Updating topographic databases:
the example of the Topographic Database of Catalonia at scale 1:5000

Maria Pla, Santi Sànchez, José L. Colomer,
Institut Cartogràfic de Catalunya
Parc de Montjuïc
E-08038 Barcelona
http://www.icc.es
mariap@icc.es
santis@icc.es
colomer@icc.es

Introduction

The 1:5.000 Topographic Database is the 2.5D vector data set that covers the whole territory of Catalonia at the largest scale. Together with the color digital orthophoto at the same scale, it provides the reference base for geographic information.

The first part of the paper reports on the experiences gained during the updating and enhancing the database from a typical mid 80's plotting-oriented product to a modern GIS-ready topographic database. The second part is centered in the updating process of this GIS-ready topographic database to introduce the changes on the territory. Finally, there are some comments about the limitations in the used methodology and technology, and suggestions for improving the updating process.

The Topographic Database of Catalonia 1:5.000

The 1:5.000-scale Topographic Base of Catalonia consists in a set of 4275 sheets, each one stored in 2.5D files[1], covering the 32,000 km2 of the territory of Catalonia, in the northeast part of Spain. Until now the base has been compiled only by photogrammetry. Next version, due to start in late 2005, will contain a richer set of features and attributes coming from fieldwork.

The first version of the database: BTC-5M v1.0

The project started in 1985. At that time vectors were compiled with IGDS – the former CAD system of Intergraph – and MicroStation as DGN files on analog and analytical photogrammetric systems. The information was never prepared for GIS applications; on the contrary, the data model was what is known now as "spaghetti model" and was intended mainly for plotting and DTM generation. Therefore, some degree of latitude was tolerated as long as the map plotted on paper was looking good. In particular, we had to made tradeoffs because of aesthetic considerations due to IGDS or MicroStation limitations.

[1] Although the topographic base consists of physically individual files, we usually refer to it as "database".

The digital terrain model (DTM) was produced from the profiles, breaklines, elevation points and contours of the topographic files as a continuous grid with one elevation post every 15 meters. This DTM was used for orthophoto production and for shading maps at smaller scales.

The first version was completed in 1995 and the updating was tasked to start immediately. While designing the new version (version 2.0), it became obvious that we should try to enhance the data model so to built a modern, robust, exhaustively documented and error-free topographic database for GIS applications and mapping, and to support automatic generalization at 1:10,000 and at 1:25,000 scales.

The second version of the database: BTC-5M v2.0

Version 2.0 is based on topographic objects and includes polygons, hydrographic and communication networks, and city blocks in urban areas (Fig. 1). Each vertex is 2.5D and no tradeoff is allowed due to aesthetical considerations. The new data model contains more topographic objects and a better classification of the features. The associated DTM is obtained from the elevation data – profiles, breaklines, elevation points and contour lines – as in the first version, and from the topographic objects collected on the terrain. In addition, the digital surface model (DSM) is obtained by including buildings and other constructions.

Figure 1.- Example of the BTC-5M version 2.0. Note the auxiliary axes that form the communications, road and hydrographic networks and the centroids inside the polygons.

A significant improvement is the complete set of documentation, which includes technical specifications and detailed guidelines for data capture, and metadata. The guidelines describe each feature in detail, the source data, the geometry, the topology and relationships with other features, and the rules and parameters for compilation. The guidelines devote special care on the treatment of the elevation of the elements.

The geometries are represented by point, line and polygon. Three coordinates define each vertex. Lines representing hydrographic objects store the orientation. Each polygon is composed by lines representing the boundary and a point representing the centroid. Although the polygons have 3D coordinates in each vertex of the boundary, they are not 3D surfaces or 3D solid models. There are no duplicated lines in the data model; therefore adjacent polygons share the common line geometry. In the line intersections there is a vertex for each line, but not a node because it can be generated automatically. Topological consistency is checked with Intergraph's MGE/MGA GIS tools, although the DGN files are never stored on MGE databases.

From version 1.0 to version 2.0

We started updating in 1996 with digital photogrammetric systems. Such systems allow for the superimposition of stereo images and vector data thus facilitating the detection of changes and the compilation of new data. At first, it seemed pretty straightforward to load the old data onto new stereo models and compile the updates. We were aware of the difficulties because of the different semantics of the features, different data capture rules, data model and poor accuracy due to the use of analog stereoplotters. In addition, we knew about the tradeoffs that we assumed because of the restrictions of the CAD systems used for data collection:

- The limitations in the parallel copy of 3D lines forced us to store such lines as planar elements.
- The irregular line patterns along non-planar lines led us to compile such lines as planar elements.
- The automatic "connect to" digitizing operation performed always in 3D. This caused, for example, the margins of a river to connect to the bridge some meters above the watercourse.

