Detection of objects composed of several regions by a region-configuration-estimation method

Vincent Pagé, Javier García,* and Philippe Réfrégier

Physics and Image Processing Group, Fresnel Institute, Ecole Nationale Supérieure de Physique de Marseille, 13397 Marseille Cedex 20, France

Received April 17, 2001; accepted June 12, 2001; revised manuscript received July 3, 2001

The task of detection of objects composed of several regions by means of statistical filters is analyzed. The target is assumed to have different unknown mean values in each of its regions. The detection is based on likelihood estimation, after performing an estimation of the actual configuration of the mean values in the target region. A simplified filter that reduces the computational complexity is also proposed. The statistical performance is analyzed theoretically and tested in computer experiments. © 2002 Optical Society of America

OCIS codes: 070.4340, 070.5010, 100.2000.

1. INTRODUCTION

The field of pattern recognition has been extensively investigated throughout the last decades. The major landmark in this research, from an optics point of view, is the introduction of matched filters[1] and the optical implementation due to VanderLugt.[2] The basic operation that can be performed in an optical system is the convolution/correlation, which makes these systems especially suitable for real-time information processing and, in particular, to pattern recognition.

Several modifications of the classical matched filter have been proposed to deal with different types of scenes. Nevertheless, the linearity of the system limits the field of application. Specific examples occur when there is an implicit nonlinearity in the input image model, such as nonadditive noise or background.[3,4] Often the linear-filtering techniques have been complemented by electronic or optical nonlinear steps to cope with these cases.[5–8] Aside from these linear approaches to pattern recognition, in electrical engineering and computer science a big effort has been devoted to techniques that rely on Bayesian estimation theory.[9] Many of the results obtained by the application of this theory involve nonlinear processing, although optoelectronic implementations are feasible.[10] The most widely used statistical methods are related to the likelihood calculations. The theory can be extended to a consideration of different situations for describing the background and the target, such as unknown gray levels in the target and background, strongly noisy images, or nonhomogeneous background.[11–14] In general, in most pattern recognition methods, the definition of the object is very precise, including the shape and the spatial intensity distribution of the object. Very often the presence of strong noise in the images promotes the simplification of the description of the target, because some object characteristics are hidden by the noise. Along these lines, a great effort has been devoted to the search for optimal algorithms that can deal with homogeneous objects with unknown gray-level values and in the presence of a nonoverlapping background. In this case the simplification of the target model reduces it to a single-region description.[12]

Recently, a new concept of statistical filters adapted to a target composed of several *a priori* known regions was introduced.[15] The basic idea was to split the object region into a set of regions that may exhibit different unknown mean values. Thus this approach provided a trade-off between a complete description of the spatial gray-level distribution (typical for linear-filter design) and a binary description of the target (as was produced by statistical-pattern-recognition filters). This situation of multiregion objects arises in many practical situations, such as three-dimensional objects under varying illumination or thermal images.

In this paper, we analyze the advantages and drawbacks of the multiregion statistical filters and propose a filter that will automatically select the number of regions that configure the object. This produces a filter more robust under the change in the gray levels of the target. The problem of a full region-configuration selection is studied, and a useful simplification is presented.

In Section 2 the basic theory for statistical pattern recognition is introduced, along with a description of the conventional statistical filters. In Section 3 the performance of the multiregion filter is compared with that of a single-region conventional filter. In Section 4 a new filter concept for region-configuration selection is introduced, and in Section 5 a useful simplification of this filter is introduced. Section 6 presents the experiments performed for testing the new filters. Finally, in Section 7 we outline the conclusions.

2. STATISTICAL FILTERS

The design of statistical filters for pattern recognition is based on a comparison of the probability of different hypotheses that describe the expected content of the input scene. In the absence of *a priori* knowledge of the distribution of probability among these hypotheses, the usual description simplifies to a likelihood comparison.

 © 2002 Optical Society of America

A basic filter is that designed for locating a homogenous object on a homogeneous background, where both the target and the background are corrupted by noise. In this case the logarithm of the likelihood l is

$$l = \log p(\theta_{\mathbf{r}}, \mathbf{s}) + \log p(\theta_{\mathbf{B}}, \mathbf{s}), \tag{1}$$

where $\mathbf{s} = \{s_i | i \in [1, N]\}$ is the input signal, N is the number of pixels, $\mathbf{r}$ is a binary window that defines the pixels that belong to the target, and $\mathbf{B}$ is the complementary window to $\mathbf{r}$ (and defines the pixels belonging to the background). $\theta_{\mathbf{r}}$ and $\theta_{\mathbf{B}}$ are the statistical parameters that describe the gray-level distributions of both the target and the background regions. A maximization of l for different locations of the window $\mathbf{r}$ will produce a maximum in the location where the probability of presence of the specified target in the specified background is maximum. This basic filter uses the full image for calculating the likelihood of the presence of the target. A significant step forward in the performance of statistical filters for pattern recognition is obtained by the maximum-likelihood-ratio test[16–18] (MLRT). This procedure introduces two major modifications. First, the estimation of likelihood is performed in a local window, not in the whole image. This reduces the influence of nonhomogeneities in the background, making it especially appropriate for detection tasks. Second, the magnitude that is being measured is not a likelihood of the hypothesis of presence of the target but the ratio between this value and the likelihood of the hypothesis of no target present in the test window. The performance in target location is similar to that with the ML filter, as long as the test window is sufficiently large.[18] To obtain the analytical expression of the MLRT, we define the target window $\mathbf{r}$ as above, and the local background window $\mathbf{b}$ in an analogous way. The union of both windows define the test window, $\mathbf{F} = \mathbf{b} + \mathbf{r}$. Then the MLRT is

$$r = \log P(\theta_{\mathbf{r}}, \mathbf{s}) + \log P(\theta_{\mathbf{b}}, \mathbf{s}) - \log P(\theta_{\mathbf{F}}, \mathbf{s}). \tag{2}$$

