MULTISCALE SPATIAL DATABASES CREATION AND CARTOGRAPHIC GENERALIZATION USING HIGH RESOLUTION SATELLITE IMAGES: CONCEPTS, PROCEDURES AND IDEAS IN THE FRAME OF TWO BILATERAL RESEARCH PROJECTS

Dimos Pantazis 1, Vassiliki Kyriakopoulou 1, Konstantina Rizopoulou 1, Panagiotis Stratakis 1, Haralambos Karathanassis 1,3, Li Yingcheng 2, Tongying Guo 2

1 SOCRATES Research Group - Surveying Department, School of Technological Applications, TEI (Technological Education Institution of Athens), Samarinas 3, Sepolia, Athens 104 43, Greece, phone: +30 210 512 95 49, fax: +30 210 512 95 49, e-mail address: pantdonz@otenet.gr, website: http://zeus.teiath.gr/

2 Department of Information Service, National Remote Sensing Center (NRSC) of China, 16 Beitaiping Road, 100039 Beijing, P.R. China, phone: +86 10 68289695, fax: +86 10 68160588, e-mail address: guotongying@vip.sina.com, website: http://www.casm.ac.cn

3 Department of Environmental Planning, Ministry for the Environment, Land Planning & Public Works, Athens, Greece

Abstract

In 2001, a bilateral project started among the Surveying Department of the TEI of Athens - Greece, the NRSC of China and “Space Imaging Europe” company with the support of the HGM Service. The objective of this collaboration was the study of the possibility of updating middle scale topographic maps, using VHR satellite images and the investigation of its practical applicability. A new two years research project was launched in June 2003 concerning the study of the improvement in the accuracy of the topographic maps using VHR satellite images. Both research groups in the frame of this second on going collaboration, agreed to conduct at the same time a common research concerning: a) the investigation of spatial databases generalization techniques and b) new procedures for multi-scale spatial databases. The objective of this article is to present all the preliminary work done in the field of the generalization and creation of spatial multi-scale databases.

1. INTRODUCTION – THE OBJECTIVES OF THE ARTICLE

The updating of topographic maps of middle scale has been realized up to now mostly with the use of aerial photos and classic photogrammetric methods. Nowadays, the launch of new and specially equipped satellites has allowed us to obtain high resolution satellite images (Zanoni and Goward, 2003) which after geometric and other corrections could achieve accuracies of 2m or more (accuracy that depends on the terrain and on the Digital Elevation Model (DEM), ground control points, the correction methods, etc) (Ganas et al., 2002). Theoretically this means, that such satellite images can be used for updating medium scale topographic maps and also generating them from scratch. In June 2001, a bilateral project started among the Surveying Department (Research group SOCRATES) of the Technological Education Institution of Athens - Greece, the Information Department of the National Remote Sensing Center of China and “Space Imaging Europe” company with the support of the Greek Geographic Military Service. The project was co-financed by the European Union (E.U.) and the Ministries of Development, Research and Technology of Greece and China, and “Space Imaging Europe” company. The objective of this collaboration was the analytical study of the above theoretical possibility and the investigation of its practical applicability, identifying at the same time its advantages and disadvantages. In the frame of this scientific collaboration both research teams (Chinese and Greek) confirmed the necessity for further investigation not only of the possibility to update the topographic maps but also to improve their accuracy without reconstructing them from scratch utilizing high resolution / accuracy satellite images (Pantazis et al., 2004). A new two years research project with the same partners and the same financial resources was launched in June 2003 concerning the analytical study of the improvement in the accuracy of the topographic maps using high accuracy satellite images, without being reconstructed from scratch.

XXII International Cartographic Conference (ICC2005) A Coruña, Spain, 11-16 July 2005
Hosted by: The International Cartographic Association (ICA-ACI)
ISBN: 0-958-46093-0
Produced by Global Congresos
In parallel both research groups (Chinese and Greek) in the frame of this second on going collaboration (which ends in November 2005), agreed to conduct a common research concerning: a) the investigation of new spatial databases generalization techniques and b) new procedures for multi-scale spatial databases creation, using the data of high resolution satellite image.

