Wavelet transform-based correlator for the recognition of rotationally distorted images

Farid Ahmed*
Mohammad A. Karim, MEMBER SPIE
University of Dayton
Department of Electrical Engineering
Center for Electro-Optics
Dayton, Ohio 45469-0245

Mohammad S. Alam, MEMBER SPIE
Purdue University
Department of Engineering
Fort Wayne, Indiana 46805-1499

Abstract. A novel wavelet-based joint transform correlator (WJTC) for rotation-invariant pattern recognition and applications in optical image processing and remote sensing is investigated. First an optimal set of filter parameters and a mother wavelet filter are selected. These are used to extract features at different resolution from a set of rotationally distorted training images. Then a composite reference feature is formulated from these features for use in the WJTC. Simulation results for both noisy and noiseless environments are presented to verify the effectiveness of this technique.

Subject terms: optical remote sensing; image processing; rotation invariances; joint transform correlators; wavelet features; filter modulations; composite reference images.

Optical Engineering 34(11), 3187–3192 (November 1995).

1 Introduction

Detection of rotationally distorted images is a challenging task in remote-sensing applications. Primarily, two different approaches are reported in the literature for rotation-invariant recognition and detection. In the first approach, geometric invariance properties are employed in the filter formulation.[1–3] Circular harmonic filter expansion for rotation invariance has been used in these works. The main limitation with this technique is the proper choice of the harmonic expansion center. For the second approach, a number of rotationally distorted images are trained to result in a filter, which is expected to work well with any other rotated image. This is the widely accepted synthetic discriminant function (SDF) approach.[4,5] The minimum-variance synthetic discriminant filter (MVSDF)[6] and the minimum average correlation energy (MACE) filter[7] are two examples of such composite filter formulations. While they have been used with significant success, their suboptimal interpolation capability and broad correlation peaks limit their performance. Recently, a filter modulation technique for extracting better features for rotation invariance has also been successfully investigated.[8–10]

The wavelet transform has been attracting increasing attention in the optical pattern recognition community for its attractive multiresolution, denoising, and feature extraction capabilities.[11–14] We here propose to use wavelet features, modulated by a filter formulation, to come up with a composite reference image, which is to be used in a joint transform correlator (JTC)[15] setup. The classical JTC and its variants[16,17] result in poor discrimination in a noisy environment. In addition, they do not yield rotation-invariant detection. We address these two problems here in an efficient way. First,the filter formulation we are using is more optimal, flexible, and adaptive in nature. Secondly, we use a simple superposition scheme for getting a composite reference image. This facilitates rotation invariance, which was not possible with the simple filter formulation. Finally, the use of wavelet features gives additional improvement in discrimination.

2 Wavelet-Feature-Based Rotation-Invariant JTC

Let R be the range of in-plane rotational distortion. This is divided into M different classes of distortions. For each of these classes, a composite reference image is found. Suppose, $w_1, w_2, \ldots, w_n$ are the wavelet features of n rotationally distorted training images in class i, given by

$$w_k(a_x,a_y,b_x,b_y) = \frac{1}{(a_x a_y)^{1/2}} \iint s(x,y)\psi\left(\frac{x-b_x}{a_x},\frac{y-b_y}{a_y}\right) \mathrm{d}x\ \mathrm{d}y \ . \quad (1)$$

Here, a_x, a_y, b_x, b_y are the scaling and dilation parameters and $\psi(x,y)$ is the mother wavelet function. The choice of the mother wavelet filter is crucial in our application, which will be discussed in the next section. In the frequency domain, the wavelet transform is basically the inner product of the spectrum Ψ of the mother wavelet and the image, given by

$$W_k(a_x,a_y,b_x,b_y) = \sqrt{a_x a_y} \iint S(\mu,\nu)\Psi^*(a_x\mu,a_y\nu) \times \exp[-j2\pi(b_x\mu + b_y\nu)]\ \mathrm{d}\mu\ \mathrm{d}\nu \ , \quad (2)$$

*Current affiliation: Western Michigan Univ., ECE Dept., Kalamazoo, MI 49008.

Paper RS-003 received Apr. 9, 1995; revised manuscript received June 8, 1995; accepted for publication June 19, 1995.

where μ, ν are the frequency domain variables scaled by a factor $2\pi/\lambda f$, λ is the wavelength of the collimating light, and f is the focal length of the Fourier transform lens. This, in general, results in a number of bandpass decompositions of the original image.

