A NON-ITERATIVE APPROACH TO INITIAL REGION ESTIMATION APPLIED TO COLOR IMAGE SEGMENTATION

Javier Portillo-García, Carlos Alberola-López†,
Lorenzo José Tardón-García, Juan Ignacio Trueba-Santander
SSR-ETSI Telecomunicación-UPM, Ciudad Universitaria s/n. 28040 Madrid, Spain
†TSCIT-ETSI Telecomunicación. Univ. Valladolid, C/Real de Burgos s/n. 47011 Valladolid, Spain
Tel: +34 1 5495700 ext. 206; fax: +34 1 3367350
e-mail: javierp@gtts.ssr.upm.es

ABSTRACT

A non-iterative segmentation approach is developed to generate a fast initial estimation of the layout of the different color textures presented in the original image mainly based on hypothesis testing. Most of the proposed methods have a tremendous computational burden which make them difficult to be implemented in a real-time working processor. We will compare our method with a known iterative clustering algorithm that guides to similar results with much higher computational cost. We present two examples that show similar results and compare the computational cost for each case. Spotty resemblance caused by pixel oriented decision is diminished in both cases by modeling regions as Markov Random Fields.

1 INTRODUCTION

In this paper we address the problem of texture segmentation by statistical decision and classification. The general problem of texture segmentation has received much attention in the last decades, and there is not a final general solution for the time being. There is a number of proposals which cover a wide spectrum of the techniques typically applied in signal and image processing: Wigner distributions, Gibbs distributions or Markov Random Fields, multiresolution and multiscale techniques, and more recently a particular case of the former by the use of Wavelet transform.

All of these proposals have proved successful in their operation, but most of them have a tremendous computational burden which makes them difficult to be implemented in a real-time working processor. Furthermore, an important subclass of them, known as model-based, needs a recursive model parameter estimation which makes the cost increase dramatically as the number of textures in the image grows.

In this contribution we concentrate in segmentation by clustering. The underlying idea in this technique is to assign a certain pixel or group of pixels to a specific class, from a known set of classes. But, as it is obvious, if no knowledge is available about the number of classes and their main features, this assignment cannot be directly done. A solution to this problem might be devised by training a processor with textures often encountered. But this solution is highly application dependent and cannot be done in an unsupervised fashion. A more elaborated solution [4] leads to an iterative (in the number of centroids) clustering algorithm (such as *k-means* or *ISODATA* [6]). A measure of dispersion within and between clusters will track the iterations until a minimum in a certain functional is achieved. This procedure is general, but quite expensive in computation. We try to avoid the need of training and to alleviate the huge need of CPU time required by the clustering just described, so we propose to bypass the computationally expensive phase of iteratively estimating the number of regions the image consists of by applying a method of statistical decision about the homogeneity of the tessels the image is divided into.

2 NON-ITERATIVE REGION ESTIMATION

The color image is divided into three independent bands. The statistical tests explained below are applied to each band, and decisions are fused by the majority criterion.

First, each band is tesselated by rectangular, non-overlapped windows. Every tessel is subdivided into four non-overlapping subwindows. For each subwindow we estimate some statistical parameters in order to carry out statistical tests to decide about the homogeneity of the window. We estimate the sample mean, the sample variance and the sample mean and covariance of (2×1) vectors selected in the window. To assure incorrelation, we estimate the two-dimensional autocovariance function of the pixels in the window $\hat{r}_{yy}$. We obtain the minimum row and column displacements, n_r and n_c that assure that the autocovariance values decrease to a 10% of its value in the origin. The samples are selected n_r rows and n_c columns apart.

In order to determine the homogeneity of a given window we perform three types of statistical hypothesis tests:

- Test for equality of means. We accept or reject the hypothesis $\mu_1 = \mu_2 = \mu_3 = \mu_4$, where μ_i is the mean of the statistical distribution of the pixels in the subwindow $i, i = 1,2,3,4$.

- Test for equality of variances. We accept or reject the hypothesis $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$, where σ_i^2 is the variance of the statistical distribution of the pixels in the subwindow $i, i = 1,2,3,4$.

- Test for equality of mean vectors and covariance matrices. We accept or reject the composite hypotesis $\vec{\mu}_1 = \vec{\mu}_2 = \vec{\mu}_3 = \vec{\mu}_4, \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}_3 = \mathbf{\Sigma}_4$ where $\vec{\mu}_i$ is the (2×1) mean and $\mathbf{\Sigma}_i$ is the (2×2) covariance matrix of the statistical distribution of the (2×1) vectors selected in each subwindow.

