27

Real-Time Pattern Recognition with Adaptive Correlation Filters

Vitaly Kober, Victor H. Diaz-Ramirez[1], J. Angel Gonzalez-Fraga[2] and Josue Alvarez-Borrego[3]
[1]*Computer Science Department, CICESE,* [2]*Faculty of Sciences, UABC of Ensenada,*
[3]*Optics Department, CICESE,* [3]*Faculty of Ingineering, UABC of Ensenada*
Mexico

1. Introduction

Since the introduction of the matched spatial filter (MSF) (VanderLugt, 1964), many different types of filters for pattern recognition based on correlation have been proposed. One of the reasons of such growing interest to design effective methods of pattern recognition stems from the need to deal with more complex images in various applications of automated image processing and from the need to process large images in real time. For some of the more critical applications, optical or hybrid optodigital techniques allow faster processing of images. This is why our approach in this area is based on correlation filters, which possess good mathematical fundamentals and can be effectively implemented digitally or optodigitally (Moreno et al., 1998).

In pattern recognition two essentially different types of tasks are distinguished: detection of a target and estimation of its exact position. When correlation filters are used, these problems can be solved in two steps. First, the detection is carried out by searching correlation peaks in the filter output, and then coordinates of these peaks are taken as position estimations. The quality of both procedures is limited by the presence of noise in an observed scene. The detection capabilities of correlation filters can be quantitatively expressed in terms of probability of detection errors (false alarms), signal-to-noise ratio, discrimination capability, peak-to-output energy ratio, etc. (Vijaya Kumar & Hassebrook, 1990). Some of the measures can be essentially improved using an adaptive approach to the filter design. According to this concept, we are interested in a filter with good performance characteristics for a given observed scene, i.e., with a fixed set of patterns or a fixed background to be rejected, rather than in a filter with average performance parameters over an ensemble of images. After the detection task has been solved we still faced with small errors of target position estimations due to distortion of the object by noise. The coordinate estimations lie in the vicinity of their actual values. So the target location can be characterized only by means of the variance of measurement errors along coordinates (Kober & Campos, 1996).

One of the most important performance criteria in pattern recognition is the discrimination capability (DC), or how well a filter detects and discriminates different classes of objects. A correlation filter with a minimum probability of anomalous detection errors (false alarms)

Source: Vision Systems: Segmentation and Pattern Recognition, ISBN 987-3-902613-05-9,
Edited by: Goro Obinata and Ashish Dutta, pp.546, I-Tech, Vienna, Austria, June 2007

referred to as the optimal filter (OF) was suggested (Yaroslavsky, 1993). An important feature of the OF is its scene-adaptivity in applications to pattern recognition or target detection because its frequency response takes into account the power spectrum of wrong objects in the observed scene or the background to be rejected. The disadvantage of the OF in optical implementation is its extremely low light efficiency. A filter with maximum light efficiency is the phase-only filter (POF) (Horner & Gianino, 1984). The drawback of the POF is its poor discrimination capability for a low-contrast target embedded into a complicated background scene. An approximation of the OF by means of phase-only filters with a quantization was made (Kober et al., 1994). There, the approximate filters with high light efficiency and discrimination capability close to that of the OF were suggested. When the object to be recognized is in the presence of disjoint background noise, the design of the optimal filter was also obtained (Javidi & Wang, 1994).

It is commonly known that the MSF is very sensitive to small distortions of the object caused by variations in scale, rotation, or point of view. One of the first attempts to overcome the problem of distortion in pattern recognition was the introduction of synthetic discriminant functions (SDFs), (Hester & Casasent, 1980; Casasent, 1984). The SDF filters use a set of training images to synthesize a template that yields a prespecified central correlation output in the response to training images. The main shortcoming of the SDF filters is appearance of sidelobes owing to the lack of control over the whole correlation plane. As a result, the SDF filters often possess a low discrimination capability. A partial solution of this problem was suggested (Mahalanobis et al., 1987). They proposed to control over the whole correlation plane by producing sharp correlation peaks for easy detection of the target as well as by minimizing the average correlation energy to suppress the presence of extraneous correlation peaks. However, these filters are not tolerant to input noise. They perform control over false alarms by an indirect way, and, finally, they are more sensitive to interclass variations than other composite filters (Billet & Singher, 2002).

This chapter treats the problems of real-time pattern recognition exploiting adaptive distortion-invariant correlation filters (González-Fraga et al., 2006; Diaz-Ramirez et al., 2006; Kober et al., 2006). The distinctive feature of the proposed methods is the use of an adaptive approach to the filters design. Specifically, we shall look at two problems: detection of known objects possessing small geometric distortions and corrupted with additive sensor's noise, and implementation of the designed filters in an optodigital setup.

The first problem is to decide on presence or absence of a distorted object. New adaptive composite filters for reliable recognition of the object in a cluttered background are presented. The information about an object to be recognized, false objects, and a known background to be rejected is utilized in iterative training procedure to design a correlation filter with a given value of discrimination capability. The synthesis of the adaptive filters also takes into account additive sensor's noise by training with a noise realization. Therefore, the filters may possess a good robustness to the noise.

The second problem concerns real-time implementation of the adaptive correlation filters. For some of the more critical applications, optical or hybrid optodigital techniques allow faster processing of images. The advantage of optical systems over computers lies in inherent ability of optical systems to process data in a parallel way. For instance the classical optical correlator allows to perform fully parallel matched filtering over an input scene containing multiple patterns. Recent progress in optical spatial light modulators gives new opportunities for creation of optodigital systems. Such modulators can be addressed

electronically that allows rapidly and flexibly change the object or the filter for real-time applications. We implemented the adaptive filters in a hybrid system using the joint transform correlator scheme. The hybrid system additionally takes into account real characteristics of used optoelectronics devices. Computer simulation and experimental results are provided and disscussed.

2. Adaptive Digital Systems

2.1. Conventional correlation filters

Consider the problem of detecting the presence and location of a known distorted target in an observed scene using the correlation operation. The correlation can be effectively implemented in a computer with the help of the fast Fourier transform. When the correlation output is obtained then coordinates of the correlation peaks can be taken as position estimations of a target.

A basic correlation filter is the MSF whose impulse response is the flipped version of a reference object. This filter is optimal with respect to the signal-to-noise ratio at the filter output when an input signal is on presence of additive white noise. A drawback of the MSF in optical implementation is its low light efficiency. A filter with maximum light efficiency is the POF. The transfer function of a basic POF (Horner & Gianino, 1984) is given by

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

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

The transfer function the OF can be approximated in the Fourier domain as

$$H_{OF}(u,v)=\frac{T^{*}(u,v)}{\left|T(u,v)\right|^2+\left|S(u,v)\right|^2}, \tag{2}$$

where $S(u,v)$ is the Fourier transform of the input scene (Yaroslavsky, 1993).

The performance of conventional correlation filters degrades rapidly with image distortions. An attractive approach to distortion-invariant pattern recognition is based on SDF filters. These filters (called composite filters) use a set of training images (patterns), which are sufficiently descriptive and representative for expected distortions. A basic SDF filter is a linear combination of MSFs for different patterns (Casasent, 1984). The coefficients of the linear combination are chosen to satisfy a set of constraints on the filter output. Two different recognition problems can be solved with the composite filters.

