Contrasted statistical processing algorithm for obtaining improved target detection performances in infrared cluttered environment

Zeev Zalevsky
David Mendlovic
Tel-Aviv University
Faculty of Engineering
69978 Tel-Aviv, Israel

Ehud Rivlin
Technion
Department of Computer Science
32000 Haifa, Israel

Stanley Rotman
Ben-Gurion University
Electrical Engineering Department
84105 Beer-Sheva, Israel

Abstract. A contrasted statistical processing approach to obtain improved probabilities of false alarm when performing automatic target detection is presented. The technique is based on analyzing each sector of the image and comparing it with surrounding windows in which the desired statistical property is calculated. The contrast of the statistical property is extracted using the prediction or the prediction-correction equations. The contrast of the statistical property is shown to be a good discriminator of the target from its background allowing the reduction of the detection threshold applied over the stationary region while maintaining a constant false alarm probability. [S0091-3286(00)02110-3]

Subject terms: statistical processing; automatic target recognition.

Paper 990041 received Jan. 27, 1999; revised manuscript received Mar. 16, 2000; accepted for publication May 12, 2000.

1 Introduction

Statistical processing is a well-known approach for detecting targets in a cluttered environment. In this approach, certain statistical parameters are calculated within a region of interest (ROI) and then are used to determine a detection threshold. A small sliding window, having a size similar to the size of the searched target, scans the background while the same statistical parameter is computed within the window and compared with the global threshold. If the threshold is exceeded, a target is declared to exist in the center position of that specific computation window. To keep a constant false alarm rate (CFAR) and to handle the nonstationarity of the clutter, the ROI is shifted along the scanning region and the threshold is updated according to the new ROI. Usually, the size of the ROI window is larger than the size of the target; however, it should not be too large or the nonstationarity of the background cannot be handled. Different types of statistical processing that use various statistical parameters were introduced, for instance, in Refs. 1–3. Among the statistical parameters introduced there were the Doyle, the probability of edge (POE), and the cooccurrence matrix. These metrics were ostensibly used for predictions of human target detection performance; however, they can also be used for automatic target detection.

The main problem of the statistical processing approach is the nonstationarity of the background. Even though the ROI window is a small window in comparison with the overall scan region, the background is not stationary. In addition, even if the background is more or less stationary, since it is random, the statistical parameter has a wide variety of values even within the stationarity region. We term this problem the nonlocality of the background. Several techniques tried to solve those problems by new approaches such as fuzzy logic.[4] However, those techniques involve heavy computations and do not immediately converge.

In this paper, we suggest a novel approach that we call contrasted statistical processing. Statistical processing is performed similarly to the already mentioned approach. However, instead of comparing the statistical parameter obtained within the search window with the threshold of the stationarity region, we calculate its contrast. The background is scanned with not one but nine sliding windows having sizes similar to the sizes of the target. Each time, the statistical parameter is computed in the central window as well as in the eight peripheral windows. The contrast of the statistical parameter is calculated and compared with a statistical contrast threshold. A decision regarding the existence or the absence of a target is reached. The suggested approach overcomes both the nonstationarity of the background and its nonlocality. Since the nine windows are close to one another, the influence of the background's nonstationarity or nonlocality is minor. For a given threshold determined by a desired probability of detection, smaller false alarm probabilities are obtained with the presented approach.

Section 2 presents the metrics. The locality problem is discussed in Sec. 3, computer simulations are shown in Sec. 4, and conclusions are presented in Sec. 5.

2 Metrics for Target Detection

In this paper, we address two main types of statistical processing parameters: the Doyle and the POE. However, the approach is general and can be applied for any other type of statistical processing.

The common statistical processing is the Doyle operator,[1] which computes the local differences existing between the target and its background. There are several versions for this statistical parameter:

$$S_{\text{Doyle}}=[(\mu_t-\mu_b)^2+k(\sigma_t-\sigma_b)^2]^{1/2}, \tag{1}$$

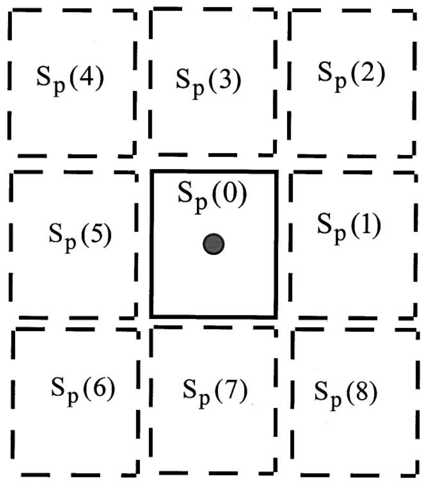

Fig. 1 Scanning processing windows.

or

$$S_{\mathrm{Doyle}}=[(\log \mu_t-\log \mu_b)^2+k(\log \sigma_t-\log \sigma_b)^2]^{1/2}, \qquad (2)$$

where μ_t and μ_b are the averages of the target and its background, σ_t and σ_b are the standard deviation of the target and its background, and k is a weighting coefficient.

Another important statistical parameter[1] is POE. This parameter correlates the input images with a high-pass filter (HPF) such as the Sobel filter and then counts the number of pixels passing a certain threshold Th:

$$S_{\mathrm{POE}}=\sum_{x,y} \mathrm{step}[f(x,y)*\mathrm{HPF}(x,y)-\mathrm{Th}], \qquad (3)$$

where * denotes a correlation operation and step(x) is the step function equal to 1 for $x>0$ and zero otherwise.

The suggested method is based upon scanning the input image with nine windows, as seen in Fig. 1. Designating the index of the central window by zero and by increasing indices of the eight peripheral windows, the suggested algorithm can be written as

$$\mathrm{SC}_p=\frac{S_p(0)-(1/8)\Sigma_{n=1}^{8}S_p(n)}{(1/8)\Sigma_{n=1}^{8}S_p(n)}, \qquad (4)$$

where SC_p is the statistical contrast of type p (the Doyle, the POE, the cooccurrence matrix, or any other parameter), and S_p is the statistical parameter of type p calculated within the n'th window. Note that the expression $(1/8)\Sigma_{n=1}^{8}S_p(n)$ is a zero-order approximation for $S_p(0)$. In this way, $S_p(0)$ is approximated by the average of its surrounding. A more general expression should be

