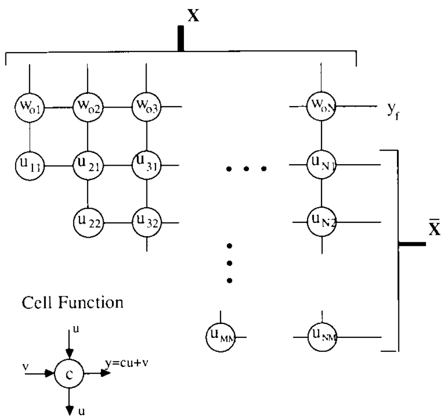

Fig. 6. Constraint preprocessor for large numbers of constraints with no permutations of input data. u_{ij} is the row i column j element of U given in (19). Systolic skew of input data not shown.

formers, each having the same number of degrees of freedom. Then one only needs to change the constants in the preprocessor to change the constraint set. This approach is applicable to any of the preprocessors in Figs. 2–6. The most efficient preprocessor depends on the feasibility of permuting the input data and the value of M relative to N. The preprocessors of Section III are more efficient for $M > N/2$ and those of Section IV for $M < N/2$.

V. Summary

Systolic preprocessing structures for linearly constrained beamforming are presented and discussed. They transform the constrained problem into unconstrained form, enabling an unconstrained adaptive processor to compute the beamformer output. The preprocessor proposed in [7] is shown to be a factored implementation of the GSC. Four alternative preprocessors are developed based on factorizations of the GSC for arbitrary constraint sets. Two are efficient when the number of adaptive degrees of freedom is close to the number of sensors, and the other two are efficient when the number of adaptive degrees of freedom is very small relative to the number of sensors, as is typical in partially adaptive beamforming. All four preprocessors implement an arbitrary nonadaptive beamformer.

References

[1] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays.* New York: Wiley, 1980.

[2] J. E. Hudson, *Adaptive Array Principles.* New York: Peter Peregrinus, 1981.

[3] K. M. Buckley, "Spatial/spectral filtering with linearly constrained minimum variance beamformers," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 249–266, Mar. 1987.

[4] J. G. McWhirter, "Recursive least-squares minimization using a systolic array," in *Real Time Signal Processing VI*, K. Bromley, Ed., *Proc. SPIE 431*, 1983, pp. 105–112.

[5] S. Haykin, *Adaptive Filter Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1986.

[6] C. R. Ward, P. J. Hargrave, and J. G. McWhirter, "A novel algorithm and architecture for adaptive digital beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 338–346, Mar. 1986.

[7] J. G. McWhirter and T. J. Shepherd, "A systolic array for constrained least-squares problems," in *Advanced Algorithms and Architectures for Signal Processing*, J. Speiser, Ed., *Proc. SPIE 696*, 1986, pp. 80–87.

[8] R. Schreiber and P. J. Kuekes, in *VLSI and Modern Signal Processing*, S. Y. Kung, H. J. Whitehouse, and T. Kailath, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[9] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 27–34, Jan. 1982.

[10] B. D. Van Veen and R. A. Roberts, "Partially adaptive beamformer design via output power minimization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1524–1532, Nov. 1987.

[11] W. F. Gabriel, "Adaptive arrays—An introduction," *Proc. IEEE*, vol. 64, pp. 239–272, Feb. 1976.

[12] L. J. Griffiths and K. M. Buckley, "Quiescent pattern control in linearly constrained adaptive arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 917–926, July 1987.

[13] D. J. Chapman, "Partial adaptivity for the large array," *IEEE Trans. Antennas Propagat.*, vol. AP-24, pp. 685–696, Sept. 1976.

[14] D. R. Morgan, "Partially adaptive array techniques," *IEEE Trans. Antennas Propagat.*, vol. AP-26, pp. 823–833, Nov. 1978.

[15] W. F. Gabriel, "Using spectral estimation techniques in adaptive processing antenna systems," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 291–300, Mar. 1986.

[16] B. D. Van Veen, "Eigenstructure based partially adaptive array design," *IEEE Trans. Antennas Propagat.*, vol. AP-36, pp. 357–362, Mar. 1988.

[17] N. L. Owsley, in *Array Signal Processing*, S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[18] R. L. Hanson and C. L. Lawson, "Extensions and applications of the Householder algorithm for solving linear least squares problems," *Math. Comput.*, vol. 23, pp. 787–812, 1969.

[19] S. Kalson and K. Yao, "A systolic array for linearly constrained least squares filtering," in *ICASSP'85*, Mar. 1985, pp. 977–980.

Binary Filters for Pattern Classification

SANTOSH S. VENKATESH and DEMETRI PSALTIS

***Abstract*—Generalized Fourier correlators imposing finite system space-bandwidth products are described, and a class of binary filters is proposed. In pattern classification and signal registration applications, it is shown that for a class of signals, the binary filters yield the same asymptotic performance as the matched filter. It is hence adduced that a dynamic range of a single bit in the filter suffices for classification purposes. The effects of statistical sidelobe fluctuations and a finite system space-bandwidth product are included in the analysis. It is demonstrated that performance improves in a natural fashion with increase in the system space-bandwidth product for both the binary filter and the matched filter.**

I. Introduction

Matched filters are commonly used in diverse applications in communication systems, signal processing, and pattern classification, where the task is typically the recognition of a particular signal or pattern immersed in noise. The principal theoretical argument supporting the use of matched filters is the classical result: among the class of all linear filters, matched filters maximize a (suitably defined) signal-to-noise ratio [1]. Practical implementations of matched filters—and linear, shift-invariant systems, in general—are greatly facilitated by the fundamental Fourier con-

Manuscript received November 8, 1985; revised August 9, 1988. This work was supported by the Army Research Office and General Dynamics.

S. S. Venkatesh was with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA. He is now with the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104.

D. Psaltis is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.

IEEE Log Number 8826103.

volution theorem wherein convolutions (or correlations) in one domain are transformed into products in the Fourier domain. As a consequence, relatively simple analog implementations such as optical Fourier-plane correlators [2], and digital implementations using algorithms such as the fast Fourier transform [3], abound.

