Rome

Road detection from aerial images : a cooperation between local and global methods

Sylvain AIRAULT
Renaud RUSKONÉ
Olivier JAMET

Laboratoire MATIS
Institut Géographique National
2, avenue Pasteur
F-94160 Saint-Mandé
(France)
E-mail : airault,rusko,jamet@rennes.ign.fr

KEYWORDS

object recognition, aerial images, road detection, road following

ABSTRACT

We present in this paper a road detection method based on the cooperation between a road following algorithm and a global method which consists of a whole image segmentation followed by a characterization of regions (using shape and texture criteria). One will find there an overview of the approach, a description of the used techniques and then a presentation of the mean we use to make both methods cooperate. Both of our methods are based on quite the same road model (roads are elongated objects with parallel edges and an homogeneous texture) and both try to extract the same kind of roads : the interest of the approach is more to improve the detection reliability for a great part of the network than to look for an exhaustive extraction.

1. INTRODUCTION

IGN is leading researches in the field of automatic interpretation of aerial images, aiming at automating (at least partially) the 3-D data capture of a topographic database. This database consists of roughly the same information as displayed on traditional topographic maps, with a 3-D geometry and a geometric accuracy in the range of one meter r.m.s. along the three axes for every object. The capture process has been fully manual until now, using analogical photographs, but the emergence of digital stereoplotters, using digital images, will permit to introduce a part of automation in the manual data capture process.
One of our fields of work is the automatic extraction of the road network from digitized aerial photographs. We are working in two main directions : on one hand trying to give the operator a real time assistance, finding tools to locally extend manually started road portions ; on the other hand looking for a fully automatic extraction of the road network, which would not require the participation of an operator. Since the geometric accuracy of the final result must be in the range of one meter, we are working with high-resolution images (50-100 cm ground resolution).

2. APPROACH

2.1. Some considerations about the necessity of the cooperation

In all the interpretation processes working with remote-sensing imagery (maybe more than with other kinds of images), the use of a single technique based on a single modelisation of the searched objects will always lead to an incomplete result with a low reliability, because of the great diversity of contexts. Two explainations can be given to the failure of one method in the detection of one object :

- the object in the image is not similar to the modelised object. When the same kind of object (roads for example) may appear under very different forms (different kinds of surface, multiple lanes, presence of obstacles, ...), it becomes necessary to complete the detection using several models.

- the object in the image is similar to the modelised object but it is not detected because of a accidental bad behavior of the algorithm. There is not a "perfect" algorithm which would be insensitive to any kind of noise, not using any thresholds, modelising completely knowledge.... In this case, it is necessary to use a new algorithm to complete the detection with an implementation of the same model.

The first case will give a preference to exhaustiveness, detecting objects even if they appear under various forms ; the second case will give a preference to reliability, possibly detecting several times each object, since the same model has been used.
From this necessity of using several techniques and several modelisations to detect one object appears the difficulty to make these different algorithms "cooperate" and to mix several results to get the best benefit.

2.2. Some examples from previous works on the subject

Mc Keown [McKEOWN88] : this is the most complete study on the use of mixed techniques which presents the way to take into account several results from different algorithms. There are several low-level methods working independently to give an hypothesis for the road axis. An intermediate level makes evaluations concerning the success of each method. Knowing that failure modalities are presumed to be independent, the result of the evaluations can lead to the stop of one method which can be restarted from the hypothesis generated by one other successful method. While all hypotheses generated by each method are identical, the process is made straightforward. Upon detection of a divergence, the different hypotheses and confidence values are used to evaluate the relative goodness of each result.

System Messie [GARNESSON90] : there are two kinds of cooperation described here. The first one appears in the high-level validation of the hypotheses. There are several "specialists", for example one for the roads extraction and one for the buildings extraction, working independently and the validation of the hypotheses is performed by a "scene controller" using knowledges on the spatial relationships between objects (for example, the following rule : there is always a road leading to a house).
The second kind of cooperation concerns the low-level extraction. Two segmentations are performed (region growing and boundaries extraction). Both results are merged into a single set of segments and are equally used by the interpretation tasks.