$$\mathrm{SC}_p=\frac{S_p(0)-\hat{S}_p(0)}{\hat{S}_p(0)}, \qquad (5)$$

where $\hat{S}_p(0)$ is the approximation for $S_p(0)$ done according to the eight windows surrounding the central scanning window. Obviously, using the average as the approximation method is not necessarily the optimal approach. Instead one may derive the optimal linear prediction equations for the statistical parameter of type p. Assuming that

$$\hat{S}_p(0)=\sum_{n=1}^{8} a_n S_p(n), \qquad (6)$$

and defining the mean square criterion of optimization for the error (where ε is the notation for the error and $E\{\}$ is the ensemble average):

$$\varepsilon=E\left\{\left[S_p(0)-\sum_{n=1}^{n=8} a_n S_p(n)\right]^2\right\}\rightarrow \min \qquad (7)$$

yields the following equations set:

$$\mathbf{R}_s(n,m)\times \mathbf{a}(n)=\mathbf{r}(n), \qquad (8)$$

where $\mathbf{R}_s(n,m)$ is a matrix whose n,m element ($1\leqslant n,m \leqslant 8$) equals $E\{S_p(n)S_p(m)\}$, $\mathbf{a}(n)$ is the vector of the desired a_n coefficients, and $\mathbf{r}(n)$ is a vector whose n element equals $E\{S_p(0)S_p(n)\}$. The coefficient vector then becomes

$$\mathbf{a}(n)=\mathbf{R}_s(n,m)^{-1}\times \mathbf{r}(n). \qquad (9)$$

Note that due to the way that the S_p parameters were defined, they are not stationary. The elements of $\mathbf{R}_s(n,m)$ and of $\mathbf{r}(n)$ can be recursively calculated during the scanning process of the region of interest:

$$R_s^{(k)}(n,m)=R_s^{(k-1)}(n,m)+\beta S^{(k)}(n)S^{(k)}(m),$$
$$r^{(k)}(n)=r^{(k-1)}(n)+\beta S^{(k)}(0)S^{(k)}(n), \qquad (10)$$

where the coefficient β is a weighting coefficient determining the adaptation process convergence rate to its steady state.

Instead of deriving the optimal prediction equations, one can derive the predictor-corrector Kalman filter[5] equations, which are quite similar except that they are expressed recursively:

$$\hat{S}_p^{(k)}(0)=\hat{S}_p^{(k-1)}(0)+\sum_{n=1}^{8} a_n S_p(n), \qquad (11)$$

where $\hat{S}_p^{(k)}(0)$ and $\hat{S}_p^{(k)}(0)$ are the predictions for $S_p(0)$ in the k and $k-1$ iteration steps, respectively.

3 Locality Problem

We term the difference between the true statistical parameters of the background and the statistical parameters evaluated within the scanning window the locality problem.

To derive the first-order statistics of the statistical parameter S_p, one must know the statistics of the background itself. The IR backgrounds can usually be represented as a first-order Markov process having the autocorrelation function of[6,7]

$$E\{z_i z_{i+k}\} = \sigma_z^2 \rho^{|k|}, \tag{12}$$

where ρ is the correlation coefficient between two areas seen by two sequential instantaneous field of views (IFOV) $(0<\rho<1)$, and k is the distance between two different IFOVs. Assume for instance, that our statistical parameter is the Doyle parameter. For this case, one must find the average and the standard deviation of the pixels within the scanning window. Since the window is size-limited (local), the average and the standard deviation extracted from the window will differ from their real values in the stationarity region of the background. The true computation of the standard deviation σ_z and of the average gray level μ_z should be done by

$$\mu_z = \frac{\sum_{i=1}^{N_w} z_i}{N_w}, \qquad \sigma_z^2 = \frac{\sum_{i=1}^{N_w} (z_i - \mu_z)^2}{N_w}, \tag{13}$$

where N_w is the number of pixels in the stationary region, and z is the pixel gray-level value. Since we do not have information about the size of the stationary region, we examine the relation between the variance and the average to be obtained in the case of prediction by a smaller scanning window. We denote by $\hat{\mu}_z$ and by $\hat{\sigma}_z$ the prediction for the average and for the standard deviation respectively, using smaller window having only N pixels:

$$\hat{\mu}_z = \frac{\sum_{i=1}^{N} z_i}{N}, \qquad \hat{\sigma}_z^2 = \frac{\sum_{i=1}^{N} (z_i - \hat{\mu}_z)^2}{N}. \tag{14}$$

Note that those equations may be corrected for small N to produce unbiased estimates. This may become especially important if the nine windows are not identical in shape and size, as the estimates will have different biases. For the following analysis we will restrict ourselves to the case where N is not too small and the nine windows are identical.

Since both $\hat{\mu}_z$ and $\hat{\sigma}_z$ depend on the random variable z, we can compute their expectancy value. For pixels in the same line one can write

$$E\{\hat{\mu}_z^2\} = \frac{1}{N^2} \sum_i \sum_j E\{z_i z_j\} = \frac{1}{N^2} \sum_i \sum_j \sigma_z^2 \rho^{|j-i|} = \frac{\sigma_z^2}{N^2} \left[\sum_{k=1}^{N} (2N-2k)\rho^k - N \right]. \tag{15}$$

Using Eq. (15) one can easily obtain

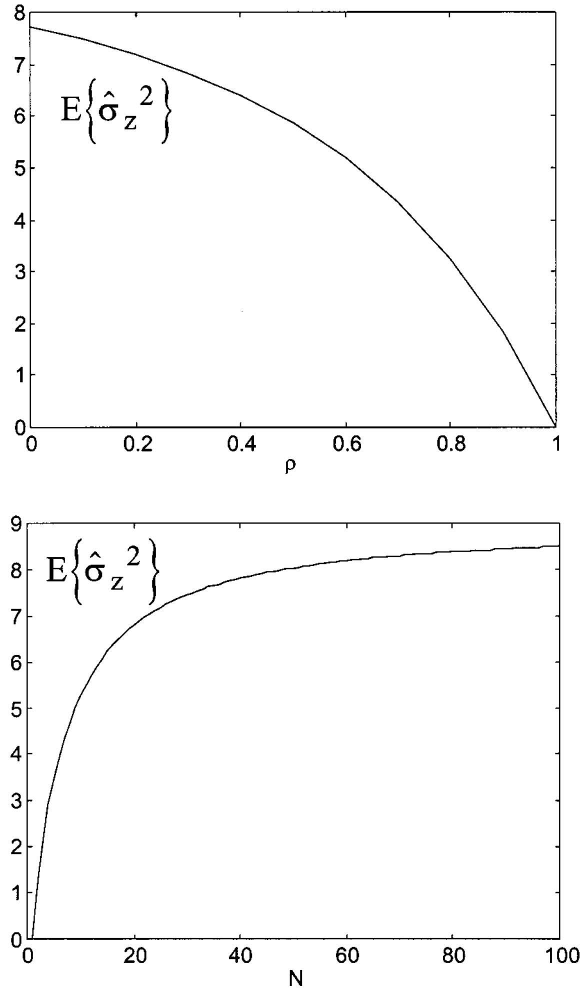

Fig. 2 Behavior of $E\{\hat{\sigma}_z^2\}$ as a function of N and ρ.