Intraclass Recognition Problem

Let $\{t_i(x,y);\ i=1,2,\ldots,N\}$ be a set of (linearly independent) training images each with d pixels. The SDF filter function $h(x,y)$ in the spatial domain can be expressed as a linear combination of a set of reference images, i.e.,

$$h(x,y)=\sum_{i=1}^{N}a_i t_i(x,y), \tag{3}$$

where $\{a_i;\ i=1,2,\ldots,N\}$ are weighting coefficients, and they are chosen to satisfy the following conditions:

$$t_i \otimes h = q_i \ . \tag{4}$$

Here the symbol $\otimes$ denotes the correlation, and $\{q_i;\ i=1,2,\ldots,N\}$ are prespecified values in the correlation output at the origin for each training image.
Let **R** denote a matrix with N columns and d rows (number of pixels in each training image), where its ith column is given by the vector version of $t_i(x,y)$. Let **a** and **u** represent column vectors of $\{a_i\}$ and $\{q_i\}$, respectively. We can rewrite Eqs. (3) and (4) in matrix-vector notation as follows:

$$\mathbf{h} = \mathbf{R}\mathbf{a} \ , \tag{5}$$

$$\mathbf{q} = \mathbf{R}^{+}\mathbf{h} \ , \tag{6}$$

where superscript $^+$ means conjugate transpose.
By substituting Eq. (5) into Eq. (6) we obtain

$$\mathbf{q} = (\mathbf{R}^{+}\mathbf{R})\mathbf{a} \ . \tag{7}$$

The (i,j)th element of the matrix $\mathbf{Q}=(\mathbf{R}^{+}\mathbf{R})$ is the value at the origin of the cross-correlation between the training images $t_i(x,y)$ and $t_j(x,y)$. If the matrix **Q** is nonsingular, the solution of the equation system is given by

$$\mathbf{a} = (\mathbf{R}^{+}\mathbf{R})^{-1}\mathbf{q} \ , \tag{8}$$

and the filter vector is

$$\mathbf{h}_{SDF} = \mathbf{R}(\mathbf{R}^{+}\mathbf{R})^{-1}\mathbf{q} \ . \tag{9}$$

The SDF filters with equal output correlation peaks can be used for intraclass distortion-invariant pattern recognition, i.e., detection of distorted patterns belonging to the true-class of objects. This can be done by setting all elements of **q** to unity, i.e.,

$$\mathbf{q} = [1\ 1\ \ldots\ 1]^{T} \ . \tag{10}$$

Multiclass Recognition Problem
Assume that there are distorted versions of a reference object and various classes of objects to be rejected. For simplicity, we consider two-class recognition problem. Thus, we design a correlation filter to recognize training images from one class (called true class) and to reject training images from another class (called false class). Suppose that there are M training images from the false class $\{p_i(x,y);\ i=1,2,\ldots,M\}$. According to the SDF approach, the composite image $h(x,y)$ is a linear combination of all training images $\{t_1(x,y),\ldots, t_N(x,y), p_1(x,y),\ldots, p_M(x,y)\}$. The both intraclass recognition and interclass discrimination problems can be solved by means of SDF filters. We can set the filter output $\{q_i=1;\ i=1,2,\ldots,N\}$ for the true class objects and $\{q_i=0,\ i=N+1,N+2,\ldots,N+M\}$ for the false class objects, i.e.,

$$\mathbf{q} = [1\ 1\ \ldots\ 1\ 0\ 0\ \ldots\ 0]^{T} \ . \tag{11}$$

Using the filter given in Eq. (9), we expect that the central correlation peak will be close to unity for the true class objects and it will be close to zero for the false class objects. Obviously, the preceding approach can be easily extended to any number of classes to be discriminated. Note that this simple procedure is the lack of control over the full correlation output because we are able to control only the correlation output at the location of cross-correlation peaks. Therefore, other sidelobes (false peaks) may appear everywhere on the correlation plane. To reduce the sidelobes, a composite correlation filter (called MACE filter) with a sharp correlation peak at the output was proposed (Mahalanobis et al., 1987). The MACE filter is synthesized in the frequency domain as follows:

$$\mathbf{H}_{MACE} = \mathbf{D}^{-1}\mathbf{P}(\mathbf{P}^{+}\mathbf{D}^{-1}\mathbf{P})^{-1}\mathbf{q}\,, \tag{12}$$

where **D** is a diagonal matrix, **P** is a matrix with N columns and d rows, where its ith column is given by the vector version of $T_i(u,v)$ (Fourier transform of $t_i(x,y)$). The entries along the diagonal are obtained by averaging the power spectrum of each image $(|T_i(u,v)|^2; i=1,2,\ldots,N)$ and then scanning the average from left to right, and from top to bottom.

2.2. Design of Adaptive Correlation Filters

To achieve good recognition of the target it is necessary to reduce correlation function levels at all false peaks except at the origin of the correlation plane, where the constraint on the peak value must be met. For a given object to be recognized, false objects, and a background to be rejected, it can be done with the help of an iterative algorithm. At each iteration, the algorithm suppresses the highest sidelobe peak and therefore monotonically increases the value of discrimination capability until a prespecified value will be reached. The discrimination capability is formally defined as ability of a filter to distinguish a target among other different objects. If a target is embedded into a background that contains false objects, then the DC can be expressed as follows:

$$DC = 1 - \frac{\left|C^{B}(0,0)\right|^{2}}{\left|C^{T}(0,0)\right|^{2}}, \tag{13}$$

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. Negative values of the DC indicate that a tested filter fails to recognize the target.

We are interested in a correlation filter that identifies a target with a high discrimination capability in cluttered and noisy input scenes. Actually in this case, conventional correlation filters yield a poor performance (Javidi & Wang, 1992). With the help of adaptive composite filters, a given value of the DC can be achieved. The algorithm of the filter design requires knowledge of the background image. Thus, we are looking for the target with unknown location in the 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, which 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 initially trained only with the target. 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 described in Section 2.1 is utilized to design a new SDF filter; that is, the true class contains only the target and the false class consists of the false class objects. The described iterative procedure is repeated till 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 adaptive SDF (ASDF) filter. A block-diagram of the procedure is shown in Fig. 1.
The proposed algorithm consists of the following steps:

1. Design ASDF filter as a conventional SDF filter trained only with the target.
2. Carry out correlation between the background and the ASDF filter.
3. Calculate the DC using Eg. (13).
4. If the value of the DC is greater or equal to the 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 filter utilizing two-class recognition problem. The true class contains only the target and the false class consists of the false class objects. Go to step 2.

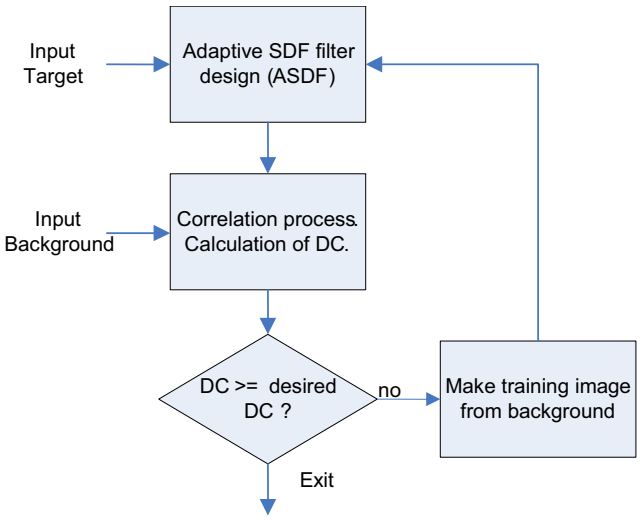


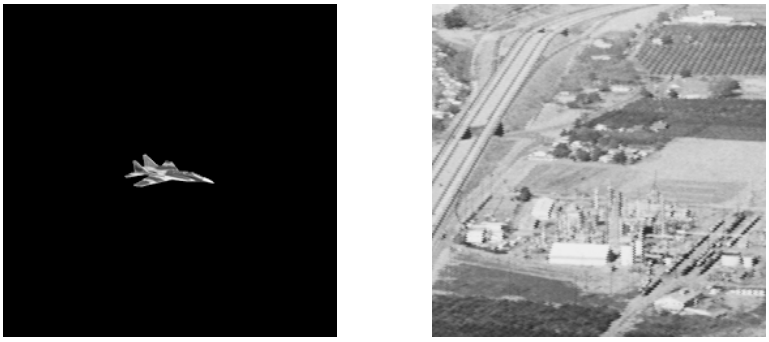
Fig. 1. Block-diagram of the iterative algorithm to design the adaptive SDF filter.

At each iteration, the algorithm chooses among all sidelobes such a peak to be suppressed in next step to ensure a monotonically increasing behavior of the DC versus the iteration index during the filter design. 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. 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 10.

2.3. 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 OF, and the MACE filters. The target is the airplane shown in Fig. 2(a).



(a) (b)

Fig. 2. Test images: (a) target, (b) real background.

The size of all images used in the experiments is 256×256 pixels. The signal range is 0 to 255. The mean value and the standard deviation over the target area are 130 and 42, respectively. The size of the target is about 69×26 pixels. In the first experiment, we use a real spatially inhomogeneous background shown in Fig. 2(b). The mean value and the standard deviation of the background are 104 and 40, respectively.

