

Limitations of a class of binary phase-only filters

Donald C. Wunsch II, Robert J. Marks II, Thomas P. Caudell, and C. David Capps

The feasibility of using certain types of binary phase-only filter (BPOF) is investigated. A critical aspect of correlation filters is not often addressed in research on BPOF's: how well do they perform as classifiers in the presence of imperfectly matching templates? It is not enough to detect a single given signal in the presence of noise; it is equally critical to make the correct classification among a number of possible templates with a low false-alarm rate. We show that $(+1, -1)$ -valued BPOF's based on the real part of a conventional matched filter can cause misclassification of simple patterns, even in the absence of noise. These are known to be suboptimal, but the seriousness of their limitations illustrates an important design issue. It is therefore concluded that other types of filters must be used for correlator-based neural network implementations and image processing in general. We also include a commentary on the potential for facing this type of problem with general POF's and BPOF's. The theoretical results are supported by computer simulation and optical experiments.

Key words: Binary phase-only filter, phase-only filter, correlator, neural networks, magneto-optical spatial light modulator.

Introduction

The computation of correlation filters has been a burgeoning topic of research over the past several years, for good reason. Optical correlators, which offer the prospect of massively parallel computation germane to various applications, potentially can be improved in performance by a wise choice of a correlation filter. Sometimes this choice is constrained by hardware; other times it is made to achieve some desirable property. Both of these reasons have led to much research on phase-only filters (POF's) and binary phase-only filters (BPOF's). While these filters have been subjected to much first-rate mathematical analysis, one important quality is frequently left unexamined. Most work concentrates on how well the filter performs in the presence of input patterns that perfectly match the stored template. A more critical question is how the filter performs in the presence of imperfectly matching inputs. The need to do this well is not just some theoretical concern; rather, it is the whole reason for doing correlations in the first place. It is important that the filter not misclassify a close, but imperfectly matching pattern,

in favor of some poorly matching pattern. Unfortunately this is precisely what some popular BPOF's are capable of doing. We discovered this in the context of using a correlator for neural network implementations, so we begin by framing the problem in that context.

Correlator as a Neural Network Hardware Implementation

Optical correlators have attracted renewed interest in recent years with the availability of spatial light modulators (SLM's) to allow real-time signal processing. Usually, applications of the devices have focused on pattern recognition or feature extraction. Another important use, however, is the implementation of neural networks. The challenge of neural networks has caused much activity in the area of hardware implementations, activity that many insist is necessary for the field to come to its full fruition. Many popular neural models use computations that are particularly amenable to hardware implementation, at least in part. The most common of these use inner products among their most critical computations. Examples include adaptive resonance theory,¹ backpropagation,² Adaline and Madaline,³ Hopfield networks,⁴ alternating projection neural networks,⁵ and many others. The point is that, as we review below, these inner products are equivalent to correlations that can be performed by optical correlators.⁶ This convenient property is augmented by the fact that one can utilize the shift invariance of the Fourier-

R. J. Marks II is with the Department of Electrical Engineering, FT-10, University of Washington, Seattle, Washington 98195. The other authors are with the Boeing Company, Seattle, Washington 98124.

Received 18 July 1991.

0003-6935/92/265681-07\$05.00/0.

© 1992 Optical Society of America.

transform magnitude to allow the simultaneous correlation of a single input with multiple templates. We show a figure that demonstrates how the VanderLugt correlator⁷ can be used for neural network implementations (see Fig. 1). The spatial form of the stored templates are placed on the first plane's SLM, which is denoted as SLM1. This permits the parallel operation needed for neural net implementation.

The correlator takes its input in the Fourier plane in order to compare it with all the templates simultaneously. This requires computing a filter based on the input pattern. Consider using the conventional matched filter in this plane. To do this, take the input to the neural network $i(x, y)$ and compute its Fourier transform $I(u, v)$. Take the complex conjugate of this $I^*(u, v)$ to get the conventional matched filter. Use of this filter in the Fourier plane results in correlations with the pattern or patterns in the plane (referred to as Memory and Learning in Fig 1). The input pattern $i(x, y)$, which is abbreviated as i in Fig. 1, is placed in the memory plane (for some applications) to obtain the autocorrelation $i \star i$. (\star is the correlation operator.) Similarly the stored templates $t_i(x, y)$, which are abbreviated as t_i , are put in the memory plane to obtain the cross correlations $i \star t_i$. The central correlation values shown in the output plane take the form

$$|O|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t_i(x, y) i(x, y) dx dy, \quad (1)$$

which is an inner product. The equivalent operation in a neural network is the sum

$$\sum_{j=1}^n t_{ij} i_j, \quad (2)$$

where \mathbf{i} is an input vector and \mathbf{t}_i is the i th weight vector. Thus a single input pattern is fed to a

