

Classification

Physics Abstracts

42.30.Tz — 42.30.Vx

## Segmentation and Real Time Control

Pierre Bonton <sup>(1)</sup>, Mustapha Derras <sup>(2)</sup>, Christophe Debain <sup>(2)</sup> and Michel Berducat <sup>(2)</sup>

<sup>(1)</sup> Université Blaise Pascal de Clermont-Ferrand, LASMEA, URA 1793 du CNRS  
(Clermont-Ferrand), 63177 Aubière Cedex, France

<sup>(2)</sup> Cemagref (Institut Français de Recherche pour l'Ingénierie de l'Agriculture  
et de l'Environnement), Domaine des Palaquins-Montoldre, 03150 Varennes sur Allier, France

**Résumé.** — Dans cet article, nous présentons une segmentation non supervisée originale en régions basée sur une modélisation markovienne d'un ensemble de sites. Chaque site (représentant une région élémentaire de  $16 \times 16$  pixels) est caractérisé par deux paramètres de texture, déduits de matrices de cooccurrences, et deux paramètres de luminance déduits d'histogrammes locaux normalisés. L'application réalisée montre la robustesse d'un algorithme de segmentation par l'analyse de la texture fauchée/non fauchée d'espaces naturels. Nous obtenons ainsi une limite optimum de séparation des deux zones qui sera la primitive de base permettant d'assurer l'asservissement visuel d'un robot de fauche. Des résultats sont fournis ; ils représentent des échantillons d'une large campagne de mesures réalisées sur le terrain sous diverses conditions météorologiques, pour divers couverts végétaux, capteurs d'images, *etc...* Une parallélisation simple de l'algorithme débouche sur une application temps réel (200 ms de temps de traitement par image).

**Abstract.** — An original region segmentation based on a Markovian modelling of a set of sites (representing a  $16 \times 16$  pixels elementary region) is presented. Each site is characterized by two texture parameters, deduced from co-occurrence matrices and two luminance parameters obtained from normalized local histograms. The application performed demonstrates the strength of a segmentation algorithm using texture analysis of mowed/unmowed natural zones. An optimum splitting limit has been obtained which is going to become the basic primitive in order to hook up a mower robot. In order to obtain a real time application (200 ms computing time per image), a simple parallelization of the algorithm and a control unit (control servoing) are realized. Lots of results which are layed out represent samples of a large land campaign measures under various meteorological conditions, with varied grass covers, image sensors, *etc...*

### 1. Introduction

The aim of this work consists in the design of a help guidance system for an auto-propelled mower. Therefore the different aspects of mobile robotics must be considered. One of these is to pick up and process information to drive a robot. In our case this information is an image that must be

split into regions to control the robot on a reape limit. In this context, the use of a texture analysis has been decided. The essential textural information is contained in small  $16 \times 16$  pixels dimension neighbourhoods: the *texel*. The descriptive elements concerning this type of texture (microtexture or random textures) can only be statistical [1].

A preliminary study [2] allowed, from a bibliographic research on reflectance physical properties of grass covers, the clearing of the most interesting zones of the electromagnetic spectrum for this application. So four parameters have been chosen. Among all methods of image texture analysis, the co-occurrence matrices take an important place [1-6]. This method assumes that all textural information is contained in a matrix which re-transcribes the spatial distribution of grey level variations between neighbouring pixels. These matrices reflect the relations existing between pixels of a region. In that way these regions form grey level primitives with specific characteristics which are accessible with texture features.

Haralick defines fourteen factors calculated from the co-occurrence matrix. Four of these parameters have been implemented to measure their efficiency. The other ten coefficients (three variance factors, three correlation factors, sum of the averages, sum of the entropies, entropy of the differences and energy) are unsuitable to our requirement [7, 8]. Moreover, we thought that the contrast and inertia coefficients (expressing grey level variations) can be replaced by parameters issued from local histograms to reduce the computing times. Therefore, only two of these characteristics have been used: the coefficient of *homogeneity* which gives us an idea of the surface state of the studied zone and the *entropy* which reflects a notion of order.

