Adaptive SDF filters for recognition of partially occluded objects

J. Ángel González-Fraga,[a] Vitaly Kober,[a] Josué Álvarez-Borrego[b]

[a]Department of Computer Science, CICESE
Km 107 Carretera Tijuana-Ensenada, Ensenada, B.C. 22860, México

[b]Optics Department, CICESE
Km 107 Carretera Tijuana-Ensenada, Ensenada, B.C. 22860, México

ABSTRACT

One of the main problems in visual image processing is incomplete information owing an occlusion of objects by other objects. Since correlation filters mainly use contour information of objects to carry out pattern recognition then conventional correlation filters without training often yield a poor performance to recognize partially occluded objects. Adaptive correlation filters based on synthetic discriminant functions for recognition of partially occluded objects imbedded into a cluttered background are proposed. The designed correlation filters are adaptive to an input test scene, which is constructed with fragments of the target, false objects, and background to be rejected. These filters are able to suppress sidelobes of the given background as well as false objects. The performances of the adaptive filters in real scenes are compared with those of various correlation filters in terms of discrimination capability and robustness to noise.

Keywords: Pattern recognition, adaptive correlation filters, partially occluded objects.

1. INTRODUCTION

Recently, a lot of efforts utilizing correlation filters were undertaken to improve pattern recognition. Various features of an object to be recognized such as the shape, texture, color etc. may effect to quality of recognition. Actually, the shape of the object to be recognized is very important for its identification with the help of a correlation filter. So when the object is partially occluded by other objects, pattern recognition becomes a very difficult task. Several methods for recognition of partially occluded object were proposed.[1-7] Detection of an object and estimation of its exact position are two important tasks in pattern recognition. Correlation filters[3-17] are powerful techniques to carry out these two tasks because they are shift invariant and they locate an object in an input scene by searching the peaks on the correlation plane.

In this paper we proposed a new method to synthesize an adaptive filter for recognition of partially occluded objects embedded in a realistic background. The method is based on synthetic discriminant functions (SDF). Recently, new adaptive correlation filters based on synthetic discriminant functions (SDF) for pattern recognition were proposed.[14-15] The filters are able to suppress sidelobes of a given background as well as false objects. In other words, the filters perform direct control over the whole correlation plane. In this paper we exploit the same idea to design adaptive filters for recognition of objects occluded by similar objects and imbedded into realistic scenes.

Further author information –
J. A. G-F.(correspondence): email: jagonzal@cicese.mx; phone: +52-646-1750500; fax: +52-6-1750593
V.K.:email: vkober@cicese.mx, J.A-B.: email: josue@cicese.mx

Application of Digital Image Processing XXIX, edited by Andrew G. Tescher,
Proc. of SPIE Vol. 6312, 63121G, (2006)

The proposed correlation filters are adaptive to an input scene, which contains fragments of a target, false objects, and background to be rejected. They also yield high correlation peaks corresponding to pieces of the target. The performance of the proposed adaptive filters is compared with that of various correlation filters in terms of discrimination capability and robustness to input additive noise.
In Section 2, we review classical correlation filters. A new design algorithm is given in Section 3. Computer simulation results are presented and discussed in Section 4. Section 5 summarizes our conclusions.

2. CLASICAL CORRELATION FILTERS

2.1. Matched filter

A basic correlation filter is the matched spatial filter[8] (MSF). This filter is optimal with respect to the signal-to-noise ratio at the filter output when an input signal is in presence of additive white noise. A drawback of the MSF in optical implementation is its low light efficiency.

2.2. Phase-only filter

A filter with maximum light efficiency is the phase-only filter (POF) introduced by Horner and Gianino.[9] A drawback of the POF is its poor discrimination capability when a low-contrast target is embedded into a complicated background scene.[10] The transfer function of the conventional POF is given by

$$H_{POF}(u,v) = \frac{T^*(u,v)}{|T(u,v)|} = \begin{cases} \exp\left(-i\Phi_t(u,v)\right), & if \ \ |T(u,v)| \neq 0 \\ 0 & \text{otherwise} \end{cases}, \tag{1}$$

where T(u,v), Φt(u,v) are the Fourier transform and the phase distribution of the target, respectively. The asterisk denotes complex conjugate.