$$E\{\hat{\sigma}_z^2\} = \frac{1}{N} \sum_i (E[z_i^2] + E[\hat{\mu}_z^2] - 2E[z_i \hat{\mu}_z]) = \sigma_z^2 - \frac{\sigma_z^2}{N^2} \left[\sum_{k=1}^{N} (2N-2k)\rho^k - N \right]. \tag{16}$$

Figure 2 presents the dependence of $E\{\hat{\sigma}_z^2\}$ on N for $\rho = 0.7$ and $\sigma_z^2 = 9$ and its dependence on ρ for $N = 7$.

A trade-off exists here for the desired window size N. On one hand, one wishes to decrease N to avoid the nonstationarity of the background. On the other hand, decreasing N too much destroys the quality of prediction. One of the advantages of the contrasted approach is that it overcomes the locality problem by observing not the value of the statistical parameter itself but the difference from its surrounding.

4 Computer Simulations

In the following simulations, the suggested approach was tested. In the simulations, we assumed a scanning window having a rectangular shape with the minimal possible dimensions that yet bound the target. The performances of the suggested technique with windows of varied dimen-

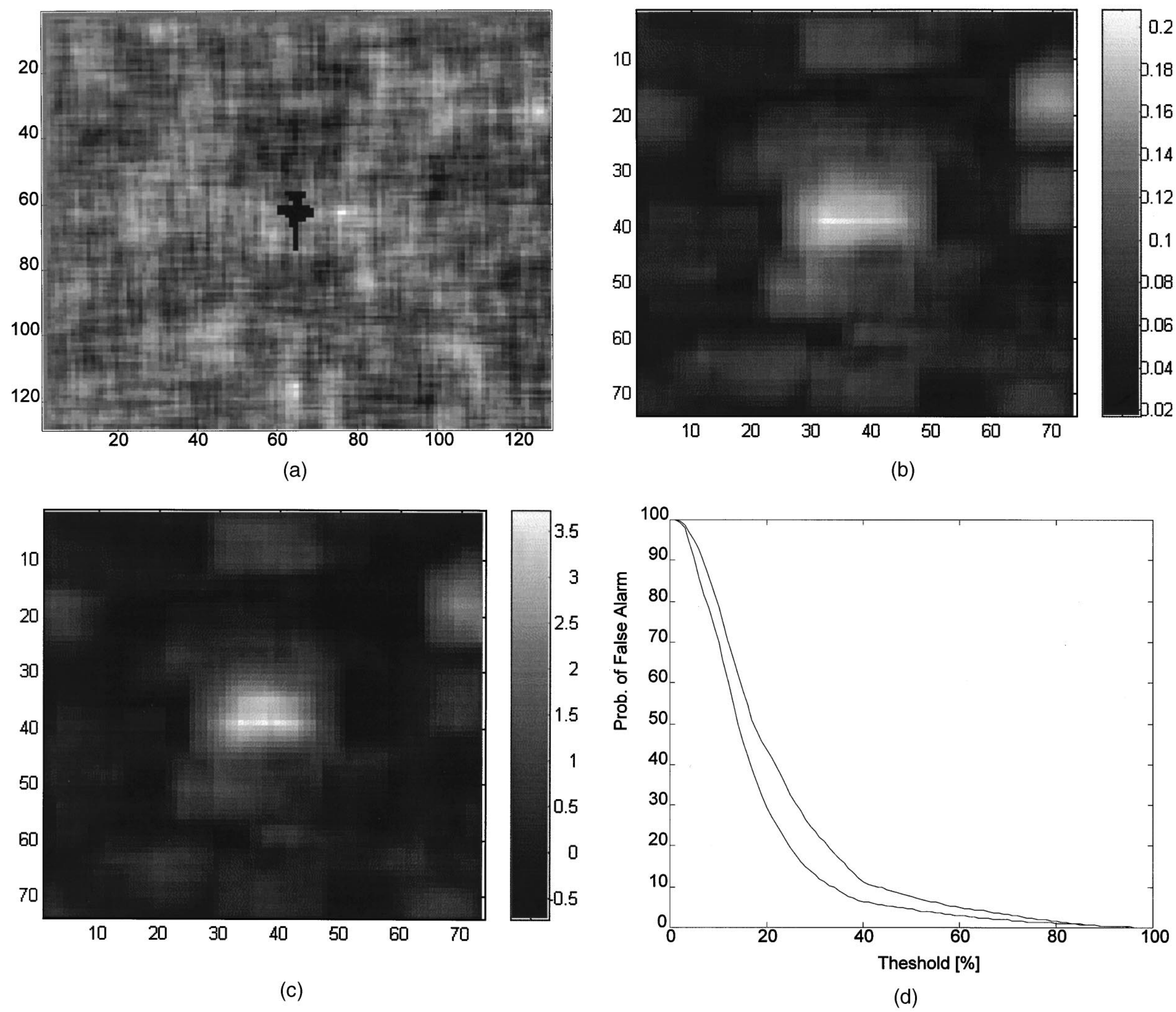

Fig. 3 Results obtained for a stationary image with a Doyle-like statistical processing type: (a) input pattern, (b) output obtained after applying a conventional Doyle like processing, (c) output obtained after applying the contrasted approach, and (d) plot of the probability of false alarm as function of the applied threshold.

sions were not tested. In addition, we assumed that the range between the sensor and the various parts of the scenes is the same.

To simulate the presented approach, an IR background was synthesized using the first-order statistical Markov model with various ρ and σ_z^2 parameters. Iterative equations were used[8] to fulfill Eq. (12):

$$z(i,j)=\rho z(i-1,j)+\rho z(i,j-1)-\rho^2 z(i-1,j-1)+w(i,j),$$

where $z(i,j)$ is the simulated IR background, ρ is the correlation coefficient, and $w(i,j)$ is a 2-D sequence of independent and identically distributed (i.i.d.) zero-mean Gaussian variables with common variance of

$$\sigma_w^2=\sigma^2(1-\rho^2)^2,$$

where σ is the standard deviation of the generated image $z(i,j)$.

