GIS REVISION BASED ON AERIAL PHOTOS : A RASTER TO VECTOR METHODOLOGY

Bruno Tellez and Sylvie Servigne
Laboratory for Information Systems Engineering (LISI)
Institut National des Sciences Appliquées de Lyon (INSA)
Bât 404, 20 avenue Albert Einstein
69621 Villeurbanne Cedex - FRANCE
E-mail : btellez,servigne@if.insa-lyon.fr

ISPRS Commission VI, Working Group 3

KEY WORDS: Cadastre, Aerial Photos, Matching , Structural Map

ABSTRACT:

This paper presents how aerial photos can be used to update automatically cadastral data. The methodology proposed allows to manipulate with a same model, vector- and raster-issued data. The main objective is to structure topologically data and objects. For cadastre, the methodology consists in introducing or recovering topology. Moreover, on photo, our methodology is used for objects extraction and identification. Based on image analysis, features are extracted to compute plane subdivision (by Delaunay triangulation). Pictorial objects so defined, integrate both geometry and topology. Objects and relations between them are used to develop the concept of *Structural Map*. For us, structural map will be considered as a common representation for cadastre and photo and it becomes the reference to process comparison and to initiate updates.

RÉSUMÉ:

Cet article présente comment les photos aériennes peuvent être utilisées dans la mise à jour de données cadastrales. La méthodologie développée permet de manipuler sous un même modèle des données raster et des données vecteur. Son objectif principal est de structurer topologiquement les données et les objets. Appliquée sur le cadastre pour introduire ou retrouver la topologie, elle permet, dans les photos, d'identifier et d'extraire les objets. Basée sur l'analyse d'images, l'extraction des caractéristiques permet de calculer une subdivision du plan (grâce à la triangulation de Delaunay). Les objets pictoriels ainsi définis intègrent géométrie et topologie. Les objets ainsi que les relations qui les lient servent à définir le concept de *Carte Structurelle (Structural Map)*. Pour nous, la carte structurelle sera considérée comme une représentation commune pour le cadastre et la photo. Elle devient la référence pour comparer et mettre à jour la base de données numériques.

INTRODUCTION

Presently, Geographical Information Systems are generally used for spatial information analysis. It is now necessary to ensure temporal validity. Evolution of the land must be considered and updating is really possible after interpretation and integration of new data.

The PHOTOPOLIS project that initiates these works, dealt with the cadastral database of Padua (Italy) in order to automatically update it. These corrections concerned legal modifications following planning permission but also detection of illegal constructions. The delay issued to capture and integrate new data is time-consuming and aerial photos are an interesting alternative. First of all, they provide some snapshots of the real world. Thus, they are well suited for treatment of urban and sub-urban zones. If aerial photo is an interesting data source, the major problem is to automatically extract information from it. Many solutions have been proposed to tackle raster data: stereomatching [4] compares two images to determine similarity (or dissimilarity), snakes deals with raster and vector [3]. Our approach chooses to transform raster to vector data using image analysis based on topological considerations. The main advantages are to consider objects rather than image features because they are more meaningful. Moreover, data so defined, can be compared at a same level.

We first present a methodology whose objective is to structure input data into a unique data model including topology. We then present how methodology based on Delaunay triangulation can be applied on cadastre but overall on aerial photo. A tessellation is managed to process high level objects extraction. Cadastral objects and pictorial objects (issued from photos) are then included into a map representation (*Structural Map*), and we detail what concepts should be integrated to proceed the comparison.

COMMON METHODOLOGY FOR COMMON MODEL

Generally, difficulties to manage both raster- and vector-based data are so important that comparison is not possible without structuring them into a common model. More, data could be incomplete or erroneous as in cadastral base or unusable as they stand in aerial photos. Our methodology allows to structure data in order to prepare them for comparison. The first step is to define what concepts can be found in our model. If it focuses first on geometry description, it could be able to describe objects.

Nevertheless describing objects geometrically is not sufficient. The model has to also include relations existing between objects or entities. These relations, called topology, prevent structural errors (overlaps that would not have to exist, polygon not closed...) and increase semantics and spatial queries capabilities. Data are generally defined according to their dimensions (punctual as Vertices, linear as Edges or areal objects as Faces) also by topological relations, which relate every entity with each of the others (see a possible representation in fig. 1 [5]).