2.3. Inverse filter

The frequency response of the inverse filter (IF) is given by

$$H_{IF}(u,v) = \frac{T^*(u,v)}{|T(u,v)|^2}, \tag{2}$$

This filter provides sharper peaks than with the previous filters. When the input scene is equal to the target, the resultant amplitude distribution is a constant. Thus, in the correlation plane a delta like function is obtained.
Campos et al.[3] suggested the use of a modified inverse filter (MIF) for recognition of partially occluded objects, with frequency response as follows,

$$H_{MIF}(u,v) = \frac{T^*(u,v)}{|T(u,v)|^2 + \sigma^2}, \tag{3}$$

The MIF is a special case of the trade-off filters proposed by Refregier.[11] A suitably chosen quantity σ^2 is used to avoid the amplification of noise when $|T(u,v)|^2$ is small. in the denominator.

2.4. Conventional SDF

When deformations of an object to be recognized are not clearly described then a set of training images consisting of distorted object versions (patterns) can be used to improve pattern recognition. In this case the designed SDF filter is a linear combination of MSFs for different patterns. The coefficients of the linear combination are chosen to satisfy a set of constraints on the filter output requiring a prespecified value for each pattern used in the filter synthesis. The SDF filters use a set of training images to synthesize a template that yields prespecified correlation outputs in response to training images. The conventional SDF filter can be expressed in space domain as[12]

$$\mathbf{h}_{SDF} = \mathbf{X}(\mathbf{X}^{T}\mathbf{X})^{-1}\mathbf{u}\,, \tag{4}$$

where superscript T means transpose.

Suppose that N training images belong to a true class, each image contains d pixels. We convert the 2D arrays of the images into the 1D column vector by lexicographical ordering. These vectors are the columns of the matrix $\mathbf{X}$ of size $d{\times}N$. The column vector $\mathbf{u}$ contains N elements, which are prespecified values of output correlation peaks corresponding to each training image. The filter $\mathbf{h}_{\mathrm{SDF}}$ is a column vector with d elements and the 2D correlation filter is obtained by reordering the column vector back to a 2D array. A detailed information about the filter design is given in ref. 12.

While the conventional SDF filter produces prespecified values at the correlation output, it also yields large sidelobes. These sidelobes sometimes are larger than the prespecified peaks values on the correlation plane. This leads to misclassification. To reduce sidelobes Mahalanobis et.al.[13] developed the minimum average correlation energy filter (MACE).

2.5. MACE filter

The MACE filter minimizes the average correlation energy of the correlation outputs due to the training images while simultaneously satisfying the correlation peak constraints at the origin. The effect of minimizing the average correlation energy is that the resulting correlation planes would yield values close to zero everywhere except at the location of a trained object, where it would produces an intense peak. In the Fourier domain the MACE filter can be expressed in vector form as follows:[13]

$$\mathbf{h}_{MACE} = \mathbf{D}^{-1}\mathbf{X}\left(\mathbf{X}^{+}\mathbf{D}^{-1}\mathbf{X}\right)^{-1}\mathbf{u}\,, \tag{5}$$

where superscript + means conjugate transpose.

Let us suppose that we have N training images of a true class, each image contains d pixels. We perform 2D Fourier Transform (FT) on these images and convert the 2D FT arrays into the 1D column vectors by lexicographic ordering. These vectors are the column vectors of the matrix $\mathbf{X}$ of size $d{\times}N$ given in Eq. (5). The column vector $\mathbf{u}$ with N elements contains prespecified correlation values of the training images and the $d{\times}d$ diagonal matrix $\mathbf{D}$ contains along its diagonal the average power spectrum of the training images (i.e., average of the magnitude squares of the columns of $\mathbf{X}$). Note that the synthesized filter $\mathbf{h}_{\mathrm{MACE}}$ is a complex column vector with d elements and the correlation filter 2D $H(u,v)$ is obtained by reordering the column vector back to a 2D array.

