Adaptive learning with joint transform correlators

Joseph Rosen
Uri Mahlab
Joseph Shamir, FELLOW SPIE
Technion—Israel Institute of Technology
Department of Electrical Engineering
Haifa 32000, Israel

Abstract. Iterative learning procedures can be employed to generate synthetic reference patterns for joint transform correlators. These patterns can be implemented directly on inexpensive spatial light modulators (SLMs) based on liquid crystal television sets. The algorithm takes into account random electronic noise in the system and compensates for SLM distortions. Experimental results demonstrate the advantages of the new approach with two algorithmic examples.

Subject terms: *optical pattern recognition; correlators; simulated annealing; iterative learning.*

Optical Engineering 29(9), 1101-1106 (September 1990).

CONTENTS

1. INTRODUCTION

Since the introduction of the holographic spatial filter[1] much research has been carried out to improve the performance of optical correlators and to overcome the difficulties involved in their practical implementation. Most of the recent progress in this field has been made by abandoning the direct photographic recording of the filters in favor of digital procedures. The beginning of this trend may be marked by the introduction of the composite filter[2] with its various follow-ups.[3]

Most of the above-mentioned procedures involve the generation of spatial filters containing high spatial frequency components. For real-time applications, it is desirable to implement these filters on spatial light modulators (SLMs), which, within the present state of the art, have quite limited resolution. To reduce this difficulty one can employ a joint transform correlator (JTC)[4,5] where a reference pattern, rather than its Fourier transform (FT), is presented at the input plane. The main advantage of the JTC architecture over the conventional 4-f correlator is that it is easier to implement. In the JTC there is no need for a coded hologram with high carrier frequency that requires high resolution recording media and tedious alignment procedures.

The conventional procedure with a JTC is to use a reference pattern identical to the one to be detected. A deviation from this standard procedure was presented in Ref. 6, where a synthetic discriminant function (SDF) was employed as the reference. The SDF was generated by linear procedures and then binarized. A second binarization was performed on the FT plane. Although

some promising results were reported, the two binarizations performed after a SDF was generated do not guarantee optimal performance.

In this paper we adopt some iterative procedures[7,8] for the JTC architecture that have proved very effective for generating spatial filters used in 4-f correlators. Two procedures developed in Sec. 2 are suitable for generating the reference patterns directly on the input SLM, taking into account all the constraints imposed on the system. Considering the limitations of presently available SLMs, the reference patterns are binary and an optical system converts them into a bipolar form that leads to better discrimination. The optical system is described in Sec. 3 and experimental results are given in Sec. 4.

2. SYNTHETIC FILTER GENERATION FOR THE JTC

Using the conventional procedure with a JTC, an input pattern $f(x,y)$ is located with its center at point (x_f, y_f), and a reference pattern $h(x,y)$, is positioned around point (x_h, y_h). The complex amplitude distribution over the output plane $P(x', y')$ can then be written in the form

$$P(x', y') = C_{ff}(x', y') + C_{hh}(x', y')$$
$$+ C_{fh} * \delta[x' - (x_h + x_f), y' - (y_h + y_f)]$$
$$+ C_{hf} * \delta[x' + (x_h + x_f), y' + (y_h + y_f)] , \quad (1)$$

where $*$ denotes convolution and

$$C_{ff}(x', y') = \int f(x,y) f^*(x - x', y - y') \, dx dy , \quad [2(a)]$$

$$C_{fh}(x', y') = \int f(x,y) h^*(x - x', y - y') \, dx dy . \quad [2(b)]$$

The functions C_{fh} and C_{hf} denote the correlation of h and f and occur at two positions, $(x_h + x_f, y_h + y_f)$ and $[-(x_h + x_f), -(y_h + y_f)]$. In our analysis we used the term correlation as is in Eq. [2(b)] only.