To use the reflectance properties of the natural surfaces, carried out during the preliminary study of this project, two parameters issued from  $16 \times 16$  pixels elementary region histograms have been defined. Firstly, the maximum histogram value (the mode value) conveys the predominant grey shade in the *texel*. Secondly, the second order moment accounts for this predominance amplitude. With these two coefficients, both a qualitative and a quantitative measure on the grey levels have therefore been realized. So these two parameters with the two previous ones, introduced in a relaxation process, are used to perform the segmentation of the image into regions.

This paper deals with the development of an image segmentation method in homogeneous regions for whom Zucker [9] and Monga [10] have defined criteria in order to obtain an optimal regions segmentation. However many data analysis methods [11-15], require, at different levels, an *a priori* choice which is incompatible with the application (element class number, comparison threshold). For example, it would be impossible to build up a representative data base of all kinds of grass (age, height, cover rate) and all meteorological conditions. So the approach moved towards unsupervised segmentation methods. One of the solutions to the problem was brought by the Markov's field modelling [16, 17]. Moreover this segmentation affords the advantage of being well adapted to the treatment of image sequences, because it offers the ability to take into account the result of the segmentation of the previous image. This is very important when the problem of motion is tackled, the final objective being a "real time" treatment. After a summary of the data and segmentation tool used, we develop our own contribution to the formulation of the problem (normalization of the *a priori* energy and calculation of the energy which determines the statistical relations between labels and data using the CHI-2 distance). A diagram, summarizing the main steps of the process, shows the different "initialization" and "image sequences" cases in the second section. The result of the segmentation of an image in similar appearance regions is then used in a last step consisting in finding the reape limit equation [2].

We have developed the effective robot control by the use of a visual servoing. To do that a virtual model of the robot and a cinematic model (as a caterpillars vehicle) were defined. Its originality comes from the trajectory calculation followed by the robot which allows us to drive this non holonome robot.

These algorithms are parallelized to permit a real time robot control. This parallelization lead to a pipeline architecture composed by three tasks:

- parameters computation,
- image segmentation (Markovian processus),
- global servoing (visual and speed control).

## 2. Image Segmentation

It is evident that increasing the number of descriptive elements, the choice of a different size of the sites or of the nature of the parameters have no influence on the following theory. This theory is supported by the example of the surface which corresponds to our application.

The goal is to obtain homogeneous textured regions of natural surface images (grass cover). Nevertheless two important constraints must be taking into account. Firstly a pre-processing of the images cannot be included because it may increase the computing times. Secondly any information about the type of texture used and the number of regions encountered in an image must be introduce because the algorithm must be auto-adaptative to face the external filming variations.

**2.1 SOLUTION WITH MARKOV'S FIELDS THEORY.** — The problem (labelling of distinct regions) lies on the fact that the image segmentation is only achieved from its different descriptive elements [18, 19]. Under a mathematics form, the problem is to maximize the *a posteriori* probability  $P(X/Y)$ , *i.e.* the probability of the searched object conditionally to the made measures [20], with:

- $X$  result of the segmentation (field of labels),
- $Y$  descriptive elements of the image.

By applying Bayes's theorem, this *a posteriori* probability can be expressed by:

$$P(X/Y) = \frac{P(X)P(Y/X)}{P(Y)} \quad (1)$$

with:

- $P(X)$  *a posteriori* probability of the field of labels,
- $P(Y/X)$ , the conditional probability of the measure in relation to one of the possible segmentation's. It describes entirely the statistical relations existing between data and labels,
- $P(Y)$ , the probability to obtain an observation. It may easily be dropped as it doesn't depend on  $X$  [21].

