K.F. Cheung, L.E. Atlas, J.A. Ritcey, C.A. Green and R.J. Marks II,
"Conventional and composite matched filters with error correction: a comparison",
Applied Optics, vol. 26, pp.4235-4239 (1987).

# Conventional and composite matched filters with error correction: a comparison

Kwan F. Cheung, Les E. Atlas, James A. Ritcey, Charles A. Green, and Robert J. Marks II

A common pattern recognition problem is finding a library object which most closely matches a received image. For additive white Gaussian input noise, optimal detection performance is obtained using a matched filter for each of the $N$ possible library objects. The use of composite matched filters (CMFs) (also called synthetic discriminant functions or linear combination filters) is one technique of reducing the number of filters required for the recognition problem. For two-level composite matched filter outputs, the reduction is from $N$ to $Q = \log_2 (N)$ filters. The CMF's performance, however, can be suboptimum. Using CMFs with bipolar (+1,−1) outputs, this paper examines the detection performance improvement obtained by using error correcting codes. Use of varying levels of error correction is shown to allow trade-off between detection probability and the number of bank filters. Also, we show that in the case of inexact processing, the CMF can perform better than the conventional matched filter.

## I. Introduction

Matched filters are commonly used in pattern recognition for detecting the presence of a known object in a received image. In many cases, the number of library objects $N$ may be quite large. Since the matched filter system requires one filter for each library signals, the number of filters grows linearly with the number of objects to be recognized. For a large system, there is some need for a suboptimal processing system to reduce the number of filters required. One manner of accomplishing this is the composite matched filter (CMF).[1–7] Each CMF is a linear combination of library objects. These filters have been simulated and implemented with some success.[2,6,7]

For additive white Gaussian input noise, the conventional matched filter system is optimal in the sense of maximizing detection probability. The CMF is thus suboptimum in this same scenario. Its performance, however, can be improved using error correction codes. Marks and Atlas[6] have applied a single-error correcting Hamming code to a CMF bank and empirically demonstrated improved performance.

The authors are with University of Washington, Interactive Systems Design Laboratory, Seattle, Washington 98195.

Received 19 March 1987.
0003-6935/87/194235-06$02.00/0.



For an orthogonal set of library vectors corrupted by additive white Gaussian noise, optimal CMF performance can be obtained by using a bipolar (−1,+1) output coding rather than binary (0,1). For bipolar coding the optimal partitioning of the CMF bank output can be performed by simply thresholding each filter output at zero. For the binary case, thresholding each filter output at one half does not result in optimal CMF performance.[1]

In this paper, we extend previous results in three ways. First, multiple-error detection block codes are applied to the CMF. Even though the resulting CMF bank does not perform optimally (the conventional matched filter is globally optimum), further error correction improves the CMF performance. Second, we demonstrate that for the bipolar case, standard block coded CMFs still are optimally decoded by simple thresholding when the orthogonal library vectors are perturbed by zero mean additive white Gaussian noise. Numerical examples are given of the performance of such codes. Last, when inexact processing is used, the conventional matched filter may no longer be optimum. Indeed, we show empirically where CMFs with and without error correction outperform the conventional matched filter.

## II. Preliminaries

The following is a common problem in pattern recognition: We have $N$ library vectors $\{\mathbf{f}_n | 0 \leq n < N\}$ each of length $L$. Given a received vector $\mathbf{g}$, which library vector is closest in the mean square sense? If each library vector has the same energy (norm), the matched filter bank provides good detection performance. One simply finds the index corresponding to the maximum of the inner product $\mathbf{g}^T\mathbf{f}_n$. Note that this is equivalent to minimizing $\|\mathbf{g} - \mathbf{f}_n\|$, where $\|\mathbf{a}\|^2 = \mathbf{a}^T\mathbf{a}$.

If $\mathbf{g}$ is equal to a library vector perturbed by zero mean white Gaussian noise, use of the matched filter maximizes detection probability. If the signal to noise ratio of the received signal is SNR, the detection probability for the *conventional* matched filter is[8–9]

$$P_D(\text{conv}) = 1/\sqrt{2\pi} \int_{-\infty}^{\infty} \exp(-x^2/2)[\text{erfc}(x + \text{SNR})]^{N-1} dx, \quad (1)$$

where

$$\text{erfc}(x) = 1/\sqrt{2\pi} \int_{x}^{\infty} \exp(-y^2/2) dy. \quad (2)$$

The SNR can be written as $\text{SNR} = \sqrt{\epsilon/\sigma^2}$, where $\sigma^2$ is the standard deviation of the input noise, and $\epsilon$ is the signal energy.