De Gunst [DEGUNST91] : the result of a classification is used to control the propagation of a road-following algorithm. The first idea was to generate a mask with areas where the presence of a road is possible : a multispectral classification is performed on the Spot images which gives a classified image and a "probability" file. This file contains the distance of each pixel to the road class center. Because of the absence of texture on the road surface, roads should have "high probabilities". The distances are thresholded and dilated to give a mask wider than the road itself. But the method is not robust enough to be sure that all roads are in the masked area and the count of changes in the classification along the road candidate is only used as additional information for roads interpretation.

Groch [GROCH82] : two different algorithms are used here but without a common use of their results : the second method is only applied when the first one fails. The extraction begins from starting points with a local method (radiometric profile analysis). At locations where large gaps interrupt the continuation of an object, a regional method is applied to bridge the gap.

2.3. Our approach

2.3.1. The road-following algorithm

As long as automatic interpretation processes do not reach an excellent reliability level, we feel important to keep an interactive mean of intervention. That is the reason why we initially choose a local method, a road-following algorithm (fig. 1), supposed to be integrated in a complete data capture process where automation would be taken as a help to a manual, operator driven interpretation [AIRAULT93].

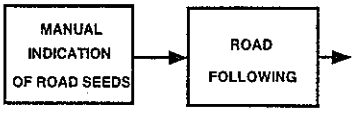

fig 1 : interactive use of road following

2.3.2. How to increase the reliability of our detection ?

Our road-following algorithm is principally based on the optimization of a textural criterion, roads being considered as homogeneous areas. Used globally on the whole image, this criterion allows to discriminate very well roads from very textured

areas as urban areas or vegetation areas but does not allow to discriminate roads from other homogeneous areas as fields. Used locally to propagate the road hypothesis from good starting points, it works quite well in many cases because of the presence of edges separating the road from the neighboring field. But, in some cases, if the road edge is not visible and if the field is more homogeneous that the road itself, it can lead to a regrettable failure.
This insufficiency of the textural criterion can be erased by the complementary use of a shape criterion (fig. 2). The shape of roads, which are very elongated objects, allows to discriminate them from fields, which are often more compact areas.

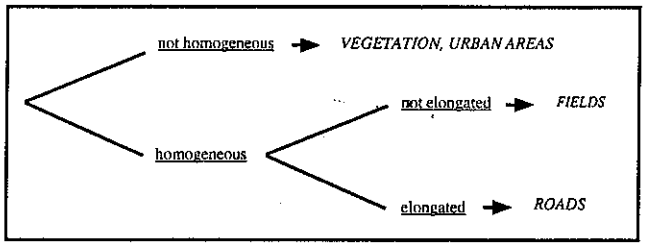

fig. 2 : discrimination of roads using a textural criterion and a shape criterion.

The analysis of shapes require a view of the image rather large to be sure to consider the whole objects. That is why global methods such as segmentation techniques performed on the whole image seems to be more adapted to the shape characterization than a too local test which could be integrated in the road-following module.

2.3.3. Integration of pre-computed data in the real-time detection task

Even in an interactive system as the one we are thinking of, it is possible to integrate any kind of pre-computed data. If the display of the detected objects is still determined by the operator, a part of the computation can have been performed before on the whole image. Figure 3 illustrates the process.