Figure 3 shows the performance of the adaptive filter in the filter design process in terms of the DC versus the iteration index.

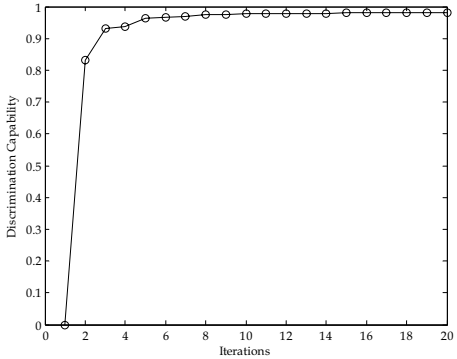


Fig. 3. Performance of the adaptive SDF filter in the design process.

After the first iteration the value of the DC is negative. After 20 iterations, the obtained ASDF filter yields DC=0.982. This means that a high level of control over the correlation plane for an input scene constructed from the background and the target can be achieved. Next, we test the recognition performance with various correlation filters when the target is imbedded into the background at arbitrary coordinates. We carried out 30 statistical trails of

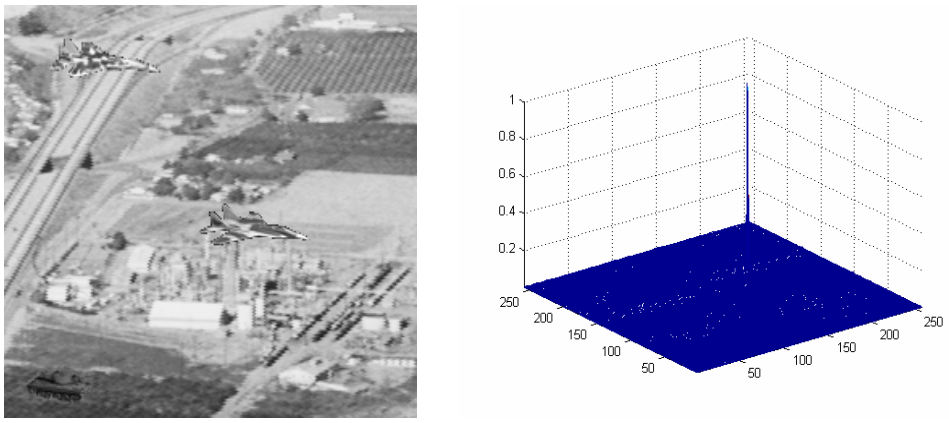
the experiment for different positions of the target. With 95% confidence the performance of the ASDF, the POF, and the OF with respect to the DC are given in line 1 of Table 1.

	POF	OF	ASDF
Scene without false target	0.35±0.22	0.66±0.10	0.95±0.01
Scene with false target	0.27±0.22	0.60±0.12	0.95±0.01

Table 1. Performance of correlation filters in terms of DC.

It can be seen that the proposed adaptive filter yields the best performance in terms of discrimination capability.

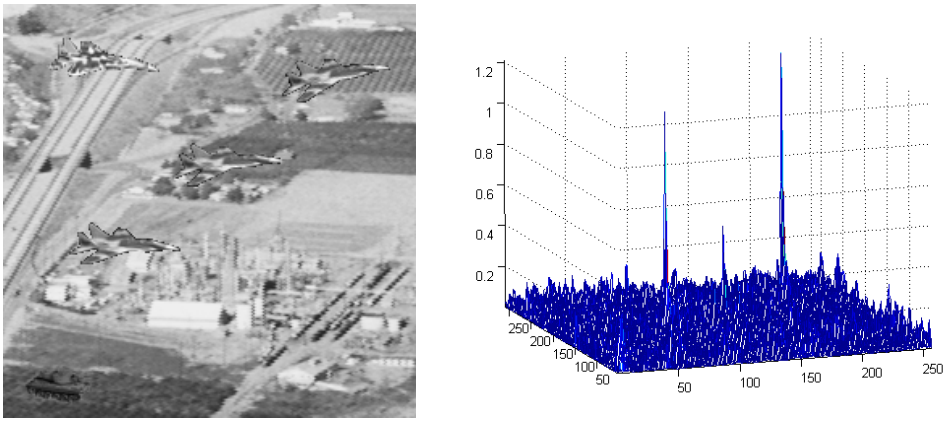
Next, we place a false object into the input scene, as it is shown in Fig. 4 (a). The performance of the correlation filters are given in line 2 of Table 1. One can observe that the ASDF filter yields the best performance with respect to the DC. Figure 4(b) shows the intensity distribution of the correlation plane obtained with the ASDF filter.



(a) (b)

Fig. 4. (a) Test scene, (b) correlation intensity plane obtained with the ASDF filter.

Now we investigate tolerance of the correlation filters to small geometric image distortions. Several methods have been proposed to improve pattern recognition in the presence of such distortions. These methods can be broadly classified into two groups. The first class concerns formally with 2-D scaling and rotation distortions. Such methods include space-variant transforms and circular harmonic functions (Arsenault & Hsu, 1983). The second class of filters uses training images that are sufficiently descriptive and representative of the expected distortions. The proposed method is based on the second approach. In our experiments, geometric distortion by means of rotation is investigated. Distorted versions of the target shown in Fig. 2(a) are used. The step and the range of object rotation are 1 deg and [0, 30], respectively. The ASDF filter is designed with seven versions of the object rotated by 0, 5, 10, 15, 20, and 25 degrees and the background scene shown in Fig. 2(b). After 30 iterations, the obtained ASDF filter yields DC=0.92. The test scene with three targets rotated by 4, 14, and 20 degrees is shown in Fig. 5(a). Figure 5(b) shows the intensity distribution of the correlation plane obtained with the ASDF filter.



(a) (b)

Fig. 5. (a) Test scene, (b) correlation intensity plane obtained with the ASDF filter.

The performance of the ASDF and MACE filters is given in Figs. 6 and 7, respectively. The MACE filter was synthesized with the same objects as the ASDF filter. Note that the conventional SDF filter fails to detect the rotated target in the cluttered background.

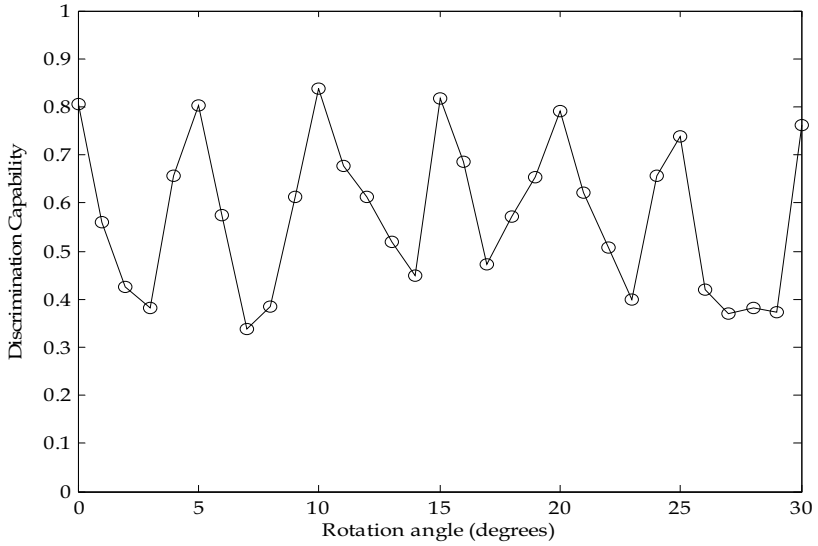


Fig. 6. Tolerance of the ASDF filter to rotation.