Note that the first two terms correspond to the likelihood of the presence of the target on the local background, while the third term corresponds to the likelihood of a uniform background with parameters $\theta_{\mathbf{F}}$. If the image is pixelated and the pixel gray levels are uncorrelated and statistically independent in the different considered regions, each probability becomes a product of the probability for each pixel. The above expression assumes a uniform definition for the target and for the background. A more complete description can be introduced by separating the target window into a set of nonoverlapping-region windows, $\{\mathbf{r}^{(k)}; k \in [0, L-1]\}$. This allows for the target having distinct distribution parameters in each region that defines it. The region windows fulfill $\mathbf{r} = \Sigma_{k=0}^{L-1} \mathbf{r}^{(k)}$. Thus the MLRT, $r_k^{(L)}$, for testing the hypothesis of having a target composed of L regions with parameters $\{\theta^{(k)}\}$ on a background with parameters $\theta_{\mathbf{b}}$ against the hypothesis of a uniform patch with parameters $\theta_{\mathbf{F}}$ is

$$r_k^{(L)} = \sum_{k=0}^{L-1} \left\{ \sum_{i \in \mathbf{w}^{(k)}} \log P[\theta^{[k]}, s_i] \right\} + \sum_{i \in \mathbf{b}} \log P(\theta_{\mathbf{b}}, s_i) - \sum_{i \in \mathbf{F}} \log P(\theta_{\mathbf{F}}, s_i). \tag{3}$$

The subindex k implies that the θ parameters of the distributions are assumed to be known. If the parameters are not known, a selection of the values based on maximizing $r_k^{(L)}$ must be performed, by obtaining the ML estimates, and the resulting values must be substituted in the previous expression.

As a simple clarifying example the MLRT for exponential noise, which appears, for instance, in synthetic-aperture-radar images[19–21,15] with unknown parameters, is defined by

$$r_u^{(L)} = -\sum_{k=0}^{L-1} N^{(k)} \log\left[\frac{1}{N^{(k)}} \sum_{i \in \mathbf{r}(k)} s_i\right] - N_b \log\left(\frac{1}{N_b} \sum_{i \in \mathbf{b}} s_i\right) + N_F \log\left(\frac{1}{N_F} \sum_{i \in \mathbf{F}} s_i\right), \tag{4}$$

where $N^{(k)}$, N_b, and N_F are the number of pixels in the target regions, in the local background, and in the test window, respectively. The total number of pixels in the target is $N = \Sigma_{k=0}^{L-1} N^{(k)}$. The filter design assumes a different gray level in every region, although it will also perform if some of the regions present the same gray level.

For comparison, the MLRT for a one-region description will be

$$r_u^{(1)} = -N \log\left(\frac{1}{N} \sum_{i \in \mathbf{r}} s_i\right) - N_b \log\left(\frac{1}{N_b} \sum_{i \in \mathbf{b}} s_i\right) + N_F \log\left(\frac{1}{N_F} \sum_{i \in \mathbf{F}} s_i\right). \tag{5}$$

It can be noted that the samples on the target in the one-region MLRT are averaged in a single value, while for the multiregion MLRT the samples are averaged in separate values and nonlinearly added.

3. ANALYSIS OF THE MULTIREGION MAXIMUM-LIKELIHOOD-RATIO TEST

The filter in Eq. 5 is designed for the detection of homogenous objects. It is expected that this filter will fail if the object contains a higher number of regions. This can be easily checked experimentally. In Fig. 1(a) an image composed of three cases of the same object is shown. Each of the objects satisfies the same four-region description, but with different gray levels in each region. As a particular case, the rightmost object has one effective region, all regions having the same gray value. This image has been corrupted with exponential noise [Fig. 1(b)]. The correlation obtained with a one-region filter [Fig. 2(a)] shows that the first two objects cannot be detected with this filter. The filter designed for four-region description, according to Eq. (4), provides good correlation peaks for all instances of the target.

Nevertheless, the use of a more complex model for the object does not necessarily lead to better detection capabilities. Several factors may occur that worsen the results. On the one hand, the finer details that make the

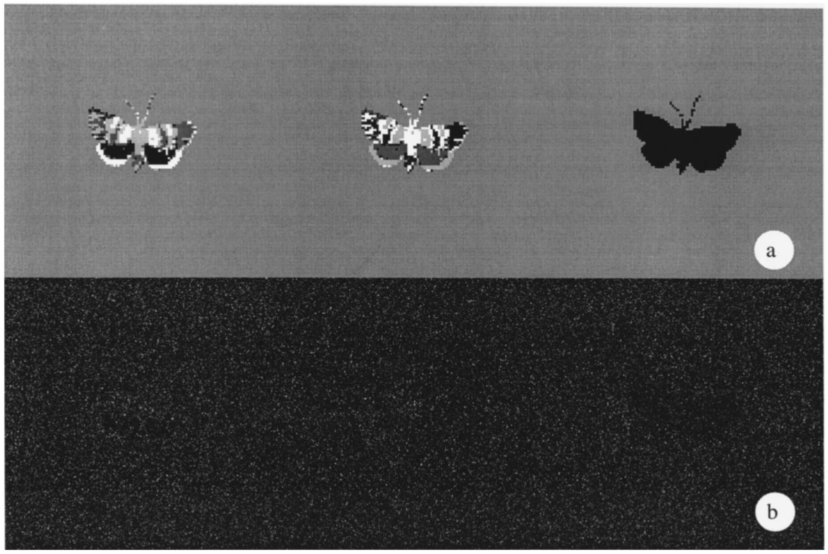

Fig. 1. (a) Noise-free image showing three objects with different gray-level combinations. (b) Same image as in (a), but corrupted with exponential noise.