As mentioned earlier, the objective of this paper is to investigate procedures to generate synthetic reference patterns. Two approaches that were found useful for FT correlators will be described. The first one is the entropy optimized filter (EOF)[7,9]

and the second one is the peak to peak ratio (PPR) method. For the process described, we find the term PPR more suitable than SNR[8] or peak to side lobe ratio (PSR).[3]

2.1. Entropy optimized filter

Assuming a set of input patterns $\{f_n(x,y)\}$, we define the goal of the EOF as the detection of the presence of patterns out of the subset $\{f_n^d(x,y)\}$ while rejecting all other patterns denoted by the subset $\{f_n^r(x,y)\}$. A reasonable criterion for detection is the appearance of a strong and narrow peak as contrasted with a uniform distribution for a pattern to be rejected. To quantify this criterion we normalize the energy distribution $|C_{f_ih}(x', y')|^2$ over the output plane for the ith input pattern and define a distribution function by the relation

$$\phi_i(x', y') = \frac{|C_{f_ih}(x', y')|^2}{\int\limits_{-\infty}^{\infty} |C_{f_ih}(x', y')|^2 dx'dy'} \quad . \tag{3}$$

Being a normalized distribution, $\phi_i(x', y')$ has all the properties of a probability density. Our criterion states that for a rejected pattern—a representative of the r subset—we would like to obtain a uniform ϕ_i over the whole output plane, while if a pattern from the d subset is presented the result should be a strong and narrow peak. The distinction between a uniform distribution and a well-defined peak may be quantified with the help of the entropy function.[10] The entropy function of ϕ_i is, by definition,

$$S_i = -\int\limits_{-\infty}^{\infty} \phi_i(x', y')\log\phi_i(x', y')dx'dy' \quad , \tag{4}$$

with its maximal value obtained for a uniform distribution and its minimal value obtained for a single sharp peak.

In previous publications[7,9] the Fourier transform of the reference function $h(x,y)$ was realized as a spatial filter in a Fourier transform correlator. In the present work $h(x,y)$ is introduced directly as the reference function for the JTC. For the implementation of the procedure with digital techniques, we display the input and reference functions on an $N \times N$ matrix of pixels, convert to a digitized form, and represent them as one-dimensional vectors of N^2 elements:

$$f_i(x,y) \rightarrow f_i(m) \ , \quad m = 1,2,... N^2 \ ,$$

$$h(x,y) \rightarrow h(m) \ , \quad m = 1,2,... N^2 \ .$$

Recognition of a pattern $f_i^d(m)$ will be represented by a steep peak on the correlation plane for which ϕ_i will have the form

$$\phi_i^d(m) = \begin{cases} 1 & m = k \\ 0 & \text{otherwise} \ , \end{cases} \tag{5}$$

where k is a pixel in the correlation plane that is not uniquely defined. In the JTC configuration we can define a correlation region within the output plane subject to certain limitations[11] and require the presence of a peak at any position within that region for the given pattern. The flexibility in the position of

the peak provides an additional degree of freedom that is useful for the optimization process.

For the rejected patterns, we would like to have a uniform ϕ over the whole output plane of $(2N-1)^2$ pixels; such that

$$\phi_i^r(m) = \frac{1}{(2N-1)^2} \ \forall \ m = 1,2,...(2N-1)^2 \ . \tag{6}$$

Measured by the entropy function this distribution has the maximum possible entropy,

$$S^r = -\log\frac{1}{(2N-1)^2} \ ,$$

while $\phi^d(m)$ of the previous equation has the minimum entropy, $S^d = 0$. If we define a cost function by the relation

$$M = \sum\limits_{\{f_n^d\}} S_n - \sum\limits_{\{f_n^r\}} S_n \ , \tag{7}$$

its global minimum value will be bounded by the above quantities. Naturally, taking into account practical limitations, this global minimum cannot be completely achieved. Considering M to be a function of the reference pattern, we employ an iterative algorithm that leads to an entropy optimized pattern $h_{\text{EOF}}(m)$ that satisfies

$$M_{\min}^{\text{EOF}} = M[h_{\text{EOF}}(m)] \ , \tag{8}$$

where $M_{\min}^{\text{EOF}}$ is as close to the ideal value of Eq. (7) as possible.

