Approximating fully complex spatial modulation with pseudorandom phase-only modulation

Robert W. Cohn and Minhua Liang

Any desired diffraction pattern can be produced in the Fourier plane by the specification of a corresponding input-plane transparency. Complex-valued transmittance is generally required, but in practice phase-only transmittance is used. Many design procedures use numerically intensive, constrained optimization. We instead introduce a noniterative procedure that directly translates the desired but unavailable complex transparency into an appropriate phase transparency. At each pixel the value of phase is pseudorandomly selected from a random distribution whose standard deviation is specified by the desired amplitude. We also derive statistical expressions and use them to evaluate the approximation errors between the desired and achieved diffraction patterns.

Key words: Optical information processing, spatial light modulators, phase-only filters, laser speckle, rough surface scattering, statistical optics, binary and diffractive optics, phased arrays.

Introduction

This study is motivated by our desire to design phase-only filters and diffractive elements with a small amount of electronic computation, and thereby permit programming of arbitrary spatial modulation at real-time rates. Popular design procedures (e.g., the Dammann grating,[1–3] simulated annealing,[4] iterative constrained optimization,[5,6] and other iterative procedures) are only practical if performed off line because of both the numerical cost of performing Fourier transforms repeatedly and the further cost of evaluating the sensitivity of the transform with respect to a large number of pixels (frequently every pixel of the input-plane spatial light modulator.) Although it can be argued that for some applications all necessary phase-only filters can be computed off line and stored in memory, in other applications either too much memory is needed to do this affordably or there may not be advance knowledge of what filter is needed.

There are many procedures in the area of computer-generated holography, especially kinoforms, that permit direct synthesis of the input plane. These presuppose that the Fourier transform pair between the fully complex-valued input and the Fourier planes is known, and they work by encoding the desired complex values to appropriate phase settings. The direct-synthesis design procedures thus permit programming at real-time rates if the desired Fourier plane pattern is known.[7] The amount of memory is also minimized if the complex-valued Fourier transform pair can be written as an easily computed function.

Most frequently the direct procedures group a few adjacent pixels together as a single pixel that approximates several discrete settings over the complex plane (i.e., cell-oriented encoding).[8] However, this reduces the space bandwidth, which is already quite small (say 128 × 128 pixels) for current spatial light modulators, as compared with traditional fixed pattern holographic and diffractive optical elements. The procedure we present here is also a direct method, but one for which a continuous value of phase is selected for each individual pixel, independent of all other settings (i.e., point-oriented encoding.)[8]

Our method directly follows from mathematical models (presented below) of far-field diffraction from arrays of randomly phased point sources. More specifically, the phases are treated as independent and *nonidentically* distributed random variables. Many previous studies have considered random phase diffusers in which the phase statistics are identically distributed across the diffuser.[9,10] Their objective was improved holographic reconstruction of the intensity of objects, for which the phase was of no concern to the observer. In our design we are interested in reconstructing full complex objects from a variably random phase-only filter plane. In our analysis and

The authors are with the Department of Electrical Engineering, University of Louisville, Louisville, Kentucky 40292.

Received 15 September 1993; revised manuscript received 23 December 1993.

0003-6935/94/204406-10$06.00/0.

© 1994 Optical Society of America.

design procedure it is convenient to represent the random variables with a single probability-density function that is varied in its mean and standard deviation. Our analysis shows that, on average, the pixels with phases drawn from distributions of larger standard deviation behave as if they have smaller amplitude transmittances. This result originally occurred to us when we noted an average contrast loss for arrays of phase-only pixels that have identically distributed phase errors.[11] We now recognize that this earlier result is just a special case of the more general result for nonidentically distributed phases.

The ideal spatial light modulator (SLM) for our procedure modulates phase continuously over a full 2π range. Regularly spaced pixels are assumed in the examples below but are not required in our theory. These characteristics are typical of those anticipated for the piston-modulating (or flexure-beam) deformable mirror device[12] and are achievable with birefringent[13] and twisted nematic[14] liquid crystals. For binary and multilevel optics the quantized values of phase can be modeled as statistical departures from the desired analog phase, and this added effect on the diffraction pattern can be estimated.[11,15]

Concept

Our design procedure specifies the degree of randomness (i.e., standard deviation) of phase at each pixel as a way of approximating arbitrary values of amplitude. This is analogous to placing a diffuser of spatially varying roughness over the input plane. Increasing roughness at a pixel decreases its coherent contribution to the far-field pattern. The remaining incoherent or diffused light is spread over the entire diffraction pattern and contributes a noise background that is frequently referred to as speckle. The design phase at any pixel is specified as the expected value of its random phase distribution. Thus the full complex input plane can be viewed as a cascade of a deterministic phase screen with a variable roughness phase screen.

This interpretation follows from the statistical expectation of a complex exponential of a random argument, which is schematically illustrated in Fig. 1 and is derived in the following section. Figure 1 shows the spatial phase and amplitude for two types

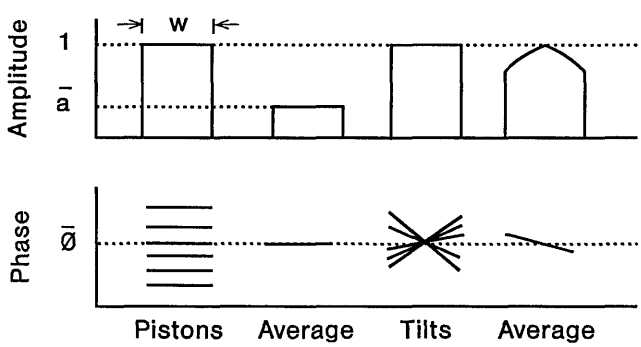


Fig. 1. Expected transmittance of random phase-only pixels (both piston-only and tilt-only pixels).

of phase-only pixels: piston-only pixels, those that produce a single value of phase across the pixel width, and tilt-only pixels, those that produce a linearly varying value of phase across the pixel width. The design procedure is concerned with the former structure, and the latter is provided as a further example of the concept. The leftmost of the four plots of amplitude and phase represents a random phase pixel. Because phase is a random variable we show an ensemble of phase values. The ensemble average of the random complex exponential produces the result in the second column: an average value of phase and a loss in amplitude transmittance. The example of pixels with random tilts shows a similar result in that the average amplitude transmittance decreases with increasing phase fluctuations that directly correspond with distance from the pivot point. Further discussion and analysis of the effects of random tilts are described in Ref. 11.

Because both expectation and Fourier transform are linear operations, the expected complex amplitude of the far-field pattern of the piston-only pixel, as well as an array of pixels, is the Fourier transform of the expected transmittance. The expected intensity of the far-field pattern consists of the magnitude squared of the complex amplitude plus a broad pedestal caused by the average intensity of speckle.