The implementation of the system transfer function for the matched filter, however, requires a large dynamic range. A question of considerable theoretical and practical import is the determination of minimal complexity filters which have minimal dynamic range requirements, and for which good classification performance still attains (vis-à-vis the matched filter). The issue here is to determine the critical information needed for classification, and to discard redundant information. In this correspondence, we propose a class of low complexity binary filters which are a step toward the resolution of this question. These filters encode information in the phase of the Fourier transform of the desired signal and require a dynamic range of just one bit.

Our principal theoretical result concerning the binary filters is the following. *For statistically uncorrelated pattern classes, the binary filters provide the same asymptotic classification performance as the matched filter.* In fact, the binary filters provide classification performance comparable to (although bounded above by) the matched filter over all ranges.

These binary filters are of considerable practical importance. The requirement of a large dynamic range for the filter (corresponding to the many bits required to represent each sample point) is obviated, and just a single representation bit is utilized per sample point. The resultant decrease in required memory storage paves the way for low cost, low complexity systems—both digital and analog—which retain good classification performance. Of particular interest in optical filter implementations is the recent availability of a two-dimensional binary spatial light modulator—the magnetooptic device. We have demonstrated good classification in experimental optical correlators with our binary filters implemented using these devices [4].

In the next section, we define a general family of bounded space-bandwidth product Fourier correlators, and formally prescribe the matched filter and the binary filter in this context. We also outline the signal statistics that we utilize, and set up a performance measure which incorporates information about both the correlation peak, and the sidelobe energy for all the pattern classes. In Section III we analyze the performance of the matched filter and the binary filter in a two-class pattern recognition problem where the patterns belong to well-defined statistical classes and are noise free. We obtain analytical results for the performance measure as a function of the system space-bandwidth product in the two cases. In Section IV we investigate the attrition in classification performance in both systems when the input patterns are corrupted by additive noise. Sections V and VI are devoted to numerical solutions and discussions of the comparative classification performance of the matched filter and the proposed binary filter. We demonstrate the monotonic improvement in performance in both systems as the system space-bandwidth product is increased, and show the asymptotic merging of the performance curves for the binary filter and the matched filter.

Notation: Let ω be some fixed (but arbitrary) positive quantity. To each real-valued function, f, of a real variable, we associate its *finite-domain Fourier transform* F_ω formally defined by

$$F_\omega(u) = \int_{-\omega}^{\omega} f(x)\, e^{-i2\pi ux}\, dx. \tag{1}$$

We will use the terminology "space" for the variable x—the domain of the input signals—and "frequency" for the variable u—the domain of the associated Fourier transform.

II. Fourier Correlators

A. Bounded Space-Bandwidth Systems

The conventional Fourier correlator of equation correlator is shift invariant and admits signals of infinite space-bandwidth product (SBP) without loss of information. In this correspondence we will analyze the effect on classification performance of imposing a *finite system space-bandwidth product*. In particular, we consider shift variant Fourier correlators which process inputs through windows $(-\omega, \omega)$ in space, and $(-\nu, \nu)$ in frequency. For a given signal, $f(x)$, and reference, $h(x)$, the output, $g(x)$, of the bounded space-bandwidth correlator is given by

$$g(x) = \int_{-\nu}^{\nu} F_\omega(u)\, \overline{H}_\omega(u)\, e^{i2\pi ux}\, du.$$

We define the *system space-bandwidth product*, which we denote by p, to be the product of the width of the spatial and frequency windows: $p = 4\omega\nu$.

We consider two representative pattern classes, $\mathcal{C}_1$ and $\mathcal{C}_2$. The input signals are real valued functions, $f(x)$, which are sample realizations (drawn from some underlying probability distribution) of one of the two pattern classes. We will denote by $f_j(x)$ the input conditioned upon being drawn from pattern class $\mathcal{C}_j$. For fixed system space-bandwidth product, p, we compare the following two classifiers for different choices of reference signal, $h(x)$.

Matched Filter: The reference signal, $h(x)$, is chosen matched to the sample realization, $f_1(x)$, of class $\mathcal{C}_1$. The correlation output, $g_j^M(x)$, for the matched filter conditioned upon class $\mathcal{C}_j$ at the input is given by

$$g_j^M(x) = \int_{-\nu}^{\nu} F_{\omega,j}(u)\, \overline{F}_{\omega,1}(u)\, e^{i2\pi ux}\, du. \tag{2}$$

If $\omega = \nu = \infty$, we have the classical matched filter. For finite p, a correlation peak is still produced for class $\mathcal{C}_1$. (Classification performance, however, deteriorates as p decreases.) Note that the matched filter above, in general, requires exponential dynamic range.

Binary Filter: The reference signal, $h(x)$, is chosen such that

$$H_\omega(u) = \operatorname{sgn}\left\{\Re\left\{F_{\omega,1}(u)\right\}\right\} = \begin{cases} 1 & \text{if } \Re\left\{F_{\omega,1}(u)\right\} \geq 0 \\ -1 & \text{if } \Re\left\{F_{\omega,1}(u)\right\} < 0. \end{cases}$$

The filter hence takes on values -1 and $+1$ only at each frequency, so that we have a dynamic range of one bit. The correlation output, $g_j^B(x)$, of the binary filter conditioned upon class $\mathcal{C}_j$ at the input is given by

$$g_j^B(x) = \int_{-\nu}^{\nu} F_{\omega,j}(u)\, \operatorname{sgn}\left\{\Re\left\{F_{\omega,1}(u)\right\}\right\} e^{i2\pi ux}\, du. \tag{3}$$

Note that the binary filter tracks the phase of $F_{\omega,1}(u)$, so that we can expect a correlation peak for class $\mathcal{C}_1$, but not for class $\mathcal{C}_2$.

In Fig. 1 we demonstrate two correlations of a random one-dimensional input sequence; in Fig. 1(a), the correlation was accomplished using a matched filter; while in Fig. 1(b), the correlation was performed using a binary filter. As seen, the correlation peaks and sidelobe fluctuation levels are essentially indistinguishable in the two cases.

