International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:2, No:5, 2008

Target Detection using Adaptive Progressive Thresholding Based Shifted Phase-Encoded Fringe-Adjusted Joint Transform Correlator

Inder K. Purohit, M. Nazrul Islam, K. Vijayan Asari, and Mohammad A. Karim

***Abstract*—**A new target detection technique is presented in this paper for the identification of small boats in coastal surveillance. The proposed technique employs an adaptive progressive thresholding (APT) scheme to first process the given input scene to separate any objects present in the scene from the background. The preprocessing step results in an image having only the foreground objects, such as boats, trees and other cluttered regions, and hence reduces the search region for the correlation step significantly. The processed image is then fed to the shifted phase-encoded fringe-adjusted joint transform correlator (SPFJTC) technique which produces single and delta-like correlation peak for a potential target present in the input scene. A post-processing step involves using a peak-to-clutter ratio (PCR) to determine whether the boat in the input scene is authorized or unauthorized. Simulation results are presented to show that the proposed technique can successfully determine the presence of an authorized boat and identify any intruding boat present in the given input scene.

***Keywords*—**Adaptive progressive thresholding, fringe adjusted filters, image segmentation, joint transform correlation, synthetic discriminant function.

I. Introduction

COASTAL areas are one of the prime regions that require constant surveillance to protect against any intrusions and attacks. Manual surveillance system to distinguish authorized vessels from unauthorized ones is usually a cumbersome process and not a practical solution for 24 hour reliable and robust security systems. Therefore, automated target detection system is an urgent need for the port authority and similar other organizations in determining whether the approaching boat is authorized and in generating alarm signal for an unauthorized boat in the secure area. Optical joint transform correlation (JTC) has been a very useful technique for real time pattern recognition [1]. However, classical JTC technique has the drawbacks, such as, poor correlation discrimination, wide side lobes, pair of correlation for a single object, and strong zero order correlation which often overshadows the desired correlation peak. In order to overcome these drawbacks, a number of modifications have been proposed for JTC technique, such as binary JTC [2], phase-only JTC [3] and fringe-adjusted JTC [4], which can overcome some of the problems mentioned above, but they still suffer from generation of duplicate correlation peaks for the same object. Recently proposed shifted phase-encoded fringe-adjusted JTC (SPFJTC) technique has been found to be effective in yielding single delta function-like correlation peak with high discrimination between the target and the non-targets [5]. As there are thousands of boats of different sizes and shapes and they may appear in the input scene with different orientations, it is very difficult to train a correlation system with many reference images and hence find the correlation. The synthetic discriminant function (SDF) has been very useful in distortion invariant pattern recognition applications [6]. A set of synthetic reference images is employed with probable distortions to train the system prior to real-time operation to make it invariant to scale and rotation variations as the target in the input scene might appear at any possible angle.

A novel target detection system is proposed in this paper using adaptive progressive thresholding (APT) based SPFJTC technique. Images of a boat with different orientations are used to generate a synthetic image, called synthetic discriminant function (SDF) image, which can be used as the reference image. In order to save the processing time and to prevent false positives, the input scene is adaptively thresholded to separate boats and other objects from the background in the given input scene and to hence reduce the search space [7]. The segmented image is then used as an input to the SPFJTC technique incorporating the synthetic reference image. A post-processing step called peak-to-clutter ratio (PCR), is performed to confirm the correlation decision regarding the presence of an authorized or an unauthorized boat in the input scene. The performance of the proposed target detection technique is evaluated in computer simulation using real-life images of coastal region.

Manuscript submitted on April 30, 2008.

Inder K. Purohit is with the Electrical and Computer Engineering Department, Old Dominion University, Norfolk, VA 23529 (phone: 757-683-3741; fax: 757-683-3220; e-mail: ipuro001@odu.edu).

M. Nazrul Islam is with the Research Foundation, Old Dominion University, Norfolk, VA 23529 (e-mail: mislam@odu.edu).

K. Vijayan Asari is with the Electrical and Computer Engineering Department, Old Dominion University, Norfolk, VA 23529 (e-mail: vasari@odu.edu).