In our design procedure, rather than average many trials, we instead select a single value of phase from the ensemble for each pixel. The expected complex amplitude and intensity of the far-field pattern for this situation is mathematically equivalent to that described in the previous paragraph. More importantly, the actual far-field pattern approximately resembles the average pattern that results from the coherent summation of a large number of random wave fronts. Thus the formation of the far-field pattern can be viewed as a physical example of the central limit theorem, i.e., the so-called law of large numbers.[16]

Mathematical Analysis

Deterministic Expressions

The following definitions for arrays of piston-only pixels are used in the development of the statistical expressions. The complex transmittance of an individual pixel centered at position x equal to x_i in the input plane will be written as $a_i(x)$, and the transmittance of the array of the N individual pixels is

$$t(x) = \sum_{i=1}^{N} a_i(x) = \sum_i r(x - x_i)\exp(j\psi_i), \qquad (1)$$

where ψ_i is the phase shift produced by the ith pixel of the SLM. The appreviation $r(x) = \mathrm{rect}(x/w)$ has been used where w is the width of each pixel. Equation (1) shows that the amplitude and phase of each pixel have been defined in local coordinates centered around $x = 0$ and then shifted to pixel locations x_i. In this and subsequent equations the

inactive area between the pixels is treated as being zero amplitude. We do this to focus on the fundamental effect. It is also a reasonable assumption in light of current SLM's that either have nearly unity fill factors (in the case of deformable-mirror devices and binary optics) or that are fabricated with opaque shadow masks (such as current liquid-crystal displays.) The intensity of the far-field diffraction pattern is written

$$I(f_x) = T(f_x)T^*(f_x) = \mathcal{F}\{t(x) \oplus t(x)\}, \tag{2}$$

where $T(f_x)$ is the Fourier transform of the transmittance $t(x)$, i.e. $T(f_x) = \mathcal{F}\{t(x)\}$, and where $\oplus$ indicates the correlation integral

$$a(x) \oplus b(x) = \int a(x' + x)b^*(x')\mathrm{d}x'. \tag{3}$$

The squared intensity spectrum expressed in terms of the fourth-order autocorrelation of SLM transmittance is

$$I^2 = TT^*T^*T = \mathcal{F}\{[t(x) \oplus t(x)] \oplus [t(x) \oplus t(x)]\}. \tag{4}$$

Unlike convolution, the order in which correlations are performed affects the result, and thus the brackets are required in Eq. (4).

Statistical Expressions

The expected value of a complex phasor of a random argument is frequently referred to as the characteristic function[16] of the random variable ψ:

$$M'(\omega) = \langle \exp(j\omega\psi)\rangle = \int_{-\infty}^{\infty} \exp(j\omega\psi)p_\psi(\psi)\mathrm{d}\psi$$

$$= 2\pi\mathcal{F}^{-1}\{p_\psi(\psi)\}, \tag{5}$$

where $\langle\ \rangle$ is the ensemble average (i.e., expectation) operator and where $p_\psi(\psi)$ is the probability-density function of ψ. The last equality in Eq. (5) emphasizes the well-known Fourier transform relationship between the probability of the density function and the characteristic function. If the phases ψ_i are independent random variables, then the expected value of pixel transmittance $a_i(x)$ is

$$\bar{a}_i(x) = \langle a_i(x)\rangle = \langle r(x - x_i)\exp(j\psi_i)\rangle,$$
$$= r(x - x_i)M_i'(1),$$
$$= r(x - x_i)\exp(j\bar{\psi}_i)M_i(1), \tag{6}$$

where M_i' is the characteristic function of ψ_i and M_i is the characteristic function of the unbiased values of phase $\psi_i - \langle\psi_i\rangle$.[17] We have written the last line of Eq. (6) to identify the magnitude and phase components, and it corresponds to the average piston case in Fig. 1 (second column).

We considered two specific probability distributions for phase ψ_i. For Gaussian distributed phase of standard deviation σ_i,

$$M_i(1) = \exp\left(-\frac{1}{2}\sigma_i^2\right), \tag{7}$$

and for uniformly distributed phase with total spread,

$$v_i = \sqrt{12}\sigma_i, \tag{8}$$

$$M_i(1) = \mathrm{sinc}\left[\frac{v_i}{2\pi}\right]. \tag{9}$$

For either distribution the average amplitude transmittance at each pixel i can be controlled explicitly by selection of the value of the standard deviation σ_i. In digital simulations we prefer the uniform distribution because most Gaussian random-number generators are derived by performing additional numerical operations on uniform samples. A further advantage of using uniform statistics is that the total spread never need exceed 2π, whereas with Gaussian statistics the standard deviation can be infinite, and which can lead to overflow and underflow errors if not handled carefully.

The expected value of Eq. (1), using either probability distribution, is the expected SLM tranasmittance

$$\bar{t}(x) = \sum_{i=1}^{N} \langle a_i(x)\rangle = \sum_i r(x - x_i)p_i^{1/2}\exp(j\bar{\psi}_i), \tag{10}$$

where we use the parameter

$$p_i = M_i^2(1) \tag{11}$$

to simplify subsequent expressions. Note that because the expectation and Fourier transform operators are both linear, the expected far-field transmittance (more precisely, the angular spectrum) is

$$\bar{T}(f_x) = \mathcal{F}\{\bar{t}(x)\}. \tag{12}$$

The nomenclature in Eq. (12) of using lowercase letters for space-domain variables and uppercase variables for frequency-domain variables will be used throughout this discussion.

The most general expectation for the intensity of the far-field diffraction pattern of SLM's with statistically independent pixels follows from the first equality in Eqs. (1) and (2). It is

$$\bar{I}(f_x) = \sum_i\sum_j \langle A_i(f_x)A_j^*(f_x)\rangle,$$
$$= \sum_{i\neq j}\sum \langle A_i\rangle\langle A_j^*\rangle + \sum_i \langle |A_i|^2\rangle,$$
$$= \sum_i\sum_j \langle A_i\rangle\langle A_j^*\rangle - \sum_i |\langle A_i\rangle|^2 + \sum_i \langle |A_i|^2\rangle,$$
$$= |\bar{T}|^2 + \sum_i [\langle |A_i|^2\rangle - |\bar{A}_i|^2], \tag{13}$$

where the intermediate results indicate how we used independence to simplify the expression. This expression shows that the expectation nearly separates into $|\langle T\rangle|^2$, except for terms involving the autocorrelation of the input-plane pixels (i.e., terms for which $i = j$). Under the same assumption of independent pixels, we find that the most general expression for the squared far-field intensity is