B. Performance Measure

In characterizing the classification performance of the two filters, we concentrate on two key measures: the strength of the correlation peak and the sidelobe structure. For specific sample realizations, not much can be said about the size of the sidelobes; however, if signal statistics are known, we can extract peak and sidelobe information from a consideration of the ensemble. In the next section we describe a specific statistical structure for the two signal classes from which we can obtain quantitative estimates of filter performance.

For $j = 1, 2$, let $g_j(x)$ denote filter output conditioned upon class $\mathcal{C}_j$ being present at the input. Define

$$\mu_j = \sup_x \left\{\left|E\left\{g_j(x)\right\}\right|\right\}.$$

$$\eta_j = \sup_x \left\{\operatorname{Var}\left\{g_j(x)\right\}\right\}.$$

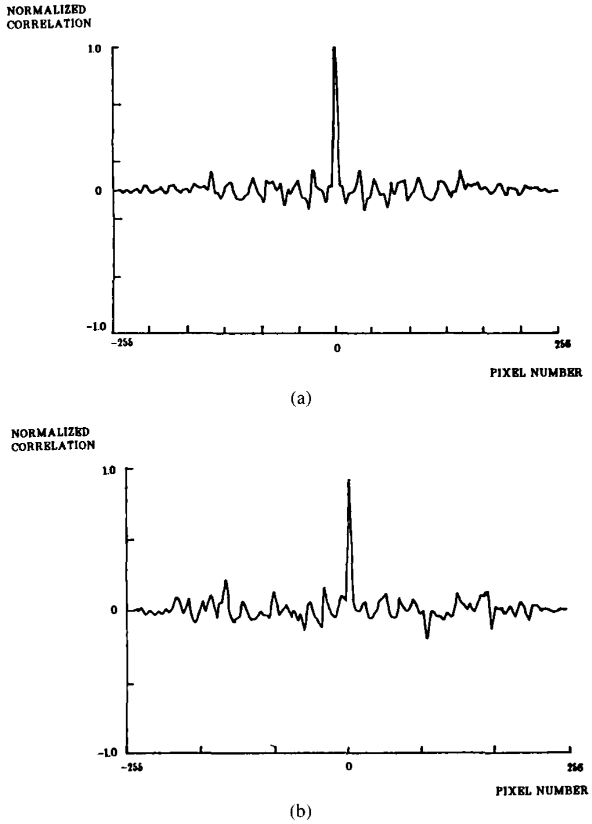

Fig. 1. (a) *Correlation of a random sequence using a matched filter* [4]. (b) Correlation of a random sequence using a binary filter [4].

We define the *performance coefficient* ρ by

$$\rho = \frac{(\mu_1 - \mu_2)^2}{\eta_1 + \eta_2}. \tag{4}$$

The term in the numerator measures the relative size of correlation peaks for the two classes, while the term in the denominator factors in the average energy in the sidelobes. The coefficient, ρ, hence is an indicator of how well the filter discriminates class $\mathcal{C}_1$ from class $\mathcal{C}_2$.

We denote by ρ^M and ρ^B, respectively, the performance coefficients for the matched filter and the binary filter. We shall take system performance to be a monotonically increasing function of the coefficient ρ, with the system with the largest ρ realizing the best performance.

Note that the form of the coefficient ρ is similar to a *signal-to-noise ratio*, the "signal" corresponding to class $\mathcal{C}_1$ and the "noise" to class $\mathcal{C}_2$. (In fact, when the output variable $g(x)$ is Gaussian, and the *a priori* probabilities of the two classes are the same, it turns out that the form of the Bhattacharyya coefficient [5] is identical to (4) for ρ.) From classical communication theory, we have that for correlational systems which are linear functionals of the input signal, the peak signal-to-noise ratio for a signal immersed in white noise is obtained for the matched filter. Hence, we expect the classification performance of the binary filter to be bounded by that of the matched filter.

C. Signal Statistics

In order to facilitate analysis, we assume a specific statistical structure for the ensemble of signals in the two classes. We assume that the signals $f_1(x)$ and $f_2(x)$ corresponding to the two classes $\mathcal{C}_1$ and $\mathcal{C}_2$ are sample realizations of mutually independent, white random processes with

$$E\{f_j(x)\} = 0,$$
$$E\{f_j(x) f_j(y)\} = \sigma_j^2 \delta(x - y). \tag{5}$$

The signal classes have been restricted to be stationary and white in order to effect some simplicity in the ensuing analysis. The stationarity constraint can be relaxed to allow correlation of functions of the form $r_j(x)\,\delta(x - y)$; the analysis for this case is essentially the same as for the case we consider. With the added constraint that the process be Gaussian, one or both constraints can be relaxed to encompass general correlation functions of the form $r_j(x, y)$.

From (1), the real and imaginary parts of $F_{\omega,j}(u)$ are given by

$$\mathcal{R}\{F_{\omega,j}(u)\} = \int_{-\omega}^{\omega} f_j(x) \cos 2\pi ux \, dx,$$
$$\mathcal{I}\{F_{\omega,j}(u)\} = \int_{-\omega}^{\omega} f_j(x) \sin 2\pi ux \, dx. \tag{6}$$

The random processes $f_j(x)$ are independent and zero mean. By virtue of the *central limit theorem* then, it can be readily seen to follow that $\mathcal{R}\{F_{\omega,1}(u)\}$, $\mathcal{I}\{F_{\omega,1}(u)\}$, $\mathcal{R}\{F_{\omega,2}(u)\}$, and $\mathcal{I}\{F_{\omega,2}(u)\}$ are mutually independent Gaussian random processes with zero mean. Some algebraic manipulation readily yields the following:

$$E\{\mathcal{R}\{F_{\omega,j}(u)\}\,\mathcal{I}\{F_{\omega,j}(t)\}\} = 0, \tag{7}$$

$$E\{\mathcal{R}\{F_{\omega,j}(u)\}\,\mathcal{R}\{F_{\omega,j}(t)\}\} = \sigma_j^2 \omega[\text{sinc } 2\omega(u - t) + \text{sinc } 2\omega(u + t)], \tag{8}$$

$$E\{\mathcal{I}\{F_{\omega,j}(u)\}\,\mathcal{I}\{F_{\omega,j}(t)\}\} = \sigma_j^2 \omega[\text{sinc } 2\omega(u - t) - \text{sinc } 2\omega(u + t)]. \tag{9}$$

