

# COMPUTER ASSISTED PRODUCTION OF PLANIMETRIC DATA : AN INDUSTRIAL APPROACH

P. NONIN

ISPRS - Commission IV

ISTAR - Espace Beethoven - Route des Lucioles  
BP 037 - 06901 SOPHIA ANTIPOLIS - FRANCE

## PURPOSE:

Space imagery is a cost effective data source for small scale mapping. However, the raster data structure has to be converted into a more suitable vector structure for geographic information processing.

A semi-automated vectorizing system, running on a standard workstation, is described. Up to date image processing techniques (spectral/textural analysis, snakes) are employed to speed up planimetric extraction.

**KEY WORDS:** Cartography, Classification, Planimetry, Vectorization, SPOT.

## 1. INTRODUCTION

The availability of high resolution remote sensing satellites, namely LANDSAT and SPOT, and the set up of efficient data delivery facilities, have enabled us to acquire and update large image data bases from any part of the earth.

More over, the geometric properties of the CCD cameras shipped in those satellites allow cartographic data to be produced.

ISTAR, as several other companies in the world, has produced more than 200 geocoded images and digital terrain models computed from stereo couples of SPOT images since 1988.

On the other hand, the lowering prices of personal computers and workstations encourage digital information processing in any domain of activity and particularly in geographic applications.

Unfortunately, most of Geographic Information Systems commercially available, cannot handle effectively such unstructured data as raster images. Computer systems are not qualified, as humans are, to analyse the semantic content of satellite images. The best way to manage efficiently geographic entities is to convert raster images in highly structured sets of vector objects describing the boundaries of the land cover map and the network of linear features.

Traditional methods require a large amount of human interaction for the extraction of both altimetric (contour lines) and planimetric features. They also require expensive photogrammetric workstations.

Our purpose was the production of vectorized planimetric features from remote sensing images for geographic information systems users.

The data should be cheaper than traditional cartographic products and above all, recently updated.

A computer assisted vectorizing system, running on standard UNIX workstations, has been designed.

A single operator can process an entire SPOT scene (60 km x 60 km) within two weeks.

The data produced describe the geometry and the topology of the main planimetric features extracted from the raster image (urban areas, forests, transportation network...).

They do not describe "external" information (i.e. information not extracted from the images, ex: administrative boundaries or demographic data).

A great demand for such cartographic data can be

found in several domains like flight simulation, transportation planning, telecommunications, developing countries cartography.

## 2. PLANIMETRIC FEATURES

In order to reduce production (and commercial) costs, a standard general purpose cartographic product is proposed. It is a collection of vector elements with a layered structure. Each layer describes a particular type of features.

All possible types are listed below:

built-up [suburban, industrial] area  
transport [road, railway, airport, navigation] facilities  
wood  
waste  
bare soil [rock, beach, sand dune, quarry]  
water [sea, flooded area, marsh]  
snow

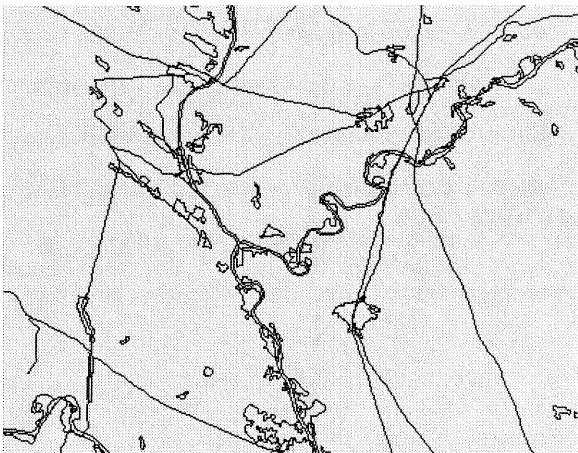
This list intends to meet the requirements of general rather than specialized needs. It is composed of two categories: linear layers (or networks) and area layers.

Each entity within a linear layer is defined as a list of bi-dimensional or tri-dimensional coordinates.

All areas are described by a single topological graph: An area entity is defined as a sorted list of arcs and nodes.



initial SPOT orthoimage



Planimetric features

Coordinates are related to a cartographic or geographic reference.

The nature and the density of the extracted features fit with those usually found in topographic 1:50 000 or 1:100 000-scale maps.

### 3. COMPUTER ASSISTED EXTRACTION

A standard product means a constant price and delay, whatever the amount of extracted features is. This goal cannot be reached with manual process: The time spent by an operator to acquire features manually is roughly linear with the amount of data. Our goal is double: Cutting down and smoothing production costs.

It is well known that fully automated extraction of relevant features in images is a hard work.

It implies a semantic analysis of the whole image.

Past and present research efforts have shown very little progress in that matter. Artificial Intelligence techniques together with image processing are involved but connection between these two domains is still difficult.

Our approach is more pragmatic, it is based on the following considerations:

A features extraction process is made of three sequential steps, identification - acquisition - quality control.

- Features identification is quite an easy task for human brain whereas automatic systems cannot achieve reliable results.