Various iterative procedures may be employed to implement the EOF[7] by minimizing the cost function of Eq. (7). In the present work we used simulated annealing[9,12] that, in principle, is capable of reaching a global minimum. Using this algorithm, we have at the lth iteration $M_l^{\text{EOF}} = M^{\text{EOF}}(h_l)$. Inducing a random perturbation over the pixels of h_l we obtain the $(l+1)$th iteration of the filter, which changes M by an amount

$$\Delta M^{\text{EOF}} = M^{\text{EOF}}(h_{l+1}) - M^{\text{EOF}}(h_l) \ . \tag{9}$$

The new filter is accepted if $\Delta M^{\text{EOF}} < 0$; otherwise, it is conditionally accepted, based on the acceptance probability

$$Pr_{\text{accept}} = \exp\left(\frac{-\Delta M^{\text{EOF}}}{T}\right) \ , \tag{10}$$

where T is the temperature parameter. This acceptance criterion is derived from an analogy with the Gibbs distribution used in thermodynamics.

The procedure is now repeated starting from a new filter. Decreasing the temperature slowly as the process continues causes M to approach its optimum value. The cooling rate and the steps of the random perturbation are important parameters that depend on the specific process implemented.

2.2. PPR procedure

In the procedure, based on Ref. 8, our goal is similar to the one we had for the generation of the EOF. We denote by

$$V_{f_ih}^{\max} = \max_{m \in R} |C_{f_ih}(m)|^2 \ , \quad m = 1,2,...(2N-1)^2 \ , \tag{11}$$

which is the maximum intensity value of the correlation between f_i and h, represented also as a one-dimensional vector. Therefore we define a cost function M^{PPR} to be

$$M^{PPR} = -\frac{\text{minimum peak value for recognized patterns}}{\text{maximum peak value for rejected patterns}}$$

$$= -\frac{\min\limits_{i \in \{d\}} V^{max}_{f_i h}}{\max\limits_{j \in \{r\}} V^{max}_{f_j h}} , \qquad (12)$$

where $i \in \{d\}$ is the detected function set and $j \in \{r\}$ is the rejected function set.

Considering M^{PPR} to be a function of a specific reference pattern, we seek a PPR optimized pattern $h^{PPR}(m)$ that satisfies

$$M^{PPR}_{min} = M[h^{PPR}(m)] . \qquad (13)$$

In principle, this procedure can also include the values of the side lobes over the correlation plane for the detected functions to improve the peak to side lobe ratio, but in our examples we found this unnecessary.

To perform the optimization process a direct search algorithm was employed.[13] Compared to simulated annealing this algorithm speeds up the iterative process, but it may become trapped in a local minimum.

Assuming the existence of a reference pattern $h(m)$, at the lth iteration we have $M^{PPR}_l = M^{PPR} (h_l)$. Inducing a change over one pixel of h, we obtain the $(l+1)$th iteration that changes M^{PPR} by an amount

$$\Delta M^{PPR} = M^{PPR}(h_{l+1}) - M^{PPR}(h_l) . \qquad (14)$$

The new pattern is accepted if $\Delta M^{PPR} \leq 0$; otherwise it is rejected. The procedure is now repeated until the minimum is achieved.

3. JTC LEARNING ARCHITECTURE

Several procedures have been proposed for producing bipolar spatial filters[14,15] but they are difficult to implement, in particular with SLMs that induce only a small rotation of the polarization. This is the case also for the inexpensive SLMs made by modifying a liquid crystal portable TV set (LCTV), such as the one used in our experiments. To overcome this difficulty we present the reference pattern on the SLM in a binary form that can be expressed as a superposition of a bipolar signal and a DC term. The DC term is optically filtered out in a preprocessing step, leaving the bipolar component. A schematic diagram of the complete processing system is shown in Fig. 1. On SLM1 we display a positive, binary function denoted by $h^+(x,y)$, with component values 1 and 0. The preprocessing is accomplished in the 4-f system with the zero order term stopped at the Fourier plane P1. The filtered image of the reference pattern $h(x,y)$, with values -1 and 1, is obtained at the P2 plane, which is the input plane to the JTC. It should be noted that the bipolar distribution will be symmetric (1 and -1) only if the original binary distribution has approximately equal numbers of zeros and ones. The input pattern is displayed on SLM2 and relayed to the input plane by a second 4-f system. This second 4-f configuration is useful for obtaining a symmetric presentation of the two patterns on plane P2.
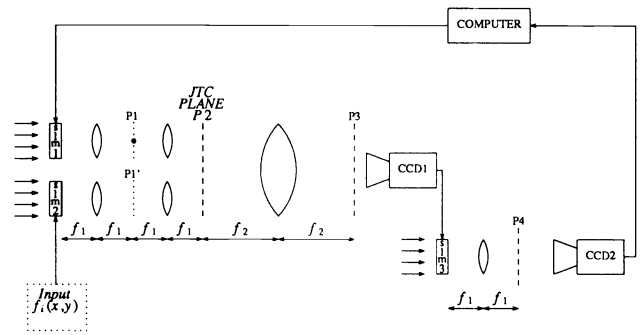

Fig. 1. Schematic diagram of JTC learning and pattern recognition system.

