

UNSUPERVISED IMAGE SEGMENTATION CONTROLLED BY MORPHOLOGICAL CONTRAST EXTRACTION

Ferran MARQUÉS, Jordi CUNILLERA, Antoni GASULL

Dept. Teoría de la Señal y Comunicaciones

E.T.S.E.T.B. - U.P.C. Apdo. 30002

08071 Barcelona, SPAIN

e-mail: ferran@tsc.upc.es

ABSTRACT

A novel approach for unsupervised image segmentation is described in this paper. This approach makes use of a Gaussian pyramid as multiresolution decomposition to analyse images. Compound random fields are used to model images at each resolution. The hierarchical image model is formed by a Strauss process in the lower level and a set of white Gaussian random fields in the upper level. This basic image model is adapted to the data present at each resolution. Segmentations at coarse resolutions are used to guide segmentations at finest resolutions. Segmentation quality is controlled, at each level, by means of morphological tools. The control procedure is based on the residue between the original image and a morphological centre transform. This procedure checks whether the current segmentation contains all the relevant regions in the scene. If not, the algorithm introduces seeds into the segmented image in order to detect the new regions.

I.- INTRODUCTION

Compound Random Fields (CRFs) [1,2] are nowadays widely used in image processing. Within the framework of image segmentation, several methods have been proposed relying on such a kind of image models. These methods are either very time demanding [3], if they use stochastic maximisation techniques, or very depending on the initial segmentation [4], if they use deterministic maximisation approaches. In order to decrease the dependency on the initial segmentation of deterministic approaches, multiple resolution analyses have been introduced in the segmentation procedure [5,6]. In these techniques, coarse resolution segmentations guide the method to better results when segmenting finest resolution images. However, multiresolution techniques still arise some problems.

In this paper, an unsupervised multiresolution segmentation technique based on compound randomfields and controlled by mathematical morphology tools is presented. The structure of the

This work has been partially supported by the project TIC-92-0800-C05-03 of the Spanish Government

paper is as follows. After this brief introduction, in Section II, the basic multiresolution segmentation scheme presented in [6] and its main drawbacks are presented. Section III deals with the unsupervised adaptation of the model parameters to the data in the image, in order to improve the performance of the segmentation technique. Section IV presents a morphological stage introduced to detect interior regions. Once an interior region has been detected, the previous technique decides whether this new region deserves to be included into the final segmentation. Finally, Section V is devoted to the presentation of some results and conclusions.

II.- BASIC MULTIREOLUTION SCHEME

In this section, the multiresolution segmentation technique presented in [6] is outlined for completeness purposes. This technique makes use of a Gaussian pyramid as multiresolution decomposition. Moreover, it relies on CRFs to model images at each resolution. The CRF upper level X (texture model) is formed by a set of independent white Gaussian random fields. On its turn, the CRF lower level Q (contour model) is assumed to be a Strauss process [7] defined on a second order neighbourhood. Parameters characterising the texture model are estimated from the image data. The contour model parameters are fixed for the whole segmentation procedure. These fixed values are obtained from an analysis of their influence on the final segmentations results [8].

The segmentation procedure starts by segmenting the coarsest resolution in the Gaussian pyramid. Once the image at a resolution (k) is segmented, this result is used as initial segmentation for the next finer resolution (k-1). This procedure is iterated down to the finest resolution (the image itself) is segmented. In this way, coarse resolution segmentations lead the procedure to good quality final segmentations. Multiple resolution decompositions allow to achieve good quality results when using deterministic maximisation techniques at each resolution. That is, the dependency of deterministic approaches on the initial segmentations is almost completely removed by multiresolution analyses.

At each resolution level, the monoresolution segmentation technique proposed in [4] is used. That is, the segmentation at each level is performed by seeking the partition $Q(k)$ of the image $X(k)$ which maximises the probability $P(Q(k)/X(k))$. This result can be achieved by maximising $P(X(k), Q(k))$:

$$P(X(k), Q(k)) = Z \exp \left(-\frac{1}{T} (n_1(k)V_1 + n_2(k)V_2) \right) \cdot \prod_n P(x_n(k) / m_n(k), \sigma_n^2(k)), \quad (1)$$

where Z is a normalising constant, T stands for the temperature, $n_1(k)$ and $n_2(k)$ are the number of cliques in the partition of the level (k) with potential V_1 and V_2 respectively [4], and $m_n(k)$ and $\sigma_n(k)$ are the parameters of the Gaussian random field characterising the region n . The maximisation procedure is carried out by a progressive boundary refinement of the initial segmentation.