$$\begin{aligned}\langle I^2\rangle &= \sum_i\sum_j\sum_k\sum_l \langle A_iA_j^*A_k^*A_l\rangle \\ &= 2[\bar{I}^2 - |\bar{T}|^4] + \left|\bar{T}^2 + \sum_i (\langle A_i^2\rangle - \bar{A}_i^2)\right|^2 \\ &\quad + 4\,\mathrm{Re}\Big[\bar{T}^*\sum_i(\langle|A_i|^2A_i\rangle - \langle A_i^2\rangle\bar{A}_i^* \\ &\quad + 2|\bar{A}_i|^2\bar{A}_i - 2\langle|A_i|^2\rangle\bar{A}_i)\Big] \\ &\quad + \sum_i [\langle|A_i|^4\rangle - 6|\bar{A}_i|^4 + 8\langle|A_i|^2\rangle|\bar{A}_i|^2 \\ &\quad - |\langle A_i^2\rangle|^2 - 2\langle|A_i|^2\rangle^2 \\ &\quad + 4\,\mathrm{Re}(\langle A_i^2\rangle\bar{A}_i^{*2} - \langle|A_i|^2A_i\rangle\bar{A}_i^*)]. \end{aligned} \tag{14}$$

It was found by substituting Eq. (1) into Eq. (4) and then taking the expectation of Eq. (4). The second quality was found by a procedure similar to that of Goodman.[18] (See Appendix A for more details.) The standard deviation of intensity $\sigma_I(f_x)$ is then directly found through the use of the well-known result

$$\sigma_I^2 = \langle I^2\rangle - \bar{I}^2. \tag{15}$$

For the specific case of piston-only, nonidentically distributed random phase, Eqs. (13)–(15) simplify to

$$\begin{aligned}\bar{I}(f_x) &= \mathcal{F}\Big\{\bar{t}(x)\oplus\bar{t}(x) + [r(x)\oplus r(x)]\sum_i q_i\Big\}, \\ &= |\bar{T}(f_x)|^2 + R^2(f_x)\sum_i q_i, \end{aligned} \tag{16}$$

where

$$q_i = 1 - p_i, \tag{17}$$

$$\begin{aligned}\sigma^2(f_x) &= \bar{I}^2 - 2|\bar{T}|^4 + |\bar{T}^2 - T_A|^2 \\ &\quad - 4\,\mathrm{Re}[\bar{T}T_B^*] - G_A. \end{aligned} \tag{18}$$

We introduce a shorthand for defining repeated autocorrelations of the rect function $r(x)$:

$$g_n(x) = [r(x)\oplus\cdots\oplus r(x)]_n, \tag{19}$$

where n indicates the number of rect functions (i.e., $n - 1$ integrals). Equation (18) is completely specified with the additional definitions for the Gaussian distribution

$$\begin{aligned} t_A(x) &= \sum_i g_2(x - 2x_i)\exp(j2\bar{\psi}_i)q_ip_i, \\ t_B(x) &= \sum_i g_3(x - x_i)\exp(j\bar{\psi}_i)q_i^2p_i^{1/2}, \\ g_A(x) &= g_4(x)\sum_i q_i^4, \end{aligned} \tag{20}$$

and for the uniform distribution

$$\begin{aligned} t_A(x) &= \sum_i g_2(x - 2x_i)\exp(j2\bar{\psi}_i)(p_i - \mathrm{d}_i), \\ t_B(x) &= \sum_i g_3(x - x_i)\exp(j\bar{\psi}_i)(q_i - p_i + \mathrm{d}_i)p_i^{1/2}, \\ g_A(x) &= g_4(x)\sum_i (q_i - 3p_i + 6p_i^2 - 4p_i\mathrm{d}_i + \mathrm{d}_i^2), \end{aligned} \tag{21}$$

where

$$\mathrm{d}_i = \mathrm{sinc}\left[\frac{v_i}{\pi}\right]. \tag{22}$$

Discussion and Interpretation

The expected intensity pattern, Eq. (16), contains the desired design intensity pattern, the magnitude square of Eq. (12), plus an additional term proportional to $G_2(f_x) = R^2(f_x)$ that we associate with the average level of speckle intensity, and which is often referred to as halo or pedestal.[19] This second term is proportional to the far-field diffraction pattern intensity of a single pixel. Examination of Eq. (18) also reveals that each of its terms has the common factor $G_4(f_x) = G_2^2(f_x)$. Thus for a standard definition of signal-to-noise ratio (SNR),

$$\mathrm{SNR}(f_x) = \frac{\bar{I}(f_x)}{\sigma_I(f_x)}, \tag{23}$$

the term $G_2(f_x)$ cancels out and is independent of spatial frequency. The reciprocal of Eq. (23) basically describes the relative approximation error (especially when σ_I is small compared with I and the contribution of pedestal can be practically ignored).

Another type of SNR that is quite common in describing the quality of an optical correlation peak is the ratio of peak intensity to background noise level; this is often referred to as the peak-to-noise ratio (PNR). The analysis above permits calculation of this as well; however, unless the pixels are point sources, the pixel element factor $G_2(f_x)$ has to be considered and so there will be different values of the SNR depending on the pixel fill factor.

Summary of the Design Procedure

The above analysis indicates that a phase-only source distribution can be treated as a full complex distribution for the purpose of approximating a desired

far-field pattern. The results from above can be used in a design procedure as follows. Specify the desired far-field pattern $T(f_x)$. Fourier transform to the desired source distribution $t(x)$. For each pixel i, invert Eq. (9) [preferred, or Eq. (7)] to find σ_i, and select a pseudorandom number, appropriately scaled by σ_i to represent the random phase ψ_i. Then Fourier transform the array of random phase-only pixels to evaluate the actual far-field pattern. If desired, compare the actual far-field pattern $I(f_x)$ with Eqs. (16), (18), and (23). In the next section we use this procedure to illustrate designs of diffraction patterns, including evaluations of the quality of the approximation.

Computer Verification and Demonstration of the Theory

General Experimental Conditions

Several designs have been simulated and analyzed, two of which are presented in some detail here. In every case the design is for a 128 × 128 pixel phase-only SLM. It is represented as a 128 × 128 array of samples embedded in the center of a 512 × 512 array of zeroes. We use a fast Fourier transform (FFT) routine to approximate the Fourier transform of the array and of the individual terms comprising the expectations in Eqs. (16) and (18).

The uniform random-number generator DRNUNF (from the IMSL numerical software library from IMSL, Inc., Houston, Tex. 77042-3020) with initialization subroutine RNOPT(6) was used. We mention this because initially there were problems when we used the random-number generator RAN1 given in Ref. 20. Specifically, when we attempted to verify Eqs. (16) and (18) by Monte Carlo estimation (i.e., ensemble averaging of a large number of identical designs generated with different random-number seeds), the estimates sometimes converged to different results than in the equations. The discrepancies were especially obvious at harmonically related sets of spatial frequencies. This may indicate that RAN1 produced correlated sequences in these experiments. These problems, however, were not discernible in individual design runs because of the magnitude of the random fluctuations.