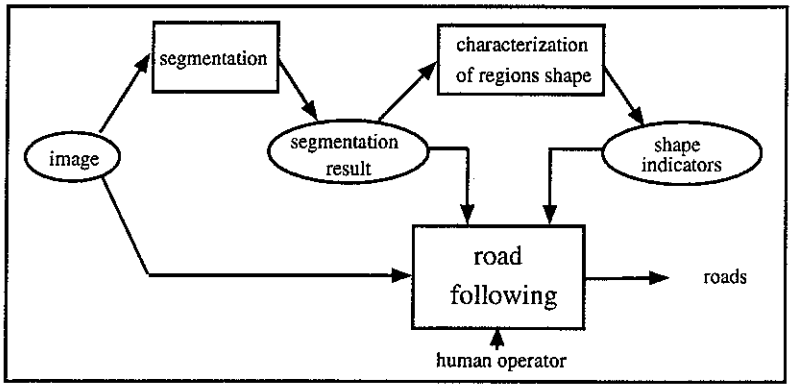

fig. 3 : integration of pre-computed data in the real-time detection task

3. LOCAL APPROACH : ROAD FOLLOWING

The road following algorithm is mainly based on the optimization of a textural criterion. To improve exhaustiveness versus reliability one tries to keep as little constraints as possible in the model, using minimal a priori knowledge (few geometrical constraints and no hypothesis on the gray-levels). Beginning with the knowledge of roads as elongated homogeneous areas, we optimize a directional homogeneity criterion [AIRAULT94] on a long enough distance to be characteristic of the specific texture of roads, which is not isotropic (roads are elongated objects and their homogeneity may be measured along the road axis). Optimizing the criterion on a important set of segments allows to fit the road shape and to cross small obstacles.

3.1. Local computation of homogeneity

The gray-level variance is computed in the possible propagation directions on elongated neighborhood with variable length (fig. 4). The next possible locations are computed according to the length and the direction of the neighborhood which minimizes the variance.

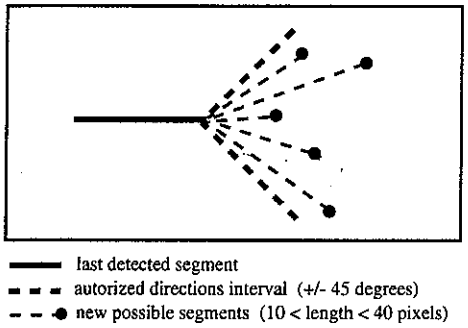

fig. 4 : local computation of the homogeneity

3.2. Choosing the best path

Using this measure of the local homogeneity, we compute at every new location on the road a tree of possible paths (a path is a set of segments) by retaining at every level of depth the most homogeneous segments around the current propagation direction. Once the tree has been built (fig. 5), the best path is determined as the one which minimizes the variance, maximizes the total length, and which is the most rectilinear (this last component of the cost function must not be considered as a geometric constraint, it rather gives a little preference to the most regular path when there are many paths with the same homogeneity). Keeping just the best path through the tree allows to make the road-following more reliable, and to cross high curvature areas (fig. 5), or obstacles (change of road surface, trees, ...).

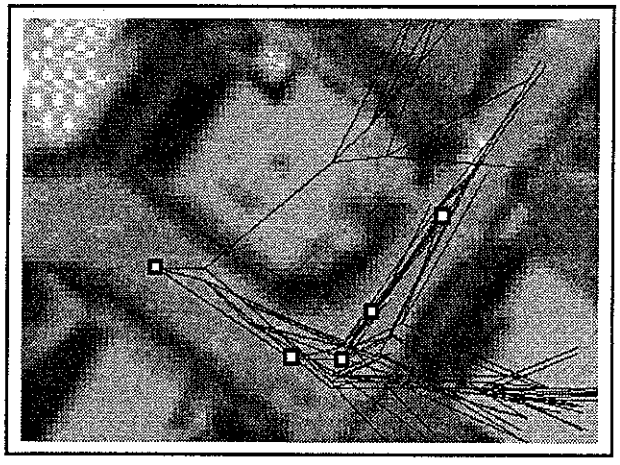

fig. 5 : tree of the possible paths
(white squares for the best path)

