Correlation postprocessing-based method for the detection of defocused images

Arturo Carnicer, Santiago Vallmitjana, and Ignacio Juvells

We present a method to detect patterns in defocused scenes by means of a joint transform correlator. We describe analytically the correlation plane, and we also introduce an original procedure to recognize the target by postprocessing the correlation plane. The performance of the methodology when the defocused images are corrupted by additive noise is also considered. © 1997 Optical Society of America

Key words: Binary joint transform correlation, optical transfer function.

1. Introduction

In a pattern-recognition process, when the scene is recorded by means of an imaging system, the image can be affected by different kinds of degradation, such as distortion, noise, defocusing, atmospheric turbulence, or relative movement between the scene and the recording system. These degradations affect the behavior of the recognition methods, yielding a lower detection capability and reduced discrimination between similar objects. However, detection difficulties derived from problems related to defocusing have not been studied in depth. These degradations can be present in the stored scene when the optical system works in low-illumination conditions. In this case the aperture of the imaging lenses should be opened, which implies a loss of depth of field.

Recently, the performance of a binary joint transform correlator (BJTC) with defocused scenes with different thresholding methods was described.[1,2] In this paper we present an original method to detect patterns in defocused images by means of a BJTC. This procedure is valid for any blurring that presents circular symmetry. This approach postprocesses the correlation to obtain reliable detections. To show the robustness of the method, it has been tested with defocused scenes that are also corrupted by Gaussian zero-mean additive noise.

2. Joint Transform Correlation Review

Joint transform correlators[3] have been shown to be powerful coherent processors for optical pattern recognition.[4] Since 1966 several improvements in this architecture have been proposed to obtain real-time detection with liquid-crystal devices.[5,6] During the past few years most authors of papers published in the pattern-recognition field have analyzed different systems to increase discrimination capability. In particular, the binarization of the joint power spectrum (JPS) has been used widely and has been shown to be a reliable method.[7] The JPS is described by the following equation:

$$I(u, v) = |F_R(u, v)|^2 + |F_S(u, v)|^2$$
$$+ 2|F_R(u, v)||F_S(u, v)|\cos[x_0 u + y_0 v$$
$$+ \phi_S(u, v) - \phi_R(u, v)], \tag{1}$$

where $|F_R(u, v)|\exp[i\phi_R(u, v)]$ and $|F_S(u, v)|\exp[i\phi_S(u, v)]$ are the Fourier transforms (FT's) of the reference $f_R(x, y)$ and the scene $f_S(x, y)$, respectively. We assume that the scene is located at $(x_0/2, y_0/2)$ and the reference at $(-x_0/2, -y_0/2)$. The binary JPS $I_b(u, v)$ is obtained by assignment of the values $+1$ or -1 to $I(u, v)$, in the manner of Eq. (2):

$$I_b(u, v) = \begin{cases} 1 & \text{if } I(u, v) \geq I_T(u, v) \\ -1 & \text{if } I(u, v) < I_T(u, v) \end{cases}, \tag{2}$$

where $I_T(u, v)$ is a predetermined threshold function. This bipolar function can be expressed as a Fourier expansion[7]:

$$I_b(u, v) = \sum_{n=1}^{\infty} A_n[u, v; I_T(u, v)]\cos\{n[x_0 u + y_0 v$$
$$+ \phi_S(u, v) - \phi_R(u, v)]\}. \tag{3}$$

An appropriate threshold function $I_T(u, v)$ is presented in Eq. (4)[8]:

$$I_T(u, v) = |F_R(u, v)|^2 + |F_S(u, v)|^2. \quad (4)$$

To remove the intraclass terms of Eq. (1), we can evaluate separately the intensity of the FT of the scene and the reference, and then obtain by computer the subtraction between Eqs. (1) and (4). Now Eq. (3) becomes

$$I_b(u, v) = \sum_{n_{\text{odd}}} \frac{1}{n} \cos\{n[x_0 u + y_0 v + \phi_S(u, v) \\ - \phi_R(u, v)]\}. \quad (5)$$