Next, we modulate this representation to come up with the corresponding training features $r_i^{(1)}$, $r_i^{(2)}$ ..., $r_i^{(n)}$ using the following filter formulation:

$$r_i^{(k)} = \mathcal{F}^{-1} \frac{e^{j\Phi}}{|W_k| + A} , \tag{3}$$

where $|W_k| e^{j\Phi} = \mathcal{F}\{w_k\}$ as described by Eq. (2). Here, the parameter A is functionally dependent on the wavelet filters we are using, and in effect it determines the relative contribution of the bandpass decompositions in the construction of features. This is to be optimized for getting more optimal features through this transformation. Note that the above formulation resembles an amplitude-modulated phase-only filter (AMPOF), which was introduced to get better discrimination in correlation.[18] Here we use it as a feature extractor for the construction of the composite reference image. The resulting filter uses the discrimination ability of the AMPOF, but in addition it provides rotation invariance. The composite reference image c_i for class i is then found using the following superposition:

$$c_i = r_i^{(1)} + r_i^{(2)} + \cdots + r_i^{(n)} . \tag{4}$$

Now, in the reference half of the input joint image (IJI) of the JTC, we have M such composite images for M classes. With the input wavelet-transformed image $t(x,y)$, the IJI is given by

$$g(x,y) = t(x,y) + \sum_{i=1}^{M} c_i(x - x_i, y - y_i) . \tag{5}$$

The joint Fourier transform of this input results in

$$G(u,v) = T(u,v) + \sum_{i=1}^{M} C_i(u,v) \exp(-jux_i - jvy_i) . \tag{6}$$

The joint power spectrum (JPS) is given by

$$\begin{aligned} |G(u,v)|^2 = {} & |T(u,v)|^2 + \sum |C_i(u,v)|^2 \\ & + \sum C_i(u,v) T^*(u,v) \exp(-jux_i - jvy_i) \\ & + \sum C_i^*(u,v) T(u,v) \exp(+jux_i + jvy_i) \\ & + \sum_{l=1}^{M} \sum_{k=1, k\neq 1}^{M} C_l(u,v) C_k^*(u,v) \\ & \times \exp[-ju(x_l - x_k) - jv(y_l - y_k)] . \end{aligned} \tag{7}$$

Note that because of the use of Eq. (4) the term $|C_i(u,v)|^2$ is very small compared to the term $|T(u,v)|^2$. In reality, with proper choice of the parameters A and the mother wavelet filter parameters, we can find a desired minimum value for this term. Likewise, the fifth term, which is the cross-correlation among the composite images, can also be neglected in the JPS. Therefore, for all practical purposes, Eq. (7) can be approximated by

$$\begin{aligned} |G|^2 = {} & |T|^2 + \sum_{i=1}^{M} [C_i T^* \exp(-jux_i - jvy_i) \\ & + TC_i^* \exp(+jux_i + jvy_i)] . \end{aligned} \tag{8}$$

The simulation results presented in the next section validate the above assumption. The output correlation is found by subtracting the zero-order term contributed by the input image from the JPS and then taking the inverse Fourier transform.

Note that in this approach, we get M output correlation peaks, each corresponding to one class of rotations. Thus the correlation output also gives an estimate of the rotational distortion of the target. This is an interesting outcome of the present work.

3 Simulation Results

To validate the performance of the proposed model, we consider two different T72 tank images and the F16 aircraft, as shown in Fig. 1. The first tank image is taken as the target, and other two images are considered as nontargets. In the simulation, each image is downsampled to 32×32 pixels, while the joint image of the JTC is of size 256×256. The training images with different in-plane rotations are first selected. We consider an interval of 5 deg in the distortion range of 0 to 90 deg. This is then divided into $M = 3$ classes of distortions of 0 to 30, 30 to 60, and 60 to 90 deg. The wavelet-feature descriptors of these training images are then found.