We also require the first and second moments of the random processes $\text{sgn}\{\mathcal{R}\{F_{\omega,j}(u)\}\}$ and $|\mathcal{R}\{F_{\omega,j}(u)\}|$. Define $r_\omega: \mathbb{R}^2 \to [-1, 1]$ by

$$r_\omega(u, t) = \frac{\text{sinc } 2\omega(u - t) + \text{sinc } 2\omega(u + t)}{(1 + \text{sinc } 4\omega u)^{1/2}(1 + \text{sinc } 4\omega t)^{1/2}}. \tag{10}$$

Note that from (8) if follows that for each u and t, $r_\omega(u, t)$ is just the correlation coefficient of the random variables $\mathcal{R}\{F_{\omega,j}(u)\}$ and $\mathcal{R}\{F_{\omega,j}(t)\}$. The following results can be readily shown (cf. [6], for instance).

$$E[\text{sgn}\{\mathcal{R}\{F_{\omega,j}(u)\}\}] = 0, \tag{11}$$

$$E\{|\mathcal{R}\{F_{\omega,j}(u)\}|\} = \frac{\sqrt{2\omega\sigma_j^2}}{\sqrt{\pi}}\sqrt{1 + \text{sinc } 4\omega u}, \tag{12}$$

$$E[\text{sgn}\{\mathcal{R}\{F_{\omega,j}(u)\}\}\,\text{sgn}\{\mathcal{R}\{F_{\omega,j}(t)\}\}] = \frac{2}{\pi}\sin^{-1} r_\omega(u, t), \tag{13}$$

$$\begin{aligned} E\{|\mathcal{R}\{F_{\omega,j}(u)\}|\,|\mathcal{R}\{F_{\omega,j}(t)\}|\} &= \frac{2\omega\sigma_j^2}{\pi}\Big[(1 + \text{sinc } 4\omega u)^{1/2}(1 + \text{sinc } 4\omega t)^{1/2} \\ &\quad \cdot (1 - r_\omega(u, t)^2)^{1/2} \\ &\quad + (\text{sinc } 2\omega(u - t) + \text{sinc } 2\omega(u + t))\sin^{-1} r_\omega(u, t)\Big]. \end{aligned} \tag{14}$$

III. Two-Class Discrimination

A. The Matched Filter

Our consideration of the matched filter as a correlational system described by (2) differs somewhat from the classical deterministic matched filter [1] in the inclusion of a finite system space-bandwidth and the representation of both input and reference signals as members of a statistical class. The performance coefficient that we derive hence reflects the relative correlation peaks, and the "noisy" sidelobe fluctuations averaged over the ensemble as a function of p (the system SBP).

We estimate the parameters, μ_j and η_j, in (4) in turn for the two classes using the results tabulated in Section II-C.

Class $\mathcal{C}_1$*:* The system output is given by

$$g_1^M(x) = \int_{-\nu}^{\nu} \left(\left[\Re\left\{ F_{\omega,1}(u) \right\} \right]^2 + \left[\Im\left\{ F_{\omega,1}(u) \right\} \right]^2 \right) e^{i2\pi ux}\, du.$$

A simple computation yields

$$\mu_1 = 4\omega\nu\sigma_1^2, \tag{15}$$

$$\eta_1 = 64\omega^2\nu^2\sigma_1^4 \int_0^1 (1 - t)\,(\text{sinc } 4\omega\nu t)^2\, dt. \tag{16}$$

Class $\mathcal{C}_2$*:* The system output is given by

$$g_2^M(x) = \int_{-\nu}^{\nu} F_{\omega,2}(u)\, \overline{F}_{\omega,1}(u)\, e^{i2\pi ux}\, du. \tag{17}$$

The correlation peak and average sidelobe energy can again be simply estimated

$$\mu_2 = 0 \tag{18}$$

$$\eta_2 = 32\omega^2\nu^2\sigma_1^2\sigma_2^2 \int_0^1 (1 - t)\,(\text{sinc } 4\omega\nu t)^2\, dt. \tag{19}$$

Defining α as a function of the space-bandwidth product p by

$$\alpha(p) = \int_0^1 (1 - t)\,(\text{sinc } pt)^2\, dt, \tag{20}$$

the performance coefficient (4) is hence given by

$$\rho^M = \frac{\sigma_1^2/\sigma_2^2}{2\alpha(p)\,[1 + 2\sigma_1^2/\sigma_2^2]}. \tag{21}$$

Asymptotic Results: The above expression can be readily evaluated for extreme values of the system space-bandwidth product. For very low space-bandwidth products, $\alpha(p)$ approaches $1/2$, so that

$$\rho^M \to \frac{\sigma_1^2/\sigma_2^2}{1 + 2\sigma_1^2/\sigma_2^2} \qquad \text{as } p \to 0.$$

For very high space-bandwidth products, on the other hand, $\alpha(p)$ asymptotically approaches the value $1/2p$, so that

$$\rho^M \to \frac{p\sigma_1^2/\sigma_2^2}{1 + 2\sigma_1^2/\sigma_2^2} \qquad \text{as } p \to \infty.$$

The asymptotic results correspond well with intuition. For very low space-bandwidth products, we expect a low processing gain for the system as not much correlation matching can be obtained. For high space-bandwidth products, on the other hand, the use of uncorrelated signals at the input yields large processing gain increasing linearly with the space-bandwidth product.