For each of the dir^{depth} paths in the tree (where *dir* is the number of retained directions at each level of depth and *depth* the total depth of the tree), the cost is calculated as follows :

$$COST_{path} = \frac{\sum_{i=1..depth} \left(variance_i \cdot \left(\Delta dir_i \cdot rigid \right) \right)}{\sum_{i=1..depth} \left(length_i \right)}$$

with :

variance$_i$ = variance along the segment of the depth i
length$_i$ = length of the segment of the depth i
Δdir$_i$ = variation of direction between the segment i and the previous one
rigid = rigidity coefficient

3.3. Road-edges detection

At the same time, a road-edges detection is performed, on one hand to adjust the axis of the road (if possible according to the quality of the detection), on the other hand to compute confidence indicators on the extracted network geometry, in view of a global adjustment. The used technique consists in assimilating locally the road-edges with line segments, and to look for the best pair of parallel lines in a gradient computed through the direction perpendicular to the propagation direction.

3.4. Stop criteria

The road-following process can stop on several criteria : too high variance, rough change of the mean radiometry, or wide dispersion of the seeking tree. The latter is the most discriminant, taking into account the anisotropic aspect of the roads homogeneity (avoiding, for example, that the plotting get lost in a field whose radiometry would be close from the one of the road).

4. GLOBAL APPROACH : CHARACTERIZATION OF SEGMENTATION GENERATED REGIONS

The method we present here consists of a characterization of the regions with the aim to quantify the possibility of a road presence. Since the textural information has already been used, we only try to separate roads from other homogeneous objects whose shape is more compact. The method is composed of three steps : segmentation, split of elongated regions and computation of shape indicators.

4.1. Segmentation

The used segmentation process is based on a watersheds delimitation in a gradient image. This computation uses a Deriche's gradient [DERICHE87]. The density of informations which appear on an aerial photograph makes the segmentation step difficult : to keep all important boundaries, it's necessary to accept an over-segmented result where each little detail of the image has been represented ; to obtain a result where only important boundaries appear, it is necessary to accept that some of them will be lost in the final result, too noisy to be kept. In our case, the second solution has been adopted (under-segmentation) to choose parameters for the two following reasons :

- we try to give a preference to reliability against exhaustiveness and it is better to loose some road edges than to keep many regions whose shape could be confused with the shape of a road.
- to compute a shape indicator, it is more significant to work on large dimensions regions.

4.2. Split of elongated regions

After the segmentation step, we try to split the elongated regions :

- to be sure that the shape of resulting regions will not be ambiguous. Without this step, if the shape indicator is computed on a region which cross the whole image, this indicator can be quite good even if the region has locally a compact shape (fig. 6).
- to make the surface area a discriminant criterion : after this step, a region whose surface area will be greater than a given threshold, will be considered as a bad region because if it had been a road region, it would have been divided before.

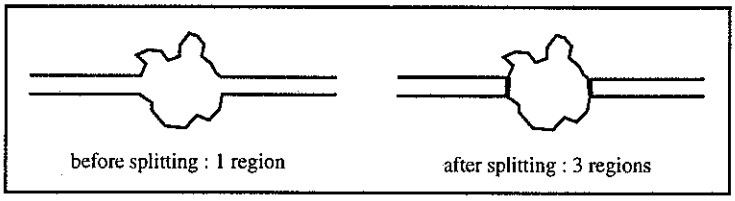

fig. 6 : split of one region into three regions with "homogeneous" shapes.

Splitting is done as follows : for each point p_i of the boundary of the region, the Euclidean distance $d_{i,j}$ to every other points p_j of the boundary (if the distance along the boundary is greater than a threshold *T1*) is computed. If the minimum of $d_{i,j}$ is lower than a threshold *T2*, the region is divided with the introduction of a new boundary between p_i and p_j. Both resulting regions are then tested to be divided again. *T1* represents the maximum width of a road and *T2* is a threshold on the maximum elongation of the final regions (giving an approximation of the surface area over which a region will never be a road).