Figure 2 shows the corresponding wavelet features of undistorted images. Here, we use the Haar wavelet transform at the highest level of decomposition. The Haar mother wavelet in 2-D is given by

$$\begin{aligned} \psi(x,y) = {} & \text{rect}(x - 0.5, y - 0.5) + \text{rect}(x + 0.5, y + 0.5) \\ & - \text{rect}(x - 0.5, y + 0.5) - \text{rect}(x + 0.5, y - 0.5) . \end{aligned} \tag{9}$$

Figure 3 shows the wavelet, $\psi(x,y)$ and its Fourier transform $\Psi(x,y)$.

The motivation for using the Haar wavelet is to extract more edge features. Another advantage is easier optical implementation because of the bipolar characteristics of the Haar wavelet. We are not interested in the localization property of the wavelet, for which Haar transform is not very good. It is a well-known fact that edge-enhanced images perform well in terms of correlation discrimination in a JTC.[19] In other words, it is the dominance of low-frequency components in the features that is responsible for broader correlation peaks.

Next, we modulate these features with the filter formulation in Eq. (3). The goal here is to obtain the sharper and more discriminating correlation performance of these filter formulations in the correlation operation. This new feature is then synthesized to form the reference image of JTC. For testing the correlation performance we use distorted images at an interval of 1 deg. Figure 4 shows the correlation peak values from the JTC, for different distorted images in the range 0 to 30 deg. The nontarget used here is the F16. It shows that better discrimination between target and nontarget

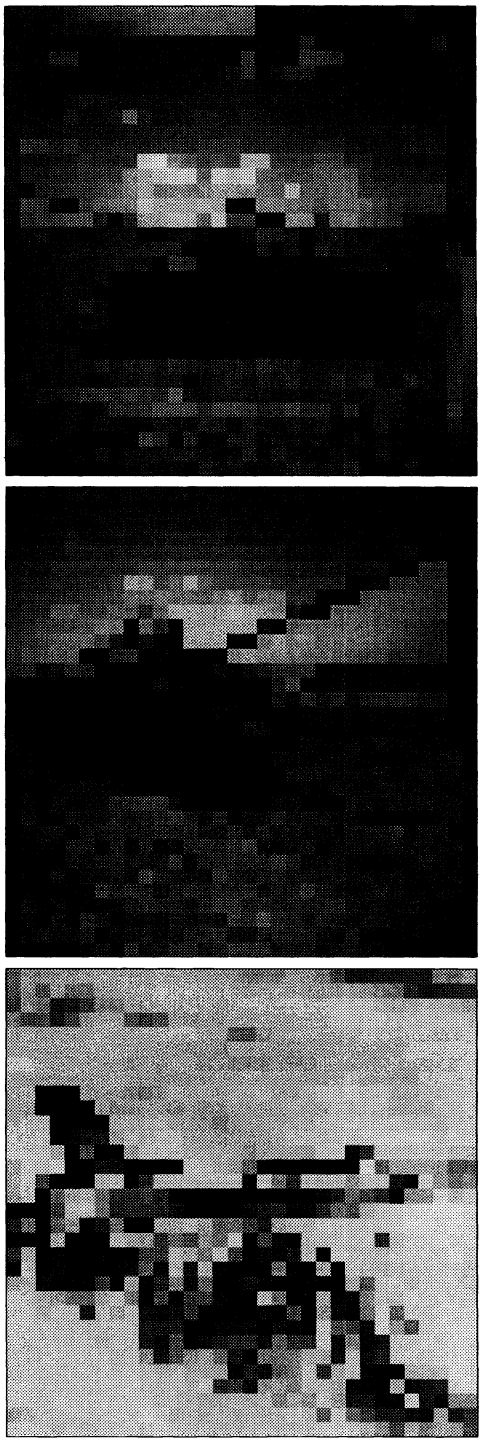


Fig. 1 Test images: two different tanks and F16 aircraft.

is obtained with high-resolution (fine-scale) decomposition [Fig. 4(b)] than with low-resolution (coarse-scale) decomposition [Fig. 4(a)]. This follows from the property of the wavelet transforms according to which increasing level of decomposition increases decorrelation.[20] But the price for high resolution is the decrease in the correlation peak values, as seen from the figure. The value of the parameter A used here is 1. To demonstrate the discrimination ability further, we take the second tank in Fig. 1 as the nontarget. A value