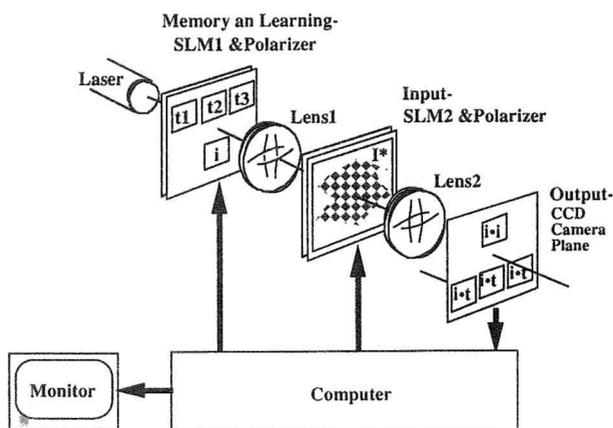


Fig. 1. Use of a VanderLugt correlator to compute inner products for neural network implementation. The correlation peaks shown on the output plane are equivalent to inner products of input pattern $i(x, y)$ with templates $t_i(x, y)$. The filter I^* must be computed in order to do this.

network with several nodes, and the output of these nodes is computed in parallel.

The neural network implementation described above is just one application of a correlator. Another operation would be an image processing application, in which a single pattern is stored as a template and searched for in a scene. (This reverses the roles of input pattern and template from what was discussed above.) For this problem, consider the Fourier-transform pairs $g(x, y)$ and $G(u, v)$. Compute the conventional matched filter $G^*(u, v)$. Then consider an input scene containing a pattern $f(x, y)$ and assess how well it matches the filter. A bright output peak at the location of $f(x, y)$ indicates a good match, and an instance of the template $g(x, y)$ has been discovered in the scene. We use this notation throughout the remainder of the paper. Below we review the POF and BPOF. We continue with a discussion of the magneto-optical SLM to motivate a discussion of the BPOF. This discussion is framed in the context of experiments with the filter. We then move to an analysis of both the POF and the BPOF. We show that the same filter can be generated by a number of significantly different templates, and discuss how this has implications for the results we have observed.

POF and BPOF

In our discussion we refer to an input pattern $f(x, y)$, its Fourier transform $\mathcal{F}[f(x, y)] = F(u, v)$, and the complex conjugate of $F(u, v) = F^*(u, v)$. We also consider a template $g(x, y)$ and the corresponding operations on it. Consider $G(u, v)$ with the amplitude and phase information given separately:

$$G(u, v) = |G(u, v)| \exp[j\phi[G(u, v)]], \quad (3)$$

where $j = \sqrt{-1}$ and ϕ indicates phase. Taking the complex conjugate gives us the conventional matched filter:

$$G^*(u, v) = |G(u, v)| \exp[-j\phi[G(u, v)]]. \quad (4)$$

The POF is given by

$$G_\phi(u, v) = \frac{G^*(u, v)}{|G(u, v)|} = \exp[-j\phi[G(u, v)]] \\ = \exp[j\phi[G^*(u, v)]], \quad (5)$$

which can be converted to a BPOF $G_{B\phi}(u, v)$ by a number of techniques. For example, we consider the popular formula⁸:

$$G_{B\phi}(u, v) = \begin{cases} 1 & \text{if } \text{Re}[G(u, v)] \geq 0 \\ -1 & \text{if } \text{Re}[G(u, v)] < 0 \end{cases} \quad (6)$$

For measuring inner products, we are really interested only in the correlations at the origin. To analyze this, consider the amplitude output O :

$$O(x, y) = f(x, y) \star g(x, y), \quad (7)$$

where \star denotes correlation. We evaluate at the origin:

$$O(0, 0) = \iint_P F(u, v)G^*(u, v)dudv \quad (8)$$

$$O(0, 0) = \iint_P |F(u, v)|\exp[j\phi[F(u, v)]] \times |G^*(u, v)|\exp[j\phi[G^*(u, v)]]dudv. \quad (9)$$