Once the correlation filter $H(u,v)$ is calculated, the input image $s(x,y)$ is tested for object recognition by forming the correlation output in the following way:

$$c(x,y) = \Im^{-1}\left\{\Im\left[s(x,y)\right]\cdot H(u,v)\right\}\,, \tag{6}$$

where $\Im$ and $\Im^{-1}$ denote the Fourier pair transforms, (u,v) are the coordinates in the frequency domain, and (x,y) are

the coordinates in the spatial domain.
Figure 1 shows a block-diagram of recognition of partially occluded objects with training.

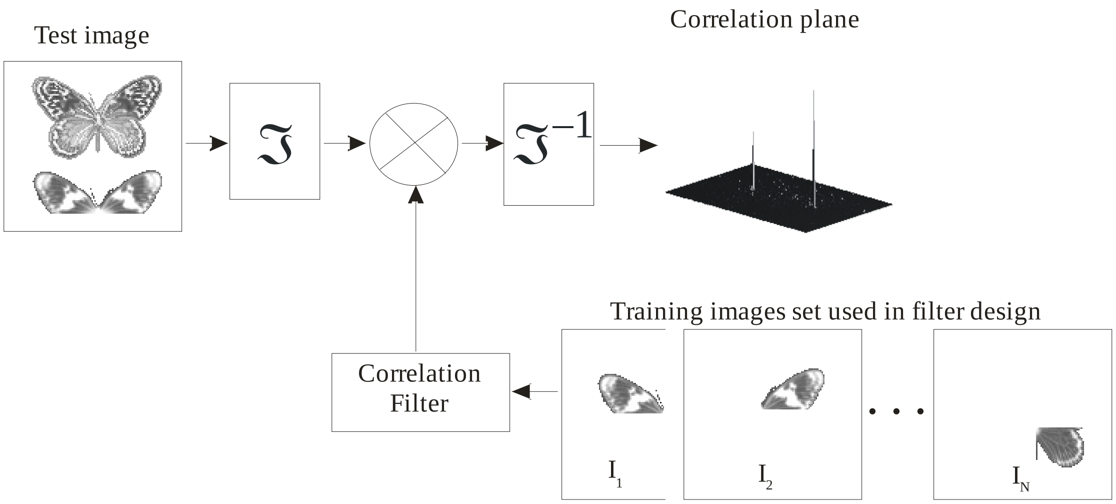


Figure 1: Block diagram showing the correlation process using correlation filters with training in the problem of recognizing partially occluded objects.

3. FILTER DESIGN AND EXPERIMENTAL RESULTS

In order to design a correlation filter that ensures a high correlation peak corresponding to fragments of the target, it is necessary to reduce correlation function levels at all false peaks except that of at the origin of the correlation plane. For a given object to be recognized, false objects, and a background to be rejected, we propose an iterative algorithm, which in each iteration suppresses the highest sidelobe peak, and, therefore, monotonically increases the value of the discrimination capability until a prespecified value is reached. The discrimination capability is formally defined[10] as ability of a filter to distinguish a target among other different objects. If a target is embedded into a background, which contains false objects, then the DC can be expressed as follows:

$$DC = 1 - \frac{\left| C^B(0,0) \right|^2}{\left| C^T(0,0) \right|^2}, \tag{7}$$

where C^B is the maximum in the correlation plane over the background area to be rejected, and C^T is the maximum in the correlation plane over the area of target position. The area of target position is determined in the close vicinity of the actual target location. The background area is complementary to the area of target position. The further accurate estimation of the target location with correlation filters can be carried out.[16] In our computer simulations the area of target position is chosen as a circle with the origin at the actual target location and with the area of about 2% of the target area. Negative values of the DC indicate that a tested filter fails to recognize the target, which means that the energy of the cross correlation is higher than that of the autocorrelation.
We are interested in a correlation filter that identifies an available portion of a target with a high discrimination capability in cluttered and noisy input scenes. Actually, in this case conventional correlation filters yield a poor performance. With the help of adaptive SDF filters a given value of the DC can be achieved. The algorithm of the filter design requires the knowledge of a background image and a target *t(x,y)*. The latter is split into *N* independent fragments

$t_i(x,y)$,

$$t(x,y)=\sum_{i=1}^{N} t_i(x,y)\,. \tag{8}$$

It is assumed that al least one of the fragments will respond to the corresponding fragment of the target embedded into a cluttered background. A test target used in our computer simulations and its portioning onto N=4 fragments are shown in Fig. 2.