Mohammad A. Karim is with the Office of Research, Old Dominion University, Norfolk, VA 23529 (e-mail: mkarim@odu.edu).

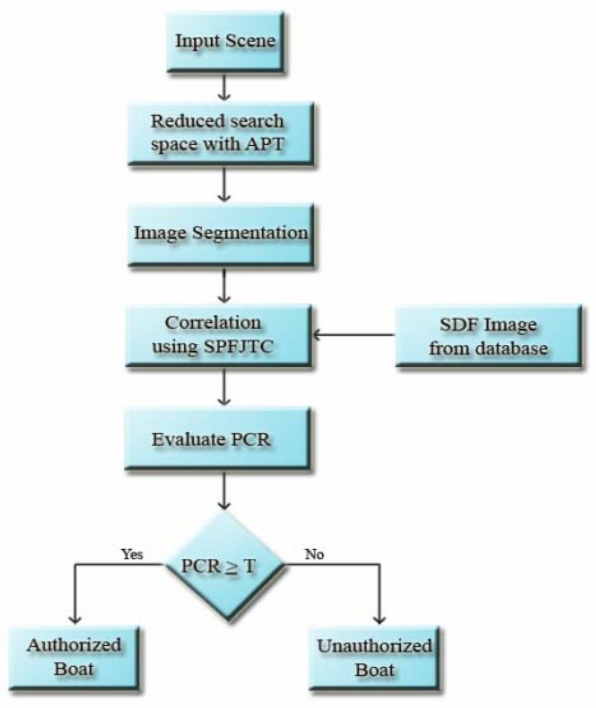


Fig. 1 An overview of the system being used for target detection

II. Analysis of the Proposed Technique

Fig. 1 shows a brief overview of the proposed system for target detection. The input image is first adaptively thresholded using the Otsu's method [7]. The binary image thus obtained is used to find the region of interest where there are any boats or objects. The segmented image is then used as the input for the SPFJTC technique which yields a sharp correlation peak for an authorized boat.

A. Adaptive Thresholding

As the input images might be taken at different times of the day, a fixed threshold may not work for different illumination conditions. For this reason, an adaptive thresholding method is incorporated based on the discriminant analysis, where the image is partitioned into two classes, C_0 and C_1, which can be classified as objects and background, respectively. If the threshold is at a gray level t, then the two classes can be given as $C_0 = \{0, 1, \ldots t\}$ and $C_1 = \{t+1, t+2, \ldots L-1\}$ where L is the number of gray levels in the image. Let σ_B^2 and σ_T^2 be the variance between classes and the total variance, respectively. Then an optimum threshold t^* can be obtained by maximizing the variance between classes as given by

$$t^* = Arg\{\max_{0\le i\le L-1}(\eta)\} \tag{1}$$

where

$$\eta = \frac{\sigma_B^2}{\sigma_T^2} \tag{2}$$

$$\sigma_B^2 = w_0 w_1 (\mu_1 - \mu_0)^2 \tag{3}$$

$$\sigma_T^2 = \sum_{i=0}^{L-1} (i - \mu_T)^2 \frac{n_i}{M} \tag{4}$$

where w_0 and w_1 are the fraction of pixels present in C_0 and C_1, respectively, and are given by

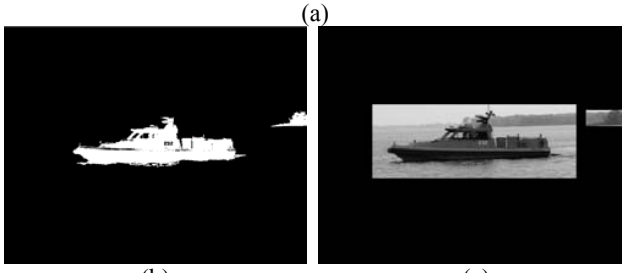
$$w_0 = \sum_{i=0}^{t} \frac{n_i}{M}, \qquad w_1 = 1 - w_0 \tag{5}$$

where n_i is the number of pixels on the i^{th} gray level, and M is the total number of pixels in the image, μ_0 and μ_1 are the class means for C_0 and C_1, respectively. The mean values can be found as

$$\mu_0 = \frac{\mu_t}{w_0} \qquad \mu_1 = \frac{\mu_T - \mu_t}{1 - w_0} \tag{6}$$

$$\mu_t = \sum_{i=0}^{t} i.\frac{n_i}{M} \qquad \mu_T = \sum_{i=0}^{L-1} i.\frac{n_i}{M} \tag{7}$$