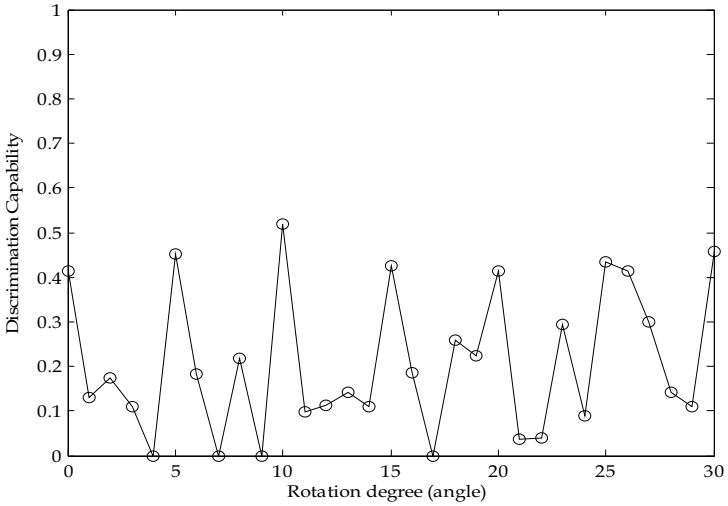
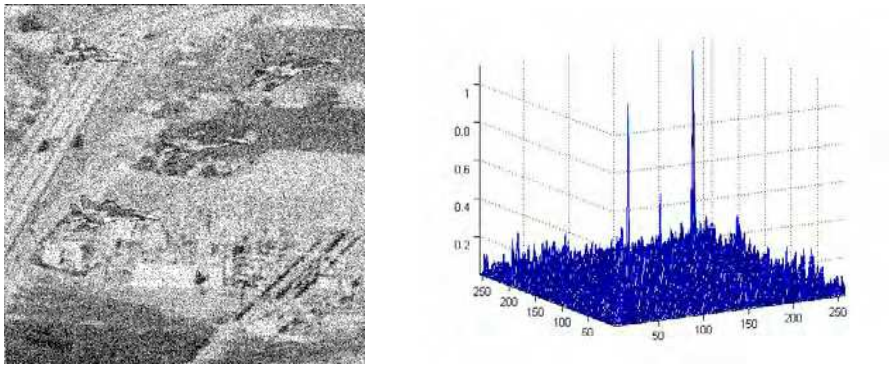


Fig. 7. Tolerance of the MACE filter to rotation.

We can see that the proposed filter possesses much better tolerance to rotation than the MACE filter. The ASDF filter adapts well by training to rotations of the target. Obviously, the preceding approach can be easily extended to any small geometric distortion of a target.
Finally we test robustness of correlation filters to additive sensor's noise that is always present in input scenes. The test scene shown in Fig. 5 (a) is used. The scene is corrupted by additive zero-mean white Gaussian noise while the standard deviation of additive noise is varied. Figure 8(a) shows the input scene corrupted by additive zero-mean white Gaussian noise with the standard deviation of 40. Figure 8(b) shows the intensity distribution of the correlation plane obtained with the ASDF filter.



(a) (b)

Fig. 8. (a) Input scene corrupted by zero-mean additive white noise with a standard deviation of 40, (b) correlation intensity plane obtained with the ASDF filter.

The tolerance of correlation filters to additive noise in terms of the DC is presented in Fig. 9.

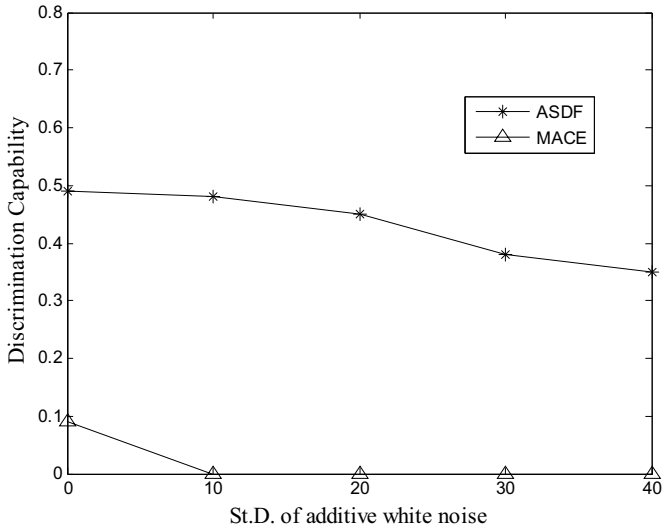


Fig. 9. Tolerance of correlation filters to additive white noise.

Since the synthesis of the ASDF filter takes into account additive noise by training with a noise realization, the filter provides a good robustness to the noise. In contrast, the performance of the MACE filter deteriorates quickly when signal noise fluctuation increases.

3. Adaptive Hybrid Optodigital Systems

Real-time pattern recognition systems based on correlation were vastly investigated in the last decades. This is because correlation filters can be implemented optically or by using hybrid (optodigital) systems exploiting the parallelism inherent in optical systems. These systems are able to carry out the recognition process at a high rate. Hybrid systems with the use of liquid crystal displays (LCDs) as spatial light modulators (SLMs) are flexible. Optodigital systems for real-time pattern recognition can be implemented on the basis of two principal architectures: 4f correlator (4FC) (VanderLugt, 1964) and joint transform correlator (JTC) (Weaver & Goodman, 1966). The advantage of the JTC compared to the 4FC is that the former is less sensitive to misalignments of an optical setup such as scale, horizontal, vertical, and azimuthal differences between the input and frequency planes. The SDF filters for distortion invariant pattern recognition were originally introduced on the basis of the 4FC. Many efforts were made to find an effective implementation of SDF filters with the JTC. In this chapter we describe an iterative algorithm to design adaptive correlation filters for the JTC architecture. The proposed algorithm takes into account calibration lookup tables of all optoelectronic devices used in real experiments.

3.1. Joint Transform Correlators

The JTC introduced in 1966 by Weaver and Goodman is shown in Fig. 10.

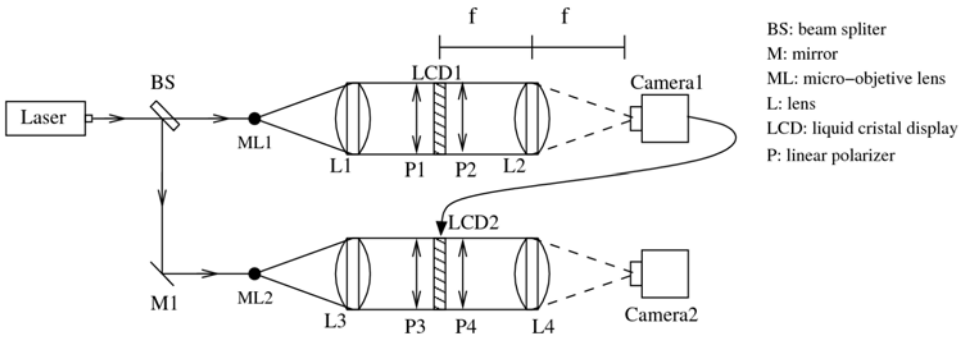


Fig. 10. Block diagram of the classical JTC.

The input plane (joint image) $f(x,y)$ is composed by the scene image $s(x,y)$ alongside the reference image $t(x,y)$ separated by a distance Δ each from origin. The joint image (displayed in LCD1, see Fig. 10) can be written as

$$f(x,y) = s(x,y+\Delta) + t(x,y-\Delta)\,, \tag{14}$$

and its Fourier transform (generated by L1)

$$F(u,v) = S(u,v)\exp(i\Delta v) + T(u,v)\exp(-i\Delta v)\,. \tag{15}$$

The joint power spectrum (captured with CCD camera 1) is given by

$$\begin{aligned} E(u,v) &= |F(u,v)|^2 = |S(u,v)|^2 + |T(u,v)|^2 \\ &+ S(u,v)T^*(u,v)\exp(i2\Delta v) + T(u,v)S^*(u,v)\exp(-i2\Delta v). \end{aligned} \tag{16}$$

Applying the inverse Fourier transform to Eq. (16) (by action of L4) we obtain

$$\begin{aligned} e(x,y) &= s(x,y)\otimes s(x,y) + t(x,y)\otimes t(x,y) \\ &+ s(x,y+2\Delta)\otimes t(x,y+2\Delta) + s(x,y-2\Delta)\otimes t(x,y-2\Delta) \end{aligned}\,. \tag{17}$$

