Application of serial- and parallel-projection methods to correlation-filter design

Tuvia Kotzer, Joseph Rosen, and Joseph Shamir

We describe generalized projection procedures for the design of arbitrary filter functions for correlators. More specifically, serial and parallel implementations of projection-based algorithms are employed. The novelty of this procedure lies in its generality and its ability to handle wide varieties of constraints by the same procedure. The procedure is demonstrated by the design of filters for the 4-f linear correlator, the phase-extraction correlator, and variants thereof. The filters are subject to a variety of constraints, including rotation-invariant pattern recognition and class discrimination. Examples are given to show the versatility, flexibility, and applicability of the design process to a variety of pattern-recognition tasks. Satisfactory results are also obtained because of the combination with the special nonlinear correlators proposed for pattern recognition.

1. Introduction

Pattern-recognition (PR) systems are usually designed with specific requirements. Examples of these requirements are rotation invariance, scale invariance, and tilt invariance. Various dedicated procedures were proposed in the past, such as circular harmonic component[1] (CHC) filters and CHC phase-only filters[2] (CHC POF's) for rotation-invariant PR, the Fourier–Mellin transform[3] for scale-invariant PR, position determination,[4] etc. The underlying characteristic of the above approaches is that they assume, *a priori,* a predefined structure for the filter. From a systems point of view, a generalized procedure for the design of arbitrary reference functions for correlators is desirable, without an *a priori* limiting structure.

In this paper we show how such requirements can be handled by general-purpose procedures. The power of the algorithms lie in (a) their simplicity, and (b) the fact that the solutions are not confined to a predetermined structure, which leaves more flexibility to arrive at not only better solutions, but also at solutions that were not previously considered possible because of a, perhaps mistaken, *a priori* confinement of the solution. The purpose of the paper is thus twofold: (a) introduce, review, and enhance some new concepts in the design of optical PR systems (for linear and nonlinear systems), and (b) demonstrate the applicability of projection-based methods for the achievement of superior performance in the above PR systems under a wide and quite stringent range of requirements.

The algorithms we employ are parallel and serial versions of the projection-onto-constraint sets (POCS's) method: when the serial-projection method[5,6] is applicable we employ it; otherwise we employ the recently introduced parallel-projection method,[7–9] based on Ref. 10, which may be employed for both linear-correlator (LC) systems as well as non-LC systems, such as the phase-extraction correlator[11] (PEC) and its variants. In Section 2, after some preliminary definitions, we describe the parallel- and the serial-projection methods and their characteristics. In Section 3 we design filters by a parallel version of POCS, for both the PEC and the 4-f LC. In Section 4 we investigate rotation-invariant filtering, based on the CHC[1] filter and introduce some energy measures according to which we can establish a fair criterion for comparison between PEC-based correlators and similar LC's. In Section 5, based on CHC decomposition theory and its application in Section 4, we design, by the serial-POCS method, special rotation-invariant filters to detect a class of objects that maintain the narrow, high-intensity, correlation peaks typical of the PEC. Conclusions are given in Section 6.

When this work was performed the authors were with the Department of Electrical Engineering, Technion, Israel Institute of Technology, Haifa 32000, Israel. T. Kotzer is now with the Algorithm Division, IMETRTIX, P.O. Box 1165, Rehovot 76110, Israel. J. Rosen is now with the Department of Applied Physics, California Institute of Technology, Pasadena, California.

2. Background

A. Serial Projections

Given a Hilbert space $\mathscr{H}$, a distance function d on $\mathscr{H}$, and a closed convex set (CCS) C in $\mathscr{H}$, projection from $\mathscr{H}$ onto C with respect to the distance function d is an operation P that associates to every element $h \in \mathscr{H}$ the (unique) element h' in C closest to h, where "close" is measured by d:

$$P(h) = h'$$
$$\text{if and only if} \quad h' \in C \text{ and } \inf_{y \in C} d(y, h) = d(h', h); \tag{1}$$

(projected vectors are henceforth marked by a prime). Usually d is derived from the prevailing Hilbert-space structure,

$$d(h, h') = \|h - h'\|_{\mathscr{H}} := \int |h(x) - h'(x)|^2 \mathrm{d}x. \tag{2}$$

If the sets C_i are closed with respect to d and are convex, the projection element exists and is unique. If the sets are not convex, procedures exist for determining the (unique) projection.[12]

Sometimes the projection operation is modified to admit relaxation. For instance, P may be replaced by the relaxed operator P_λ defined by

$$P_\lambda(h) = P(h) + \lambda[P(h) - h], \tag{3}$$

where λ is a real relaxation parameter with $|\lambda| < 1$.

Given N CCS's, C_i, $i = 1, \ldots, N$, $C_i \subset \mathscr{H}$, with a nonempty intersection $C_0 = \cap_{i=1}^N C_i$, we can associate a separate relaxed projection P_{i,λ_i} with each set C_i and corresponding projection P_i. To obtain an element in C_0 we iterate the composed operator T, defined by,

$$T = P_{N,\lambda_N} P_{N-1,\lambda_{N-1}} \cdots P_{1,\lambda_1}, \tag{4}$$

by using the following algorithm:

Algorithm 1: Given an arbitrary initial function $h^0(x)$,

$$h^{k+1} = T(h^k), \qquad k \geq 0. \tag{5}$$

For any arbitrary initial function h^0 we ensure that the infinite sequence $\{h^0, h^1, h^2, \ldots\}$ generated by algorithm 1 converges weakly[13] to an element in C_0, provided all projections are performed with respect to the same distance function[14] and that all the N sets are CCS's (in finite dimension, e.g., $\mathscr{H} = \mathbb{C}^n$, weak and strong convergence are the same). If some of the N sets are not convex, we are assured of a monotonic nonincrease of some error function along the iterates, provided $N \leq 2$. If $N > 2$ this is not guaranteed, even if only one set is not convex.

B. Parallel Projection

We start by defining generalized weighted, L^2, norm-squared distance functions, with weight W_i:

$$d_i(h_1, h_2) := \|H_1 - H_2\|_{W_i}^2$$
$$:= \int_{-\infty}^{\infty} |H_1(u) - H_2(u)|^2 W_i(u) \mathrm{d}u \quad h_1, h_2 \in \mathscr{H}, \tag{6}$$

where $W_i(u)$ is an essentially positive and essentially bounded weighting function, and uppercase letters denote the Fourier transform (FT) of the lowercase functions, e.g., $H_i(u) := \mathscr{F}\{h_i(x)\}$. We also define a cost functional:

$$\hat{J}(h)^2 := \sum_{i=1}^{N} \beta_i d_i[P_{C_i}^{d_i}(h), h]$$
$$= \sum_{i=1}^{N} \beta_i \| \mathscr{F}\{P_{C_i}^{d_i}(h)\} - \mathscr{F}\{h\}\|_{W_i}^2, \tag{7}$$

where $\beta_i > 0$ attributes an importance to the projection, $P_{C_i}^{d_i}(h)$ denotes the projection of h onto the set C_i with respect to the distance function d_i, i.e.,

$$P_{C_i}^{d_i}(h) = h' \quad \text{if and only if} \quad \inf_{h_1 \in C_i} d_i(h_1, h) = d_i(h', h),$$
$$h' \in C_i. \tag{8}$$

We also denote by P_{i,λ_i} the relaxed projections, as above [where we omit the superscript $(\cdot)^{d_i}$ for brevity]. If the sets C_i are closed with respect to d_i and convex, the projection element exists and is unique. If the sets are not convex, procedures exist for determining the (unique) projection.[12] With these definitions we can state the parallel-projection algorithm in its space (time) representation, generating the sequence of successive estimates $\{h^0, h^1, \ldots,\}$. Although the algorithm operates in an infinite-dimensional Hilbert space as well,[7,15] we assume finite dimension (as it is implemented on a digital computer).

Algorithm 2, space domain: Given an arbitrary initial function $h^0(x)$, calculate,

$$v_i^{k+1}(x) := P_{i,\lambda}[h^k(x)], \quad \text{for all } i = 1, 2, \ldots, N, \tag{9a}$$

$$h^{k+1}(x) = \mathscr{F}^{-1}\left\{ \frac{\sum_{i=1}^{N} \beta_i W_i(u) \mathscr{F}\{v_i^{k+1}\}(u)}{\sum_{i=1}^{N} \beta_i W_i(u)} \right\}, \tag{9b}$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ denote the FT and its inverse, respectively. For an equivalent frequency representation, see Ref. 16.

A detailed mathematical justification of this algorithm is provided in Refs. 15 and 16, which is briefly

reviewed in appendix A. Here we note only that iterates generated by this parallel algorithm converge weakly to C_0, provided that all sets are CCS's and $\lambda \in (-1, 1)$, and that the individual projections may be defined with respect to different distance functions, in contrast to the serial algorithm. Also, even if some, or all, of the sets are nonconvex, the cost function $\hat{J}$ is nonincreasing along the iterates, provided that $\lambda \in (0, 1)$, assuring us of improved estimates along the iterates. This holds for an arbitrary number of sets, as opposed to the serial algorithm in Subsection 2.A.

