Detection of multiregion objects embedded in nonoverlapping noise

Javier García,* Vincent Pagé, and Philippe Réfrégier

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

Received April 2, 2001

The concept of a statistical filter for objects that comprise several regions is introduced. The process is optimal in the presence of nonoverlapping noise for the target and may perform independently of variations in the mean value in every region. The basic performance of the filter is described, and a comparison with other types of processing is made.

OCIS codes: 070.5010, 070.4550.

The first optimal filters for pattern recognition were matched filters[1] and their optical implementations. These filters produce the optimal signal-to-noise ratio in the output in the presence of additive noise and when the gray levels of the target are known. Some heuristic pattern-recognition methods have been introduced for cases when the gray levels of the target are not completely specified[3] or when the noise is nonoverlapping.[4] We have filters based on statistical estimation theory, however, most of which were obtained by the maximum-likelihood (ML) approach, that optimize the probability of correct detection relative to the probability of a false alarm. The statistical methods where applied mainly for simple objects that exhibit homogeneous statistics,[5,6] although complete knowledge of the object's gray levels can also be introduced.[7] The gap between complete knowledge of the object and a simplified model of a homogeneous target may be excessive for some practical detection–recognition tasks. We introduce partial knowledge of the object gleaned by a definition of multiple regions and used in a statistical filter approach.

It is likely that the target is composed of a few homogeneous regions but with relative mean value in every region that may vary. A clear example is an object under various kinds of illumination. Depending on the direction of illumination or the number of light sources, the mean value for every region of the target will change. The inclusion of several regions in a ML processor is natural for segmentation but seldom has been introduced in pattern recognition. In Ref. 8 the likelihood of existence of three regions is taken into account, but these regions (target bulk, border between target and background, and background) are not part of the target model but are the result of preprocessing of the input scene.

Let us consider a scene image composed of N pixels, $\mathbf{s} = \{s_i \,|\, i \in [1, N]\}$. We use one-dimensional notation without loss of generality. A similar definition applies for the target $\mathbf{r}$ and for the background $\mathbf{b}$. We assume that the gray levels of the target and the background are random, spatially uncorrelated, and distributed with different probability-density functions. The target is defined inside a support window $\mathbf{w}$, which takes a value of 1 for pixels inside the target and of 0 otherwise to determine the target shape. One can boost the performance of a ML filter by considering not the whole background of the image but only a local background surrounding the target. The corresponding algorithm, the ML ratio test (MLRT),[9] tests for every pixel shift the hypothesis (H_1) that a target with shape $\mathbf{w}$ is present and is surrounded by background noise in a test window $\mathbf{F}$ against the hypothesis (H_0) that inside $\mathbf{F}$ there is nothing but noise. For a test window $\mathbf{F}$ a few times larger than the target, it has been shown[10] that the probability of correct location of the target is similar to that obtained with a simple ML test, whereas the MLRT filter performs much better with nonhomogeneous noise in the background. Denoting by $\theta_{\mathbf{w}}$ ($\theta_{\mathbf{b}}$) the parameters that define the probability-density function of the gray-level statistical distribution of the target (background) yields the MLRT expression

$$r = \log P(\theta_{\mathbf{w}}, \mathbf{s}) + \log P(\theta_{\mathbf{b}}, \mathbf{s}) - \log P(\theta_{\mathbf{F}}, \mathbf{s}), \qquad (1)$$

where the two first terms correspond to the log likelihood of hypothesis H_1, whereas while the third term stands for the log likelihood of hypothesis H_0. We define the MLRT for every pixel in the image merely by shifting all windows.

To define the object regions we let L be the number of regions and let $\{\mathbf{w}^k; k \in [0, L-1]\}$ be a set of binary nonoverlapping windows that defines the pixels that belong to every region. The window set must fulfill $\mathbf{w} = \sum_{k=0}^{L-1} \mathbf{w}^k$. Assuming that the noise on the image is uncorrelated and statistically independent in every region and in the background, the probabilities can be obtained as the product of the probability for every pixel. Thus the MLRT becomes

$$r_{\text{known}}^{(L)} = \sum_{k=0}^{L-1} \sum_{i \in \mathbf{w}^k} \log P(\theta_{\mathbf{w}^k}, s_i) + \sum_{i \in \mathbf{b}} \log P(\theta_{\mathbf{b}}, \mathbf{s}_i) - \sum_{i \in \mathbf{F}} \log P(\theta_{\mathbf{F}}, s_i), \qquad (2)$$

where the superscript L denotes the filter designed for L regions.