[Henceforth we suppress the $O(0, 0)$ and simply call this O . We take the pupil function to be the maximum width available on the SLM. This is not the only way to choose the region of integration, an issue that is treated in detail by Kumar and Bahri.^{9]} Furthermore, measure the intensity only:

$$|O|^2 = \left| \iint_P |F|\exp[j\phi(F)]|G^*|\exp[j\phi(G^*)]dudv \right|^2. \quad (10)$$

We now consider the signal-to-noise ratio (SNR) achieved by the filter G :

$$\text{SNR} = \frac{\left| \iint_P |F|\exp[j\phi(F)]|G^*|\exp[j\phi(G^*)]dudv \right|^2}{\iint_P P(u, v)|G(u, v)|^2dudv}, \quad (11)$$

where $P(u, v)$ is the power spectral density of the input noise. This is easily shown to be a maximum when $F/P = G$. If we are constrained to the $(+1, -1)$ filter described above, maximizing the SNR is equivalent to maximizing the output at the origin, since $|G| \equiv 1$ implies that the denominator of Eq. (11) is independent of $G(u, v)$ (see Ref. 10). If we set $P \equiv 1$ (white noise), we get the conventional matched filter of Eq. (4), i.e., $G^* = F^*$ gives us the best SNR.

Now consider the output when we use the POF $G_\phi(u, v)$ instead of G^* . Measuring the output, we find that

$$|O|^2 = \left| \iint_P |F|\exp[j\phi(F)]\exp[j\phi(G^*)]dudv \right|^2. \quad (12)$$

In the analogous case for the BPOF $G_{B\phi}(u, v)$, we have

$$|O|^2 = \left| \iint_P |F|\exp[j\phi(F)]G_{B\phi}dudv \right|^2. \quad (13)$$

We return to these expressions in our analysis section.

Magneto-Optical Spatial Light Modulator

The following discussion is intended as background to our subsequent experimental section, and closely

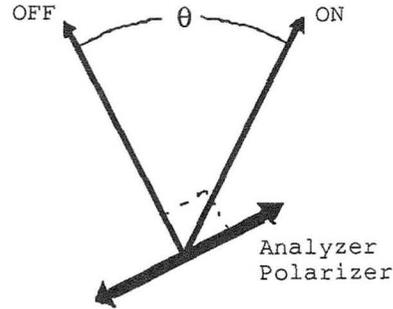


Fig. 2. Orientation of a magneto-optical SLM's analyzer polarizer to achieve binary ON-OFF $(+1, 0)$ modulation.¹¹

follows Davis and Waas.¹¹ The SLM used for this experiment is a 48×48 magneto-optic SLM commercially available from Semetex Corp. (Sightmod, SMD48I). This device rotates the polarization axis for linearly polarized light as the light passes through any electrically activated SLM element. Activated elements are defined as the ON state. Light passing through unactivated (OFF) elements of the SLM does not have its polarization axis rotated. This is depicted in Fig. 2. A polarizer (called the analyzer polarizer) is placed in the beam path before the SLM (as shown in Fig. 3), in this case with an orientation that is perpendicular to that of the OFF state. Thus light will only pass through the ON elements.

The SLM can also be configured for bipolar phase modulation. This is obtained by orienting the analyzer polarizer perpendicular to the bisector of the NEG and POS polarization states, as shown in Fig. 4. The result is a π phase shift of the E -field vector of the transmitted light between the ON and OFF states. This leads to the definition of an angle ϕ_c , which is the bisector of the angle θ in Fig. 4. (We measured our SLM's value of θ to be 7° .) The BPOF is then given by

$$G_{B\phi}(u, v) = \begin{cases} 1 & \text{if } \phi(u, v) > \phi_c \\ -1 & \text{if } \phi(u, v) \leq \phi_c \end{cases}, \quad (14)$$

which reduces to Eq. (6) if $\phi_c = 0$. Of course, these are not the only possible orientations of the analyzer polarizer. Various complex filters can be realized by choosing other analyzer polarizer orientations.¹² If we are not constrained by the rotation angle θ , we get the general unconstrained BPOF. Farn and Goodman¹⁰ have shown how to compute that filter. Our results are consistent with theirs in that they do not limit themselves to the device constraints explained here.