Fourier transforming Eq. (5) and taking $n = 1$, we obtain the correlation:

$$c(x, y) = \delta(x - x_0, y - y_0) + \delta(x + x_0, y + y_0). \quad (6)$$

3. Defocused-Image Joint Transform Correlation Analysis

The intensity of a scene acquired by means of an optical system can be described by $f_S(x, y) = f_I(x, y) * h(x, y)$, where $f_S(x, y)$ is the registered image, $f_I(x, y)$ is the ideal image, $h(x, y)$ is the point-spread function (PSF) of the optical system, and the asterisk represents the convolution operator. This assumption is valid if we suppose that an imaging system is space invariant. By Fourier-transforming the former expression we obtain $F_S(u, v) = F_I(u, v)H(u, v)$, where $F_S(u, v)$ and $F_I(u, v)$ are the respective FT's of $f_S(x, y)$ and $f_I(x, y)$, and $H(u, v)$ is the optical transfer function (OTF). In an on-focus aberration-free system, if we ignore diffraction effects, it is obvious that $h(x, y) = \delta(x, y)$. For a defocused optical system the JPS is described by

$$I(u, v) = |F_R(u, v)|^2 + |F_I(u, v)H(u, v)|^2 \\ + 2 \, \text{sign}[H(u, v)]|F_R(u, v)||F_S(u, v)| \\ \times \cos[x_0 u + y_0 v + \phi_S(u, v) - \phi_R(u, v)], \quad (7)$$

where $\text{sign}(x)$ is a function defined as

$$\text{sign}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}. \quad (8)$$

Note that, if the system is affected by only defocusing, the OTF is real. Binarizing the JPS with the threshold function of Eq. (4) and taking the first term we obtain

$$I_b(u, v) = \text{sign}[H(u, v)]\cos[x_0 u + y_0 v + \phi_S(u, v) \\ - \phi_R(u, v)], \quad (9)$$

and finally the correlation is

$$\text{FT}\{\text{sign}[H(u, v)]\} \\ * \text{FT}\{\cos[x_0 u + y_0 v + \phi_S(u, v) - \phi_R(u, v)]\}. \quad (10)$$

It is the term $\text{FT}\{\text{sign}[H(u, v)]\}$ that is responsible for the degradation of the correlation. If the elimination of this term were possible, the correlation process would be successful.

In this research we obtained synthetic defocused scenes by convolving an on-focus image with a PSF. Stockseth[9] derived an analytic approximation of the OTF specially suitable for moderate and large amounts of defocusing. The approximation has the form

$$H(s) = 2(1 - 0.69s + 0.0076s^2 \\ + 0.043s^3) \frac{J_1(a - 0.5as)}{a - 0.5as}, \quad (11)$$

where $a = 4\pi ws/\lambda$, and s is the reduced spatial frequency that varies in the range $0 \le s \le 2$. For $|s| > 2$ Eq. (11) is equal to zero and $s = 2$ corresponds to the cutoff frequency of the optical system. The parameter w is the optical distance between the optical system's emergent wave front and the reference sphere centered on the axis at the out-of-focus point, measured along the extreme pupil ray. Equation (11) is easier to calculate, and for $w > 1$ its use is fairly accurate. We use this equation to simulate different degrees of defocusing in realistic conditions, as described earlier in this section.

As is shown in Eq. (10) the form of the correlation is described by $\text{FT}\{\text{sign}[H(u, v)]\}$. Notice that in this case $H(u, v) = H(\rho)$ presents circular symmetry, so $\text{sign}[H(\rho)]$ can be described in terms of $\text{circ}(\rho)$ functions and the zeros of $H(\rho)$. The $\text{circ}(x, y)$ distribution is defined as

$$\text{circ}(u, v) = \text{circ}(\rho) = \begin{cases} 1 & \text{if } \rho = (u^2 + v^2)^{1/2} \le 1 \\ 0 & \text{if } \rho = (u^2 + v^2)^{1/2} > 1 \end{cases}. \quad (12)$$

Let $R_1 < R_2 < R_3 < \cdots < R_k$ be the values where $H(\rho) = 0$. Note that the number of poles is finite owing to the existence of the cutoff frequency in the Fourier domain. Then we can write $\text{sign}[H(\rho)]$ as an expansion as

$$\text{sign}[H(\rho)] = 2 \, \text{circ}\left(\frac{\rho}{R_1}\right) - 2 \, \text{circ}\left(\frac{\rho}{R_2}\right) \\ + 2 \, \text{circ}\left(\frac{\rho}{R_3}\right) - \cdots + (-1)^{k+1} \, \text{circ}\left(\frac{\rho}{R_k}\right), \quad (13)$$

by Fourier-transforming the latter expression:

$$\text{FT}\{\text{sign}[H(\rho)]\} = 2 \frac{J_1(2\pi R_1 r)}{2\pi R_1 r} - 2 \frac{J_1(2\pi R_2 r)}{2\pi R_2 r} \\ + 2 \frac{J_1(2\pi R_3 r)}{2\pi R_3 r} - \cdots \\ + (-1)^{k+1} \frac{J_1(2\pi R_k r)}{2\pi R_k r}, \quad (14)$$
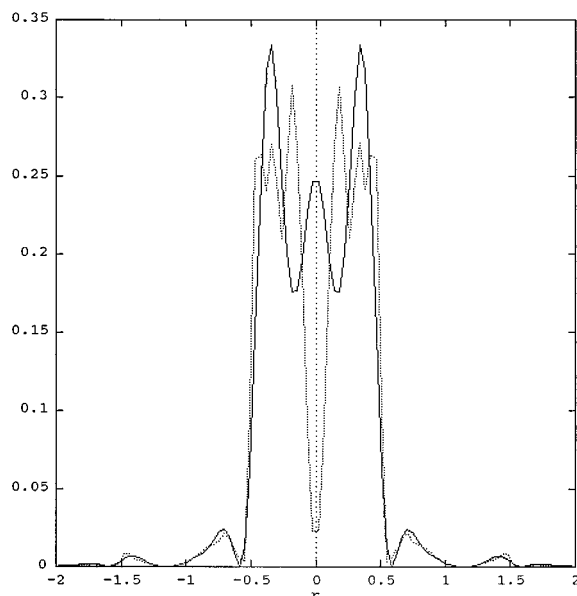
Fig. 1. Section of the correlation when using defocused scenes. Solid curve: $k = 3$; dotted curve: $k = 10$. The vertical coordinate is in arbitrary units.

where $r = (x^2 + y^2)^{1/2}$. Finally, by combining Eqs. (6), (10), and (14) and assuming a diffractometer constant $\lambda f'$, we can describe the correlation plane by

$$c(x, y) \propto \left| \frac{1}{\pi r} \left[\sum_{i=1}^{k-1} \frac{(-1)^{i+1}}{R_i} J_i \left(\frac{2\pi R_i r}{\lambda f'} \right) \right. \right.$$
$$\left. \left. + \frac{(-1)^{k+1}}{2R_k} J_1 \left(\frac{2\pi R_k r}{\lambda f'} \right) \right] \right|^2 * [\delta(x - x_0, y - y_0)$$
$$+ \delta(x + x_0, y + y_0)]. \quad (15)$$

To illustrate the performance of the correlation described by Eq. (15), we computed a particular case considering $R_i = i$ and $\lambda f' = 1$. Because Eq. (11) is nearly periodic, the assumption of $R_i = i$ seems suitable. And because of a severe difficulty in dealing with the JPS with defocused images—the dynamic range of current detector arrays is limited to two or three decades and at least 3 orders of magnitude are required to cover the intensity of the first zeros of the OTF—two situations are considered. Fig. 1 shows a plot of the correlation section with Eq. (15). The solid and dashed plots correspond to set $k = 3$ and set $k = 10$, respectively, illustrating a comparison between two very different situations of recording the JPS. In both cases the correlation section is confined around the detection point.

4. Results: Correlation Postprocessing

To obtain defocused images we calculated the convolution between the scene depicted in Fig. 2 and a PSF we obtained starting from the OTF pointed out in Eq. (11). We applied this equation for simulating realistic conditions: the imaging system is a standard photographic objective of $f' = 50$ mm, working with an aperture of $f/11$ and wavelength of $\lambda = 600$ nm.
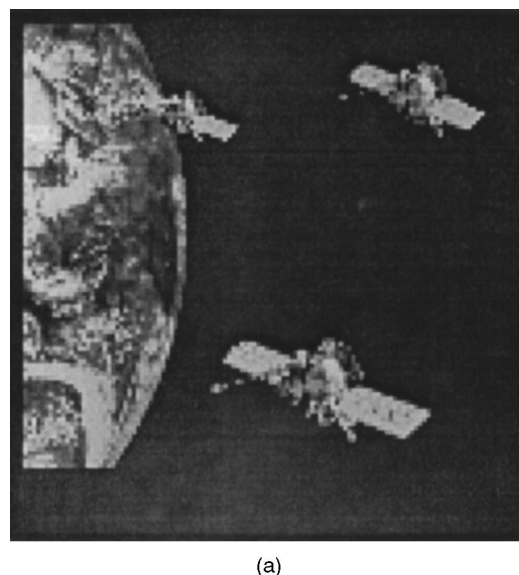

