Perceptual Grouping and Active contour functions for the extraction of roads in satellite pictures

Laurent Alquier, Philippe Montesinos
LGI2P - Parc scientifique G. BESSE
NIMES, F-30000
email : alquier@eerie.eerie.fr

December 6, 1996

Conference title : Image and Signal Processing
Conference chair : Desachy J. at satellite remote sensing symposium

ABSTRACT

We present a new method for Perceptual Grouping of pixels into roads after crest lines detection in satellite pictures. First the visual properties expected from the groupings are modelled as a quality function similar to active contour functions. They involve curvature, grey levels and co-circularity.

This function is computed recursively and optimized from a local to global level with an algorithm related to dynamic programming. The final groupings are then selected according to their global quality.

Applied to satellite images, the method proved its adaptability and its robustness to noisy environments. The results showed how the use of visual properties can provide an effective segmentation with no prior knowledge of the scene. This segmentation can be used to initialize a high level interpretation process or give a first description of the scene to an interactive decision system.

Keywords: Perceptual Organization, Dynamic Programming, Roads Extraction, Segmentation, Remote Sensing.

1. Introduction

In aerial or low resolution satellite pictures, roads or other thin networks such as river beds correspond to crest lines of the image intensity function. These salient linear structures can be detected by low level processings. However, this detection is biaised by noise or gaps, introduced by the influence of textures or noise. We propose in this paper a method for the extraction of these salient structures after crest lines detection. This method uses Perceptual Grouping techniques with the optimisation of active contour energy functions.

The use of Perceptual Grouping techniques in Computer Vision is not a recent idea. In his theory of Vision, Marr[7] suggested in 1976 the idea of a "primal sketch" involving both information from contour segmentation and grouping of primitives such as curves or lines.

Perceptual Grouping[6 12 14 11] in Computer Vision applies some observations made by Gestalt Psychology,[13] or psychology of shape. With the help of psychovisual experiments, Gestalt Psychology showed how human vision structures the representation of visual information in a picture instead of seeing it as independant primitives. Human Vision tends to group naturally the visual primitives according to basic geometric relations before even starting to analyse the shapes represented in the image. These relations can range from continuation or symmetry to similarity or object-background separation. When translated to Computer Vision, this complex combinatorial problem is mostly expressed as grouping together image primitives into larger structures (such as chains or regions). Algorithmic and optimization techniques have been developped according to the complexity of the primitives to group.

In order to apply Perceptual Grouping experiments to Computer Vision, Lowe[6] defined a statistic feature for each grouping, representing the probability to be accidental. This probability is low for salient and organised structures in the scene. Lowe's work showed how Perceptual Grouping could be used to efficiently structure images produced by poorly accurate or biased low-level processing, and effectively prune the search space of curves in images. Later work involved functions taking into account the saliency of curves as the human eye would do.[9 12 3] In the domain of optimization techniques, models of active contours or Snakes[4] have been developed for the detection of structures containing weak edges ; an energy function composed of opposed terms representing internal and external influences, controls the shape of a contour on the image. As the energy function is not convex, there is no direct way to find its minimum and an initialization close to an optimal solution is required.[1] Finally, an interesting iterative scheme related to Dynamic Programming was presented by Sha'ashua[11] for the optimization of quality functions in Perceptual Grouping. This scheme allows a local to global optimization for some classes of functions.

The method described in this paper falls within the scope of Perceptual Organization with optimization techniques. First in section 2., visual properties expected for groupings are defined as 'Snakes like' functions taking into account internal and external influences. These functions represent the quality of a grouping. Our method is based on a two-stage algorithm which first aggregates the primitives with regard to the quality function. A procedure related to dynamic programming[2 11] optimizes this quality function from a local to a global level. The main groupings are then selected semi-automatically according to their global quality. Finally, results on synthetic and real images are presented in section 5..

2. Perceptual grouping with active contour functions

Our goal is to extract salient curved lines from pictures after crest lines detection, using only geometric properties. The pertinence of crest lines detection is increased by organizing pixels according to a quality function relating their saliency in the image. Solutions to this optimization problem are the best matchings between the low level primitives and quality functions representing the visual properties expected from these groupings.