(a)



(b) (c)

Fig. 2 Adaptive progressive thresholding process: (a) input scene, (b) thresholded image, and (c) segmented image

The threshold value is set at the gray level t^* found from (1), using which the image is further divided into two classes. Thus a new image $I^{(1)}$ is formed such that all the pixels in the original image that are higher than t^* are excluded from $I^{(1)}$. Hence, the pixels contained in $I^{(1)}$ will have a range of $C^{(1)}$ given by $\{0, 1, 2 \ldots t^*\}$. This procedure is applied recursively on the input scene, and the cumulative limiting factor (CLF) is used to find an appropriate threshold after each iteration. The CLF value for the Δ^{th} iteration can be obtained as

$$\mathrm{CLF}(\Delta) = \frac{\sigma_B^2(\Delta)}{\sigma_T^2}, \quad \text{for } \Delta \ge 1 \tag{8}$$

where $\sigma_B^2(\Delta)$ is calculated as in (3) by taking w_0, w_1, μ_0 and μ_1 from the progressive image $I^{(\Delta)}$. Then the appropriate threshold $t^*(\Delta)$ is obtained by maximizing the value of CLF(Δ) for the image $I^{(\Delta)}$ [8, 9]. The iteration stops when CLF(Δ) becomes smaller than an empirically determined value, such as

$$\mathrm{CLF}(\Delta) \le \alpha \frac{\mu_T}{\sigma_T^2} \tag{9}$$

where α is the limiting parameter, which can be obtained by repeated experiments on the images. The thresholded image thus obtained is then inverted in order to show the boat, trees and other cluttered regions as white blocks on a black background. The portions of the input scene corresponding to the connected regions in the thresholded image are extracted and used as input to the next stage. Due to thresholding, some of the edges of the connected regions might be lost, so a few additional pixels are picked up in all directions of the connected region. Fig. 2(a) shows a sample input scene under investigation. The thresholded image and the image with the reduced search space are shown in Figs. 2(b) and 2(c), respectively. The plot for the gray-levels versus between-class variance is shown in Fig. 3.

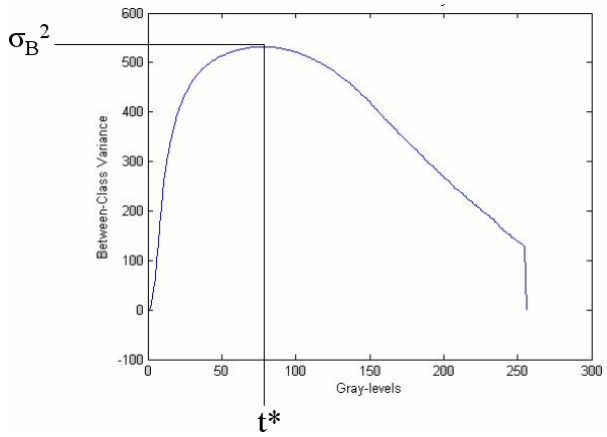


Fig. 3 Gray-levels vs between-class variance plot

B. Synthetic Discriminant Function

There are two images that are used as input to the SPFJTC technique, one is the input scene at hand and the other is the reference image. The reference image is formed as a synthetic discriminant function (SDF) from a set of training images with possible distortions. If there are N training images $r_1(x, y)$, $r_2(x, y)$ … $r_N(x, y)$ containing the possible distorted features of the object to be detected, the spatial SDF, $r_{\mathrm{SDF}}(x,y)$, can be synthesized as a weighted average function of the training images as

$$r_{SDF}(x,y) = \frac{\sum_{i=1}^{N} a_i r_i(x,y)}{\sum_{i=1}^{N} a_i} \tag{10}$$