4.3. Computation of the shape indicator

The method we use consists in the computation of the centered second order moments which allows to know the orientation of the shape and gives a measure of the inertia along the principal axis and along the orthogonal axis. The ratio between both measures of inertia will give an indicator of the shape elongation.

If p and q represent the order of the moment (the order is p+q) :

$$M_{pq} = \iint (x-\bar{x})^p (y-\bar{y})^q dx\, dy\ F$$

The second order moments (M_{20} ,M_{02} and M_{11}) characterize the shape orientation :

$$\theta = \frac{1}{2} . \arctan\left(\frac{2\ .\ M_{11}}{M_{20} - M_{02}}\right)$$

where *theta* is the angle between the principal axis of inertia and the x axis.

We can then compute the moments according to the new axes :

$$M = \cos^2\theta . M_{20} + \sin^2\theta . M_{02} + \sin 2\theta . M_{11}$$

$$M' = \sin^2\theta . M_{20} + \cos^2\theta . M_{02} - \sin 2\theta . M_{11}$$

and the elongation indicator :

$$elong = \frac{max(M, M')}{min(M, M')}$$

On top of this, little regions and large regions are eliminated from the computation : little ones because their shape is considered as not significant and large ones because we already know that they are not good road candidates (see 4.2.).

Remark ;

Our initial idea was to use for the regions classification not only a characterization of their shape but also to take into account properties related to their neighborhood relationships, defining rules to describe the spatial relationships between geographical objects. For example, a presumed "road" region will get a better reliability if a neighboring object of the same class is found with a similar orientation.
We chose a model (as defined before) with three classes : roads, compact homogeneous regions and textured regions with few rules describing their neighborhood relationships. The computation of this classification was based on a probabilistic relaxation algorithm [ROSENFELD76]. At every iteration the probability for a given region to belong to a given class is updated, taking into account the region characteristics and a compatibility measure related to neighboring regions. That gives for every region, when the convergence process is complete, a measurement of its probability to belong to each class of the model.
But it gave not so good results as expected for several reasons :

- model too simple.
- difficulty of defining rules.
- process too dependent of the segmentation quality.

5. COMMON USE OF ROAD FOLLOWING AND SEGMENTATION

5.1. Three questions

Trying to make techniques of different nature work together is quite an issue. The difficulties outlined are :

- to select the level at which outputs will be collated.
- to evaluate the success rate of each method.
- to define decision rules to select (should a conflict arise) the output to be retained.

As for the collation level of outputs, one may try to match either low-level segments (using boundaries issued from several segmentations [GARNESSON89]) or higher level objects which would be road portions (collation of several road extractors [McKEOWN88]).

As for the evaluation of the success rates, two approaches are possible : either the different algorithms involved evaluate their own level of performance by means of a quantification attached to the criterion they optimized (with an often low reliability level), or the evaluation is managed externally by a higher level routine able to match the outputs with an independent optimized criteria model or to cross-compare the results [McKEOWN88].

As for the decision rules, one may either drop the unmatched extraction outputs (retaining only matched one), or retain only the evaluation-wise best result, or generate a common output taking divergences into account.

5.2. Pieces of answer

Our problem is such : we have a road following algorithm well adapted to the semi-automatic approach we aim at and, on the other hand, which offers a good balance between exhaustiveness and reliability. We hold as well a segmentation output with shape indicators which is believed to be able to further enhance the detection reliability. How should we use that material ?

The approach we retained is to keep the interactive aspect of road following (with real time display to allow for operator control) and thus to collate our two methods at low level, every "path" (road following wise) taking segmentation data into account. We will use the cost value attached to the road following paths together with the "probability" of the regions they cross to be roads, according to their shape. With so local data it is impossible to get a workable validity criterion and, taking into account the uncertainty pertaining to the input sources, it is not desirable to enforce any kind of consistency between them. Better than a too strict decision rule, we would rather balance the cost value along road following paths against the probability of being a road along the path, according to the regions shape.. Since some road-edges are lost in the segmentation result (see 4.1.), it becomes better to use the shape criterion as a relative measure to order several road hypotheses than to take it as an absolute indication of the presence of a road.