In some scenarios, the number of library vectors $\mathbf{N}$ can become prohibitively large. The CMF then becomes an attractive alternate detection algorithm.[1–3,6]

The following example shows the construction of a composite matched filter for a system of $\mathbf{N} = 16$ library vectors. We form the bipolar matrix corresponding to the binary representation of the sixteen indices ($+ = 1$ and $- = -1$):

$$\Delta = \begin{bmatrix} - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + \\ - & - & - & - & + & + & + & + & - & - & - & - & + & + & + & + \\ - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + \\ - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + \end{bmatrix}.$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \end{matrix}$$

We will assume that the library vectors are orthonormal:

$$\mathbf{F}^T\mathbf{F} = \mathbf{I}, \quad (3)$$

where the library matrix is defined by $\mathbf{F} = [\mathbf{f}_0|\mathbf{f}_1|\ldots|\mathbf{f}_{N-1}]$ and the CMF filter bank as

$$\mathbf{S} = \mathbf{F}\Delta^T. \quad (4)$$

Then, for example, $\mathbf{S}^T\mathbf{f}_9 = [+ - - +]^T$, and we have identified the input by its index.

Note that each row of the $\Delta$ matrix is orthogonal to every other row; that is, $\Delta\Delta^T = L\mathbf{I}$. Thus, if white Gaussian noise is added to the input, the output will also be accompanied by white Gaussian noise. The optimal decoding procedure is then to threshold the output of each filter at zero.[1] The probability of having one filter output correct is

$$p = \text{erfc}(-\text{SNR}/\sqrt{N}). \quad (5)$$

Thus the overall probability of detection for the CMF is

$$P_D(\text{CMF}) = p^Q. \quad (6)$$

In the absence of noise, both the conventional and CMF perform without error. Due to its established optimality, the conventional matched filter detection probability in Eq. (1) will always exceed the CMF's performance in Eq. (6) in the presence of additive white zero mean Gaussian input noise.

## III. CMFs with Error Correction

The performance of CMFs can be improved by using additional error correction filters.[6] Let $M$ be the number of additional CMFs so that the total number of filters in our CMF bank is now

$$Q_+ = Q + M. \quad (7)$$

The resulting code will be able to correct up to $E$ errors, where $E$ is the largest integer satisfying the inequality[10,11]

$$2^M \le \sum_{i=1}^{E} \binom{Q_+}{i}. \quad (8)$$

The most efficient of the block error correcting codes are the BCH codes.[10,11] (Hamming codes are BCH codes for $E = 1$.)

For general block codes, a bipolar vector of length $Q$ is coded into a vector of length $Q_+$ according to a bipolar generator matrix G.[10,11] Let $\mathbf{d}_n$ denote the $n$th column of the $\Delta$ matrix. The corresponding coded vector is

$$\mathbf{c}_n = \mathrm{G} \otimes \mathbf{d}_n, \quad (9)$$

where multiplication and addition in the matrix operation are performed modulo two:

$$\begin{array}{c|cc} \oplus & -1 & 1 \\ \hline -1 & -1 & 1 \\ 1 & 1 & -1 \end{array} \qquad \begin{array}{c|cc} \otimes & -1 & 1 \\ \hline -1 & -1 & -1 \\ 1 & -1 & 1 \end{array}. \quad (10)$$

Equivalently, Eq. (9) can be written as

$$\Delta_+ = \mathrm{G} \otimes \Delta, \quad (11)$$

where $\mathbf{c}_n$ is the $n$th column of $\Delta_+$. Our augmented CMF matrix follows as $\mathrm{S}_+ = \mathbf{F}\Delta_+^T$. A received vector is multiplied by $\mathrm{S}_+$. Each element of the resulting vector is thresholded at zero to form a bipolar vector. The bipolar vector is then analyzed for errors.

In the Appendix, we demonstrate that the $\Delta_+$ matrix is composed of mutually orthogonal rows, i.e., $\Delta_+^T\Delta_+ = L\mathbf{I}$. This result insures that if the input to the augmented CMF bank is corrupted with additive white Gaussian noise, the output will also be corrupted by additive white Gaussian noise.[9] The probability of correct detection for a given filter in the bank is the same as the probability of correct detection for a given filter in the unaugmented CMF bank and is given in Eq. (5).