We propose two different procedures to decide about the homogeneity of a window. In the first procedure, for each band and window, we perform six hypotesis tests for the equality of means:

$$H_0 : \mu_i = \mu_j, i,j = 1,2,3,4; i \neq j \tag{1}$$

$$H_1 : \mu_i \neq \mu_j, i,j = 1,2,3,4; i \neq j \tag{2}$$

and six hypothesis tests for the equality of variances:

$$H_0 : \sigma_i^2 = \sigma_j^2, i,j = 1,2,3,4; i \neq j \tag{3}$$

$$H_1 : \sigma_i^2 \neq \sigma_j^2, i,j = 1,2,3,4; i \neq j \tag{4}$$

The thresholds of the critical regions can be obtained, for a given level of significance α, using the normal distribution of the sample mean differences, and the Snedecor's F distribution of the sample variance ratio (assuming in this last case gaussian statistic for the pixel values) [2]. In the case of the test about variances, we have the additional problem of approximating the Snedecor's F function, since we do not know a priori the number of uncorrelated points that we are going to select in each subregion. We approximate this function using the incomplete Beta function [3]. We declare, for each band, a region to be homogeneous if in five hypotesis tests for the means and in five hypothesis tests for the variances, the hypothesis H_0 is not rejected. This criterion makes difficult a non-homogeneous zone to be declared as homogeneous, since this type of error affects dramatically to the assignment of features to the clusters, and therefore to the post-processing phase. Obviously, this increases the probability that a homogeneous window is declared as heterogeneous. However, this type of error can be solved in the postprocessing phase.

The second procedure can be seen as a generalization of the preceding one. We only have an hypothesis test on the (2×1) vector of sample means and (2×2) covariance matrices:

$$H_0 : \vec{\mu}_1 = \vec{\mu}_2 = \vec{\mu}_3 = \vec{\mu}_4, \mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}_3 = \mathbf{\Sigma}_4 \tag{5}$$

$$H_1 : \vec{\mu}_i \neq \vec{\mu}_j, \text{ for some } i,j; i \neq j, \text{ or}$$
$$\mathbf{\Sigma}_i \neq \mathbf{\Sigma}_j, \text{ for some } i,j; i \neq j \tag{6}$$

In this case, assuming a bivariate normal distribution for the sample (2×1) vectors, it is possible to find an asymptotic distribution for the generalized likelihood criterion ratio, λ, for the composite hypothesis H_0, as a linear combination of Chi-square distributions [1], which can in turn be approximated using the incomplete Gamma function [3].

The tessels considered as homogeneous are used to properly estimate the number of regions avoiding the use of iterative clustering algorithms. Care is taken about the problem of space-separated windows corresponding to the same texture pattern, which may appear, for instance, when we segmentate aerial photographs of a road in a field or a river in a forest. The problem is solved applying the same hypothesis tests to the disjoint homogeneous regions estimated in the first phase of the method.

A non-iterative pixel assignment to the previously determined clusters is carried out over those tessels considered as non-homogeneous in order to label all pixels in the image. This step is based on the determination of local statistics both in homogeneous and heterogeneous windows, specifically, the statistics used in this stage are means $(\hat{\mu}_{i,j})$, variances $(\hat{\sigma}_{i,j})$ and covariances $(\hat{C}_{i,j}(0,1),\ \hat{C}_{i,j}(1,0))$. Features estimated in homogeneous regions are used to determine the position of the clusters (centroids) denoted $\vec{\eta}(k)$, of the K regions that compose the image so that unclassified pixels can be assigned to the region with the nearest cluster.

In this stage, a non-euclidean distance measure is used to determine the nearest cluster to each pixel; provided that all selected features are non-null, we weight each of them so that their gravity center is always 1 in order to eliminate the efect of the different relative magnitudes between the features used in this phase. This makes the distance between a region labeled k and a pixel at position $(i,\ j)$ be defined as:

$$d_k^{(i,j)} = \sqrt{\sum_{f \in features} \frac{1}{gc_f} (\vec{\eta}_f(k) - \vec{y}_f^{(i,j)})^2} \tag{7}$$

where gc_f is the gravity center of feature f and $\vec{y}_f^{(i,j)}$ and $\vec{\eta}_f(k)$ the component f of the feature vectors for a pixel at position $(i,\ j)$ and for region k respectively.

3 SMOOTHING STAGE

A final smoothing phase is driven by the result of the previous stage after accomplishing the pixelwise classification in heterogeneous windows, so that the spotty resemblance caused by this pixel-oriented assignment is diminished. This smoothing is based on the assumption that the unobservable field of regions can be modeled as a simple Markov Random Field with a first order neighbourhood such that a single parameter conditional probability density function is used [7]. We give the same weight to each of the regions provided that no further knowledge about their shape, extension or relative importance is available.