It is instructive to compare the performance measure given by (21) with the classical matched filter result for the signal-to-noise ratio (SNR) of a deterministic signal immersed in white noise. The *processing gain* of a classical system (defined to be the ratio of the output SNR to the input SNR) is given essentially by the signal space-bandwidth product [1]. If we define σ_1^2/σ_2^2 to be a measure of the input SNR for the statistical case under consideration, then the processing gain of our system, in the limit of large p and small input SNR, is given by $1/2\alpha(p) \approx p$, which is precisely the classical result. (The additional input SNR dependent term present in the denominator of (21) arises because the statistical sidelobe fluctuations are also taken into account in our performance measure; this term will not be significant for low input SNR scenarios.) Finally, the presence of a finite system space-bandwidth product manifests itself in a loss of processing gain; the larger the space-bandwidth product, the more the processing gain realized by the system.

B. The Binary Filter

The system output conditioned upon class $\mathcal{C}_j$ being present at the input is given by (3). We again estimate the parameters, μ_j and σ_j, for the two classes in turn.

Class $\mathcal{C}_1$*:* The output of the system with $f_1(x)$ at the input is given by substitution in (3)

$$g_1^B(x) = \int_{-\nu}^{\nu} \left| \Re\left\{ F_{\omega,1}(u) \right\} \right| e^{i2\pi ux}\, du + i \int_{-\nu}^{\nu} \Im\left\{ F_{\omega,1}(u) \right\} \text{sgn}\left\{ \Re\left\{ F_{\omega,1}(u) \right\} \right\} e^{i2\pi ux}\, du. \tag{22}$$

From (7) and (12), it then follows that

$$\mu_1 = \sup_x \left\{ \left| E\left\{ g_1^B(x) \right\} \right| \right\} = \frac{\sqrt{8\omega\nu^2\sigma_1^2}}{\sqrt{\pi}} \int_0^1 (1 + \text{sinc } pt)^{1/2}\, dt, \tag{23}$$

where, as before, p is the space-bandwidth product $4\omega\nu$.

Now, in (22), set

$$k_1(x) = \int_{-\nu}^{\nu} \left| \Re\left\{ F_{\omega,1}(u) \right\} \right| e^{i2\pi ux}\, du,$$

$$k_2(x) = \int_{-\nu}^{\nu} \Im\left\{ F_{\omega,1}(u) \right\} \text{sgn}\left\{ \Re\left\{ F_{\omega,1}(u) \right\} \right\} e^{i2\pi ux}\, du.$$

Then

$$g_1^B(x) = k_1(x) + ik_2(x).$$

Clearly, $k_1(x)$ and $k_2(x)$ are uncorrelated complex random processes with $k_2(x)$ being zero mean. Hence,

$$\text{Var}\left\{ g_1^B(x) \right\} = E\left\{ \left| k_1(x) \right|^2 \right\} + E\left\{ \left| k_2(x) \right|^2 \right\} - \left| E\left\{ k_1(x) \right\} \right|^2.$$

Using (9) and (12)–(14), we obtain, after some algebraic manipulation, that

$$\text{Var}\left\{ g_1^B(x) \right\} = \frac{4\omega\nu^2\sigma_1^2}{\pi} \int_{-1}^{1} \int_{-1}^{1} \left| \text{sinc}\left\{ \frac{p}{2}(u - t) \right\} \cdot \sin^{-1} r_{p/4}(u, t) - \frac{1}{2}(1 + \text{sinc } pu)^{1/2} \cdot (1 + \text{sinc } pt)^{1/2} \left(1 - \sqrt{1 - r_{p/4}(u, t)^2} \right) \right| \cdot \cos 2\pi(u - t)\,\nu x\, du\, dt. \tag{24}$$

Note that $r_\omega(\nu u, \nu t) = r_{p/4}(u, t)$, which can be verified by direct substitution in defining (10) with $p = 4\omega\nu$.

No analytic expression is available in general for $\eta_1 = \sup_x \{\text{Var}\ g_1^B(x)\}$, and we have to resort to numerical evaluation for specified parameters p, σ_1^2, and σ_2^2. (Note that, in general, the supremum does not occur at $x = 0$.)

Class $\mathcal{C}_2$*:* From (3), the output for class $\mathcal{C}_2$ is given by

$$g_2^B(x) = \int_{-\nu}^{\nu} F_{\omega,2}(u)\, \text{sgn}\left\{ \Re\left\{ F_{\omega,1}(u) \right\} \right\} e^{i2\pi ux}\, du.$$

Again, having recourse to Section II-C, we can show that

$$\mu_2 = \sup_x \left\{ \left| E\left\{ g_2^B(x) \right\} \right| \right\} = 0, \tag{25}$$

$$\text{Var}\left\{ g_2^B(x) \right\} = \frac{4\omega\sigma_2^2}{\pi} \int_{-1}^{1} \int_{-1}^{1} \text{sinc}\left\{ \frac{p}{2}(u-t) \right\} \sin^{-1} r_{p/4}(u, t) \cdot \cos 2\pi(u-t)\,\nu x\, du\, dt. \tag{26}$$

Again, no analytic expression can be found for $\eta_2 = \sup_x \{\text{Var}\, g_2^B(x)\}$, in general, and we must resort to numerical evaluation.

Define β_0, β_1, and β_2 as functions of the space-bandwidth product p by

$$\beta_0(p) = \left[\int_0^1 (1 + \text{sinc}\, pt)^{1/2}\, dt \right]^2, \tag{27}$$

$$\beta_1(p) = \sup_x \int_{-1}^{1} \int_{-1}^{1} \text{sinc}\left\{ \frac{p}{2}(u-t) \right\} \cdot \sin^{-1} r_{p/4}(u, t) \cos 2\pi(u-t)\,\nu x\, du\, dt, \tag{28}$$

$$\beta_2(p) = \sup_x \int_{-1}^{1} \int_{-1}^{1} \Bigg\{ \text{sinc}\left\{ \frac{p}{2}(u-t) \right\} \cdot \sin^{-1} r_{p/4}(u, t) - \frac{1}{2}(1 + \text{sinc}\, pu)^{1/2} \cdot (1 + \text{sinc}\, pt)^{1/2} \left\{ 1 - \left(1 - r_{p/4}(u, t)^2\right)^{1/2} \right\} \Bigg\} \cdot \cos 2\pi(u-t)\,\nu x\, du\, dt. \tag{29}$$

Combining the results of (23)–(26), and using the defining (27)–(29), we obtain the performance coefficient, ρ^B, of the binary filter to be

$$\rho^B = \frac{2\beta_0(p)\,\sigma_1^2/\sigma_2^2}{\beta_1(p) + \beta_2(p)\,\sigma_1^2/\sigma_2^2}. \tag{30}$$

We will return to a comparative analysis of the expressions (21) and (30) in Section V.