This choice is not optimal in terms of cooperation as it does not allow for a reliable validation of outputs as it may be expected from such a technique. Data of very different natures however are allowed to be taking into account : a mix of output from different extraction techniques is allowed, as well as integration of pre-computed data in a real time process.

The new cost function we use is a linear combination of the old cost (variance computation, see 3.2.) and of a function using the shape elongation measure of the crossed regions. This choice is totally empirical. We have shown, using a rather large set of points belonging to the three classes (roads, urban & vegetation areas, fields), that this combination allows to discriminate quite well roads from other areas (see 6.3.).

6. RESULTS

To make the evaluation of the discrimination possibilities for each criterion (texture and shape) and for the combination of both, we have chosen around 800 points belonging to our three classes (roads, urban & vegetation areas, fields ; around 250 points each) from a 2000 x 2000 image. We have then computed both criteria and drawn the histograms class by class for each criterion independently and for the combination of both.

6.1. The textural criterion

A local measure of the gray-values homogeneity (see 3.1.) allows to see immediately roads and fields being separated from urban and vegetation areas (fig. 7). But, even optimized on a large distance from a good point on the road, the criterion can sometimes show its insufficiency when the road-edges are not well visible and when the neighboring region is a field more homogeneous than the road itself (fig. 8). The left part of the figure shows the already detected portion of the road, which is adjacent to a field. The right part show the tree of possible paths with the best path (accordind to the variance measure) going in the field. This confusion appears on the histogram computed on the set of points (fig. 9), where there are two peaks in the low values, one for roads, one for fields.

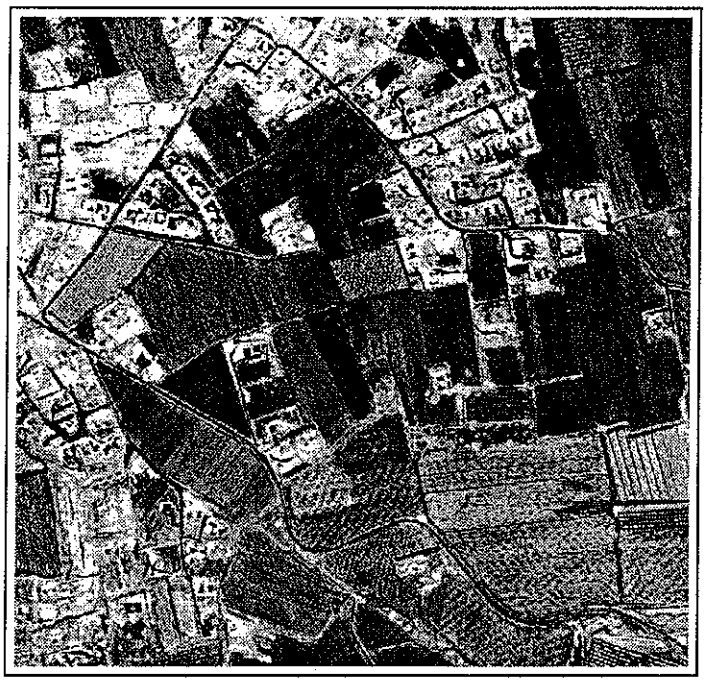

fig. 7 : gray-levels variance computed on the whole image.
(minimum with 30 pixels length neighborhoods in 32 directions)

fig. 8 : insufficiency of the homogeneity criterion.