When using DRNUNF we did find that plots of Monte Carlo estimates of the expected value and standard deviation of intensity, except for small fluctuations, appear to match our expressions closely when 1000 design runs are used in the estimate. We also checked that the relative error between the theory and the estimate decreases with an increasing number of design runs and that this improvement is roughly proportional to the square foot of the number of design runs used in the estimate, as would be expected. Our closest result, using 10^5 design runs, converged to the theoretical expressions with a fluctuation of less than 0.4% standard deviation.

Pseudorandom Encoding of Binary Amplitudes

Figures 2 and 3 illustrate the design of a phase modulation to approximate an elliptically shaped aperture of eccentricity 6:1. Figure 2(a) shows the desired amplitude modulation, and Fig. 2(b) shows the corresponding phase-only modulation. The mottled region in Fig. 2(b) corresponds to phase that is totally random (uniform statistics with spread $v_i = 2\pi$), and the white region corresponds to constant phase ($v_i = 0$). Figure 2(c) shows the expected intensity, Eq. (16) from the random modulated SLM, and Fig. 2(d) shows the resulting intensity pattern. It is basically an elliptical version of the Airy pattern (the exact far-field pattern of an elliptical aperture) on top of a uniform intensity (dark gray) background. The actual diffraction pattern [Fig. 2(d), the far-field pattern of the phase-only modulation] also resembles the elliptical Airy pattern on top of a speckle pattern. Figure 3 presents quantitative information along the vertical axis of Figs 2(c) and 2(d). The thin curves depict the error bars of $\pm\sigma_I$ added to the expected intensity. These provide a rough idea of the magnitude of the actual fluctuations, both of the coherent peak and the speckle-dominated sidelobe region.

It should be apparent that fluctuations and thus approximation errors will be smaller for those binary amplitude designs that use a greater number of nonrandom pixels. This relationship can be seen in Eq. (23), which simplifies to

$$\mathrm{SNR}(0) = \frac{N_N^2 + N_R}{(2N_R N_N^2 - N_R + N_R^2)^{1/2}} \approx \frac{N_N}{\sqrt{2N_R}} \quad (24)$$

at zero frequency where the diffraction intensity peaks.[21] Here N_R is the number of randomly modulated pixels on the SLM and N_N is the number of nonrandomly modulated pixels, and their sum is the total number of SLM pixels, 16,384. The approximation in Eq. (23) is valid except for SNR close to unity; for instance, for 1000 nonrandom pixels the relative error is roughly 1% and the SNR is 5.7. Figure 4 plots relation (24) against 500 run Monte Carlo estimates of SNR(0). These estimates correspond to elliptical aperture designs, with a major axis of 128 pixels and eccentricity ranging from 1 to 13, and circular apertures of diameter from 42 to 128 pixels. For reference, there are 2016 nonrandom pixels in the ellipse of eccentricity 6:1. This data point on Fig. 4 is nearly indistinguishable from the point for a circular aperture with a diameter of 50 pixels and that contains 1976 pixels.

A second performance measure describing the quality of the far-field diffraction pattern is the PNR,[22–24] which we choose to define here as

$$\mathrm{PNR} \equiv \frac{\bar{I}(0)}{\sigma_I(f_x)} \approx \frac{N_N^2}{N_R} \approx 2\mathrm{SNR}^2(0) \approx \frac{\bar{I}(0)}{\bar{I}(f_x)}, \quad (25)$$

where f_x is assumed to be a frequency in the sidelobe region. The approximations follow by assuming the expected transmittance is negligible, i.e. $\langle T(f_x)\rangle = 0$ in Eqs. (16) and (18), with respect to the speckle intensity in the sidelobe region. Additionally, we set the ratio of $G_2(0)/G_2(f_x)$ to unity, mainly to indicate more