Figure 3(a) presents a stationary background with an airplane target inserted within it. A processing window having a size of 15 pixels was slid over the input image. The statistical processing that we have investigated first was a Doyle-like parameter. Since the target had low gray levels (in comparison to the gray levels of the background) and the processing window is a bit longer than target's dimensions, the following statistical processing will result in a high value for the target and in lower values for the background:

$$S_p=\frac{\hat{\mu}_z}{\hat{\sigma}_z}=\frac{\sum_{n=1}^N z_n}{[\sum_{n=1}^N (z_n-\sum_{n=1}^N z_n)^2]^{1/2}}.$$

Figure 3(b) presents the image obtained after applying this statistical processing. In Fig. 3(c), one can see the image obtained after applying the contrasted statistical processing.

One can see that in both Figs. 3(b) and 3(c) high values were obtained in the center, which corresponds to the location of the target. Figure 3(d) presents a plot of the probability of false alarm as function of the threshold level in percentages of the difference between the maximum and the minimum obtained values. The upper curve corresponds to the probabilities obtained for the conventional statistical processing and the lower curve corresponds to the contrasted statistical processing. These probabilities were obtained from Figs. 3(b) and 3(c) by applying a threshold and counting the number of pixels passing this threshold value. One can see the improvement obtained in using the contrasted approach.

Figure 4(a) presents an airplane placed on a center of a nonstationary background generated by a linear combination of several Markov processes. Once again, the Doyle-like approach was applied. The results obtained in Fig. 4(b) correspond to conventional statistical processing. The higher values are located in the center, which corresponds to the position of the target. For this case, a threshold of 13% of the difference between the obtained maximal and minimal values is required in order to detect the target. For this threshold the false alarm probability is 13.14%. Figure 4(c) presents the contrasted statistical processing which results in a 5.2% probability of false alarm, with a threshold that is 42% of the difference between the maximal and the minimal values. This threshold is required to obtain the detection of the target. Obviously here as well, higher values are obtained in the center, which corresponds to the location of the target. Indeed, the results are much improved since lower probabilities of false alarm are obtained for thresholds required for detection.

In Fig. 5, the POE statistical processing was applied over the nonstationary background of Fig. 4(a). The threshold Th of Eq. (3) for obtaining the edge image was 70% of the average level in the central processing window (the zero number window). The image obtained by the conventional statistical processing is presented in Fig. 5(a), and the image obtained by the contrasted processing is seen in Fig. 5(b). In both, the higher values are seen in the center, which correspond to the location of the target. The curve expressing the probability of false alarm as a function of the threshold (percentages of the difference between the maximal and the minimal obtained values) applied over the processed image is seen in Fig. 5(c). The upper curve corresponds to the conventional processing approach and the lower curve corresponds to the contrasted processing. One can see the improvement—for similar threshold values much lower probabilities of false alarm are obtained.

Figures 6 and 7 present the potential of the suggested approach with real IR (3 to 5-μm spectral band) images. Figure 6(a) presents an IR image containing a target in the center of the image. A Doyle-like processing was applied. Figure 6(b) is the result of conventional processing and Fig. 6(c) of the contrasted processing. Figure 6(d) once again presents the false alarm probability as function of the applied threshold, in percentages of the difference between the maximal and the minimal values. The upper curve in

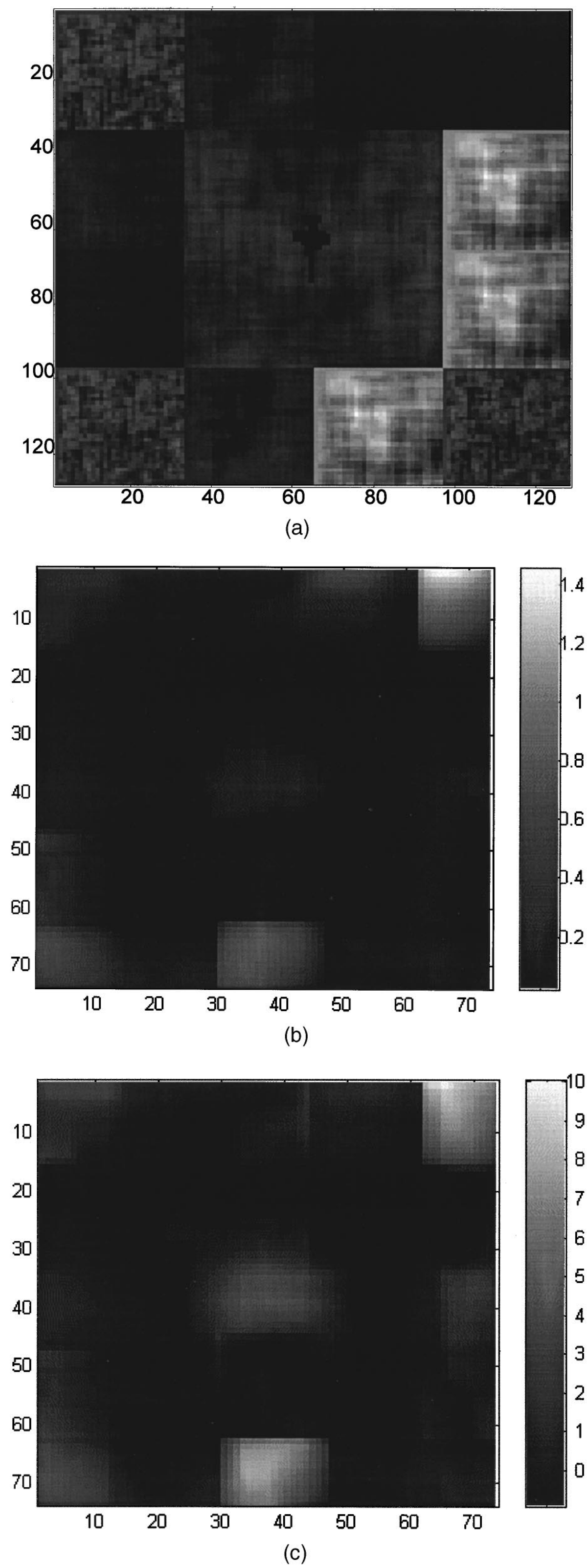

Fig. 4 Results obtained for a nonstationary image with a Doyle-like statistical processing type: (a) input pattern, (b) output obtained after applying a conventional Doyle-like processing, and (c) output obtained after applying the contrasted approach.