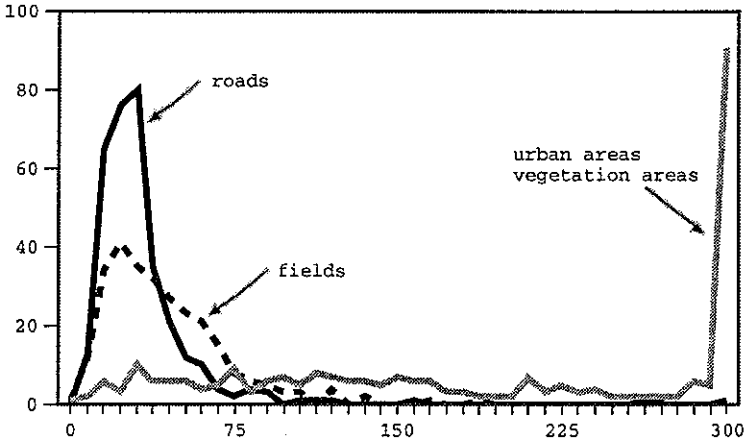

fig. 9 : histogram of the variance cost values. (x axis : variance ; y axis : population)

6.2. The shape criterion

The segmentation result is not complete since some road-edges have been lost. But the shape indicator (fig. 10) allows to separate well-segmented roads from the fields which are always characterized by a compact shape or by a too large surface area which has not be divided (see 4.2.).

fig. 10 : shape elongation indicator computed on the whole image
(white : compact shape ; black : elongated shape)

On the histogram, there are two populations for roads : one in the low values which represents well segmented roads and one in the high values which represents roads with at least one missing edge (fig. 11).

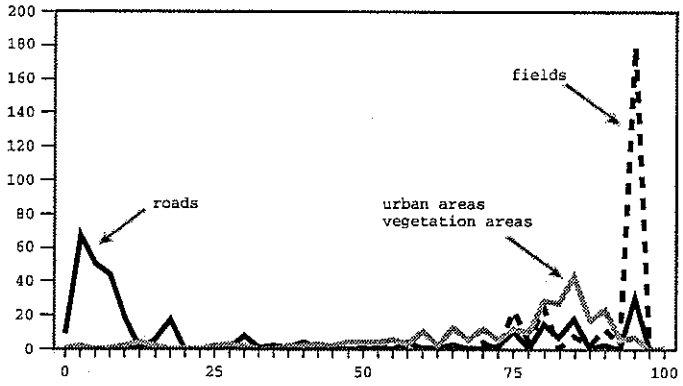

fig. 11 : histogram of the shape elongation indicator
(x axis : shape indicator = (10 - elong)*10 ; y axis : population)

6.3. Combination of both criteria

The histogram of the combination of both criteria shows three distinct peaks, corresponding to our three classes (fig. 12). A second peak appears on the roads histogram, corresponding to the wrong-segmented roads with a bad shape indicator. But the residual confusions are not really a problem since the criteria are used as a cost function to evaluate relatively several hypotheses.

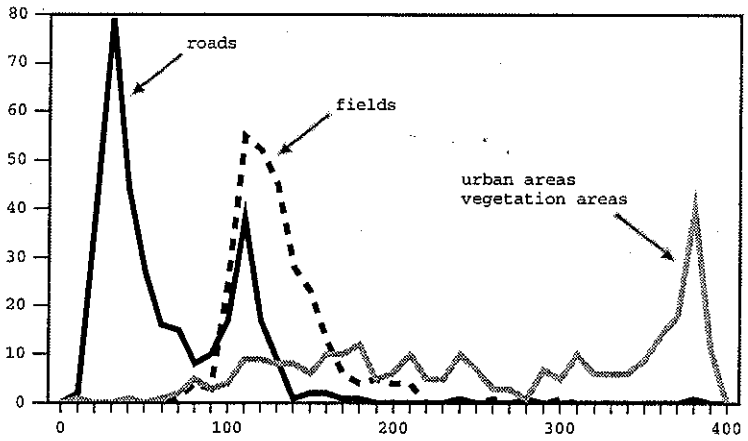

fig. 12 : histogram of the linear combination of both indicators
(x axis : homogeneity indicator + shape indicator ; y axis : population)