As a sample of application, for images corrupted with exponential noise[11] the parameters θ are just the mean values of every region. A little algebra shows that the MLRT when the parameters are known is

$$r_{\text{known}}^{(L)} = \sum_{k=0}^{L-1} N^k \log \theta_{\mathbf{w}^k} - N_b \log \theta_{\mathbf{b}} + N_F \log \theta_{\mathbf{F}} - \sum_{k=0}^{L-1} N^k \frac{\hat{\theta}_{\mathbf{w}^k}}{\theta_{\mathbf{w}^k}} - N_b \frac{\hat{\theta}_{\mathbf{b}}}{\theta_{\mathbf{b}}} + N_F \frac{\hat{\theta}_{\mathbf{F}}}{\theta_{\mathbf{F}}}, \qquad (3)$$

where the values marked with circumflexes are the actual estimates in the ML sense. N^k, N_b, and N_F are the numbers 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 $N_t = \sum_{k=0}^{L-1} N^k$. When the parameters are unknown, they are simply replaced by their estimates in Eq. (3), defining the following MLRT:

$$r_{\text{unknown}}^{(L)} = -\sum_{k=0}^{L-1} N^k \log \hat{\theta}_{\mathbf{w}^k} - N_b \log \hat{\theta}_{\mathbf{b}} + N_F \log \hat{\theta}_{\mathbf{F}}. \qquad (4)$$

The filters in Eqs. (3) and (4) will permit the detection of objects composed of an arbitrary number of regions with, respectively, unknown or known mean values in every region.

In what follows, we evaluate the performance of the new filter by means of several experiments. Figure 1a shows a synthetic test image used in the first experiments. The use of a synthetic image allows access to the true values of the gray levels before noise corruption as well as controlling the contrast of the image. The scene is composed of three versions of the same object with different distribution of the gray levels between the four possible regions. At left is a single region object of mean 25; the other two objects comprise four regions, the means of which are, from darker to brighter, 5, 10, 30, and 35. The mean of the background is 20, so the objects in the center and the right parts of Fig. 1b present a low average contrast with the background. The total number of pixels within the target is 2568. Figure 1b shows the same image corrupted by exponential noise.

Figure 2 illustrates the output of three versions of MLRT applied to the images of Fig. 1b by means of normalized profiles along the maximum value. In all cases, the test window is a square of 91 × 91 pixels. The evaluation of the new filter in relation to other MLRT filters can be obtained from a simple ranking of the available *a priori* information and contrasted with these figures. The performance of the single-region filter is superior for the one-region object but at the price of losing detection capabilities for the four-region objects (Fig. 2a). The filter in Eq. (3) will perform if the object matches the multiregion model in shapes and the mean values of the regions (the filter was prepared for the true mean values of the second object). It will fail if the object does not have the same region description, either in number of regions or in their mean values, as a result of incorrect *a priori* information (Fig. 2b).

The MLRT for four regions with unknown mean values [Eq. (4)] can correctly detect instance of both four-region objects and a single-region object (Fig. 2c). In the latter case the performance is worse than that of a one-region MLRT because the *a priori* information is not complete.

It is worth noting that the background for the output planes (Figs. 2b and 2c) exhibits a larger variation because of the smaller structures involved in the MLRT for these cases. Actually, analyzing the statistics of a homogeneous region through a patch of several small areas instead of a single large region necessarily leads to an additional fluctuation of the output. Such appears to be the case when the test window contains only homogeneous background noise.

We tested the robustness and wide range of application of the new filter in a second experiment involving real images. The image, shown in Fig. 3a, is a collection of low-light-level snapshots taken with a

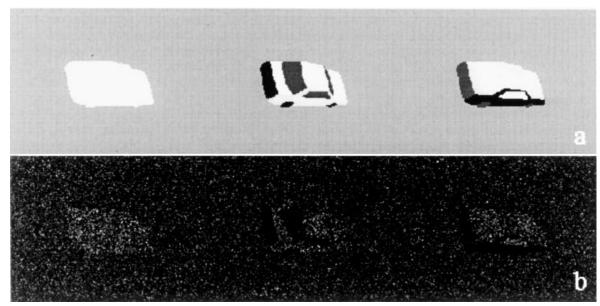

Fig. 1. Input image used for the experiments: a, noise-free image; b, image corrupted with exponential noise.

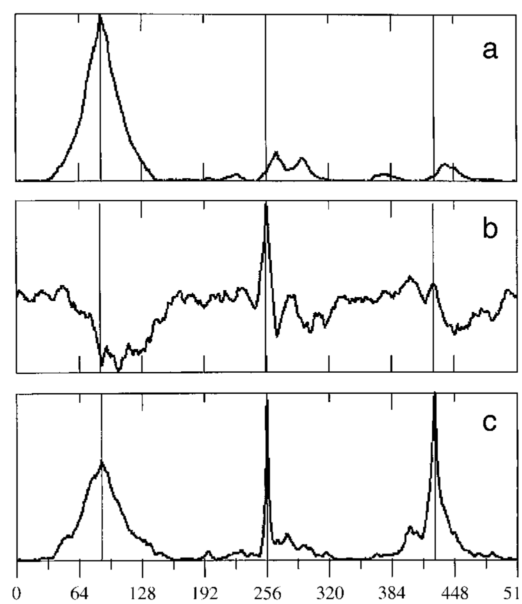

Fig. 2. Normalized profiles of the output plane for a, a single-region MLRT with an unknown mean; b, a four-region MLRT with known mean values; and c, a four-region MLRT with unknown mean values. Vertical lines indicate the target locations.