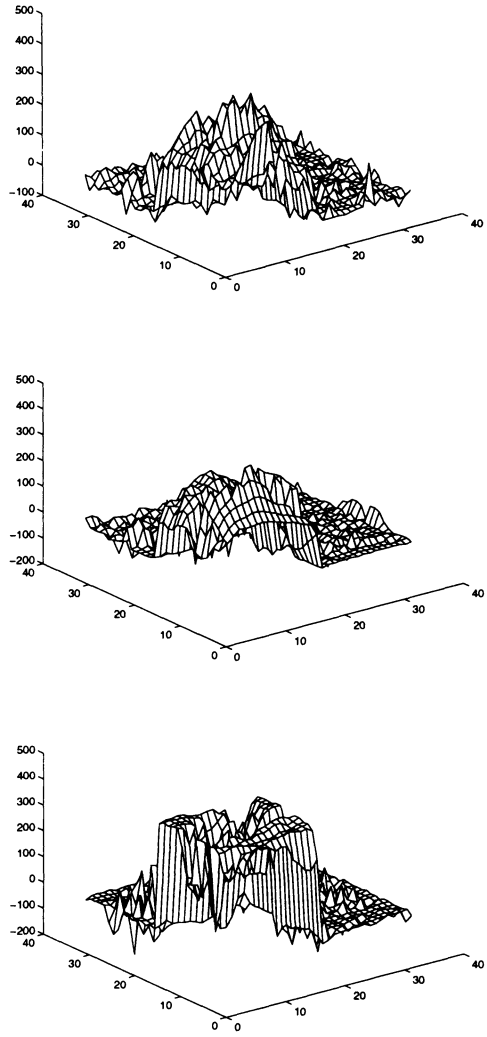


Fig. 2 Wavelet feature of test images.

$A = 10$ here results in highly discriminatory correlation output, as shown in Fig. 5.

For the remaining part of simulation we will be using the high-resolution wavelet features and the value $A = 1$. Figures 6(a) and 6(b) illustrate the corresponding correlation performance in the distortion ranges of 30 to 60 and 60 to 90 deg. Note that in these cases we can set a threshold value of the correlation peak, above which the input can be treated as a target and below which it is a nontarget. Next, we place the three composite reference images for the three classes of distortions in the input joint image. The input comprises 91 distorted images of the target in the range 0 to 90 deg in steps of 1 deg. We get three correlation peaks, corresponding to the three classes of distortions. Figure 7 shows the results. Note particularly how class 1 (0–30) performs well in the distortion range 0–30 deg, but its correlation performance is poor for the range 30–90 deg. In the same way, each class gives better performance in its domain of distortion. We then combine these ranges of better performance. Figure 8 shows the resulting discrimination between target and nontarget. Figure 9 gives the output correlation plot for the particular distortion value of 90 deg. Note that the peak corresponding to class 3 (60–90), which is at the upper right-hand corner

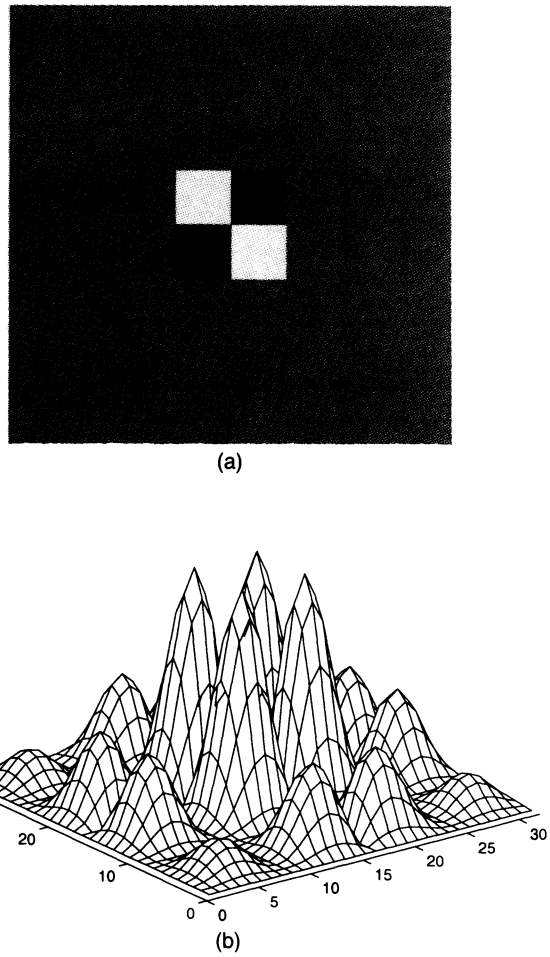


Fig. 3 (a) Haar mother wavelet and (b) Fourier transform of wavelet.