The final result (fig. 13) we obtained is the semi-automatic road network capture on the whole image from manually indicated starting points (around 20 starting points are necessary for this capture). The combination of both criteria gives to the road-following algorithm a more "determinist" behavior, less sensitive to the starting-points location and to the noise on the gray values.

fig. 13 : semi-automatic data capture on the whole image

Another criterion combining texture and shape

The same test has been performed (same classes, same set of points) with the criterion we use as stop criterion in the road-following algorithm (see 3.4.) which is the spatial dispersion of the tree of possible paths. This measure combine texture and shape because each segment of the tree is found according to the textural constraint and the dispersion of all segments shows if the texture is isotropic or not. The measure is the root mean square of the distance from the tree to the mean path. As we can see on the histogram (fig. 14), this measure could be very interesting but we are still working to find the best way to use it during the propagation not only as a control but to to generate and to compare hypotheses.

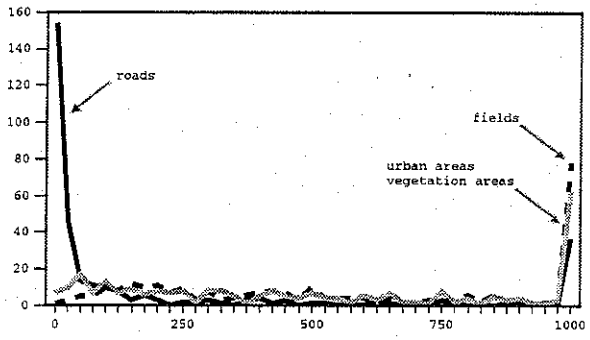

fig 14 : histogram of the spatial dispersion of the tree
(x axis : r.m.s. on the whole tree ; y axis : population)

7. CONCLUSION

The proposed approach do not give a very important answer concerning the cooperation between algorithms. We rather presented the common use of multiple criteria for an operationnal approach where an operator would be present. In this context, the low-level criteria combination is a more important problem than a real cooperation using high-level knowledge, since the final decision to validate a road segment can be taken by the operator. In this way, our method give good results.

An extension of the proposed approach would be to integrate other pre-computed data in the interactive process. The road following algorithm may be only used to generate many hypotheses with the homogeneity criterion and these hypotheses may be evaluated using independently several pre-computed data sets. "Independently" means that it would be interesting, for an interactive use, not to try to combine all criteria but rather to keep as many road hypotheses as criteria and to let the operator choose between them.

REFERENCES

[AIRAULT93] S. AIRAULT & O. JAMET. Automatisation de la saisie des routes sur images aériennes. Bul. S.F.P.T.,n.131, pp. 12-23, December 1993.

[AIRAULT94] S. AIRAULT & O. JAMET. Détection et restitution automatiques du réseau routier sur images aériennes. In: RFIA's proceedings, Paris, France, Vol. 1, pp. 519-531, January 1994

[DEGUNST91] M. E. DE GUNST, "Automatic extraction of roads from Spot images", Graduation report, Delft University of Technology, july 1991.

[DERICHE87] R DERICHE. Using Canny's criteria to derive a recursively implemented optimal edge detector. International journal of Computer Vision, may 1987.

[GARNESSON89] P. GARNESSON, G. GIRAUDON & P. MONTESINOS. Messie: un système multi-spécialiste en vision. Application en imagerie aérienne. INRIA research report, april 1989.

[GROCH82] W. D. GROCH, "Extraction of line shaped objects from aerial images using a special operator to analyze the profiles of functions", Computer graphics and image processing, n. 18, pp. 347-358, 1982.

[McKEOWN88] D. McKEOWN & J. DELINGER, "Cooperative methods for road tracking in aerial imagery",Proceedings of Computer Vision and Pattern Recognition, Ann Arbor, june 5-9, 1988.

[ROSENFELD76] A. ROSENFELD, R.A. HUMMEL & S.W. ZUCKER, "Scene labeling by relaxation operations", IEEE transaction on systems, man, and cybernetics, vol. 6, n. 3, pp. 420-453, june 1976.