CODING-ORIENTED SEGMENTATION BASED ON GIBBS-MARKOV RANDOM FIELDS AND HUMAN VISUAL SYSTEM KNOWLEDGE.

Ferran MARQUÉS   Antoni GASULL
Dept. Teoría de la Señal y Comunicaciones
E.T.S.E.T.B. - U.P.C. Apdo. 30002
08080 Barcelona, SPAIN

Todd R. REED   Murat KUNT
Laboratoire de Traitement des Signaux
E.P.F.L. - Ecublens (DE)
CH - 1015 Lausanne, Switzerland

ABSTRACT

In this paper a new segmentation algorithm for still black and white images is introduced. This algorithm forms the basis of a region-oriented sequence coding technique, currently under development. The algorithm models the human mechanism of selecting regions both by their interior characteristics and their boundaries. This is carried out in two different stages: with a preprocessing that takes into account only gray level information, and with a stochastic model for segmented images that uses both region interior and boundary information. In the stochastic model, the gray level information within the regions is modelled by stationary gaussian processes, and the boundary information by a Gibbs-Markov Random Field (GMRF). The segmentation is carried out by finding the most likely realization of the joint process (Maximum a Posteriori criterion), given the preprocessed image. For decreasing the computational load while avoiding local maxima in the probability function, suboptimal versions of the algorithm are proposed.

1.- Introduction

Currently, the image coding approaches used most frequently are based on transform techniques, and mainly on the Discrete Cosine Transform. Usually, these techniques lie on block-oriented schemes that avoid the transmission of boundary information and have low computational complexity. On the other hand, two main problems arise when using a fixed block partition of the images: the information within each block is not always stationary which results in a loss of coding performance, and each block is treated separately which results in an obviously false contour effect (blocking-effect) when decoding the image. Moreover, the compression ratios achieved by these techniques seem to have reached their maximum.

Two main ideas can be applied for increasing the compression ratio in image sequence coding: the reduction of inter- and intra-frame redundancy, and the use of knowledge of the human visual system (Second Generation techniques) [1]. A way to combine these two ideas and to avoid the main drawbacks of a block-oriented scheme is to use object-oriented sequence coding techniques. These techniques make use of segmentation to provide the coding algorithm for regions, in which a given criterion of stationarity should be fulfilled. Furthermore, in this case the independent treatment of each region does not lead to a false contour effect but highlights the boundaries of the segmented image. Thus, in order to obtain a high visual quality of the decoded image, the segmentation performed should be as close as possible to the visually perceived segmentation.

In this paper, we introduce a new segmentation algorithm for still images which forms the basis of a sequence coding technique. The segmentation takes into account the human mechanism of selecting regions both by their interior and by their boundary characteristics. The first aspect is considered in a preprocessing that will be described below. Moreover, both characteristics are used jointly in a stochastic model for segmented images. The gray level information within the regions is modelled by stationary gaussian processes and the boundary information by a GMRF [2]. The segmentation is carried out by using the Maximum a Posteriori (MAP) criterion. Hence, the algorithm seeks the most likely realization of the process, given the preprocessed image. For the sake of computational simplicity, suboptimal versions of the algorithm are proposed.

The organization of the paper is as follows: the second section deals with the preprocessing method. In the third section, the stochastic model for segmented images is defined. A MAP criterion to perform the segmentation, using this model, is presented. The fourth section describes two suboptimal algorithms that overcome the problems of the optimum algorithm. Results obtained using these algorithms are shown. The last section is devoted to comments and conclusions.

2.- Preprocessing stage

When comparing classical with visual segmentations, it can be observed that classical methods have a tendency to split visually homogeneous light areas into several regions, as well as to merge visually inhomogeneous dark areas, producing overgrown regions. These effects lead to segmentations in which the regions seem to correspond poorly to objects in the scene.

In order to avoid these problems, the first stage of our segmentation algorithm preprocesses the image to account for the greater sensitivity of the human visual system (HVS) to gradients in dark areas as compared to light ones. This has been implemented following Stevens' law, recently used to build a new visual model of brightness perception [3]. The algorithm changes the gray level of each pixel within the image following the relation

$$y = K * x^n,\qquad(1)$$

where n is an exponent found to be about 0.33 [3]. This algorithm allows us to obtain regions that conform more closely to those detected visually. Therefore, the coding performance may improve, even if the number of regions is higher than the number obtained when using classical methods.

