Improved-fidelity error diffusion through blending with pseudorandom encoding

Li Ge, Markus Duelli,* and Robert W. Cohn

The ElectroOptics Research Institute, University of Louisville, Louisville, Kentucky 40292

Received December 17, 1999; revised manuscript received May 8, 2000; accepted May 17, 2000

Error diffusion (ED) and pseudorandom encoding (PRE) methods of designing Fourier transform holograms are compared in terms of their properties and the optical performance of the resulting far-field diffraction patterns. Although both methods produce a diffuse noise pattern due to the error between the desired fully complex pattern and the encoded modulation, the PRE errors reconstruct uniformly over the nonredundant bandwidth of the discrete-pixel spatial light modulator, while the ED errors reconstruct outside the window of the designed diffraction pattern. Combining the two encoding methods produces higher-fidelity diffraction patterns than either method produces individually. For some designs the fidelity of the ED–PRE algorithm is even higher over the entire nonredundant bandwidth than for the previously reported [J. Opt. Soc. Am. A **16**, 2425 (1999)] minimum-distance-PRE algorithm. © 2000 Optical Society of America [S0740-3232(00)00609-8]

OCIS codes: 230.6120, 090.1760, 070.2580, 030.6600.

1. INTRODUCTION

This paper continues an ongoing study on the development of new procedures for designing Fourier transform holograms.[1–8] The focus of the study has been to develop algorithms that can be computed in real or near-real time and that demonstrate good (rather than optimal) optical performance within the available time constraints for practical spatial light modulators (SLM's). These constraints frequently include that the SLM represents only a limited range of complex values (e.g., phase-only, quantized phase-only, coupled amplitude-phase) and that the SLM has a relatively small number of pixels [or equivalently, space–bandwidth product (SBWP)] compared with fixed-pattern diffractive optical elements and holograms. The motivation behind this research is the development of programmable optoelectronic systems that can automatically design and implement Fourier transform holograms in response to the unanticipated data presented by real-world situations.[9] Examples of such proposed systems include multispot-laser beam-steering systems, multitarget laser designator systems, and distortion-invariant pattern recognition systems implemented with composite filters in a coherent optical correlator.[9]

On the basis of these constraints of the SLM and of the real-world environment, we have avoided computationally intensive global searches for an optimal modulation, and instead we have considered both (1) noniterative encoding of the desired complex-valued function into a spatial modulation pattern and (2) iterative encoding in which one, or a few, free parameter(s) are varied to produce a suboptimal design.

Single-pixel encoding[8] (rather than group-oriented encoding)[10,11] is used. It maps each desired complex value into the modulation value of a corresponding single pixel, without consideration of the modulation values of the surrounding pixels. Single-pixel encoding has the advantages of low computational overhead, and the encoded modulation can have a SBWP that is identical to the SBWP of the desired (spatially sampled) signal. Since the SBWP of the SLM is identical to the number of pixels in the SLM, the far-field diffraction pattern can be reconstructed anywhere within the nonredundant bandwidth (NRB), i.e., the reciprocal of the pixel pitch. Group-oriented methods produce unwanted diffraction patterns within the NRB, which reduce the usable bandwidth to less than the SBWP of the SLM. Examples of single-pixel encoding include

1. Minimum-distance encoding[2] (MDE), in which each desired complex value is mapped to the closest modulation value produced by the SLM. For continuous-phase, phase-only SLM's this corresponds to the classical kinoform[13] or, in the case of pattern recognition filters, the phase-only matched filter.[14]
2. Pseudorandom encoding[1,4–6] (PRE), in which multiple modulation values are randomly selected on a percentage basis so that the expected value of the modulation equals the desired complex value.
3. Blended methods (referred to as MD–PRE)[2,3,8] that encode some desired values by MDE and other values by PRE.

Error diffusion[15–20] (ED) is yet another way to encode complex-valued functions in much less computation time than is required with global search methods. As with single-pixel methods, each desired complex value is encoded as a modulation value of a corresponding SLM pixel. However, the modulation value is also determined by the encoding errors (the difference between a desired value and an encoded value) of a few nearby pixels that have previously been encoded. The weighting factors for encoding errors are chosen so that the Fourier transform of the encoding errors is spatially separated from the desired diffraction pattern. Thus, as with group-oriented encoding, the ED reconstruction is limited to a bandwidth that is a fraction of the NRB of the SLM. This limitation does not restrict the location of the reconstruction within

the NRB, since the error weighting factors can be changed for each ED design to maintain separation between the error and the desired reconstruction.[18] However, the design of such weightings can be numerically complex or even impractical for desired patterns that span the NRB (e.g., a pattern of a few widely separated, randomly located spots.)

