmental Metropolitan Committee (CAM), which is a compilation of the amounts of gases (Carbon monoxide, CO; Sulfur dioxide, SO₂; Nitrogen Oxides, NOx) and pollutants (Total of suspended particles and hydrocarbons) emitted by industries, commerce and service establishments, houses, vehicles, aircraft, soil and vegetation among others.

2. OBJECTIVES

The aim of this work consists of the systematization of the information using EI for every area and point sources. EI allows to the decision-makers and researchers to analyze quantitatively the distribution of pollutant sources, their emissions and its relationship in the MCMA.

3. METHODOLOGY

The available information is used to carry out this work; the PREDATAIND94 database for industry has been developed from a survey conducted by the environmental authorities. In this case, INE and the governments of Distrito Federal and the Mexico State (PROAIRE, 1995). To validate the data, several industries contain bad coordinates or they are located outside of the limits of MCMA. Therefore, they have been excluded. Also these industries did not report the energy consumption and the emissions have been put away.

The computation procedures for the EI of the area sources, ozone precursors and carbon monoxide; population, and housing data; and the total values of the emissions of gas LP have been estimated by PEMEX-GAS and PETROQUÍMICA BÁSICA (1997).

For the data integration, the software ArcView v.3.1 has been used (ESRI, 1998). It had a great utility to analyze and solve problems related to the scale of the spatial problems at MCMA. Using this application, it is easier to make spatial analysis, to obtain the data, to handle the information and to represent it by means of maps. (SENDRA, 1991). This system brings to this work, a powerful tool to handle the spatial data manage for specific areas and point sources, which integrate different databases in a single set of data.

For the database design the entity-relationship model has been used, in which delegations and municipalities form one entity and the attributes of each entity corresponds to the industry and area data. A unique identifier often expressed in a numeric value represents the link between the entity and the data. A specially designed program, SIE-ROX, is used to make the consistency of the data, the joining process is considered by Dbase (DBF) format or texts (TXT). Once the consistency of the data are obtained by the analysis process, which is made into ArcView. The most databases came from economic surveys realized by the INEGI (1999), relating housing, expenditures, services and population by geographic units.

4. RESULTS

4.1. INPUT DATABASE FOR GIS IN POINT SOURCES

Spatial analysis for industry is carried out from 4,135 valid records. The information that has been used for the EI is provided by the INE database, which contains 5019 registered enterprises. Despite this information in some special cases does not present valid records that can be able to define the total emissions in comparison to PROAIRE document, e.g., thermoelectric plants.

4.2. INPUT DATABASE FOR GIS INTO THE POINT SOURCES

Spatial analyses for these sources have been performed at delegation and municipality level. Another database was designed at the scale of simulation grid.

4.3. ADDITIONAL DATABASE

Several files with population and its density and the service establishments in MCMA have been built to gather them in the GIS. The highway and communications layers are used to explain the industry and services locations provided by the Mexican Petroleum Institute (IMP).

The Federal District concentrates the most industries and workers. However, it uses less part of the energy and produces less than the half of total emissions for MCMA. For surrounding municipalities is the most of the emissions and energy consumption. When the total emissions are desegregated, the municipalities contribute with the most of the TSP, SO₂, CO, y NO_x, while only in the Federal District the emissions for HC surpass these in the municipalities.

In general, for NO_x, the energy generation industry counts on the most of the emissions. Printing companies represent the greatest contribution for HC. Carbon monoxide is presented in everybody except for metallurgic industry as well as TSP. The SO₂ gas is concentrated in mineral-metallic industry, meals, wood and timber, vegetables and animals processing and dressing industries.

Figure 2 shows a punctual distribution of the 4135 industries by municipalities and they follow a pattern as given by roads and main avenues. Figure 3 shows the spatial distribution of the total emissions according to the simulation grid has been made. This grid feeds an air quality simulation model and contains its origin at coordinates 2110000 N and 450000 E, Zone 14 UTM system. Each cell has 5 Km of the length.