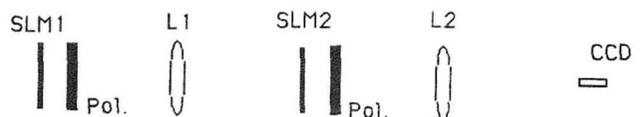


Fig. 3. Diagram of a VanderLugt correlator showing the polarizers (Pol.).

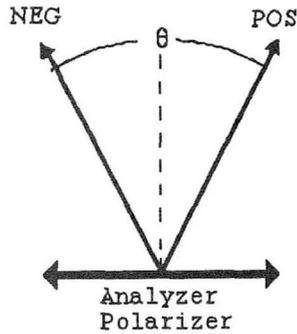


Fig. 4. Orientation of the SLM's analyzer polarizer to achieve bipolar (+1, -1) modulation.¹¹

Experimental and Simulated Results

We ran several experiments and simulations to implement an unsupervised learning neural network architecture.⁶ This generated many different templates in response to alphabet input patterns. Templates took the form of conjunctions of various sets of alphabet patterns over time. This yielded many different examples of inner products between various patterns that had been measured by the system. This functionality is what we desire for any of the neural net models mentioned earlier. We report here some key findings that led us to suspect BPOF limitations of presenting a problem. In the plots below, theoretical inner product is the correct pixel overlap between two images. For example, the 5×5 patterns **G** and **E** have an overlap of 16 pixels (out of 25 possible), so 16 is their theoretical inner product. In Fig. 5 the BPOF computed inner product is the simulated correlator's computation of inner products by using the BPOF without any attempt to normalize the output numbers. Ideally, the result would be some kind of monotonic increasing function. As expected, the actual result has some relationship to the correct inner product values, but does not provide an accurate computation of these.

In Fig. 6 the optically computed inner product is the inner product measured by a VanderLugt correlator in our laboratory. Figures 5 and 6 qualitatively agree in that the correlator fails to correctly order the inner products. The principal difference between the data in Figs. 5 and 6 is the number of plot points. The following figures, however, exhibit significant

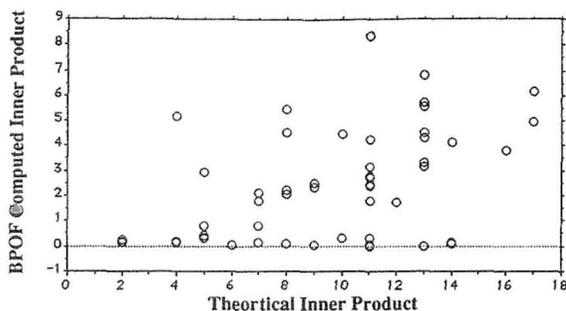


Fig. 5. Simulation of correlator as an inner product processor without normalization.

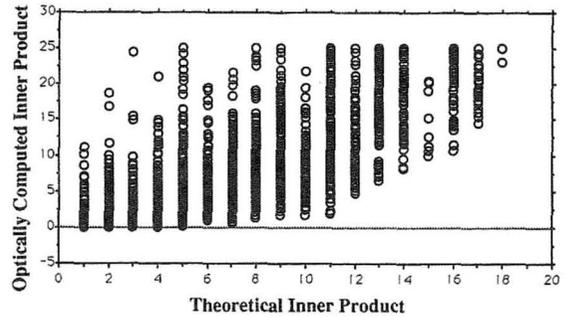


Fig. 6. Experimental correlator data for inner products measured without normalization.

qualitative differences because of different normalizing schemes. This is particularly significant in the case in which the input pattern is a strict subset of more than one template. Each template will correlate well with the input because the latter is a strict subset, and the strength of the output peak will be determined by the intensity allowed through by the template; i.e., the larger template will always win. This can be compensated for by inversely scaling the measured output by the (already known) size of the template pattern.

We show this in Fig. 7. The reason for choosing an inversely scaled normalizing factor is that the smaller templates let through less energy, so they obtain smaller inner products than larger templates even if they have a similar number of pixels overlapping with the input pattern. While mathematically sound, this reasoning fails in practice because of noise sensitivity. When the template size is small, dividing by it results in enhancing the noise. The figure reflects this in the large number of small templates that give saturated inner product readings.