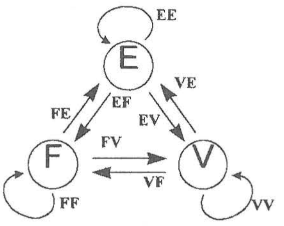

Figure 1. WINGED Representation

Topological relations provide important information for each entity (e.g Faces located at the left and right of an edge). As in planar graph, all parts of the space could be considered and so this organisation is fully adapted for plane subdivision. Due to lack of data quality, structuring could be erroneous or objects dispersed. No information could be obtained to deal with empty space between objects. To produce plane subdivision with cadastre or photo, a tessellation is designed. Our solution is based on Delaunay triangulation that decomposes space into triangles. It ensures a construction based on vertices proximity. This will be used to detect topological situations in cadastre and to process object extraction in aerial photos. Delaunay construction is used in this classic way to proceed aerial photo (raster) and manage edges (vector) in the cadastre constrained triangulation. However, it is fully adapted to the model described before, even data are incomplete. They can be recovered incrementally to extract a topological and complete structure - triangulation processing defines only FV and FF relations - as defined in figure 1.

This methodology allows to define the entire space with entities geometry and topological relations existing between them. The tessellation obtained by Delaunay triangulation will be used at a second step to reconstruct objects with high level of abstraction, with high level of semantics. They represent a view of real world objects features: parcel geometry in cadastre, building detected in aerial photo for example...

CADASTRE RECONSTRUCTION

Difficulties to deal with cadastre data is that generally topology was not taken into account. So if object description is complete from a geometric point of view, nothing certifies validity structure. Most often, data are issued from manual digitising and errors in location as small as they are, could lead to inconsistencies in topology.

Two kinds of errors can be found in cadastre data. If cadastre is considered as a tessellation, errors modify this ideal state. Due to bad vertex location, objects overlaps or lost of adjacency represent situations which do not affect objects definition but prohibit spatial queries about relative dispositions between objects. If the objective for cadastre is to deal with topology, the model previously defined is fully adapted. Data are introduced in our model in order to define topological relations existing between objects. An adapted process, driving by model methodology, is applied to obtain data that detect constraints linked to cadastre specifications.

The application of methodology on cadastre consists in an objects triangulation. To simplify approach, we only preserve few entities: parcels, buildings without distinctions, roads and city blocks. They are described by polygons (faces) but only edges and vertices are inserted into triangulation (Constrained Delaunay Triangulation takes edges into accounts). So we can find initial objects from triangles, with additional information about topology. The result is a set of triangles where topology is well defined. The goal is now to analyse edges and triangles because they bear errors cases (intersection on edges or bad adjacency suggested by thin triangles). Finally, topological relations are integrated to be linked with cadastre semantics to define relations between geographical objects (parcels adjacent with parcels, city-blocks surrounded by roads,...).

Even the model can help us to tessellation reconstruction and topology maintaining, it is not obvious to correct by merging or moving some points. The important fact at this point is to suggest objects disposition.

OBJECTS EXTRACTION IN AERIAL PHOTOS

Finding objects in aerial photos is a central problem in our confrontation. Indeed, it represents the source that gives new data for comparing and maintaining quality in cadastral base. But extracting information as vector objects is not trivial. Generally, these methods are based on a same principle: detection of features and reconstruction of high levels objects. Examples are numerous: stereomatching [4] where images are compared features by features, snakes [3] use initial boundaries that are modified to cling at best to real boundaries of objects. The methodology was designed with three important objectives: preserve geometry, introduce "vector topology" and if possible, interpret objects semantics. Therefore, it can deal with these concepts: geometry comes from features extracted from photo, topology by triangulation of these features and semantics is interpreted before reconstruction. The objective for features extraction -geometry- is to propose set of points that will be used to compute Delaunay triangulation and so, to generate plane tessellation including strong topology.

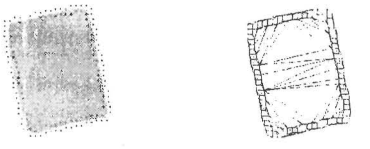

Figure 2. Object and its double definition. Triangulated object

The approach to extract points relies on classical image processing techniques. The main difficulty with objects extraction is its complexity due to texture. Our approach does not study texture at this point and we suppose that studied objects are relatively homogeneous in colour. So the first step defines homogeneous zones of colour (using tresholding and non-linear filtering).

The interesting way is to detect regions not by a single boundary that separates two homogeneous zones but rather by cutting off each colour zone as shown in fig. 2. Finally, regions obtained are now described by two boundaries: first one comes from its own detection and the second from boundaries of objects that surround it. More conceptually, resulting boundary is defined by a buffer zone that includes the theoretical boundary.

Only few points are registered to describe boundaries. These called *contour points* are extracted following two considerations [6]: they must belong to interior zone and be near another homogeneous zone (at a distance *d*). The advantage is that topology construction (Delaunay triangulation) relies on a principle of proximity. So we can rightfully imagine that points belonging on these two boundaries will be connected. The set of triangles thus defined constitutes a *boundary zone* (see fig. 2).

