Segmentation of thin networks using Perceptual Organization with active contour functions

Laurent Alquier

LGI2P - Parc scientifique G. BESSE
NIMES, F-30000

Philippe Montesinos

LGI2P - Parc scientifique G. BESSE
NIMES, F-30000

December 9, 1996

ABSTRACT

This paper describes a new method of perceptual organization of thin networks using geometric properties. The key point of our approach is to consider perceptual organization as a problem of optimization : solutions to this problem are the best matchings between continuous curves and the low level primitives.

First the quality of a grouping is defined with a class of functions related to the energy functions of active contours optimization. Such functions are computed recursively, and optimized from a local to a global level with an algorithm related to dynamic programming. This is followed by a selection procedure which rates and extracts principal groupings automatically and gives a new segmentation of image primitives, based on smoothed continuation.

This segmentation is used to initialize a high level interpretation process involving projective reconstruction of 3D contours in sequences of images. The adaptability and robustness of this method have been tested on various situations, such as the extraction of ellipses from indoor scenes, roads from satellite pictures or blood vessels from medical images.

Keywords: Perceptual Organization, Dynamic Programming, Non convex Objects Extraction, Segmentation.

1. Introduction

The detection of lines or thin networks is an important problem of image analysis. In this paper we propose a method of Perceptual Grouping for the extraction of such networks, as it applies to the detection of roads in aerial images or blood vessels in medical images.

The psychological origin of Perceptual Grouping[6,11,13,10] in Computer Vision was Gestalt Psychology,[12] or psychology of shape. Based on psychovisual experiments, Gestalt Psychology showed how human vision structures the representation of information in a picture instead of seeing it as independant primitives. Such groupings organize the visual primitives according to basic geometric relations. These relations can range from continuation or symmetry to similarity or object-background separation. This complex combinatorial problem is mostly expressed as grouping together interest points into larger structures (such as chains or regions). Other approaches are possible, according to the nature and complexity of the primitives used for grouping. These primitives can be grouped according to undetermined shapes (such as segments, curves, or regions) or parametric shapes (circles, squares, ellipses). Finally, these groupings are used to more easily initiate a model-based shape recognition system. Very few attempts have been made to group larger structures.

The use of Perceptual Grouping techniques in computer vision is not a recent idea. Since 1976, Marr[7] suggested the idea of a "primal sketch" involving both information from contour segmentation and grouping of primitives such as curves or lines. However, this idea has never been implemented and Perceptual Grouping has not been really used in Computer Vision before Lowe's work. Lowe[6] defined for each grouping its own statistic feature, representing the probability to be an accidental grouping. This probability is low for salient 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[9,11,3] and an interesting iterative scheme allowing a local to global optimization for some classes of functions was presented by Sha'ashua.[10] In the domain of optimization techniques, models of active contours or Snakes[4] have been developed for the detection of structures containing weak edges. According to these models, the trace of a contour is modified iteratively to minimize an energy function, split into independent and yet opposed terms which represent internal and external influences. It also controls the motion of the contour trace. This model, however, shows an important weakness. As the energy function is not convex, there is no direct way to find its minimum. Iterative methods (such as Gradual Non Convexity Algorithm[1]) are available, but they require an initialization close to an optimal solution. Finding this initialization automatically is not easy for most applications, particularly when different objects are present in a scene at the same time.

The method described in this paper falls within the scope of Perceptual Organization with optimization techniques. First in section 2., the quality of a grouping is qualified with 'Snakes like' functions taking into account internal and external influences. 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[2,10] optimizes this quality function from a local to a global level. The optimal groupings are then selected according to their global quality. Section 3. discusses the behaviour of the algorithm and the use of the different parameters. Finally, results on synthetic and real images are presented in section 4..

This method is able to recognize large linear structures in noisy images. The results show how it is possible to extract most of the networks with no prior high level knowledge of the scene. Only very general geometric relations are used, such as curvature and co-circularity. The results also show the applicability of our approach and the interest of our method for initiating a scene interpretation process.