We can see that the autocorrelations of the scene and target images mainly contribute at the origin, whereas the cross-correlation terms, which are the terms of interest, are placed at the distances $\pm 2\Delta$. A drawback of the classical JTC is its low tolerance to geometrical distortions of objects and to noise when objects are embedded in a nonstationary background noise. Assumes that the input image $f(x,y)$ contains the input objects $s(x,y)$ (desired and nondesired) and the non-overlapping background $b(x,y)$:

$$f(x,y) = s(x,y+\Delta) + \tilde{b}(x,y+\Delta) + t(x,y-\Delta), \tag{18}$$

where

$$\tilde{b}(x,y) = w(x-x_0, y-y_0)b(x,y), \tag{19}$$

and (x_0, y_0) are unknown coordinates of the target in the input scene; $w(x-x_0, y-y_0)$ is a binary function defined as

$$w(x-x_0, y-y_0) = \begin{cases} 0, & \text{within the object area} \\ 1, & \text{otherwise} \end{cases}. \tag{20}$$

The joint power spectrum is given by

$$\begin{aligned} |F(u,v)|^2 &= |S(u,v)|^2 + |T(u,v)|^2 + |\tilde{B}(u,v)|^2 \\ &+ \left[T(u,v)S^*(u,v) + T(u,v)\tilde{B}^*(u,v) + S(u,v)\tilde{B}^*(u,v) \right] \exp(i2\Delta v) \\ &+ \left[T^*(u,v)S(u,v) + T^*(u,v)\tilde{B}(u,v) + S^*(u,v)\tilde{B}(u,v) \right] \exp(-i2\Delta v). \end{aligned} \tag{21}$$

Note that the joint power spectrum contains the Fourier transforms with phase the factors of $\exp(\pm i2\Delta v)$ corresponding to the cross-correlation terms between the target and the input objects, the target and the background, and the input objects and the background. The later correlation term severely affects the DC.

To improve the correlation performance of the JTC, several partial solutions were proposed: the nonlinear JTC (Javidi, 1989) and the fringe-adjusted JTC (Alam & Karim, 1993). In the former a nonlinear element-wise transformation of the joint power spectrum is carried out before applying the inverse Fourier transform. In the latter the joint power spectrum is multiplied by the frequency response of a real-valued filter before applying the inverse Fourier transform. These two approaches yield a better performance compared to that of the classical JTC in terms of correlation peak intensity, correlation width, and discrimination capability.

3.2. Adaptive Joint Transform Correlator

We wish to design a JTC that ensures a high correlation peak corresponding to the target while suppressing possible false peaks. To achieve a good recognition of the target, it is necessary to reduce correlation function levels at all sidelobes except at the origin of the correlation plane, where the constraint on the peak value must be meet. For a given object to be recognized and for false objects and background to be rejected, an iterative algorithm is used. At each iteration, the algorithm suppresses the highest sidelobe peak and therefore monotonically increases the value of discrimination capability until a prespecified value is reached. With the help of adaptive SDF filters, a given value of the DC can be achieved.

The first step is to carry out the joint transform correlation between the background and a basic SDF filter, which is initially trained only with the target. Next the intensity 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. The created object is added to the false class of objects. Now a two-class recognition problem is utilized to design a new SDF filter; that is, the true class contains only the target and the false class consists of the false-class objects. The described iterative procedure is repeated until a given value of DC is obtained. Note that if other false-objects are known, they can be directly included in the false class and used for the design of the adaptive filter. A block diagram of the procedure is shown in Fig. 11.

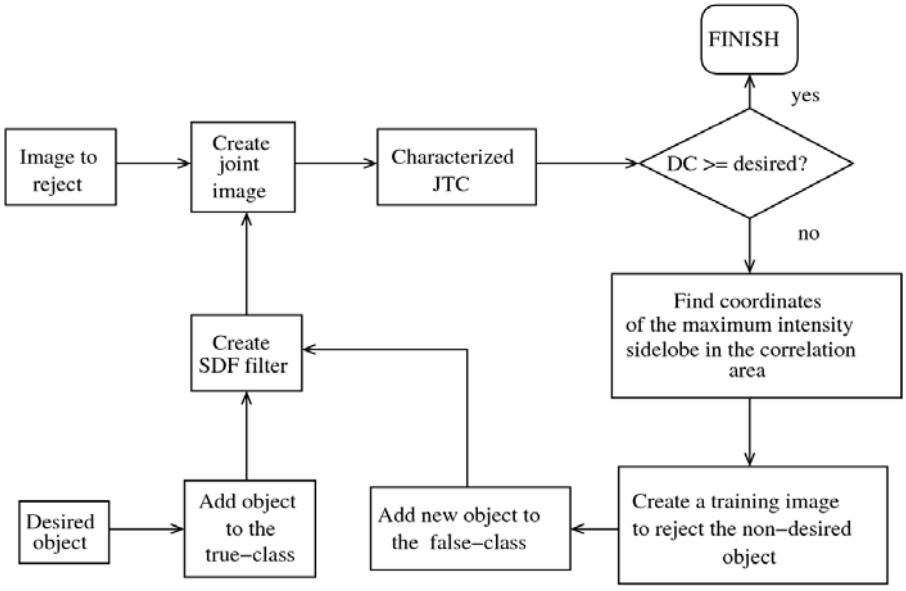


Fig. 11. Block diagram of the iterative algorithm for the design of the adaptive JTC.

The proposed algorithm consists of the following steps:

1. Create a basic SDF filter trained only with the target.
2. Create the input image (see Eq. (14)) by composing the designed SDF filter and the image to be rejected (nondesired objects or a background).
3. Carry out the joint transform correlation including calibration lookup tables of all optoelectronics devices such as a real SLM and a CCD camera.
4. Calculate the DC using Eq. (13).
5. If the value of the DC is greater or equal to the desired value, then the filter design procedure is finished; otherwise, go to the next step.
6. Create a new object to be rejected from the background. The origin of the object is at the highest sidelobe position in the intensity correlation plane. The region of support of the new object is the union of the shapes of all objects involved in the process (desired and non-desired objects). The object is included in the false class of objects.
7. Design a new SDF filter utilizing the two-class recognition problem. The true class contains only the target and the false class consists of the false class objects. Go to step 2.

3.3. Optodigital Implementation

Twisted nematic LCDs are widely used for real-time pattern recognition. Their important characteristics are as follows:

1. They are electrically controlled with standard video signals.
2. They can operate as amplitude-only or phase-only modulators by changing the direction of the polarization vector of the incident light (Lu & Saleh, 1990).
3. They operate at the speed of conventional television standards.

4. They can handle a dynamic range of $[0,255]$ for amplitude modulation and a phase range of $[-\pi,\pi]$ for phase modulation.

In general, the impulse response of SDF filters is a bipolar image. To introduce these kinds of images into spatial light modulators we use two methods.
First method is called bipolar decomposition method. Assume that $h(x,y)$ is a bipolar impulse response:

$$h(x,y)=h^{+}(x,y)-h^{-}(x,y), \tag{22}$$

where

$$h^{+}(x,y)=\begin{cases} h(x,y), & h(x,y)>0 \\ 0, & otherwise \end{cases}, \tag{23}$$

and

$$h^{-}(x,y)=\begin{cases} h(x,y), & h(x,y)\le 0 \\ 0, & otherwise \end{cases}. \tag{24}$$

The intensity cross-correlation between $s(x,y)$ and $h(x,y)$ may be written as follows:

$$\begin{aligned} c(x,y)&=\left|e(x,y)\right|^2=\left|s(x,y)\otimes\left[h^{+}(x,y)\otimes h^{-}(x,y)\right]\right|^2 \\ &=\left|s(x,y)\otimes h^{+}(x,y)\right|^2+\left|s(x,y)\otimes h^{-}(x,y)\right|^2-2\sqrt{\left|s(x,y)\otimes h^{+}(x,y)\right|^2}\sqrt{\left|s(x,y)\otimes h^{+}(x,y)\right|^2}. \end{aligned} \tag{25}$$