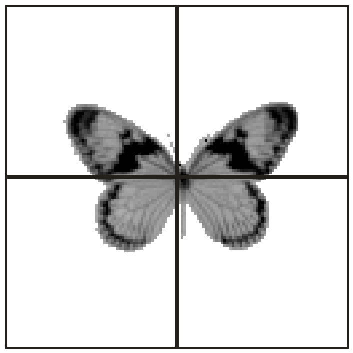


Figure 2. Target split into four fragments.

So we are looking for the fragments of the target with unknown locations in a known input scene background. The background can be described either stochastically, for instance, it can be considered as a realization of a stochastic process or deterministically, that can be a picture. The background can also contain false objects with unknown locations. The first step is to carry out correlation between the background and a basic SDF filter, which is trained only with $\{t_i(x,y;\ i=1,2,..,N\}$. Next, the maximum of the filter output is set as the origin, and around the origin we form a new object to be rejected from the background. This object has the region of support equals to that of the target. The created object is added to the false class of objects. Now two-class recognition problem is utilized to design a new SDF filter; that is, the true class contains only the fragments of the target and the false class consists of the false class objects. The described iterative procedure is repeated untill a given value of the DC is obtained. Finally, note that if other objects to be rejected are known, they can be directly included into the false class and used for the design of the adaptive filter. A block-diagram of the procedure is shown in Fig. 3. The proposed algorithm consists of the following steps.

1. Design a basic adaptive filter (ASDF) as a SDF filter, which is trained only with the fragments $t_i(x,y)$ of the target.

2. Carry out correlation between the background and the ASDF.

3. Calculate the DC.

4. If the value of the DC is greater or equal to a desired value then the filter design procedure is finished, else go to the next step.

5. Create a new object to be rejected from the background. The origin of the object is at the highest sidelobe position in the correlation plane. The object is included into the false class of objects.

6. Design a new ASDF utilizing two-class recognition problem. The true class contains the fragments $t_i(x,y)$ of the target and the false class consists of the false class objects created in step 5. Go to step 2.

At each iteration the algorithm chooses among all sidelobes a such peak to be suppressed in the next step to ensure a monotonically increasing behavior of the DC versus the iteration index. As a result of the procedure the adaptive composite filter is synthesized. The performance of the filter in recognition process is expected to be close to that of in the synthesis process. Obviously, since the problem is to recognize fragmented objects, the DC in test scenes will be lower than that in the training process. Extensive computer simulations showed that for complicated input scenes with real and stochastic cluttered backgrounds the number of iterations needed to achieve the value of the DC higher than 0.9 is about ten. Note that other iterative optimization algorithms frequently used for composite filter design such as simulated annealing and genetic algorithms require significantly higher computational complexity.[17]

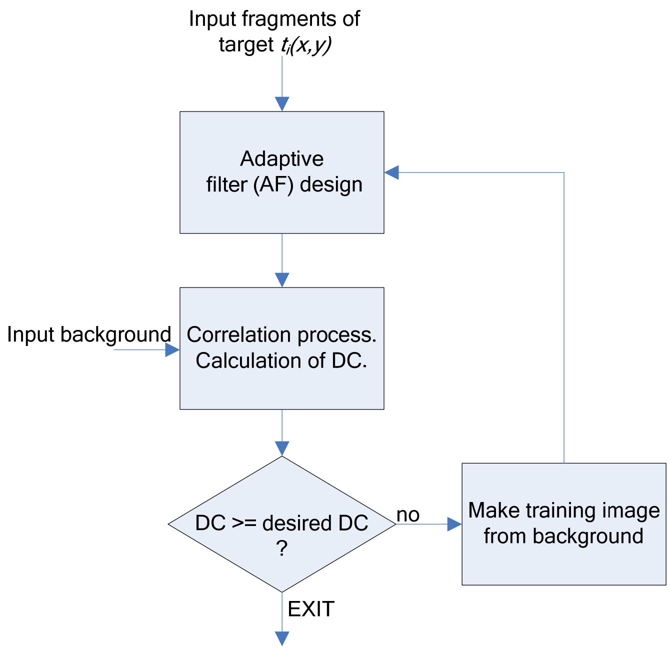


Figure 3. Block-diagram of the iterative algorithm for design of adaptive SDF filters.