Using optical processing one may present several patterns of the training set simultaneously. Presenting K such patterns, the distribution over the input plane P2 can be written in the form

$$s(x,y) = h(x,y)*\delta(x-x_h, y-y_h)$$

$$+ \sum_{i=1}^{K} f_i(x,y)*\delta(x-x_i, y-y_i) , \qquad (15)$$

where (x_h,y_h) denote the center coordinates of the reference pattern and (x_i,y_i) is the center position of the ith input pattern. The FT of this pattern is obtained over plane P3, which is observed by the camera CCD1. The intensity distribution $|S(w_x,w_y)|^2$ is displayed on SLM3 and Fourier transformed to plane P4. Denoting the Fourier transform operator by $\mathcal{F}$, the distribution over the correlation plane is given by

$$P(x', y') = \mathcal{F}|S(w_x,w_y)|^2 \qquad (16)$$

and is observed by camera CCD2. It should be noted that in a practical system one may use time sharing[16] to implement the two parts of the system with a single SLM and CCD camera. The computer samples the proper region for each correlation intensity $|C_{f_i h}(x', y')|^2$ and computes Eqs. (3), (4), and (7) to search for an EOF, or Eqs. (11) and (12) to search for a maximum PPR reference. Thus the optimization algorithm can be implemented within the electro-optical system that will be used afterwards for the recognition tasks.

The main advantage of using this system in order to generate synthetic reference patterns as compared to pure digital process is a substantial reduction in learning time. This is achieved by the instantaneous Fourier transform and parallel processing of the whole correlation function $C_{f_i h}$. In addition, this procedure takes into account the actual system parameters, and distortions are automatically corrected. Some of the experimental results are described in the next section.

4. SYNTHESIS AND PERFORMANCE EXPERIMENTS

In the preliminary experiments described here part of the system in Fig. 1 was implemented digitally, leading to intensive computations. Therefore, in these demonstrative experiments, relatively small training sets were used. The realization of the complete electro-optical architecture with much larger training sets is under way and will be reported in a later publication.
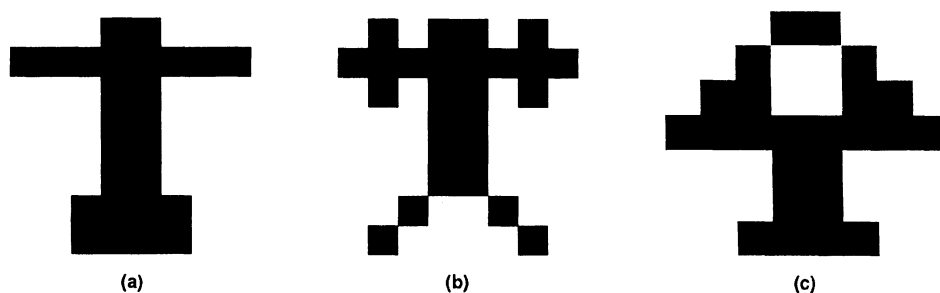
(a) **(b)** **(c)**

Fig. 2. Patterns of training set. Pattern in Fig. 2(a) is to be detected against the others.
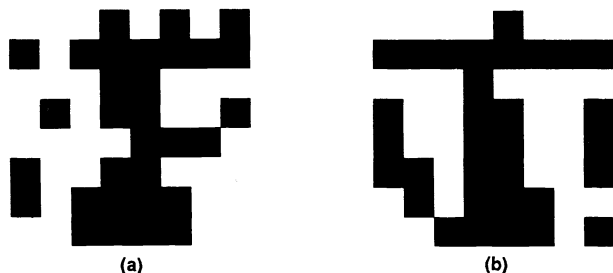

(a) **(b)**

Fig. 3. Reference patterns generated by (a) EOF method and (b) PPR method.