IV. Classification in Additive Noise

In practice, the issue of *system robustness* in the face of signal degradations and noise becomes important. We illustrate how noisy signals result in performance attrition in the two correlator systems.

We consider the case where the input signal $f(x)$ is contaminated by an additive noise term $n(x)$. (We assume that the reference signal, $h(x)$, being known *a priori*, can hence be represented in a reasonably accurate and noise-free manner.) We take $n(x)$ to be an independent noise process which is additive and white with

$$E\{n(x)\} = 0,$$
$$E\{n(x)\, n(y)\} = \sigma_n^2 \delta(x - y).$$

The input signal term is then $f_j(x) + n(x)$, and the reference signal term (matched to class $\mathcal{C}_1$) is $f_1(x)$.

A. The Matched Filter

Let $g_{j,n}^M(x)$ denote the (noisy) correlation output of the system when the input signal is a noisy realization of class $\mathcal{C}_j$, viz., $f_j(x) + n(x)$. Then

$$g_{j,n}^M(x) = \int_{\nu}^{-\nu} F_{\omega,j}(u)\, \overline{F}_{\omega,1}(u)\, e^{i2\pi ux}\, du + \int_{-\nu}^{\nu} N_\omega(u)\, \overline{F}_{\omega,1}(u)\, e^{i2\pi ux}\, du = g_j^M(x) + g_n^M,$$

where the first term, $g_j^M(x)$, is the noise-free system response of (2) and the second term, g_n^M, is the additive noise term in the output correlation. The noise term is independent of the signal term, and is zero mean with peak variance at the origin

$$\text{Var}\left\{ g_n^M(0) \right\} = 2p^2\sigma_1^2\sigma_n^2\alpha(p),$$

identical in form to (19). Hence, using (15)–(20), we have

$$\mu_{j,n} = \sup_x \left\{ \left| E\left\{ g_{j,n}^M(x) \right\} \right| \right\} = \mu_j,$$

$$\eta_{1,n} = \sup_x \left\{ \text{Var}\left\{ g_{1,n}^M(x) \right\} \right\} = 2p^2\sigma_1^2\alpha(p)\,(2\sigma_1^2 + \sigma_n^2),$$

$$\eta_{2,n} = \sup_x \left\{ \text{Var}\left\{ g_{2,n}^M(x) \right\} \right\} = 2p^2\sigma_1^2\alpha(p)\,(\sigma_2^2 + \sigma_n^2).$$

The performance coefficient ρ_n^M for the matched filter when input noise is present is hence given by

$$\rho_n^M = \frac{(\mu_{1,n} - \mu_{2,n})^2}{\eta_{1,n} + \eta_{2,n}} = \frac{\left(\dfrac{\sigma_1^2}{\sigma_2^2 + 2\sigma_n^2} \right)}{2\alpha(p)\left(1 + \dfrac{2\sigma_1^2}{\sigma_2^2 + 2\sigma_n^2} \right)} \tag{31}$$

where $\alpha(p)$ is as defined in (20).

A comparison of (21) and (31) shows that the presence of additive input noise is equivalent to an additive increase in the variance (or spread) of class $\mathcal{C}_2$ by exactly twice the spread of the noise.

B. The Binary Filter

Tracing through an analogous analysis yields the performance coefficient ρ_n^B for the binary filter when the input is degraded by additive noise. In general, however, it turns out that the form of ρ_n^B is not conducive to a convenient representation as in (30) for the noise-free case; specifically, in (29), the functional $\beta_2(p)$ has to be replaced by a more complicated supremum taken over the sum of two integrals, the coefficient of one being σ_1^2, and of the other being σ_n^2. (The supremum is now a function of not only the space-bandwidth product p, but also of the signal and noise variances.) Using $\sup\{A + B\} \leq \sup\{A\} + \sup\{B\}$, we can arrive at the following convenient lower bound estimate for ρ_n^B for the sake of comparison:

$$\rho_n^B \geq \frac{2\beta_0(p)\, \dfrac{\sigma_1^2}{\sigma_2^2 + 2\sigma_n^2}}{\beta_1(p) + \beta_2(p)\, \dfrac{\sigma_1^2}{\sigma_2^2 + 2\sigma_n^2}} \tag{32}$$

with the functionals $\beta_0(p)$, $\beta_1(p)$, and $\beta_2(p)$ given by (27)–(29).

On comparing (30) and (32), we see that the effect of additive noise is to create a larger effective spread for class $\mathcal{C}_2$ just as in the case of the matched filter. In both cases, the noise effectively reduces the ability of the system to pick out class $\mathcal{C}_1$ by increasing sidelobe energy, and at the same time increasing the correlation spread of class $\mathcal{C}_2$.

V. Numerical Solutions and Discussion

Let σ^2 denote the ratio $\sigma_1^2/(\sigma_2^2 + 2\sigma_n^2)$. We will refer to σ^2 as the *class spread ratio*; in essence, σ^2 is a statistical measure of the relative strengths of "signal" (class $\mathcal{C}_1$) and "noise" (class $\mathcal{C}_2$, and additive noise) at the input of the correlational system. Recapitulating the expressions for the performance coefficients for easy reference, we have

$$\rho^M = \frac{\sigma^2}{2\alpha(p) + 4\alpha(p)\,\sigma^2},$$

$$\rho^B = \frac{2\beta_0(p)\,\sigma^2}{\beta_1(p) + \beta_2(p)\,\sigma^2},$$

where the functionals $\alpha(p)$, $\beta_0(p)$, $\beta_1(p)$, and $\beta_2(p)$ are defined, respectively, in (20) and (27)–(29).