Nowadays vector cartographic databases as well analogue maps exist for different scales. Given the fact that for the first time the alternation of the view process (to zoom in and zoom out) in a satellite image is related with the correspondent pixel resolution from scales 1:5000 to 1:500.000 with the same dataset (raster), a set of new concepts, ideas and procedures for spatial multi-scale databases creation has started to be developed. Generally speaking their creation needs aerial photographs of different scales based on different aerotriangulations and overall photogrammetric techniques. In addition different updating process must be realised for different scales cartographic maps. Our objective is the creation of multiscale cartographic databases with interrelated objects permitted the continuous and simultaneous interrelated updated at all scales of all cartographic objects. A multi-resolution, multi-scale/representation database can be described as a spatial/cartographic database which can be used to store the same real-world-phenomena at different levels of precision, accuracy and resolution. The objective of this article is to present all the preliminary work done in the field of the generalization and creation of spatial multi-scale databases and the innovative concepts and ideas investigated in the frame of both bilateral research projects.

2. STRUCTURE OF THE ARTICLE

The article is composed of six paragraphs. After the short introduction that presented the framework, the problem and the article’s objective; a set of definitions, basic concepts and a literature review are given in paragraph 3. Innovative concepts, ideas and processes developed or to be developed and implemented in the frame of our bilateral research are discussed in paragraph 4. In paragraph 5 the basic points of the framework for multi-scale cartographic database creation are presented. The last paragraph addresses the conclusions and perspectives based on the results achieved and the experience gained up today.

3. DEFINITIONS, BASIC CONCEPTS AND LITERATURE REVIEW

There are a lot of applications that require the representation of geographic information in different levels of detail and generalization. A multi-scale database aims at the representation of the same geographic area at different levels of detail. There are several reasons to create multi-scale cartographic database (MCDB) : a MCDB allows a multi-scale analysis of the cartographic/spatial data. This means that information in one resolution/scale can be analysed with respect to information given in another resolution. In addition a MCDB gives the possibility of propagating updates between scales, which called “incremental generalisation”. The automatic reliable updating of the multi-scale cartographic database creates the problem of consistency. There are also different conceptualizations of geographic entities that need to taken into account (Kavouras and Kokla, 2002). In our case we try to elaborate solutions based on use of constraints describing the allowable and prohibited cartographic data evolutions from one scale to another.

The multi-scale cartographic database creation is a very complex process integrating many parameters. Generalization is a key mechanism in this process, filtering out information that is needed for particular scales or layers. The objective of generalization is to remove detail while retaining important information content and recognisable characteristics of the reality being represented. A related objective, particularly meaningful to cartographers (Robinson et al, 1984; Keates 1989), is to assure a legible representation. It is thus essential to define importance in order to standardise before being able to automate the generalization process. Defining importance in an objective manner (does exist one ?) is not an easy task; therefore the generalization process tends to be a subjective operation (Kavouras, 1999).

Rendering map data at scales smaller than their source leads to competition for space amongst map objects exhibiting graphic conflict. At a reduced scale objects are either too small to be seen or too close to each other to be distinguishable, while at the same time, symbol size is exaggerated, in order to maintain size relations and legibility. Cartographic map generalization is the process by which complexity of the map, due to scale reduction, is reduced and any graphic conflict is resolved, emphasizing the essential while suppressing the unimportant, maintaining logical relations between map objects, and preserving aesthetic quality. The main objective is to create maps of high graphical clarity, so that the map image can be easily perceived and the message the map intends to deliver can be readily understood (Weibell and Dutton, 1999). The goal of digital cartographic generalization of a multiscale spatial database is to maintain clarity of presentation at the target scale.