To illustrate, the generator matrix for a $(Q_+,Q) = (15,5)$ BCH code[10,11] capable of correcting $E = 3$ errors is

$$\mathrm{G} = \begin{bmatrix} + & - & - & - & - & + & - & + & - & - & + & + & - & + & + \\ - & + & - & - & - & + & + & + & + & - & + & - & + & + & - \\ - & - & + & - & - & - & + & + & + & + & - & + & - & + & + \\ - & - & - & + & - & + & - & - & + & + & - & + & + & + & - \\ - & - & - & - & + & - & + & - & - & + & + & - & + & + & + \end{bmatrix}.$$

The $\Delta$ and $\Delta_+$ matrices for a $Q = 5$ and the augmented ($Q_+ = 15$) CMF are

$$
\Delta = \begin{bmatrix}
- & - & - & - & - & - & - & - & - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\
- & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + & - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + \\
- & - & - & - & + & + & + & + & - & - & - & - & + & + & + & + & - & - & - & - & + & + & + & + & - & - & - & - & + & + & + & + \\
- & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + \\
- & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & +
\end{bmatrix}
$$

$$
\Delta^{+} = \begin{bmatrix}
- & - & - & - & - & - & - & - & - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\
- & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + & - & - & - & - & - & - & - & - & + & + & + & + & + & + & + & + \\
- & - & - & - & + & + & + & + & - & - & - & - & + & + & + & + & - & - & - & - & + & + & + & + & - & - & - & - & + & + & + & + \\
- & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & + & + \\
- & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + \\
- & - & + & + & - & - & + & + & + & + & - & - & + & + & - & - & + & + & - & - & + & + & - & - & - & - & + & + & - & - & + & + \\
- & + & - & + & + & - & + & - & + & - & + & - & - & + & - & + & - & + & - & + & + & - & + & - & + & - & + & - & - & + & - & + \\
- & - & - & - & + & + & + & + & + & + & + & + & - & - & - & - & + & + & + & + & - & - & - & - & - & - & - & - & + & + & + & + \\
- & - & + & + & + & + & - & - & + & + & - & - & - & - & + & + & - & - & + & + & + & + & - & - & + & + & - & - & - & - & + & + \\
- & + & + & - & + & - & - & + & - & + & + & - & + & - & - & + & - & + & + & - & + & - & - & + & - & + & + & - & + & - & - & + \\
- & + & - & + & - & + & - & + & + & - & + & - & + & - & + & - & + & - & + & - & + & - & + & - & - & + & - & + & - & + & - & + \\
- & - & + & + & + & + & - & - & - & - & + & + & + & + & - & - & + & + & - & - & - & - & + & + & + & + & - & - & - & - & + & + \\
- & + & + & - & - & + & + & - & + & - & - & + & + & - & - & + & - & + & + & - & - & + & + & - & + & - & - & + & + & - & - & + \\
- & + & + & - & + & - & - & + & + & - & - & + & - & + & + & - & + & - & - & + & - & + & + & - & - & + & + & - & + & - & - & + \\
- & + & - & + & + & - & + & - & - & + & - & + & + & - & + & - & + & - & + & - & - & + & - & + & + & - & + & - & - & + & - & +
\end{bmatrix}
$$

## IV. Performance Analysis

We now contrast the performance between the conventional and composite matched filter in three ways. First, we compare the detection probabilities of the detectors for inputs contaminated with zero mean additive white Gaussian noise. Second, we examine the same scenario for high SNRs. Last, Monte Carlo simulations are used to analyze filter performance of the case where processing is inexact. Such inaccuracy, for example, is characteristic of optical processors.

### A. Performance with White Gaussian Input Noise

For the CMF bank augmented for error correction, the probability of correct detection in a filter is still given by Ref. 5. The number of filter banks, however, has been increased from $Q$ to $Q_+$, and we now can tolerate $E$ errors by the constraint in Eq. (8). The probability of detection immediately follows as[10]

$$P_D(\text{CMF}+) = \sum_{i=0}^{E} \binom{Q_+}{i} p^{Q_+-i}(1-p)^i. \tag{12}$$

A plot of this expression is shown in Fig. 1 for $N = 32$ orthonormal library vectors and a composite matched filter of $Q = 5$ output bits. The BCH codes used were capable of correcting $E = 3, 7,$ and 9 errors, corresponding to $Q_+ = 15$, 25, and 29, respectively. Clearly, increasing $E$ improves the overall performance at the cost of additional filters. No matter how large we make $E$, however, performance of the CMF will always be below that of the conventional matched filter.