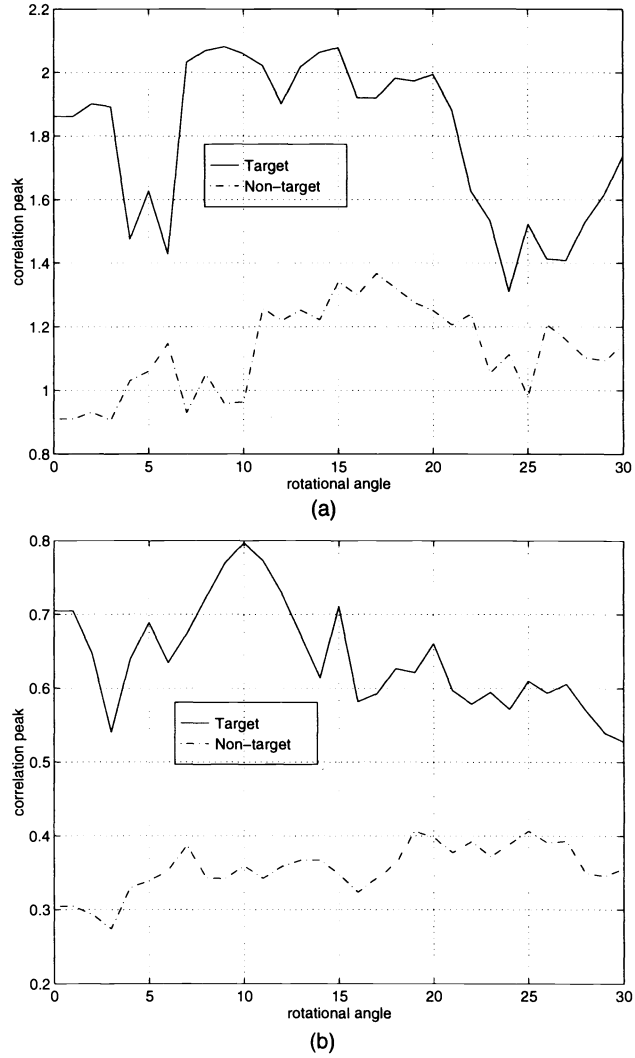


Fig. 4 Correlation performance in the 0 to 30-deg distortion range with (a) low-resolution and (b) high-resolution wavelet decomposition feature. Here $A = 1$.

of the correlation plane, is higher than the other two peaks. This tells us that the input distorted image is in class 3, in addition to the result that it can be discriminated from a nontarget.

Finally, we perform the simulation with zero-mean, Gaussian white noise of different variances. With 10-dB noise the discrimination is still good, as seen from Fig. 10(a). The corresponding noise variance is 3000, resulting in a standard deviation of about 58 in the pixel gray values. Performance deteriorates as the noise variance is increased, as demonstrated by Fig. 10(b).

4 Conclusion

A wavelet-based joint transform correlator for rotation-invariant optical target detection has been presented. The scheme uses a simple technique for the extraction of a composite reference image from the wavelet features. With the choice of proper thresholding in the output correlation, the proposed technique is shown to result in highly robust and discriminating rotation-invariant detection.

References

1. Y. N. Hsu and H. H. Arsenault, "Optical pattern recognition using circular harmonic expansion," *Appl. Opt.* **21,** 4016–4019 (1982).
2. F. T. S. Yu, X. Y. Li, S. Jutamalia, and D. A. Gregory, "Rotation in-