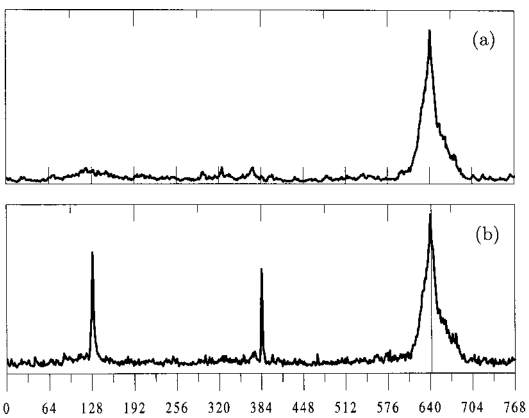

Fig. 2. Output profiles obtained for (a) one-region MLRT (b) four region MLRT. Input image is shown in Fig. 1(a). Profiles show the maximum of every column in the output MLRT images.

object definition will tend to provide a higher signal even in regions where there is nothing but noise. On the other hand, the target itself may not convey the full-region definition. The introduction of a multiregion model is based on the fact that some objects may exhibit different parameters in separate regions. One particular case occurs when all the regions have the same parameters (that is, the target is homogeneous) or, in general, when several regions in the object definition have the same parameters or are indistinguishable because of the uncertainty of the estimation of the parameters. In this situation the number of regions has been overestimated. In this section we analyze the loss of performance that may occur for these reasons.

The effect of an increase in the order of the model (i.e., the number of regions) on the likelihood is, in the case of uniform samples, a well-known problem in statistics theory. Following Ref. 22 or Ref. 20, if the target-support region contains only homogenous samples, the difference in the output for a MLRT filter for L and P regions' models is

$$r_u^{(P)}(\text{uniform}) - r_u^{(L)}(\text{uniform}) = X/2, \tag{6}$$

where X is a random variable that approaches a chi-square (χ^2) distribution with $P - L$ degrees of freedom when the number of samples tend to infinity. The relation holds independently of the mean value in the uniform patch where the filter is applied and independently of the number of pixels in every region, provided that they are large enough to fit the noise model. The quantification of the minimum number of pixels in each region was checked by a simple experiment. Figure 3 shows the results obtained for the difference between MLRT computed with four and one regions, as defined by Eq. (4). The probability density function (pdf) corresponding to a third-degree χ^2-distribution random variable divided by 2 is shown along with the pdf's for the true difference obtained for three different numbers of pixels in each region. This test shows that the fit between the theoretical result and the true result is excellent for 20 pixels per region, although this condition could be relaxed further.

Equation (6) gives the key to the penalty introduced by adding an arbitrary number of unnecessary regions to the target definition. It shows that for the case of a uniform target, the value of the output will increase with the complexity of the model that is used in the filter design.

To disclose the transcendence of this concept for detection, let us assume a test object composed of one homogenous-region target on a uniform background and compare the output for a one-region and a P-region filter. The homogenous object is just a particular case of a multiregion target, and thus the processing should be able to detect it. When passing from a one-region filter to a P-region filter, the output will suffer an addition of a χ^2-distributed variable, with order $P-1$, as corresponds to the splitting of the target window in the P regions. This addition to the output will occur for object locations when the object is homogenous and for background locations that are assumed to be always homogenous areas. Then the expected values of the output for both the target and the background have suffered an increase with the same statistical distribution. As a result the difference of output mean values for target and background remains

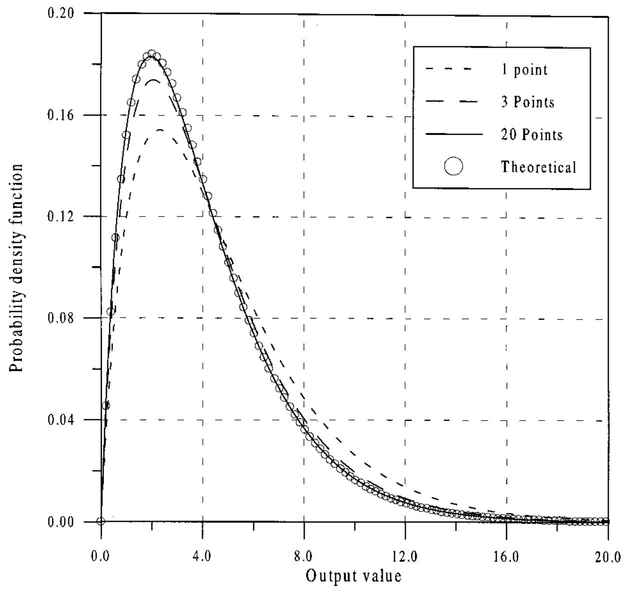

Fig. 3. Pdf of the difference between the MLRT for four region- and MLRT for one region. The four regions are the same size, taking values of 1, 3, and 20 points per region. Theoretical asymptotic pdf is shown for comparison.

unchanged, whatever the number of regions is considered, preserving the absolute separation between the output for the two classes. Nevertheless, the spread of the values for the background and for the target around their mean values is larger, as corresponds to adding a random variable to the previous output. Therefore output values for both background and target will become harder to separate. A simple, although incomplete, way to characterize this loss is by means of the Fisher ratio.[9] Let us call μ_{1T}, σ_{1T}^2 the mean and variance of the output when a target is present, and μ_{1B}, σ_{1B}^2 the mean and variance of the output when there is only background, in both cases using a one-region filter. Then the Fisher ratio, which describes the separability between the background and the target classes, will be

$$F_1 = \frac{|\mu_{1T} - \mu_{1B}|^2}{\sigma_{1T}^2 + \sigma_{1B}^2}. \tag{7}$$

If the output distribution for the P-region filter is assumed to be independent of the one-region filter output,

$$F_P = \frac{|\mu_{PT} - \mu_{PB}|^2}{\sigma_{PT}^2 + \sigma_{PB}^2} = \frac{|\mu_{1T} - \mu_{1B}|^2}{\sigma_{1T}^2 + \sigma_{1B}^2 + \sigma_X^2/2}, \tag{8}$$

where σ_X^2 is the variance of the χ^2 distribution. As can be easily checked, the output for a single-region filter will degrade as the number of regions in the model is increased. In addition to this global parameter, given the contrast target/background (ratio of means), the output value pdf's for both the background and the target can be estimated, provided that the noise model is known.