### B. Performance for High SNR

Direct calculation of the performance measures in Eqs. (1), (6), and (12) for high SNR results in numerical underflow. To get around this problem, we reformulate those equations by using an asymptotic expansions. In all cases, approximation of the probability of erroneous detection, $P_E = 1 - P_D$, is expedient.

A bound corresponding to Eq. (1) can be found in Van Trees[9]:

$$P_E(\text{conv}) \leqq [(N-1)/\sqrt{\pi}\text{SNR}] \exp(-\text{SNR}^2/4), \tag{13}$$

from which it follows that

$$\log P_E \leqq \log[(N-1)/\sqrt{\pi}\text{SNR}] - \text{SNR}^2/4. \tag{14}$$

For high SNR, this bound is sufficiently tight to use as an approximation.

To approximate similarly the CMF performance for high SNR, we rewrite Eq. (6) as

$$P_E(\text{CMF}) = \sum_{i=1}^{Q} \binom{Q}{i} p^{Q-i} q^i,$$

where $q = 1 - p$. Keeping only the first term gives

$$P_E(\text{CMF}) \geq Q p^{Q-1} q.$$

For high SNR, $q \ll 1$, and again the bound can be used as a good approximation. A further simplification results from recognizing that, for moderate $Q$, $P^{Q-1} \approx 1$. Thus $P_E \approx Qq$. From Van Trees[9]

$$\begin{aligned} q &= \text{erfc}(\text{SNR}/\sqrt{N}) \\ &\leq \sqrt{N/2\pi}\, \exp(-\text{SNR}^2/2)/\text{SNR}. \end{aligned} \tag{15}$$

Using this as a good approximation for a high SNR, we obtain our final performance expression for the CMF

$$\log P_E(\text{CMF}) = \log[(\sqrt{N/2\pi})(1/\text{SNR})] - \text{SNR}^2/2N + \log Q. \tag{16}$$

The approximation for the error correcting composite matched filter can be similarly obtained. From Eq. (12), we have

$$P_E(\text{CMF}+) = \sum_{i=E+1}^{Q_+} \binom{Q_+}{i} p^{Q_+-i} q^i.$$

Keeping only the first term gives

$$P_E(\text{CMF}+) \geqq \binom{Q_+}{E+1} p^{Q_+-E-1} q^{E+1}.$$

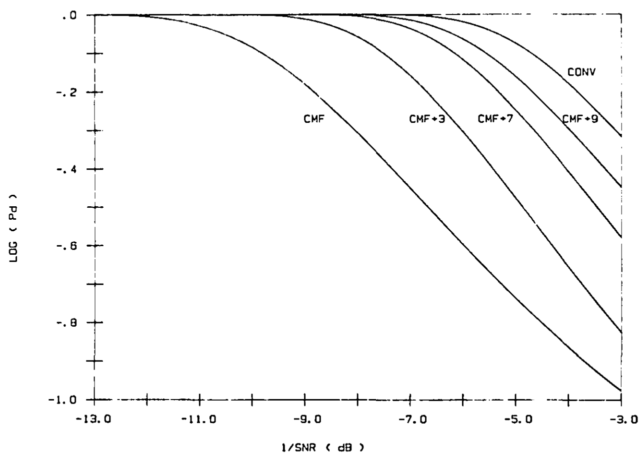


Fig. 1. Performance comparison between the conventional and composite matched filter for $N = 32$ library vectors. Performance of the CMF is shown for $E = 0$, 3, 7, and 9, where $E$ is the maximum number of correctable errors. This corresponds to a total of $Q_+ = 5$, 15, 25, and 29 filters in the CMF bank. The input was corrupted by additive white Gaussian noise with variance $\sigma^2$. The SNR is $\sqrt{\epsilon/\sigma^2}$, where $\epsilon$ is the signal energy.

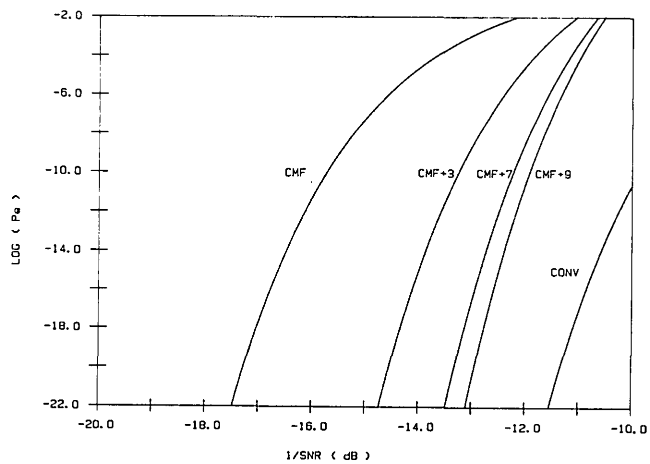


Fig. 2. Performance comparison between the conventional and composite matched filters for high SNRs. The parameters are the same as those in Fig. 1.