Next step concerns triangles. They are labelled according to colour vertices: if all are similar, the triangle is regarded as an homogeneous zone (we call it *zone triangle*), if only two have the same colour then the triangle lies on boundary (*boundary triangle*) or the three are different and a triangle connects three different objects.

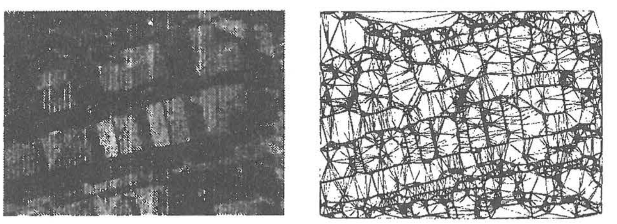

Figure 3. Image and its tessellation of triangles

Aerial photo is now considered as a tessellation of triangles (fig 3.) respecting topology and where all entities are classified according to their membership of a boundary or a zone. The final step in image processing is to obtain high level objects resulting on triangles combination. Using data model and its topology, sides (edges) are connected thanks to left/right triangle labels (essentially boundary triangles). By designing several closed polygon, the level of abstraction was increased and the last step is to evaluate what polygons have to be combined, to obtain objects (see fig. 4). Objects are defined by a boundary zone, and boundary zones are defined by two polygons. Topology helps us too to determine precisely internal and external contours. Combining with Minimum Bounding Rectangle (*MBR*), these concepts are derived to obtain a hierarchy, formed between objects [7]. The *double contour* (internal and external) is used to discover the three essential topological relations for an image: inclusion, adjacency and disjunction. Each situation results from a contour combining.

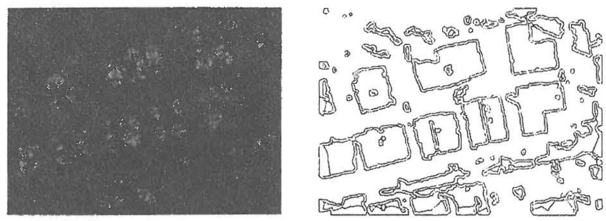

Figure 4. Labelling and contour extraction

The hierarchy represents a good approach to structure aerial photos but it has to be semantically enriched in relations, in definitions of objects and so, it will be extended to be more adapted for comparison processes.

DEFINING MAP REPRESENTATION FOR COMPARISON

Our methodology allows to define and integrate data in a common structure. It combines both objects and topology but the model is not sufficient in its conceptual representation. Indeed, until now, topology was considered from a geometrical point of view. In the new representation, we want to integrate the semantics of these relations. The *Structural Map* (fig. 5) is characterised by an abstraction of vector- and raster-based data. The model defined previously becomes the layer more dedicated to geometrical topology. Objects and topological relations are more developed in semantics in order to facilitate the two essential functions of map: **Represent** and **Compare**, the comparison lying on representation.

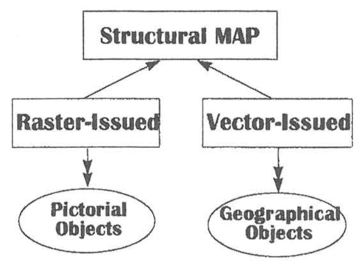

Figure 5. Structural Map Representation

Object representation does not integrate only one geometrical representation but many (for example, pictorial could be represented by two polygons as boundaries or by a set of triangles). Each representation has its own utility in map management. Topologic dimension can also be extended to integrate specificity of each domain. For example, semantics of cadastre lead some rules about disposition of objects that simplifies representation. Indeed, if we consider entities: {*Building, Road, Parcel, City Block*}, only relations {*Adjacent, Over*} can be used and rules are easily definable {e.g. *Parcel Adjacent Parcel, Building Over Parcel* ...}. But in a cadastre/photo comparison, it is not obvious to verify these relations. As example, cadastre parcels can not be detected in aerial photo because parcel limits are not visible in photo. To obtain more information about relations and objects, we must consider more than geometrical aspects. Many ways exist to represent all situations between objects. Egenhofer model [2] studies relations as combination of entities (point, line and polygon) according to three object characteristics : interior, boundary and exterior. It can express all relations that can occur between them. As far as relations between region concern us, eight cases are distinguished as *overlap, meet, disjoint*... Allen [1] relations are another example of a mean to represent objects disposition. It is based on temporal intervals notion. The objective with these new information is to precise relative position of objects by direction (as Frank's sense), proximity, and at last, obtain a global structure. Representing semantics is a larger task. It implies to be able to describe real world what seems impracticable. It is certainly more interesting to develop what objects (or parts of objects) are encountered in aerial photo and how to describe their structure (house = roof + chimney) to recognise them.