As an example, let us consider a scene corrupted by exponential noise. The object is composed of one homogenous region with mean value 10, and the background has a mean of 13. The number of pixels for the target is 400 and the local background has 600 pixels. We will consider the case of splitting the target into four regions with 100 pixels each. This number was already shown to provide a sufficient number of samples. Figure 4 depicts the pdf's for the target and for the background for the one- and four-region filters. As stated before, the pdf's broaden as the number of regions is increased, although they keep the same separation. Therefore the overlapping of the output for target and background classes enlarges, which increases the cases where the classes can be confused.

For a more complete characterization of the detection properties, Fig. 5 shows the plot of probability of correct detection versus the probability of false alarm [receiver operating characteristic (ROC) curve] for a one-region object for a varying number of regions in the filter definitions. It is obvious that the performance of the filter decreases, for this case, as the number of regions in the filter design increases. The same argument is applicable for the expected difference between the output of the MLRT filters with different numbers of regions if the number of regions have been overestimated in the filter design. If the object fulfills an L-region definition, the difference between the output of an $L + P$ region's filters and an L region's filter will obey a χ^2 pdf with order P.

On the other hand, if the splitting of one region really matches the object definition, the expected value of the increase will be much larger than the one predicted in Eq. (6) (for instance, see Ref. 20 for the expression for one region split in exponential noise).

In the analysis of the filters throughout this paper, the actual shapes of the regions are not relevant, as we do not consider the spatial characteristics of the output MLRT but only the central value. In a real case, the output of the filters will also include points that are not located at the center of the target or in a homogenous patch, giving rise to sidelobes. The importance of sidelobes will depend on the spatial characteristic of the regions. The in-

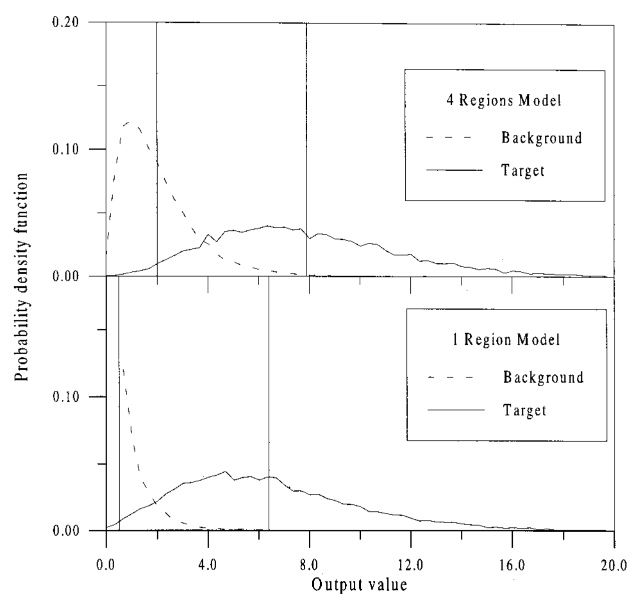

Fig. 4. Pdf's at target and background locations obtained with one-region MLRT filter (upper graph) and four-region MLRT filter (lower graph). The mean values for background and target are marked with vertical lines. Note that the separation of the background target remains constant, but the variance of both classes increases, as seen in the overlapping of the pdf's.

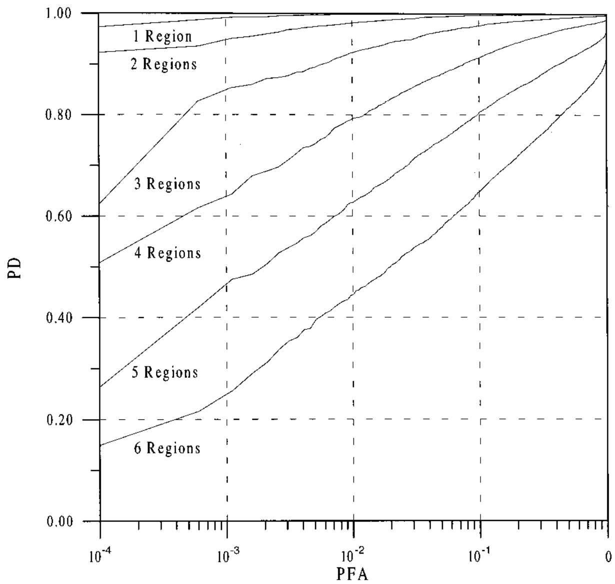

Fig. 5. Losses in performance when the MLRT is evaluated with an increasing number of regions of equal size for a homogeneous object.

crease in the number of regions will produce a smaller size of the regions and, as a result, the spatial-frequency contents of the filter will spread. The primary effect will be the increase in partial matches, and therefore sidelobes, outside the central location of the object. The quantification of the effect is complex, as it depends on the specific structure of the object.

We can summarize that the performance of the filter is closely linked to the match between the region configuration in the target definition used to prepare the filter and in the actual object that is being processed. It is tempting to design a detection algorithm that will estimate the region description and include this information in the detection process. The difference in the behavior when splitting a uniform or a nonhomogeneous region can serve as the key for deciding the object definition that better fits the actual object that is being inspected, as analyzed in the Section 4.

4. REGION SELECTION MAXIMUM-LIKELIHOOD-RATIO TEST

A problem similar to the one posed at the end of the previous section has been addressed in the fields of segmentation,estimation, and classification. The basic problem addressed is to decide whether a pair of regions are indeed the same region. Arising from different conceptual points of view, the most common solution is to estimate the best region description by maximizing a modified likelihood function. The modification uses a penalty term that will penalize the increasing complexity of the model.[23,24]

Following Ref. 23, the fundamental reason for this penalty is the bias in the log likelihood according to the number of free parameters. This bias is found to be the number of parameters needed to describe the model of the random process divided by 2. For the actual case, this means that the likelihood ratio should be corrected by subtracting the number of regions of the target divided by 2. This correction is in accordance with the mean values of the random variable added to the MLRT that was shown above [see Eq. (6)].

A conceptually different approach is the minimum description length (MDL) concept. Here the penalty term is directly connected to the complexity of the description.[24] The underlying philosophy is not to take a more complicated model unless a net gain in the goal function is achieved. The basic problem of MDL is the definition of the complexity of the object. Although a basic trend is usually easy to estimate, the exact value of the complexity of a configuration is, in general, complex to derive.

