Application of the projection-onto-constraint-sets algorithm for optical pattern recognition

Joseph Rosen and Joseph Shamir

Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel

Received January 23, 1991

Projection-onto-constraint sets is an efficient algorithm for constructing synthetic discriminant functions to be employed in pattern-recognition systems. The algorithm is implemented by a digital procedure based on a simulated joint-transform correlator.

Projection-onto-constraint sets (POCS) has been used for many years in several areas of signal processing. Recently the use of the POCS was extended to new fields, such as computer-generated holograms[1] and neural networks.[2] In this Letter we show that it is also possible to employ the POCS for the generation of synthetic discriminant functions[3] (SDF's) for pattern recognition. Many algorithms[3] used in computing a SDF consider only the value at one point, usually the origin of the correlation plane. Other algorithms[4,5] are made more global by using a cost function that contains information about the whole correlation plane, but they still do not control the complete correlation distribution. The POCS is a process that creates, in every iteration, the total correlation distribution and that may form any desired shape of correlation function. It has other advantages as well, as it is a relatively rapid iterative process, but it usually achieves suboptimal solutions.

The POCS is an iterative process that transfers a function or a vector from one domain to another and vice versa. In every domain it is projected onto one or several constraint sets. The convergence of the process, if it exists, is achieved when the function satisfies all the constraints in every domain simultaneously. The exact conditions on the constraint sets that are sufficient to guarantee weak convergence of the POCS process are described in Ref. 6. If all the sets are closed and convex, and they have at least one common term, then the process weakly converges.

In a typical pattern classification task we assume two object classes to be classified. In class A there are N patterns, and in class B there are M patterns. Our goal is to find a SDF that produces a sharp peak in the correlation plane when an object from class A is in the input of the system and a diffused distribution when an object from class B is present. Since the correlator is space invariant, we can handle the problem with all the $M + N$ objects presented at the input plane simultaneously. It is convenient to split this plane into two regions, where one contains the N objects of class A and the other contains the M objects of class B. As a result the correlation plane is also split into two regions that contain the respective correlation functions of the objects from class A and class B. The constraint set in this plane requires the appearance of only N bright correlation peaks corresponding to the centers of the correlations with the objects of class A. Hence, the central points of the correlation functions related to the objects of class A, designated region R_1, will be equal to or above a predetermined threshold level denoted by T_1. The distribution in a ring around these central peaks, designated region R_2, will remain unconstrained. All other points of the correlation function, designated region R_3, which are above a second threshold level T_2 ($T_2 < T_1$), will have the value T_2. A schematic representation that describes the various regions of the correlation plane is given in Fig. 1. Formally, a constraint set C_1 may be defined by the relation

$$C_1 = \{c(x') : c(x') \geq T_1 \text{ if } x' \in R_1; \; c(x') \leq T_2 \text{ if } x' \in R_3\}, \quad (1)$$

where $c(x')$ is the correlation function. The ratio between T_1 and T_2 determines the discrimination ratio

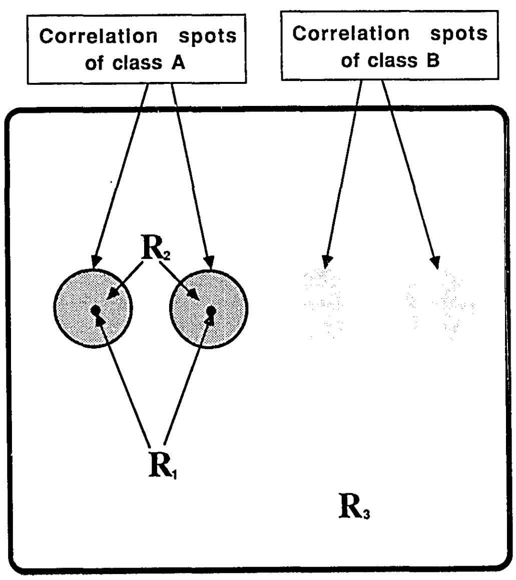


Fig. 1. Various regions in the correlation plane to be operated on by the operator P_1.