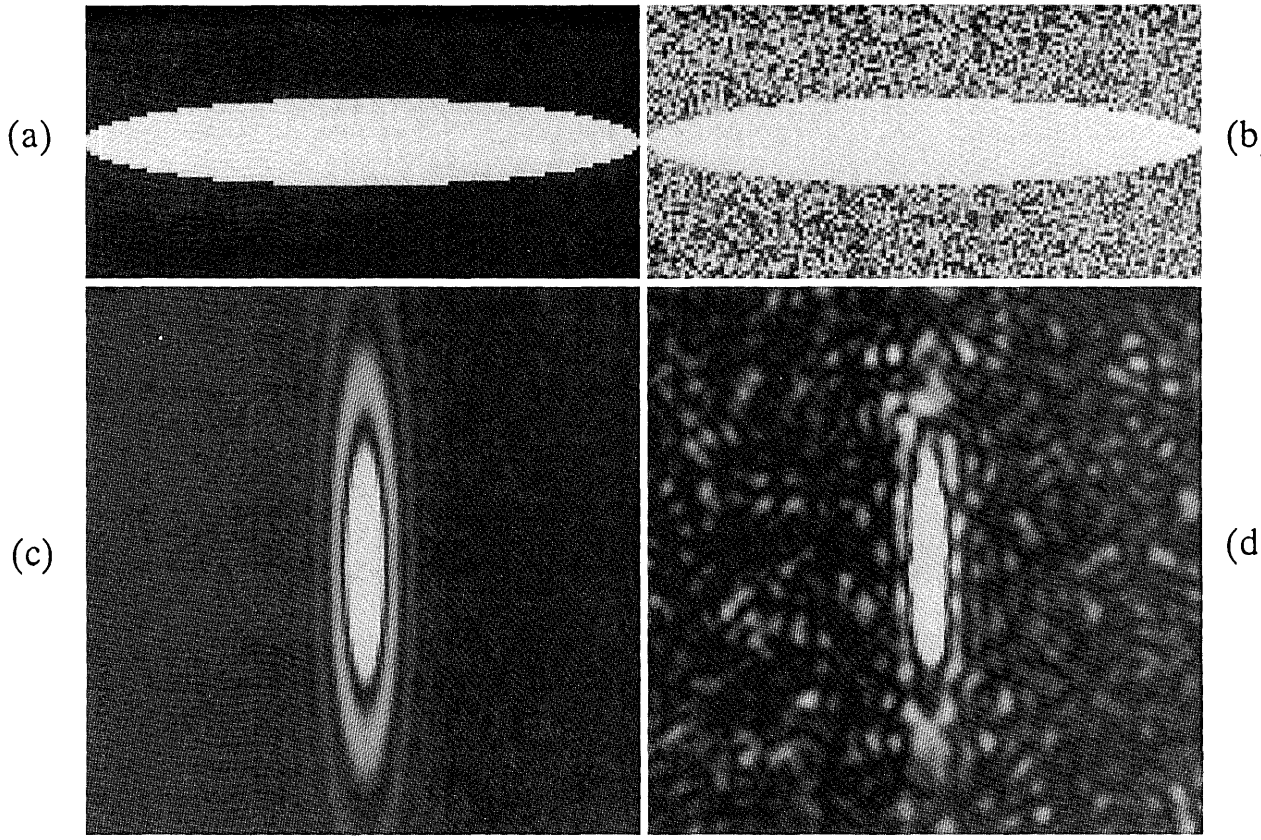


Fig. 2. Pseudorandom phase-only design of elliptical aperture. (a) Desired amplitude modulation. (b) Phase-only approximation to (a). (c) Expected value of (d) the intensity of the far-field diffraction pattern. Only the central 128 × 64 pixels of the 128 × 128 modulation are shown in (a) and (b). Only the central 128 × 128 pixels of the 512 × 512 FFT are shown in (c) and (d). We have nonlinearly transformed the intensities in (c) and (d) by the exponent 1.3 (i.e., gamma) to increase the contrast of low-lying sidelobes of the Airy pattern.

dramatically the simple relationship between PNR and SNR(0). In our simulations, for which we chose f_x as the Nyquist frequency [i.e., the point halfway between the (0, 0) and the (1, 1) diffraction order], the first approximation is within 1.5% of the exact value of the PNR for 1000 or more nonrandom pixels. The third approximation indicates that the average level of speckle in the sidelobe regions is equal to the standard deviation of intensity. This result is not unexpected based on previous observations that the intensity of fully developed speckle patterns is exponentially distributed and that the standard deviation of this distribution is equal to its mean.[11,19]

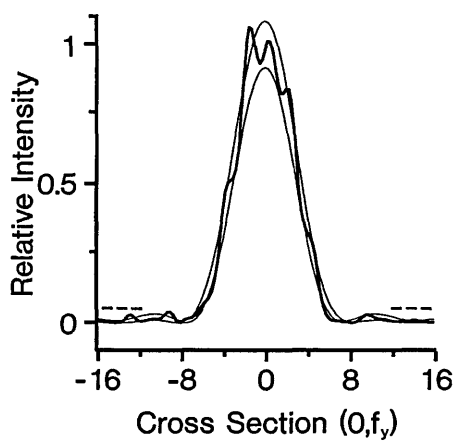


Fig. 3. Diffraction pattern of phase-only approximated elliptical aperture. The actual pattern (thick curve) is compared with the expected intensity plus and minus the error limits of one standard deviation. Horizontal dashed lines indicate the saturation (full-white) level of corresponding figures, Figs. 2(c) and 2(d). The spatial coordinates are normalized so that one unit corresponds to one of the 128 resolvable scan positions of the SLM or four samples of the FFT window.

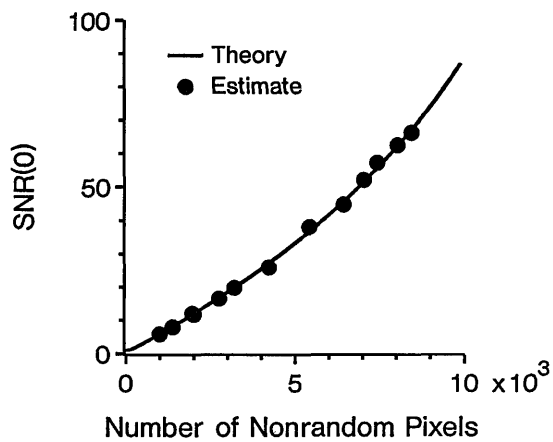


Fig. 4. SNR at diffraction peak as a function of number of nonrandom pixels for pseudorandom encoding of binary amplitudes.

We have also evaluated the design procedure for the approximation of an apodized input plane. Two-dimensional Gaussian, Airy pattern, and sinc functions with varying aspect ratios have been synthesized. We originally chose the sinc and Airy functions to produce brick-wall and top-hat diffraction patterns. However, the finite spatial extent of the SLM severely truncates the ideal functions and produces significant in-band ripple (Gibbs phenomena) and large sidelobes. The only feature that recommends this design is the extremely rapid transition band.

The transition bandwidth can be readily traded off with ripple and sidelobe level by the window design method that is widely used in the design of finite impulse response digital filters.[25] In this method one controls truncation effects by multiplying the ideal, infinite extent functions by an amplitude-tapered window in place of a rect function. One of the best windows for the design of uniform intensity patterns is the Dolph–Chebyshev function. It is considered optimal in the sense that its Fourier transform produces the minimum mainlobe width for a given sidelobe level.[26] The parameter α specifies the sidelobe level of 20α in decibels.

We have used the window method to design a pseudorandom phase modulation that will diffract into a close approximation of a brick-wall shape. The design function is a one-dimensional sinc function multiplied by a Dolph window in both coordinates. Our general design goal for this example is to produce a diffraction pattern with the largest aspect ratio (or eccentricity) possible that still reasonably resembles a brick-wall shape. We feel that for a 128×128 SLM this goal is met by the function $\mathrm{sinc}(4x)$ multiplied by Dolph functions in x and y that each have α equal to 1.3.

This result is presented in Figs. 5 and 6. In order to show the correspondence between amplitude in Fig. 5(a) and random phase in Fig. 5(b) more clearly, the gray-scale image shows, rather than phase error, the absolute phase error

$$\delta\psi(x_i) = |\psi_i - \bar{\psi}_i|. \tag{26}$$

The value of α has been chosen just large enough so that the sidelobes disappear in the speckle background in Fig. 5(d). They are still somewhat apparent in the expected intensity, Fig. 5(c). Figure 6(a) makes clear that increasing α further also increases the intensity of speckle, and overall passband ripple is not reduced because the random fluctuation is already larger than the ripple for the Dolph function. The aspect ratio of this diffraction pattern is 6.7:1 at