Therefore, the image comes in the form of a data field (four parameters for each elementary region  $16 \times 16$  pixels) that a label field must describe. This last field, in the end, will be the image segmentation. The use of the Markov's fields theory permits to limit the effects of each element of the lattice representing the image to a local interaction between neighbouring sites. These sites are linked by different "cliques" which determine the shape and size of the neighbourhood system we examine.

Then this first Bayes's equation must be expressed under another form more suitable to be implemented. The Hammersley-Clifford theorem enables the use of Gibbs's distributions which are expressed as followed:

$$P(X = \omega) = 1/Z \exp\{-U(\omega)\} \quad \omega \text{ belongs to } \Omega \quad (2)$$

with:

- $Z$ , the normalization constant,
- $U(\omega)$ , the energy function. This distribution describes the stability of our system,
- $\omega$ , the particular state of the system,
- $\Omega$ , the set of possible states.

Consequently, each of the terms of the equation (1) can be given under the form of a Gibbs's distribution. Thus:

$$P(X/Y) = P(X)P(Y/X). \quad (3)$$

By using the natural logarithms of  $P(X/Y)$  and  $P(X)$ , the expression can be reduced to:

$$U(X/Y) = U(X) + U(Y/X). \quad (4)$$

The purpose is to calculate the two terms of the function to minimize the energy  $U(X/Y)$  with:

- $U(X)$  constraints tied to the field of labels (*a priori* information),
- $U(Y/X)$  energy function in relation to statistical relationships between labels and data.

**2.2 SEGMENTATION DEVELOPMENT.** — The implements of the method of image segmentation by using the Markovian modelling, have been inspired by the deterministic relaxation algorithm I.C.M. (Iterated Conditional Mode) [8]. For each site, the different possible values of the energy function are calculated and only the state corresponding to the minimal energy is retained.

The two terms of the *a posteriori* energy function  $U(X/Y)$  will be given by:

- the *a priori* energy,
- the energy function related to the statistical relationships "data-labels".

The originality of this algorithm comes from the particular calculation of these two energies.

**2.2.1 A Priori Energy.** —  $U(X)$  is an *a priori* function. If the neighbourhood of a studied site is considered, an equation can be defined which takes into account the neighbouring sites label:

$$U(X) = \exp\{-Af(e_s)\}$$

In the following example  $A = 1$ , the potential of the sites are identical and equal to one.

$e_1$	$e_1$	$e_1$
$e_2$	**	$e_1$
$e_2$	$e_3$	$e_3$

here  $f(e_1) = 4$ ,  $f(e_2) = 2$  and  $f(e_3) = 2$ .

To compute the *a priori* energy value, a law of the form  $(1/x)$  has been firstly applied, performing probative segmentation's. The problem with such a law is the existence of an asymptote in 0 which makes the function tend towards infinity. Consequently, the form given by  $(1/x)$  is conserved and the asymptote is eliminated by a penalty function of the type  $\exp(-x)$ . The evolution of these two functions are given by Figure 1.

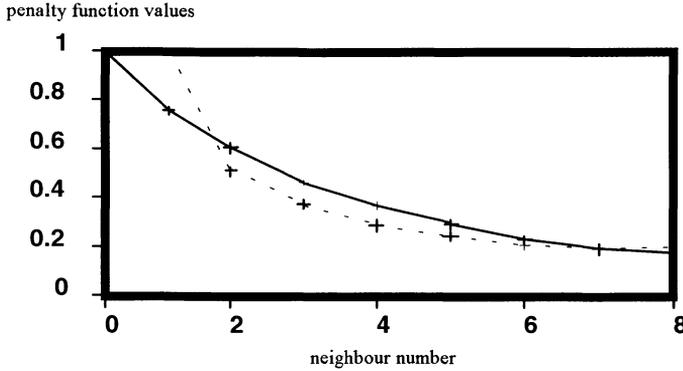


Fig. 1. — Penalty function evolution related to the number of neighbours: ( - - - ) function of the form  $(-1/x)$ ; ( — ) function of the form  $\exp(-x)$ .