A more viable approach is to normalize by using a known factor. In our case, one of the computations made by the correlator is cheaper than all the others to compute electronically. In the particular neural network model we are implementing, we need the autocorrelation of the input. This term required only N ADD operations, instead of N MULTIPLY and N ADD operations for the cross correlations. We computed this electronically in order to find a normalizing factor to use for all the other terms. As Fig. 8 shows, this works better, with none of the small inner products giving saturated readings. Other ap-

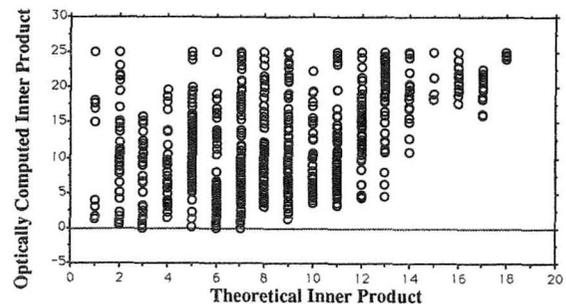


Fig. 7. Correlator data for inner products measured with normalization by template size.

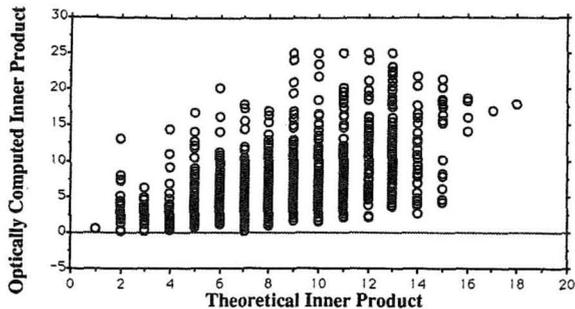


Fig. 8. Correlator data for inner products measured by using a known normalization factor.

proaches to normalization have been dealt with by Dickey and Romero.¹³

We are still seeing less accuracy than we would like for neural network implementations. As we will discuss, this is not the fault of the normalization scheme. If we had used the conventional matched filter instead for this computation, the correct theoretical values would be found in the absence of noise. This particular (1, -1) BPOF cannot do this reliably. Furthermore it cannot even correctly obtain the relative order reliably within a set of measurements. This is really more critical than obtaining the correct inner product measurements. The correlator may make significant mistakes regarding the value of inner products so long as it still classifies the best match first, the second best next, and so on. In fact, some distortion may be beneficial, such as excessive values for the best match at the expense of the others. Unfortunately for this BPOF, it does not even meet this relaxed criterion.

Figure 9 shows the ordering distortion caused by this (1, -1) BPOF in a noise-free simulation. This information was acquired by running a simulation of the device trying to classify various templates. The

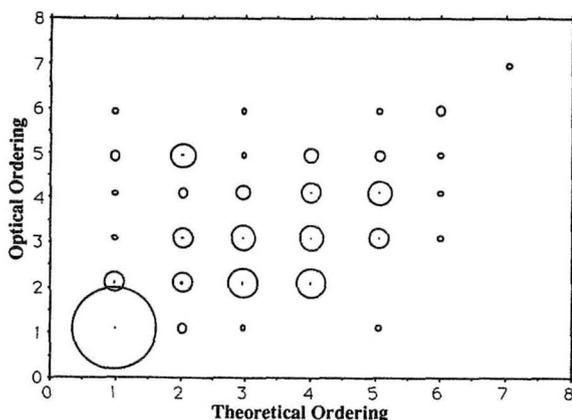


Fig. 9. Ordering distortion caused by the BPOF. Larger circles indicate a greater overlap of data at the same point. While the BPOF is more likely to be correct than incorrect, it makes a significant number of serious misclassifications, even in the absence of noise. It is especially susceptible to misclassifying the highest priority categories. The ideal result would be a set of large circles forming a line of unit slope. The lower left data is of much greater importance than the upper right in this regard.

known values of the template matching scores were compared with the optical systems' scores. Actual measured inner product values were ignored; the only performance criterion was correct ordering from largest to smallest. In Fig. 9, 1 means largest, 2 is second largest, and so on. Ties were handled by repeating a number twice and skipping the following number, e.g., 1, 2, 2, 4. The larger circles correspond to overlapping points; the more data at a single point, the bigger the circle.

The x axis of this figure is the correct ordering for each data point. The y axis shows the optically computed ordering. The data points are computed for various experiments with anywhere from two to seven classes. An ideal result would be a single line of large circles extending diagonally across the plot. Note that the classifications are independent of normalization. We are looking at only the order of the winners, not their relative magnitudes, and we are not making comparisons between classifications with varying sets of patterns. No normalizing scheme can cure the problem illustrated here.