Whatever the basic foundation is, the likelihood should be biased by a penalty term, depending on the number of regions of the object model. In most of the literature, this penalty term is basically proportional to the number of regions.

In the above discussion only the number of regions in the model for the target have been considered. An object may fulfill this model but with a number of distinct gray levels smaller than the maximum number of regions. Then some regions are fused, and the target presents a lower number of effective regions. If the number of effective regions L is larger than one and smaller than the maximum number of regions P, there are several possible region configurations. Figure 6 shows this situation by means of an example. A region configuration is described by the number of distinct gray levels in the object and by the actual gray level in each region. We will assume that the penalty term is independent of the specific region configuration, depending only on the effective number of regions.

Thus we propose a detection filter, taking into account the possibility of fused regions; the procedure is divided in two steps. First, the MLRT of every configuration is computed and a penalty term that depends on the number of effective regions is subtracted. This step gives the estimation of the actual region configuration of the object

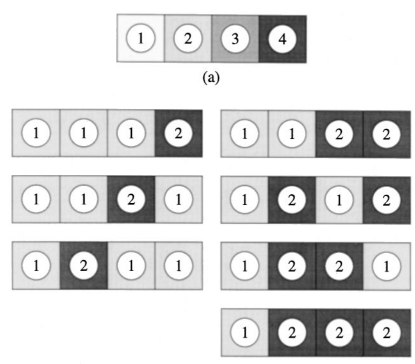

Fig. 6. Example of the different region configurations for a given number of effective regions. (a) Four-region target model defined by the spatial distribution and the maximum number of regions. (b) Different region configurations with two effective regions. The number of effective regions is defined by the number of distinct gray levels. Every region configuration is defined by four digits, where each digit is the index of the effective region, and the order in the row is the region index in the target definition, using the order in the target model description.

and, in particular, the number of estimated effective regions in the object. The region-selection (RS) number of regions is thus obtained as

$$L_{\mathrm{RS}} = \arg\max_{M}\{r_u^{(M)} - \alpha M\}. \tag{9}$$

This operation is performed for each pixel displacement in the image, so that the estimated number of regions may vary pixelwise. In a second step, the output of the RS filter (RS-MLRT) is given by the MLRT of that configuration which gives the maximum in the compensated MLRT. The RS-MLRT is

$$R = r_u^{(L_{\mathrm{RS}})}. \tag{10}$$

The detection process will conclude with a thresholding of this output to separate the background and the target classes.

The parameter α controls the penalty term. A high value of α will tend to keep regions joined unless a large difference exists, while a low value will tend to split regions more frequently with a small or even null difference in the likelihood.

As we can see, a delicate point is the selection of the α parameter. The Akaike information criterion[23] would mean the use of $\alpha = 0.5$. In the MDL framework, the value is selected according to the complexity of the definition of a region configuration. The complexity is evaluated as the number of bits needed to describe the configuration.[24] For the case of gamma noise, a region is described just by its mean value, and the standard deviation of the estimated mean is the mean divided by the square root of the number of pixels. Thus the number of distinct values needed to describe the region mean equals $\sqrt{N}$, the result of dividing the mean by its uncertainty. Then the penalty term will be

$$\alpha_{\mathrm{MDL}} = \log\sqrt{N}. \tag{11}$$

Nevertheless, note that this value does not take into account the spatial distribution of the regions.

An additional factor to be considered in the region-estimation approach is the number of possible region configurations. If the object is composed of P regions, we have to consider in the test for region-configuration estimation all possible combinations. The number of possible configurations that must be checked can be obtained by combinatorial analysis. The number of possibilities of partitioning a set of P regions into L subsets is given by the Stirling number[25] of the second kind, $S_P^{(L)}$. The number of possible configurations grows exponentially as the number of regions increases, making the use of a number of regions larger than 6 or 7 difficult in practice (see Table 1 for the listing of the number of configurations for the first ten values of the maximum number of regions). For a sufficiently small number of regions in the object definition, the algorithm for calculating the RS-MLRT will consist of calculating the MLRT of every configuration and subtracting the penalty term, depending on the number of effective regions. The MLRT of the configuration that gives the maximum value produces the desired output.

Table 1. Number of Possible Region Configurations for the First Ten Values of the Maximum Number of Regions[a]

Maximum Number of Regions	Number of Region Configurations
1	1
2	2
3	5
4	15
5	52
6	203
7	877
8	4,140
9	21,147
10	11,5975

[a] For a maximum number of regions the number of configurations is given by $NC(P) = \Sigma_{L=1}^{P} S_P^{(L)}$, where $S_P^{(L)}$ is the Stirling number of the second kind.

The calculation of the RS-MLRT filter involves the estimation of the actual region configuration. The accuracy of the estimation will depend on the number of pixels in each region, because a larger number of pixels will make the likelihood estimation more precise. It also depends on the contrast among the target regions, because a well-contrasted target is easier to discriminate. Finally, the α value will strongly influence the estimation, because it changes the tendency of the algorithm to select configurations of higher or lower complexity. For a given detection task the number of pixels in every region are known, but the contrast may vary among different target instances. Therefore the α value should be chosen without taking the contrast into account.

A way of choosing the parameter α is to check the accuracy in the estimation of region selection. In Fig. 7 the average estimated number of regions is plotted for different values of contrast and α. The object is composed of two regions of equal size with 200 pixels each, and the region description is four equal regions 100 pixels in size. The contrast is defined as the quotient between the mean values in the two regions of the object. The behavior for other region configurations has been observed to be similar. Several facts can be concluded from this figure. For all values of the contrast, the estimated number of regions decrease monotonically from four to one as α increases. For a given α, increasing the contrast implies a larger or equal estimated number of regions. For low contrast, a continuous decay is observed while for higher contrasts, the curve stabilizes at the right number of regions (two in the example), eventually falling to one for high enough α value. As a main conclusion from these graphs, a low α value will tend to overestimate the number of regions, while a large α value will correctly estimate it, provided that the contrast is high enough. To keep the widest range of useful contrast, the lowest value of α that brings a correct estimation of region number should be considered. For the case depicted in the graph, this means a value ~3. In Fig. 8, the same plot is shown for three different numbers of regions in the target, for the case of infinite contrast. Albeit the position of the elbow of the curves is not exactly the same, a value of $\alpha = 3$ will perform properly for all cases.