where a_i's are the associated coefficients for the respective training images. The coefficients are selected such that the SDF produces uniform correlation peak with each of the training images. The values of the coefficients are iteratively selected, with all the coefficients starting with unity value for the first iteration. After each iteration, the resultant SDF is used as the reference image with which each of the training images are compared to form a correlation matrix given by

$$corr_i^k(x,y) = r_{SDF}^k(x,y) \oplus r_i(x,y) \tag{11}$$

where k is the iteration number and $\oplus$ is the correlation operator. The correlation peak intensities corresponding to each training image are evaluated as

$$C_i^k \in \max(corr_i^k) \tag{12}$$

Then the maximum peak intensity is estimated as

$$C_{\max}^k \in \max(C_i^k) \tag{13}$$

In order to have equal correlation peak intensity for all images in the training set, the coefficients a_i of (10) are updated using the empirical relation given by

$$a_i^{k+1} = a_i^k + (C_{\max}^k - C_i^k)\delta \tag{14}$$

where δ is the relaxation factor which determines the rate of changing the coefficients from one iteration to the next and the coefficients are so chosen that $a_i^0 = 1$. The iterative process is continued until the difference between the maximum and the minimum correlation peak intensities becomes less than an error limit ξ^k given by

$$\xi^k = \frac{(C_{\max}^k - C_{\min}^k)}{C_{\max}^k} \tag{15}$$

where $C_{\max}^k$ and $C_{\min}^k$ are the maximum and the minimum correlation peaks, respectively, for the k^{th} iteration. A sample case of SDF formation is depicted in Fig. 4. A total of 12 different images for the same boat are employed as shown in Figs. 4(a) – 4(l). The SDF image thus formed is shown in Fig. 4(m), which gives equal correlation peak for the detection of the boat in any one of the orientations.

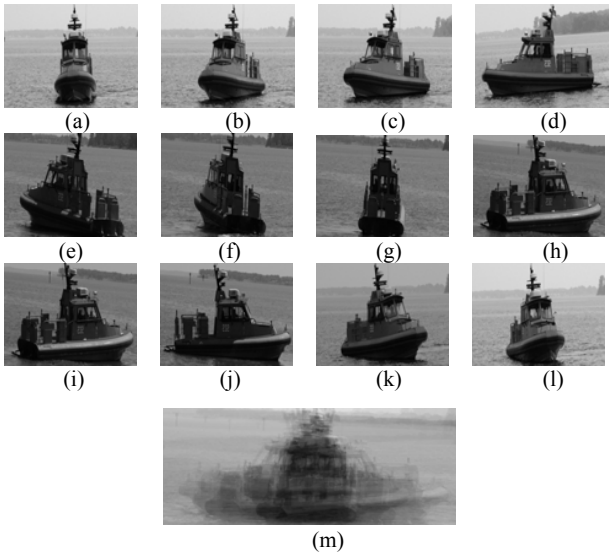


Fig. 4 Formation of SDF image: (a) – (l) training images, and (m) the SDF image

C. Shifted Phase-Encoded Fringe-Adjusted Joint Transform Correlation

The SPFJTC technique is used for distortion-invariant class-associative pattern recognition. The SDF image developed is Fourier transformed and fed into two parallel

channels, where one channel introduces a 180 degree phase shift. A random phase mask, φ(*x*, *y*), is applied to the two channels individually and the two resultant signals are given by

$$S_1(u,v) = |R_{SDF}(u,v)| \exp[j\phi_r(u,v)] \times \phi(u,v) \quad (16)$$

$$S_2(u,v) = |R_{SDF}(u,v)| \exp[j\pi] \exp[j\phi_r(u,v)] \times \phi(u,v) \quad (17)$$

where $R_{SDF}(u,v)$ and $\varphi_r(u,v)$ are the amplitude and phase, respectively, of the Fourier transform of $r_{SDF}(x, y)$, $\varphi(u,v)$ is the Fourier transform of $\varphi(x, y)$, u and v are mutually independent frequency domain variables scaled by a factor of $2\pi/f\lambda$, where λ is the wavelength of collimating light, f is the focal length of the Fourier lens. The inverse Fourier transform of these signals gives the phase-encoded reference SDF images as

