Unsupervised Regions Segmentation: Real Time Control of an Upkeep Machine of Natural Spaces

M. DERRAS, C. DEBAIN, M. BERDUCAT
CEMAGREF
CLERMONT-FERRAND CENTRE
Domaine des Palaquins - Montoldre
03150 Varennes sur Allier, France
Tél.: 70 45 03 12 Téléfax : 70 45 19 46

P. BONTON, J. GALLICE
LASMEA,
U.R.A. 830 of the C.N.R.S.
Les Cézeaux, 24 Avenue des Landais
63177 Aubière cedex, France
Tél.: 73 40 72 63 Téléfax : 73 40 72 62

ABSTRACT. An original image segmentation based on a Markovian modeling of a set of four parameters is presented. The application performed demonstrates the strength of an algorithm using texture analysis of natural scenes. A splitting limit is obtained which is going to become the basic primitive in order to hook up a mower robot. To obtain a real time application we also present a simple parallelization of the algorithm and a control servoing.

I INTRODUCTION

The aim of our project consists in the design of a help guidance system for an auto-propelled mower using texture analysis over 16x16 pixels dimension neighbourhoods: the texel [13]. A preliminary study [9] 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 two parameters have been chosen using the co-occurrence matrices [7], [12], [13], [21] which re-transcribes the spatial distribution of grey level variations between neighbouring pixels of a region. To use the reflectance properties of the natural surfaces, two parameters issued from 16x16 pixel region histograms have been defined. These four descriptors, introduced in a relaxation process, are used to perform the segmentation of the image into regions. The first section of this paper deals with the development of an image segmentation method in homogeneous regions [15], [22]. However many data analysis methods [1], [14], [17] require, at different levels, an a priori choice which is incompatible with the application (element class number, comparison threshold). So the approach moved towards unsupervised segmentation methods. One of the solutions to the problem was brought by the Markov's field modelling [2], [6], [11], [19]. Moreover this segmentation has 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. The second section explain the different steps of the process [10]. Then we develop the effective robot control by the use of a visual servoing [5], [16]. These algorithms are parallelized to permit a real time robot control. This parallelization leads to a pipeline architecture..

II - IMAGE SEGMENTATION

It is obvious that by 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 on the example of the surface which corresponds to our application. The goal is to obtain homogeneous textured regions of natural surface images (grass cover).

Lecture Notes in Computer Science, Vol. 801
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94

II.1 - Solution with Markov's fields theory

The problem (labelling of distinct regions) lies in the fact that the image segmentation is only achieved from its different descriptive elements [8], [11]. Under a mathematic form, the problem is to maximize the a posteriori probability *P(X/Y)*, i.e. the probability of the searched object conditionally to the measures made [4]. *X* represent the result of the segmentation (field of labels) and *Y* represent the descriptive elements of the image. By applying Bayes's theorem, this a posteriori probability can be expressed by :

$$P(X/Y) = \frac{P(X).P(Y/X)}{P(Y)} \quad (1) \qquad\qquad P(X=\omega) = \left(\tfrac{1}{Z}\right).\exp\{-U(\omega)\} \quad \omega \in \Omega \quad (2)$$

Here *P(X)* is the a posteriori probability of the field of labels. *P(Y/X)* is the conditional probability of the measure in relation to one of the possible segmentations. It describes entirely the statistical relations existing between data and labels and *P(Y)* is the probability to obtain an observation. It may easily be dropped as it doesn't depend on *X* [3]. Therefore, the image comes in the form of a data field that a label field must describe. 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. Then this first Bayes's equation must be expressed under another form more suitable for implementation. The Hammersley-Clifford theorem enables the use of Gibbs's distributions which are given by (2).

Here *Z* is a normalization constant and $U(\omega)$ an energy function. This distribution describes the stability of our system. ω is then a particular state of the system and Ω the set of possible states. Consequently, each terms of the equation (1) can be given under the form of a Gibbs's distribution. By using the natural logarithms of *P(X/Y)* and *P(X)*, the expression can be reduced to $U(X/Y) = U(X) + U(Y/X)$ (3). The purpose is then to calculate *U(X)* and *U(Y/X)* to minimize the energy function *U(X/Y)*.