Research in cartographic generalization has taken several directions including algorithmic design and testing, the design of models and conceptual frameworks, the application of expert systems, and the modelling of cartographic features.
In an attempt to formalise the generalization process, many such as Muller et al. (1995) have distinguished the following views: object generalization (from reality to representation), model or database generalization (controlled reduction of data for storage saving and computational efficiency) and cartographic generalization. The first two views are often faced as one by the name “conceptual generalization”. Thus far, most of the work in generalization has focused on what is termed “cartographic” generalization, which involves the graphical considerations associated with scale change.

A second, less researched, area is in “model” generalization, by which generalization operators (simplification, smoothing, aggregation, agglomeration, and others) are applied to an original digital landscape model (DLM) in order to create secondary representations of the database, called digital cartographic models (DCM)* (Lam et al, 2004). These terms, DLM and DCM, are taken from the European cartographic literature, where a significant amount of work has been completed in this area.

Intelligent automated generalization will permit the encoding of raw observations in digital form to derive more responsive, application-specific representations. The VHR satellite images could be a useful framework to this. Both semantic and geometric generalisation may be related with spatial resolution changes in a automatic/computerised way in most cases. Jones (1997) argued that successful automation “[...] of generalization concerned with caricature requires a computer program to simulate knowledge of geographical and cartographic structure in order to recognise the cultural of geomorphological nature of the map feature”. VHR satellite images provide in a certain degree the basis for this knowledge.

Generalization, the range of procedures entailed in spatial database handling and the cartographic presentation of spatial data, is one of the more difficult cartographic processes to automate but at the same time a key factor / process in the creation of multi-scale cartographic database. In developing automated techniques, the challenge is to embed the contextual and semantic knowledge into the software.

A lot of work has been done in the field of automatic generalization and multiple representation and more generally in multiscale cartographic databases (see http://ica.ign.fr/ for more than 100 references in all relative field and subfields as also Anne Ruas, 2004). One of most important achievement in the field is the German project ATKINS (Authoritative Topographic-Cartographic Information System) which is a multi-resolution representation database (MRDB). The target of this system is quite similar with our target, nevertheless the process and the concepts, and theoretical background of our work is quite different. The main literature that constitutes the base for further research is mentioned below.

Kilpeläinen (1997) focuses on the maintenance and generalization of geo-databases, with particular emphasis on topographic data. His work is based on a two-fold conceptual model for generalization in databases, in which the first part consists of model generalization and the second of cartographic generalization. The current work also supports the utilization of the model within the concept of multiple representation. A model for multiple representation database (MRDB) system is presented, discussed and applied to topographic data. The concept has been developed of incremental generalization for automatic updating of generalized versions in an MRDB environment. Three different generalization tests were performed, which resulted in the discovering of four different categories of rule type: geometric, topological, context-related, and culture-related. The central outcome of this research has been combining these rules, incrementality, and modularization into the generalization processes for utilization and maintenance of an MRDB system. The central requirement for an MRDB, bidirectional connectivities between the various object representations for geographic data, are modeled and implemented in an object-oriented environment.

Sester (1999) has given an overview of the generalization problem in spatial databases. As a major research topic, the establishing of links between different levels of detail has been identified. In order to provide these links, an object-oriented approach in conjunction with a machine learning technique has been presented. A prototypic implementation showed the potential of this approach, namely the flexibility and the creation of dynamic knowledge bases. Machine learning techniques are very useful in situations where no explicit knowledge is available. Their main feature is to reveal explicit knowledge from implicit data. In this case, existing maps serve as a knowledge store, which capture a lot of implicit information. A further advantage is the fact that the result can be evaluated and validated, since it is given in a form which is understandable, e.g. in terms of rules or decision trees.