4. COMPUTER SIMULATIONS

In this section computer simulation results obtained with adaptive SDF filters are presented. The results are compared with those of the POF, the MIF and the MACE filters. The target used in our experiments is the butterfly shown in Fig. 4(a). Various parts of the target to be recognized are shown in Fig. 4(b)-(h) and these objects are used for generate the test scenes. The false butterfly has the same shape but with different information contents as the target. The energies of the complete objects are the same. The size of all images used in the experiments is 256×256 pixels and the signal range is [0-255]. The mean value and the standard deviation over the target area are 108 and 51, respectively. The size of the target is about 90×60 pixels. In the experiments we use a spatially inhomogeneous real background shown in Fig. 5. The mean value and the standard deviation of the background are 89 and 47, respectively. The mean and standard deviation of the false object are 103 and 53, respectively.

An ASDF filter was designed to recognize the target fragments and to reject the false butterfly as well as the background. Figure 6 shows the performance of the ASDF in the design process in terms of the DC versus the iteration index. After 30 iterations the obtained ASDF yields DC=0.97. This means that a high level of control over the correlation plane for the input scene constructed from the background and the fragments of the target can be achieved. Next, we test the ASDF when the input scene contains the objects shown in Fig. 4(b)-(h). These objects are imbedded into the background at arbitrary coordinates. The performance of the ASDF, the POF, the MIF and the MACE with respect to false alarms are given in Table 1. The MACE filter was synthesized with the same four fragments. It can be seen that the proposed filter yields the best performance in terms of false alarms. The number of statistical trials is 30.