The input patterns were generated on a matrix with $N \times N = 8 \times 8$ pixels within an input plane containing 64×64 pixels. The training set contained the three patterns shown in Fig. 2, with the reference pattern generated to detect Fig. 2(a) and reject the other two. The substantially different pattern [Fig. 2(c)] was included in the learning process to demonstrate that the system is able to cope with a larger training set. Distinction between this pattern and Fig. 2(a) is naturally not difficult. Thus in the following description of the results no correlations with that pattern are presented. The initial reference pattern for the two iterative processes described in Sec. 2 was a constant,

$$h_0(x,y) = 1 \ ,$$

which is a choice as good as any other. This assumption was supported by experiments with different initial conditions, such as a conventional matched filter.

Figures 3(a) and 3(b) are the two reference patterns generated by computer simulation of the system in Fig. 1, according to Eqs. (7) and (12), respectively. A black pixel represents 1 and a white pixel represents -1. For comparison, the result of a conventional JTC with Figs. 2(a) and 2(b) as inputs and Fig. 2(a) as reference is shown in Fig. 4(a). The improved discrimination using the pattern of Fig. 3(a) is shown in Fig. 4(b), with a similar result for the reference pattern in Fig. 3(b). Quantifying the discrimination by the ratio between the detection peak and the rejection peak, we obtained 2.5 for the conventional JTC [Fig. 4(a)] and 6.25 for the synthetic reference pattern [Fig. 4(b)].

To test the performance in the laboratory we built the system shown in Fig. 5. In this preliminary experiment the learning stage to generate the reference pattern h was performed on a digital computer. The final binary reference pattern was plotted by a laser printer and reduced photographically. Both the input and reference patterns were presented on slides and projected onto the JTC input plane using a Mach-Zehnder configuration.
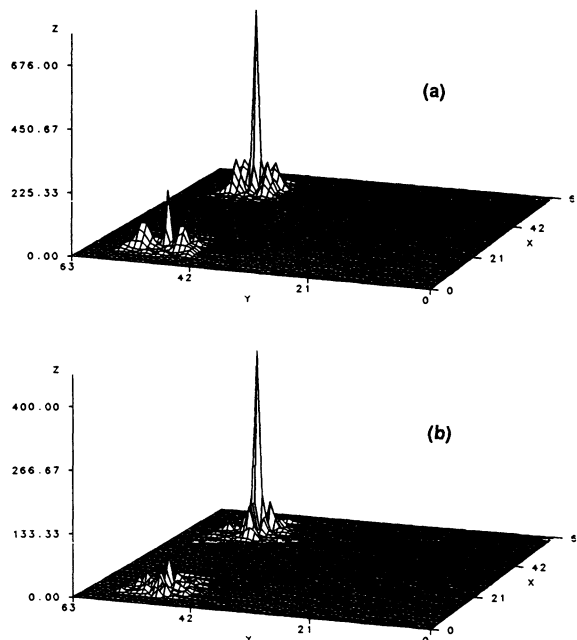

Fig. 4. Computer simulations with Figs. 2(a) and 2(b) on the input plane with (a) Fig. 2(a) as reference and (b) Fig. 3(a) as reference.
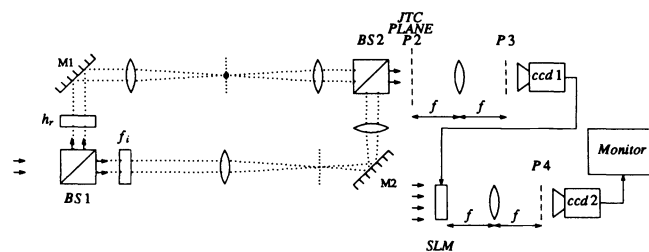

Fig. 5. Experimental system. Mach-Zehnder configuration superposes the input plane F_i and reference pattern h_r. The joint transform is recorded by camera CCD1 and displayed on the SLM. Correlation plane is observed by CCD2.

The layout appearing on the JTC plane is shown in Fig. 6. Figure 7 illustrates the spatial spectrum observed by CCD1 (plane P3).