0146-9592/91/100752-03$5.00/0

and can be adjusted by the user during the synthesis process. The projection operator of the correlation domain P_1 is defined as

$$P_1[c(x')] = \begin{cases} T_1 & \text{if } x' \in R_1 \text{ and } c(x') < T_1 \\ T_2 & \text{if } x' \in R_3 \text{ and } c(x') > T_2. \\ c(x') & \text{otherwise} \end{cases} \quad (2)$$

A convenient architecture to implement the POCS is the joint-transform correlator[7] (JTC), since a single plane includes the input as well as a reference function. A pattern to be recognized is represented by a function $f(x)$ at a distance b from the center of the input plane. A reference pattern $h(x)$ of width w_h is introduced at a distance $-b$ from the center. Thus the overall distribution at the input plane is $s(x) = f(x - b) + h(x + b)$. The squared magnitude of the Fourier transform of $s(x)$, $|S(u)|^2$, is Fourier transformed again to yield the complex amplitude distribution over the output plane,

$$\begin{aligned} c(x') &= f(x') \star f(x') + h(x') \star h(x') \\ &\quad + f(x') \star h(x' - 2b) + h(x') \star f(x' + 2b), \end{aligned} \quad (3)$$

where $\star$ denotes correlation. This distribution contains three terms that are spatially separated, with two of them representing the desired correlation between $f(x)$ and $h(x)$. In our POCS procedure we perform the operation P_1 [Eq. (2)] on the two correlation regions simultaneously.

The second domain of the POCS is the JTC input plane. In the first example the constraint set over the input plane contains all the patterns of the training set arranged around a point at a distance b from the origin and a space-limited real reference function around the point $-b$. This can be written as

$$C_2 = \left\{ s(x) : s(x) = g(x - b) + h(x + b); \right.$$
$$\left. h(x) \in \mathcal{R};\, h(x) = 0 \text{ if } |x| > \frac{\omega_h}{2} \right\}. \quad (4)$$

$g(x)$, having a width w_g, represents all the training set: $g(x) = \sum_{i_A=1}^{N} f_{i_A}(x - d_{i_A}) + \sum_{i_B=1}^{M} f_{i_B}(x - d_{i_B})$, where $d_{i_{A,B}}$ is the distance of $f_{i_{A,B}}$ from the origin. The operation of the projection operator P_2 on $s(x)$ is given by

$$P_2[s(x)] = \begin{cases} s(x) & \text{if } \left(-b - \dfrac{w_h}{2}\right) < x < \left(-b + \dfrac{w_h}{2}\right) \\ g(x - b) & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}. \quad (5)$$

The complete POCS process as implemented on a JTC-like configuration is described in Fig. 2. We start the process with an input plane containing all the patterns of the training set clustered in two subregions (for the two classes, A and B) and a random real space-limited reference function. In the kth iteration, $s_k(x)$ is Fourier transformed to $S_k(u)$. The phase distribution of $S_k(u)$ is kept in the memory, while the magnitude is squared and inverse Fourier transformed to $c_k(x')$. In this stage the operator P_1 [Eq. (2)] operates.

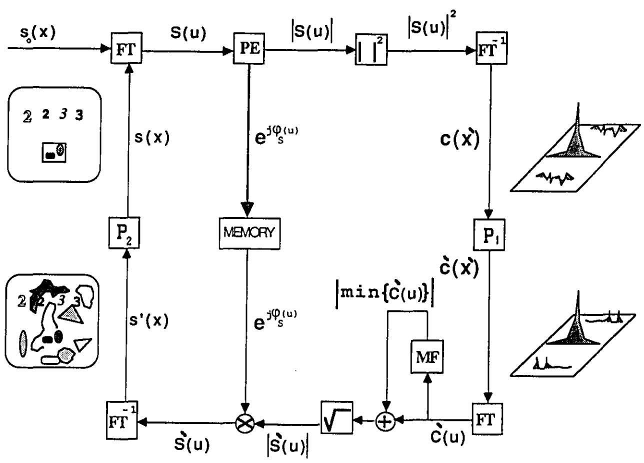
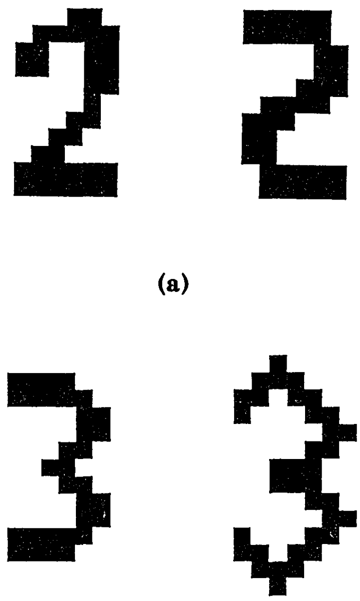


Fig. 2. Block diagram of the POCS process (see the text for details). FT, Fourier transform; FT^{-1}, inverse Fourier transform; PE, phase extractor; MF, minimum finder.



(a)

(b)

Fig. 3. Input training set. (a) Patterns to be detected, (b) patterns to be rejected.