Using the structural map, an approach is proposed to compare the cadastre structure and aerial photo. Our representation is physically defined by a graph in which nodes are objects and edges relations between them. The problem of comparison becomes an evaluation of matching (for corresponding zone) or subgraph detection (modified zone).

At objects level, the comparison will be performed by researching similar objects. The criteria are size, general orientation, intrinsic topology, lengthening... But it is important to note that due to imprecision of pictorial objects extraction, general aspect of objects can be far from real objects and first analysis in geometrical aspect detect easiest cases (zones wrongly merged, shapes abnormally distorted...). So if matching cannot be completed, relations focus attention. There are two reasons to use topological relations in our confrontation: first, comparison of relations is equivalent to evaluate the similarity of maps structure. This information is essential to determine if part of cadastre evolved. Second, the cadastre disposition is an indication to lead the process. If we detect (or show) similar objects as initial step, they could be used to reconstitute gradually cadastre from aerial photo.

Topological relations are interpreted to find correspondence in objects location and mixed approach using Allen and Egenhofer relations will be integrated. The last aspect deals with semantics used in comparison. Semantics are the way to obtain a description of real objects. This implies that aerial photo interpretation is processed to discover what objects have to be combined together to match geographical objects. Here, the textural information is used to make assumptions on features nature.

If the representation could bring large information, the problem of graph comparison is a complex problem. So as in [8], graph must be well defined to allow incremental comparison of graph parts.

CONCLUSION AND FUTURE WORKS

The problem of updating GIS and more precisely cadastre necessitates information that can be difficult to obtain. Aerial photos are an interesting source because they fix the study land, and scale is well adapted to urban zones. But if recognising objects seems easy for human, automatic process comes up against a problem of incompatible data sources. The methodology described in this paper focuses on problem of similar representation onto a same model. Raster and vector data (photo and cadastre) are used to construct tessellations (by Delaunay triangulation). With a topological model, the topological relations can be recovered and used to answer spatial queries but also for the reconstruction phase that enables to deal with high level objects.

The object-based approach linked to a strong topology constitutes the first base to a common representation, the *Structural Map* concept. This concept is the common structure that includes vector- and raster-issued data. Geographical and pictorial objects are described in a same way to be compared. The relations (hierarchical, topological, directional) define objects disposition in the map and they allow comparison in a more global level.

Future work would proceed by managing the advanced relations to move the comparison to objects level towards global structure of the map. If we imagine farther than differences detection, a complete updating process should propose scenarios to notify at an end-user the cadastre modifications. A knowledge level should integrate semantics in order to interpret objects as geographical objects rather than geometrical entities.

References

[1] Allen, J.F (1983). *Maintaining Knowledge about Temporal Intervals.* Communications of the ACM, 26(11). pp 832-843

[2] Egenhofer, M.J (1993). *A Model for Detailed Binary Topological Relationships.* Geomatica, 47 (3&4). pp 261-273

[3] Elomary, Y. (1994). Modèles déformables et Multi-résolutions pour la détection de contours en traitement d'images. Ph Dissertation (in French).

[4] Fleck, M (1991) *A Topological Stereo Matcher. International Journal of computer Vision* 6:3. Kluwer Academic Publishers. pp 197-226

[5] de Floriani, L., Puppo, E., Magilla, P. (1995). *Technical Aspects of Spatial Data.* Geographic Information Systems - Materials for a Post-Graduate Course. Vol 2: GIS Technology. Andrew U. Franck. Editor. pp 395-425

[6] Gold, C. (1996) *A Topological Approach to Processing Scanned Urban Maps.* Proc. 1st International Conference on Geographic Information Systems in Urban Regional and Environmental Planning. Samos. pp 156-166

[7] Laurini, R., Thompson D. (1992). *Fundamentals of Spatial Information Systems.* The APIC Series, Number 37. Academic Press.

[8] Nabil, M., Sheperd J., Ngu A.H.H (1995). *2D Projection Interval Relationships: A symbolic Representation of Spatial Relationships.* Proc. 4th International Symposium, on Advances in Spatial Databases (SSD '95). Portland. Lecture Notes in Computer Science (951). pp 292-309

[9] Tellez, B., Servigne, *Maintaining Cadastral Maps Using Aerial Photographs.* In: Proceedings of Jec'97 : Joint European conference on Geographical Information, Vol. 2. Pp. 956-965. Vienne, Autriche, 16-18 Avril 1997.