**2.2.2 Energy Related to Statistical Relationships “Data-Labels”.** — The distance between each site and a general reference in the image (for example the region prototype) must be defined. The Euclidian one has been used but the well known difficulty of this distance is the effect of the variations of the amplitude of the considered elements on a statistic problem. That is the reason why the function  $U(Y/X)$  (relation between data and labels) is given by a CHI-2 distance [14] between sites and region vectors (described by four parameters).

**2.2.3 Process Description.** — The process breaks up into four steps summarized in Figure 2. As it can be seen, the process is formed of two distinct parts. The first one (full lines) describes the “initialization” part of the algorithm. Nevertheless, one of the interests of this segmentation in an image sequence, is inherent in the use of the result issued from the previous image. An *a priori* information is taken into account in the image segmentation. This constitutes the “dynamic” aspect (dashed lines), which consists to relate relation the extracted data of the new image with the label field of the previous image. Obviously, it has been considered that there are few changes between two images, in the motion meaning, and a noticeable gain of time is obtained for an equivalent efficiency segmentation.

The stage n° 1, which corresponds to the initialization of the process, starts without knowing anything about the shape or the number of the different kinds of texture that can be found in an image. Consequently, the image is supposed to be composed of a single region at the beginning. The field of labels is therefore initialized at zero and a first cut-out is realized from the comparison of the local energy of each site with the value of the average global energy calculated on the entire image. A general reference  $P_1$  made of elements which correspond to the averages of each descriptor on the whole of the image is used:

$$P_1(H_1, E_1, M_1, M_{01})$$

$$U_1(X/Y) = (1/N) \sum_{i=1}^{32} \sum_{j=1}^{32} \{U_{ijP_1}(Y/X) + U_{ij}(X)\} \tag{5}$$

with:

- $H_1, E_1, En_1, M_1, M_{01}$ , respectively the average values of homogeneity, entropy, local histogram mode and second order moment,