The distribution at the central region, of width $(w_h + w_g)$, remains unchanged. The result of the projection, $c_k'(x')$, is Fourier transformed back to the spatial frequency domain. Since we started from this domain with $|S_k(u)|^2$, the new function $C_k'(u)$ must be positive, which is achievable by subtracting the minimum value of $C_k'(u)$. The next step is to take the square root of $C_k'(u)$ to obtain $|S_k'(u)|$, which is then multiplied by the phase function stored in memory. The product is inverse Fourier transformed back to the input plane to

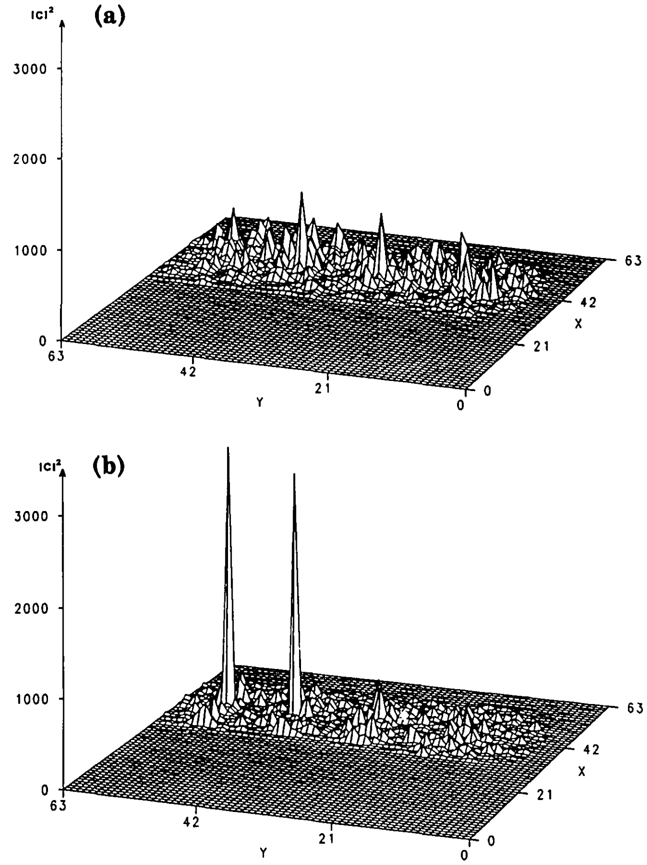


Fig. 4. Output correlation planes (in arbitrary units). (a) After the first iteration, (b) after 60 iterations.

be operated on by the operator P_2. The $(k + 1)$th reference function is obtained by reducing to zero all the function $s_k'(x)$ except for the limited region of width w_h, centered at a distance $-b$ from the origin. The input plane for the next iteration contains this reference function and the original training set. The process proceeds in the same way until we obtain a space-limited reference function that produces the desired correlation distribution, determined by T_1 and T_2. Both C_1 [Eq. (1)] and C_2 [Eq. (4)] are convex. However, the problem is that a solution does not always exist for every training set.

In our initial experiment we chose two versions of the digit 2 for class A and two versions of the digit 3 for class B, as shown in Fig. 3. In Fig. 4(a) we see the correlation plane after the first iteration. The same region is shown in Fig. 4(b), which after 60 iterations reveals the two strong correlation peaks related to the digits 2.

As a second example we tried to achieve, directly by the POCS algorithm, a binary reference function. For the case of binary (0, 1) function, C_2 can be rewritten as

$$C_2' = \Big\{ s(x) : s(x) = g(x - b) + h(x + b); \quad h(x) \in \{0, 1\}; h(x) = 0 \text{ if } |x| > \frac{w_h}{2} \Big\}. \tag{6}$$

It is easy to see that the modified constraint set is no longer a convex set, and, in fact, the algorithm did not converge under this constraint set in any trial that we made. Our conclusion is that it is better to calculate a gray-level reference function by the POCS algorithm and then to code it, if needed, into a binary form by one of the noniterative methods, rather than to obtain a binary reference function directly.

References

1. J. P. Allebach and D. W. Sweeney, Proc. Soc. Photo-Opt. Instrum. Eng. **884,** 2 (1988).
2. R. J. Marks II, Appl. Opt. **26,** 2005 (1987).
3. C. F. Hester and D. Casasent, Appl. Opt. **19,** 1758 (1980).
4. R. R. Kallman, Appl. Opt. **25,** 1032 (1986).
5. M. Fleisher, U. Mahlab, and J. Shamir, Appl. Opt. **29,** 2091 (1990).
6. D. C. Youla and H. Webb, IEEE Trans. Med. Imag. **MI-1,** 81 (1982).
7. C. S. Weaver and J. W. Goodman, Appl. Opt. **5,** 1248 (1966).