This technique yields good quality segmentation results [6]. However, final segmentations may overlook details in the original image. The reason for this lack of detail is mainly twofold. First, the use of the same parameters to characterise images through the whole decomposition does not completely exploit the possibilities of the multiresolution analysis. In addition, given that the maximisation relies on a boundary refinement procedure, totally interior regions (that is, regions whose contours do not touch other contour regions) which are not present in the initial segmentation cannot be detected by this segmentation technique.

III.- IMAGE MODEL ADAPTATION

Image data at different decomposition levels do not share the same features, due to the filtering procedure involved in the computation of the Gaussian pyramid. Thus, different model parameters should be used in order to characterise the data at different decomposition levels. Since upper level parameters are already constantly updated in the basic multiresolution scheme, only lower level model parameters have to be adapted. The three parameters used in (1) for modelling the lower level can be reduced to only two:

$$\frac{1}{T} [n_1(k)V_1 + n_2(k)V_2] = \frac{1}{T^*} [n_1(k) + n_2(k)V^*] \quad (2)$$

Actually, only the T^* parameter adaptation results in an improvement of the procedure performance [8].

This parameter controls the relationship between the lower level model and the upper level model in the CRF. In order to adapt its value to the data contained at

each resolution, the information of the Laplacian pyramid is introduced into the procedure. Each level of the Laplacian pyramid (k) contains the difference between two consecutive levels of the Gaussian pyramid. Assuming that level $(k+1)$ has been correctly segmented, the information contained in the level (k) of the Laplacian pyramid gives an estimation of how far this segmentation is from that of level (k) . That is, if most of the pixels at level (k) of the Laplacian pyramid have zero value, segmentation at level $(k+1)$ of the Gaussian pyramid is very close to that at level (k) . On the other hand, to have several pixels with high values in the Laplacian level translates into the necessity of performing several corrections in the segmentation of level $(k+1)$ so that the segmentation of level (k) is obtained. That is, new regions have to be detected and contours have to be refined in order to conform to those of level (k) . The way to carry out such tasks is by giving priority to the texture information with respect to the contour information in the CRF [8]. The value of the T^* parameter at each resolution is therefore given by the following expression:

$$T^*(k) = \frac{\sum \text{abs} [L_{ij}(k)]}{\sum \text{abs} [L_{ij}(0)]}, \quad (3)$$

where $L_{ij}(k)$ is the gray value of the pixel (i, j) at level (k) of the Laplacian pyramid. This approach leads to good results, as illustrated in Figure 1. In this figure, the original Cameraman image (Figure 1.a), its Gaussian pyramid as well as their segmentations are shown. The final segmentation (264 regions) is presented as a contour image (Figure 1.c) and as a mosaic image (Figure 1.d). In mosaic images, each region is filled with its mean gray value. It has to be highlighted that these segmentations present good quality and improve those obtained by the basic multiresolution technique [6]. In addition, the $T^*(k)$ values yielded by (3) are very close to these found manually to be optimal. Actually, even the different behaviour found for textured and non-textured images is respected by this estimation procedure [8].

However, it has to be noticed that, in spite of the above commented improvement, some details are still overlooked in the final segmentation (e. g.: the sticks in the tripod, the aeriels or some parts of the camera and the face). This drawback is owing to the presence of interior regions in the image, which cannot be extracted by means of the previous algorithms. To solve this problem, a stage controlling that, at each level of the decomposition, the final segmentation does not overlook interior regions has been introduced in the algorithm.