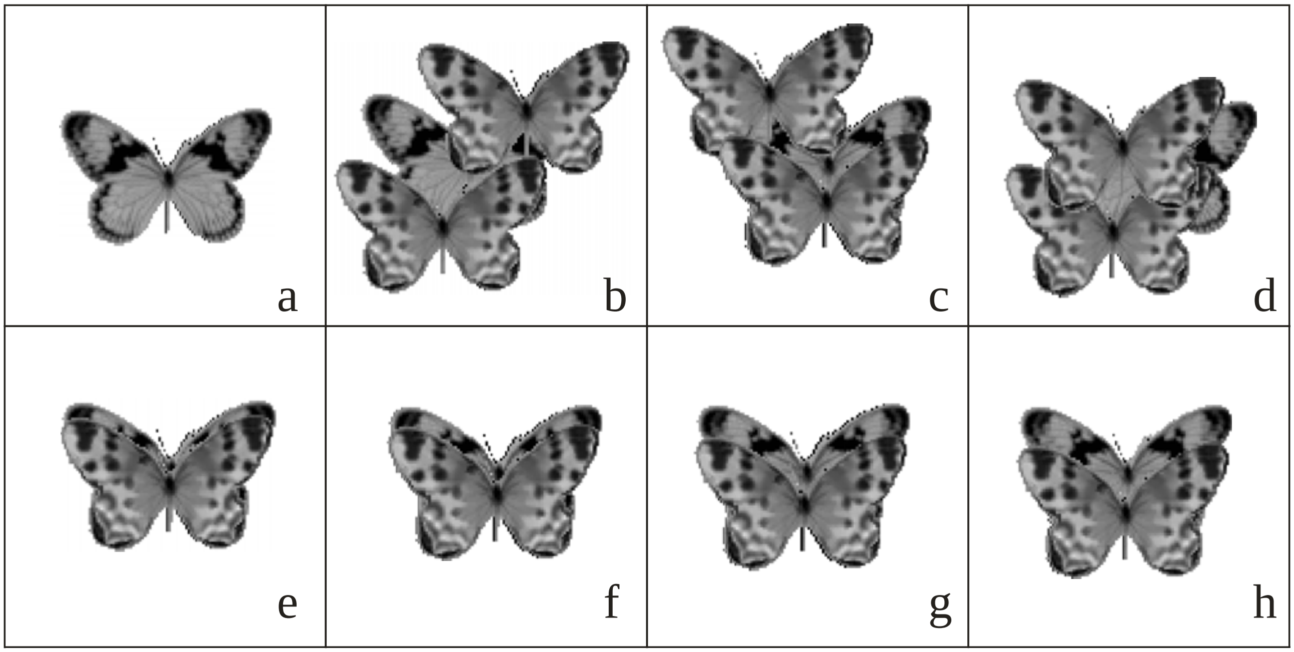


Figure 4. (a) Target, (b-g) target partially occluded by another similar object.

Figure 5. Real background used in experiments.

Fig. 7(a) shows an example of the input scene for the classification problem. The test objects sre shown in Fig. 4. (b). For the used fragment of the target, the POF, the MIF and MACE filters yeild the highest rate of misclassifications. The proposed ASDF is able correctly to recognize the objects. To guarantee statistically correct results, 30 statistical trials of the experiments for different positions of the target and non-targets are performed, and with 95% confidence the DC is equal to 0.62±0.01 for the objects shown in Fig. 3 (b). Fig. 7(b) demonstrates the intensity distribution obtained with the ASDF.

Table 2 gives the performance of the adaptive filters in terms of the DC for classification problem. We also use 30 statistical trials for different positions of the target fragments and non-target.

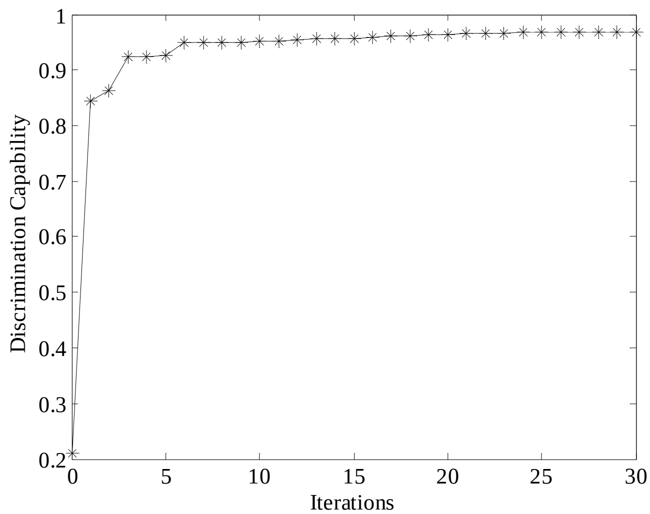
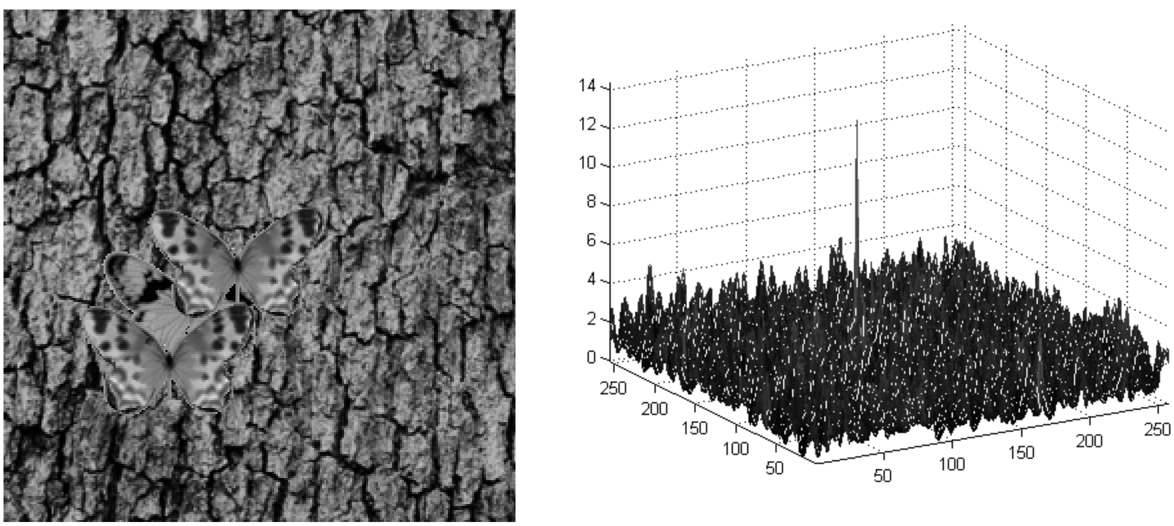