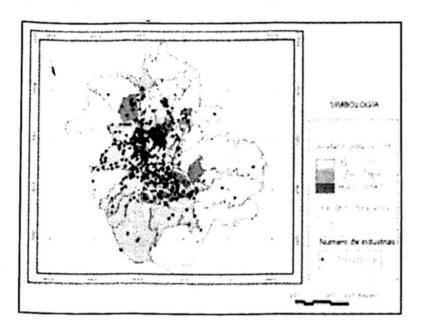


Figure 2. Spatial distribution of industry in MCMA, 1994.

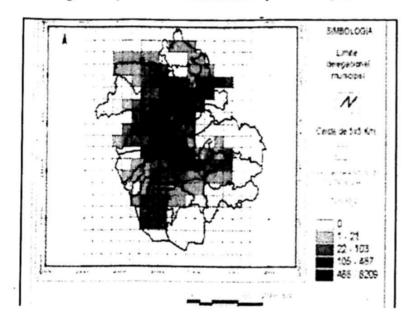


Figure 3. Gridded total emissions of the industry distribution, 1994.

5. CONCLUSIONS

The characterization of a particular region is cartographically represented by common features such as population density, number of sources, energy consumption, and emission generation. They will provide a better planning for measures, which will reduce the emissions, as an emissions inventory makes. The difference is that the map will predict the behavior of pollutants as a variable depending on social, political and economical variables. Thus, the information is useful as input for air quality models.

Using the EI systematization for point sources, the information of 4135 industries have been integrated. These industries count on the 3% of total emissions in

1994. When data have been validated, the found error is 6% with respect to data published by PROAIRE.

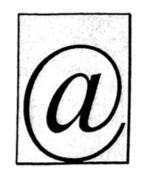
REFERENCES

- [1] IMP, Comisión Ambiental Metropolitana (CAM), Proyecto para el diseño de una estrategia integral en la gestión de la calidad del aire en el valle de México 2001 2010, (México, D.F., 2000).
- [2] MIT, Integrated Program on Urban, Regional, and Global Air Pollution. (USA, 2000).
- [3] Comisión Ambiental Metropolitana (CAM). Inventario de emisiones a la atmósfera de la Zona Metropolitana del Valle de México 1996, (México, D.F., 1999).
- [4] Departamento del Distrito Federal (DDF), Gobierno del Estado de México (GEM), Secretaría del Medio Ambiente, Recursos Naturales y Pesca (SEMARNAP), Secretaría de Salud (SS), Programa para Mejorar la Calidad del Aire en el Valle de México (PROAIRE) 1995 2000, (México, D.F., 1996).
- [5] Instituto Nacional de Estadística Geografía e Informática (INEGI); GDF, Estadísticas del medio Ambiente del Distrito Federal y Zona Metropolitana 1999, (México, D.F., 1999).
- [6] Instituto Mexicano del Petróleo (IMP), Sistemas de Información geográfica aplicados a la información cartográfica del proyecto IMADA, Instituto Mexicano del Petróleo, GCA-097085, (México, D.F., 1997).
- [7] E. Ortiz, Modelación del efecto en la calidad del aire en la Zona Metropolitana de la Ciudad de México, (Informe GCA 98-002, (México, D.F.: Instituto Mexicano del Petróleo, 1998).
- [8] PEMEX-GAS Y PETROQUÍMICA BÁSICA, Efecto de los Componentes del Gas Licuado de Petróleo en la Acumulación de Ozono en la Atmósfera de la Zona Metropolitana de la Ciudad de México, (México, D.F.: Memoria técnica del estudio, 1997).
- [9] Secretaría de Energía (SE), Balance Nacional de Energía 1996, (México, D.F., 1997).
- [10] Secretaría del Medio Ambiente (SMA), CAM, Hacia un programa de calidad del aire 2000 2010, (México, D.F.: Comisión Ambiental Metropolitana (CAM), 2000).
- [11] SEMARNAP Y RADIAN, Mexico Emissions Inventory Program Manuals. Vol. V. Area source inventory development. Vol. VI Mobile source inventory development, 1997.