- On the other hand, acquiring accurately a lot of coordinates with a digitizing device is a rather slow and boring job. A computer assistance should certainly speed up the process if properly designed.

- Two kind of errors may be done while extracting features, geometric errors (a coordinate does not fit with the true feature position) and syntactic errors (topological graph inconsistency).

Geometric errors are easily detected when extracted data are superimposed on the raster image displayed on the screen.

Syntactic errors cannot be fully detected by visual control. An automated checking process is needed to get reliable results.

Those simple remarks have led us to design semi-automated tools making the most of both humans

and computers performances.

These tools perform supervised classification of land cover areas, delineation of linear networks and edition of topological relations.

General purpose computers have been chosen instead of specialized hardware because the performance growing rate of standard hardware is higher and its lower price enables a constant improvement of equipment.

#### 3.1 Land cover extraction

Well known spectral classification techniques produce good performances for vegetative and hydrographic items extraction (ref. Haralick).

They are quite inefficient for man-made features (like built-up areas) where spatial texture is much more significant than spectral information.

These techniques can be extended to texture classification if a texture analyser is provided (ref. Zhang). That seems to be useful for fine and regular urban structures.

Supervised or unsupervised segmentation methods based on region growing or split and merge algorithms have been tested on SPOT images (ref. Bretaudeau). The requirement of finely tuned parameters, depending on image properties, makes these methods impracticable in an industrial context.

In order to unify procedures and thus reduce operating complexity, we have implemented a single criterion for land cover extraction, that is maximum likelihood classification with normal distributions assumption and supervised training.

We think it is a rather general criterion, even if real spectral distributions are not often normal nor unimodal, because any statistical distribution can be approximated by a finite union of normal distributions. In that case, operator must define sub-classes during the training phase and then, merge the corresponding classified pixels.

The most important point is the training phase, the samples designation should be as fast and accurate as possible.

Our system uses interactive designation with visual feedback:

The operator draws, with a mouse, polygonal samples on the raster image displayed in a window. The classified pixels are displayed on request in real time (delay ~ 1 second). So, the operator can, at any time, test the fit of the samples set.

Classification may be operated in global or local mode. In local mode, the operator extracts interactively a single area entity pointed at, using the mouse.

In any case, the classification process can be thought as an explicit extension of a (implicit) description of the land cover given by human visual interpretation.

#### 3.2 Networks extraction

Manual vectorization, one point after another, of linear features is extremely time consuming.

Human operators, for bio-mechanical reasons, cannot achieve both precision and speed.

Automatic methods have been investigated in many

research laboratories (ref. Fischler, Groch, McKeown, Yee). The best known proceed hierarchically:

Pieces of linear elements, extracted by contrast or line detectors are progressively merged into more complex objects under topologic and morphologic constraints (connectivity, alignment...).

These methods have missed their goal, mainly because they only deal with geometric properties and they lack higher level knowledge about objects semantic and real world organisation.

This high level knowledge gives to the human visual cortex the ability to identify, at first sight, roads or railways in an image.

The snake concept take advantage of this ability in the following way:

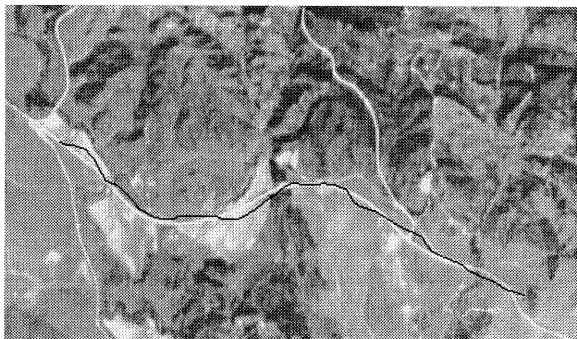
A feature, detected by an operator, is approximately drawn on the screen. This initial polygonal line is automatically moved (snake analogy) until it meets some optimal condition, depending on actual position of the real feature in the raster image.

Feature identification is human driven and geometric accuracy is ensured by the computer.

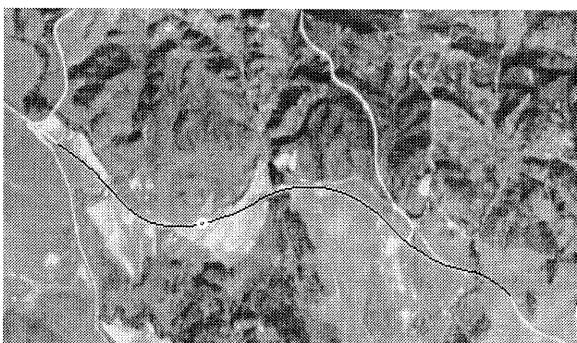
This concept has been implemented as a computer assisted interactive graphic editor.

Every polygonal element can be automatically adjusted to the nearer feature in the underlying image.