II.2 - Segmentation development

The implementation 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) [11]. 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 and the energy function related to the statistical relationships "data-labels".

II.2.1 - A priori energy

U(X) is an a priori function. If the neighbourhood of a studied site is considered, the equation (4) can be defined which takes into account the neighbouring sites label where V_s is the neighbourhood of the studied site, *A* a weighting factor, $V_{c.}$ the potential assigned to the site e_s.

$$U(X) = \exp\left\{A.\left(\sum_{s \in V_s} V_c(e_s)\right)\right\} \quad (4) \qquad\qquad d^2(m,m') = \sum_{i=1}^{p}\left(\frac{K}{K_{d.}}\right).\left(\frac{K_{im}}{K_s} - \frac{K_{im'}}{K_{.s'}}\right)^2 \quad (5)$$

II.2.2 - Energy related to statistical relationships "data-labels"

A distance between each site and a general reference in the image (for example the region prototype) must be defined. The Euclidian one have 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's the reason why the function

$U(Y/X)$ (relation between data and labels) is given by the CHI-2 distance (5) between sites and region vectors (described by four parameters).

II.2.3 - Process description

The process breaks up into four steps. 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 take into account in the image segmentation. This constitutes the "dynamic" aspect, which consists in putting in relation the extracted data of the new image with the label field of the previous image. Obviously, it have been considered that there are few changes between two succiding images, and a noticeable gain of time is obtained for an equivalent efficiency segmentation.

The initialization step 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 $P_I\ (H_I, E_I, M_I, Mo_I)$ constituted by elements which correspond to the averages of each descriptor on the whole of the image is used to calculate :

$$U_I(X/Y) = \left(\frac{1}{N}\right) \cdot \sum_{i=1}^{32} \sum_{j=1}^{32} \left\{ U_{ijPI}(Y/X) + U_{ij}(X) \right\} \qquad (6)$$

Where H_I, E_I, M_I and Mo_I are the average values of homogeneity, entropy, local histogram mode and second order moment. P_I is the initial prototype. $U_{ijPI}(Y/X)$ represent the distance of the site *(ij)* to the prototype P_I. $U_{ij}(X)$ is the penalty function of the site *(ij)*. N is the site number defined in an image (here 1024) and $U_I(X/Y)$ is the global average energy function value of an image. For each site the local energy function is calculated by $U_{ij}(X/Y) = U_{ijPI}(Y/X) + U_{ij}(X)$ (7). $U_{ij}(X/Y)$ is then compared to $U_I(X/_Y)$. If $U_{ij}(X/Y) \leq U_I(X/Y)$ 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 Initialization stabilization helps to suppress the very small regions 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 is no more change of labels. Step three and four which can be named "new textured class research" and "new class stabilization" are generalizations of the two first stages to all detected regions. The aim consists in detecting the presence of possible new regions inside those that have been defined by stages one and two (discontinuities are searched in these regions). As in step 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, we proceed as follows. The local energy values corresponding to the labels detected in the neighbourhood of the studied site are calculated. Only the lowest ones is kept. Then $U_{ijk}(X/Y)$ (local minimum energy value of the site ij, corresponding to the labelled region k) and $U_k(X/Y)$ (average global energy value of the labelled region k) are compared. 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 does not change any further.

III - 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). We present one example of result of segmentation of grass surfaces.

Picture n°1 : A turn of cut grass between two unmowed areas is detected.

In a dynamic use (image sequences), the Markov fields theory is entirely justified. The idea consists in using again, as an a priori information, the field of labels issued from the image segmentation at the time T to process the next image at the time $T + \Delta T$. The main purpose is to limit in one hand the computing time with a reactualization of the label field and, in the other hand, the errors of detection, because the changes between two successive images are supposed to be very slight. In the approach presented here, only the first image of the sequence is entirely processed (initialization step). Then, in order to improve the computing time, the segmentation is localised on an area of interest (half of the image) sufficient to ensure the future guidance of the upkeep machine. We must insist on the fact that luminance in natural scenes, vibrations of the engine, kind of grass, etc..., can not be controlled.