The accuracy of the region-configuration estimation relies on statistical considerations. In the case of segmentation, a region is split into two, and the likelihoods before and after the splitting are compared. The regions are dynamically modified (by exchanging a few pixels) and compared, which produces a large amount of redundancy in the checks because the theoretical number of checks in every split-merge step is related to the number of pixels in the region. In the RS-MLRT case, the regions are fixed by the object definition, so a single check for every region configuration is performed. This may lead easily to incorrect splitting of the regions in the target. That is, the process of detection is based on a previous step of estimation of the region configuration that may have a large uncertainty.

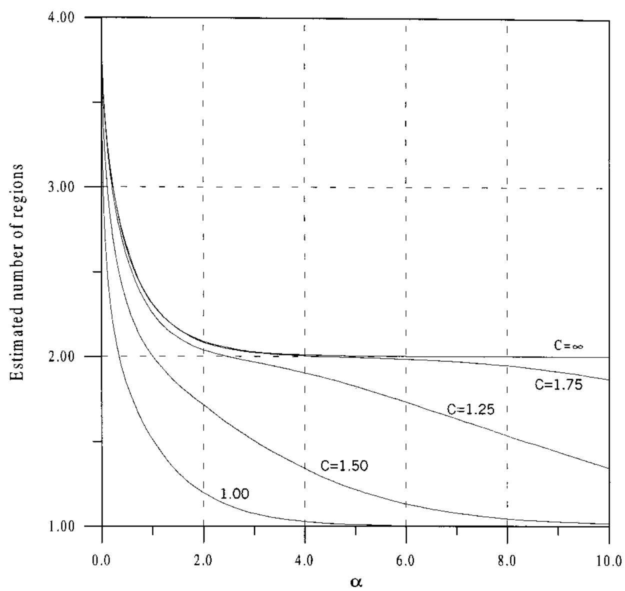

Fig. 7. Estimated number of regions as a function of the parameter α for different contrasts. The object is formed by two regions of equal size, and the RS-MLRT, filter model is four regions.

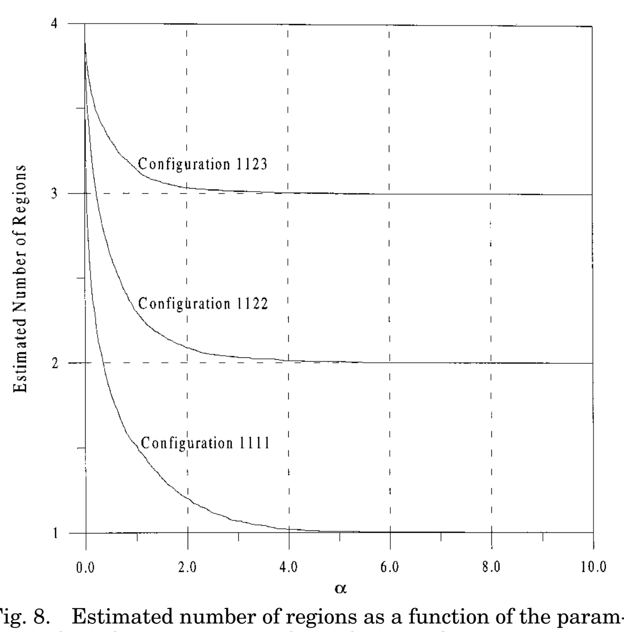

Fig. 8. Estimated number of regions as a function of the parameter α for infinite contrast. Three object configurations are considered.

5. SIMPLIFIED REGION-SELECTION ALGORITHM

As stated in Section 4, the selection of the proper configuration for an unknown object can be too complex to be calculated. A significant simplification of the algorithm can be made, specially matched for detection, relaxing the requirements of the region-estimation step. In the actual posing of the problem, only two classes are considered to be distinguished. One is the target, with an arbitrary number of distinct regions taken from the complete region description. The other is the background, which is assumed to be uniform. A difficulty in the RS algorithm is that the value of MLRT in every region configuration can be evaluated only once. Therefore the possible increase due to the unnecessary split of a region may be hidden by the uncertainty of the MLRT. This problem can be reduced by considering not one splitting at a time, but the total splitting between one region and the maximum number of regions.

If the region is indeed uniform, the average of the difference in likelihood in all steps will be closer to the value obtained from pdf's [Eq. (6)]. The same applies for a uniform target. On the contrary, if the target is multiregion, there will be a significant increase of the likelihood in the P-region's MLRT with respect to the one-region MLRT, no matter in what intermediate step (or steps) the increase is significant. To take advantage of this fact, we propose the following simplified-region-selection (SRS-MLRT) filter for detection of multiregion objects. First, the number of regions (L_{SRS}) is estimated to be one or the maximum number of regions P just by the difference in likelihoods:

$$L_{\mathrm{SRS}} = \begin{cases} P & \text{if } r_u^{(P)} - r_u^{(1)} > \alpha_{\mathrm{SRS}} \\ 1 & \text{otherwise} \end{cases}; \tag{12}$$

then the SRS-MLRT is given by

$$R_{\mathrm{SRS}} = r_u^{(L_{\mathrm{SRS}})}. \tag{13}$$

The selection between $r_u^{(P)}$ and $r_u^{(1)}$ is made independently for every pixel in the input image. With a proper choice of the parameter α_{SRS} the filter should select one region output for the background and P regions for the target. An exception will occur if the target is indeed one region. Then one region will be selected, as corresponds to the object case. A thresholding of the output will yield the final binary detection output.

The selection of the value of α for this filter is not as critical as for the RS filter, as a much larger difference is expected between the likelihoods for one and P regions than in the stepwise region-selection method. In a uniform patch, the difference will be just a random variable (divided by 2) with χ^2 distribution with $P - 1$ degrees of freedom. The probability of wrong splitting in the background can be easily obtained from this pdf.