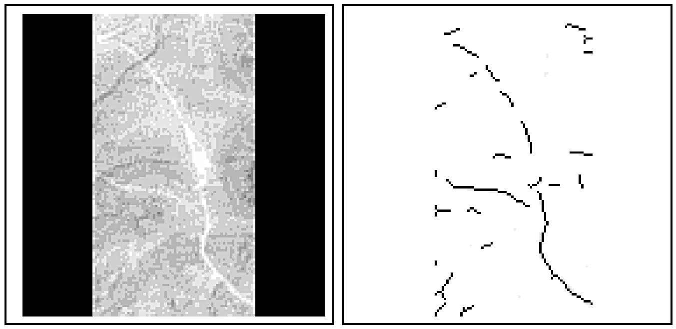

Figure 1: Crest lines detection

First, quality terms have to be defined. They represent the visual properties stressed for "good" groupings of pixels. The kind of property suggested by the groupings depends on the type of scene ; in the case of satellite pictures, roads are expected to be long, continuous and smooth curves. These quality terms follow the same formalism as the energy functions in models of "snakes". The final quality function is composed of an internal contribution (the own visual quality of the grouping) and an external contribution (the shape constraints imposed by the image on the grouping). The opposition of these terms gives its stability and robustness to the method. However, in contrast to the snakes approach, our optimization scheme is done locally.

The main external influence of the image on the solutions is represented by $(\mathcal{G})$, a function of the grey levels of the pixels along the curves. This matching can be reinforced if the tangents along the solution follows orientations $(\mathcal{O})$ of the low level primitives when available. However the attraction imposed by pixels and local orientations is very important and has to be counter-balanced by shape constraints. Curvature $(\mathcal{C})$ is an elementary way to describe the shape of a curve. In order to obtain stable results, early experiments lead us to add a co-circularity term $(\mathcal{K})$. According to the importance given to this term, it is possible to preferentially extract loops or open curves. The complete quality function is defined as a linear combination of four quality terms and can be written, for a given pixel P:

$$\mathcal{Q}(P) = \alpha \cdot \mathcal{C}(P) + \beta \cdot \mathcal{K}(P) + \gamma \cdot \mathcal{G}(P) + \delta \cdot \mathcal{O}(P) \tag{1}$$

Where ($\alpha, \beta, \gamma, \delta \in [0, 1]$) are parameters rating the influence of the different terms. They control the class of curves to extract (open curves or loops for instance). The importance and sensitivity of each parameter are discussed in section 3.2..

3. Optimization and grouping

Each quality term is written as a bi-lateral function associating a pixel to a pair of its neighbours, in order to define a trace *coming in* and *going from* this pixel. The following recursive expression of these terms allows more efficient computation and a global optimization (a detailed discussion about the quality terms and algorithms can be found in[5]).

For a curve arriving in a pixel P, a quality term $\mathcal{F}$ is defined as the sum of local terms along the path of the curve crossing P. A factor $0 \leq \rho \leq 1$ weights these local terms and represents the attenuation of the quality with distance. Written as a bi-lateral function, with a trace $\mathcal{F}_l(P)$ coming in P and a trace $\mathcal{F}_r(P)$ going from P, the quality becomes:

$$\mathcal{F}(P) = (\mathcal{F}_r(P) + \mathcal{F}_l(P)) \tag{2}$$

with :

$$\begin{aligned} \mathcal{F}_l(P) = \quad & \tfrac{1}{2} \cdot Q(P) \\ & + \rho \cdot Q_P(P-1) \\ & + \rho^2 \cdot Q_{P-1}(P-2) + ... \end{aligned} \tag{3}$$

and :

$$\begin{aligned} \mathcal{F}_r(P) = \quad & \tfrac{1}{2} \cdot Q(P) \\ & + \rho \cdot Q_P(P+1) \\ & + \rho^2 \cdot Q_{P+1}(P+2) + ... \end{aligned} \tag{4}$$

This term written in a recursive way gives, for a distance n starting from P:

$$\mathcal{F}_l{}^{(n)}(P) = Q_P(P) + \rho \cdot \mathcal{F}_l{}^{(n-1)}(P-1) \tag{5}$$

where $Q(P)$ is the local quality term for the pixel P and $Q_P(P-1)$ represents the evaluation of a contribution from the pixel $(P-1)$ viewed from P. Each term of this quality function is representative of a long distance measure of the quality of the curve. The distance n is given by the number of iterations.