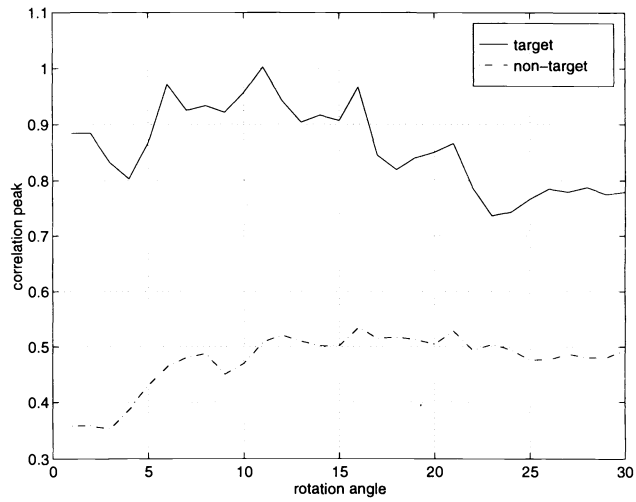
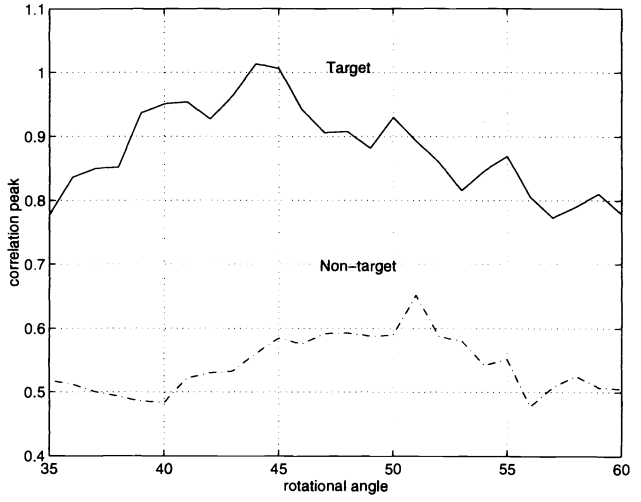
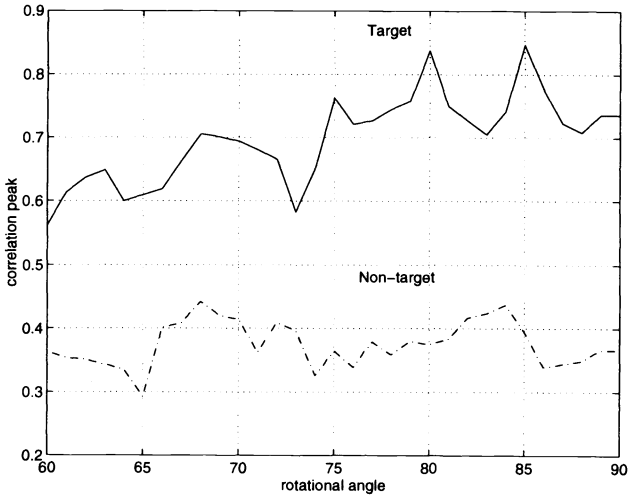


Fig. 5 Correlation performance in the 0 to 30-deg distortion range for $A = 10$.



(a)



(b)

Fig. 6 Correlation performance in the distortion range of (a) 35 to 60 deg and (b) 60 to 90 deg.

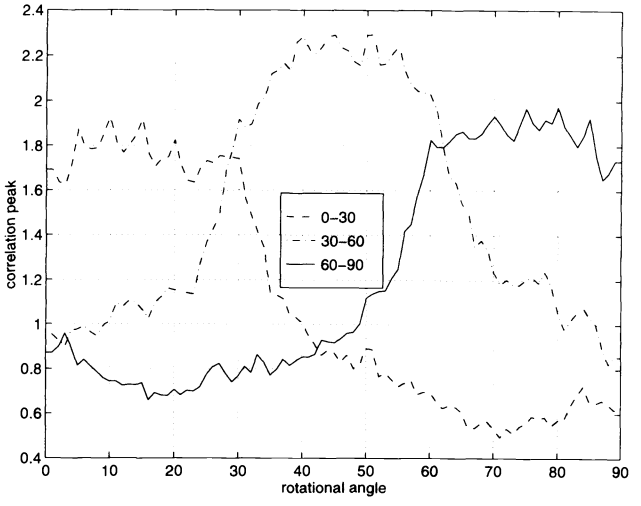


Fig. 7 Correlation performance of the composite JTC for the target for the whole distortion range of 0 to 90 deg.