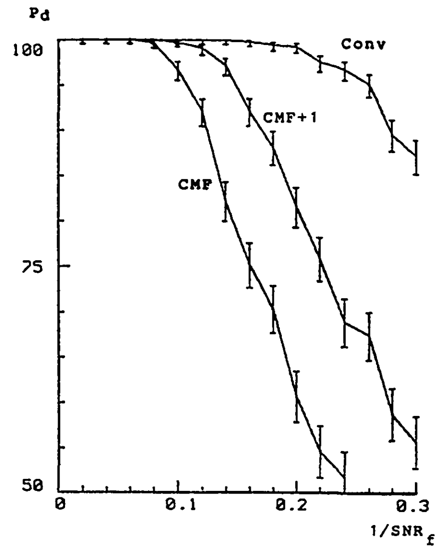


Fig. 3. Monte Carlo performance simulation results for the CMF and conventional matched filter when inexact processing is used. For $N = 16$ library vectors, the CMFs require $Q_+ = 4$ and 7 filters corresponding to $E = 0$ and 1 error correction capability. 90% confidence intervals are shown. To simulate inexact processors, white Gaussian noise with a standard deviation of $\sigma_s = 0.1$ ($SNR_s = 10$) was added to each element of each filter vector.

Making the approximation $p_{Q_+-E-1} \simeq 1$ and applying Eq. (15) give

$$\log P_E(\mathrm{CMF+}) \cong \log\binom{Q_+}{E+1} + \log[(\sqrt{N/2\pi})(1/\mathrm{SNR})] - (E+1)\mathrm{SNR}^2/2N.$$

As before, this bound becomes very tight for high SNR.

Plots of Eqs. (13), (16), and (17) are shown in Fig. 2 for the same parameters used in Fig. 1. ($N = 32$, $Q = 5$, and $Q_+ = 15$, 25, and 29 for $E = 2$, 7, and 9.)

## C. Effects of Processor Inaccuracy

Monte Carlo simulations were performed on a $Q = 4$ ($N = 16$) CMF using single-bit-error correction ($Q_+ = 7$). The inputs were vectors of length $L = 16$:

$$\mathbf{f}_n = \delta_{n+1}; \quad 0 \le n < 16,$$

where $\delta$ has a 1 in its $n$th position and is otherwise zero. White Gaussian noise was added to both the input and each element of the augmented CMF matrix, $S_+ = \mathbf{F}\delta\Delta_+^T$. The signal-to-noise ratios are $SNR_s$ and $SNR_f$, respectively. Since $\epsilon = 1$, we have $SNR_s = 1/\sigma_s$, where $\sigma_s$ is the standard deviation of the input noise. Similarly, $SNR_f = 1/\sigma_f$.

Results are shown in Figs. 2 and 3. Each point estimate was generated by 800 trials; 90% confidence intervals were computed under the assumption that each simulation was a Bernoulli trial.[12]

Figure 2 ($SNR_s = 10$) shows performance much like that observed with exact processing. The conventional matched filter outperforms both CMF banks and single error correction outperforms simulation with no error correction.

As seen in Fig. 4 ($SNR_s = 3.33$), however, the performance becomes significantly different as the SNR of the processor is reduced. The CMF, with and without error correction, outperforms the conventional matched filter for a high SNR.

## V. Conclusions

Matched filters are commonly used in pattern recognition for detecting the presence of a known object in the presence of noise. In this paper, the technique of composite matched filtering has been augmented to include error correction. This has been shown to improve the performance of the composite matched filter bank in the presence of additive white Gaussian noise.