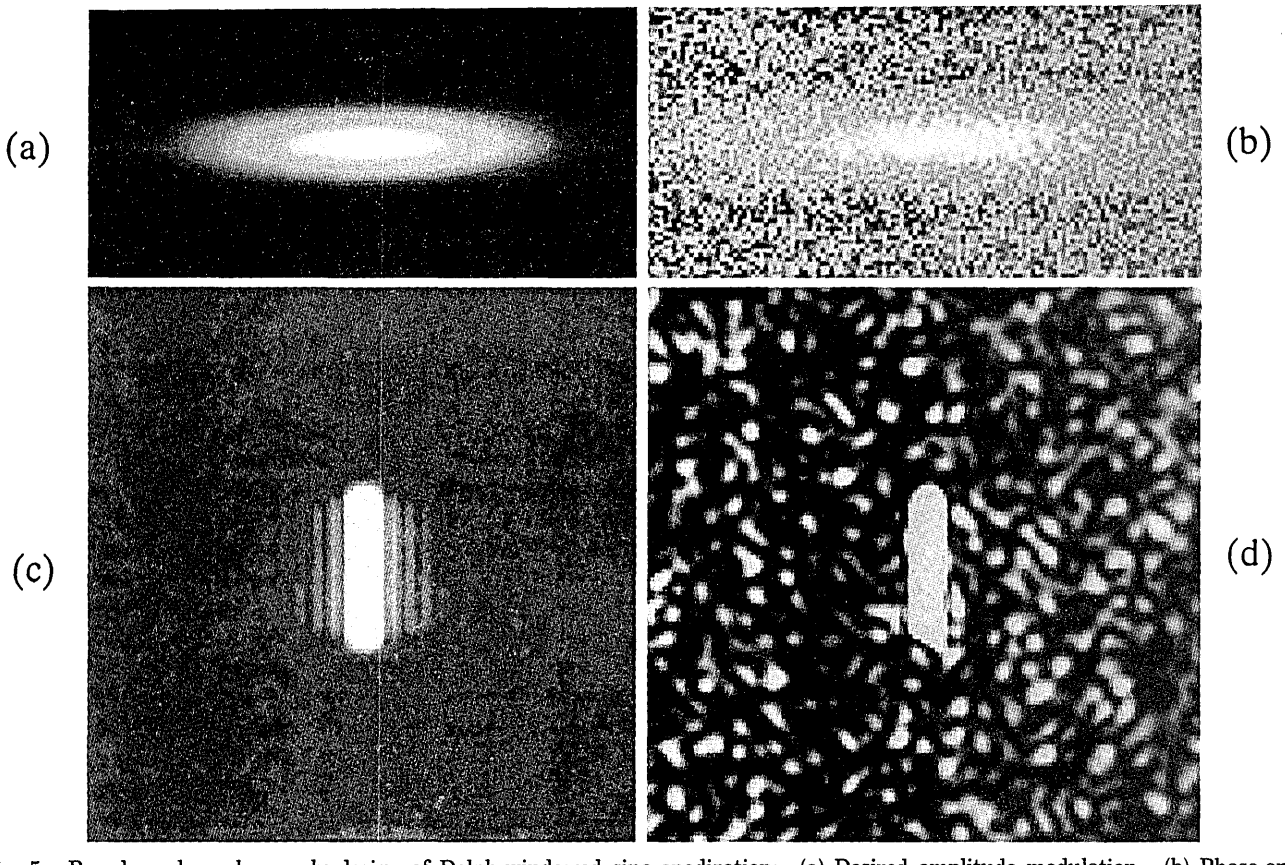


Fig. 5. Pseudorandom phase-only design of Dolph-windowed sinc apodization: (a) Desired amplitude modulation. (b) Phase-only approximation to (a). (c) Expected value of (d) the intensity of the far-field diffraction pattern. All units and settings are the same as in Fig. 2 except the recording gamma, which is unity for (c) and (d).

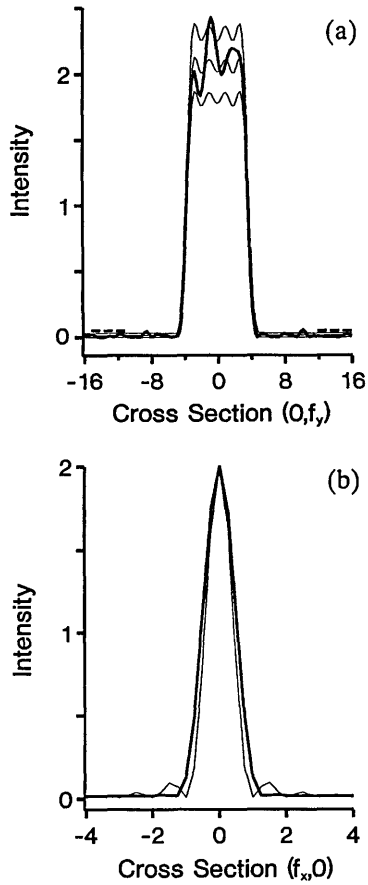


Fig. 6. Diffraction pattern of a phase-only approximated Dolph-windowed sinc apodization. (a) Cross section across vertical axis $(0, f_y)$. The actual pattern (thick curve) is compared with the expected intensity and the expected intensity plus and minus the errors limits of one standard deviation. Horizontal dashed lines indicate the saturation (full-white) level of corresponding figures, Figs. 5(c) and 5(d). Legend and units identical to those in Fig. 3. (b) Cross section across horizontal axis $(f_x, 0)$. The expected intensity (thick curve) is compared with the expected intensity $\text{sinc}^2(f_x)$ if no Dolph windowing had been employed.

the half-power points as measured from the plots in Fig. 6. Figure 6(b) also shows that the Dolph windowing in the vertical direction has widened the diffraction pattern by approximately 25% at the half-power points.

To compare the performance of various design functions with the phase-only encoding procedure, we introduce the definition of the effective number of random pixels,

$$\hat{N}_R \equiv \sum_i^N q_i. \tag{27}$$

Equation (27) is identically the amplitude of the noise pedestal term in Eq. (16), and it is identical to N_R for the case of binary modulation. Figure 7 compares the SNR for the various functions studied. The curves demonstrate a rough correspondence, with the apodized functions being more closely clustered than

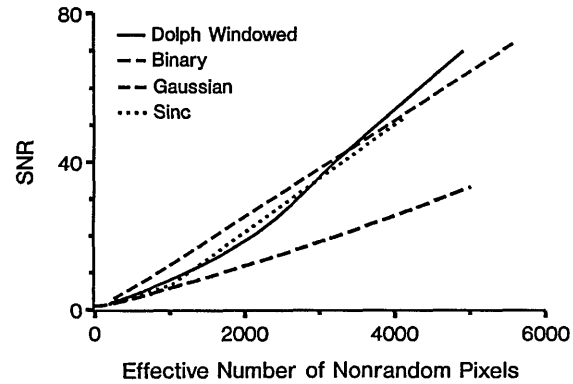


Fig. 7. SNR as a function of effective number of nonrandom pixels for pseudorandom encoding of continuous amplitudes. SNR(0) is plotted for Gaussian and binary, and average SNR across passband is plotted for Dolph and sinc designs. The sinc is a rectangularly symmetric function, and the Gaussian is a circularly symmetric function.

the binary amplitude functions. The circularly symmetric approximation to an Airy pattern on the input plane is not shown, but its performance nearly overlays that for the sinc curve in Fig. 7. For reference, the SNR of the design in Figs. 5 and 6 has a SNR of roughly 8, which corresponds to an effective number of 1000 nonrandom pixels.