Another approach to tackle problems involved in automated map generalization (Yang and Gold, 1999) combines database generalization and dynamic object generalization capabilities and is coupled with a map agent on top of a map object. The map agent is for the construction of navigating maps, performing tasks on behalf of, and communicating with, users. The object classes are topologically and geometrically structured with the dynamic VMO-tree (Voronoi Map Object Tree). In the context of map generalization, the primary role of the map agent is to control, schedule, and validate dynamic generalization operations enacted by objects. Another role of the map agent is to aid map uses. Experiences show that such a system looks promising for providing desirable solutions.
Li et al (2001) apply a **cell-based model** to GIS generalization, which represents a new methodology within the GIS context. Cellular Automata (CA) has found a place in many interesting real world applications, including the modeling and simulation of numerous systems, across many disciplines. In this study, which is based on CA techniques, an extended neighbourhood algorithm is used in the cell-based model to automatically generalize thematic maps derived from classified satellite images. An example of generalizing a land use map of Lisbon Bay in Portugal is given, which gives satisfactory results.

Database generalization is a process which aims at producing a geographic database from an initial geographic database in order to satisfy new needs. Han-Sze-Chuen et al. (no dated) from the study of an actual database generalization process developed at the Minister de Ressources Naturelles of Quebec (MRNQ), tried to formalise the cartographic knowledge contained in the process to make it reusable. MRNO’s process has the form of a workflow which contains a list of actions applied to object classes. They introduced the notions of problem and operation; then they proposed a **causal graph** which associates a set of problems to a set of operations. The utility of such a model is that it offers an explanation for each action contained in the workflow and allows an anticipation of future problems when creating a new workflow.

**Data mining methods** are used to extract such knowledge from a multiple representation database. An unsupervised classification of the objects in the large scale data set is performed using the ISODATA algorithm. For each class the probability that members of the class are represented in a certain way in the small-scale data set is calculated. This knowledge is used for the automated propagation of updates. The method is currently being tested on topographical data sets provided by the Swedish National Land Survey on a scale of 1:10 000, 1:50 000 and 1:100 000 (Duncars, 2004).

The MurMur (Multiple Representations, Multiple Resolutions) project aims at establishing and implementing a representation paradigm allowing each information fact to be recorded through multiple, consistent, and possibly irreducible representations (i.e., not derivable from one another). Such a capability gives full freedom to designers to choose the representations they want to implement, and to users to define which representations they want to query, as well as a personalized outcome for their queries (Spaccapietra et al., 1999).

Some of the previous concepts are partially implemented in different software e.g. **Clarity** from Laser-Scan. Laser-Scan and Intergraph are working to bring together the best generalization technology and best geospatial technology available today in the marketplace. With a collaborative approach and using interoperability principles of open standards-based data exchange, significant advances can be made in the field of generalization. eCognition is the first commercially available product for object based and multiscale image analysis. Going far beyond the methodical limits of pixel-based approaches, eCognition leads to higher classification accuracy and to better semantic differentiation. DynaGen automates the generation of map products at varying scales from a single high resolution database, providing cost savings in long-term data maintenance, providing large number of operators and algorithms for generalization of feature geometry for the purpose of either the reduction of data or cartographic presentation, dynamic displays enable the user to see the results of the generalization before committing the changes, possibility for saving generalization parameters for reuse later and organized for batch processing, validation of all results to ensure that changes do not violate topological and real world feature-to-feature relationships, calculation and updating attribute information for modified or newly created generalized features.

### 4. MULTIPLE CARTOGRAPHIC DATABASES BASED ON VHR – CONCEPTS: IDEAS, PROPOSED PROCESSES

One of the most evident examples where multiple representations are needed is in cartographic applications. Map producers need to build maps of the same geographic area at different levels of abstraction (different scales/resolutions). Typically, they maintain one database per scale, with no interrelationship.