Figure 6. Performance of the adaptive SDF filter in the design process.



(a) (b)

Figure 7. (a) Test scene containing the objects shown in Fig. 4 (b), (b) correlation intensity distribution obtained with the ASDF.

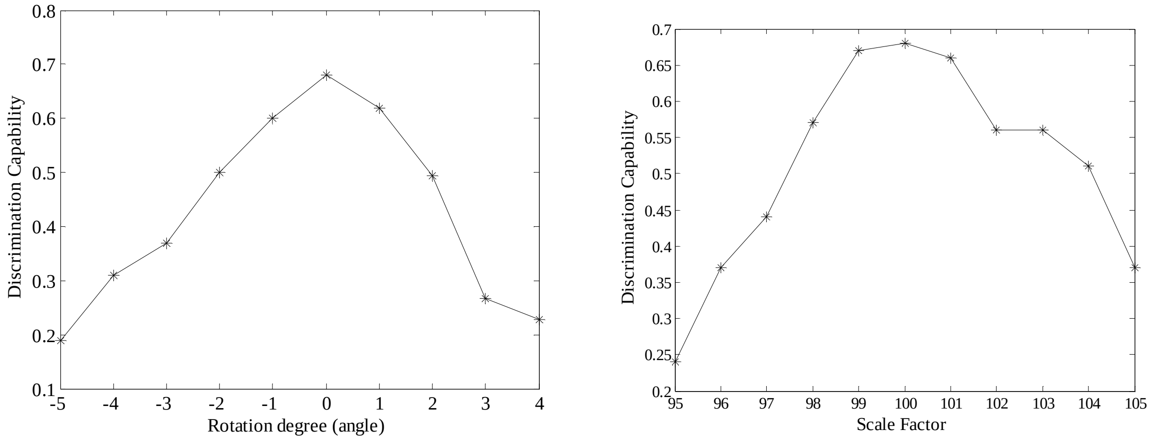
Finally we test robustness of the correlation filters to small distortions. The test scene shown in Fig. 7 (a) is used. Figure 8 presents the tolerance of the filters to rotation and scaling in terms of the discrimination capability. The proposed adaptive filter possesses the best tolerance to small distortions. In contrast, the performance of the POF, the MIF and MACE deteriorates quickly when signal distortions are presented.

Table 1. Performance of different correlation filters in terms of false alarms (misclassifications). Number of trials is 30.

Objects in test scene	POF	MIF	MACE	ASDF
b	7	0	7	0
c	30	23	30	0
d	18	8	22	0
e	6	2	8	0
f	1	0	0	0
g	0	0	0	0
h	0	0	0	0

Table 2. Performance of adaptive filter (ASDF) in terms of DC

Objects in test scene	Available part of the target	Discrimination Capability with confidence of 95% for 30 statistical trials
b	43%	0.62±0.014
c	29%	0.30±0.017
d	29%	0.28±0.016
e	19%	0.26±0.022
f	25%	0.45±0.017
g	38%	0.68±0.011
h	46%	0.70±0.010



(a) (b)

Figure 8. Tolerance of the ASDF filter to small distortions of test scene in Fig. 7 (a) in terms of DC. (a) Rotation, (b) Scaling.

5. CONCLUSIONS

In this work we proposed a new method for recognition of partially occluded objects by means of correlation filters with training. This problem is important in robotics, vision and artificial intelligence. The adaptive SDF filters improve recognition of a fragmented target embedded into a known cluttered background. It was shown that the proposed filter design algorithm with a few training iterations takes the control over the whole correlation plane. The computer simulation results demonstrated superiority in the performance of the proposed filter for pattern recognition comparing with that of the POF and the MACE filters. The proposed filters are able to discriminate noisy similar objects when available information of a target is about 19%.