$$S_1(x,y) = r_{SDF}(x,y) \otimes \phi(x,y) \quad (18)$$

$$S_2(x,y) = -r_{SDF}(x,y) \otimes \phi(x,y) \quad (19)$$

The segmented input scene containing the different objects, $t_i(x, y)$, is now introduced to both the channels to form two joint images given by

$$f_1(x,y) = S_1(x,y) + \sum_{i=1}^{k} t_i(x,y) = \sum_{i=1}^{M} r_{SDF_i}(x,y) \otimes \phi(x,y) + \sum_{i=1}^{k} t_i(x,y) \quad (20)$$

$$f_2(x,y) = S_2(x,y) + \sum_{i=1}^{k} t_i(x,y) = -\sum_{i=1}^{M} r_{SDF_i}(x,y) \otimes \phi(x,y) + \sum_{i=1}^{k} t_i(x,y) \quad (21)$$

where M is the number of objects in the reference class, k is the number of objects in the input scene. Fourier transform of (20) and (21) gives the joint power spectra (JPS) given by

$$\begin{aligned} |F_1(u,v)|^2 &= \sum_{i=1}^{M} |R_{SDFi}(u,v)\phi(u,v)|^2 + \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{M} R_{SDFi}(u,v). \\ & R^*_{SDFi}(u,v) + \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{M} R^*_{SDFi}(u,v) R_{SDFi}(u,v) + \\ & \sum_{i=1}^{K} |T_i(u,v)|^2 + \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{K} R_{SDFi}(u,v) T^*_j(u,v)\phi(u,v) + \\ & \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{K} R^*_{SDFi}(u,v) T_j(u,v)\phi^*(u,v) + \sum_{i=1}^{K}\sum_{j=1,j\neq i}^{K} T_i(u,v). \\ & T^*_j(u,v) + \sum_{i=1}^{K}\sum_{j=1,j\neq i}^{K} T^*_i(u,v) T_j(u,v) \end{aligned} \quad (22)$$

$$\begin{aligned} |F_2(u,v)|^2 &= \sum_{i=1}^{M} |R_{SDFi}(u,v)\phi(u,v)|^2 + \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{M} R_{SDFi}(u,v). \\ & R^*_{SDFi}(u,v) + \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{M} R^*_{SDFi}(u,v) R_{SDFi}(u,v) + \\ & \sum_{i=1}^{K} |T_i(u,v)|^2 - \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{K} R_{SDFi}(u,v) T^*_j(u,v)\phi(u,v) - \\ & \sum_{i=1}^{M}\sum_{j=1,j\neq i}^{K} R^*_{SDFi}(u,v) T_j(u,v)\phi^*(u,v) + \sum_{i=1}^{K}\sum_{j=1,j\neq i}^{K} T_i(u,v). \\ & T^*_j(u,v) + \sum_{i=1}^{K}\sum_{j=1,j\neq i}^{K} T^*_i(u,v) T_j(u,v) \end{aligned} \quad (23)$$

The two JPSs in the above equations contain a number of terms that produce autocorrelation and cross-correlation terms. Equation (23) is subtracted from (22) and the resultant signal is multiplied by the same phase mask that was used earlier. So we get,

$$P(u,v) = [|F_1(u,v)|^2 - |F_2(u,v)|^2]\phi(u,v) \quad (24)$$

The inverse Fourier transform of (24) gives the correlation output. The modified JPS in (24) will contain two terms to produce cross-correlation output. However, one of the terms will be scattered in space because of the random nature of the phase mask. Therefore, only a single term will produce the final correlation output. In order to enhance the correlation performance of the technique, a modified fringe-adjusted filter (FAF) is developed, given by

$$H(u,v) = \frac{C(u,v)}{D(u,v) + \sum_{i=1}^{M} \alpha_i |R_{SDF_i}|^2} \quad (25)$$

where $C(u,v)$ and $D(u,v)$ are either constants or functions of u and v. The parameter $C(u,v)$ is adjusted to avoid having an optical gain greater than unity, while $D(u,v)$ is used to overcome the pole problem associated with a normal filter. The constants a_i's are adjusted such that $\Sigma\alpha_i = 1$. The JPS given in (24) is multiplied by the FAF transfer function which results in an enhanced JPS given by

$$P_f(u,v) = P(u,v).H(u,v) \quad (26)$$

Inverse Fourier transformation of the signal in (26) results in a single and very sharp correlation peak of uniform height for each potential target object of the reference class present in the input scene.