2. Perceptual grouping with active contour functions

Due to noise, weak strength of edges, occlusions or textures, low level processing, such as edge segmentation or crest-lines detection, is often spoiled with gaps and false detection. In order to increase the pertinence of this information, we want to organize this data into smooth and continuous groupings. This is described as an optimization problem ; solutions to this problem are the best matchings between the low level primitives and quality functions representing the geometric properties expected from these groupings.

As stated before, a quality function composed of an internal term and an external term is defined, following the same formalism used for the energy functions in models of "snakes". However, in contrast to the snakes approach, our optimization scheme is done locally. Primitives are organized iteratively together all over the image during the optimization process. The method can be applied to any image primitive (segments, edge elements, ...) as long as quality terms can be defined to group primitives together. As showed in the results, the method has been tested for groupings of pixels and of chains of pixels.

The first step in this method is the choice of quality terms. Each quality term represents a gometric property expected to be high for a "good" grouping between a primitive and its neighbors. 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, whereas straight lines and corners would be more expected for indoor scenes. It also depends on the choice of primitive to group together (the length of chains is an important factor for the grouping of chains, while there is no such information available for the grouping of pixels).

For example, the grouping of pixels for medical and satellite applications leads to the following choice. The main external influence of the image on the solutions is represented by a function of the grey levels of the pixels along the curves in the segmented image after crest lines detection. This matching can be reinforced if the tangents along the solution follow orientations of the low level primitives when available.[8] However the attraction imposed by pixels and local orientations is very important and has to be counter-balanced by shape constraints. Curvature 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. According to the importance given to the cocircularity term, it is possible to preferentially extract loops or open curves.

In the case of a grouping of chains, external influences of the image are functions of the length of chains and also the distance between the extremes of the chains to group together. The shape constraint is represented by the curvature of the link between the chains (a polynomial curve generated with respect to the orientations of each extreme to link together - see fig.2 about the grouping of chains).

The complete quality function is defined as a linear combination of the quality terms. Each quality term is normalized in order to allow a better control of their influence by a parameter. The importance and sensitivity of each parameter are discussed in section 3.

2.1. Optimization of the quality function

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

A quality term $\mathcal{F}$ of a curve arriving in a pixel P is defined as the sum of local terms along the trace of the curve entering in and the curve exiting from P, with a factor $0 \leq \rho \leq 1$ representing the attenuation of the quality with distance. If we write the relation as a bi-lateral function of the trace, with a trace $\mathcal{F}_l(P)$ coming in P and a trace $\mathcal{F}_r(P)$ going from P, the quality becomes:

$$\mathcal{F}(P) = (\mathcal{F}_r(P) + \mathcal{F}_l(P)) \tag{1}$$

with :

$$\mathcal{F}_l(P) = \begin{array}{l} \frac{1}{2} \cdot Q(P) \\ +\rho \cdot Q_P(P-1) \\ +\rho^2 \cdot Q_{P-1}(P-2) + ... \end{array} \tag{2}$$

and :

$$\mathcal{F}_r(P) = \begin{array}{l} \frac{1}{2} \cdot Q(P) \\ +\rho \cdot Q_P(P+1) \\ +\rho^2 \cdot Q_{P+1}(P+2) + ... \end{array} \tag{3}$$

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

$$\mathcal{F}_l^{(n)}(P) = Q_P(P) + \rho \cdot \mathcal{F}_l^{(n-1)}(P-1) \tag{4}$$

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.