It can be seen from Eq. (25), that with the help of decomposition and simple postprocessing, how to obtain the output of the JTC when the reference image has positive and negative values. Note that with the bipolar decomposition method two independent optical correlations are needed.
The second method is referred to as constant addition method. The idea of the method is to transform the input composed bipolar image into an input composed nonnegative image. It can be easily done by adding a bias value to the input bipolar image. Next the joint transform correlation with the input composed nonnegative image is performed. Simple postprocessing is required to obtain the output of the JTC. Note that we need only one optical correlation. The transformed nonnegative joint image can be written as

$$f(x,y)=\tilde{s}(x,y+\Delta)+\tilde{h}(x,y-\Delta), \tag{26}$$

where $\tilde{s}(x,y)=s(x,y)+c$ and, $\tilde{h}(x,y)=h(x,y)+c$, $s(x,y)$ is the scene image, $h(x,y)$ is the bipolar image, and $c=MIN\left[h(x,y)\right]$ is a constant value. The intensity output of the JTC with the new joint image is given by

$$c(x,y) = \left|\tilde{s}(x,y) \otimes \tilde{s}(x,y)\right|^2 + \left|\tilde{h}(x,y) \otimes \tilde{h}(x,y)\right|^2 + \left|\tilde{s}(x,y+2\Delta) \otimes \tilde{h}(x,y+2\Delta)\right|^2 + \left|\tilde{h}(x,y-2\Delta) \otimes \tilde{s}(x,y-2\Delta)\right|^2. \quad (27)$$

The two latter terms of Eq. (27) are the terms of interest. The intensity of the cross-correlation between $s(x,y)$ and $h(x,y)$ can be computed from the intensity of the cross correlation between nonnegative images as follows:

$$\begin{aligned}\left|s(x,y) \otimes h(x,y)\right|^2 &= \left|\left[s(x,y)+c-c\right] \otimes \left[h(x,y)+c-c\right]\right|^2 = \left|\left[\tilde{s}(x,y)-c\right] \otimes \left[\tilde{h}(x,y)-c\right]\right|^2 \\ &= \left|\tilde{s}(x,y) \otimes \tilde{h}(x,y)\right|^2 + \left|\tilde{h}(x,y) \otimes c\right|^2 + \left|\tilde{s}(x,y) \otimes c\right|^2 + \left|c \otimes c\right|^2 \\ &-2\left\{\left[\tilde{s}(x,y) \otimes \tilde{h}(x,y)\right]\left[\tilde{h}(x,y) \otimes c\right]\right\} - 2\left\{\left[\tilde{s}(x,y) \otimes \tilde{h}(x,y)\right]\left[\tilde{s}(x,y) \otimes c\right]\right\} \\ &+2\left\{\left[\tilde{s}(x,y) \otimes \tilde{h}(x,y)\right]\left[c \otimes c\right]\right\} + 2\left\{\left[\tilde{h}(x,y) \otimes c\right]\left[\tilde{s}(x,y) \otimes c\right]\right\} \\ &-2\left\{\left[\tilde{h}(x,y) \otimes c\right]\left[c \otimes c\right]\right\} - 2\left\{\left[\tilde{s}(x,y) \otimes c\right]\left[c \otimes c\right]\right\}.\end{aligned} \quad (28)$$

Further simplifying, we can write

$$\begin{aligned}\left|s(x,y) \otimes h(x,y)\right|^2 &= \left|\tilde{s}(x,y) \otimes \tilde{h}(x,y)\right|^2 + C_1^2 + C_2^2 + C_3^2 - 2\left\{\left[\tilde{s}(x,y) \otimes \tilde{h}(x,y)\right]C_1\right\} \\ &-2\left\{\left[\tilde{s}(x,y) \otimes \tilde{h}(x,y)\right]C_2\right\} + 2\left\{\left[\tilde{s}(x,y) \otimes \tilde{h}(x,y)\right]C_3\right\} + 2\left(C_1C_2\right) - 2\left(C_1C_3\right) - 2\left(C_2C_3\right).\end{aligned} \quad (29)$$

Here, $\tilde{s}(x,y) \otimes \tilde{h}(x,y)$ can be obtained by applying the pointwise square root to the intensity $\left|\tilde{s}(x,y) \otimes \tilde{h}(x,y)\right|^2$, constants $C_1 = \tilde{h}(x,y) \otimes c$, $C_2 = \tilde{s}(x,y) \otimes c$, and $C_3 = c \otimes c$ are computed in the following way:

$$\begin{aligned}C_1 &= \alpha\left[\tilde{h}(x,y) \otimes c\right] = \alpha\left[c \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \tilde{h}(x+\tau_x, y+\tau_y)\,d\tau_x d\tau_y\right] \approx \alpha\left\{c\sum\left[\tilde{h}(x,y)\right]\right\}, \\ C_2 &\approx \alpha\left\{c\sum\left[\tilde{s}(x,y)\right]\right\} \approx \alpha c^2,\end{aligned} \quad (30)$$

where, α is a normalization factor and the symbol "$\sum[\ ]$" denotes the summation of all elements of the image.

3.4. Experimental Results

First we characterized optoelectronics devices such as a twisted neimatic LCD of 800x600 pixels and a monochrome CCD camera of 640x480 pixels. The LCD worked in the amplitude-only modulation regime. Figure 12 shows the experimental calibration lookup table of the intensity response of the LCD captured with the CCD camera.

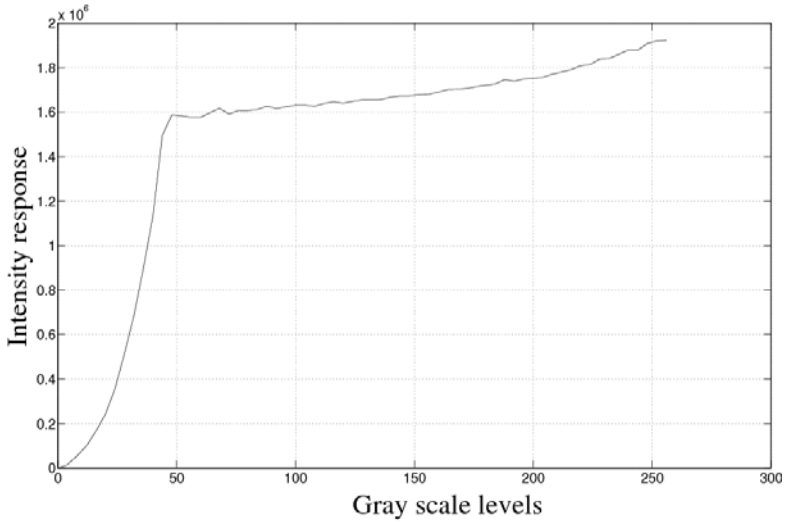
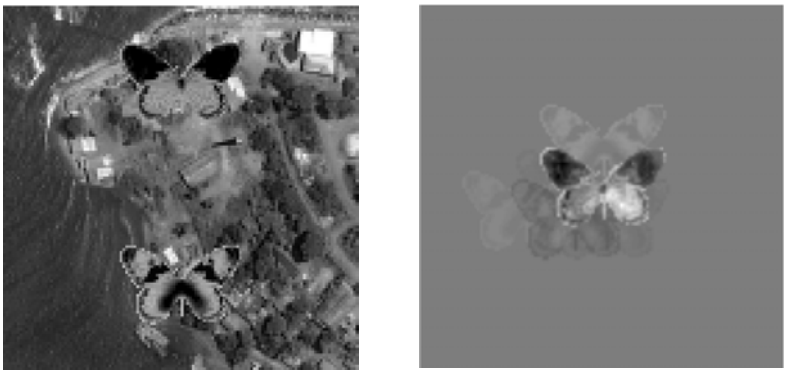


Fig. 12. Intensity response of a twisted neimatic LCD captured with a CCD camera.

It can be seen from Fig. 12 that a gray-scale dynamic range is [0-48]. It is interesting to note that in this range the plot is nonlinear due quantization effects, and it is well approximated with a *k*th-law nonlinearity $\text{Output} = \left(|\text{Input}|^2 \right)^{-k}$ when $k = 0.7$. We used this information in the iterative process of the adaptive JTC design.

The size of the input images used in our experiments is 128 x 128 pixels. The signal rage is [0, 255]. The input scene is shown in Fig. 13 (a).