In the first experiment, the final FT (to plane P4 of Fig. 1) was performed digitally. The output correlation regions are shown in Fig. 8, where Figs. 8(a) and 8(b) correspond to the correlations with the synthetic patterns of Fig. 3(a) and Fig. 3(b), respectively, presented in their binary form over the input plane. The
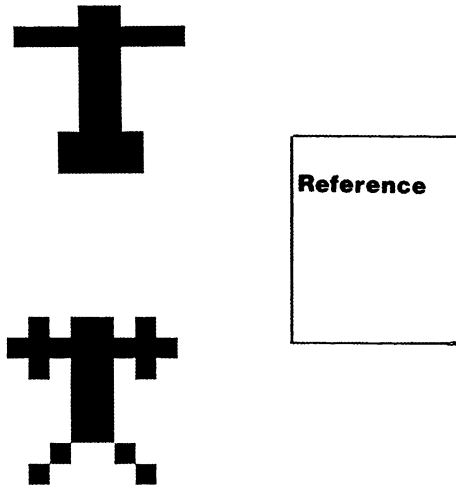
Reference

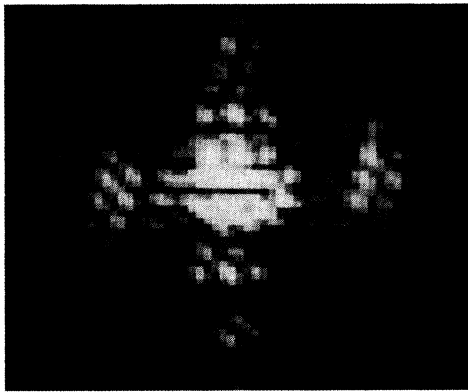**Fig. 6. Input and reference layout on JTC input plane of Fig. 5.**

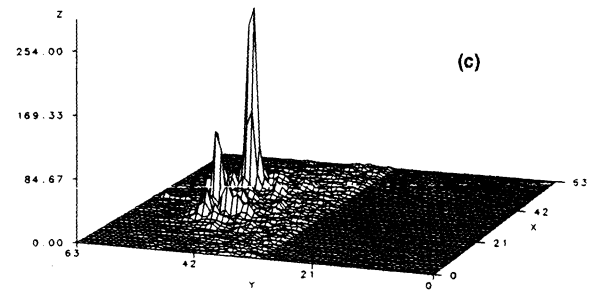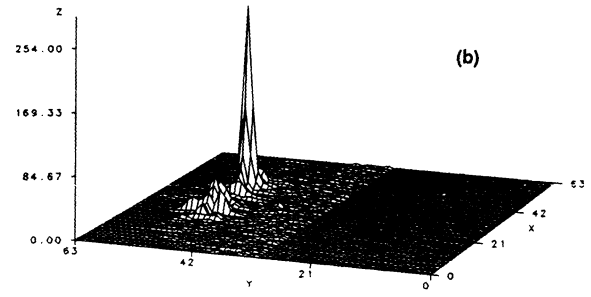**Fig. 7. Intensity distribution on the Fourier plane P3 observed by CCD1.**
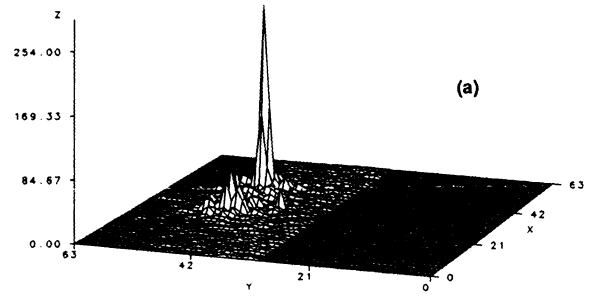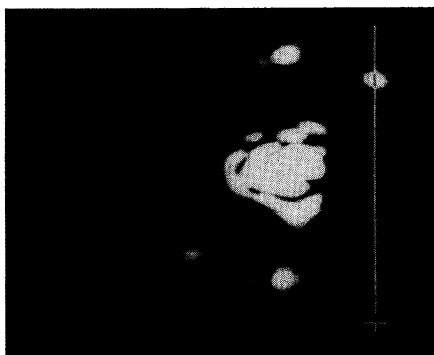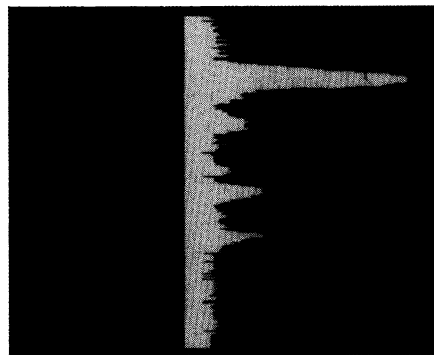

Fig. 8. Output plane distribution with the last FT performed digitally: Reference patterns were (a) Fig. 3(a), (b) Fig. 3(b), and (c) Fig. 2(a).