(a)
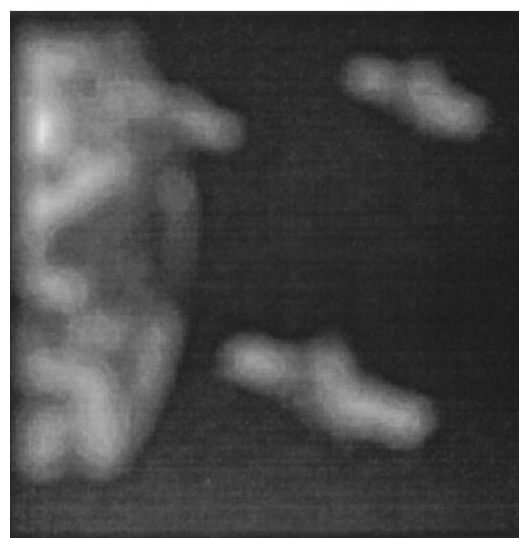
(b)

Fig. 2. (a) Scene used as a test. (b) Scene affected by a defocusing of 6λ.

The cut-off frequency ($s = 2$) is approximately 300 line pairs/mm (lp/mm). The image is registered by means of a commercial CCD with 512×512 pixels and a size of 8 mm × 6 mm. The effective resolution of the electronic device is more than 30 mm^{-1}.

Figure 2(a) shows a scene used in the simulations, and Fig. 2(b) shows the same scene affected by severe defocusing of $w = 6\lambda$. Figure 3(a) shows the BJTC between the defocused scene and the larger satellite. We produced Fig. 3(b) by taking the smaller satellite as a reference. As described in Section 3, a circular-symmetry distribution appears centered at the detection point and, in both cases, detection is almost impossible. Nevertheless, in Fig. 3(b), the circular-symmetry distribution is not as evident, owing to the presence of background clutter in the neighborhood of the target.

To deal with this correlation and to determine with precision the position of the target, we propose the
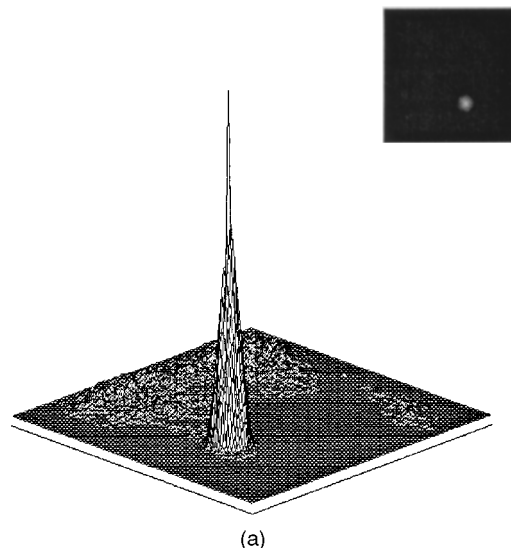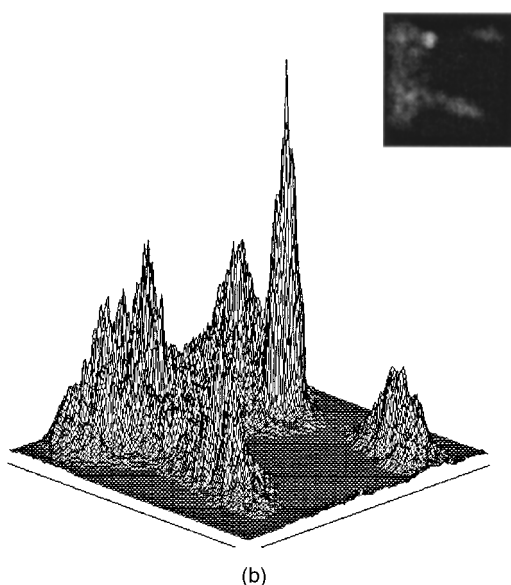
(a)

(a)