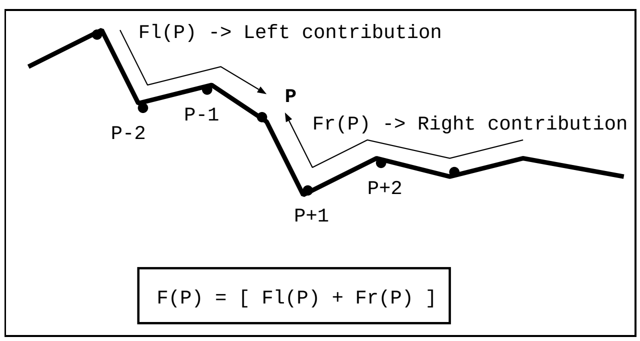

Figure 1: Notations used for a quality term on a dynamic curve during the optimization

2.2. Optimization of the quality function

The quality function is optimized iteratively, from a local to a global level, using a method related to Dynamic Programming[2].[10] For each primitive, a connection is defined by a pair of neighbours representing the directions of arrival and departure for a possible grouping between the primitives. The pair of neighbors giving the best value for the quality function is selected.[5] 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. Along the iterations, the importance of individual primitives decreases with regard to primitives included in large structures.

2.3. Computation of the connections

For each primitive and each of its neighbours (called *input*) we first select the neighbor giving the best quality. The neighbor selected is 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 :

$$Q_o = exiting(Q_i, P) \tag{5}$$

  such as :

$$\mathcal{F}^{(n)}_{Q_i, Q_o}(P) = \max_{Q_o{}' \in V(P) \setminus \{Q\}} \mathcal{F}^{(n)}_{Q_i, Q_o{}'}(P) \tag{6}$$

  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.

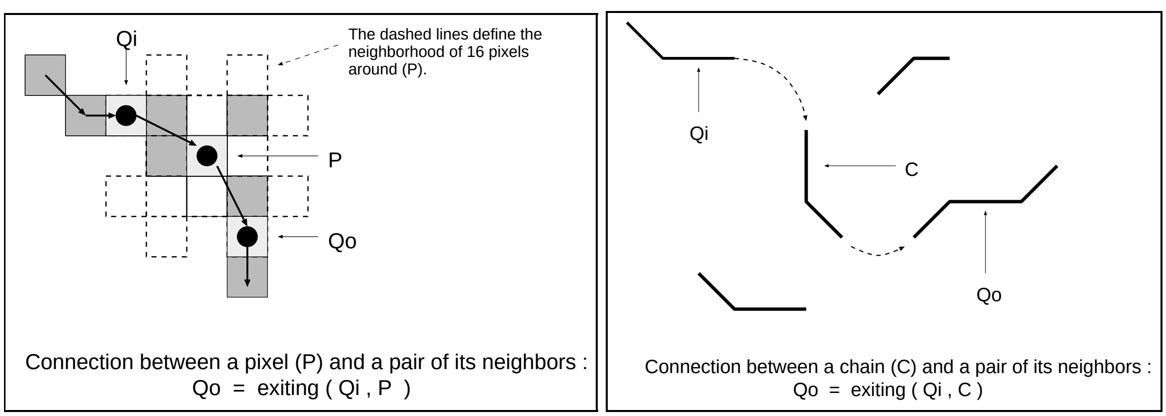

Figure 2: Examples of groupings depending on the choice of primitive : Pixels and Chains

- **Connecting outputs toward inputs**

  To be sure to select 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.

  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 :

$$entering\,(Q_o, P) = \begin{cases} Q_m & \text{if } \mathcal{F}^{(n)}_{Q_m,Q_o}(\,P\,) > \mathcal{F}^{(n)}_{Q,Q_o}(\,P\,) \\ Q & \text{if not} \end{cases} \tag{7}$$

The complete optimization algorithm is described in Algorithm 1. The algorithmic complexity is directly related to the number of objects of interest in the image after low level processing, as well as the number of neighbors around each object. In the case of grouping of pixels, the optimization is applied to each pixel in the image, with a constant neighborhood. In the case of grouping of chains, the grouping process is reduced to the number of chains in the image, with a dynamic number of neighbors (depending on a fixed area around each extreme of the chains).

2.4. Selection of the best solutions

Once the optimization has been performed, the curves are extracted by following the connections from one primitive to another until the grouping 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 any other primitive from the segmented image).