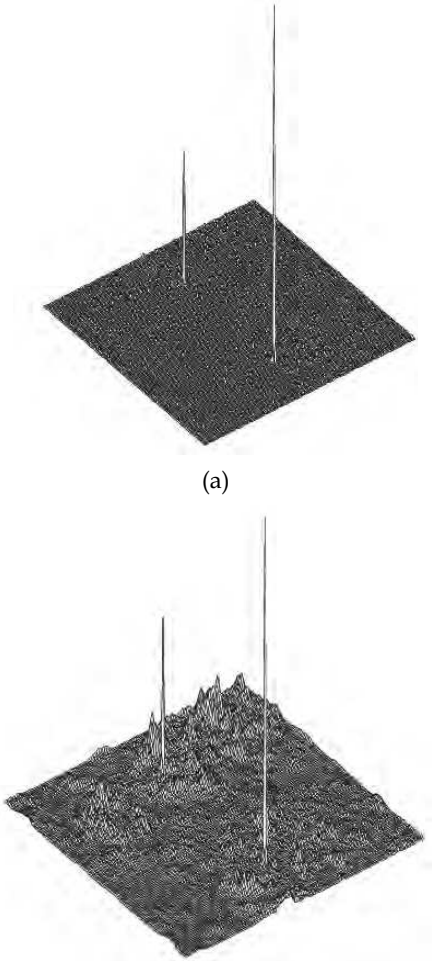


(a) (b)

Fig. 13. (a) Input scene containing two objects with similar shapes but with different information content; (b) bipolar reference image obtained with the proposed method.

The scene contains two objects with a similar shape and size (approximately 44x28 pixels) but with different gray-level contents. The target is the upper butterfly with black- wings. The objects are embedded into an aerial picture at unknown coordinates. The performance

of the adaptive JTC in the design process after eight iterations reaches DC = 0.95. The obtained bipolar reference image is shown in Fig. 13 (b).



(a)

(b)

Fig. 14. Computer simulation results obtained for the input scene in Fig. 13 (a) with: (a) binary JTC, (b) fringe-adjusted JTC.

We compare the performance of proposed adaptive JTC with those of the binary JTC and the fringe-adjusted JTC. The intensity correlation planes obtained with latter two systems are shown in Fig. 14. We see that the binary JTC and the fringe-adjusted JTC fail to discriminate the target against the false object with a similar shape. Next we test digitally the recognition performance with the adaptive JTC. The correlation intensity plane obtained with the adaptive JTC for the input scene in Fig. 13 (a) is shown in Fig. 15.

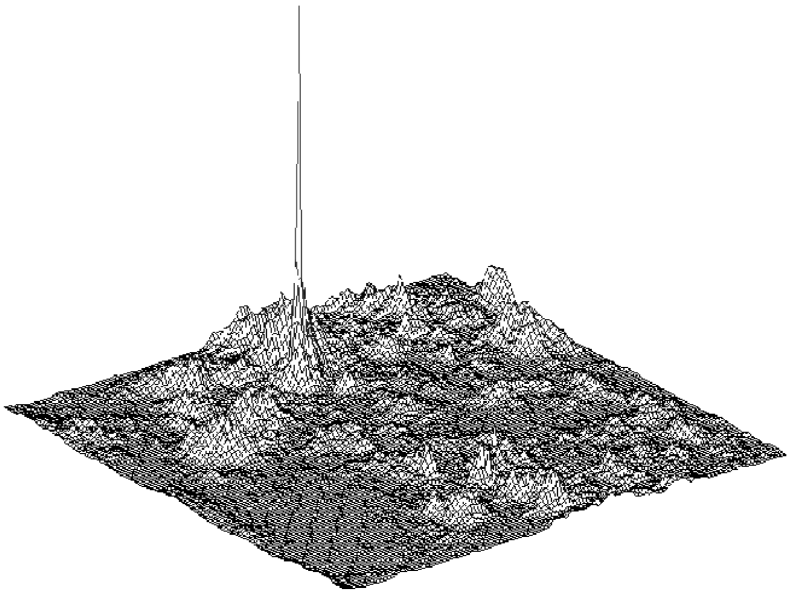


Fig. 15. Computer simulation result obtained for the input scene with the adaptive JTC.

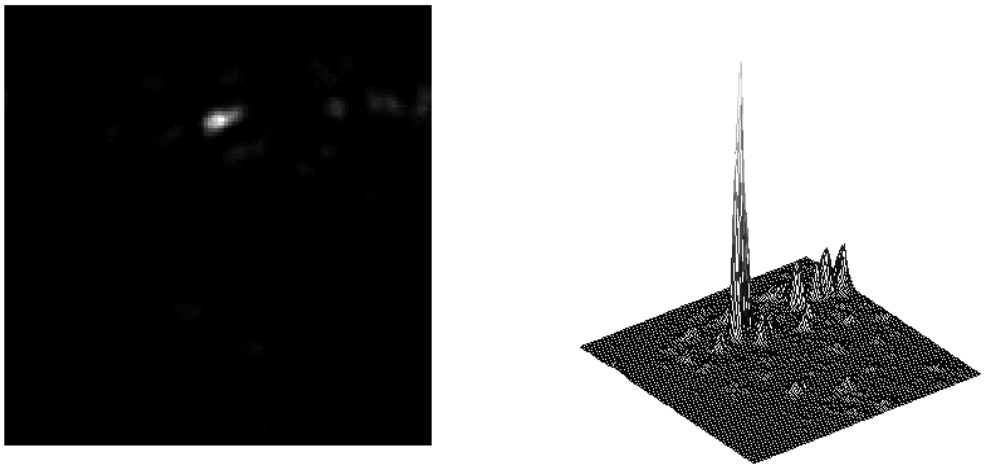
Note that the target is clearly detected. The adaptive JTC architecture can reliably detect a target embedded in a noisy background even if the target presents small geometric image distortions. We used 50 statistical trials of our experiment for different positions of the target. With 95% confidence, the DC obtained in computer simulation is equal to 0.82±0.003.

Bipolar Decomposition Method Results

The first optodigital experiment is based on the bipolar decomposition method. The reference image in Fig. 13(b) has real positive and negative values. We decompose this image into two nonnegative images (see Eqs. (22)-(24)). Two experiments are performed. In the first experiment the input scene is composed with the positive part of the reference image and the joint transform correlation is carried out. The experiment is repeated with the negative part of the reference image. The intensity correlation plane obtained after the postprocessing given in Eq. (25) is shown in Fig. 16. The DC obtained in the experiment is equal to 0.78.

Constant Addition Method Results

The second optodigital experiment is based on the constant addition method described. We use the input image and the reference image shown in Fig. 13. The SLM has a finite size (less than the size of the optical lens), and, after adding a high constant bias to the joint image, the signal at the plane of the SLM may be considered as a signal masked by a rectangular window.



(a) (b)

Fig. 16. Cross-correlation intensity plane obtained with bipolar decomposition method: (a) intensity plane, (b) intensity distribution.

The joint image formed for the constant addition method is shown in Fig. 17.

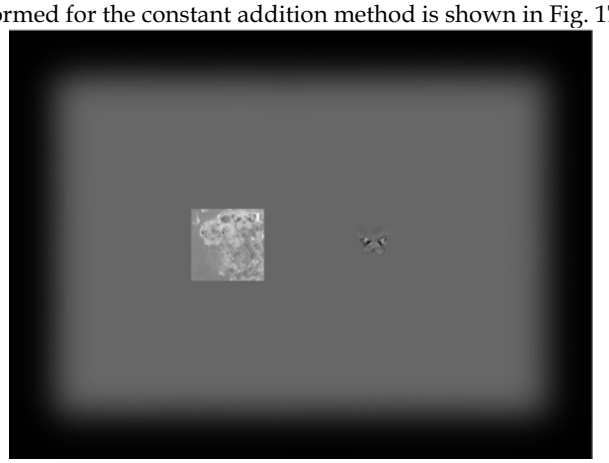
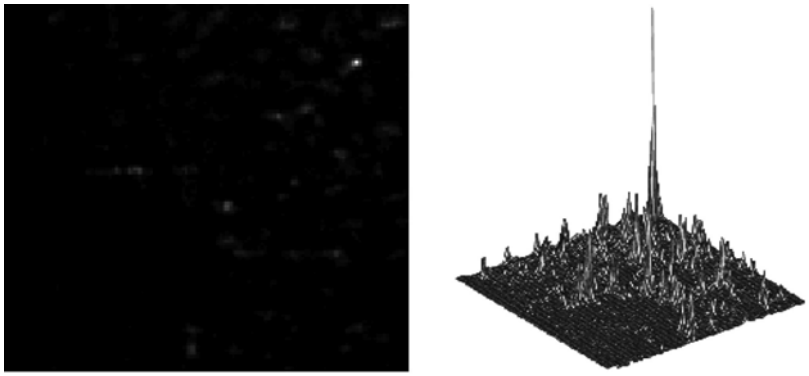


Fig. 17. Joint image formed for the constant addition method.