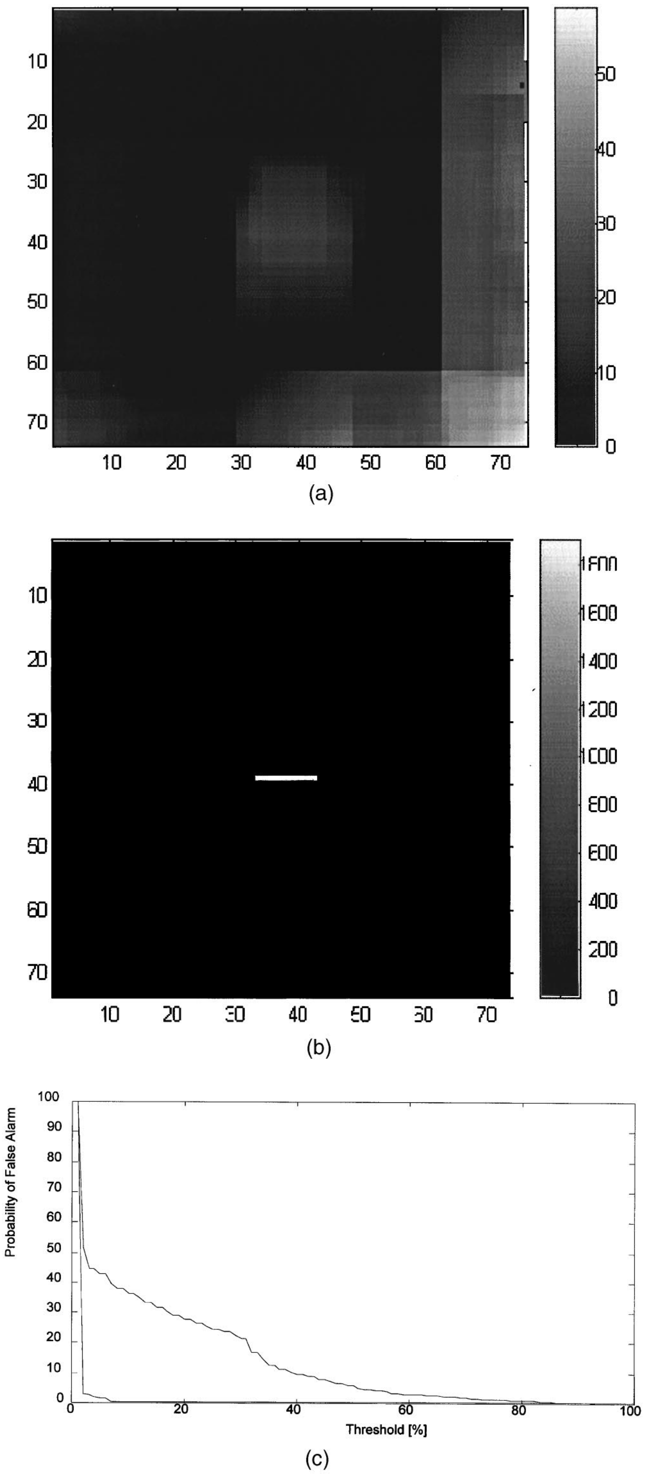

Fig. 5 Results obtained for a nonstationary image with a POE statistical processing type: (a) output obtained after applying a conventional POE processing, (b) output obtained after applying the contrasted approach, and (c) plot of the probability of false alarm as function of the applied threshold.

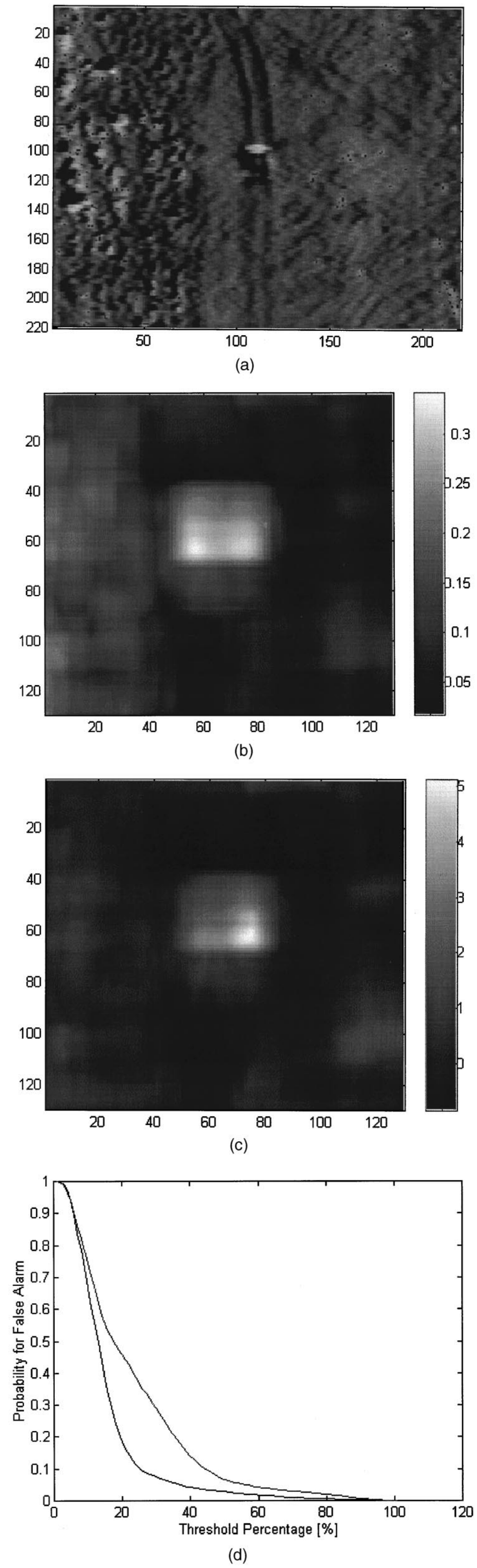

Fig. 6 Results obtained for a real IR image with a Doyle-like statistical processing type: (a) input pattern, (b) output obtained after applying a conventional Doyle-like processing, (c) output obtained after applying the contrasted approach, and (d) plot of the probability of false alarm as function of the applied threshold.