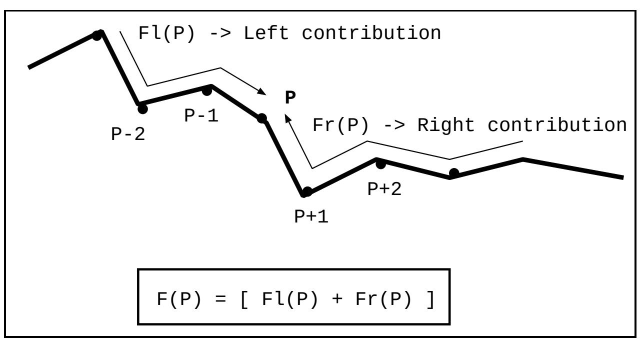

Figure 2: Notations used for a quality term on a dynamic curve during the optimization

A method related to Dynamic Programming[2 10] is presented here to optimize the quality function iteratively, from a local to a global level. For each pixel, a pair of neighbors defines a connection ; they represent the directions of arrival and departure for a possible curve crossing the pixel. The pair of neighbors giving the best value for the quality function is selected. The recursive expression of the quality functions makes it possible to compute their values with a local part (defined by the local characteristics of the connection) and a global contribution provided by each neighbor. This global contribution takes into account more distant pixels for each iteration.

Along the iterations, the importance of individual pixels of impulse noise decreases with regard to pixels included in large structures. The quality of pixels within gaps is increased by influence of neighboring structures.

This mechanism increases the global level for the quality functions along the iterations, filling gaps and smoothing solutions to the optimization as they take into account more and more distant pixels.

3.1. Computation of the connections

For each primitive and each of its neighbours (called *input*) we first select the neighbor giving the best quality (called an *output* for the connection). There can be multiple inputs associated to the same output. More formally, the connections are computed in two steps: 'inputs toward outputs' then 'outputs toward inputs'.

- **Connecting inputs toward outputs**

  For each primitive P and each input Q_i, we look for an exiting primitive Q_o which maximizes the quality function at step n in P : $\mathcal{F}^{(n)}_{Q_i,Q_o}(\,P\,)$

  By construction, we have the following relationship between Q_i and Q_o :

$$Q_o \;= exiting(\; Q_i,\; P\;) \tag{6}$$

  such as :

$$\mathcal{F}^{(n)}_{Q_i,Q_o}(\;P\;) \;=\; \max_{Q_o{}' \in V(P)\backslash\{Q\}} \mathcal{F}^{(n)}_{Q_i,Q_o{}'}(P) \tag{7}$$

  One can notice that the connection of inputs toward outputs is not symmetrical: for an entering direction Q_i and a primitive P, $exiting(\;Q_i,\;P\;)$ is the only exiting direction, the quality function of which is maximum but the opposite is false.

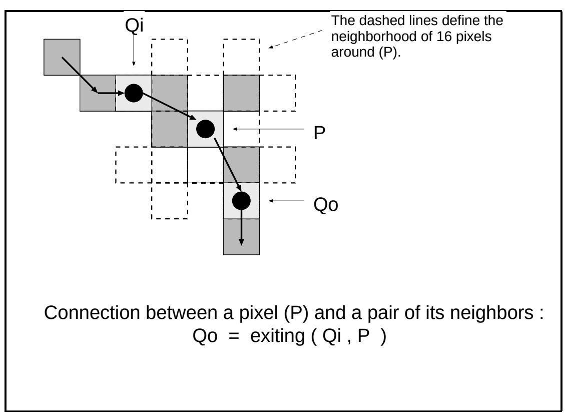

Figure 3: Connections for the grouping of pixels

- **Connecting outputs toward inputs**

  To be sure to select always the best connection, we associate, when possible, each output with the input giving the best quality. The influence of inputs comes from a long distance as the choice of output is influenced by the local noise. Thus, connecting outputs with inputs results in less sensitivity to noise than would exist by connecting inputs to outputs only. This can also be viewed as a correction of the previous relationship in order to chose always the best pairs of neighbors.