The Fourier transform of such a signal is the convolution between the spectrum of the joint image and a sinc function (Fourier transform of the rectangular window). Actually, the sinc function possesses high sidelobes that may severely affect the joint power spectrum. To avoid these effects, the input joint image is masked by a window with smoothed edges. Next we calculate all needed constants C_1, C_2, and C_3 given in Eq. (30). Figure 12 gives the relationship between a dynamic range of the used optodigital LCD and CCD camera and a digital range of a signal. Whereas digital images possess a range of [0-255] gray-scale levels, the signals in the optodigital domain have a range of [0-48] levels. We need to scale all images and the constant bias involved in the optodigital setup. The needed constants are equal to $C_1 = 31.75$, $C_2 = 23.03$, and $C_3 = 40$. The α value can be estimated as $\alpha = 1/cs$,

where s is the number of image pixels. The cross-correlation intensity plane obtained in the optodigital JTC after postprocessing is shown in Fig. 18.



(a) (b)

Fig. 18. Cross-correlation plane obtained with constant addition method; (a) intensity plane, (b) intensity distribution.

One can observe that the target is successfully recognized with DC=0.648. Finally, note that this method requires only one optical correlation, whereas the bipolar decomposition method uses two correlations to reconstruct the desired output.

4. Conclusion

Adaptive pattern recognition is still in state of rapid evolution. In this chapter we proposed digital and hybrid optodigital systems designed on the base of adaptive correlation filters to improve recognition of objects in cluttered backgrounds. It was shown that the proposed iterative filter design algorithms with a few training iterations helps us to take the control over the whole correlation plane. The digital systems are based on iterative training of the SDF filters. The hybrid systems additionally take into account real characteristics of used optoelectronics devices. The digital systems can be easily implemented in a computer, whereas the hybrid systems are able to provide real-time pattern recognition. The computer simulation and experimental results demonstrated a good performance of the proposed filters for pattern recognition comparing with known correlation filters. The suggested filters possess high scene-adaptivity, good robustness to small geometric image distortions and input noise.

5. References

Alam, M. S. & Kairm, M. A. (1993). Fringe adjusted joint transform correlator. *Applied Optics*, Vol. 32, No. 23, (August 1993) (4344-4350), ISSN 0003-6935.

Arsenault, H. & Hsu, Y. (1983). Rotation invariant discrimination between almost similar objects. *Applied Optics*, Vol. 22, No. 1, (January 1983) (130-132), ISSN 0003-6935.

Billet, O. & Singher, L. (2002). Adaptive multiple filtering. *Optical Engineering*, Vol. 41, No. 1, (January 2002) (55-68), ISSN 0091-3286.

Casasent, D. (1984). Unified synthetic discriminant function computational formulation. *Applied Optics*, Vol. 23, No. 10, (May 1984) (1620-1627), ISSN 0003-6935.

Diaz-Ramirez, V. H.; Kober, V. & Alvarez-Borrego, J. (2006). Pattern recognition with an adaptive joint transform correlator. *Applied Optics*, Vol. 45, No. 23, (August 2006) (5929-5941), ISSN 0003-6935.

González-Fraga, J. A.; Kober, V. & Alvarez-Borrego, J. (2006). Adaptive synthetic discriminant function filters for pattern recognition. *Optical Engineering*, Vol. 45, No. 5, (May 2006) (0570051-05700510), ISSN 0091-3286.

Horner, J. L. & Gianino, P. D. (1984). Phase-only matched filtering. *Applied Optics*, Vol. 23, No. 6, (March 1984) (812-816), ISSN 0003-6935.

Hester, C. F. & Casasent, D. (1980). Multivariant technique for multiclass pattern recognition. *Applied Optics*, Vol. 19, No. 11, (June 1980) (1759-1761), ISSN 0003-6935.

Javidi, B. (1989). Nonlinear joint power spectrum based optical correlation. *Applied Optics*, Vol. 28, No. 12, (June 1989) (2358-2366), ISSN 0003-6935.

Javidi, B. & Wang, J. (1994). Design of filters to detect a noisy target in nonoverlapping background noise. *Journal OSA (A)*, Vol. 11, No. 10, (October 1994) (2604-2612), ISSN 1084-7529.

Kober, V. & Campos, J. (1996). Accuracy of location measurement of a noisy target in a nonoverlapping background. *Journal OSA (A)*, Vol. 13, No. 8, (August 1996) (1653-1666), ISSN 1084-7529.

Kober, V.; Yaroslavsky, L.P.; Campos, J. & Yzuel, M.J. (1994). Optimal filter approximation by means of a phase only filter with quantization. *Optics Letters*, Vol. 19, No. 13, (July 1994) (978-980), ISSN 0146-9592.

Kober, V.; Mozerov. M & Ovseevich I.A. (2006). Adaptive correlation filters for pattern recognition. *Pattern Recognition and Image Analysis*, Vol. 16, No. 3, (2006) (432-431), ISSN 1054-6618.

Lu, K. & Saleh, B. E. A. (1990). Theory and design of the liquid crystal TV as an optical spatial phase modulator. *Optical Engineering*, Vol. 29, No. 3, (March 1990) (240-247), ISSN 0091-3286.

Moreno, I; Campos, J; Yzuel, M.J. & Kober, V. (1998). Implementation of bipolar real-valued input scenes in a real-time optical correlator: application to color pattern recognition. *Optical Engineering*, Vol. 37, No. 1, (January 1998) (144-150), ISSN 0091-3286.

Mahalanobis, A.; Vijaya Kumar, B.V.K. & Casasent, D. (1987). Minimum average correlation filters. *Applied Optics*, Vol. 26, No. 17, (September 1987) (3633-3640), ISSN 0003-6935.

VanderLugt, A. B. (1964). Signal detection by complex spatial filtering, *IEEE Trans. Inf. Theory*. Vol. 10, No. 2, (April 1964) (139-145), ISSN 0018-9448.

Vijaya Kumar, B.V.K. & Hassebrook, L. (1990). Performance measures for correlation filters. *Applied Optics*, Vol. 29, No. 20, (July 1990) (2997-3006), ISSN 0003-6935.

Waver, C. S. & Goodman, J. L. (1966). Technique for optically convolving two functions. *Applied Optics*, Vol. 5, No. 7, (1966) (1248-1249), ISSN 0003-6935.

Yaroslavsky, L.P. (1993). The theory of optimal methods for localization of objects in pictures. In: *progress in Optics XXXII*, E. Wolf, (Ed.), (145-201), Elsevier, ISBN: 0-444-86923-9, North-Holland.

Vision Systems: Segmentation and Pattern Recognition

Edited by Goro Obinata and Ashish Dutta

ISBN 978-3-902613-05-9
Hard cover, 536 pages
Publisher I-Tech Education and Publishing
Published online 01, June, 2007
Published in print edition June, 2007

Research in computer vision has exponentially increased in the last two decades due to the availability of cheap cameras and fast processors. This increase has also been accompanied by a blurring of the boundaries between the different applications of vision, making it truly interdisciplinary. In this book we have attempted to put together state-of-the-art research and developments in segmentation and pattern recognition. The first nine chapters on segmentation deal with advanced algorithms and models, and various applications of segmentation in robot path planning, human face tracking, etc. The later chapters are devoted to pattern recognition and covers diverse topics ranging from biological image analysis, remote sensing, text recognition, advanced filter design for data analysis, etc.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Vitaly Kober, Victor H. Diaz-Ramirez, J. Angel Gonzalez-Fraga and Josue Alvarez-Borrego (2007). Real-Time Pattern Recognition with Adaptive Correlation Filters, Vision Systems: Segmentation and Pattern Recognition, Goro Obinata and Ashish Dutta (Ed.), ISBN: 978-3-902613-05-9, InTech, Available from: http://www.intechopen.com/books/vision_systems_segmentation_and_pattern_recognition/real-time_pattern_recognition_with_adaptive_correlation_filters

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821