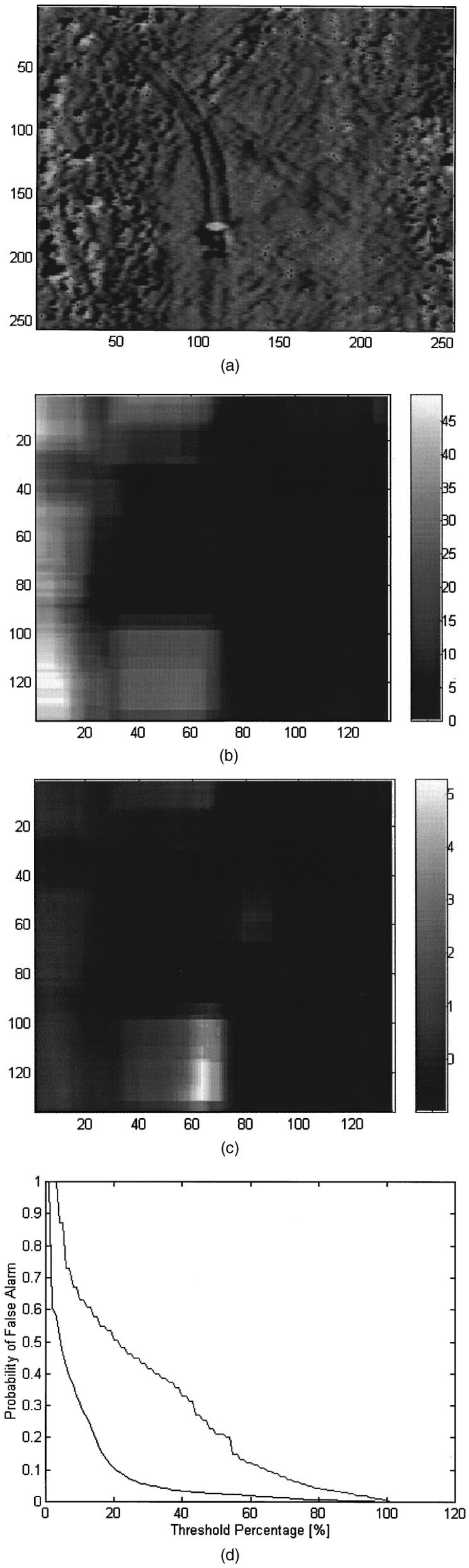

Fig. 7 Results obtained for a real IR image with a POE statistical processing type: (a) input pattern, (b) output obtained after applying a conventional POE processing, (c) output obtained after applying the contrasted approach, and (d) plot of the probability of false alarm as function of the applied threshold.

Fig. 6(d) corresponds to the conventional approach and the lower curve to the contrasted approach. One can see the improvement.

Figure 7(a) presents similar image while the target is located in the lower middle region of the input scene. A POE processing was applied. Figure 7(b) presents the results obtained using the conventional approach. Figure 7(c) corresponds to the contrasted approach. One can see that much higher values were obtained using the contrasted approach in the region where the target is supposed to be located. Figure 7(d) presents a plot of the probability of false alarm as function of the threshold in percentages of the difference between the maximal and the minimal values. The upper curve in Fig. 7(d) corresponds to the conventional approach and the lower curve to the contrasted approach. Once again, a significant improvement is revealed—much lower probabilities of false alarm are obtained for equal thresholds.

To avoid confusion, let us note that the presented false alarm methodology does not particularly indicate the decrease of detection clusters but rather the decrease in their sizes. In many automated detection configurations where computation complexity plays a major role, the clustering operation is skipped and then the detection is done, as presented, per pixel in the output plane. However, in more redundant applications, the number of clusters, rather than their dimensions, is an important parameter. For instance, returning again to Fig. 7 and applying a threshold which equals $\min + 0.65^*(\max - \min)$, where min and max are the minimal and the maximal values of the output plane respectively. This results in one cluster in the suggested approach, which appears on the location of the target, and three clusters in the conventional approach, which does not appear at the correct position of the target. Note that since in the conventional approach more false pixels pass the threshold, they merge a cluster whose location does not coincide with the position of the true target.

This is not the case of Fig. 6. Here for the same threshold a single cluster appears in both cases. In the conventional case, the cluster is only bigger.

To further justify the suggested approach, additional test scenes were input. Figure 8(a) presents the input. A Doyle-like processing was applied with a processing window size of 30 pixels. The output obtained after applying the conventional approach is presented in Fig. 8(b) and the contrasted technique may be seen in Fig. 8(c). After applying the same 65% threshold the number of clusters in the conventional approach was 14 while in the contrasted approach it was 8. In addition, the clusters in the contrasted approach were smaller. Similar processing was applied to the scene of Fig. 9. In this case, the conventional processing resulted with 7 clusters while the contrasted approach yields only 3. Following the same path in Fig. 10 yields 3 clusters for the conventional approach and 2 smaller clusters for the contrasted.

A POE processing with Th of 70% was applied over the scene of Fig. 9(a) and resulted in the conventional and contrasted outputs presented in Figs. 11(a) and 11(b) respectively. Applying a threshold equal to $\min + 0.65^*(\max - \min)$ yields 4 and 2 clusters in the conventional and contrasted approaches, respectively.