The figure shows that this BPOF, while usually correct or at least close, occasionally makes disastrous misclassifications, regarding the best match as the worst, and vice versa, in the most extreme cases. This may be acceptable for some applications in which mistakes are tolerable (and this may be the case for some neural network implementations) but it certainly would cause performance degradation in many applications. It must be emphasized that these results are in the absence of noise. Performance would degrade further with noisy inputs, with noise in the filter values, or with noisy detectors.

Analysis

To further assess these results, we derive some useful properties of POFs and BPOFs. We begin with the description of an output peak achieved by using the POF. Refer to Eq. (12). At first glance, it is not apparent that G_ϕ is a good approximation to G^* . To investigate this, consider the approximation of G^* by an arbitrary phase function $\exp[j\theta(u, v)]$. The L_2 norm, or mean square error, in such an approximation is

$$\epsilon^\phi = \iint_P \left| |G^*(u, v)| \exp[j\phi(G^*)] - \exp[j\theta(u, v)] \right|^2 du dv. \quad (15)$$

Expanding, we find that

$$\begin{aligned} \epsilon^\phi &= \iint_P \{ |G^*(u, v)| \exp[j\phi(G^*)] - \exp[j\theta(u, v)] \\ &\quad \times [|G^*(u, v)| \exp[-j\phi(G^*)] - \exp[-j\theta(u, v)]] \} du dv, \end{aligned} \quad (16)$$

$$\begin{aligned} \epsilon^\phi &= \iint_P [|G^*(u, v)|^2 \\ &\quad + 1 - 2|G^*| \cos[\phi(G^*) - \theta]] du dv. \end{aligned} \quad (17)$$

Clearly $\theta(u, v) = \phi(u, v)$ minimizes ϵ^ϕ over all possible choices of θ . This proves the following theorem:

Theorem 1. The filter $G_\phi = \exp[j\theta(u, v)]$ is the closest approximation (in L_2 norm) to G^* over the entire class of possible POF's.

This theorem leads us to suspect that G_ϕ will be the optimal filter over all POF's, in the sense that it maximizes the output peak and, thus, the SNR. This was indeed proven by Kumar et al. in 1988. (See Kumar *et al.* for various relevant contributions.^{9,14-19})

We now consider the BPOF case by comparing the filter to the conventional matched filter. Below we consider an arbitrary BPOF $B(u, v)$. We show that choosing this to be anything other than the $(+1, -1)$ filter $G_{B\phi}$ from Eq. (6) will result in a greater error. The error in approximating the conventional matched filter by the BPOF is:

$$\epsilon^{B\phi} = \iint \left| G^*(u, v) |\exp[j\phi(G^*)]| - B(u, v) \right|^2 dudv, \quad (18)$$

$$\begin{aligned} \epsilon^{B\phi} = & \iint |G(u, v)|^2 dudv + \iint dudv \\ & - 2 \iint |G(u, v)| \cos[\phi[G(u, v)]] B(u, v) dudv, \end{aligned} \quad (19)$$

where $\phi[G(u, v)]$ is the phase of $G(u, v)$. The first two integrals are unaffected by $B(u, v)$. Then Eq. (19) is minimized by choosing $B(u, v)$ to have the same sign as $\cos[\phi[G(u, v)]]$, i.e., as the real part of $G(u, v)$. Thus the use of any $(+1, -1)$ filter other than $G_{B\phi}$ yields a greater error.

Theorem 2. This theorem is proved by the above. The filter $G_{B\phi}$ is the closest (in L_2 norm) approximation to G^* over the entire class of possible BPOF's with values of 1 or -1.

Nonuniqueness and Its Implications

Because we are considering problems related to misclassifications, it is important to note that many highly dissimilar functions have the same POF or BPOF. This is true because, if we consider

$$\begin{aligned} \text{POF}[g_1(x, y)] = G_{\phi_1}(u, v) &= \frac{G_1(u, v)}{|G_1(u, v)|} \\ &= \exp[j\phi[G_1(u, v)]], \end{aligned} \quad (20)$$

where g is an arbitrary pattern, we can construct the two-lens processor shown in Fig. 10, which optically

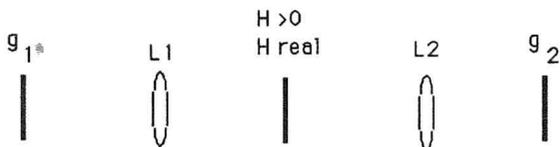


Fig. 10. Optical construction of a different function with the same POF.

constructs the function

$$g_2(x, y) = \mathcal{F}[G_1(u, v)H(u, v)], \quad (21)$$

where $\mathcal{F}[\cdot]$ is the Fourier-transform operator.