The optimization reduces the number of possible groupings on the image to a single optimized grouping for each possible starting point. Primitives 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 grouping is the sum of local qualities for each of its points. This definition divides the possible solutions into classes of groupings with equivalent qualities. It is possible to reduce the number of solutions in each class by weighting their quality with an additional factor.

This factor is obtained by considering the quality of the primitives encountered along a grouping with respect to the quality of the sections introduced by the optimization process in order to fill gaps in this grouping. For example, in the case of grouping of pixels, this factor should be the number of pixels from the original image encountered by the the grouping with respect to the number of pixels added by filling gaps. A good grouping

Algorithm 1: Algorithm of optimization and selection.

```
begin
   Initialization of quality functions and connections at step 0 ( n = 0 )
   %
   % Iterations
   For each Primitive 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 Primitive P do
      Follow the connection from P
      Compute quality of the grouping
      Update GlobalQuality( P )
   endfor
   For each Primitive P ∉ Groupings already selected do
      if ( GlobalQuality( P ) > Threshold ) then Select Current Grouping
   endfor
end
```

should have a higher ratio of primitives coming from the image with respect to the primitives added by the optimization.

A simple thresholding of this weighted function gives good results to automatically select a set of optimized groupings as a first description of the scene.

3. Convergence

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, in the case of pixels, the length of the widest gap gives us the minimum of iterations required to 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.

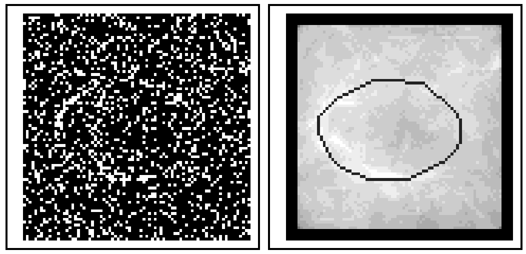

Figure 3: Ellipse 80x80 with %20 of noise - 25 iterations (26 sec / iteration)

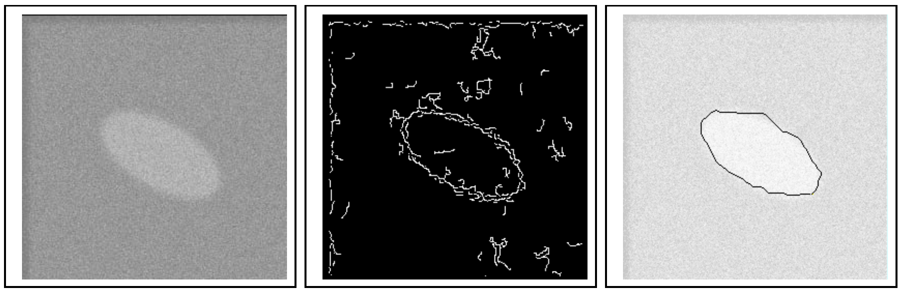

Figure 4: Ellipse 256x256 with Gaussian noise - Segmented Images and Grouping - 10 iterations (5 mins / iteration)

4. Results

This method has been developed and tested first for groupings of pixels. The good quality of results for groupings of pixels motivated the implementation of this method for the grouping of chains of pixels, mostly to reduce time and computer resources. Early results about groupings of chains are presented here. They use distance, curvature, and the length of chains as quality functions. With the same picture of 256x256 pixels, the computing time goes from one hour for the grouping of pixels, down to less than 10 minutes for the grouping of chains.

In the case of groupings of pixels, we applied the method to Synthetic images (fig 3) with different classes of noise. For both cases, the Signal to Noise ratio (SNR) has been computed as follows :

$$SNR_{dB} \;=\; -10\log\left(\frac{\sum_i\sum_j I(i,j)^2}{\sum_i\sum_j N(i,j)^2}\right)$$

Where $I(i,j)$ represents the image noise free, and $N(i,j)$ the altered image.