C. Correlator and Related Definitions

Using one-dimensional notation for brevity, we define the correlation between an input function $f(x)$ and a reference (filter) function $h(x)$ by

$$\Phi(x) = \mathscr{F}^{-1}\{N_l\{\mathscr{A}[f(x)]\}N_l\{\mathscr{A}[h(x)]\}\} \tag{10}$$

where N_l is a, possibly nonlinear, operator defined by

$$N_l[R(u)] = |R(u)|^l \exp[i\varphi(u)],$$
$$R(u) = |R(u)|\exp[i\varphi(u)], \qquad 0 \le l \le 1. \tag{11}$$

We further define $f_p(x)$ and $h_p(x)$ by $f_p(x) := \mathscr{F}^{-1}\{N_{l=0}\{\mathscr{A}[f(x)]\}\}$, $h_p(x) := \mathscr{F}^{-1}\{N_{l=0}\{\mathscr{A}[h(x)]\}\}$, which correspond to the phase parts of the functions $f(x)$ and $h(x)$, respectively.

With these definitions, we are in a position to state the following three correlators that are considered in this work:

LC:

$$\Phi(x) = h(x) \ f(x). \tag{12}$$

PEC:

$$\Phi(x) = h_p(x) \ f_p(x). \tag{13}$$

Generalized PEC (GPEC):

$$\Phi(x) = h(x) \ f_p(x), \tag{14}$$

Other degrees of nonlinearity, (monitored by l) can be tried as well, leading to nonlinear correlators similar to the nonlinear joint transform correlator,[17–19] as indicated in Ref. 20. Also, note that both the PEC and the GPEC are nonlinear correlation systems.

In the rest of the paper we employ the serial- and the parallel-projection methods for the design of filters for the LC, the PEC, and the GPEC. This is performed subject to a variety of demands (constraints) including class discrimination and class recognition with rotation invariance. We note that the serial POCS has already been applied successfully to the design of filters that are both rotation and shift invariant, as well as having a predetermined, limited scale range for which the response is constant too.[21] This was possible because of the flexibility of the method. The optical implementations of the various

PEC's and other non-LC's[11,18–20,22] and LC's[21] were given elsewhere and are not repeated here, for brevity.

3. Applications

Throughout the following sections, we use various projection algorithms to design filters for optical LC and non-LC systems. In the design process the constraints are basically composed of discrimination and peak energy (amplitude) constraints. Noise constraints, e.g., noise robustness, can be easily incorporated into the design process as well, at least for the parallel algorithm, as shown in Ref. 7 (in Ref. 7 the noise is taken into account for image-restoration purposes and the idea is similar for PR purposes).

Our interest here is concerned mainly with non-LC systems like the PEC and the GPEC that provide better discrimination than the LC and are barely affected by noise up to a certain level. Moreover, it was shown in Ref. 20 that the presence of noise actually assisted in the case of multiple-object inputs. Thus, for brevity, noise problems are not considered further, nor is shift invariance, which was demonstrated in Refs. 11 and 20.

A. Class Discrimination by a Linear Correlator

For a class-discrimination problem we define a training set consisting of two classes. The class to be detected is placed in a region of space R_1, and the class to be rejected is situated in the region R_2. The task is to design a filter, h, such that

(1) Its correlation with a given input function, f, will satisfy some correlation constraint C_1. Namely, in the detection region, R_1, the correlation peaks will be larger than some predetermined value T_1, whereas in the rejection region, R_2, the correlation will be lower than some predetermined value T_2. If the complete training set is presented simultaneously over the input plane, then R_1 corresponds to regions in the correlation plane that correspond to the positions of objects to be detected, whereas the regions R_2 represent the location of objects to be rejected and empty regions surrounding the correlation peaks in R_1. During the learning stage the correlation peak is assumed to be contained in a single pixel. Because this is physically not possible, some of the peak energy will leak out into neighboring pixels, constituting the background that should be below T_2. Also, T_1 and T_2 are appropriately chosen threshold values to provide sufficient discrimination (at least T_1/T_2) as well as sufficient energy in the peak (high absolute value of T_1). The appropriate values will depend on the specific application and the level of similarity between both classes.

(2) Its FT, $\mathscr{A}[h]$, corresponds to a passive element (C_2).

(3) It should have finite support, say $[-a, a](C_3)$.

Any filter h that satisfies all three constraints above is considered a solution. More specifically, the con-

straints are given by the following definitions:

$$C_1 := [h \,|\, (h \ f)(j) \in \hat{C}_1, \forall j], \tag{15a}$$

$$\hat{C}_1 := [\Phi(j) \,|\, |\Phi(j)| \leq T_2, \quad \text{for } j \in R_2;$$
$$\Phi_{\text{re}}(j) \geq T_1 \text{ and } \Phi_{\text{im}}(j) = 0, \quad \text{for } j \in R_1], \tag{15b}$$

$$C_2 := [h \,|\, |\mathscr{A}[h(j)]| \leq 1], \tag{15c}$$

$$C_3 := [h \,|\, h(j) = 0, \quad \text{for } j \ [-a, a]; \quad a > 0], \tag{15d}$$

where

$$\Phi(j) := (h \ f)(j), \qquad \Phi(j) := \Phi_{\text{re}}(j) + i\Phi_{\text{im}}(j);$$
$$\Phi_{\text{re}}(j) = \text{Re}[\Phi(j)], \qquad \Phi_{\text{im}}(j) = \text{Im}[\Phi(j)].$$

Actually, the measured quantity is $|\Phi|^2$ and not its imaginary or real values. However, the constraint, $|\Phi(j)|^2 \geq$ const. is not a convex constraint set and convergence is then not guaranteed. This is not the case for C_2, where the constraint is $|I(j)| \leq$ const. Thus with our choice, C_1, $\hat{C}_1$, C_2, and C_3 are CCS's.

Projections onto C_2, C_3 with respect to the distance function given by Eq. (6) with unity weighting $[W_i(u) = 1, \quad i = 2, 3]$, i.e., the Euclidean norm, are simple and are given by

$$P_{C_2}^{d_2}[h(x)] = \mathscr{F}^{-1}[H'(u)],$$

where

$$H'(u) = \begin{cases} H(u) & \text{if } |H(u)| \leq 1 \\ \exp[i\varphi_H(u)] & \text{otherwise} \end{cases},$$

$$P_{C_3}^{d_3}[h(x)] = \begin{cases} h(x) & \text{if } x \in (-a, a) \\ 0 & \text{otherwise} \end{cases}, \tag{16}$$

This careful choice of the weighting function results in a simple projection, $viz.$,

$$\{\mathscr{F}^{-1}[V_1] =\}$$

$$v_1 := P_{C_1}^{d_1}(h), \qquad \text{where } V_1(m) = \frac{\mathscr{A}[\Phi'(j)](m)}{F(m)},$$

$$\Phi'(j)$$

$$= \begin{cases} T_2 \exp[i\varphi_\Phi(j)], & \text{if } j \in R_2 \text{ and } |\Phi(x)| > T_2 \\ \Phi(j), & \text{if } j \in R_2 \text{ and } |\Phi(j)| \leq T_2 \\ T_1, & \text{if } j \in R_1 \text{ and } \Phi_{\text{re}}(j) < T_1, \\ \Phi_{\text{re}}(j), & \text{if } j \in R_1 \text{ and } C_{\text{re}}(j) \geq T_1 \\ \Phi(j), & \text{otherwise} \end{cases}$$

where

$$\Phi(j) = |\Phi(j)| \exp[+i\varphi_\Phi(j)]$$

$$= \mathscr{F}^{-1}[H(m)F(m)] = h(j) \ f(j). \tag{18}$$

For details see Ref. 16.

Algorithm 2 allows projections with respect to several different distance functions, and, therefore, projecting different quantities in domains where both the constraint set and the distance function are simple is possible (see Ref. 14, Section VI); hence it is employed for this filter synthesis task. The sequence $[h^k]_{k=0}^\infty$ generated by algorithm 2 converges to a function in C_0, satisfying all constraints, and is given by [see Eqs. (9a) and (9b)] $h^{k+1}(j) = \mathscr{F}^{-1}[H_{k+1}(m)]$, where

$$H^{k+1}(m) = \frac{\mathscr{A}[P_{C_1}^{d_1}\{\mathscr{F}^{-1}[H^k]\}](m)W_1(m) + \mathscr{A}[P_{C_2}^{d_2}\{\mathscr{F}^{-1}[H^k]\}](m) + \mathscr{A}[P_{C_3}^{d_3}\{\mathscr{F}^{-1}[H^k]\}](m)}{W_1(m) + 1 + 1}, \tag{19}$$

where $H(u) = \mathscr{A}[h(x)] = |H(u)| \exp[i\varphi_H(u)]$. Unfortunately, projection onto C_1 with respect to the Euclidean norm is complicated and is a typical constrained deconvolution problem in itself.[14,16]

To perform the projection onto C_1 easily we follow the idea proposed in Ref. 14. We perform the projection onto C_1, with respect to the distance function induced by a weighted norm squared, with the appropriate weighting given by $W_1(m) := |\mathscr{A}[f(j)]|^2 = |F(m)|^2$:

$$d_1(H_1, H_2) = \sum_m W_1(m)|H_1(m) - H_2(m)|^2$$

$$= \sum_j |\mathscr{F}^{-1}[[W_1(m)]^{1/2}](j) \ [h_1(j) - h_2(j)]|^2. \tag{17}$$

and where d_1 is given by Eq. (17), a zero-relaxation parameter $(\lambda = 0)$ is employed, and $d_2(h_1, h_2) = d_3(h_1, h_2) = \|h_1 - h_2\|$ (the Euclidean norm).