(a)

(b)

Fig. 9. Output plane distributions, as in Fig. 8(a), but with the last step performed optically. Intensity distribution is shown along scan line.
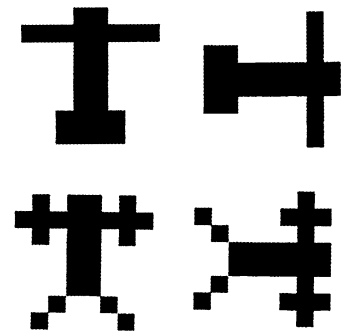

Fig. 10. Second training set. Reference pattern was generated to discriminate upper class against lower class.

comparison with a conventional procedure [Fig. 2(a) as reference] is shown in Fig. 8(c). The measured ratio between the intensity of the correlation peak and the cross-correlation peak was about 3.5, while with the synthetic reference patterns the corresponding ratio was about 6.3. The performance of the complete system, with the power spectrum of Fig. 7 presented on a

well-corrected SLM and Fourier transformed optically, is indicated by the output plane distribution shown in Fig. 9.

The capability of the system to cope with phase distortions in the SLM and a more difficult training set was tested by computer simulation experiments. In these experiments the training set of Fig. 10 was used to generate a reference pattern for
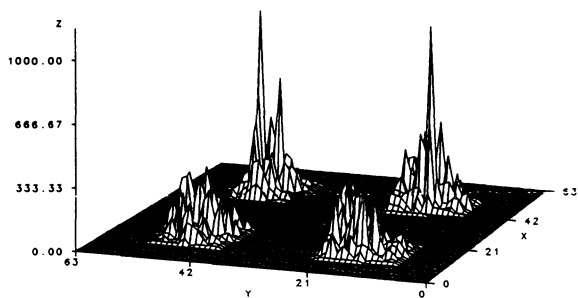
Fig. 11. Output plane distribution with phase-distorted reference pattern (computer simulation).

the detection of the two patterns in the upper row and rejection of those in the lower row. Assuming a very bad SLM, two values of a uniform random distribution of phases were assigned to each pixel independently. One phase value corresponded to a pixel in the "1" state, while the second phase value (independent of the first one) corresponded to the pixel in the "−1" state. The resulting phase distribution over the whole SLM multiplied the ideal transfer function during each iteration and the final state. Using the PPR method to generate the reference pattern, the output plane distribution of Fig. 11 was obtained. As expected for this difficult training set, a lower discrimination was obtained than in the earlier experiment. The detection/rejection ratio was 2.6. It is interesting to note that the same experiment with no phase distortions resulted in a ratio of only 2.25.

5. CONCLUSIONS

We introduced an iterative learning system for the design of synthetic reference patterns to be employed in a JTC. The high discrimination capability of these reference patterns was demonstrated by computer simulations and laboratory experiments. In our preliminary experiments the iterative generation of the reference pattern was done by digital computer, but the pattern recognition process was performed optically in real time. Further investigation, now under way, involves the implementation of the whole learning process on the JTC itself as indicated in Fig. 1.

The hybrid learning architecture is very promising because it is very fast, performing all the Fourier transforms optically, and takes into account the actual system parameters. Thus aberrations and distortions are automatically corrected, as was demonstrated by our computer simulation experiments. Initial laboratory tests also support this conclusion.