A numerically generated family of performance curves for the two systems is depicted in Figs. 2 and 3. In each figure, the performance coefficient, ρ, is plotted as a function of the class spread ratio, σ^2, and the family of curves is generated by varying the space-bandwidth parameter p between 8 and 256. In order to facilitate

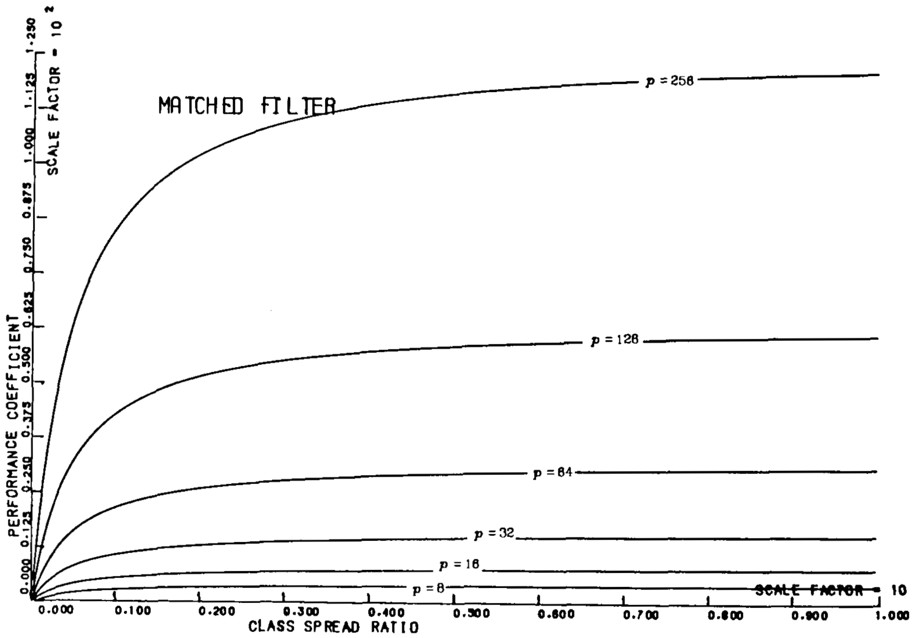

Fig. 2. Plot of the performance coefficient, ρ_m, of the matched filter versus the class spread ratio, σ^2, with the system space-bandwidth product, p, as a parameter.

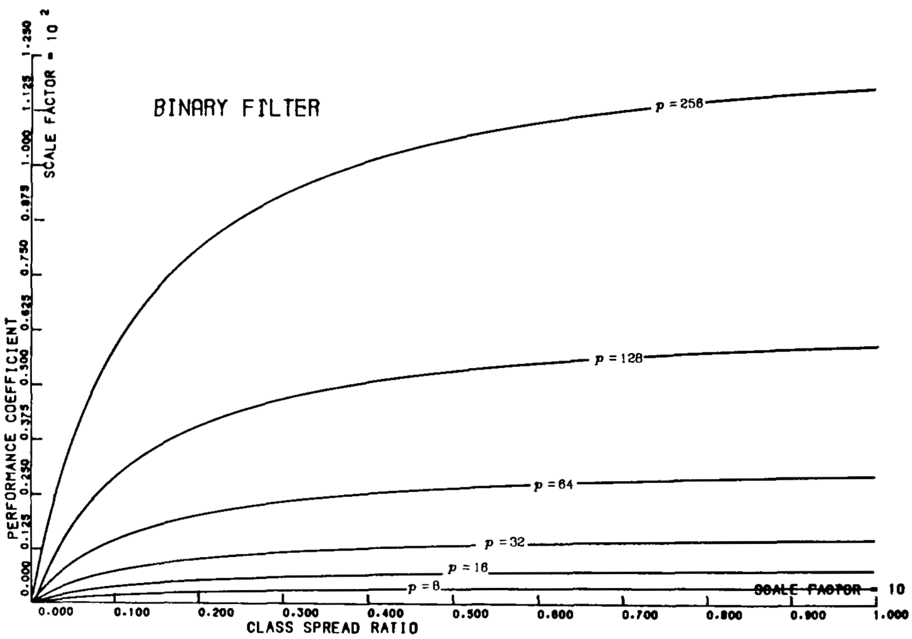

Fig. 3. Plot of the performance coefficient, ρ_b, of the binary filter versus the class spread ratio, σ^2, with the system space-bandwidth product, p, as a parameter.

comparison between the matched filter and the binary filter, for values of $p = 8$, and $p = 256$, the corresponding performance curves of the two systems are extracted from Figs. 2 and 3, and plotted on the same graph in Figs. 4 and 5.

It can be immediately seen from the figures that, all other things being held constant, *the performance coefficient* ρ *is a monotonically increasing function of the system space-bandwidth product* for both filtration systems. This is clearly in accordance with our expectations as increasing the system space-bandwidth product is equivalent to increasing the size of the windows in the space and frequency domains, so that a greater degree of correlation matching can be obtained.

Now, when the class spread ratio, σ^2, is large, we have a situation where the noise power, σ_n^2, and the class $\mathcal{C}_2$ spread, σ_2^2, are both much smaller than the class $\mathcal{C}_1$ spread, σ_1^2. This can be viewed as essentially saying that patterns of class $\mathcal{C}_1$ can take on values