The algorithm is based on dynamic programming techniques. The vertices of the polygonal line are locally moved until radiometric properties along the line realize the best fit to a given model of linear feature.



initial hand drawn line



automatic adjustment

### 3.3 Topological relationship

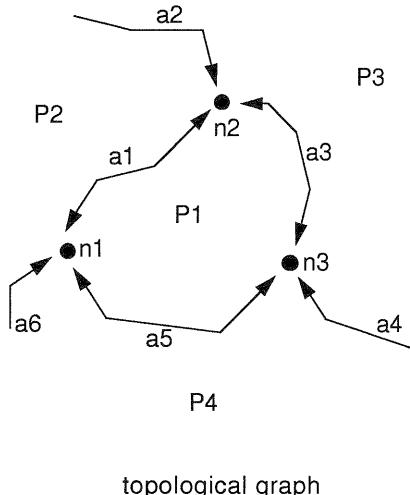
A topological graph is made of arcs (or lines), nodes and polygons.

An arc describes a single linear element (without crossing).

A node defines connection between two or more arcs.

A polygon is a sorted list of connected arcs defining a single closed area.

Each arc contains references to the polygon on its right and the polygon on its left. So, any spatial relation between polygons can be easily computed.



topological graph

The interest of topologically structured data is obvious for geographic information processing: It enables efficient spatial reasoning and filtering.

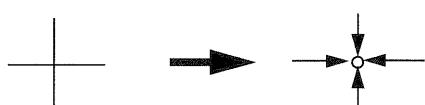
Up to date geographic information systems make use of topologic data files.

From a data producer point of view, topological relationship makes syntactic data control easier:

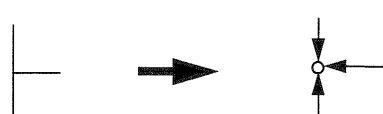
Network connectivity, land cover completion and areas boundaries consistency can be automatically checked.

For productivity sake, relation edition should not burden operators with an additional task.

We have integrated a topologic structuring capability in our graphic editor: The land cover graph is automatically generated, so are "X-crossing" connections within networks. "T-crossing" connections still have to be manually edited.



X-crossing connection



T-crossing connection

#### 4. FROM ALGORITHMS TO PRODUCTION LINE

The requirement of industrial data production implies the integration of the tools described above in a hardware and software environment including all steps of a production line from initial data input to sending of the final product.

Such an environment is operated by ISTAR since 1991.

The hardware configuration is based on a network of UNIX workstations, 10 Gbytes of SCSI disk storage and various I/O devices (CCT, streamer and 8 mm cartridges tapes drives).

The software environment is built on a raster/vector processing kernel with advanced visualization capabilities and a semi-automatic graphic editor.

Features data are extracted from orthoimages produced by ISTAR altimetric production line (ref. Renouard).

Since 1991, ten SPOT scenes have been processed. The cartographic data produced cover approximately 30 000 km<sup>2</sup> mainly in Europe and Middle-East countries.

This first experience provides a good estimation of production costs:

Land cover extraction from a geocoded SPOT scene (~ 3000 km<sup>2</sup>) requires about 1 man-week, whatever the density of features is.

For the same input, networks extraction requires from 1 to 1.5 man-week, slightly depending on the features density.

#### 5. CONCLUSION

The joint availability of high resolution raster data from SPOT or LANDSAT TM, and powerful geographic information systems (GIS), has induced demand for a new concept: data conversion from raster to planimetric vectors files fit for use in GIS.

The new product has to preserve the same marketing features as original data, that is set prices and delays. This is possible only if a computer assistance is supplied.

In order to quickly settle a computer assisted production line, we have designed a two level interactive system:

- High level tasks, implying semantic data analysis, are managed by human operators.
- Low level tasks, requiring a lot of accurate data processing, are done by the computer.

The performances allow the commercial production of cartographic data bases.

Future upgrades will tend to increase automation of networks extraction, especially in aerial images with higher resolution.

#### 6. REFERENCES

Bretaudau F. Une approche pour augmenter la séparabilité des textures. 7eme congrès de reconnaissance des formes et d'intelligence artificielle, Paris, November 1989.

Fischler M.A., Tenenbaum J.M., Wolf H.C. Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration

technique, Computer Graphics and Image Processing 15 : 201-223, 1981.

Groch W.D. Extraction of line shaped objects from aerial images using a special operator to analyse the profiles of functions Computer Graphics and Image Processing 18 : 347-358, 1982.

Haralick R.M., King-Sun Fu Pattern recognition and classification, Manual of remote sensing, American society of photogrammetry, pp.793-804.

McKeown M., Harvey W.A., McDermott J. Rulebased Interpretation of aerial imagery IEEE Workshop on principles of knowledge based systems, 1983.

Renouard L. Experiences with automated terrain extraction from SPOT data, Proceedings of the 10th Earsel symposium, Toulouse, June 1990.

Yee B. An expert system for planimetric feature extraction IGARSS 87 symposium, Ann Arbor, May 1987.

Zhang Zi-Jue, Shimoda H., Fukue K., Sakata T. A new spatial classification algorithm for high ground resolution images, IGARSS 88 symposium, September 1988, pp. 509-512.