In the case of white noise, the image used represents an ellipse where %40 of the pixels have been removed by white noise. %20 of pixels of white noise have then been added to this image (the SNR is $11.4db$). The elliptic shape is still recovered even with high level of noise

In the second situation, the image represents a grey level ellipse where Gaussian noise has been introduced. Edges extraction in this case cannot be performed efficiently by low level edge detection but circular shapes are recovered after applying the perceptual grouping. The SNR is $7.6db$ for this ellipse.

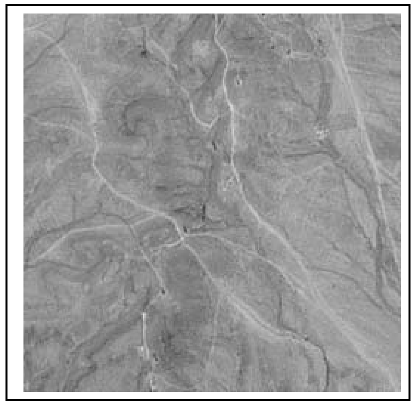

Figure 5: Satellite Picture 256x256 pixels - Original Image

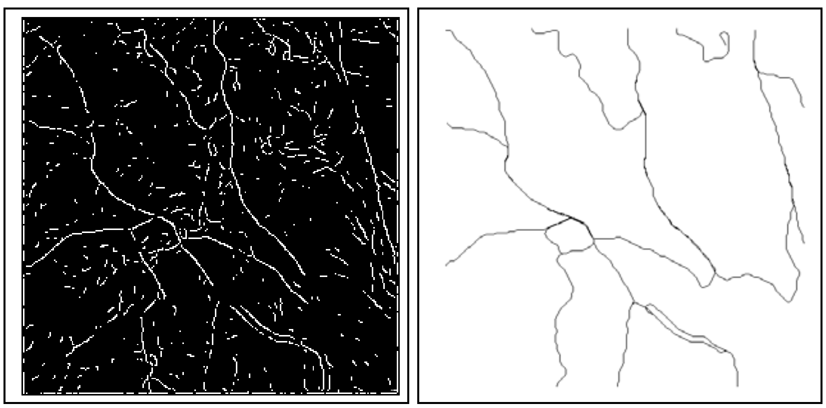

Figure 6: Satellite Picture 256x256 pixels - Segmented image and Final selection of 14 main groupings of pixels

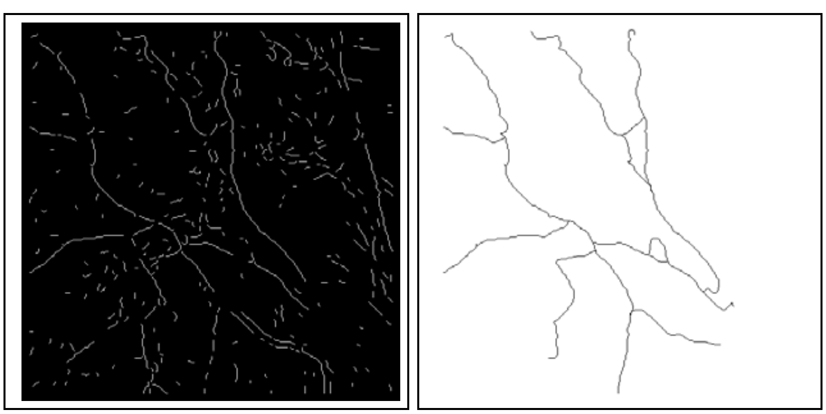

Figure 7: Satellite Picture 256x256 pixels - Segmented image and Final selection of 10 main groupings of chains

Grouping has been applied to crest lines detection on both satellite and medical images. The segmented images in fig 7 and fig 10 show the result of thin network extraction.[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.