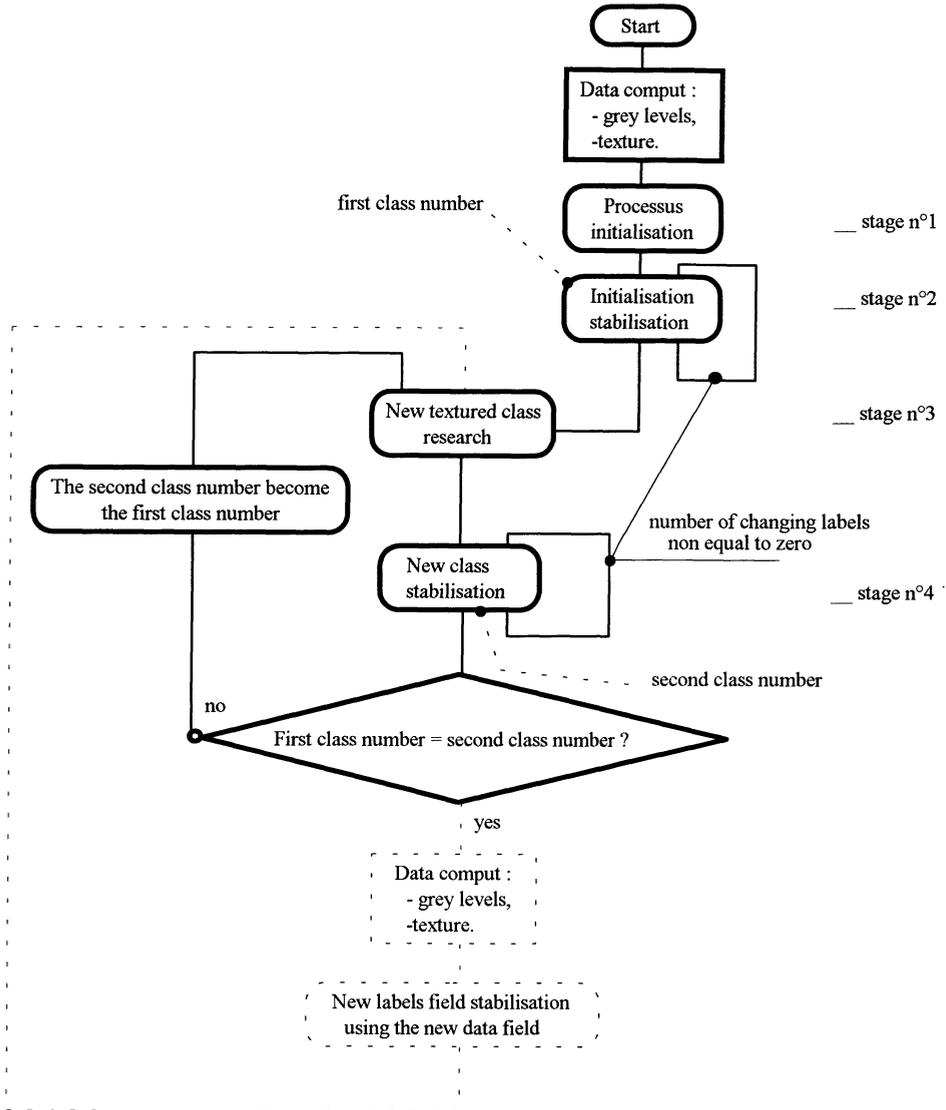


Fig. 2. — Description of the segmentation process.

- $P_1$ , the initial prototype,
- $U_{ijP_1}(Y/X)$ , the distance of the site  $ij$  to the prototype  $P_1$ ,
- $U_{ij}(X)$ , the penalty function of the site  $ij$ ,
- $N$ , the site number defined in an image (here 1 024),
- $U_1(X/Y)$ , the global average energy function value of an image.

From  $P_1$  and  $U_1(X/Y)$ :

- the local energy function is calculated for each site by (6):

$$U_{ij}(X/Y) = U_{ijP_1}(Y/X) + U_{ij}(X) \quad (6)$$

- $U_{ij}(X/Y)$  is then compared to  $U_1(X/Y)$ . If  $U_{ij}(X/Y) \leq U_1(X/Y)$  the site  $(ij)$  is assumed to belong to the region zero and remains unchanged. Otherwise the label of the site  $ij$  is set to one (remember that all sites are initialized to zero).

At the end of the initialization, one region is obtained in the case of an homogeneous image, and two regions in other cases. The stage n° 2 (Initialization stabilization) helps to suppress the very small regions (noise) and to redistribute the labels more regularly. The energy values in relation to the different labels are calculated and the state of each studied site is then modified according to the region giving the lowest energy. This step is repeated during several scans of the image, until there are no more change of labels.

The stages three and four (respectively New textured class research and New class stabilization) are generalizations of the two first stages to all detected regions. The purpose consists in detecting the presence of possible new regions inside those that have been defined by the stages one and two (discontinuities are searched in these regions). Conform to in the stage one, the average global energy values of each region are calculated as well as their prototype. If the segmentation provides  $k$  regions, a prototype  $P_k$  is obtained for each one as well as an average global energy value  $U_k(X/Y)$ . With these elements, one proceeds as follows:

- the local energy values corresponding to the labels detected in the neighbourhood of the studied site are calculated. Only the lowest ones are kept,
- then  $U_{ijk}(X/Y)$  and  $U_k(X/Y)$  are compared.
  - $U_{ijk}(X/Y)$  local minimum energy value of the site  $ij$ : it corresponds to the labelled region  $k$ ,
  - $U_k(X/Y)$  average global energy value of the labelled region  $k$ . If  $U_{ijk}(X/Y) > U_k(X/Y)$  then the label of the site  $ij$  is set to a new value in the label field, otherwise its label remains unchanged.