IV - MACHINE CONTROL [18, 20, 24]

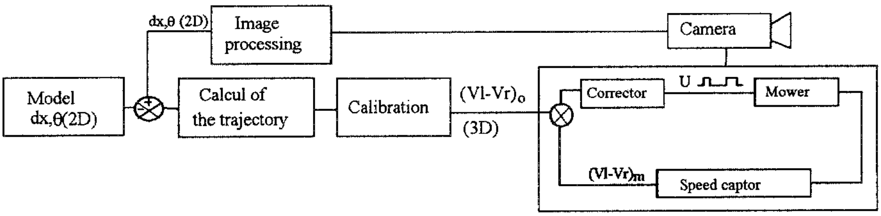

Figure n°5 : Complete system.

The totality of the automation embedded on the robot is presented on the figure n°5. From these visual information and the computing time, the algorithm of control determines the trajectory. It calculates the order, that compared to real speeds, allows a numerical corrector to control the speed of the wheel. As the machine is moving to an average speed of 3 kilometres an hour, it makes slightly more than 80 centimeters in one second. So to obtain a periodicity of its control which should be compatible with its dynamics and to limit risks of interface losses in the image, we have developed a multi-processor architecture adapted to the chosen algorithms of image processing and control. The solution consists in sharing the algorithm in several tasks executed in parallel on three cards organized according to the model of the pipeline.

V- ARCHITECTURE AND PARALLELIZATION

In spite of the technological evolution of the microprocessors and especially of their speed, it is necessary to realize a parallel approach of the algorithm on an adapted architecture to obtain a real time application. If its not a new parallelism approach [14] we can remark that the multiprocessor calculators and components (such as the transputers) have appeared too recently. Nonetheless few of these machines are available on the market. M.J. FLYNN [23] has proposed a classification of the parallel calculators in four groups based upon the nature of the instructions and the data stream. However, lots of other criteria exist which allow finer differentiation of the parallelization of the algorithms. These criteria depend often on the kind of memory use (shared or distributed) or the "grain" size (parallelization at the level of the instructions or at higher levels of groups of instructions).

In our case the knowledge of an a priori information (the previous image segmentation) and the algorithm parallelization permit us to reach our real time goal (230 ms average computing time per image). We have implemented an architecture composed by three VME microprocessors boards. The problem caused by such a pipeline construction consists in the repartition and the organization of all the concurrent tasks witch constitute the parallelized algorithm. Three tasks have been determined and attributed to each processor. To manage all the problems arising from by the shared resources between the different processors we have chosen the operating system OS9. Among all the different mechanisms of synchronization and communication between tasks offered by this environment we have particularily used the signals, the communication pipes and the data modules.

IV - CONCLUSION

The objective of this article was to demonstrate the important potential offered by the markovian modelling and above all its adaptability to face the various conditions of external scene. The method proposed here provides good results compared to the techniques using thresholds 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. The implementation of the application on a parallel architecture has permit to reduce the computing time to 230 ms. At last this approach can be extended to various kinds of agricultural tasks (harvesting for example). Although non formalised, the visual servoing approach of control seems to be similar the principle developed for the CRV of the manipulator robots. This theory is presently studied in our team. This paper describes an engine which actually runs at the CEMAGREF of Clermont-Ferrand. The complete architechture is embedded on the machine. The first tests carried out give satisfaction. A finer and deeper perfecting will increase the strength of the system.

REFERENCES

[1] - **S.C. Ahalt, A.K. Krishnamurthy, P. Chen, D.E. Melton**, "Competitive learning algorithm for vector quantization", Neural Networks, vol. 3, 1990, pp 277-290.

[2] - **M. Bernard, M. Sigelle**, "Champs de Gibbs et Champs markoviens", Télécom Paris, Département Images, Groupe Image, mai 1990.

[3] - **M. Berthod, G. Giraudon and J.P. Stromboni**, "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, (ed.) G. Sandini, may 1992, pp 67-71.