We are going to define the *entering* function ($i = entering(Q_o, P)$) which gives, for any exiting direction Q_o of a primitive P, a corresponding entering direction i. There are two possible cases: one or several entering directions may exist, or no corresponding input exists.

Case of no input:

In this case, a given output Q_o has no corresponding input. We consider Q_o as a possible input and we keep the connection already defined between inputs towards outputs. We define the function *entering* as:

$$entering(Q_o, P) = Q_i \quad with \quad \mathcal{F}^{(n)}_{Q_i,Q_o}(P) = \max_{Q_o' \in V(P)\setminus\{Q\}} \mathcal{F}^{(n)}_{Q_i,Q_o'}(P)$$

Case of multiple inputs:

In the case of multiple inputs, we have to make a choice between L possible inputs: we define the *entering* function, such that the quality function in P with this output Q_o is maximal for the input Q_k among the possible inputs.

We keep the pair (Q_o, Q_k) such that $\mathcal{F}^{(n)}_{Q_k,Q_o}(P)$ is maximum.

We can now define the input Q_k, the optimal neighbour entering in P. Between this connection "output toward input" (Q_k, Q_o) and the previous connection "input toward output" $(Q_k, exiting(Q_k, P))$, we keep the connection that gives the best quality function.

Let: $Q = exiting(Q_k, P)$
We define eventually :

$$entering(Q_o, P) = \begin{cases} Q_m & \text{if} \quad \mathcal{F}^{(n)}_{Q_m,Q_o}(P) > \mathcal{F}^{(n)}_{Q,Q_o}(P) \\ Q & \text{if not} \end{cases} \tag{8}$$

The complete optimization algorithm is described in Algorithm 1. The algorithmic complexity is directly related to the number of pixels of interest in the image after low level processing, as well as the number of neighbors around each pixel (to give a better set of possible directions around each pixel, a neighborhood of 16 pixels is used. See fig.3).

3.2. Control and behaviour of the algorithm

The quality function can be controlled by two sets of parameters.

The parameters for each term in the final linear combination represent the importance given to these terms in the quality function. They are respectively related to the influence of curvature, co-circularity, strength of crest lines and orientation terms on the shapes of the selected curves. For a high value given to Grey Levels, the curves selected will tend to be attracted more by detected pixels. A high influence of Cocircularity will give a better quality to loops instead of open curves. It can take a certain number of trials to optimize exactly the class of groupings expected, but once the correct settings are found, the detection remains significantly good for different images with the same parameters.

The second set of parameters is used in the recursive expression of each term. They represent the influence of distant contribution for the term (for example, ρ in (2) and (3)). A value of 0 will reduce the corresponding term to its local value only.

During the optimization process, the number of iterations is related to the distance between contour elements we want to connect. For example, the length of the widest gap gives us the minimum of iterations required to

Algorithm 1: Algorithm of optimization and selection.

```
begin
    Initialization of quality functions and connections at step 0 ( n = 0 )
    %
    % Iterations
    For each Pixel do
        For each entry do
            Local optimization of the quality function : computation of the connections
            Update of the quality function
        endfor
    endfor
    %
    % Following and selection the best paths
    For each Pixel P do
        Follow the connection from P
        Compute quality of the grouping
        Update GlobalQuality( P )
    endfor
    For each Pixel P ∉ Groupings already selected do
        if ( GlobalQuality( P ) > Threshold ) then Select Current Grouping
    endfor
end
```

fill gaps along the curves of the image. We must remember that the optimized groupings tend to be smoothed as they receive more global contributions. Thus, a high number of iterations means a loss in the precision of the selected curves. It can be interesting to select the best groupings at various levels of precision as they represent increasingly more global results.