In one of our simulation experiments we started from a filter h such that $h \ C_1$, $h \in C_2$, $h \ C_3$. Figure 1 shows the input distribution.

The task is to detect the letter F and reject all others. Figure 2(a) shows the correlation distribution with a LC, where the filter is a POF[23] matched to the letter F. Fig. 2(b) shows the correlation distribution with the filter generated by algorithm 2. The improvement in both recognition and discrimination is obvious.

In the case of the GPEC, for which the correlation is given by $\Phi(x) := h(x) \ f_p(x)$, we may employ the serial-projection algorithm (POCS) with equal ease. This was already treated in Ref. 14 and is not discussed here.
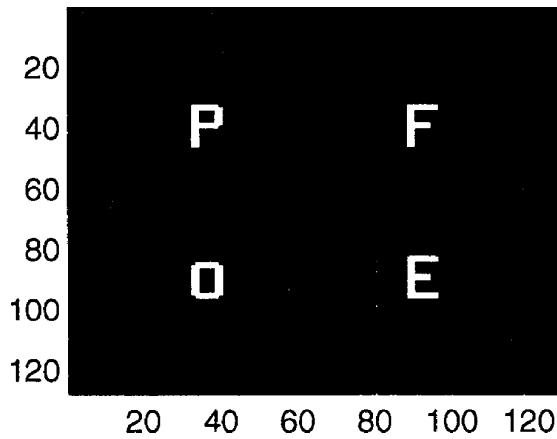
Fig. 1. Input distribution.

Fig. 3. Correlation results with the PEC. The input is Fig. 1, where the filter is (a) a POF matched to letter F, (b) generated by algorithm 2 for the PEC.

B. Class Discrimination by the Phase-Extraction Correlator

For the PEC, for which the correlation is given by $\Phi(x) := f_p(x) \ h_p(x)$, the convex constraint set C_2 in Eq. (15a) must be replaced by the nonconvex constraint set $C_{2-\text{nc}}$:

$$C_{2-\text{nc}} := \{h || \mathscr{A}\{h(j)\}| = 1\}. \tag{20}$$

In this case h is a solution if $h \in C_1 \cap C_{2-\text{nc}} \cap C_3$, i.e., $h = h_p$. Thus, to design a suitable POF h_p, it is necessary to iterate the operator $T := P_{C_1} P_{C_{2-\text{nc}}} P_{C_3}$. However, because one of the sets is nonconvex and we have more than two sets to project onto, we are not assured of any monotonic behavior of the iterates of algorithm 1.[12] However, the parallel algorithm, algorithm 2, may be employed, with assured monotonic reduction of the cost function $\hat{J}$ of Eq. (7).

Figure 3(a) shows the result of the PEC, according to Eq. (13), where the input is given by Fig. 1 and the filter is the POF, matched to the letter F. Figure 3(b) shows the result of the PEC with the same input, with the POF generated by algorithm 2. The improvement is again quite evident. Also, note that the correlation peaks are sharper in the PEC com-
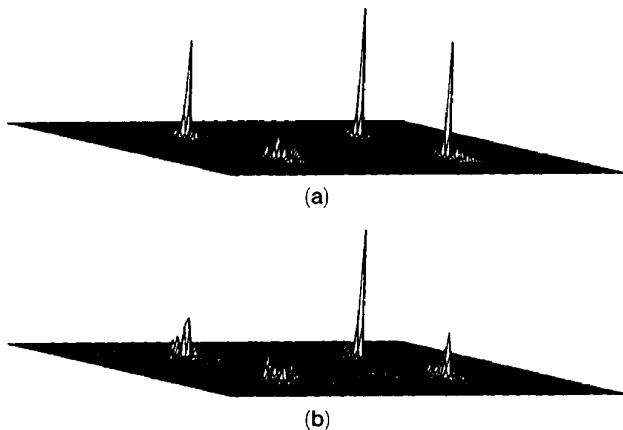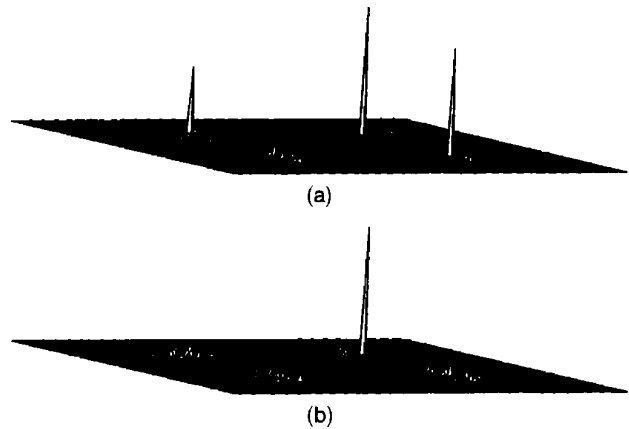
pared with those of the LC. This is due to the intrinsic high-frequency amplification of the PEC. However, as noted above, there may be some shift variance. To minimize this, we confined the impulse response of the filter to be narrow in the space domain (constraint C_3). The impulse response of the filter is shown in Fig. 4(a). Indeed, when taking the input shown in Fig. 1 and interchanging the positions of the letters F and E, we obtain the correlation function shown in Fig. 4(b), which is similar to that of Fig. 3(b) (when noting the interchange of letters), demonstrating approximate shift invariance. We also note that another approach for the design of filters for the PEC may be to do a phase-only operation on the filter generated for the GPEC (thus avoiding the problematic nonconvex constraint). However, clearly, it is better to incorporate the phase-only requirement in the design procedure, ensuring that all parameters involved are being optimized and designed according to them. Also, naturally, the parallel-projection algorithm is stopped after a finite number of iterations [basically when $\hat{J}(h)$ is considered to be small enough] and thus not all constraints are (yet) strictly satisfied. Nevertheless, in our simu-

Fig. 2. Correlation results with a LC. The input is Fig. 1, where the filter is (a) a POF matched to letter F, (b) generated by algorithm 2 for the LC.
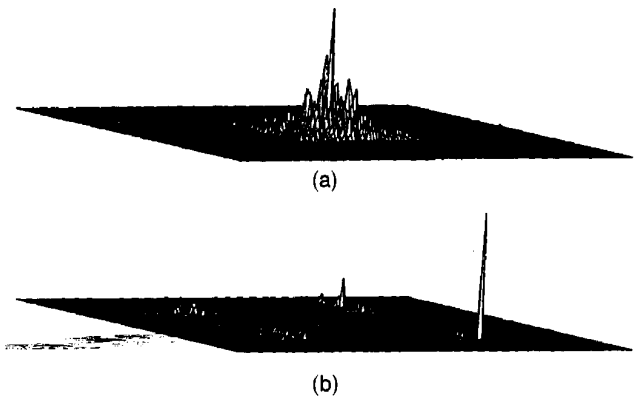

Fig. 4. (a) Impulse response of the filter. (b) As Fig. 3b, but with the letters of the input (from Fig. 1) F and E interchanged.

lation, almost all constraints were satisfied, i.e., the filter passed over 99% of the incident energy.