[4] - **R. Canals, P. Bonton**, "Segmentation spatio-temporelle : régularisation par champs markoviens", rapport interne au Laboratoire d'Electronique de l'Université Blaise-Pascal, Clermont II, avril 1991, 31 p.

[5] - **F. Chaumette**, "La commande référencée vision : une approche aux problèmes d'asservissements visuels en robotique", Thèse de doctorat de l'Université de Rennes, juillet 1990, 171 pages.

[6] - **Chaur Chin Chen**, **R.C. Dubes**, "Experiments in fitting discrete Markov random fields to texture", CVPR (Computer Vision and Pattern Recognition) 89, San-Diego, june 1989, pp 248-303.

[7] - **P.C. Chen, T. Pavlidis**, "Segmentation by texture using a cooccurence matrix and a split and merge algorithm", Computer Graphics and Image Processing, vol. 10, 1979, pp 172-182.

[8] - **H. Derin**, **H. Elliot**, "Modeling and segmentation of noisy and textured image using Gibbs random fields", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-9, n°1, january 1987, pp 39-55.

[9] - **M. Derras**, **M. Berducat** et **P. Bonton**, "Vision guided mower for the upkeep of natural environment", Poceedings 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, 10 p.

[10] - **M. Derras, M. Berducat, P. Bonton, J. Gallice** et **R. Canals,** "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, 4 pages.

[11] - **D. Genam, S. Genam, C. Graffigne, Ping Dong**, "Boundary detection by constrained optimization", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 12, n°7, july 1990, pp 609-628.

[12] - **R.M. Haralick**, "Statistical and structural approches to texture", Proceeding of the IEEE, vol. 67, n°5, may 1979, pp 786-804.

[13] - **C.A. Harlow, M. Trivedi, R.W. Conners**, "Use of texture operators in segmentation", Optical Engineering, vol. 25, n°11, november 1986, pp 1200-1206.

[14] - **S.L. Horowitz and T. Pavlidis**, "Picture segmentation by a directed split and merge procedure", Proc. second international joint conference on Pattern Recognition, 1974, pp 424-433.

[15] - **O. Monga** "Segmentation d'images : où en sommes-nous ?", Rapport de recherche INRIA n°1216, avril 1990, 63 pages.

[16] - **R. Pissard-Gibollet et P. Rives**, "Asservissement visuel appliqué a un robot mobile : état de l'art et modélisation cinématique", Rapport de Recherche INRIA n°1577, unité de recherche de Sophia Antipolis, décembre 1991, 42 pages.

[17] - **M. Popovic, F. Chantemargue, R. Canals, P. Bonton**, "Several approaches to implement the merging step of the split and merge region segmentation", EUROGRAPHICS, Vienne, september 1991.

[18] - **C. Samson and K. Ait-Abderrahim**, "Mobile robot control part 1 : feedback control of nonholonomic wheeled cart in cartesian space", Rapport de Recherche INRIA n°1288, unité de recherche de Sophia Antipolis, octobre 1990, 50 pages.

[19] - **J.R. Sullins**, "Distributed learning of texture classification", Computer Vision, ECCV 90, Antibes 1990, pp 349-358.

[20] - **P. Tournassoud**, "Planification et contrôle en robotique : application aux robots mobiles et manipulateurs", Hermes, Paris 1992, 248 pages.

[21] - **R. Yokoyama, R.M. Haralick**, "Texture synthesis using a growth model", Computer Graphics and Image Processing, vol. 8, n°3, 1978, pp 369-381.

[22] - **S.W. Zucker**, "Region growing : childhood and adolescence", Computer graphics and Image processing, n°5, 1976, pp 382-399.

[23] - **M.J. Flynn**, "Some computer organization and their effectiveness", IEEE Transaction on Computer, vol. C21, n°9, september 1972, pp 948-960.

[24] - **N. Kehtarnavaz, N. C. Grisword, J. S. Lee**, "Visual control of an autonomous vehicle (BART) - The vehicle-following problem", IEEE Transaction on Vehicle Technologie, vol. 40, n°3, August 1991.