The process stops when the number of regions change any further.

**2.3 APPLICATION: REAP LIMIT SEARCH.** — This theory has been applied to the search of the reap limit in an image in order to ensure the working of an upkeep mobile machine (a mower).

Two examples of segmentation of grass surfaces are presented (Figs. 3 and 4).

### 3. Conclusion

The objective of this paper was to demonstrate the important potential offered by the Markovian modelling and above all its adaptability, because in the case of natural scenes the problem is the variety of images acquired in various conditions. The method proposed here in provides good results compared to the techniques using threshold of comparison. The other advantage is the use of an *a priori* information which allows the solution of a part of the problems of the real time analysis of image sequences.

In a first version, the algorithm runs on a SUN Sparc Station 1E, in a UNIX environment, with a sequential program written in C language. The average of the calculating time was near to 500 ms including read-out of data (texture and grey levels), segmentation process and result

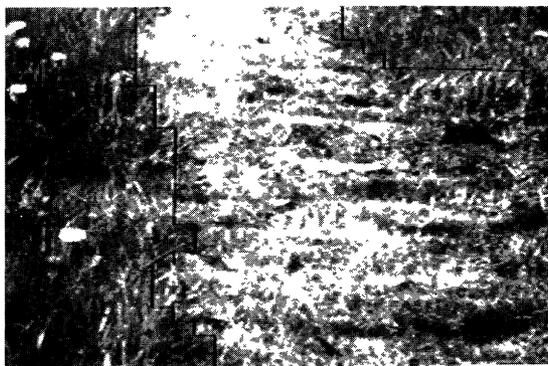


Fig. 3.

Fig. 3. — Result of segmentation.

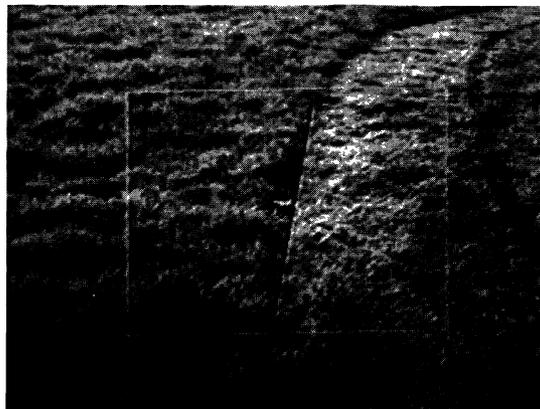


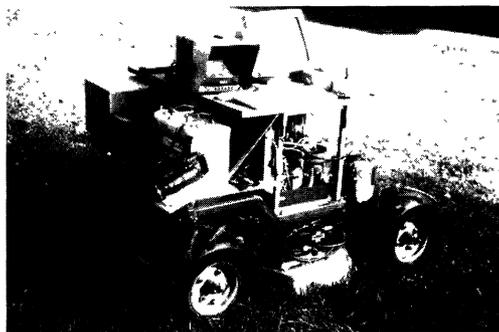
Fig. 4.

Fig. 4. — Analysis result after segmentation.

analysis. The implementation of the application on a new parallel architecture have allowed to reduce the computing time to 230 ms. It is now possible to consider another parallelization more “fine” of the data texture extraction because these calculations are repetitive. Nonetheless it is necessary to realize an architecture very different. The Markovian segmentation seems to be hard to parallelize but some actual works [6] permit to consider a solution. At last this approach can be extended to various kind of agricultural tasks (harvesting for example).



Before



After!

#### References

- [1] Haralick R.M., Statistical and structural approaches to texture, *Proc. IEEE* 67 (1979) 786-804.
- [2] Derras M., Berducat M., Bonton P., Gallice J. et Canals R., Segmentation texturale originale appliquée au guidage visuel d'un engin d'entretien des espaces naturels, Quatorzième colloque sur le traitement du signal et des images, GRETSI, Juan-les-Pins, 13-16 septembre 1993.