4. Distortion-Invariant Filter Design

Distorted views of an object are of major concern in optical correlators and were widely treated in the literature. The procedures described in this paper are particularly suitable to treat distortion problems. As an example we demonstrate the efficiency of a new concept in rotation-invariant signal processing by combining the PEC[11,22] and its generalization, the GPEC, with the rotation invariance of the CHC POF.[2]

A schematic diagram of the rotationally invariant PEC is shown in Fig. 5. For simplicity, in the analysis below we assume a single object in the input plane to which the filter is matched. The complications arising from multiple objects are seldom observable.[20] We denote the complex amplitude distribution of a given object (in polar coordinates) by $a(r, \theta)$, and its version rotated by an angle α is given by $a(r, \theta + \alpha)$. The respective FT's are denoted by $A(\rho, \varphi) \equiv \mathscr{F}[a(r, \theta)]$ and $A(\rho, \varphi + \alpha) = \mathscr{F}[a(r, \theta + \alpha)]$. Putting this single object as an input, we have $f(r, \theta) = a(r, \theta)$, and write its FT in the form

$$F(\rho, \varphi) \equiv |A(\rho, \varphi)|\exp[i\gamma(\rho, \varphi)].$$

Taking a filter function matched to our present input, we may write $h(r, \theta) = a(r, \theta + \pi)$ and obtain $H(\rho, \varphi) = |A(\rho, \varphi)|\exp[-i\gamma(\rho, \varphi)]$, which, after the phase-extraction operation (which corresponds to the operator $N_{l=0}$), turns into $H'(\rho, \varphi) = \exp[-i\gamma(\rho, \varphi)]$. Expanding $\exp[i\gamma(\rho, \varphi)]$ into its CHC[1] yields

$$\exp[i\gamma(\rho, \varphi)] = \sum_{N=-\infty}^{\infty} A_N(\rho)\exp(iN\varphi), \qquad (21a)$$
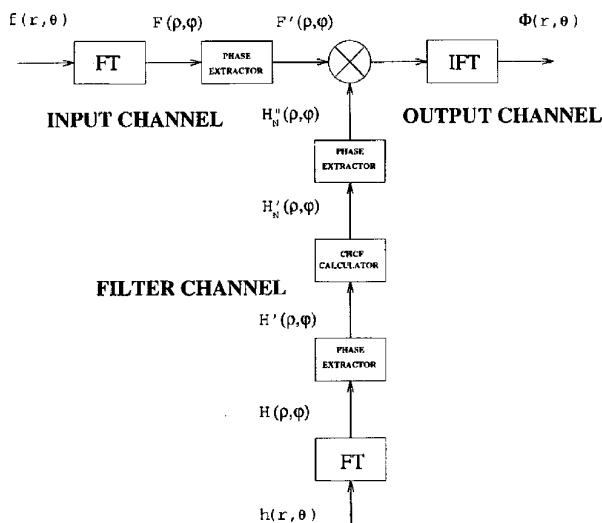

Fig. 5. Block diagram of the phase-extraction rotation-invariant correlator: IFT, inverse FT; $f(x, y)$, $h(x, y)$ the input and the filter functions, respectively; CHCF, CHC filter.

where

$$A_N(\rho) = |A_N(\rho)|\exp[i \, \arg[A_N(\rho)]]$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \exp[i\gamma(\rho, \varphi)]\exp(-iN\varphi)d\varphi. \qquad (21b)$$

First we analyze the output signal quality of the rotation-invariant PEC and compare it with its linear counterpart by using the measure of peak sharpness,[24] i.e., the peak-to-correlation energy (PCE), which is defined by the relation

$$\text{PCE} = \frac{|\Phi(0, 0)|^2}{\iint_{S_0} |\Phi(x, y)|^2 dxdy}, \qquad (22)$$

where $\Phi(x, y)$ is the output correlation function as above, at position (x, y), and S_0 is the aperture size. Choosing the Nth-order CHC for the filter calculator in Fig. 5, we get $H'_N(\rho, \varphi)$. After performing the phase-extraction operation, we obtain the final filter function:

$$H''_N(\rho, \varphi) = \exp[i \, \arg[H'_N(\rho, \varphi)]]$$

$$= \exp(-i[N\varphi + \arg[A_N(\rho)]]). \qquad (23)$$

Note that this is the phase distribution in the Nth component of the CHC decomposition of the phase part of the FT of the filter function. With the input $a(r, \theta + \alpha)$ constrained by a circular aperture of radius R_0 and using the orthogonality of CHC's, we obtain

$$\Phi(0, 0) = \left(\frac{1}{2\pi}\right)^2 \int_0^{2\pi} \int_0^{R_0} F'(\rho, \alpha + \varphi)H''_N(\rho, \varphi)\rho d\rho d\varphi$$

$$= \frac{\exp(iN\alpha)}{2\pi} \int_0^{R_0} |A_N(\rho)|\rho d\rho, \qquad (24)$$

the intensity of which is independent of the rotation angle.

The denominator in Eq. (22) is the energy over the whole correlation plane. Thus, because the filter is a POF, we may use Parseval's theorem to write

$$\iint_{S_0} |\Phi(x, y)|^2 dxdy = \frac{1}{2\pi} \sum_{M=-\infty}^{\infty} \int_0^{\rho_0} |A_M(\rho)|^2 \rho d\rho. \quad (25)$$

Substituting Eqs. (24) and (25) into the sharpness criterion [Eq. (22)] we obtain

$$\text{PCE}_N^P = \frac{1}{2\pi} \frac{\left[\int_0^{R_0} |A_N(\rho)|\rho d\rho\right]^2}{\sum_{M=-\infty}^{\infty} \int_0^{\rho_0} |A_M(\rho)|^2 \rho d\rho}, \qquad (26)$$

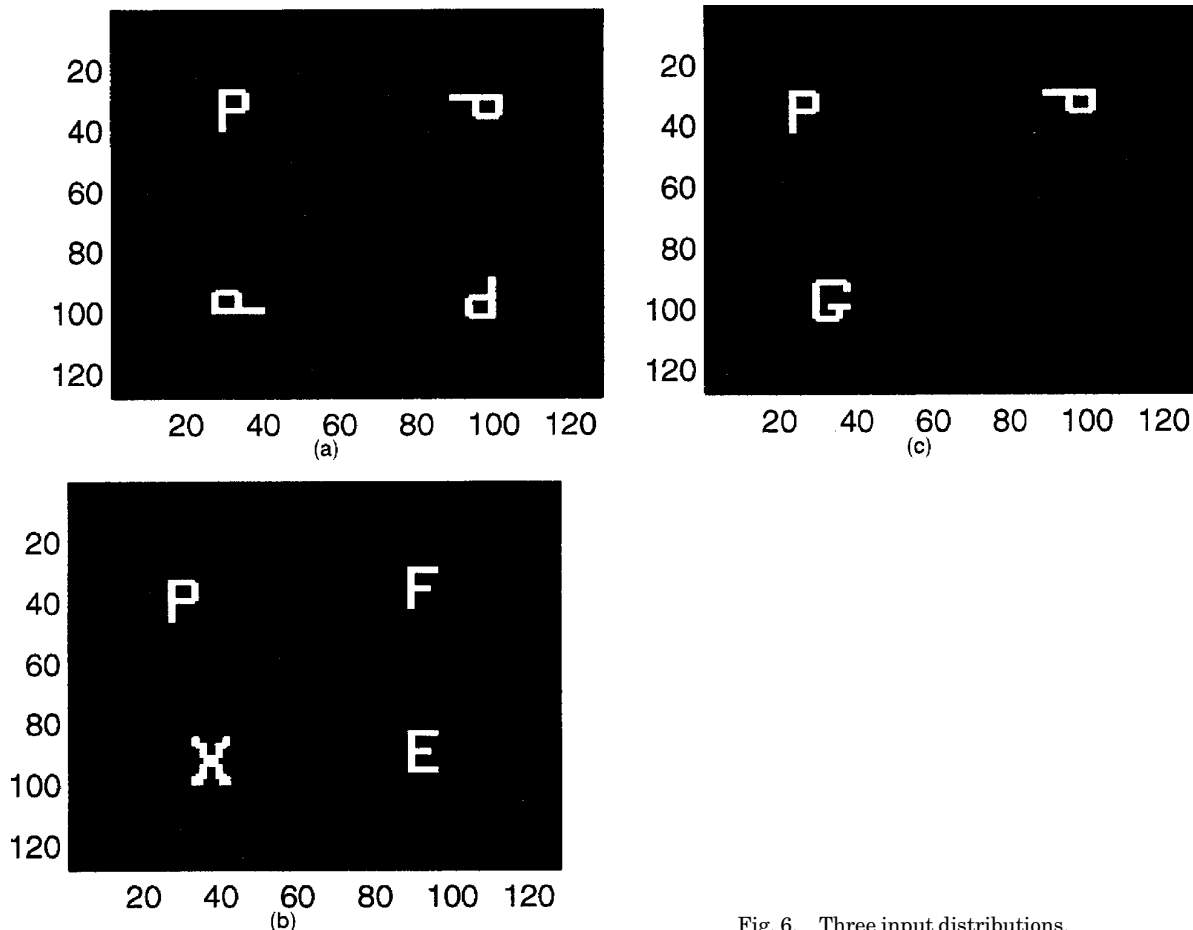where the subscript denotes the order of the filter and

Fig. 6.   Three input distributions.

the superscript P denotes that we are dealing with a PEC system.

To compare this result with the conventional LC, we take the phase-only CHC filter.[2]   Defining

$$B_N(\rho) = |B_N(\rho)| \exp[i \ \arg[B(\rho)]]$$

$$= \frac{1}{2\pi} \int_0^{2\pi} |A(\rho, \theta)| \exp[i\gamma(\rho, \theta)] \exp(-iN\theta) d\theta,$$

$$(27)$$

we obtain the linear phase-only CHC filter distribution as

$$H_N''(\rho, \varphi) = \exp(-i[N\varphi + \arg[B_N(\rho)]]), \qquad (28)$$

which is, in general, different from the filters for the PEC.