Conclusions

We have presented a method, based on properties of random phase having spatially varying statistics, that approximates fully complex input-plane modulation. Diffraction patterns from pseudorandom phase-only modulation can be as energy efficient as any passive, fully complex modulation, with performance loss arising from the addition of a nearly uniform level speckle background. Designers can use the method to specify diffractive optical elements directly from Fourier transform relationships between the input plane and far field. The performance of any design can be readily evaluated, and the quality of the far-field patterns can be anticipated from the effective number of nonrandom pixels in the input plane. In the remainder of this section we consider potential applications of the method. The method is especially useful in that phase and amplitude are specified independently of each other. Although the examples in the last section demonstrate the approximation of amplitude-only inputs, the mathematical analysis shows that any value of phase (specified as the average value of phase) is permissible. For example, the far-field pattern can be translated by adding a phase slope to the design values of phase values. Therefore, one can use a single phase-only SLM to perform simultaneous and independent beam shaping and beam steering. Two-dimensional scanners can be envisaged that have much more flexibility than current ones. Scanning is not limited to rastered formats, and multiple spots can be formed. For such applications, it should generally not be necessary for

one to use a new set of random numbers each frame. Instead of generating them on the fly, one can store a single frame of random numbers in a video memory. A simple class of functions can also be selected for beam shaping that use a small number of function calculations and memory.

It may be possible to apply this flexibility in scanning to pattern recognition, tracking, and sensing. Currently raster scanning of lasers is used as a way to sense outlines of objects.[27] An edge point is identified by a sudden change in reflectance while scanning. A phase-only SLM-based scanner can locate edges and then adapt the beam shape to enhance the return from the edge. Spots could be contoured and positioned to fit over multiple edges and corners, and this could be used as a feature-based method[28] of tracking or recognition.

The pseudorandom design procedure may have various applications to the design of free-space optical interconnects. Certainly designs can be directly synthesized, and their performance can be readily evaluated; however, their optimality is not guaranteed. We believe that the approach may nonetheless reduce the computation time of other design procedures by providing a better initial guess for the iterative design procedures, especially simulated annealing approaches[4] that start with a pseudorandom array of weights anyway.

What our design procedure does not address is the numerous local minima that one forms by placing the constraint of a phase-only input plane on the design statement. For example, the Dammann grating has a multiplicity of unique solutions consisting of evenly spaced diffraction peaks of equal intensity and phases of 0 or π.[2] One finds the most diffraction-efficient design by exhaustively solving for every possible combination of phase. However, we already know that the most diffracation-efficient designs correspond to those that have a large number of nonrandom pixels. It may be possible for one to use this information to guide such global searches more efficiently.

In optical correlators, pseudorandom phase modulation can represent (typically) real-valued images in the input plane and complex filters in the filter plane with an accuracy that is reasonably modeled. We have proposed a compact phase-only correlator that uses a single phase-only SLM to perform both input- and filter-plane modulation.[29] The pseudorandom encoding is an alternative to nonlinearly transforming real-valued image data to phase. In the filter plane one may use pseudorandom modulation to approximate the full complex matched filter without employing a full complex SLM. Even though the performance will be reduced from that from full complex filters, it may provide useful experimental information for researchers in advance of practical full complex devices.

Pseudorandom phase-only design may well prove useful in many optical processing applications because it is a direct noniterative procedure, and straightforward mathematical expressions are available for the determination of performance bounds of the procedure.

Appendix A

We present a brief sketch of the derivation of Eq. (14), the second-order statistical moment of the far-field intensity $I(f_x)$. Under the assumption of statistically independent pixels, we find that the expectation simplifies to expectation over each pixel when subscripts i, j, k, and l are not equal; however, there are many cases where some or all of the subscripts are equal. In fact, there are 15 possible combinations of equal and nonequal subscripts. These are enumerated in Goodman.[18] The 15 terms can be grouped into seven distinct terms as

$$\begin{aligned}\langle I^2\rangle &= \sum_i\sum_j\sum_k\sum_l \langle A_iA_j{}^*A_k{}^*A_l\rangle \\ &= \sum\sum_{i\neq j\neq k\neq l}\sum \overline{A}_i\overline{A}_j{}^*\overline{A}_k{}^*\overline{A}_l + \sum_i \langle |A_i|^4\rangle \\ &\quad + 4\,\mathrm{Re}\sum_{i\neq j}\sum \langle |A_i|^2A_i{}^*\rangle\overline{A}_j \\ &\quad + 2\sum_{i\neq j}\sum \langle |A_i|^2\rangle\langle |A_j|^2\rangle + \sum_{i\neq j}\sum \langle A_i{}^2\rangle\langle A_j{}^{*2}\rangle \\ &\quad + 4\sum_{i\neq j\neq k}\sum\sum \langle |A_i|^2\rangle\overline{A}_j\overline{A}_k{}^* \\ &\quad + 2\,\mathrm{Re}\sum_{i\neq j\neq k}\sum\sum \langle A_i{}^2\rangle\overline{A}_j{}^*\overline{A}_k{}^*. \end{aligned} \tag{A1}$$

Each multiple summation with unequal subscripts can be rearranged as a few summations for which there is no inequality constraint on the subscripts. An example of this is presented in the derivation of Eq. (13) for the double summation that describes the expected value of the far-field intensity. The summations in Eq. (A1) over two, three, and four subscripts are handled in a similar manner, only the algebra is more tedious. One can handle the algebra more easily by using the following tensor shorthand:

$$\begin{aligned} T_{ijkl} &= \sum_i\sum_j\sum_k\sum_l a_ib_jc_kd_l, \\ T_{\substack{ijkl\\ i\neq j\neq k\neq l}} &= \sum\sum_{i\neq j\neq k\neq l}\sum a_ib_jc_kd_l. \end{aligned} \tag{A2}$$

Summations over two and three subscripts are similarly defined as T_{ij} and T_{ijk}. There should be no confusion between the use of the tensor symbol T_{ijkl} and transmittance $T(f_x)$ in this discussion.