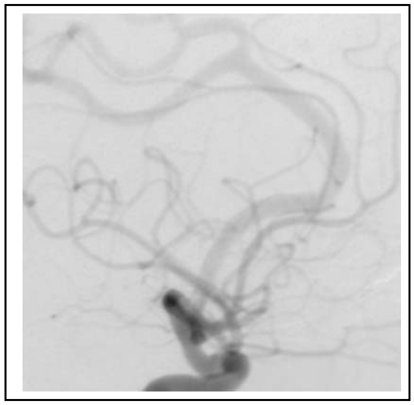

Figure 8: Medical Image 400x400 pixels - Original Image

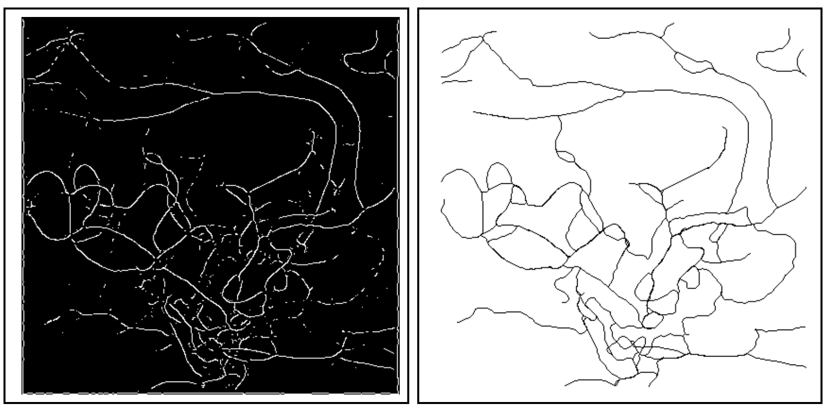

Figure 9: Medical Image 400x400 pixels - Segmented image and Selection of the 50 main groupings

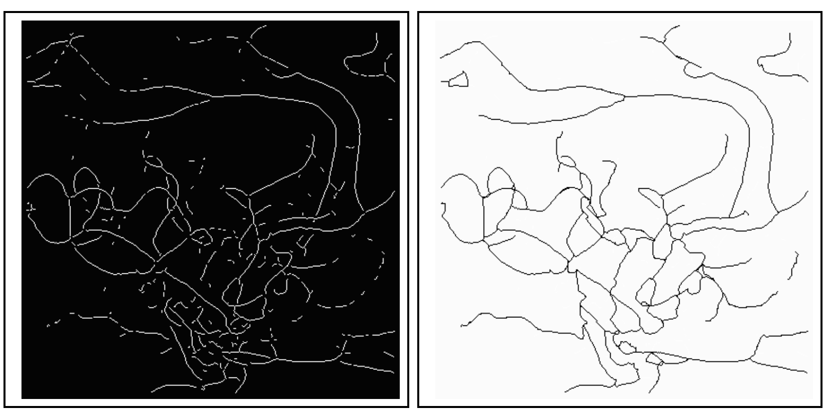

Figure 10: Medical Image 400x400 pixels - Segmented image and Selection of the 43 main groupings

More accurate 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). Without prior knowledge about the shapes being examined, the extraction of main groupings 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. We are currently improving the method in order to increase the quality of final groupings of chains (with new quality functions) and to provide a description of the scene using these groupings.

5. Conclusion

We have presented a new optimization technique for the Perceptual Organization of thin networks using dynamic programming and 'snake like' quality functions. This technique has been applied to the groupings of pixels as well as chains of pixels. We proposed also a method for the automatic selection of the best groupings.

The high quality of the results for the grouping of pixels shows how the method is robust to noise. Our method can be easily applied to real situations with large images such as roads or blood vessels detection. The good quality of results for groupings of pixels motivated the implementation of this method for the grouping of chains of pixels, mostly to reduce time and computer resources and make it more suitable for real applications. It can be used to initialize a high level interpretation process or to propose optimized solutions to a human expert (such as a physician with medical images). Possible applications range from closing edges to the initialization of active contours or the extraction of unknown shapes in very noisy images.

To further develop the method, we are focusing on the high level description of the scene into simple shapes in order to orient our work toward a 3D description from multiple views.

6. 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. Grouping contours elements using a locally connected network. In *Neural Information Processing Systems*, 1990.

[11] 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.

[12] 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.

[13] 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.