The peak sharpness measure for this linear filter is given by

$$\text{PCE}_N^L = \frac{1}{2\pi} \frac{\left[\int_0^{\rho_0} |B_N(\rho)| \rho d\rho \right]^2}{\sum\limits_{m=-\infty}^{\infty} \int_0^{\rho_0} |B_m(\rho)|^2 \rho d\rho} . \qquad (29)$$

Figure 6 shows some input distributions.   In one of the simulation experiments performed, the input distributions of Figs. 7(a) and 7(b) were used with CHC filters of various orders $(N = 0, 1, 2)$ prepared for the letter P.   For illustrative purposes, the correlation-plane distribution for the different filters, of the order of $N = 1$, with the corresponding input patterns are shown in Fig. 7 for the LC and in Fig. 8 for the PEC.

A comparison of the two figures clearly demonstrates that both the rejection, where rejection =
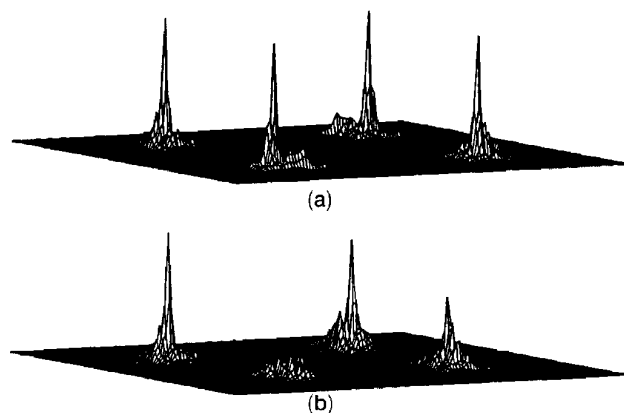

(a)

(b)

Fig. 7.   Output correlation distributions for the LC corresponding to the input patterns of Figs. 6(a) and 6(b), with the appropriate CHC POF matched to the letter P of the order of $N = 1$.
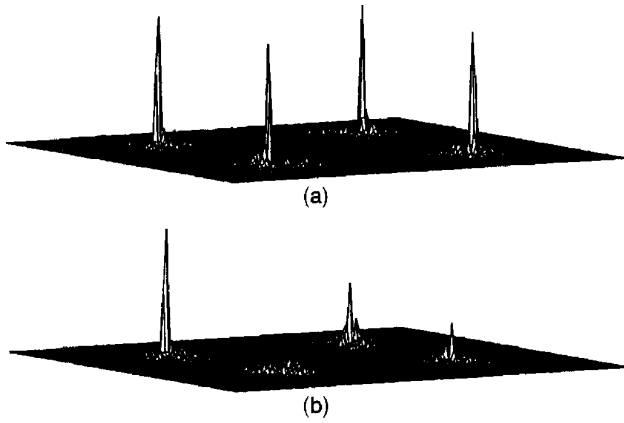
(a)

(b)

Fig. 8. As Fig. 7, but for the PEC.

(peak corresponding to P)/(largest other peak), and the sharpness of the peak are substantially superior for the PEC compared with those of the conventional LC. The numerical comparison of the two results for $N = 0, 1, 2$ is summarized in Table 1.

It is also interesting to compare the two correlators with respect to the energy distribution among the various harmonic components. For the sake of comparison we normalize the function at the Fourier plane, $|A(\rho, \theta)| \leq 1$, and obtain, by Parseval's theorem,

$$\int_0^{2\pi} \int_0^{R_0} |A(\rho, \varphi)\exp[i\gamma(\rho, \varphi)]|^2 \rho d\rho d\varphi$$

$$= 2\pi \sum_{N=-\infty}^{\infty} \int_0^{R_0} |B_N(\rho)|^2 \rho d\rho$$

$$< \int_0^{2\pi} \int_0^{R_0} |\exp[i\gamma(\rho, \varphi)]|^2 \rho d\rho d\varphi$$

$$= 2\pi \sum_{N=-\infty}^{\infty} \int_0^{R_0} |A_N(\rho)|^2 \rho d\rho. \qquad (30)$$

This relation shows that under our present normalization the energy in the correlation plane generated by the PEC is larger than that of the LC. Thus, effectively, the phase-extraction operation amplified the energy contained in the sum of all the harmonics. This is not necessarily true for each individual harmonic, and it may very well be that some of the

harmonics have higher energy in the LC than in the PEC. Thus, in order to compare on a fair basis the performances of the PEC and LC, we look for harmonics with similar fractional energy, which is defined by the energy measure EM:

$$\text{EM} = \frac{\text{the energy in the } N\text{th order}}{\text{the total energy}} . \qquad (31)$$

For the LC this measure is given by

$$\text{EM}_N^L = \frac{\displaystyle\int_0^{\rho_0} |B_N(\rho)|^2 \rho d\rho}{\displaystyle\sum_{M=-\infty}^{\infty} \int_0^{\rho_0} |B_M(\rho)|^2 \rho d\rho} , \qquad (32)$$

whereas, for the PEC, we have

$$\text{EM}_N^P = \frac{\displaystyle\int_0^{\rho_0} |A_N(\rho)|^2 \rho d\rho}{\displaystyle\sum_{M=-\infty}^{\infty} \int_0^{\rho_0} |A_M(\rho)|^2 \rho d\rho} . \qquad (33)$$

Because for both correlators we used POF's matched to a certain CHC, it cancels the phase of that CHC. The result is that most of the energy contained in that component is concentrated into the correlation peak, and thus the EM establishes a fair measure for comparison between the PEC and the LC.

Table 2 shows that for $N = 1$ the EM of the LC is similar to the EM of the PEC for $N = 2$. Nevertheless, the sharpness, as defined by the PCE criterion, is 40% better in the PEC. The rejection of the PEC is 2.6 times as much as that offered by the LC. Another interesting example illustrating the superiority of the PEC over the LC is for $N = 0$. Despite the fact that the EM of the LC is almost 3 times larger than the EM of the PEC, both the PCE and the rejection are better for the PEC. As expected, it is also evident from the tables that the lower the EM, the lower the corresponding PCE in the given correlator. Another interesting point is that the energy in the PEC is far less concentrated in the lower orders than in the LC. Thus in the PEC we may use higher-order CHC's, which leads to better discrimination (as the differences between objects are usually in the higher spatial frequencies) without severe degradation in the peak intensity. It should also be noted

Table 1. PCE and Rejection Measurements for the LC and the PEC for Various Orders of CHC's[a]

	PCE		Rejection	
N	LC	PEC	LC	PEC
0	0.117	0.148	1.33	3.36
1	0.05	0.14	1.17	2.09
2	0.012	0.069	1.06	3.14

[a]Rejection = (peak corresponding to P)/(largest other peak).

Table 2. EM for the LC and the PEC for Various Orders of the CHC's

	EM	
N	LC	PEC
0	0.61	0.22
1	0.11	0.16
2	0.0029	0.12

that in our simulations the problem of proper center[1] was much less severe than in the LC. Whereas in the LC, displacements of the expansion center by ~20% of the pattern size caused up to 300% variation in the correlation-peak intensity, the variation for a similar change observed in the PEC was only 30%. Thus in the PEC it is not so important to search for a proper center.

5. Rotation-Invariant Class Discrimination

Section 4 described a rotation-invariant non-LC based on circular harmonic decomposition, with extremely high discrimination. However, if this method is applied to the problem of class discrimination, in which different objects should generate similar correlation peaks, performance is substantially reduced. In this section we demonstrate that this and other difficult problems can be handled by our procedures demonstrated in Section 3. As an example, we design a rotation-invariant filter to discriminate between a class containing the letters P, F, and X from the class containing the letter E by using the GPEC. Note that the members of the first class have much less features in common among them than with the second class (X is completely distinct from P and F, whereas E is quite similar to the latter two and is difficult to discriminate by conventional spatial filtering). Looking at Fig. 8(b), we find that the task is not simple. It may be easily shown that using a filter of the form $H(\rho, \varphi) = H(\rho)\exp(iN\varphi)$, where $H(\rho)$ is not necessarily any of the Nth CHC's of the Fourier phases of the letters, still leads to an invariant correlation response. What is desired is that the filter will, basically, detect the Nth harmonic of P (constraint $C_{p-\text{det}}$) and be orthogonal to the differences in the Nth harmonic between P and F (constraint $C_{pf-\text{rej}}$) and P and X (constraint $C_{px-\text{rej}}$), where differences and orthogonality are defined by integrals.