(b)
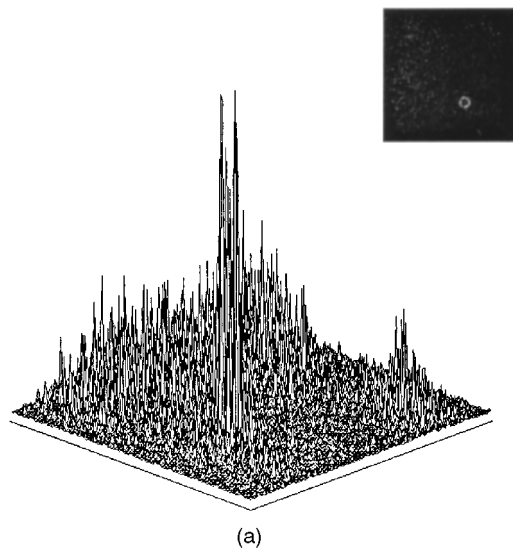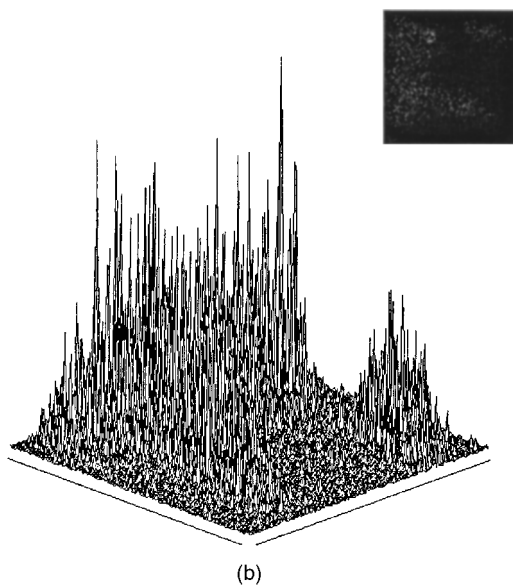
(b)

Fig. 3. (a) BJTC between scene of Fig. 2(b) and the larger satellite. (b) BJTC between scene of Fig. 2(b) and the smaller satellite.

Fig. 4. (a) Postprocessed correlation of Fig. 3(a). (b) Postprocessed correlation of Fig. 3(b).

following procedure based on a postprocessing of the correlation plane: perform a local convolution between the correlation plane and a circle of suitable radius. We can evaluate a proper radius by considering the first zero R_1 in the degraded JPS. This decision is based in the discussion regarding the plots of Fig. 1. By using the Airy formula we can generate a circle related to the real defocusing of the system.

Finally, Figs. 4(a) and 4(b) show the simulated crosscorrelation between the results presented in Figs. 3(a) and 3(b) and a circle obtained from the first zero of the degraded JPS. In both cases the procedure is successful.

5. Detection of Defocused Scenes in Presence of Noise

To study the robustness of the method, a numerical simulation that involves defocused scenes corrupted by Gaussian zero-mean additive noise has been carried out. To test the postprocessing we took into account several defocused versions of the larger satellite, from $w = 2\lambda$ to $w = 6\lambda$. The variance of the noise added to the image varies from $\sigma^2 = 10$–10,000. To establish a ratio between the signal and the noise, we considered the input signal-to-noise ratio (SNR) we obtained by computing the square of the average of the pixel of the scene between the noise variance. Fifty noise realizations were generated for each image to secure reliable results. Two metrics to assess the quality of the correlation were considered: the SNR and the detection probability.