3.- Segmented image model

Throughout this paper, all images are defined on an $N_1 * N_2$ rectangular lattice

$$I = \{ (i, j): 1 \le i \le N_1, 1 \le j \le N_2 \}.$$

The given data (preprocessed image) $y = \{ y_{i,j} \}$ is a realization from a stochastic process $Y = \{ Y_{i,j} \}$. The segmented image is assumed to be a set of regions $R = \{ R_n \}$. The partition of the segmented image $q = \{ q_{i,j} \}$ is modelled as a sample from a homogeneous Gibbs Random Field (GRF), $Q = \{ Q_{i,j} \}$. This partition models the shape of the contours within the segmented image. The set of gray level values within a region R_n, $y_n = \{ y_{i,j}: (i, j) \in R_n \}$, is assumed to be a sample from a stationary gaussian process.

There is a one-to-one relation between GRF and Markov Random Fields [5]. This relation allows the use of local characteristics to cope with a process which has been modelled by means of global characteristics. A second order GMRF is used to model the boundaries of the regions. That is, for each point in the lattice (i, j), the neighbourhood system η_{ij}, on which the process is defined, consists of its eight nearest neighbours. In order to define a Gibbs distribution on this region of support, the concept of clique has to be introduced. A clique, denoted by c, is a set of points. This set may be
- a single point, or
- if $(i, j) \ne (k, l)$, if $(i, j) \in c$ and if $(k, l) \in c$ then $(i, j) \in \eta_{kl}$.
In figure 1, a second order neighbourhood η and the set of all its possible cliques is shown.
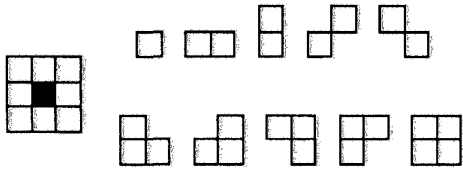

Figure 1.-   A second order neighbourhood and its ten associated cliques.

Each of these cliques has an associated potential V_c which must be defined. In our case, these potentials do not depend on the position of the clique on the lattice, but only on the type of clique (the GMRF is said to be homogeneous). Since the GMRF will be used to model the shape of the boundaries in partition Q, the clique potential values are defined such that:
- a clique wholly contained in a region has potential zero, in particular one-point cliques. For the rest of cliques:
- the potential of three-points and four-points cliques can be set to zero, thanks to the fact that they can be determined by the potential of two-point cliques;
- the process is supposed to be isotropic. Thus, vertical and horizontal cliques must have the same potential (V_1), as must both diagonal cliques (with potential V_2).

The general expression of a Gibbs distribution

$$P(Q) = \frac{1}{Z} \exp \left(- \frac{1}{T} U(Q) \right) \quad \text{with} \quad U(Q) = \sum_{c \in Q} V_c , \quad (2)$$

where Z and T are constants, under the above constraints

becomes

$$P(Q) = \frac{1}{Z} \exp \left(- \frac{1}{T} (n_1 V_1 + n_2 V_2) \right), \quad (3)$$

where n_1 and n_2 are the number of cliques in the given partition with potential V_1 and V_2, respectively.

Segmentation algorithm

The objective now is to find, by a MAP criterion, the most likely realization of the process Q. That is, to maximize the conditional probability $P(Q/Y)$ with respect to q. Using the Bayes' rule:

$$P(Q/Y) = \frac{P(Q) \, P(Y/Q)}{P(Y)} , \quad (4)$$

where

$$P(Y/Q) = \prod_n P(y_n / m_n, \sigma_n^2). \quad (5)$$

The parameters that define each gaussian process (m_n, σ_n^2) are estimated by means of a maximum likelihood criterion:

$$\hat{m}_n = \frac{1}{N_n} \sum_{R_n} y_{i,j} \qquad \hat{\sigma}_n^2 = \frac{1}{N_n} \sum_{R_n} y_{i,j}^2 - \hat{m}_n^2. \quad (6)$$

where N_n is the number of pixels of the region R_n. This model has been already used [4] in a contour relaxation algorithm. Although the assumption of gaussian processes is a coarse approximation, it allows computationally simple algorithms while yielding good results. As the maximization is not affected by the given image y, the joint distribution $P(Y,Q)$ can be maximized, instead of $P(Q/Y)$:

$$P(Y,Q) \propto \exp \left(- \frac{1}{T} (n_1 V_1 + n_2 V_2) \right) \prod_n P(y_n / \hat{m}_n, \hat{\sigma}_n^2). \quad (7)$$