This leads us to the following set of constraints:

$$C_{p-\text{det}} := \left\{ H(\rho) \middle| \int H(\rho)\alpha_p(\rho)\rho d\rho = \text{const}, \quad \text{const} \in \text{IR} \right\},$$
(34a)

$$C_{pf-\text{rej}} := \left\{ H(\rho) \middle| \int H(\rho)\alpha_{pf}(\rho)\rho d\rho = 0 \right\},$$
(34b)

$$C_{px-\text{rej}} := \left\{ H(\rho) \middle| \int H(\rho)\alpha_{px}(\rho)\rho d\rho = 0 \right\},$$
(34c)

where $\alpha_p(\rho) = A_N^p(\rho)$, $\alpha_{pf}(\rho) = A_N^p(\rho) - A_N^f(\rho)$, $\alpha_{px}(\rho) = A_N^p(\rho) - A_N^x(\rho)$, and the superscript over A_N indicates what letter the expansion of the phase corresponds to, where A_N is as given in Eq. (21b). The other constraint required in the design process is that the filter correspond to a passive element C_{pas} [this constraint was described above in Eq. (15c) and is not repeated]. We note that the set of constraints in Eqs. (34a)–(34c)

is equivalent to

$$C_{p-\text{det}} := \left\{ H(\rho) \middle| \int H(\rho)\alpha_p(\rho)\rho d\rho = \text{const} \right\}, \quad (35a)$$

$$C_{f-\text{det}} := \left\{ H(\rho) \middle| \int H(\rho)\alpha_f(\rho)\rho d\rho = \text{const} \right\}, \quad (35b)$$

$$C_{x-\text{det}} := \left\{ H(\rho) \middle| \int H(\rho)\alpha_x(\rho)\rho d\rho = \text{const} \right\}, \quad (35c)$$

and $\alpha_f(\rho) = A_N^f(\rho)$, $\alpha_x(\rho) = A_N^x(\rho)$.

As all constraints are convex (and the projections are simple to perform with the Euclidean norm as a metric) one may employ the serial-projection algorithm described in Subsection 2.A. Hence, iterating the composition operator,

$$T[H(\rho)] := P_{C_{\text{pas}}} P_{C_{p-\text{det}}} P_{C_{pf-\text{rej}}} P_{C_{px-\text{rej}}}[H(\rho)], \quad (36)$$

generates iterates converging to a solution. The projections onto the sets described by Eqs. (34a)–(34c) are readily solvable by the use of Lagrange optimization techniques. These projections are derived and given in appendix B [see Eqs. (B10) and (B13) for $P_{C_{p-\text{det}}}$ and Eqs. (B10) and (B14) for $P_{C_{pf-\text{rej}}}$ and $P_{C_{px-\text{rej}}}$].

We generated a filter by iterating the composition operator T defined by Eq. (36). When placing it as the filter in the GPEC with the input shown in Fig. 6(a), we obtained the correlation output distribution shown in Fig. 9(a), demonstrating the full rotation invariance required. When changing the input to that of Fig. 6(b), we see that the letters P, F, and X are detected, with a strong rejection of the letter E, i.e., the ratio of the lowest peak from the detection class to the largest peak from the rejected class is (peak of letter F)/(peak of letter E) = 2.22, as shown in Fig. 9(b). Clearly, the more dissimilar the objects from the rejection class (from those of the recognition class), the greater the rejection, as is demonstrated in Fig. 9(c).

We note that in the design process, as given by the constraint sets in Eqs. (34a)–(34c), it was implicitly assumed that each letter would be presented in the input alone. However, the actual inputs, e.g., Fig. 6(b), which is composed of multiple inputs presented simultaneously, do not adhere to this implicit assumption. Therefore, because of the lack of strict shift invariance of the GPEC, the output correlation distribution is not the superposition of each of the correlation distributions generated by each input alone. Hence, despite the synthesis algorithm that arrives at a solution that satisfies Eqs. (34a)–(34c), the actual correlation result from the designed filter and the input of Fig. 6(b) does not generate identical correlation peaks in the center of positions of the letters P, F, and X; they are only approximately the same [see Fig. 9(b)], as expected.[20]
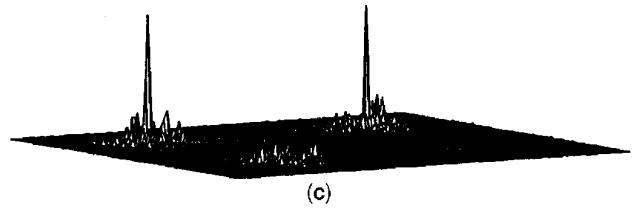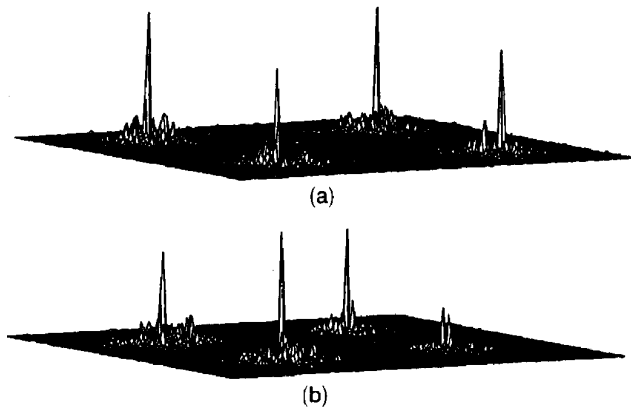
(a)

(b)

(c)

Fig. 9. (a)–(c) Output correlation distributions for the GPEC that correspond to the input patterns of Figs. 6(a)–6(c), respectively, but with the filter generated by the serial POCS algorithm, for the GPEC.

This last synthesis example demonstrates that, despite the wide class of letters to be recognized, adequate discrimination against other letters, although similar, is still maintained. Clearly, discrimination constraints could have been employed in the design process as well, enhancing the rejection capabilities. Also note that, although the appropriate CHC POF, matched to A_N^P, discriminated quite strongly against the letter F (and of course against letters which are more dissimilar, like X) [see Fig. 8(b)], we were able to design a filter that recognizes all letters from the recognition class similarly [Fig. 9(b)].

6. Conclusions

The powerful procedure of some new enhanced projection algorithms was shown to be suitable for the design of a wide variety of spatial filters, either for linear or nonlinear PR processes. The described procedures can, in principle, be used to design filters with arbitrary requirements, as long as they do not violate physical principles. Moreover, if the requirements are not consistent among themselves or with the physical principles, the designed filter will approach the requirements, i.e., be at the smallest average distance from the requirements.[15] The average distance is defined in terms of the proper distance functions used in the algorithm, i.e., the solution generated by algorithm 2 is a global minimizer of $\hat{J}$, defined by Eq. (7).[15]

The procedures were demonstrated by several PR cases, some of which were not achieved earlier.

Appendix A.   Multidistance Product-Space Formalism

1.   Introduction and Basic Definitions

The product-space formalism, originally due to Pierra,[25] was subsequently generalized by Censor and Elfving [10], although in a finite dimensional Hilbert space. For definity we assume $\mathscr{H} = L^2(\mathrm{IR})$ consisting of square integrable functions, N convex constraint sets C_i, $C_i \subseteq \mathscr{H}$, and N, possibly different, weighted norm-squared generalized distance functions d_i, where each set C_i is closed with respect to the corresponding generalized distance function d_i. Throughout, for the analogous construction for practical purposes, i.e., l^2 for the discrete case and IR^N for

the finite-dimensional case, integration should be replaced by summation.

2.   Construction of the Product Space

Denoting quantities that correspond to the product space by boldface type, we define the product space:

$$\mathbf{H} := \prod_{i=1}^{N} \mathscr{H}_i, \tag{A1}$$

which consists of compound functions of the form $\mathbf{h} = (h_1, h_2, \ldots, h_N)$, $h_i \in \mathscr{H}_i$. We also define the convex set $\mathbf{C} := \Pi_{i=1}^{N} C_i$; $\mathbf{C} \in \mathbf{H}$, where the N constituent constraint sets are usually different. A natural way to embed a vector $h \in \mathscr{H}$ in the product space is by associating with h an element in the product space that consists of N copies of h, namely $\boldsymbol{\tau}(h)$, where

$$\boldsymbol{\tau}(h) := \underbrace{(h, h, \ldots, h)}_{N \text{ times}}. \tag{A2}$$

This provides a convenient means for studying the properties of h. However, in general, the N components of functions in the product space, which are not necessarily of the form $\mathbf{h} = \boldsymbol{\tau}(h)$, will be different. It is useful to define the subspace $\boldsymbol{\Delta}$, which consists of functions that have N identical components, i.e., $\boldsymbol{\Delta} = \{\mathbf{h} | \mathbf{h} = \boldsymbol{\tau}(h); \quad h \in \cap_{i=1}^{N} \mathscr{H}_i\}$.