The transform of $g_2(x, y)$ clearly has the same phase as $G_1(u, v)$ because, as shown in Fig. 10, we have chosen H to be purely real and positive. H can therefore take a myriad of different forms. This suggests that the filters cannot see the difference between arbitrarily different patterns, yet can miss other patterns that are essentially similar. It would be possible, for instance, to construct a set of patterns such that Fig. 5 or Fig. 9 would have a straight horizontal line of points, which is far from the ideal of some kind of monotonic increasing function. The constraints that H be real and positive are weak, so a large class of functions can be constructed with the same POF or BPOF. Thus we anticipate a possibility of constructing sets of templates that cannot be discriminated, given a particular POF or BPOF algorithm.

From the observation above it follows that the set of functions with the same POF forms a convex set since, for an arbitrary constant α with $0 < \alpha < 1$,

$$\begin{aligned} [g_1(x, y) \rightarrow G_\phi, g_2(x, y) \rightarrow G_\phi] \\ \Rightarrow \alpha g_1(x, y) + (1 - \alpha)g_2(x, y) \rightarrow G_\phi, \end{aligned} \quad (22)$$

where \rightarrow is the POF operator and \Rightarrow is the logical implication operator. The same argument applies to the BPOF. The convex nature of the POF or BPOF set may prove useful in determining the classification performance of the filter because it is a straightforward matter to determine the distance between two or more nonintersecting convex sets.⁵

The nonuniqueness observation of this section, however, is more important than the result on convexity of the POF set. This is because

- (1) The convexity observation depends on the nonuniqueness observation;
- (2) The nonuniqueness has implications on limitations of POF's and BPOF's, including BPOF's computed in a number of different ways.

We anticipate that the nonuniqueness observation could lead to a choice of useful examples of the misclassification potential of various filters, and that the convexity observation could lead to characterization of the discrimination properties of these filters.

Discussion

The discussion of nonuniqueness above broadens the scope of filters that may be affected by the problems we consider. Thus we have come full circle and suggest that misclassification problems occur in the BPOF itself.

Our results apply to a popular kind of BPOF, but not to all of them. It is worth noting the cases left unexamined. We have not shown misclassification errors for the unconstrained complex BPOF's dealt

with by Farn and Goodman,¹⁰ although the non-uniqueness analysis applies to these. It is worth noting that if one is willing to add hardware complexity, one can achieve full unconstrained complex modulation with two one-parameter SLM's.²⁰ Juday and Florence have also shown how to obtain full complex modulation out of a single deformable mirror device SLM, at the cost of halving its resolution.²¹ We predict that the most fruitful work will choose the unconstrained complex values rather than the $(+1, -1)$ or $(+1, 0)$ filters that have been popular in the past. One final caveat: various measures of the quality of a BPOF have been devised over the last few years,²² some of which were succinctly summarized recently by Horner.²³ For another recent germane discussion, see Downie.²⁴ It is entirely possible that some filter might provide better classification performance and perform well within one respect but not in another.

Conclusion

A class of BPOF's examined for neural network implementation is found to have some fundamental limitations. We show that this kind of BPOF misclassifies patterns even in the absence of noise. This is a serious limitation on the efficacy of this kinds of BPOF for image processing and neural network implementation problems. These results are especially practical for work involving the magneto-optical SLM, which has constraints similar to those analyzed here. We point out that a variety of functions look equivalent to POF's and BPOF's as a possible explanation for some of the problems encountered.

We thank the reviewers for many insightful comments that have significantly improved the final version of this paper. Among other things, one reviewer suggested a simplification to our proof of theorem 2, while another pointed out how our observation regarding the nonuniqueness of functions yielding a given BPOF relates to the potential for classification problems. We are grateful for these and other helpful comments.