oracida como de la coma acada, en contenedo, acongrazada

en en de deux en en el la completa proposición de la completa del completa de la completa de la completa del completa de la completa del la completa del la completa de la completa de la completa de la completa de la completa del la completa de la completa de la completa de la completa del la compl

species and process of the constraint



AUTHOR INDEX

		Page			Page
A			\mathbf{H}		
,	Alexandrov, V.	2, 11		Han, S.Y.	42
	Angulo, J.	59		Herrera, J.M.	245
				Heynen, M.	110
В				Hofmann, P.	110
	Balley, S.	160			
	Benz, U.	110	${f L}$		
	Bodansky, E.	67		Lee, D.	82
				Lee, J.Y.	140
\mathbf{C}				Levachkine, S.	2, 11, 22, 34, 42, 102, 169
	Chai, W.Y.	205		Li, Z.	92
	Cheng, T.	92		Lingenfelder, I.	110
	Chung, Y.K.	140			
			\mathbf{M}		
\mathbf{F}			114	Manrique, P.	127
	Felicísimo, A.	238		Moreno, M.A.	102
	Felipe, E.	52		Mukai, N.	197
	Feng, J.	197		Muñoz, J.	238
		•		Mustière, S.	152
\mathbf{G}					
	Gelbukh, A.	42	O		
	Gesbert, N.	152	U	Ortiz, E.	245
	González, J.	215		,	
	Gould, M.	179, 188	D		•
	Granell, C.	179	P	Pech, M.	215
	Guevara, A.	229		Pogrebnyak, O.	127
				Poveda, J.	188
				i o rodu, J.	100

		Page			Page
Q			T		
	Quintero, R.	145		Theng, L.B.	205
				Torres, M.J.	169
R					
	Ramírez, E.	245	${f V}$		
	Ramos, M.E.	52		Velázquez, A.	34
	Razo, A.F.	226		Villaseñor, E.	245
	Ren, J.	133			
	Rodríguez, A.	245	\mathbf{W}		
	Ruas, A.	76		Watanabe, T.	197
				Willhauck, G.	110
S					
	Serra, J.	59	${f Z}$		
	Sheeren, D.	152	2	Zepeda, C.	221
	Shen, Y.	133			
	Sol, D.	215, 221, 226	1		

Haramini)

- Amedinas

1.4vi 20m, 2.

A. compression

G 143

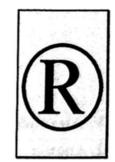
discount collections.

11.

0.00

Surfereity for

ALA MAS



REVIEWERS

V. Alexandrov

SPIIRAS, Russia

J. Angulo

Ecole des Mines de Paris-Centre de Morphologie Mathématique, France

U. Benz

Definiens Imaging/eCognition, Germany

D. Blostein

University of Kingston, Canada

E. Bodansky

Environmental System Research Institute, USA

M. Bye

GEO-IMAGE, Australia

J. Chevallier

Laval Université, Canada

J. L. Díaz de León Centre for Computing Research, IPN, Mexico

M. Egenhofer

National Center for Geographic Information and Analysis at the University of Maine, USA

A. Felicísimo

Extremadura University, Spain

A. Frolov

University of Maryland at College Park, USA

M. Gould

Universitat Jaume I, Spain

A. Guevara

GTT NetCorp., Inc., USA

A. Guzmán

IEEE/ACM, Mexico

V.-B. Kwon

Chung-Ang University, Korea Environmental System Research Institute, USA

D. Lee

Centre for Computing Research, IPN, Mexico

S. Levachkine

Centre for Computing Research, IPN, Mexico

M. Moreno

Real Jardín Botánico de Madrid, CSIC, Spain

J. Muñoz T. Poiker

Simon Fraser University, Canada

A. Ruas

IGN, France

J. Serra

Ecole des Mines de Paris-Centre de Morphologie Mathématique, France

R. Suárez

Mexican Petroleum Institute/PIMAYC, Mexico

K. Tombre

LORIA, France

M. Torres

Centre for Computing Research, IPN, Mexico Centre for Computing Research, IPN, Mexico