4. Selection of the best solutions

Once the optimization has been performed, the curves are extracted by following the connections from one pixel to another until the chain of pixels crosses its own trace, reaches a boundary of the image or comes to a dead end (for example, when the curve has traced a certain distance without encountering pixels from the crest lines detection). As the chains are smoothed by the optimization, they are replaced by C-Splines curves for a better description. These splines can be used later to initialize active contours and provide an optimized selection of the solutions. The optimization reduces the number of possible curves on the image to a single optimized curve for each possible starting point. Pixels of high local quality are most likely to belong to large structures ; they give the starting points of a first selection of good solutions.

This selection is refined with regard to the global quality of the solutions. The global quality of a curve is the sum of local qualities for each of its points. This definition divides the possible solutions into groups of curves with equivalent qualities. It is possible to reduce the number of curves in each group by weighting their quality with an additional factor. This factor is obtained by considering the amount of pixels from the original image encountered by the chain with regard to the amount of pixels followed by filling gaps in this image. A simple thresholding of this weighted function is generally enough to automatically select the best groupings.

The difficulty of the selection of the best grouping comes from the presence of local maxima of quality (if we consider for example the function associating each pixel to the quality of the curve extracted starting from this pixel). To improve this extraction, future developments should involve an “hypothesis-verification” loop between the original image and solutions extracted at different iterations. Independant groupings can't pretend to describe roads completely at this point as they don't take into account topological aspects. However, the set of selected groupings gives a first description of the scene to help the interpretation process.

5. Results and Conclusion

Grouping has been applied first applied to synthetic pictures to show the effect of noise on the method (fig 10 and fig 11). It has been then applied to satellite images (256x256 pixels). To show the robustness of the method they have been applied with the same parameters (the choice of parameters is dependant more on the kind of curves expected than the image itself). The segmented images in fig 5 and 8 show the result of crest lines detection[8] . It is important to remember that the selection algorithm extracts salient solutions according to their quality. This explains the missing groupings one can notice on the results images. It's always possible to obtain more groupings with a lower selection threshold.

More appropriate results to each scene can be obtained by further tuning of these parameters (for example, increasing the importance of curvature and reducing the importance of grey levels gives better shapes). The extraction of main groupings automatically with no prior knowledge about the shapes produces a first description of the scene and a good initialization for further higher level processing such as model based shape recognition or active contours optimization.

The high quality of the results shows how the method is robust to noise. Our method can be easily applied to real situations with large images such as roads or rivers detection. However, the computer time and memory required can become large enough to require the use of an hybrid involving the grouping of pixels for dense areas of large images and the grouping of chains of pixels for more simple areas.

It can be used to initialize a high level interpretation process, for the identification of the area by matching the groupings with a model of a larger picture. It can also propose optimized solutions as a first description of the scene to a human expert, allowing local decisions on the solutions.

In the future, we plan to focus on more complex groupings such as chains of pixels and adapt this method to the extraction of 2D and 3D curves.