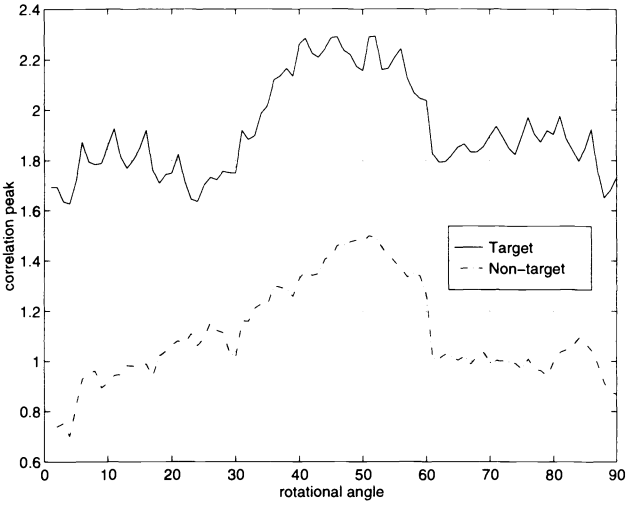


Fig. 8 Correlation discrimination of the target and nontarget for the composite JTC.

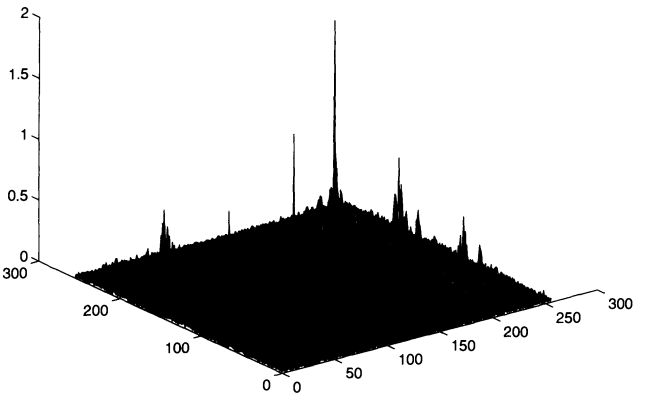


Fig. 9 Correlation output for a target rotated by 90 deg.

variant pattern recognition with a programmable joint transform correlator," *Appl. Opt.* **28,** 4725–4727 (1989).
3. E. Elizur and A. A. Friesem, "Rotation-invariant correlation with incoherent light," *Appl. Opt.* **30,** 4175–4178 (1991).
4. C. F. Hester and D. Casasent, "Multivariant technique for multiclass pattern recognition," *Appl. Opt.* **19,** 1758–1761 (1980).
5. D. P. Casasent, "Unified synthetic discriminant function computational formulation," *Appl. Opt.* **23,** 1620–1627 (1984).
6. B. V. K. Vijayakumar, "Minimum variance synthetic discriminant functions," *J. Opt. Soc. Am A* **3,** 1579–1584 (1986).
7. A. Mahalanobis, B. V. K. Vijayakumar, and D. Casasent, "Minimum average correlation energy filters," *Appl. Opt.* **26,** 3633–3640 (1987).
8. D. Jared and D. Ennis, "Inclusion of filter modulation in the synthetic discriminant function construction," *Appl. Opt.* **28,** 232–239 (1989).
9. M. B. Reid, P. W. Ma, J. D. Downie, and E. O. Choa, "Experimental verification of modified synthetic discriminant function filters for rotation invariance," *Appl. Opt.* **29,** 1209–1214 (1990).
10. R. K. Wang, C. R. Chatwin, and M. Y. Huang, "Modified filter synthetic discriminant functions for improved optical correlator performance," *Appl. Opt.* **33,** 7646–7654 (1994).
11. D. Casasent, J. Smokelin, and A. Ye, "Wavelet and Gabor transforms for detection," *Opt. Eng.* **31,** 1893–1898 (1992).
12. D. P. Casasent and J. S. Smokelin, "Neural net design of macro Gabor wavelet filters for distortion-invariant object detection in clutter," *Opt. Eng.* **33,** 2264–2271 (1994).
13. X. J. Lu, A. Katz, E. G. Kanterakis, and N. P. Caviris, "Joint transform correlator that uses wavelet transform," *Opt. Lett.* **17,** 1700–1702 (1992).
14. G. Jin, Y. Yan, W. Wang, Z. Wen and M. Wu, "Image feature extraction with the optical Haar wavelet transform," *Opt. Eng.* **34,** 1238–1242 (1995).