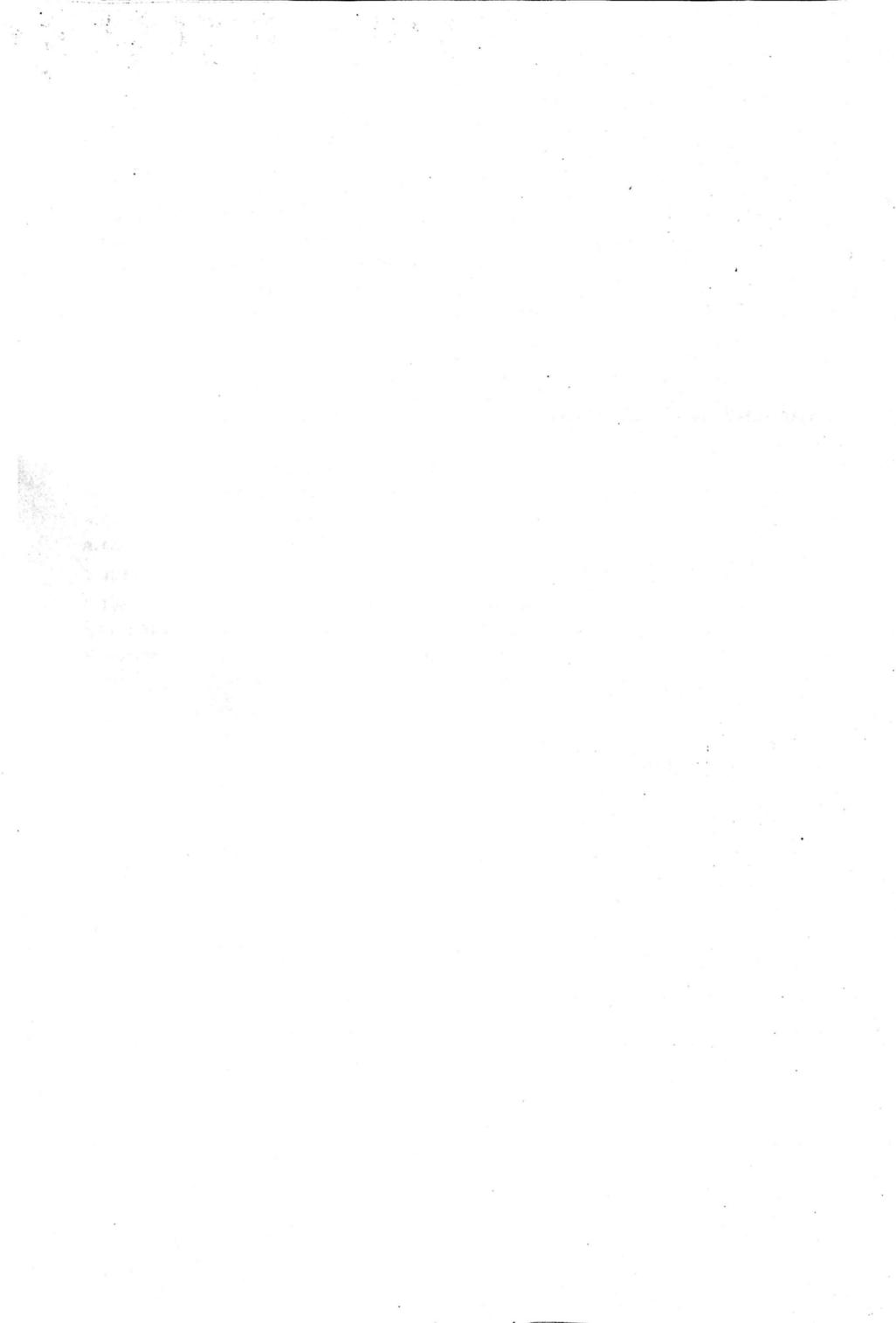
R. Quintero A. Velázquez

Mexican Petroleum Institute/PIMAYC, Mexico

. 4

*

×



. .

SEMANTIC PROCESSING OF SPATIAL DATA

Serguei Levachkine, Jean Serra and Max Egenhofer (Eds.)

This book contains selected papers submitted to the International Workshops on Semantic Processing of Spatial Data (GEOPRO 2002 & 2003), which were held at the Centre for Computing Research (CIC) of the National Polytechnic Institute (IPN), Mexico City