D. Post Processing

A post processing step is performed to confirm the correlation decision regarding the presence of an authorized or an unauthorized boat. The peak-to-clutter ratio (PCR) is measured from the correlation output which is obtained from the inverse Fourier transform of (26). The PCR value can be determined as [10],

$$PCR = \frac{CPI_t}{CPI_c} \quad (27)$$

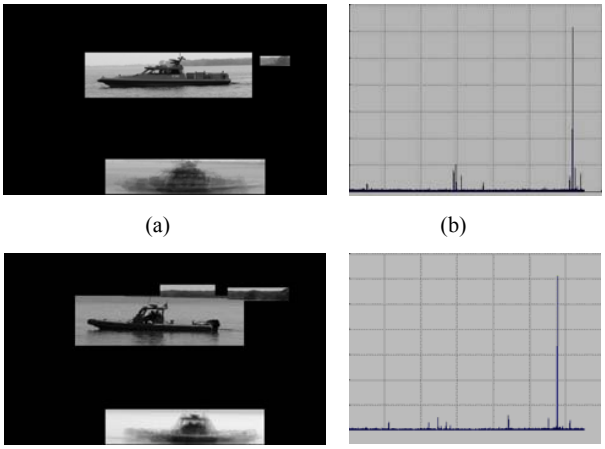
where CPI_t is the correlation peak intensity (CPI) of the desired target and CPI_c is the CPI of the background clutter. The value of PCR is then compared with a threshold, T, to determine whether there is an authorized boat present in the input scene using the following hypotheses,

$$\begin{aligned} PCR \geq T &\rightarrow \text{Authorized Boat} \\ PCR < T &\rightarrow \text{Unauthorized Boat} \end{aligned} \quad (28)$$

III. Experimental Results

One of the major advantages of the proposed technique is that the processing time is significantly reduced, as the background of the input scene need not be used for the correlation purpose. The proposed algorithm has been investigated using computer simulation program developed in MATLAB 7.1 software. The experiments were done on a PC with Pentium 4 processor (2.4GHz clock and 1.0GB RAM) in Windows XP platform. Fig. 5(a) and 5(c) show two joint

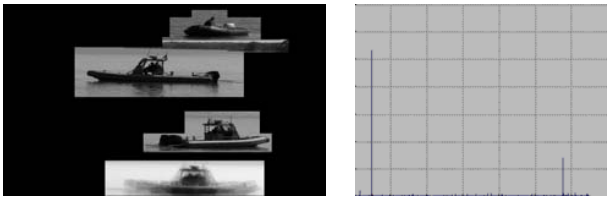
images where the top part contains the segmented input image and the bottom part contains the SDF image. The correlation output shown in Fig. 5(b) and 5(d), respectively, verify that the proposed technique can successfully detect the boat by generating a high and sharp peak, while rejecting other objects of the input scene with almost insignificant correlation peaks as compared to the target peak. Fig. 6(a) contains multiple targets in the input scene. Out of these, two boats are authorized, while one boat is unauthorized. This fact is confirmed by the correlation output shown in Fig. 6(b).



(a) (b)

(c) (d)

Fig. 5 Detection performance: (a),(c) joint image with the segmented input and SDF reference image, and (b),(d) correlation output

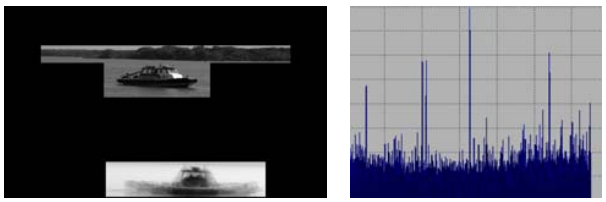


(a) (b)