Define the weighted generalized distance function in the product space:

$$\mathbf{D}(\mathbf{h}_1, \mathbf{h}_2) = \sum_{i=1}^{N} \beta_i d_i(h_{1i}, h_{2i});$$

$$h_{1i}, h_{2i} \in \mathscr{H}_i; \quad \mathbf{h}_j := (h_{j1}, h_{j2}, \ldots, h_{jN}), \quad j = 1, 2, \tag{A3}$$

where $\{\beta_i\}_{i=1}^{N}$ are as in algorithm 2. Then $\sqrt{\mathbf{D}}$ is the norm in the product space, and $\mathbf{H}$ is a closed inner product space with respect to $\mathbf{D}$, i.e., a Hilbert space (Ref. 15, Section 4). $\mathbf{C}$ is a closed convex subset of $\mathbf{H}$ with respect to $\mathbf{D}$ that is due to the closedness and the convexity of C_i with respect to d_i, respectively. By a general Hilbert-space theory, $\boldsymbol{\Delta}$ is a closed convex

subset of **H** with respect to **D** as well, in fact, a closed linear subspace (Ref. 15, Section 4).

Using this formalism, one may define a projection in the product space onto an arbitrary set **S**, closed with respect to **D**, as follows. $\mathbf{h}' \in \mathbf{S}$ is the projection of **h** onto **S**, denoted by

$$\mathbf{P_S^D}(\mathbf{h}) = \mathbf{h}' \qquad \text{if and only if}$$

$$\inf_{\mathbf{h_1} \in \mathbf{S}} \mathbf{D}(\mathbf{h_1}, \mathbf{h}) = \mathbf{D}(\mathbf{h}', \mathbf{h}). \quad \text{(A4)}$$

If **S** is **C** (or **Δ**) then the projection onto **C** (or **Δ**) exists and is unique because of the closedness and convexity of the sets.[26]

It follows from Eq. (A3) that, for any $h \in \mathscr{H}$,

$$\inf_{\mathbf{h_1} \in \mathbf{C}} \mathbf{D}[\mathbf{h_1}, \boldsymbol{\tau}(h)] = \sum_{i=1}^{N} \beta_i \big[\inf_{h_{1i} \in C_i} d_i(h_{1i}, h) \big]. \quad \text{(A5)}$$

Using Eqs. (A5) and (8) it is not difficult to derive an explicit expression for a projection in the product space onto **C** (see Lemma 4.1 in Ref. 10).   This is obtained when parallel projections are performed on the individual sets C_i, i.e.,

$$\mathbf{P_C^D}[\boldsymbol{\tau}(h)] = [P_{C_1}^{d_1}(h), P_{C_2}^{d_2}(h), \dots, P_{C_N}^{d_N}(h)]. \quad \text{(A6)}$$

The projection $\mathbf{P_\Delta^D}$ can be expressed in a similar way with the relation (see Lemma 4.2 in Ref. 10 and the derivation in Sections 3 and 4 in Ref. 15)

$$\mathbf{h}' = \mathbf{P_\Delta^D}(\mathbf{h}) \qquad \text{if and only if} \qquad \mathbf{h}' = \boldsymbol{\tau}(h')$$

$$\text{where } \nabla_{h_1} \mathbf{D}\left[\boldsymbol{\tau}(h_1), \mathbf{h}\right]\big|_{h_1 = h'} = 0. \quad \text{(A7)}$$

This leads to

$$\mathbf{h}' = \boldsymbol{\tau}(h'), \quad \text{where } h' := \mathscr{T}^{-1}\left\{ \frac{\sum_{i=1}^{N} \beta_i W_i(u) V_i(u)}{\sum_{i=1}^{N} \beta_i W_i(u)} \right\}.$$

Here ∇_{h_1} denotes the gradient.

3.   Projections onto Nonconvex Sets

Define the relaxed projection operators by

$$P_{C,\lambda}(h) := P_C(h) + \lambda[h - P_C(h)]. \quad \text{(A8)}$$

We have the following theorem, which is from Levi and Stark, Ref. 12, p. 934:

Theorem 1:   Given a Hilbert space $\mathscr{H}$ with an inner product $\langle h_1, h_2 \rangle$, $h_1, h_2 \in \mathscr{H}$.   Let C_1, C_2 be two closed subsets of $\mathscr{H}$, not necessarily convex, with relaxed projection operators given by $T_1 := T_{C_1, \lambda_1}(h)$, $T_2 := T_{C_2, \lambda_2}(h)$, respectively.   Then a recursion of the form

$$h^{k+1} := T_1[T_2(h^k)]; \qquad k \geq 0, \quad \text{(A9)}$$

has the property that the summed-distance error

functional, J, given by

$$J(h) := \sum_{i=1}^{2} \|h - P_{C_i}(h)\|, \quad \text{(A10)}$$

where $\|\cdot\|$ is the appropriate norm in the Hilbert space $\mathscr{H}$, is nonincreasing for any λ_1, λ_2 in the interval $[0, 1]$.

Now consider the following algorithm:
Algorithm 3:   Initialization: Let $\mathbf{h}^0 := \boldsymbol{\tau}(h^0)$, $h^0 \in \mathscr{H}$ arbitrary.
Iterative step: given the function $\mathbf{h}^k := \boldsymbol{\tau}(h^k)$, calculate

$$\mathbf{v}^{k+1} = \mathbf{P_{C,\lambda}^D}(\mathbf{h}^k),$$

and then, set

$$\mathbf{h}^{k+1} = \mathbf{P_\Delta^D}(\mathbf{v}^k), \quad \text{(A11)}$$

where $\mathbf{P_{C,\lambda}^D}(\mathbf{h}) := \mathbf{P_C^D}(\mathbf{h}) + \lambda[\mathbf{h} - \mathbf{P_C^D}(\mathbf{h})]$.

This algorithm performs two alternating operations in the product space: a relaxed projection onto **C** and a projection onto **Δ**.   Even if all sets C_i are not convex (implying that **C** is not convex), in any event, in the product space we have only two sets: **C** and **Δ**. Thus we may apply the theorem of Levi and Stark,[12] cited above, in the product space, where the Hilbert space is **H**, the distance function is the norm in the product space $\sqrt{\mathbf{D}}$, and the appropriate summed-distance error functional in the product space is

$$\mathbf{J}(\mathbf{h}) := [\mathbf{D}[\mathbf{P_C^D}(\mathbf{h}), \mathbf{h}]]^{1/2} + [\mathbf{D}[\mathbf{P_\Delta^D}(\mathbf{h}), \mathbf{h}]]^{1/2}. \quad \text{(A12)}$$

Because it can easily be shown that algorithms 2 and 3 are equivalent and that $\mathbf{h}^k$ generated by algorithm 3 is equal to $\boldsymbol{\tau}(h^k)$, where h^k is the sequence generated by algorithm 2, we have the following corollary.

Corollary 2:   The functional $\hat{J}(h^k)$ given by Eq. (7), which is equal to $\mathbf{J}(\mathbf{h}^k)$ given by Eq. (A12), is (monotonically nonincreasing) convergent along the iterates of the respective algorithm 2, i.e., $\{\hat{J}(h^k)\}_{k \geq 0}$ is a convergent sequence, for any sequence $\{h^k\}_{k \geq 0}$ generated by algorithm 2, irrespective of the sets C_i being convex or not.   Moreover, if all sets C_i are convex, then **C** is convex, and if $\cap_{i=1}^{N} C_i$ is nonempty then $\mathbf{C} \cap \mathbf{\Delta}$ is nonempty, and hence h^k converges to C_0.

Appendix B.   Some Specific Projections

Below we develop the projections onto the sets given by Eqs. (34a)–(34c).   Define the distance function by

$$d^2[H(\rho), H'(\rho)]$$

$$= \int_0^\infty [[H_r(\rho) - H'_r(\rho)]^2 + [H_i(\rho) - H'_i(\rho)]^2] \rho \, d\rho, \quad \text{(B1)}$$

which is just the Euclidean norm.

Let

$$A = \int_0^\infty [H_r(\rho)f_r(\rho) - H_i(\rho)f_i(\rho)]\rho d\rho, \qquad (B2)$$

$$B = \int_0^\infty [H_r(\rho)f_i(\rho) + H_i(\rho)f_r(\rho)]\rho d\rho, \qquad (B3)$$

where the subscripts r and i denote the real and imaginary parts, respectively:

$$H(\rho) = H_r(\rho) + iH_i(\rho); \qquad H'(\rho) = H'_r(\rho) + iH'_i(\rho),$$

$$f(\rho) = f_r(\rho) + if_i(\rho). \qquad (B4)$$

Writing the constraints of Eqs. (34) in full yields requirements of the following form:

$$\int_0^\infty [H'_r(\rho)f_r(\rho) - H'_i(\rho)f_i(\rho)]\rho d\rho = 0,$$

$$\int_0^\infty [H'_r(\rho)f_i(\rho) + H'_i(\rho)f_r(\rho)]\rho d\rho = 0 \qquad (B5)$$

for $C_{pf-\mathrm{rej}}$ and $C_{px-\mathrm{rej}}$;

$$\int_0^\infty [H'_r(\rho)f_r(\rho) - H'_i(\rho)f_i(\rho)]\rho d\rho \geq T_1,$$

$$\int_0^\infty [H'_r(\rho)f_i(\rho) + H'_i(\rho)f_r(\rho)]\rho d\rho = 0 \qquad (B6)$$

for $C_{p-\mathrm{det}}$.