Unfortunately, we ran into hidden problems that were neither known nor foreseen at all. For example, the inconsistencies because of local adjustments of geodetic network and also because of weak aerotriangulation methodology along the coastline. The sum of both effects produced problems when tying photogrammetric blocks together. This was solved in 1996 with a global compensation of the network and with a much robust aerotriangulation methodology.

After several attempts, no automatic methods were found for repairing and converting version 1.0 data to version 2.0. Manual removal of errors and inconsistencies and adapting to the new data model proved to be too arduous and time consuming. Note that because of the inconsistencies in accuracy and the incorrect elevations (artificial planar elements, effects of the "connect to", etc.) the old vectors were not superimposing well on top of the new stereoimages. The planar vectors represented, by large, the worse case because it was affecting directly the stereovision of the operators leading to discomfort, tiredness and reduced productivity. In such cases, the operator has a tendency to delete completely the vectors and collect them again from scratch. Differences in planimetry and in the amount of detail collected was not so problematical.

After evaluating the situation we finally decided to compile the version 2.0 from scratch with the exception of few object classes that could be migrated to version 2.0 directly.

Quality control

We decided that accuracy, quality and quality control would play a central role in version 2.0. Consequently we have set up quality control procedures at every job step, amounting a total of 25% of the effort devoted to the project.

The decision proved to be the correct one. Among other benefits such as improved customer satisfaction and less repetitions because of errors, it has enabled to put in operation a semi automatic generalization workflow that converts the 1:5000 database into a 1:25,000 2.5D topographic database (Baella, 2003). It is also interesting to note that during the benchmarking of systems for the generalization project, the database –actually DGN files– could be loaded without problems into different systems such as Dynamo (Intergraph), ArcSDE (ESRI), Lamps2 (LaserScan) and Smallworld.

Updating the 2.0 version

Version 2.0 started slowly in 1996 using photos from a 1995 flight. We had a slow start because of the training of the operators in the new workflow and in the new data model.

As a result, the area flown in 1995 changed so fast that, in 1998, the areas compiled with the 1995 photos were clearly outdated. Consequently, we tasked this area for update with a brand new flight. The same control and pass points used for the 1995 flight were used for the aerotriangulation – thus co-registering both flights together – and the same photo scale was used. Therefore, the new block of photos for aerial triangulation had the same shape and distribution as the previous one.

Both requirements –same photo scale, same block shape– have been abandoned recently when we had to add a new highway from a strip of images. The photo scale was around 40% larger and photos were in color instead of in black and white. We didn't use the points of the first triangulation at smaller scale but the ones from a database containing aerotriangulation points of large-scale projects – typically at 1:1000 – stored as image patches. Even the strip was aerotriangulated independently of the block, stereo-superimposition of the old vectors with the new images was very good.

Note that only updating –not repairing– had to be performed this time: data was now clean and consistent and the data model was the same. So it was the occasion to know if the whole methodology was good and enough to endure updating.

Updating operations

So far, we are using very simple criteria for updating: in parts with many changes, everything is deleted and collected again from the new image (Fig. 2). Note that if a map has become too aged, updating is practically equivalent to start from scratch.

Figure 2.- Left image shows data collected with aerial photoimages from a 1995 flight. Right image shows the updated data from the new aerial photoimage of the same area showed in the center.

In areas with less changes, the updates are performed by adding, modifying or removing objects as shown below:

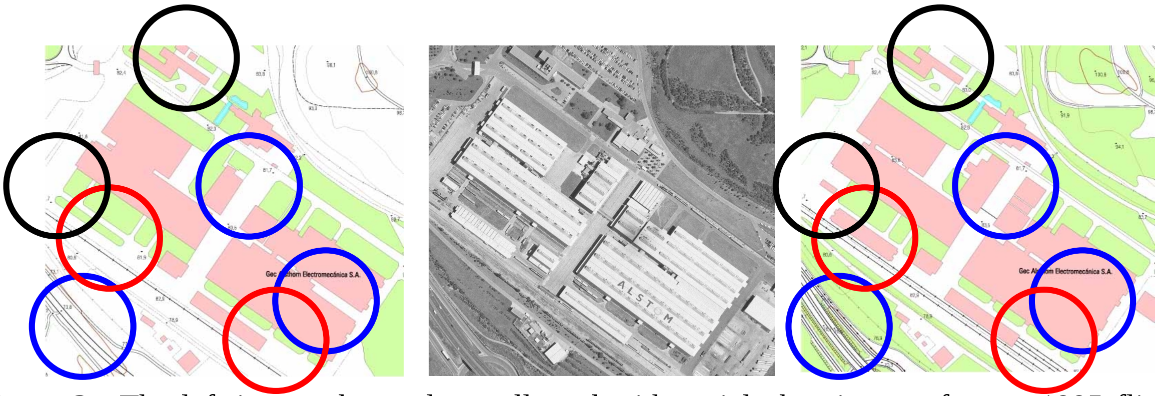

Figure 3.- The left image shows data collected with aerial photoimages from a 1995 flight. The image on the right shows the updated data from the new aerial photoimage of the same area showed in the center. Red circles indicate added information, blue circles indicate modified data and black circles indicate removed objects.

Productivity

The stereoplotting effort measured in hectares/hour with images spaced out 4 years is almost twice. As far as the editing and topological checking is concerned, results are far from being the same because no differences in productivity are observed. Editing of the base consists on validating the topological relationships by checking connectivity, correcting of overshoots and undershoots, etc. followed by the generation of polygons.

In summary, with our current tools and for a relatively outdated map, updating is two times faster in stereoplotting but the editing effort remains the same. The reason is that MicroStation doesn’t preserve topology and therefore the topological checking and cleaning must be repeated again.

As a result of all these experiences, we have started to look for a system that could store the topographic information in a seamless database and with a stereoplotting interface that could preserve the topology at updating. The system should maintain the ergonomics, comfort and productivity of the current CAD-based photogrammetric systems. High reliability and robustness and low prize per seat is also desired. At the time of this writing, there are announcements from different vendors that seem to accomplish these requirements. We are eager to test them and see if they can fulfill our needs.

The updating policy: selective updating

It is clear that an “outdated map” does not relate to time but to the amount and importance of the changes. In accordance with this, we look for important changes in the territory and update accordingly. More drastically, we are now updating selectively even before the country is completely covered by version 2.0. This is the case of updating the database with a new highway.

The main difficulty here is to decide what is to be considered “important”. Normally, changes in the communications network get top priority for updating, but while this might be enough for maps at medium scale –1:25,000 and 1:50,000– the larger 1:5,000 scale asks for a more ample concept which that shall include also changes in the type of coverage of the territory. This is the case, for example, of a complete change in a sector of a city. We have realized that the territorial planning agency is pretty aware of changes and, consequently, an excellent partner for helping us decide on their importance.

The ideal case is to update areas with a high rate of change almost every year. This happens, as expected, around the Barcelona metropolitan area. GIS-based applications for public transportation management including “what if” type of simulations are now based on the 1:5000 topographic base stored in an ArcSDE/Oracle environment. Even though it is used for display –there are neither user object classes sharing geometry with the base nor additional user defined attributes added to the original objects– the requirement is to update the database yearly.

Our current mechanism for handling updates is rather primitive: on demand, we deliver 3 separate files with data that has been kept untouched, the deleted data and the new data. A change in one object because of a previous error makes it appear in both the “deleted” and the “new” files. Of course we also distribute a file with all the changes included. It is easy to foresee that a better mechanism shall be put in place before customers do need something more intelligent than just bulk loading the updated version. Therefore, to the requirement of providing a photogrammetric interface, we have added versioning capabilities to our future system.

Conclusions

The paper describes the transition of the Topographic Base of Catalonia 1:5000 from a mid-80’s plotting oriented data set to a modern GIS and mapping oriented database. The main conclusion is that it is easy to be overoptimistic with technology and start planning for automatic data conversion procedures assuming that the quality of data is homogeneous and that all errors are known. What happens in reality is that a mix of different type of errors because of diverse reasons sum up together and to repair them may costs more than compiling from the beginning. Therefore, the decision of starting all over again is a real option. Note that the stereo photogrammetric environment doesn’t conceal any existing problem on data and that stereo-superimposition is very tiresome for operators if old vectors and new stereo-models don’t fit well together.

The second conclusion is that a photogrammetric interface to a rich GIS-oriented topographic database must exist because of the cost of re-checking topology or other relationships between objects at every updating cycle. Versioning and transaction-based distribution system for the updates are also mandatory.

References

Baella, B.; Pla, M. An example of database generalization workflow: the Topographic Database of Catalonia at 1:25.000. ICA Workshop on Map Generalization, Paris, April 2003.