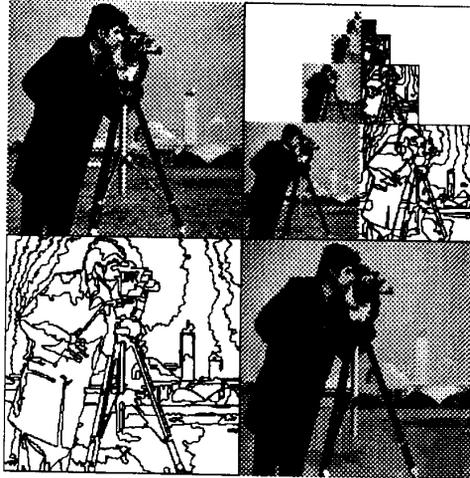


Figure 1.- Segmentation of the Cameraman image

IV.- SEED EXTRACTION BY MORPHOLOGICAL TOOLS

To analyse the presence of non-detected interior regions within a segmentation, the error image is built. This image contains the difference between the original image and the mosaic representation of its segmentation. Assuming that segmentations have been correctly performed, error image information is mainly related to textures in the original image and non-detected interior regions. Therefore, the main task is to discriminate between texture and non-detected interior region information. Interior regions appear in the error image as nearly flat, contrasted areas; whereas textured information appears as dense, contrasted fluctuations. Since textured areas may yield pixels with larger absolute values than those produced by non-detected interior regions, classical thresholding techniques cannot be used. Thus morphological tools [9] have been applied, given that they are known to efficiently perform in discriminating between these kinds of information.

Morphological filters do not solve the above problem. Hence, the necessity of using other kind of transforms, such as morphological residues. From the set of different residues that can be defined, the residue between the original signal and a morphological centre [10] (contrast extraction transform) turns out to be very efficient for detecting interior regions. The centre is computed from the open_close, the close_open and the identity operators:

$$\text{Contrast extractor: } |I - \text{centre}(\gamma\phi, \phi\gamma, I)| \quad (4)$$

where γ and ϕ stand for the open and close filters respectively, and the centre is defined as

$$\text{centre}(\gamma\phi, \phi\gamma, I) = \text{Min}(\gamma\phi, \text{Max}(\phi\gamma, I)) \quad (5)$$

An example using a monodimensional signal is shown in Figure 2 for illustrating the performance of the morphological contrast extractor. This signal contains two contrasted, almost flat areas which are related to interior regions as well as three contrasted zones with rapid fluctuations, related to textured areas. Note that the contrast extractor detects both interior regions, yielding only a single connected component for each one of them (concept of seed). On the other hand, textured information is almost completely removed. The small remaining components can be withdrawn by performing a cleaning step relying on size and gray level value criteria [8].

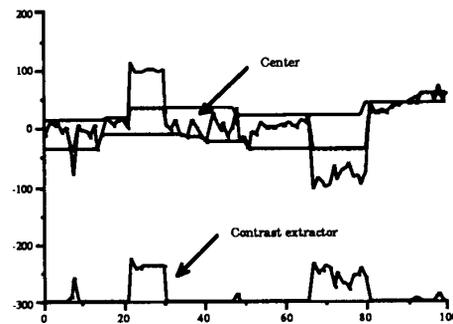


Figure 2.- Monodimensional example

After the cleaning step, the remaining components are used as seeds for detecting new regions in the segmentation procedure. That is, seeds are added to the previous segmentation and this new partition is used as initial segmentation in the procedure described in Section III. In this way, seeds do not directly conform new regions but the monoresolution segmentation technique decides whether they deserve to form a new region or not. If a seed is considered to be related to an overlooked interior region, the segmentation method finds its actual shape.

V.- RESULTS AND CONCLUSIONS

The use of morphological tools allows removing rapid variations of the signal while providing with seeds for interior regions. Therefore, in textured zones, new regions do not appear and their oversegmentation is avoided. On the other hand, interior regions are clearly detected. This effect can be seen in Figure 3, where the segmentation of the Cameraman image is performed using the above technique. Figure 3.a shows,

in its left-hand side, the segmentation at each level of the pyramid, while, in its right-hand side, the set of seeds obtained at each level. On its turn, Figure 3.b shows the seeds obtained at the bottom of the pyramid. Finally, Figures 3.c and 3.d show the contours of the final segmentation (353 regions) and its mosaic representation, respectively.