The algorithm is initialized by taking each point in the image as a single region. The segmentation is carried out by computing the ratio of the current joint distribution and the joint distribution when changing the label (flag of the region in which a point is included) of a boundary point (point belonging to, at least, one clique with potential different from zero). This ratio is calculated for all the boundary points. The change of label that increases the most the likelihood of the process is performed. The algorithm is iterated until its convergence. The convergence is guaranteed, given that only label changes that increase the joint probability are allowed. Hence, the iteration stops when none of the current possible label changes increases the likelihood of the joint distribution. At this point a maximum in the probability function has been reached.

In order to avoid local maxima that do not yield a correct segmentation, the parameter T of the joint distribution can be modified, starting from $T \to \infty$, each time that a maximum is achieved, until a given value of T is reached. The modification of this parameter can be viewed as a simulated annealing problem [2]. Even though parallel implementations of the simulated annealing algorithm exist [6], the solution of this problem still demands a huge computational effort.

4.- Suboptimal algorithm

One way to reduce the computational complexity of the implementation is to provide the algorithm with not only the preprocessed image, but also with an estimate of the desired segmentation. This first estimate should be an oversegmented image, since the algorithm of section 3 is not allow to create new regions. Moreover, in order to be consistent with the model of segmented images presented in section 3, the initial estimation should be performed taking into account the homogeneity in mean and variance of the regions. Thus, this first segmentation could be obtained as the first local maximum achieved in the simulated annealing, for T → ∞ in (7). Nevertheless, an algorithm based on MAP criterion using mean and variance parameters performs only a few merging operations, and the eventual segmentation is still far from the desired result. Therefore, another criterion should be used.

Obtention of the first estimation.

The algorithm proposed to obtain the initial estimate is based on a region growing technique. This kind of method has been chosen because it can be easily extended to image sequence segmentation. The algorithm uses as input the preprocessed image and has three different steps. At each step the merging criterion becomes more complex. In the first step, two pixels are merged if the difference between their gray level is lower than a fixed threshold.

The criterion in the second step uses the concepts of size and mean value of the regions. That is, a region smaller than a given size is merged accounting for the similarity of its mean value with the mean values of the regions of its neighbourhood. The mergings of regions with closest mean values are performed first, and mean values updated after each merging. The reason for removing small regions is twofold. First, as the model of section 3 uses the estimated value of the variance of the regions, it is worth having a fair number of points in each region, in order to avoid inaccurate estimations. Second, this segmentation is the basis of a sequence coding technique. Thus, small details in the scene can be blurred by overgrowing regions, to reduce the amount of information to code. However, the loss of small details at this step of the algorithm is not an irreversible process, and relevant small regions are recovered when introducing the boundary information in the region model.

In the third step, the gray level values within each region are modelled as in section 3. That is, they are assumed to be a sample from a stationary gaussian process. The merging criterion at this step checks all the possible mergings of two neighbour regions, by means of a test of similarity of their gray level distributions. This is carried out with the Generalized Likelihood Ratio (GLR) test, due to its greater performance over other tests. As in the previous step, the merging of the most similar regions is performed first, and the GLR tests are updated after each merging. This algorithm finishes when a segmentation with a given number of regions is obtained. At this point, a first estimation of the desired segmentation has been reached. Hence, the algorithm of section 3 can be used to find the maximum closest to this estimation.

As the process of simulated annealing has been shortcut with the use of this first estimation, the parameter T in (7) is redundant and can be set to one. In order to fix the potentials V_1 and V_2, two concepts can be applied: the boundaries of the regions should conform to the boundary of natural objects, and the resulting contours should be as easy as possible to code. Both concepts lead to a smooth contour solution, which can be achieved penalizing diagonal cliques within the partition. Paying

attention to this conclusion, several pairs of values have been studied and it has been concluded that a large range of values yield similar results (values of V_1 between 0.1 and 10 have been used, as well as ratios of V_2 and V_1 between 0.1 and 0.5).