Fig. 6 Detection performance with multiple targets: (a) joint image and (b) correlation output

TABLE I
PERFORMANCE EVALUATION

Number of Boat Images	Authorized/ Unauthorized	Rightly Classified	Accuracy
120	100/20	111	92.50%

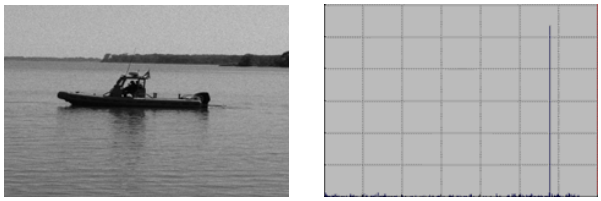


(a) (b)

Fig. 7 (a) Joint input scene with an unauthorized boat and (b) correlation output

As seen from Fig. 7, when there is an unauthorized boat in the input scene, the correlation output consists of noise and no distinct peak is seen for it (the noise level has been amplified for displaying). Then a noisy input scene is investigated where also the proposed algorithm can detect the authorized boat as shown in Fig. 8.

The simulations were carried out on 120 boat images consisting of various orientations of 12 boats. As listed in Table I, of these 120 boat images, 100 belonged to authorized boats and 20 belonged to unauthorized ones. The algorithm could successfully classify the boats as either being authorized or unauthorized with an accuracy level of 92.5%.



(a) (b)

Fig. 8 (a) A noisy input scene and (b) Correlation output

Although most of the boats are rightly classified, some of the boat images generated false positives. One reason for this error could be the vast number of shapes and sizes of these boat images, which make the training of the SDF image a difficult task. Further work is being carried out in achieving a better procedure to train the system that tackles these problems and helps in rightly classifying the boats and reducing the misclassification.

ACKNOWLEDGMENT

The authors of this paper would like to thank US Army Night Vision Lab for providing the boat images used for the testing of the proposed technique.

REFERENCES

[1] C. S. Weaver and J. W. Goodman, “Technique for optically convolving two functions,” *Applied Optics*, vol. 5, pp. 1248 – 1249, 1966.
[2] B. Javidi and C. Kuo, “Joint transform image correlation using a binary spatial light modulator at the Fourier plane,” *Applied Optics*, vol. 27, pp. 663 – 665, 1988.
[3] A. K. Cherri and M. S. Alam, “Reference phase-encoded fringe-adjusted joint transform correlation,” *Applied Optics*, vol. 40, pp. 1216 – 1225, 2001.
[4] M. S. Alam and M. A. Karim, “Fringe-adjusted joint transform correlation,” *Applied Optics*, vol. 32, pp. 4344 – 4350, 1993.
[5] M. R. Haider, M. N. Islam, M. S. Alam and J. F. Khan, “Shifted phase-encoded fringe-adjusted joint transform correlation for multiple target detection,” *Optics Communications*, vol. 248, pp. 69 – 88, 2005.
[6] X. W. Chen, M. A. Karim, and M. S. Alam, “Distortion-invariant fractional power fringe-adjusted joint transform correlation,” *Optical Engineering*, vol. 37, no. 1, pp. 138–143, 1998.
[7] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms”, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, pp. 62-66, 1979.
[8] K.V. Asari, T. Srikanthan, S. Kumar, and D. Radhakrishnan, "A pipelined architecture for image segmentation by adaptive progressive thresholding," *Journal of Microprocessors and Microsystems* , vol. 23, no. 8-9, pp. 493-499, December 1999.

[9] S. Kumar, K.V. Asari, and D. Radhakrishnan, "Real-time automatic extraction of lumen region and boundary from endoscopic images," *IEE Journal of Medical & Biological Engineering & Computing*, vol. 37, no. 5, pp. 600-604, September 1999.

[10] M. S. Alam, A. Bal, E. H. Horache, S. F. Goh, C. H. Loo, S. P. Regula and A. Sharma, "Metrics for evaluating the performance of joint-transfor-correlation-based target recognition and tracking algorithms," Optical Engineering, vol. 44, no. 6, pp. 067005-1 – 067005-12, 2005.