The high quality of the final segmentation can be observed. It deals correctly with large textured areas without withdrawing any relevant small region. It is worth noticing that the morphological step detects relevant interior regions without introducing new ones in textured areas (e. g.: the grass). This effect is even more noticeable in the example of Figure 4. In it, the image presents a large textured area which has been gathered into a few regions. Furthermore, no main detail in the image has been overlooked and, thus, the representation of the hand and the racket is of very high quality. The final segmentation of this image contains 75 regions.

It has to be highlighted that, in spite of the quality of the results achieved by this technique, it is not time demanding. Actually, the average computational load required for segmenting 256x256 pixel images with a Sun Sparc II workstation is of 27 seconds.

The current work is mainly twofold. First, it copes with the extension of the whole segmentation technique to the case of segmenting image sequences. In parallel, we are applying this segmentation method to the problem of image coding. That is, a region-based image coding scheme is being developed.

REFERENCES

[1] F.-C. Jeng, J. W. Woods "Compound Gauss-Markov Random Fields for Image Estimation", *IEEE Trans. on Signal Proc.*, vol. 39, pp. 683-697, 1991.
 [2] S. Geman, D. Geman, "Stochastic Relaxation, Gibbs Distribution and the Bayesian restoration", *IEEE Trans. PAMI*, vol. 6, pp. 721-741, Nov. 1984.
 [3] S. Lakshmanan, H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing," *IEEE Trans. PAMI*, vol. 11, pp. 799-813, 1989.
 [4] F. Marqués, A. Gasull, T. Reed, M. Kunt, "Coding-oriented segmentation based on Gibbs-Markov Random Fields and human visual system knowledge", *Proc. ICASSP-91*, pp. 2749-2752, Toronto, May 1991.
 [5] C. Bouman, B. Liu, "Multiple resolution segmentation of textured images," *IEEE Trans. PAMI*, vol. 13, pp. 99-113, 1991.
 [6] F. Marqués, J. Cunillera, A. Gasull, "Hierarchical segmentation using Compound Gauss-Markov Random Fields", *Proc. ICASSP 92*, pp. III. 53-56, San Francisco USA, March 1992.

[7] H. Derin, P. Kelly, "Discrete-index Markov-type random processes", *Proc. IEEE*, vol. 77, pp. 1485-1510, 1989.
 [8] F. Marqués, *Multiresolution image segmentation based on compound random fields. Application to image coding*, PhD Thesis, Univ. Pol. Cat., Barcelona, 1992.
 [9] J. Serra, *Image Analysis and Mathematical Morphology*, vol. I, Academic Press, London, 1982.
 [10] P. Salembier, J. C. Serra, "Morphological multiscale image segmentation", *Visual Comm. and Image Proc. '92*, pp. 620-631, Boston, 1992

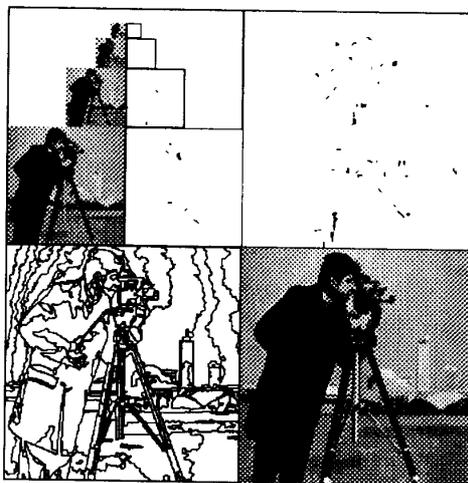


Figure 3.- Final segmentation of the Cameraman image

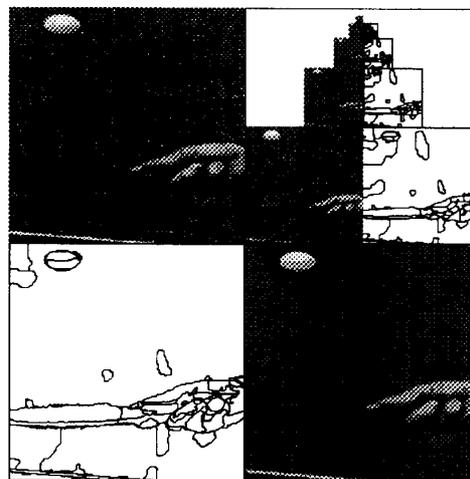


Figure 4.- Segmentation of a textured image