A value of

$$\alpha_{\mathrm{SRS}} = (P - 1)\alpha, \tag{14}$$

where α is the value used for RS-MLRT, is consistent with the previous considerations.

Note the dramatic simplification in the computation of the MLRT in this case in comparison with the RS-MLRT filter. There is no need for calculating the MLRT for every possible configuration, but only for two. Moreover, the two configurations that must be calculated (namely one region and P regions) are unique and directly determined by the target-model definition.

6. RESULTS

The performance of the two newly introduced algorithms were tested by means of computer experiments. The target is in all cases made of four regions with 100 pixels each. The tests were performed for the most significant situations of contrast among the target regions and overall contrast between target and background.

Figure 9 shows the COR curves for the case of a four-region object with mean values 10, 12.5, 17.5, and 20. The background has a mean value of 15. Owing to the null overall contrast between the target and the background, the one-region filter fails in the detection of the target. The region definition matches a four-region description, which enables the four-region MLRT to perform the best in the test. The RS-MLRT closely follows the four-region filter, indicating that the number of regions have been correctly estimated. The SRS-MLRT shows a lower performance than the RS-MLRT, although it is still valuable.

In Fig. 10 the case of a one-region object with poor contrast with the background is depicted. The target is homogenous with mean value 10, and the mean of the background is 12.5. It can be seen that the one-region filter performs the best, similar to the SRS filter. The four-

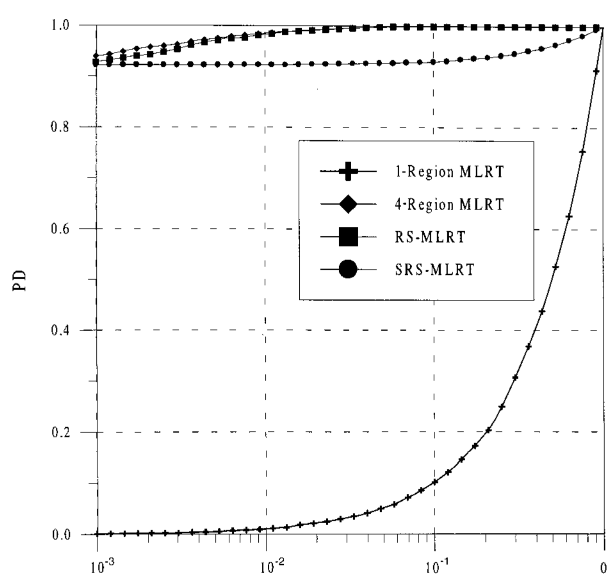

Fig. 9. COR curves for one-region MLRT, four-region MLRT, RS-MLRT, and SRS-MLRT. The object is a four-region object with null contrast with the background.

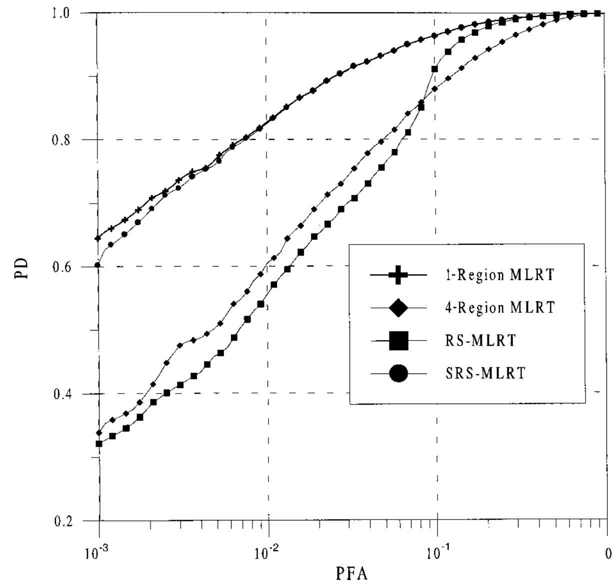

Fig. 10. Same as Fig. 9, but the object is a one-region object with low contrast with the background.

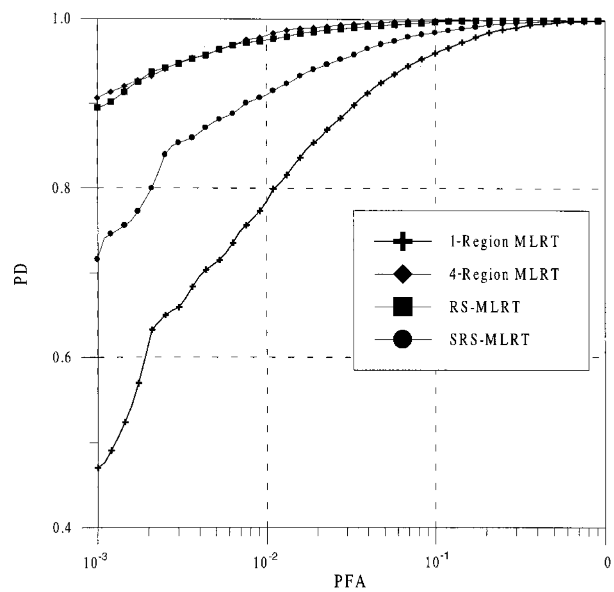

Fig. 11. Same as Fig. 9, but the object is a two-region object with low contrast with the background.

region filter has a low performance due to the losses discussed in Section 2. The RS-MLRT exhibits a mixed behavior. For low thresholds (right-hand side of the graph), it fits the one-region filter, while it is close to the four-region filter for high thresholds (left-hand side of the graph). This behavior is due to the separation that the COR curve representation makes of the incorrect estimated values. The low values of the output will, obviously, have a higher probability of producing a one-region estimation, and for high values in the output the trend will be to estimate a four-region object.

Finally, Fig. 11 is an intermediate case. The object is composed of two regions of 200 pixels each, with mean values 10 and 15. The background has a small contrast with the target (mean value 15.7). The COR curves show a good performance for the four regions and RS-MLRT filters. The one-region filter is less effective, and the SRS filter is in the middle.