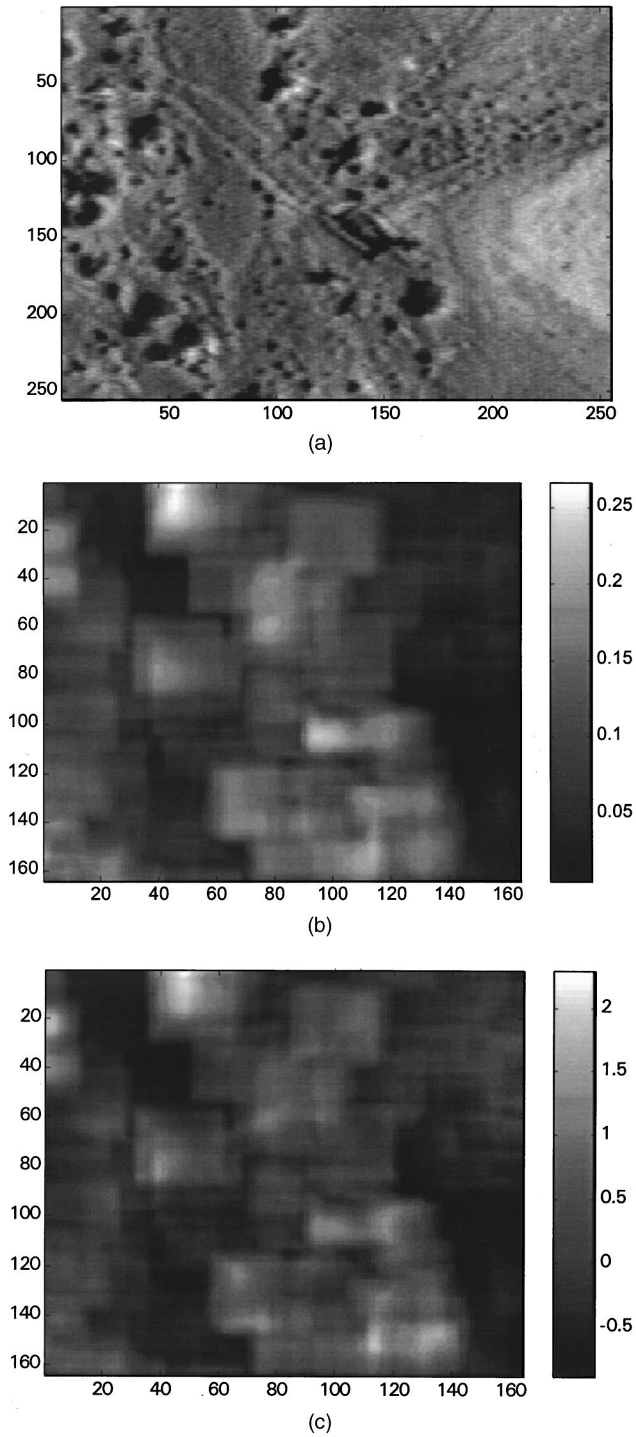

Fig. 8 Results obtained for a real IR image with a Doyle processing type: (a) input pattern, (b) output obtained after applying a conventional processing, and (c) output obtained after applying the contrasted approach.

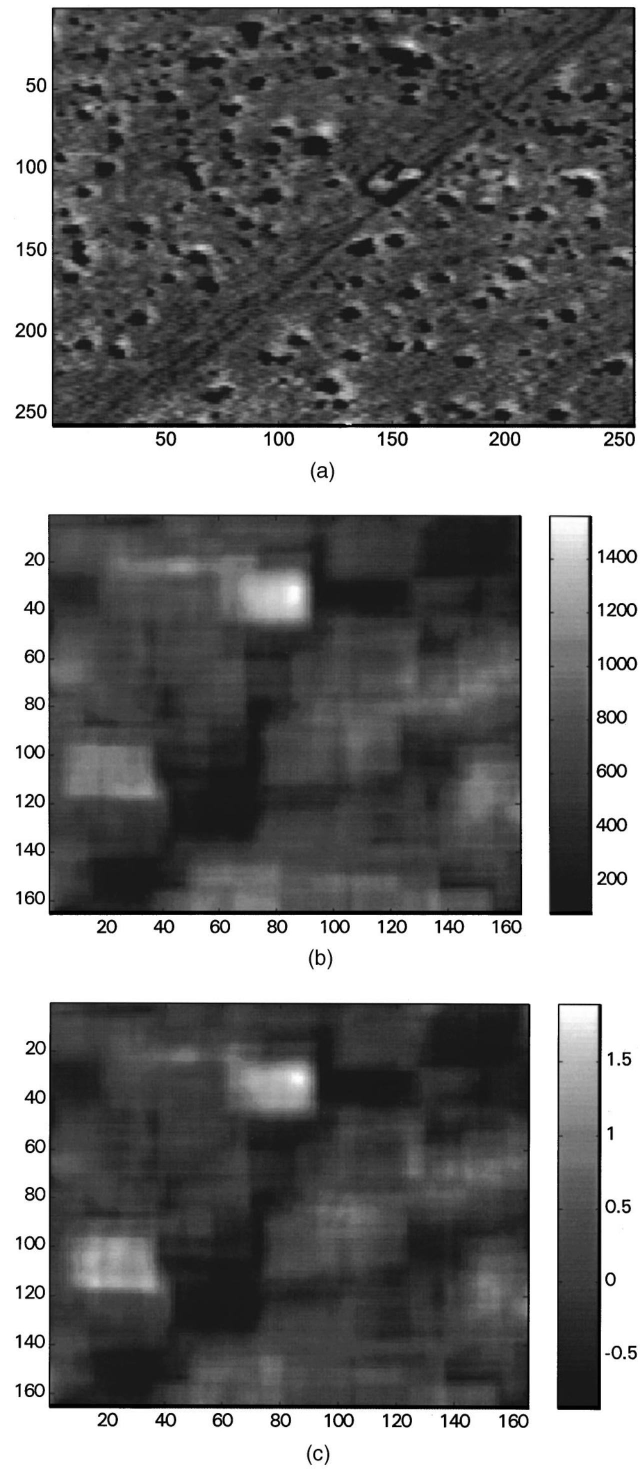

Fig. 9 Results obtained for a real IR image with a Doyle processing type: (a) input scene, (b) output obtained after applying a conventional approach, and (c) output obtained after applying the contrasted approach.

5 Conclusions

This paper presented a new approach based on a contrast computation of a desired statistical property. Prediction and a prediction-correction (Kalman) equations were applied over the calculated statistical property, while the statistical contrast was evaluated. The suggested approach enabled an improved probability of false alarm since it further enhanced the relationship between the statistical property obtained in the processing windows of the target and the background, respectively. This overcomes both the nonstationary and the locality problem of IR backgrounds exposed to a statistical processing type algorithm. Computer