Acknowledgment

This work was supported by the grant 46154 from the CONACYT.

REFERENCES

1. K.V. Mardia, W. Quian, D. Shah and K.M.A. de Souza, “Deformable template recognition of multiple occluded objects”, *IEEE Trans. on Pattern and Anal. Machine Intell*, 1997, vol. 19, pp. 1035-1042.
2. L. Wiskott and C. Malsburg , “A neural system for the recognition of partially occluded objects in cluttered scenes (A Pilot Study)”, *Int. J. of Pattern. Recognition and Artificial Intelligence*, 1993, vol. 7, pp. 935-948.
3. J. Campos, K. Styczynski, M. J. Yzuel and K. Chalasinska-Macukow, “Recognition of partially occluded objects by correlation methods”, *Optics Communication*, 1994, vol. 106, pp. 45-51.
4. J. Khoury, P. D. Gianino and C. L. Woods, “Adaptive optimal filters for correlators operating on obscured inputs”, *Optical Engineering*, 1998, vol. 37, pp. 112-122.
5. J.A. Gonzalez-Fraga, V. Kober and J. Alvarez-Borrego, “Recognition of partially occluded objects using correlation filters with training”, 2005. *SPIE. The International Society for Optical Engineering, 50th Annual Meeting, Applications of Digital Image Processing XXVIII*, 31 July-03 August in San Diego, California, USA
6. C.E.,Villalobos-Flores, J. Alvarez-Borrego, V. Kober, G. Cristóbal and E. Castro-Longoria, "Study of 21 fragmented fossil diatoms using a digital invariant correlation". 2002. *SPIE-The International Society for Optical Engineering, The International Symposium on Optical Science and Technology*. 7-10 July, Seattle Washington, Vol. 4790, pp. 528-533..
7. B. Javidi, R. Ponce-Diaz and S. Hong, “Three-dimensional recognition of occluded objects by using computational integral imaging”, *Optics Letters*, 2006, vol. 31, pp. 1106-1108.
8. A. VanderLugt, “Signal detection by complex filters”, *IEEE Trans. Inf. Theory*, 1964, Vol. IT-10, pp. 139-145.
9. J.L. Horner, P.D. Gianino, “Phase-only matched filtering,” *Applied Optics*, 1984, vol. 23, pp. 812-816.

10. L.P. Yaroslavsky, “The theory of optimal methods for localization of objects in pictures,” *in progress in Optics XXXII, E. Wolf, Ed., Elsevier,* 1993, pp. 145-201.
11. Ph. Refregier, “Filter design for optical pattern recognition: multicriteria optimization approach”, 1990, *Optics Letters,* vol. 15, pp. 854-856
12. B.V.K. Vijaya-Kumar, “Tutorial survey of composite filter designs for optical correlators”, *Applied Optics*, 1992 vol. 31, pp. 4773-4801.
13. A. Mahalanobis, B.V.K. Vijaya-Kumar and D. Casasent, “Minimum average correlation energy filters”, *Applied Optics*, 1987, vol. 26, pp. 3633-3640.
14. J.A. Gonzalez-Fraga, V.H. Diaz-Ramirez, V. Kober, J. Alvarez-Borrego, “Improving the discrimination capability with an adaptive synthetic discriminant function filter”, *Lecture Notes in Computer Science*, 2005, vol. 3523, pp. 83-90.
15. J.A. Gonzalez-Fraga, V. Kober and J. Alvarez-Borrego, “Adaptive synthetic discriminan function filtres for pattern recognition”, *Optical Engineering,* 2006, vol. 45, pp. 057005-1-057005-10.
16. V. Kober and J. Campos, “Accuracy of location measurement of noisy target in a nonoverlapping background”, *J. Opt. Soc. Am. A.*, 1996, vol. 13, pp. 1653-1666.
17. O. Billet, L. Singher, “Adaptive multiple filtering,” *Optical Engineering*, 2002, 41(1), 55-68.