Figure 5(a) shows the evolution of the correlation SNR as a function of the input SNR (SNR_i). As we expected, the plot shows a decreasing trend. For small values of noise (high SNR_i) and depending on the defocusing degree, the values of the SNR vary
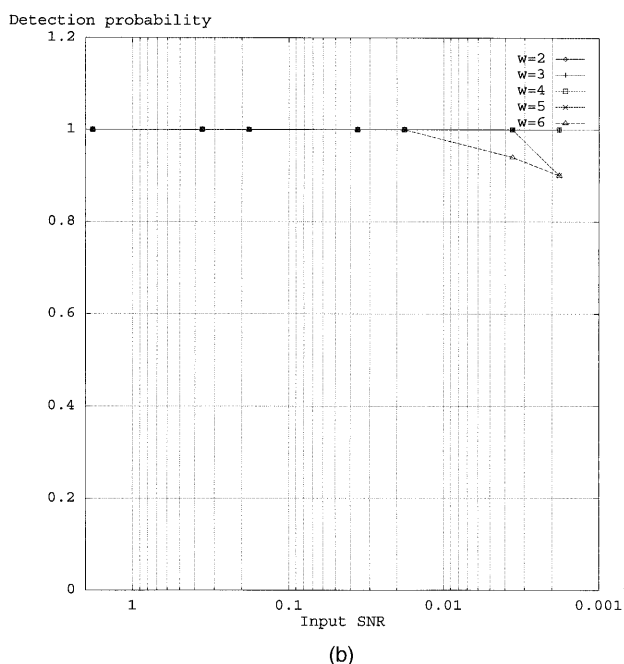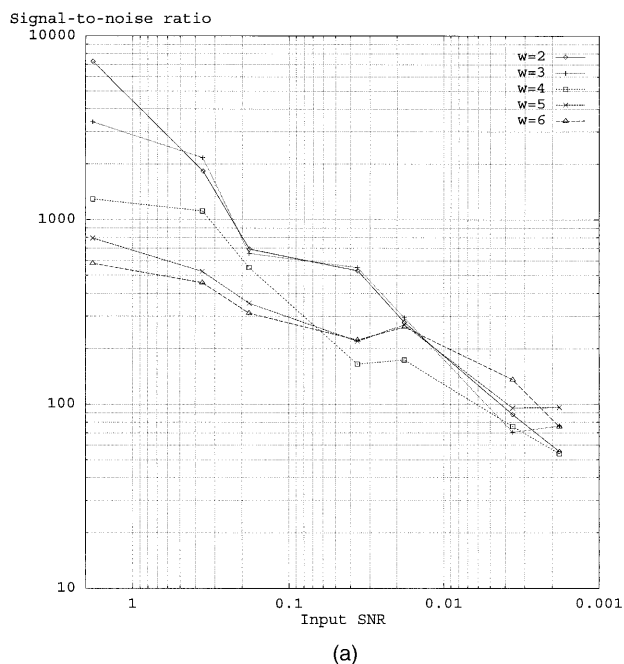
at a distance above 1 pixel of the true position are considered errors. For small and medium defocused inputs ($w < 5\lambda$), the detection is always correct with independence of the noise. For $w = 5\lambda$ or $w = 6\lambda$, allocation errors sometimes appear for $\mathrm{SNR}_i < 0.01$. Moreover, no false alarms due to sidelobes far from the detection position have been detected in all the simulations carried out. In conclusion, these results show the robustness of the method.

6. Conclusion

A procedure for the recognition of objects in defocused scenes by means of a BJTC has been demonstrated. The correlation is analytically described and a method for solving the problem is studied. The method consists of convolving the crosscorrelation term with a circle of a suitable radius. This is obtained starting from the first zero of the defocused JPS.

To test the method when a defocused scene is corrupted by additive noise, a numerical simulation is carried out. The analysis of the SNR and, in particular, the study of the allocation errors as a function of the noise in the signal show that reliable detections are obtained even when a defocused scene is corrupted with additive noise.

References

1. B. K. Karch, M. A. Karim, and D. L. Flannery, "Binary joint transform correlation of defocused images," Opt. Eng. **32,** 2709–2719 (1993).
2. A. Carnicer, S. Vallmitjana, J. R. de F. Moneo, and I. Juvells, "Implementation of an algorithm for detecting patterns in defocused scenes using binary joint transform correlation," Opt. Commun. **130,** 327–336 (1996).
3. C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," Appl. Opt. **5,** 1248–1249 (1966).
4. F. T. S. Yu, Q. W. Song, Y. S. Cheng, and D. A. Gregory, "Comparison of detection for VanderLugt and joint transform correlator," Appl. Opt. **29,** 225–232 (1990).
5. F. T. S. Yu and X. J. Lu, "A real time programmable joint transform correlator," Opt. Commun. **52,** 10–16 (1984).
6. F. T. S. Yu, S. Jutamulia, T. W. Lin, and D. A. Gregory, "Adaptive real-time pattern recognition using a liquid crystal television based joint transform correlation," Appl. Opt. **26,** 1370–1372 (1987).
7. B. Javidi, "Nonlinear correlation joint transform correlation," Appl. Opt. **28,** 2358–2367 (1989).
8. B. Javidi, J. Wang, and Q. Tang, "Multiple-object binary joint transform correlation using multiple level threshold crossing," Appl. Opt. **30,** 4234–4244 (1991).
9. A. Stockseth, "Properties of a defocused optical system," J. Opt. Soc. Am. **59,** 1314–1321 (1969).

Fig. 5. (a) Correlation SNR versus SNR_i. (b) Detection probability versus SNR_i.

from 10^3 to 10^4. For small SNR_i, the correlation SNR falls to values between 50 and 100. Figure 5(b) shows the detection probability as a measure of the accuracy of the detection. Correlation peaks located