Considering the above results, the RS-MLRT filter behaves the best in all cases except when the object is purely one region. In this case the performance is bad for low false-alarm probability but good for high detection probability. The SRS-MLRT is a trade-off between one-region and four-region MLRT, and exhibits good to medium detection capabilities.

7. CONCLUSIONS

The performance of statistical filters that are designed for multiregion objects has been studied. The gains of a multiregion filter are shown to be maximal if the object fulfills the multiregion definition and especially if the contrast between target and background is low. The conventional MLRT filter, matched to one single region, will behave better only if the object fits into the one-region definition. The loss of performance when an object is tested with a number of regions higher than the effective number of regions has been analyzed by means of the Fisher ratio and with COR curves.

Two different ways of combining the output corresponding to the possible region configurations of the target have been tested. The first way is to use a penalty function according to the number of regions and select the configuration out of all possible ones that maximizes the corrected likelihood. The strength and the weakness of this approach is the need for an estimation of the configuration of the regions in the target. If this intermediate step does not provide accurate results, the final output will have a reduced performance.

We proposed a second method for distinguishing between the target and the background: a selection between one-region and a maximum-region definition. This considerably simplifies the computing of the filter and produces a better estimation in the background regions. Experiments show a good compromise in the detection capabilities of the new proposed filters with respect to the previously described filters, as they do not present any severe failure for any type of objects, for which where previous filters have failed.

The correct estimation of the region configuration has shown to be the critical issue in the filter design. A deeper investigation of this subject is to be developed.

ACKNOWLEDGMENTS

Javier García acknowledges a grant from the Spanish Secretaría de Estado de Educación y Universidades. The authors thank the Direcció General de Recerca of the Generalitat de Catalunya (project XT99-0014).

*Permanent address, Departamento de Óptica, Universitat de València, Dr. Moliner, 50, 46100 Burjassot, Spain.

REFERENCES

1. G. L. Turin, "An introduction to matched filters," IRE Trans. Inf. Theory **IT-6**, 311–329 (1960).
2. A. VanderLugt, "Signal detection by complex filtering," IEEE Trans. Inf. Theory **IT-10**, 139–145 (1964).
3. B. V. K. Vijaya Kumar, D. P. Casasent, and A. Mahalanobis, "Correlation filters for target detection in a Markov model background clutter," Appl. Opt. **28**, 3112–3117 (1989).
4. B. Javidi and J. Wang, "Design of filters to detect a noisy target in nonoverlapping background noise," J. Opt. Soc. Am. A **11**, 2604–2612 (1994).
5. J. M. Hereford and W. T. Rhodes, "Nonlinear optical image filtering by time-sequential threshold decomposition," Opt. Eng. **27**, 274–279 (1998).
6. B. Javidi, J. Li, A. Fazlolahi, and J-L. Horner, "Binary nonlinear joint transform correlator performance with different thresholding methods under unknown illumination conditions," Appl. Opt. **34**, 886–896 (1995).
7. P. García-Martinez, D. Mas, J. García, and C. Ferreira, "Nonlinear morphological correlation: optoelectronic implementation," Appl. Opt. **37**, 2112–2118 (1998).
8. P. García-Martinez, H. H. Arsenault, and S. Roy, "Optical implementation of the sliced orthogonal nonlinear generalized correlation for images degraded by nonoverlapping background noise," Opt. Commun. **173**, 185–193 (2000).
9. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis* (Wiley, New York, 1973).
10. D. Prévost, Ph. Lalanne, J. C. Rodier, and P. Chavel, "Video-rate simulated annealing for stochastic artificial retinas," Opt. Commun. **132**, 427–431 (1996).
11. A. H. Fazlollahi, B. Javidi, and P. Willet, "Minimum-error-probability receiver for detecting a noisy target in colored background noise," J. Opt. Soc. Am. A **14**, 845–852 (1997).
12. F. Goudail and P. Réfrégier, "Optimal detection of a target with random gray levels on a spatially disjoint noise," Opt. Lett. **21**, 495–497 (1996).
13. P. Réfrégier, F. Goudail, and C. Chesnaud, "Statistically independent region models applied to correlation and segmentation techniques," in *1999 Euro-American Workshop on Optoelectronic Information Processing*, Ph. Réfrégier and B. Javidi, eds., Vol. CR74 of SPIE Critical Review Series (SPIE, Bellingham, Wash., 1999), pp. 193–224.
14. B. Javidi, P. Réfrégier, and P. Willet, "Optimum receiver design for pattern recognition for nonoverlapping target and scene noise," Opt. Lett. **18**, 1660–1662 (1993).
15. J. García, V. Pagé, and P. Réfrégier, "Detection of multiregion objects embedded in nonoverlapping noise," Opt. Lett. **26**, 1394–1396 (2001).
16. P. H. Garthwaite, I. T. Jolliffe, and B. Jones, *Statistical Inference* (Prentice-Hall, London, 1995).
17. H. Sjöberg, F. Goudail, and P. Réfrégier, "Optimal algorithms for target detection in nonhomogeneous binary images," J. Opt. Soc. Am. A **15**, 2976–2985 (1998).
18. V. Pagé, F. Goudail, and P. Réfrégier, "Improved robustness of target location in nonhomogeneous backgrounds by use of the maximum-likelihood ratio test location algorithm," Opt. Lett. **24**, 1383–1385 (1999).
19. M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions* (Wiley, New York, 1993).
20. C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images* (Artech House, Norwood, Mass., 1998).
21. F. Guérault and P. Réfrégier, "Unified statistically independent region processor for deterministic and fluctuating targets in nonoverlapping background," Opt. Lett. **22**, 630–632 (1997).
22. A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, Boston 1991).
23. Y. Sakamoto, M. Ishiguro, and G. Kitagawa, *Akaike Information Criterion Statistics* (KTK Scientific, Tokyo 1986).
24. J. Rissanen, *Stochastic Complexity in Statistical Inquiry* (World Scientific, Singapore, 1989).
25. M. Abramowitz and I. Stegun, eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (U.S. GPO, Washington, D.C., 1972).