Second, we have shown that the introduction of multiple error correcting linear block codes does not affect the optimal decoding of the bipolar CMF bank

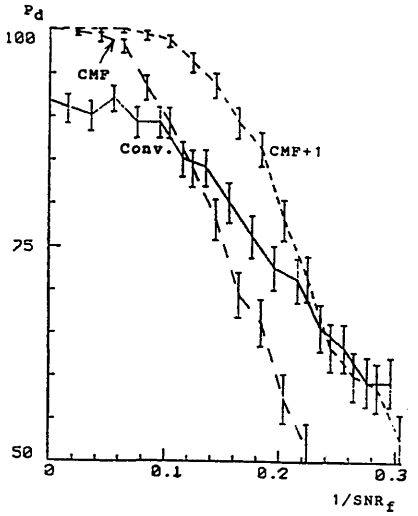


Fig. 4. Monte Carlo performance simulation results. The parameters are the same as were used in Fig. 3, except that $\sigma_s = 0.3$ ($SNR_s = 3.33$). Here the CMF with and without error correction outperforms the conventional matched filter at high input SNRs.

by simple thresholding. Additionally, the composite matched filter has been empirically shown to perform better than the conventional matched filter in conditions of inexact processing.

## Appendix A: Independence of Error Correcting Composite Matched Filter Output

Here we demonstrate that the rows of the bipolar matrix $\Delta_+$ are orthogonal. Specifically, we show that if the bipolar vector $\mathbf{u}$ contains the same number of 1 and $-1$ terms, then, for any other bipolar vector $\mathbf{v}$, the bipolar vector $\mathbf{w} = \mathbf{v} \oplus \mathbf{u}$ is orthogonal to $\mathbf{v}$.

This theorem relates directly to our problem. From Eq. (11), each row of $\Delta_+$ is a modulo 2 sum or difference of the rows of $\Delta$. The rows of $\Delta$ are orthogonal.[1] The only vector in the set of mutually orthogonal bipolar vectors that does not have the same number of 1 and $-1$ terms is the vector that contains all 1 terms (or $-1$ terms). This vector does not appear as a row in $\Delta$. By induction, then, the modulo 2 sum of a number of orthogonal bipolar vectors must result in a vector either orthogonal to each summand or proportional to a summand. (Note that the proportionality constants are limited to 1 and $-1$.) In a coding scheme, we clearly avoid proportional rows in the $\Delta_+$ matrix since inefficient redundant bit information results.

To prove our theorem, we write

$$\mathbf{w}^T\mathbf{v} = \sum_{k=1}^{L} w_k v_k$$

$$= \sum_{k=1}^{L} (v_k \oplus u_k) v_k.$$

Since

$$(v_k \oplus u_k) v_k = -u_k; \quad v_k = \pm 1,$$

and there are the same number of 1 and $-1$ terms in $\mathbf{u}$, we conclude that

$$\mathbf{w}^T\mathbf{v} = -\sum_{k=1}^{L} u_k,$$

$$= 0,$$

and the proof is complete.

## References


1. R. J. Marks II, J. A. Ritcey, L. E. Atlas, and K. F. Cheung, "Composite Matched Filter Output Partitioning," Appl. Opt. *26*, 2274 (1987).
2. Q. B. Braunecker, R. Hauk, and A. W. Lohmann, "Optical Character Recognition Based on Correlation Measurements," Appl. Opt. **18**, 2746 (1979).
3. B. V. K. V. Kumar, "Efficient Approach to Designing Linear Combination Filters," Appl. Opt. **22**, 1445 (1983).
4. D. Casasent, "Unified Synthetic Discriminant Function Computational Formulation," Appl. Opt. **23**, 1620 (1984).
5. H. J. Caulfield and R. S. Putnam, "Fault Tolerance and Self-Healing in Optical Systolic Array Processors," Opt. Eng. **24**, 65 (1985).
6. R. J. Marks II and L. E. Atlas, "Composite Matched Filtering with Error Correction," Opt. Lett. **12**, 135 (1987).
7. J. Riggus and S. Butler, "Simulation of Synthetic Discriminant Function Optical Implementation," Opt. Eng. **23**, 721 (1984).
8. A. Papoulis, *Probability, Random Variables and Stochastic Processes* (McGraw-Hill, New York, 1984).
9. H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I* (Wiley, New York, 1968).
10. R. E. Blahut, *Error-Correcting Codes* (Addison-Wesley, Reading, MA, 1983).
11. W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes* (MIT Press, Cambridge, MA, 1972).
12. W. W. Hines and D. C. Montgomery, *Probability and Statistics in Engineering and Management Science* (Wiley, New York, 1980).



This research was supported by the SDIO/IST's ultrahigh speed computing program administered through the U.S. Office of Naval Research in conjunction with the Optical Systems Laboratory at Texas Tech University and was also partially supported by a National Science Foundation Presidential Young Investigator Award and a Physio-Control career development grant.

The authors thank Ken Ogami for generating the plots for the first two figures.