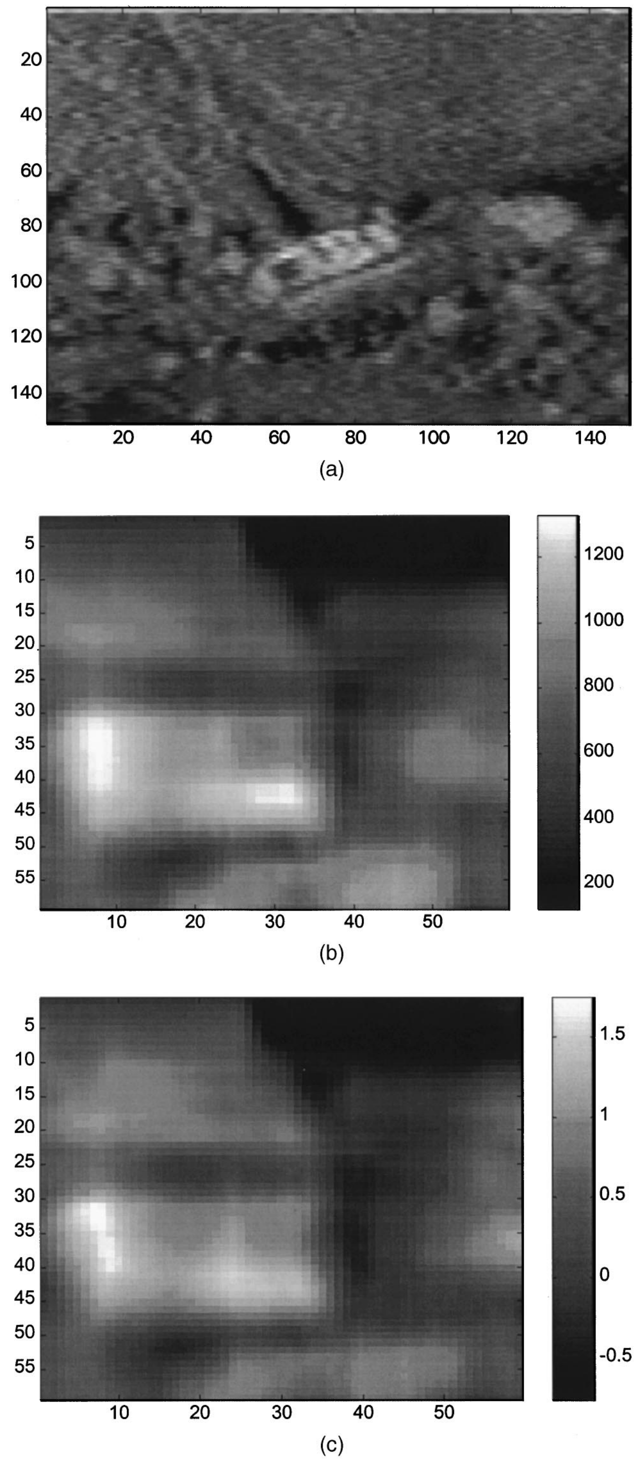

Fig. 10 Results obtained for a real IR image with a Doyle statistical processing type: (a) input pattern, (b) output obtained after applying a conventional processing, and (c) output obtained after applying the contrasted approach.

simulations demonstrated the capabilities of the suggested approach on synthesized as well as real backgrounds. The obtained results were compared favorably with the conventional statistical processing approach.

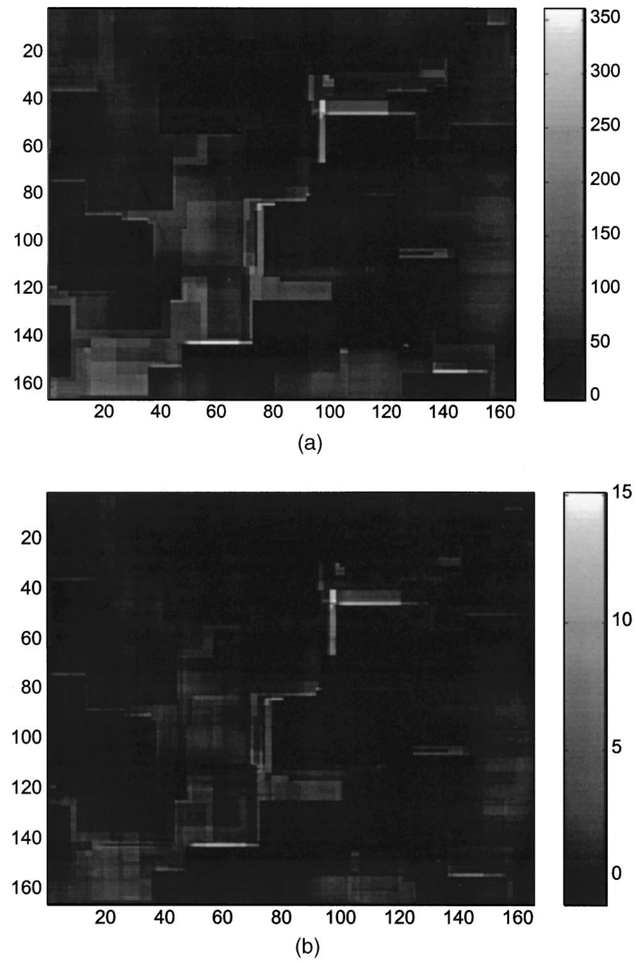

Fig. 11 Results obtained for a real IR image with a POE statistical processing type: (a) output obtained after applying a conventional POE processing and (b) output obtained after applying the contrasted approach.

References

1. S. R. Rotman, G. Tidhar, and M. L. Kowalezky, ''Modeling human search and target acquisition performance: IV. Detection probability in the cluttered environment,'' *Opt. Eng.* **33**, 801–808 (1994); S. R. Rotman, D. Hsu, A. Cohen, D. Shamay, and M. L. Kowalczyk, ''Texture metrics for clutter affecting human target acquisition,'' *Infrared Phys. Technol.* **37**, 667–674 (1996).
2. P. C. Chen and T. Pavlidis, ''Segmentation by texture using a co-occurrence matrix and a split-and-merge algorithm,'' *Comput. Graph. Image Process.* **10**, 172–182 (1979).
3. J. Parkkinen, K. Selkainaho, and E. Oja, ''Detecting texture periodicity from the co-occurrence matrix,'' *Pattern Recogn. Lett.* **11**, 43–50 (1990).
4. T. J. Meitzler, S. Harpreet, A. Labib, S. Euijung, and G. R. Grant, ''Predicting the probability of target detection in static infrared and visual scenes using the fuzzy logic approach,'' *Opt. Eng.* **37**, 10–17 (1998).
5. H. L. Van Trees, ''Detection estimation and modulation theory, part I,'' John Wiley & Sons, New York (1968).
6. Y. Itakura, S. Tsutsumi, and T. Takagi, ''Statistical properties of the background noise for the atmospheric windows in the intermediate infrared region,'' *Infrared Phys.* **14**, 17–29 (1974).
7. G. Tidhar and Z. Zalevsky, ''Performance evaluation methodology of missile warning and IRST systems,'' in *Infrared Technology and Applications XXIV, Proc. SPIE* **3436**, 891–902 (1998).
8. J. W. Modestino and R. W. Fries, ''Edge detection in noisy images using recursive digital filtering,'' *Comput. Graph. Image Process.* **6**, 409–433 (1977).