Let $z_{i,j}$ denote the state of a pixel at position $(i,\ j)$, and $z_{\eta_{i,j}}$ the first order neighbourhood of pixel $(i,\ j)$, then the probability density function (pdf) used that follows the general Gibbs distribution [7], is

$$f(z_{i,j} = k | z_{\eta_{i,j}}) = \frac{e^{\beta \delta_{i,j}(k)}}{\sum_{l=0}^{K-1} e^{\beta \delta_{i,j}(l)}} \tag{8}$$

where $\delta_{i,j}(l)$ is the number of neighbours of pixel $(i,\ j)$ in state l. Finally, β is the parameter that contains information of the degree of spatial clustering observable in the state process. Note that we will be interested in values of $\beta > 0$ so that spots will have a high probability of being cleared out if they happen to be inside a large different region.

Indeed we are using an annealing procedure [5] so that in each iteration the process is guided by a stronger clustering pdf. The starting parameter β is estimated applying Newton Raphson method [3] to the initial segmentation.

4 COMPARATIVE ANALYSIS

To establish the computational burden savings of this method we present a comparison of performance between the presented method and an iterative one based on *k-means* and the iterative estimation of the number of regions through the maximization of a specified parameter.

4.1 Iterative Initial Clustering

As said above, the iterative segmentation algorithm is based on *k-means* which is an iterative unsupervised clustering algorithm that generates our initial region estimation provided that the number of regions and the features to clusterize are given [6], by minimizing the global distance (Euclidean distance in this case) of the feature vector of each pixel $(i,\ j)$ to its corresponding region.

But it is neccesary to obtain the number of regions the image consists of. So we start with a minimum fixed number of textures, obtain a pre-segmentation using *k-means* and then get and estimate of the dispersion of the features of pixels with respect to their corresponding regions and the dispersion of the centroids with respect to the image's feature vector. [4].

So we make use of the within-cluster scatter matrix defined as:

$$S_w = \frac{1}{K} \sum_{k=0}^{K-1} \frac{1}{N(k)} \sum_{y^{(i,j)} \in \omega(k)} (\vec{y}_v^{(i,j)} - \vec{\eta}(k))(\vec{y}_v^{(i,j)} - \vec{\eta}(k))^t \tag{9}$$

and the between-cluster scatter matrix:

$$S_b = \frac{1}{K} \sum_{k=0}^{K-1} (\vec{\eta}(k) - \vec{\eta}_0)(\vec{\eta}(k) - \vec{\eta}_0)^t \tag{10}$$

where K is the number of regions, $N(k)$, the number of pixels that belong to region k according to the present segmentation, $\omega(k)$ represents the set of pixels in the image that belong to region k, $\vec{\eta}(k)$ is the feature vector for region k as it was defined previously, and finally, $\vec{\eta}_0$ represents the feature vector for the whole image. So we try to find a maximum for

$$\rho = tr\{S_b\} \cdot tr\{S_w\} \tag{11}$$

where $tr\{\cdot\}$ stands for the trace matrix operator.

The parameter ρ is expected to have a good behavior so that its value increases from 0 (when the number of regions equal 1 then $tr\{S_b\} = 0$) to the maximum and then decreases again to 0.

Notice that the number of regions obtained this way must be optimum from this mathematical point of view which may not correspond to a good visual discrimination of textures.

4.2 Computational Cost Comparison

Now we present an example showing results obtained with the non-iterative proposed method and the iterative one and a second example which reflects the intermediate stage after homogeneous region detection and the final result. Figure 1 a) shows an arrangement of two textures taken from Brodazt's album [8] and figure 1 b) the results of the segmentation procedure using iterative initial clustering (*k-means* [6]) with previous estimation of the number of regions. Figure 1 c) presents the final result of applying the proposed improved algorithm.

Figure 1 d) shows a composition of two images of grinded coffee under two different illuminations, figure 1 e) shows the homogeneous tessels (black and white) found and figure 1 f) presents the final segmented image.

Table 1 shows the computational burden savings of the proposed method versus the iterative version in execution time in the segmentation of these 256x256 images. In the first case two regions where detected by both algorithms, but execution time is much smaller for our non-iterative version, even in the best case for iterative method (only two regions are detected). In the second case (segmentation results using iterative clustering are not presented here) iterative segmentation decides the presence of three different regions making the execution time dramatically increase both with the number of regions and the number of features used in the segmentation when treating the three bands of the color image.