6. Figures

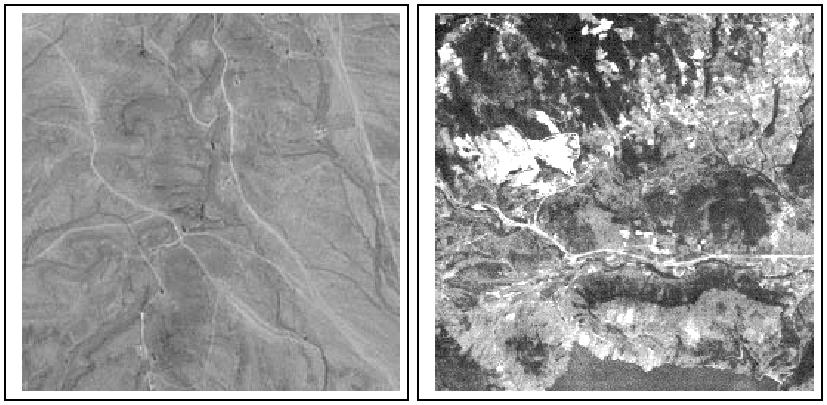

Figure 4: Mountain scenes - Original Images

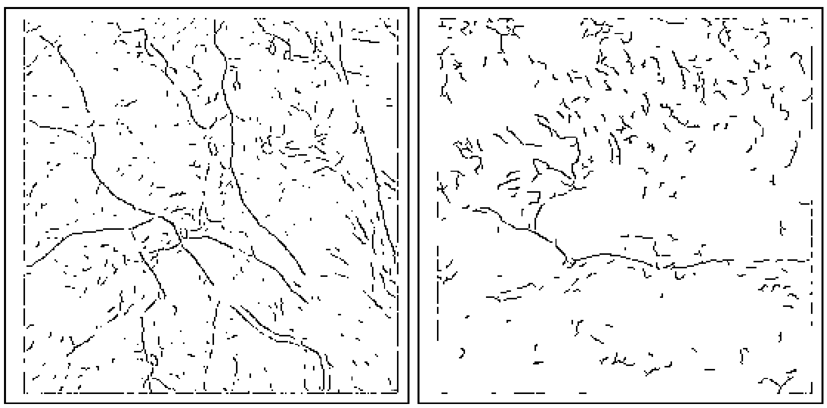

Figure 5: Mountain scenes - After crest lines detection

Figure 6: Mountain scenes - Final selection of the main groupings

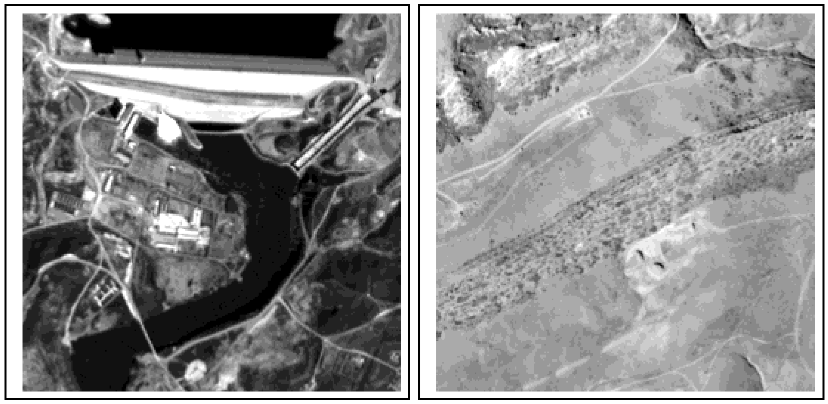

Figure 7: Dam and Desert scenes - Original Images

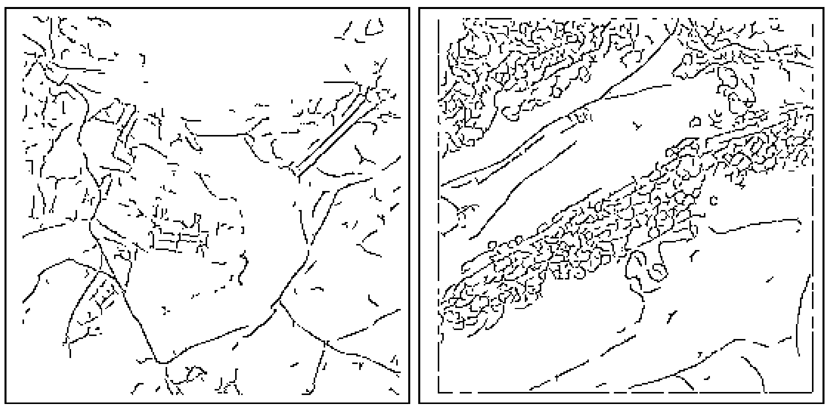

Figure 8: Dam and Desert scenes - After crest lines detection

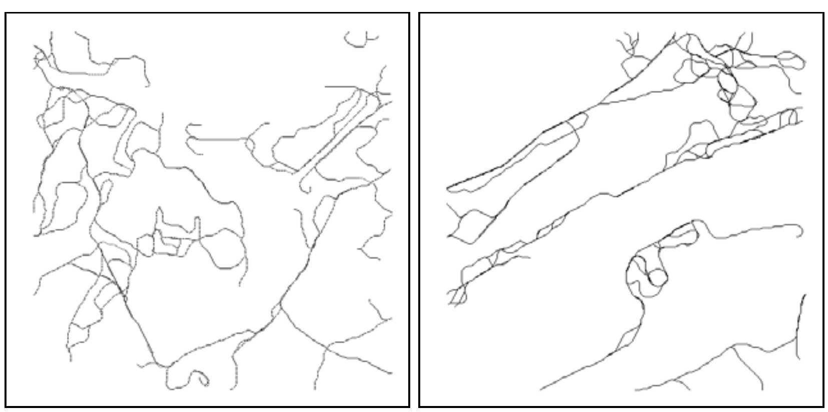

Figure 9: Dam and Desert scenes - Final selection of the main groupings

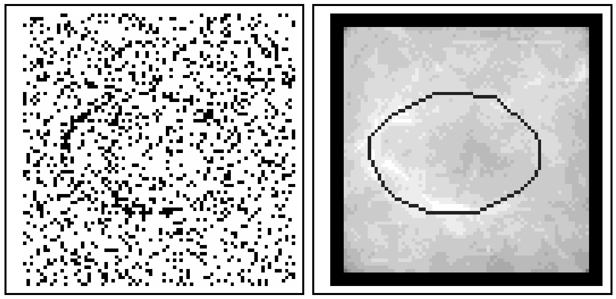

Figure 10:

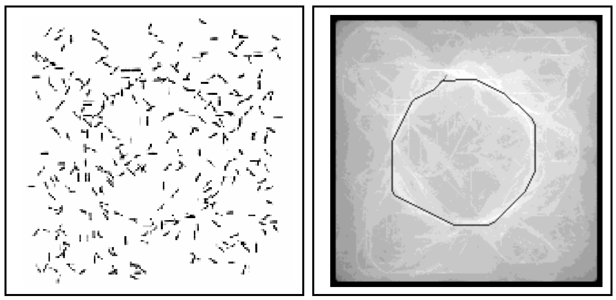

Figure 11:

7. REFERENCES

[1] M. O. Berger. *Contours actifs : modélisation, comportement et convergence.* PhD thesis, Institut National Polytechnique de Lorraine, 1991.

[2] D. P. Bertsekas. *Dynamic Programming : Deterministic and Stochastic Models.* Prentice-Hall, INC., Englewood Cliffs, N.J. 07632, 1987.

[3] L. Hérault. *Réseaux de neurones récursifs pour l'optimisation combinatoire.* PhD thesis, Institut National Polytechnique de Grenoble, Février 1991.

[4] M. Kass, A. Witkins, and D. Terzopoulos. Snakes : Active contour models. In *Third International Conference on Computer Vision*, pages 259–268, June 1987.

[5] L.Alquier and P. Montesinos. Perceptual organization with active contour functions : application to aerial and medical images. Technical report, Laboratoire de Génie Informatique et d'Ingénierie de Production, Nîmes, France, 1996.