- [3] Chen P.C. and Pavlidis T., Segmentation by texture using a cooccurrence matrix and a split and merge algorithm, *Comp. Graphics and Image Process.* **10** (1979) 172-182.
- [4] Harlow C.A., Trivedi M. and Connors R.W., Use of texture operators in segmentation, *Opt. Eng.* **25** (1986) 1200-1206.
- [5] Lumia R., Haralick R.M., Zuniga O., Shapiro L., Ting Chuen Pong and Farpeing Wang, Texture analysis of aerial photographs, *Pattern Recogn.* **16** (1983) 39-46.
- [6] Yokoyama R. and Haralick R.M., Texture synthesis using a growth model, *Comput. Graphics and Image Process.* **8** (1978) 369-381.
- [7] Haralick R.M., Shanmugam K. and Dinstein I., Textural features for image classification, *IEEE Trans. Systems Man Cybernet.* **3** (1973) 610-621.
- [8] Trivedi M.M., Harlow C.A. and Connors R.W., Object detection based on gray level cooccurrence, *Comput. Vision Graphics and Image Process.* **28** (1984) 199-219.
- [9] Zucker S.W., Region growing: childhood and adolescence, *Comput. Graphics Image Process.* **5** (1976) 382-399.
- [10] Monga O., Segmentation d'images : où en sommes-nous ?, Rapport de recherche INRIA n° 1216 (avril 1990).
- [11] Ahalt S.C., Krishnamurthy A.K., Chen P. and Melton D.E., Competitive learning algorithm for vector quantization, *Neural Networks* **3** (1990) 277-290.
- [12] Derras M., Berducat M. and Bonton P., Vision guided mower for the upkeep of natural environment, Proceedings of the 1st International seminar of on-machine vision systems for the agricultural and bio-industries (3rd and 6th technical Sections of the CIGR), Montpellier, September 1991. bibitem13
- [13] Horowitz S.L. and Pavlidis T., Picture segmentation by a directed split and merge procedure, Proc. second international joint conference on Pattern Recognition (1974) pp. 424-433.
- [14] Harlow C.A., Trivedi M. and Connors R.W., Use of texture operators in segmentation, *Opt. Eng.* **25** (1986) 1200-1206.
- [15] Popovic M., Chantemargue F., Canals R. and Bonton P., Several approaches to implement the merging step of the split and merge region segmentation, EUROGRAPHICS, Vienne, September 1991.
- [16] Chen P.C. and Pavlidis T., Segmentation by texture using a cooccurrence matrix and a split and merge algorithm, *Comput. Graphics Image Process.* **10** (1979) 172-182.
- [17] Sullins J.R., Distributed learning of texture classification, Computer Vision, ECCV 90 (Antibes, 1990) pp. 349-358.
- [18] Derin H. and Elliot H., Modeling and segmentation of noisy and textured image using Gibbs random fields, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-9** (1987) 39-55.
- [19] Genam D., Genam S., Graffigne C. and Ping Dong, Boundary detection by constrained optimization, *IEEE Trans. Pattern Anal. Mach. Intell.* **12** (1990) 609-628.
- [20] Canals R. and Bonton P., Segmentation spatio-temporelle : régularisation par champs Markoviens, rapport interne au Laboratoire d'Électronique de l'Université Blaise-Pascal, Clermont II, avril 1991.
- [21] Berthod M., Giraudon G. and Stromboni J.P., Deterministic pseudo-annealing: optimization in Markov random fields. An application to pixel classification, Proceedings of the ECCV'92, second european conference on computer vision, Santa Margherita Ligure, Italy, G. Sandini, Ed. (May 1992) pp. 67-71.
- [22] Boutheymy P., Détermination du mouvement apparent sans une séquence d'images. Extraction de primitives locales, structuration intermédiaire, estimation du champ des vitesses, Rapport de Recherche INRIA n° 619, unité de recherche de Rennes (février 1987).