The VHR satellite images provide in practice many conceptual arguments for modeling and inter-related cartographic data of all scales. The basic idea of our efforts is based on the simple fact/observation that the VHR satellite image contains or are itself the geographic / cartographic data of a wide range of scales: from 1:5000 (and even bigger) to 1:500.000 (and even smaller). Consequently the preliminary / basic form of the cartographic data we need for the cartographic production in all those scales are there (see also Bianchin and Bravin, 2004)). It is obvious that the overall geographic / cartographic base for any scale cartographic data / map is contained in VHR satellite images. This “way to conceptualize things” in the field of cartography naturally leads somehow to a “Russian dolls” system for cartographic data of different scales.

A multi-scale cartographic database must be based in a geodetic network which will guarantee the geographic reference and positional accuracy in each scale and at the same time will relate the coordinates, the accuracy and resolution of the raster data. This means that in each scale the accuracy of the control geodetic points will refer in specific surface which
could be represented in a given scale. From this point of view the importance of a geodetic network interrelated with the concepts of resolution, scale and accuracy is high (Troispoux, 2005).

Until now the problem of multi-scale cartographic / geographic database has been analyzed only from a static point of view. This means static representations of the “reality” in different scales somehow interrelated. The concept / technique we try to develop is rather a fluid representation of the reality changing from a starting point (scale) to another (scale) continuously.

The scales covered in the frame of our project in a first phase are in the range of 1:5,000 to 1:500,000. It is clear that the nature of the data (raster data) does not permit for to talk about cartographic objects to being identified since the data are registered in pixel values. So the first problem to resolve is the transformation of the raster data to “objects”. A second problem is to define all the relationships between all objects in all scales. Finally, a third problem is to define the “symbols” which will replace “objects” when necessary (during the scale change) defining at the same time all their relations (topological and others) with all others objects in different scales.

Definiens Imaging’s eCognition (Baatz & al, 2003), a multi-resolution object-oriented image analysis software, could meet in a certain degree the above two first problems. The concept behind eCognition is that important semantic information necessary to interpret an image is not represented in single pixels, but in meaningful image objects and their common relationships. It supports different supervised classification techniques and different methods to train and build up a knowledge base for the classification of image objects which are extracted in a previous image segmentation step. Using the possibility to produce image objects in different resolutions, a project can contain an hierarchical network with different object levels of different resolutions. This structure represents image information on different scales simultaneously. Each object “knows” its neighbors and its sub-objects. Thus, different object levels can be analyzed in relation to each other. For instance, image objects can be classified as to the detailed composition of sub-objects. Class hierarchy contains all classes of a classification scheme, allowing the passing down of class descriptions to child classes on the one hand, and meaningful semantic grouping of classes on the other. Moreover, eCognition allows for the automated extraction of polygons (vectors), with an attached attribute list, based on image objects, while after vectorization, it simultaneously holds image objects in raster and vector representations.

The basic idea of our efforts based in the creation of buffers of specific wide for the points, lines and polygons representing objects in different scales. Each object could “move” during its different representations only inside of the specific limits of the buffer. All relations of all buffers of all scales must be analytically described. The different cartographic symbols, in addition of the relation with the objects which represent have also a topological relation with the specific buffer of each cartographic object. “Basic” points, lines and polygons (of different types: Hydrographic network, road network, trigonometric points, etc) will provide the “skeleton” which will be presented with different spatial resolution in many scales.

Another important step in developing raster cartographic databases is the establishing of the relation between scale and spatial resolution. This relation most of time is based on empirical rules. As a starting point in our case we accepted the graphic minimal error which is 0.2 mm at the map and this will be the base of any further calculation.

Conceptualize space by cartographic objects needs the “conceptualization» of their relationships in different scales, and this observation underlying the need for a different approach in the way that maps, cartographic objects and scales are inter-related in different scales. The conceptual modeling of multi-scales cartographic databases will include not only the objects of specific scales and their relations – in the same scale but also in different ones- (in our case 1: 5000, 1 : 10,000, 1 : 25.000, 1; 50.000, 1: 100.000, 1: 500.000) but also the mathematical equations providing the continuous transformations from one scale to another. The life of an “object” could be present in a continuous frame of scales with the same of different representations and with or without its participation in the creation of other objects (objects which exist in smaller scales). Consequently the identification of rules concerning cartographic objects and their relations and transformations (appearance, disappearance, changes,…) is necessary in order to assure a natural inheritance on each change from big scales to smaller and vice-versa. In addition changes from middle scale towards smaller or bigger scales must be possible to be realized.