[6] D. G. Lowe. *Perceptual Organization and Visual Recognition.* Kluwer Academic publisher, Hingham MA 02043, USA, 1985.

[7] D. C. Marr. *Vision.* Freeman, Oxford, 1982.

[8] O. Monga, N. Armande, and P. Montesinos. Crest lines and thin net extraction. In *SCIA*, volume 1, pages 287–295, June 1995.

[9] P. Parent and S. W. Zucker. Trace inference, curvature consistency, and curve detection. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 11, August 1989.

[10] A. Sha'ashua and S. Ullman. Structural saliency: The detection of globally salient structures using a locally connected network. In *IEEE, Second International Conference on Computer Vision*, Tampa Florida, 5 dec 1988.

[11] A. Sha'ashua and S. Ullman. Grouping contours elements using a locally connected network. In *Neural Information Processing Systems*, 1990.

[12] Y. C. Shiu. Experiments with perceptual grouping. In *SPIE, Proc. Intelligent Robots and Computer Vision IX : Algorithms and Techniques,*, volume 1381, Boston, Massachusetts, 5-7 Nov 1990.

[13] M. Wertheimer. Untersuchungen zur lehe von der gestalt ii, translated as: "principles of perceptual organization". In *Readings in Perception, 1958*, pages 115–135, Princeton, N.J, 1923.

[14] S. W. Zucker, A. Dobbins, and L. Iverson. Two stages of curve detection suggest two styles of visual computation. In *Neural Computation*, volume 1, pages 68–81, Massachusetts Institute of Technology, 1989.