We note that the functional form of the requirements in Eqs. (B5) and (B6) are similar. Therefore we develop the projection onto $C_{p-\mathrm{det}}$ first and then arrive by inspection at the projection onto $C_{pf-\mathrm{rej}}$ and $C_{px-\mathrm{rej}}$.

Rewriting the requirement of Eqs. (B6) yields

$$g(\rho) = \int_0^\infty [H'_r(\rho)f_r(\rho) - H'_i(\rho)f_i(\rho)]\rho d\rho - (T_1 + \xi) = 0,$$

$$(B7)$$

$$q(\rho) = \int_0^\infty [H'_i(\rho)f_r(\rho) + H'_r(\rho)f_i(\rho)]\rho d\rho = 0, \qquad (B8)$$

where we introduce the surplus variable ξ, which is confined by $\xi \geq 0$.

To determine the projection we must minimize Eq. (B1). By defining $\nabla^T \ (\partial/\partial H'_r, \ \partial/\partial H'_i)$, then, with the above notation, the Lagrange requirement becomes [remembering that our task is to minimize Eq. (B1) subject to the requirements given by Eqs. (B7) and (B8)],

$$\nabla d^2(H, \ H') = \lambda \nabla g + \mu \nabla q. \qquad (B9)$$

Performing the necessary derivatives given by the

∇ operator in Eq. (B9) and applying Eqs. (B7) and (B8) yield

$$H'_r(\rho) = H_r(\rho) + \frac{\lambda f_r(\rho) + \mu f_i(\rho)}{2}, \qquad (B10a)$$

$$H'_i(\rho) = H_i(\rho) + \frac{-\lambda f_i(\rho) + \mu f_r(\rho)}{2}. \qquad (B10b)$$

To determine λ, μ, we substitute Eqs. (B10a) and (B10b) into Eqs. (B7) and (B8) to find the λ, μ values that satisfy the necessary constraints. After some algebra we finally arrive at

$$\lambda = \frac{-2(A - T_1 - \xi)}{E}, \qquad (B11a)$$

$$\mu = \frac{-2B}{E},$$

$$E = \int_0^\infty |f(\rho)|^2 \rho d\rho. \qquad (B11b)$$

Substituting the values of λ, μ from Eqs. (B11) into Eqs. (B10) and minimizing Eq. (B1) yield

$$d(H, H') = \min \frac{\lambda^2 + \mu^2}{4} E. \qquad (B12)$$

Thus λ is the value that will minimize λ^2 subject to the constraint that $\xi \geq 0$. μ is given by Eq. (B11b). To determine λ, we examine two cases:

(a) $A \leq T_1$.
In this case, λ^2, as given by Eq. (B11a), is a monotonic increasing function of ξ for $\xi \geq 0$, λ^2 attaining its minimum with $\xi = 0$, $\lambda = -2(A - T_1)/E$.

(b) $A > T_1$.
In this case we may choose $\xi = A - T_1$ and $\lambda = 0$. Hence we have

$$\lambda = 0, \qquad \text{if } A > T_1$$

$$\lambda = \frac{+2(T_1 - A)}{E} \qquad \text{if } A \leq T_1 \qquad , \qquad (B13a)$$

$$\mu = \frac{-2B}{E}. \qquad (B13b)$$

Thus, finally, $P_{C_{p-\mathrm{det}}}[H(\rho)] = H'(\rho)$ is given by Eqs. (B10), with λ, μ given by Eqs. (B13), where the values of A and B are determined by Eqs. (B2) and (B3) and $f(\rho) = \alpha_p(\rho)$.

By analogy we identify that $P_{C_{pf-\mathrm{rej}}}$, $P_{C_{px-\mathrm{rej}}}$ follows the same analysis with $T_1 = 0$ and $\xi = 0$; hence we get

$$\lambda = \frac{-2A}{E}, \qquad (B14a)$$

$$\mu = \frac{-2B}{E}. \qquad (B14b)$$

Thus finally $P_{pf-\mathrm{rej}}[H(\rho)] = H'(\rho)$ and $P_{px-\mathrm{rej}}[H(\rho)] = H'(\rho)$ are given by Eqs. (B10) with λ, μ given by Eqs. (B14), where the values of A and B are determined by Eqs. (B2) or (B3) and $f(\rho) = \alpha_{pf}(\rho)$ or $f(\rho) = \alpha_{px}(\rho)$ for $P_{pf-\mathrm{rej}}$ and $P_{px-\mathrm{rej}}$, respectively.

References

1. Y. Hsu, H. H. Arsenault, and G. April, "Rotation invariant digital pattern recognition using circular harmonic expansion," Appl. Opt. **21,** 4012–4015 (1982).
2. J. Rosen and J. Shamir, "Circular harmonic phase filter for efficient rotation invariant pattern recognition," Appl. Opt. **27,** 2895–2899 (1988).
3. D. Casasent and D. Psaltis, "New optical transforms for pattern recognition," Proc. IEEE **65,** 77–84 (1977).
4. M. Fleisher, U. Mahlab, and J. Shamir, "Target location measurement by optical correlators: a performance criterion," Appl. Opt. **31,** 230–235 (1992).
5. J. Rosen and J. Shamir, "Application of the projection-onto-constraint-sets algorithm for optical pattern recognition," Opt. Lett. **16,** 752–754 (1991).
6. J. Rosen, "Learning in correlators based on projections onto constraint sets," Opt. Lett. **18,** 1183–1185 (1993).
7. T. Kotzer, N. Cohen, and J. Shamir, "Image reconstruction by a novel parallel projection onto constraint sets method," EE Pub. 919 (Dept. of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, June 1994).
8. T. Kotzer, N. Cohen, J. Shamir, and Y. Censor, "Multi-distance, Multi-projection, parallel projection method," presented at the International Conference on Optical Computing-OC94, Edinburgh, Scotland, August, 1994.
9. T. Kotzer, N. Cohen, and J. Shamir, "Signal synthesis and reconstruction by projection methods in a hyper space," presented at the Annual Meeting of the Optical Society of America, Dallas, Tex., October 1994.
10. Y. Censor and T. Elfving, "A multiprojection algorithm using Bregman projections in a product space," Numerical Algorithms **8,** 221–239 (1994).
11. T. Kotzer, J. Rosen, and J. Shamir, "Phase extraction pattern recognition," Appl. Opt. **31,** 1126–1137 (1992).
12. A. Levi and H. Stark, "Image restoration by the method of generalized projections with application to restoration from magnitude," J. Opt. Soc. Am. A **1,** 932–943 (1984).
13. D. C. Youla and H. Webb, "Image restoration by the method of convex projections: part 1—theory," IEEE Trans. Med. Imag. **TMI-1,** 81–94 (1982).
14. T. Kotzer, N. Cohen, and J. Shamir, "Extended and alternative projections onto convex constraint sets: theory and applications," EE Pub. 900, (Dept. of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, November 1993).
15. T. Kotzer, N. Cohen, and J. Shamir, "A projection algorithm for consistent and inconsistent constraints," EE Pub. 920 (Dept. of Electrical Engineering, Technion Israel Institute of Technology, Haifa, August 1994).
16. T. Kotzer, N. Cohen, J. Shamir, and Y. Censor, "Summed distance error reduction of simultaneous multiprojections and applications," EE Pub. 909 (Dept. of Electrical Engineering, Technion—Israel Institute of Technology, Haifa, Israel, August 1994).
17. B. Javidi, "Nonlinear joint power spectrum based optical correlation," Appl. Opt. **28,** 2358–2367 (1989).
18. B. Javidi and J. Wang, "Binary nonlinear joint transform correlation with median and subset thresholding," Appl. Opt. **30,** 967–976 (1991).
19. W. B. Hahn and D. L. Flannery, "Design elements of binary joint transform correlation and selected optimization techniques," Opt. Eng. **31,** 896–905 (1992).
20. T. Kotzer, J. Rosen, and J. Shamir, "Multiple-object input in nonlinear correlation," Appl. Opt. **32,** 1919–1932 (1993).
21. E. Silvera, T. Kotzer, and J. Shamir, "Adaptive pattern recognition with rotation, scale and shift invariance," Appl. Opt. **34,** 1891–1900 (1995).
22. J. Rosen, T. Kotzer, and J. Shamir, "Optical implementation of phase extraction pattern recognition," Opt. Commun. **83,** 10–14 (1991).
23. J. L. Horner and P. D. Gianino, "Phase only matched filtering," Appl. Opt. **23,** 812–816 (1984).
24. B. V. K. V. Kumar, W. Shei, and C. Hendrix, "Phase only matched filters with maximally sharp correlation peaks," Opt. Lett. **15,** 807–809 (1990).
25. G. Pierra, "Decomposition through formalization in a product space," Math. Prog. **28,** 96–115 (1984).
26. N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space* (Ungar, New York, 1961), Vol. I.