The relationships between tensors with subscripts that can be equal and those that cannot be equal are

$$\begin{aligned} T_{\substack{ij\\ i\neq j}} &= T_{ij} - T_{ii}, \\ T_{\substack{ijk\\ i\neq j\neq k}} &= T_{ijk} - T_{iii} - T_{\substack{iij\\ i\neq j}} - T_{\substack{iji\\ i\neq j}} - T_{\substack{ijj\\ i\neq j}}, \\ &= T_{ijk} + 2T_{iii} - T_{iij} - T_{iji} - T_{ijj}. \end{aligned} \tag{A3}$$

The second line of the second equation in Eq. (A3)

follows from the first equation. Similarly, we use Eq. (A3) to derive

$$\sum_{\substack{ijkl \\ i\neq j\neq k\neq l}} T_{ijkl} = T_{ijkl} - 6T_{iiii} + T_{iijj} + T_{ijij} + T_{ijji} + 2(T_{iiij} + T_{iiji} + T_{ijii} + T_{ijjj}) - (T_{iijk} + T_{ijik} + T_{ijjk} + T_{ijki} + T_{ijkj} + T_{ijkk}). \quad \text{(A4)}$$

We use Eq. (A4) to represent the first term of Eq. (A1). When $\langle A_i \rangle = a_i = b_i^* = c_i^* = d_i$, as is the case in the first term of Eq. (A1), we can use this symmetry to simplify Eq. (A4) to

$$\sum_{\substack{ijkl \\ i\neq j\neq k\neq l}} T_{ijkl} = T_{ijkl} - 6T_{iiii} + 2T_{iijj} + T_{ijji} + 8\,\mathrm{Re}\,T_{iiij} - 4T_{iijk} - 2\,\mathrm{Re}\,T_{ijki}. \quad \text{(A5)}$$

We use Eqs. (A3) and (A5) to evaluate the six terms in Eq. (A1) that have unequal subscripts. Certain terms are recognized as $\langle I \rangle$ and $\langle T \rangle$ [see Eq. (10), (12), and (13)], which then leads to Eq. (14).

The authors are grateful to M. Hudak of Codonics, Inc. and S. Nowell of Alden Electronics, Inc. for providing hard copies of our gray-scale images. This research is sponsored by the Advanced Research Projects Agency and monitored by Rome Laboratory through contract F19628-92-K0021.

References and Notes

1. H. Dammann and K. Gortler, "High-efficiency multiple imaging by means of multiple phase holograms," Opt. Commun. **3,** 312–315 (1971).
2. J. Jahns, M. M. Downs, M. E. Prise, N. Streibel, and S. J. Walker, "Dammann gratings for laser beam shaping," Opt. Eng. **28,** 1267–1275 (1989).
3. F. B. McCormick, "Generation of large spot arrays from a single laser beam by multiple imaging with binary phase gratings," Opt. Eng. **28,** 299–304 (1989).
4. M. P. Dames, R. J. Dowling, P. McKee, and D. Wood, "Efficient optical elements to generate intensity weighted spot arrays: design and fabrication," Appl. Opt. **30,** 2685–2691 (1991).
5. N. C. Gallagher and B. Liu, "Method for computing kinoforms that reduces image reconstruction error," Appl. Opt. **12,** 2328–2335 (1973).
6. R. W. Gerchberg and W. O. Saxton, "Practical algorithm for the determination of phase from image and diffraction plane pictures," Optik (Stuttgart) **35,** 237–250 (1972).
7. If the complex Fourier transform pair is not known, then one FFT would have to be computed. Although such direct methods are much faster than the iterative methods, it is impractical to include electronic calculation of the FFT in real-time optoelectronic processors. Optical Fourier transform and interferometric detection with video cameras can be performed at real-time rates and might be practically employed in the latter case. See, for example, R. W. Cohn, "Adaptive real-time architectures for phase-only correlation," Appl. Opt. **32,** 718–725 (1993).
8. O. Bryngdahl and F. Wyrowski, "Digital holography—computer-generated holograms," in *Progress in Optics,* E. Wolf, ed. (Elsevier, Amsterdam, 1990), Vol. 28, pp. 1–86, and references therein, including reviews by W. J. Dallas and W. -H. Lee.
9. C. B. Burckhardt, "Use of random phase mask for the recording of Fourier transform holograms of data masks," Appl. Opt. **9,** 695–700 (1970).
10. R. Brauer, U. Wojak, F. Wyrowski, and O. Bryngdahl, "Digital diffusers for optical holography," Opt. Lett. **16,** 1427–1429 (1991).
11. R. W. Cohn and R. J. Nonnenkamp, "Statistical moments of the transmittance of phase-only spatial light modulators," in *Miniature and Micro-Optics: Fabrication,* C. Roychoudhuri and W. B. Veldkamp, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **1751,** 289–297 (1992).
12. R. M. Boysel, J. M. Florence, and W. R. Wu, "Deformable mirror light modulators for image processing," in *Optical Information Processing Systems and Architectures,* B. Javidi, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1151,** 183–194 (1989).
13. J. Amako and T. Sonehara, "Kinoform using an electrically controlled birefringent liquid-crystal spatial light modulator," Appl. Opt. **30,** 4622–4628 (1991).
14. N. Konforti, E. Marom, and S. -T. Wu, "Phase-only modulation with twisted nematic liquid-crystal spatial light modulators," Opt. Lett. **13,** 251–253 (1988).
15. R. W. Cohn, "Random phase errors and pseudorandom phase modulation of deformable mirror spatial light modulators," in *Optical Information Processing Systems and Architectures IV,* B. Javidi, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1772,** 360–368 (1992).
16. A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* 1st ed. (McGraw-Hill, New York, 1965).
17. The overbar in Eq. (6) is distinguished from the $\langle\ \rangle$ operator as a way to indicate that the overbarred variable is a result of ensemble averaging. At times, when the meaning is clear, we will interchange the use of the overbar and the $\langle\ \rangle$ operator in order to improve the readability.
18. J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985), App. C.
19. J. C. Dainty, ed., *Laser Speckle and Related Phenomena,* 2nd ed. (Springer, Berlin, 1984); Ref. 18, Chap. 8.
20. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge U. Press, Cambridge, England, 1986).
21. Equation (24) is identical for both Gaussian and uniform distributions.
22. B. V. K. V. Kumar and L. G. Hassebrook, "Performance measures for correlation filters," Appl. Opt. **29,** 2997–3006 (1990).
23. J. L. Horner, "Metrics for assessing pattern-recognition performance," Appl. Opt. **31,** 165–166 (1992).
24. R. W. Cohn and J. L. Horner, "Effects of systematic phase errors on phase-only correlation," Appl. Opt. (to be published).
25. A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing* (Prentice-Hall Englewood Cliffs, N.J., 1975), Sec. 5.5.
26. F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," Proc. IEEE **66,** 51–83 (1978).
27. H. Kamoda, H. Goto, and K. Imanaka, "Two dimensional optical position sensor using a dual axis miniature optical scanner," in *Miniature and Micro-Optics: Fabrication,* C. Roychoudhuri and W. B. Veldkamp, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **1751,** 280–288 (1992).
28. R. C. Gonzalez and P. Wintz, *Digital Image Processing,* (Addison-Wesley, Reading, Mass., 1977), Sec. 6.5.
29. R. W. Cohn, "Adaptive real-time architectures for phase-only correlation," Appl. Opt. **32,** 718–725 (1993).