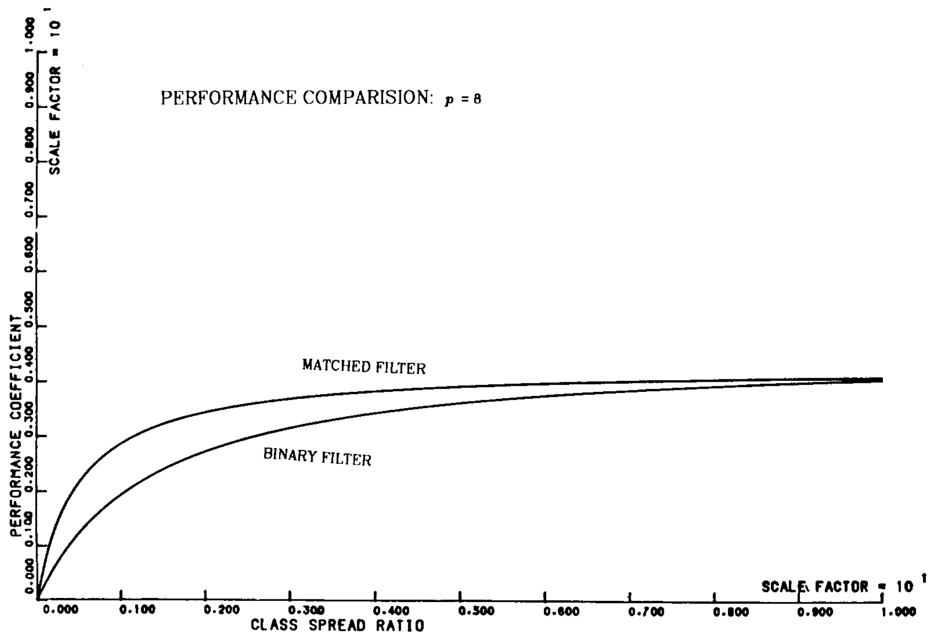

Fig. 4. Plots of the relative performance of the matched filter and the binary filter for a given system space-bandwidth product $p = 8$.

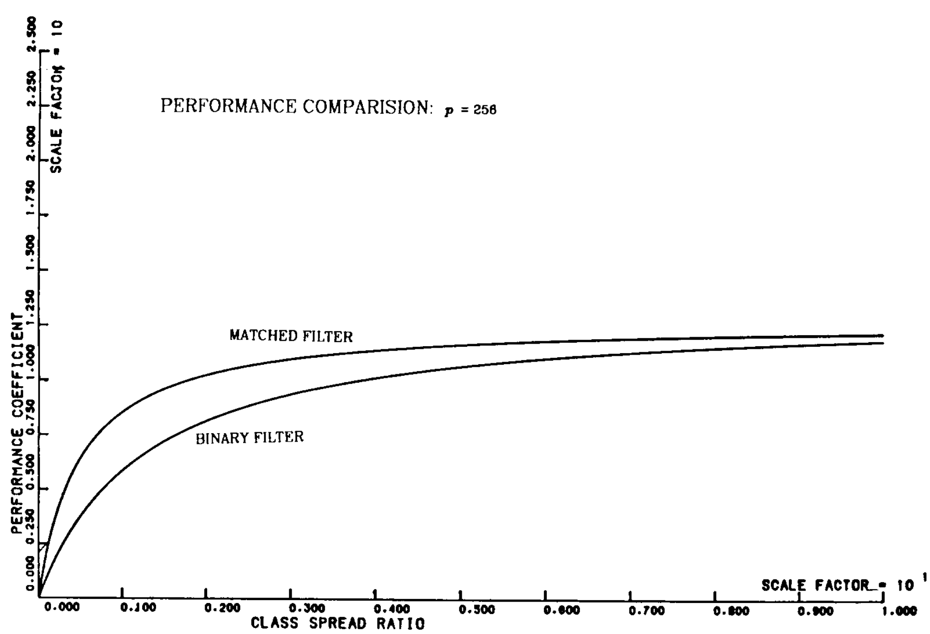

Fig. 5. Plots of the relative performance of the matched filter and the binary filter for a given system space-bandwidth product $p = 256$.

from a much wider set than can patterns of class $\mathcal{C}_2$ and the noise patterns. The probability of significant cross correlation in any particular case is then quite small, so that we expect good classification performance for large values of σ. This intuitive expectation is echoed in Figs. 2–5, where we see that for the matched filter and the binary filter, *the performance coefficient* ρ *is a monotonically increasing function of the class spread ratio*, σ^2, for each performance curve (corresponding to fixed p).

For the matched filter, a close examination of the asymptotes and the slope near the origin of each performance curve reveals that "*large* p" *behavior holds for relatively small values of the system space-bandwidth product* (as small as $p = 8$). The asymptote of the performance curve for the matched filter is approximately $p/2$, and the graph near the origin is a straight line with positive slope p.

Although ρ^B is always bounded from above by ρ^M, *for large class spread ratios the performance curve of the binary filter approaches the same asymptote*, $p/2$, *as the matched filter*, so that their performance is virtually identical. An examination of their relative performance for each p in the range considered indicates that when the class spread ratio is unity (i.e., the two classes have the same variance), we have $\rho^B \approx 2\rho^M/3$.

VI. Conclusion

These numerical simulations, coupled with the prior success of experimental systems utilizing binary filters [4], tend to bolster the intuitive notion that the phase of the Fourier transform contains most of the information content in the signal. The significance of the results lies in the demonstration that, for classification purposes, most of the information content in the signal can be extracted with filters of low complexity. Specifically, the binary filters of this correspondence require only a single bit dynamic range, but provide classification performance comparable to the matched filter which is much more prodigious in its dynamic range requirements. While the success of these schemes is very encouraging, some questions remain. We have demonstrated binary correlator structures based on heuristic algorithms; however, it is not immediately obvious whether we can specify *optimum* binary correlator structures for a given problem. As a specific instance, we can obtain filters which maximally separate pattern classes in that the filter is orthogonal to all unwanted patterns, while yielding a significant correlation only if the desired pattern is present. It is not clear, however, whether an algorithm can be specified which yields the binary filter which is the best approximation to any such maximally separating filter.

References

[1] G. L. Turin, "An introduction to matched filters," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 311-329, 1960.
[2] A. Vander Lugt, "Signal detection by complex spatial filtering," *IEEE Trans. Inform. Theory*, vol. IT-10, pp. 139-145, 1964.
[3] A. V. Oppenheim and R. V. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.
[4] D. Psaltis, E. G. Paek, and S. S. Venkatesh, "Optical image correlation using a binary spatial light modulator," *Opt. Eng.*, vol. 23, pp. 698-704, 1984.
[5] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun. Technol.*, vol. COM-15, pp. 52-60, 1967.
[6] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IRE Trans. Inform. Theory*, vol. IT-4, pp. 69-72, 1958.