Considering that the algorithm presented in the previous section has to be extended to a sequence segmentation, a simplification in terms of data structure and computational load is in order. Thus, two different algorithms are proposed. These algorithms do not seek the label change that increases the likelihood in the whole image at each step, but for either a restricted area or a given step.

Recursive (sequential) algorithm

The image is scanned line by line, from top to bottom and from left to right. For each boundary point, the label that maximizes the likelihood of the realization of the process is assigned to it. After each label change, the parameters of the model are updated. To speed up the algorithm, a second scan in the opposite direction is carried out. This double scan is iterated until no label change is performed.

Non-recursive (parallel) algorithm

For a given partition, all the label changes that initially increase its likelihood are made without updating the parameters of the model. Once this step is finished, the entire set of parameters is updated. Since the convergence of this method is not guaranteed (label changes that do not maximize the current partition may occur), it should be iterated only a fixed number of times. However, in all the cases in which this algorithms has been tested without fixing the number of iterations, at each iteration the probability function has increased, and eventually a maximum has been reached.

It should be remarked that the results obtained with both algorithms are similar. In figure 2, the final segmentation of frame number 6 from the sequence "Miss America" is shown. Note that hair and background are successfully separated thanks to the preprocessing stage. Figure 3 shows this segmentation with each region filled with its mean value. Figure 4 shows a coarse segmentation of the image "Cameraman". This segmentation has not been obtained with the method proposed above, but it is useful to highlight that even using such a segmentation as initial approximation, the proposed algorithms are able to recover the details of the original image, and to perform a correct segmentation (see figures 5 and 6). The number of iterations necessary in this case is greater than when using the first segmentation proposed above.

5.- Conclusions

A new segmentation algorithm for still images, that forms the basis of a region-oriented sequence coding technique (currently under development), has been presented. The algorithm introduces knowledge of the HVS in order to obtain more natural segmentations. The segmentation is performed using a MAP criterion applied to a stochastic process that models segmented images, using both boundary and gray level information jointly. Since the interior of each region is modelled by a gaussian process, the spatial correlation of the points in the lattice is not fully exploited. In order to make a better use of the spatial correlation information, other kinds of models may be applied (e.g.: a GMRF for the interior of the region). The current research deals with the application of this kind of model, as well as the extension of the algorithm to image sequence segmentation.

References

[1] Kunt, Ikonomopoulos, Kocher, "Second-Generation image coding techniques", Proc. IEEE vol. 73 num. 4 April 1985.

[2] Geman, Geman, "Stochastic Relaxation, Gibbs Distribution and the Bayesian restoration", IEEE Trans. PAMI vol 5 1984.

[3] Xie, Stockham, "Toward the unification of three visual laws and two visual models of brightness perception", IEEE Trans. Sys. Man and Cyb. vol. 19 num. 2 March 1989.

[4] Mester, Franke, "Statistical model based image segmentation using region growing, contour relaxation and classification" SPIE Conference on Visual Comm. and Image Proc. '88.

[5] Derin and Kelly, "Discrete-Index Markov-Type Random Processes", Proceedings IEEE, vol. 77, 1989 pp. 1485-1510.

[6] Murray, Kashko, Uxton, "A parallel approach to the picture restoration algorithm of Geman and Geman on an SIMD machine", Im. & Vis. and Comp., Vol. 4, 1986 pp.133-142.

Figure 4.- Coarse segmentation of "Cameraman".

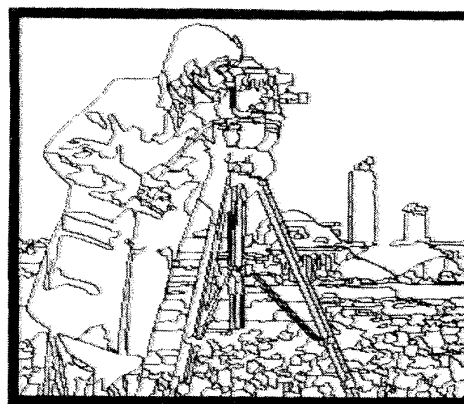Figure 2.- Segmentation obtained by means of the recursive algorithm (256 regions).

Figure 5.- Segmentation obtained by means of the non-recursive algorithm (350 regions).

Figure 3.- Segmentation of figure 2 with each region filled with its mean value.

Figure 6.- Segmentation of figure 5 with each region filled with its mean value.