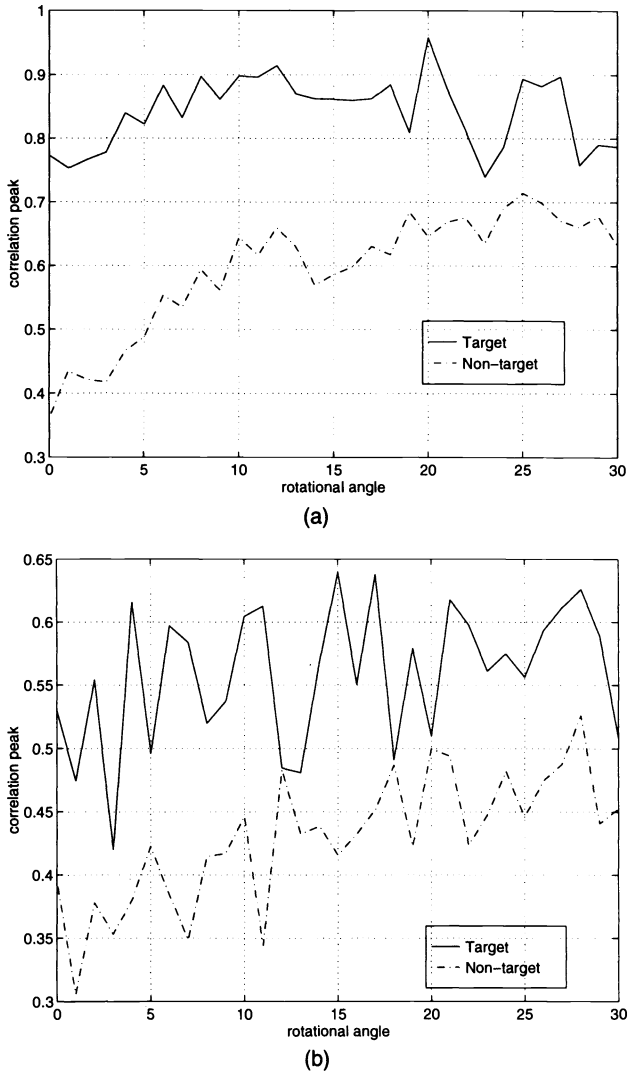


Fig. 10 Correlation performance in noise: (a) SNR = 10 db, (b) SNR = 5 dB.

15. F. T. S. Yu and J. E. Ludman, "Microcomputer based programmable joint transform correlator for automatic pattern recognition and identification," *Opt. Lett.* **11,** 395–397 (1986).
16. B. Javidi and C. Kuo, "Joint transform image correlation using a binary spatial light modulator at the Fourier plane," *Appl. Opt.* **27,** 663–665 (1988).
17. M. S. Alam and M. A. Karim, "Fringe-adjusted joint transform correlator," *Appl. Opt.* **32,** 4351–4356 (1993).
18. A. A. S. Awwal, M. A. Karim, and S. R. Jahan, "Improved correlation discrimination using an amplitude-modulated phase-only filter for optimum recognition," *Appl. Opt.* **29,** 233–236 (1990).
19. M. S. Alam, O. Perez, and M. A. Karim, "Preprocessed multiobject joint transform correlator," *Appl. Opt.* **32,** 3102–3107 (1993).
20. I. Daubechies, *Ten Lectures on Wavelets,* SIAM, Philadelphia, 1992.

Farid Ahmed received his BS degree in electrical and electronic engineering in 1988 from Bangladesh University of Engineering and Technology (BUET) and his MS degree in computer science and engineering in 1993 from Wright State University, Dayton, OH. He is currently a PhD student in the Electrical Engineering Department of the University of Dayton, Dayton, OH. He was a Lecturer and Assistant Professor in the Department of Computer Science and Engineering, BUET, Bangladesh, from 1988 to 1990. His research interests include pattern recognition, optical and digital image processing, and neural networks, wavelet transforms, and digital design.

Mohammad A. Karim: Biography and photograph appear with the special section guest editorial in this issue.

Mohammad S. Alam: Biography and photograph appear with the paper "Fractional power fringe-adjusted joint transform correlation" in this issue.