Note that we are showing just process execution time required by the segmentation routine, for comparison

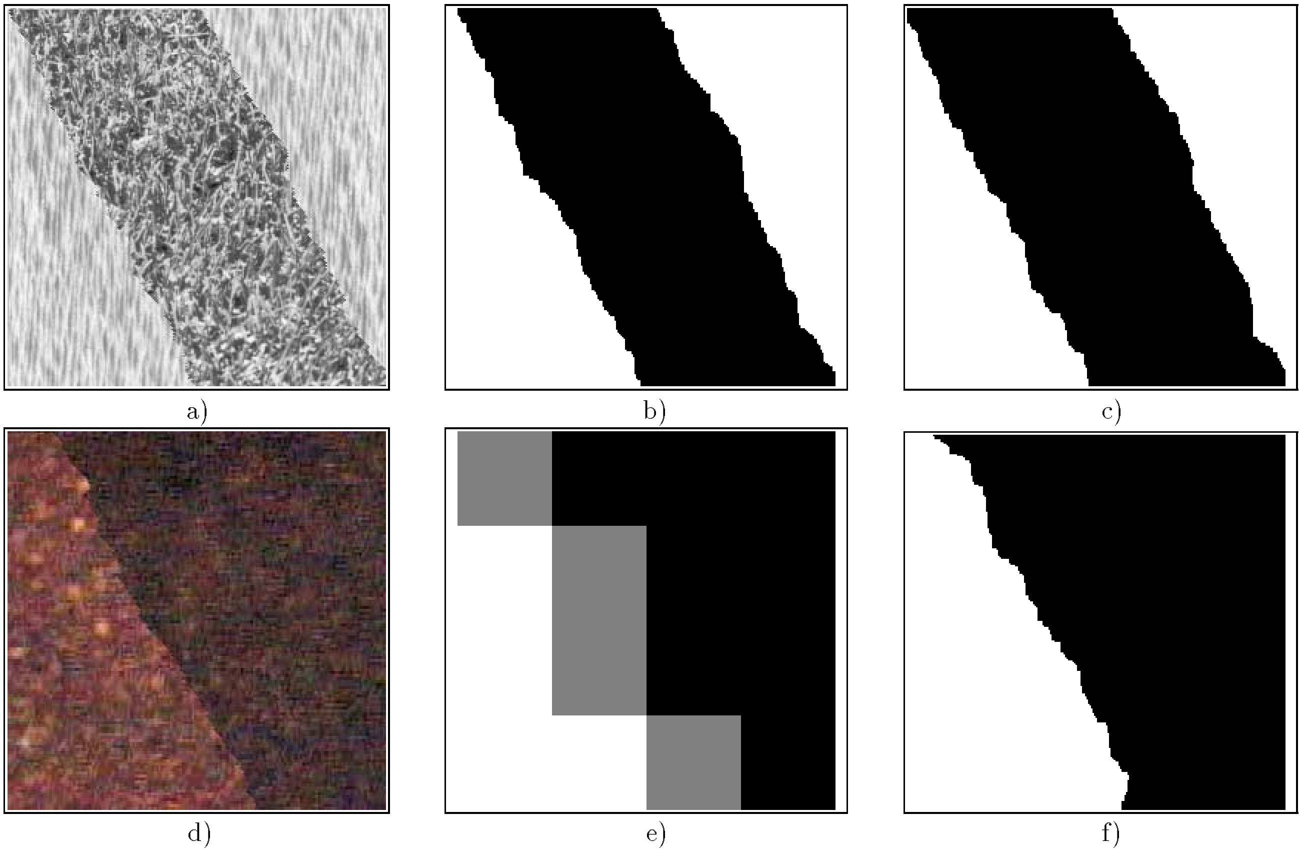

Figure 1: Segmentation examples.

Process time (sec.)	Non-iterative	Iterative
Image 1	39	115
Image 2	36	1063

Table 1: Initial region estimation.
Computational cost on a Sparc station 2.

purposes; neither process time needed to obtain the features used by both segmentation methods nor time spent during the smoothing stage are included.

5 CONCLUSIONS

A non-iterative algorithm to perform an initial segmentation on a color image is presented that greatly reduces the computational cost when compared with iterative procedures. The method is able to determine in an unsupervised way the number of regions that the final segmentation must contain avoiding both training and iterative maximization of any functional.

Homogeneous regions are found by the use of hypothesis testing whereas pixels in heterogeneous windows are classified according to the distance of their associated feature vector to the region's feature vector, using a non-euclidean distance measure.

References

[1] T.W. Anderson, An Introduction to Multivariate Statistical Analysis, *John Wiley and Sons*, 1984.

[2] Arnold O. Allen, Probability, Statistics, and Queueing Theory, *Academic Press*, 1978.

[3] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, Numerical Recipes in C, *Cambridge University Press*, 1990.

[4] G. Coleman, H. Andrews, Image Segmentation by Clustering, *Proc. of the IEEE*, Vol. 67, No. 5, 1979.

[5] S. Geman, D. Geman, Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No. 6, November 1984.

[6] C.W. Therrien, Decision, Estimation and Clasification,*John Wiley and Sons*, 1989.

[7] J. Zhang et al, Maximum Likelihood Parameter Estimation for Unsupervised Stochastic Model-Based Image Segmentation, *Trans. on Image Processing*, Vol. 3, No. 4, July 1994.

[8] P. Brodazt, Textures: A Photographic Album for Artists and Designers. New York: Dover 1966.