6. ACKNOWLEDGMENTS

7. REFERENCES

1. A. B. VanderLugt, "Signal detection by complex spatial filtering," IEEE Trans. Inf. Theory IT-10, 139–145 (1964).
2. H. J. Caulfield and W. T. Maloney, "Improved discrimination in optical character recognition," Appl. Opt. 8, 2354–2356 (1969).
3. B. V. K. Vijaya Kumar, Z. Bahri, and L. Hassebrook, "Review of synthetic discriminant function algorithms," in Real-Time Signal Processing for Industrial Applications, B. Javidi, ed., Proc. SPIE 960, 18–28 (1988).
4. C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," Appl. Opt. 5, 1248–1249 (1966).
5. F. T. S. Yu and X. J. Lu, "A real-time programmable joint transform correlator," Opt. Commun. 52, 10–16 (1984).
6. B. Javidi, "Synthetic discriminant function-based binary nonlinear optical correlator," Appl. Opt. 28, 2490–2493 (1989).
7. M. Fleisher, U. Mahlab, and J. Shamir, "Entropy-optimized filter for pattern recognition," Appl. Opt. 29, 2091–2098 (1990).
8. R. R. Kallman, "Optimal low noise phase-only and binary phase-only correlation filters for threshold detectors," Appl. Opt. 25, 4216–4217 (1986).
9. U. Mahlab and J. Shamir, "Phase-only entropy optimized filter generated by simulated annealing," Opt. Lett. 14, 1168–1170 (1989).
10. R. Kikuchi and B. H. Soffer, "Maximum entropy image restoration. I: The entropy expression," JOSA 67, 1656–1665 (1977).
11. F. T. S. Yu, F. Cheng, T. Nagata, and D. A. Gregory, "Effects of fringe binarization of multiobject joint transform correlation," Appl. Opt. 28, 2988–2990 (1989).
12. P. J. M. van Luarhoven and E. H. L. Aarts, Simulated Annealing: Theory and Applications, D. Reidel Publishing Co., Dordrecht, The Netherlands (1987).
13. M. A. Seldowitz, J. P. Allebach, and D. W. Sweeney, "Synthesis of digital holograms by direct binary search," Appl. Opt. 26, 2788–2798 (1987).
14. A. Furman and D. Casasent, "Bipolar incoherent pattern recognition by carrier encoding," Appl. Opt. 18, 660–665 (1979).
15. D. Psaltis, E. G. Paek, and S. S. Venkatesh, "Optical image correlation with a binary spatial modulator," Appl. Opt. 23, 698–704 (1984).
16. B. Javidi, J. L. Horner, and S. Odeh, "Single SLM joint transform correlator," Appl. Opt. 28, 1027–1032 (1989).

Joseph Rosen received his B.Sc. and M.Sc. in electrical engineering from Technion—Israel Institute of Technology in 1985 and 1987, respectively. He is studying for his D.Sc. in the field of optical pattern recognition in the Department of Electrical Engineering at Technion.
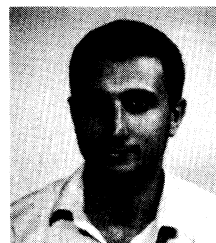
Uri Mahlab received his B.Sc. in electrical engineering cum laude from Ben-Gurion University in 1986 and his M.Sc. from the Technion—Israel Institute of Technology in 1988. He is studying for his D.Sc. in the field of optical pattern recognition in the Department of Electrical Engineering at Technion.
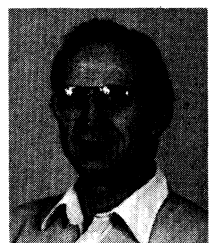
Joseph Shamir received the B.Sc. (1962), M.Sc. (1963), and D.Sc. (1968) from the Department of Physics, Technion-Israel Institute of Technology, Haifa, Israel. He has been on the faculty of the Department of Electrical Engineering at Technion since 1969. Dr. Shamir was a visiting scientist at Siemens AG, Munich (1971–1972), visiting professor in the Department of Electrical Engineering at Texas Tech University (1981–1982), visitor and lecturer in the "Window on Science Program" at various U.S. institutions (1979, 1983), and visiting scientist at Heinrich Hertz Institute, Berlin (1983). Since 1985 he has been partially affiliated with the Center for Applied Optics, University of Alabama in Huntsville, and is currently Senior Research Scientist. Dr. Shamir's research interests include optical signal processing and computing, nondestructive testing, holography, interferometry, ellipsometry, and light scattering. He has over 200 publications to his credit. Dr. Shamir was a member of the Foundation Committee (1969), Secretary (1981–1982), and since 1983 has been President of the Israel Laser and Electro-Optics Society. He is a Fellow of SPIE and OSA and a Senior Member of IEEE.