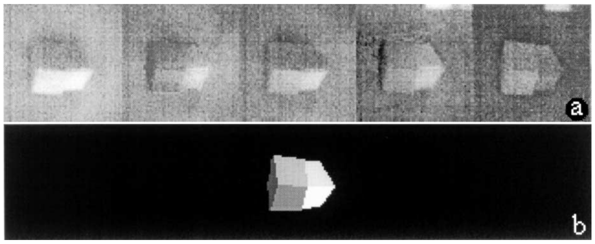

Fig. 3. a, Images from a CCD camera at low light levels, showing five cases of a three-dimensional object under different illumination conditions. b, Description of the object's regions, indicated with gray levels.

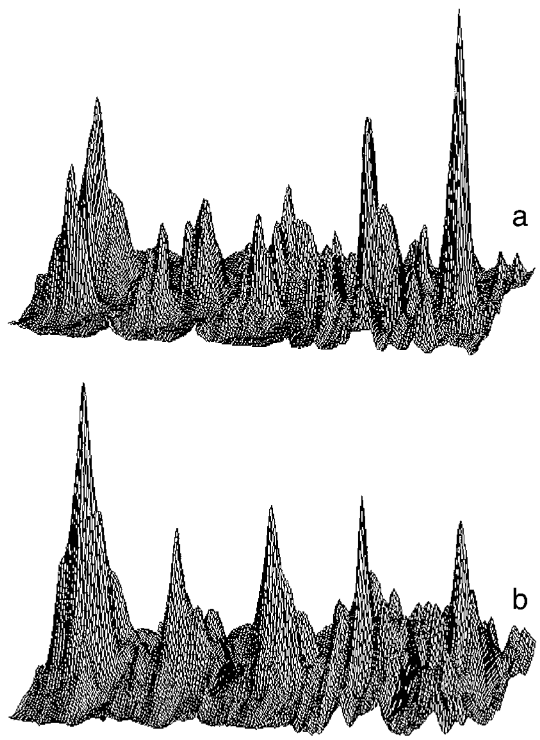

Fig. 4. a, Output for the one-region MLRT, b, output for the five-region MLRT.

CCD camera of a three-dimensional object with different backgrounds and under different illumination conditions. The object shows five planar facets to the camera, which constitute the object regions (Fig. 3b). Owing to the low number of photodetections, the noise fits a Poisson model, for which the sufficient statistic is the mean value. For real images with unknown illumination, known parameters cannot be evaluated, as the true values are not defined. The MLRT for Poisson noise with unknown parameters is

$$r'_{\mathrm{unknown}} = \sum_{k=0}^{L-1} N^k \hat{\theta}_{\mathbf{w}^k} \log \hat{\theta}_{\mathbf{w}^k} + N_b \hat{\theta}_b \log \hat{\theta}_b - N_F \hat{\theta}_F \log \hat{\theta}_F \,. \tag{5}$$

Figure 4a shows the output obtained with the conventional one-region MLRT. For the low-contrast objects, a low maximum value of the MLRT is obtained and the maximum is displaced from the true object location. The filters tend to give high output values in locations where large, high-contrast patches are found in the image. Thus, for the first three objects a double peak that corresponds to the two upper and lower halves of the object is obtained.

The output for the MLRT designed for five regions (Fig. 4b) behaves in a more robust way. It can properly detect all the instances of the object. The more-detailed structure of the five-region MLRT increases the importance of sidelobes in the output plane. As a result, the one-region filter provides for the two last objects (which are basically homogeneous) sharper correlation peaks than does the proposed new filter. Nevertheless, the five-region filter provides in all cases the location of the output peak that corresponds to the true location.

In conclusion, a new statistical filter for pattern detection that can cope with objects composed of several regions has been introduced. This new filter expands the variety of situations in which a MLRT can be successfully used. Experiments performed on both synthetic and real images, with two different noise statistics, confirm that the filter will perform even when the number of regions has been overestimated.

J. García acknowledges a grant from the Spanish Secretaría de Estado de Educación y Universidades. P. Réfrégier's e-mail address is refregie@enspm009.u-3mrs.fr.

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

References

1. G. L. Turin, IRE Trans. Inf. Theory **IT-6,** 311 (1960).
2. A. VanderLugt, IEEE Trans. Inf. Theory **IT-10,** 145 (1964).
3. P. Garcia-Martinez, H. H. Arsenault, and S. Roy, Opt. Commun. **173,** 185 (2000).
4. B. Javidi and J. Wang, J. Opt. Soc. Am. A **11,** 2604 (1994).
5. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis* (Wiley, New York, 1973).
6. F. Goudail and P. Réfrégier, Opt. Lett. **21,** 495 (1996).
7. B. Javidi, P. Réfrégier, and P. Willet, Opt. Lett. **18,** 1660 (1993).
8. F. Guérault and P. Réfrégier, Opt. Commun. **142,** 197 (1997).
9. P. H. Garthwaite, I. T. Jolliffe, and B. Jones, *Statistical Inference* (Prentice-Hall, London, 1995).
10. V. Pagé, F. Goudail, and P. Réfrégier, Opt. Lett. **24,** 1383 (1999).
11. C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images* (Artech House, Norwood, Mass., 1998).