5. BASIC POINTS OF A MULTISCALE CARTOGRAPHIC DATABASES CREATION THEORETICAL FRAMEWORK

The development of a new framework of theoretical concepts concerning the continuous multi-scale representation is developing in parallel with the practical efforts for their implementation. The basic ideas of the start needed to be followed by a stronger theoretical background and to be implemented by the actual or future technology. At the same time practical efforts and experiments give the area to continue and / or change direction in our theoretical investigations. The basic points of this framework are summarized below:
1. Independently of the different representations of the reality in different scales (in cartographic maps / databases) and their fundamental differences all representations will have the same basic structure.

2. The VHR satellite images provide finite means for infinite number of representations of the reality.

3. The overall concept developed is based on a relatively simple rule and simple structural relation between all cartographic objects of the same nature through all the scales but also between different objects in different scales.

4. Representations changes of the reality doesn’t necessarily means changes of the cartographic objects, e.g. the airplane classic example: more the airplane gains altitude more the object we see from the window becomes less clear taking another form, the objects does not nor really change, change only that we can see (the objects representations). This practically means that an object of a smaller scale could contain objects of bigger scales or even could be created by some parts of those objects.

5. The continuous representation’s change of the reality (e.g. when the airplane takes off) and the same cartographic objects must be taking into account in the frame of a multi-scale cartographic database. This practically means that an infinite number of scales must be represented by a finite number of processes (mathematical equations/algorithms which will describe the representation change of each object and their inter-relations).

6. Points, lines and polygons which will represent basic features will be presented at different spatial resolution at all scales of the multi-scale cartographic database.

7. All objects will be inter-related in all scales.

8. The accuracy of representation of each object will follow the rules of the graphic representation’s cartographic error and also the visual capabilities of each scale representation.

9. All the representations of an object will be encapsulated in this object.

10. All the relations (topological and others) between all objects in all scales will be defined.

6. CONCLUSIONS AND PERSPECTIVES

A new perspective based on a preliminary research work concerning the multi-scale cartographic databases developing in the frame of a bilateral research project between China and Greece was summarily presented. Specific solutions related to the problems such as the metadata development (of multi-scale cartographic databases), the conceptual modeling, the data dictionary etc are under investigation. Our first conclusions underline the need for integrated conceptual engagement on multi-scale database design.

From technological point of view it is of great importance the editions of the GDI (Graphical Device Interface), the possibility of simultaneous use of all kind of libraries (16 bits, 32, etc), the aero / Longhorn / Indigo and others similar technologies and also the new graphic cards with extended capabilities in graphic displays.

In the field “reality” the establishment of GPS receivers of high accuracy (e.g. France, network Teria which include GPS, Glonass and in the future Galileos) can procure the basic geodetic network related with spatial resolution, scale and coordinates accuracy for interrelated cartographic data of all scales.

REFERENCES


Han-Sze-Chuen, D., Mustiere, S. and Moulin, B. (no dated), Formalising the Geographic Database Generalization Process by means of a conflicts/operations graph. Computer Science Department and Geomatics Research Centre, Laval University, GIK 7P4, Ste-Foy, Quebec, Canada.


Ruas Anne, 2004. Le changement de niveau de detail dans la representation de l’information geographique, Memoire d’habilitation a diriger des recherches, Universite de Marne - la - Vallee


Yang, W., Gold, C., 1999. A System Approach to Automated Map Generalization. Centre de recherche en géomatique, Université Laval, Québec, Canada G1V 7P4.