References and Notes

- G. Carpenter and S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine," *Comput. Vision Graphics Image Process.* **37**, 54-115 (1987).
- D. E. Rumelhart, G. Hinton, and R. J. Williams, *Parallel Distributed Processing* (MIT, Cambridge, Mass., 1986), pp. 318-364.
- B. K. Widrow and M. E. Hoff, "Adaptive switching circuits," presented at the 1959 WESCON Convention.
- J. Hopfield and D. Tank, "'Neural' computation of decisions in optimization problems," *Biol. Cybernetics* **52**, 141-152 (1985).
- R. J. Marks, II, S. Oh, and L. E. Atlas, "Alternating projection neural networks," *IEEE Trans. Circuits Syst.* **36**, 846-857 (1989).
- D. C. Wunsch, II, T. P. Caudell, C. D. Capps, R. J. Marks, II, and R. A. Falk, "An optoelectronic implementation of the adaptive resonance neural network," *IEEE Trans. Neural Networks* (to be published).
- A. B. Vanderlugt, "Signal detection by complex spatial filtering," *IEEE Trans. Inf. Theory* **IT-10**, 139 (1964).
- D. Psaltis, E. G. Paek, and S. S. Venkatesh, "Optical image correlation with a binary spatial light modulator," *Opt. Eng.* **23**, 698-704 (1984).
- B. V. K. V. Kumar and Z. Bahri, "Efficient algorithm for designing a ternary valued filter yielding maximum signal to noise ratio," *App. Opt.* **28**, 1919-1925 (1989).
- M. W. Farn and J. Goodman, "Optimal binary phase-only matched filters," *Appl. Opt.* **27**, 4431-4437 (1988).
- J. A. Davis and J. M. Waas, "Current status of the magneto-optic spatial light modulator, in *Spatial Light Modulators and Applications III*, U. Efron, ed., *Proc. Soc. Photo-Opt. Instrum. Eng.* **1150**, 27-43 (1989).
- J. D. Downie, "Design of optimal binary phase and amplitude filters for maximization of correlation peak sharpness," *Opt. Lett.* **16**, 508-510 (1991).
- F. M. Dickey and L. A. Romero, "Normalized correlation for pattern recognition," *Opt. Lett.* **16**, 1186-1188 (1991).
- B. V. K. V. Kumar and Z. Bahri, "Optimality of phase-only filters," in *Computer-Generated Holography II*, S. H. Lee, ed., *Proc. Soc. Photo-Opt. Instrum. Eng.* **884**, 146-152 (1988).
- B. V. K. V. Kumar and J. M. Connelly, "Effects of quantizing the phase in correlation filters," in *Optical Information Processing Systems and Architectures*, B. Javidi, ed., *Proc. Soc. Photo-Opt. Instrum. Eng.* **1151**, 166-173 (1989). This reference is especially relevant to the present work because it shows how to calculate the loss in SNR caused by using a BPOF or, for that matter, an N -ary POF.
- B. V. K. V. Kumar and L. Hassebrook, "Performance measures for correlation filters," *Appl. Opt.* **29**, 2997-3006 (1990).
- B. V. K. V. Kumar, W. Shi, and C. Hendrix, "Phase-only filters with maximally sharp correlation peaks," *Opt. Lett.* **15**, 807-809 (1990).
- B. V. K. V. Kumar and Z. Bahri, "Phase-only filters with improved signal to noise ratio," *Appl. Opt.* **28**, 250-257 (1989).
- R. D. Juday, B. V. K. V. Kumar, and P. K. Rajan, "Optimal real correlation filters," *Appl. Opt.* **30**, 520-522 (1991).
- R. D. Juday and J. M. Florence, "Full complex modulation with two one-parameter SLMs," in *Wave Propagation and Scattering in Varied Media II*, V. K. Varadan, ed., *Proc. Soc. Photo-Opt. Instrum. Eng.* **1558**, 499-504 (1991).
- R. D. Juday and J. M. Florence, "Full complex spatial filtering with a phase mostly DMD," in *Wave Propagation and Scattering in Varied Media II*, V. K. Varadan, ed., *Proc. Soc. Photo-Opt. Instrum. Eng.* **1558**, 487-498 (1991).
- B. V. K. V. Kumar and L. Hassebrook, "Performance measures for correlation filters," *Appl. Opt.* **29**, 2997-3006 (1990).
- J. L. Horner, "Metrics for assessing pattern-recognition performance," *Appl. Opt.* **31**, 165-166 (1992).
- J. D. Downie, "Optimizing binary phase and amplitude filters for PCE, SNR, and discrimination," in *Aerospace Sensing*, *Proc. Soc. Photo-Opt. Instrum. Eng.* **1701** (to be published).