For wide-bandwidth diffraction patterns it may be simpler to use PRE methods for which the encoding errors reconstruct as a uniform-level noise pattern over the entire NRB. Thus, rather than using spatial separation between the desired pattern and the error pattern to obtain good performance, PRE attempts to distribute the error energy over the entire NRB, which results in low average error intensity everywhere. However, the maximum-intensity noise peak can be on the order of 10× larger than the average noise intensity. This is a consequence of the error pattern statistics, which are identical to the statistics of laser speckle;[1] specifically, the error pattern intensities are exponentially distributed, which makes possible a few noise peaks that are much more intense than the average. MDE also tends to produce a few bright noise spikes that appear at sum and difference frequencies of the desired pattern. Particular blendings of MDE with PRE have been demonstrated to produce lower peak noise (and also more accurate approximation of the intensities of the desired diffraction pattern) than either method individually.[3,8,21] We note that the reduction of noise spikes, particularly in spot array generators designed with a minimum-distance criterion, has been a motivation both for ED by Kirk *et al.*[19] and for MD–PRE.[3,8]

In this paper we present, to our knowledge for the first time, comparisons of ED and PRE in terms of their properties and optical performance. The general differences that we discussed above are brought out further by application of each encoding method to the same desired fully complex function and comparison of the resulting diffraction patterns. In addition to reviewing the ED and PRE algorithms, we show that a hybrid algorithm can be constructed out of the individual ED and PRE algorithms. We show that the ED–PRE blended algorithm outperforms both ED and PRE in terms of two fidelity metrics that measure noise spikes and accuracy between the desired and the resulting diffraction pattern. We also include comparisons of the blended ED–PRE method with the earlier MD–PRE method.

Section 2 presents a mathematical description of each algorithm evaluated in the study. Their performance is evaluated by computer simulation. The simulation procedure is described in Section 3, and the results and the performance comparisons are presented in Section 4.

2. DESCRIPTION OF THE ENCODING ALGORITHMS

The algorithms presented in this paper are specialized for phase-only SLM's that produce any value of phase continuously over 360°. Encoding algorithms for many other modulation characteristics are possible. A few of these include encoding for amplitude-phase coupled,[3,5,12] binary quantized,[16] and m-ary quantized characteristics.[6,8,12,17,21] Although specifying and evaluating such a variety of algorithms is beyond the scope of this paper, this section would also help in the development of blended algorithms for modulation characteristics other than phase-only. We first present the three individual algorithms MDE, ED, and PRE, followed by blended MD–PRE and the new ED–PRE.

When algorithms are encoded a desired modulation is specified. The desired function will be considered to be a discretely sampled two-dimensional array of, in general, complex numbers. The x coordinate will be associated with the index i and the y coordinate with index j. For the simulations in this paper the desired function and the SLM pixels always are equally spaced. A fully complex value from the desired function is written as $\mathbf{a}_{c\,ij}$ where boldface indicates a complex-valued quantity. After encoding, the function is mapped into the SLM modulation values $\mathbf{a}_{ij}$.

A. Minimum-Distance Encoding

MDE is presented because (1) it is a limiting case of ED when all the nearby error weighting coefficients are set to zero, (2) it is part of the blended MD–PRE algorithm, and (3) for phase-only SLM's it has the interesting property of producing the maximum diffraction efficiency for any encoding of a given complex-valued function.[22] An equivalent statement is that MDE produces the smallest total encoding error.

MDE is illustrated in Fig. 1(a). The unit circle represents the complex modulation characteristic of a phase-only SLM. The encoding algorithm is a direct point-by-point mapping of the desired value (along radial lines) to the closest point on the modulation characteristic, which is identical to kinoform design. The MDE algorithm for continuous phase-only modulation is written

$$\mathbf{a}_{ij} = \exp[j \arg(\mathbf{a}_{c\,ij})], \tag{1}$$

which has error

$$\mathbf{e}_{ij} = \mathbf{a}_{c\,ij} - \mathbf{a}_{ij}. \tag{2}$$

Desired values of any magnitude $[0,\infty]$ are mapped to the unit circle, as illustrated in Fig. 1(a).

B. Error Diffusion

ED [Fig. 1(b)] can be viewed as a modified version of MDE in that the value

$$\mathbf{a}_{ij} = \exp[j \arg(\mathbf{b}_{ij})], \tag{3}$$

where

$$\mathbf{b}_{ij} = \mathbf{a}_{c\,ij} + (\varepsilon_{i-1,j} + \varepsilon_{i,j-1})/2, \tag{4}$$

is a perturbed value of $\mathbf{a}_{c\,ij}$ and where the perturbations are the errors

$$\varepsilon_{ij} = \mathbf{b}_{ij} - \mathbf{a}_{ij} \tag{5}$$

for two nearest-neighbor samples to $\mathbf{a}_{c\,ij}$. The specific error samples and their weights (1/2 each) are identical to those used by Weissbach *et al.*[17] Other combinations of weights and error samples have been used to vary the error reconstruction pattern.[18] The specific algorithm given here will be used in the computer simulations that follow.

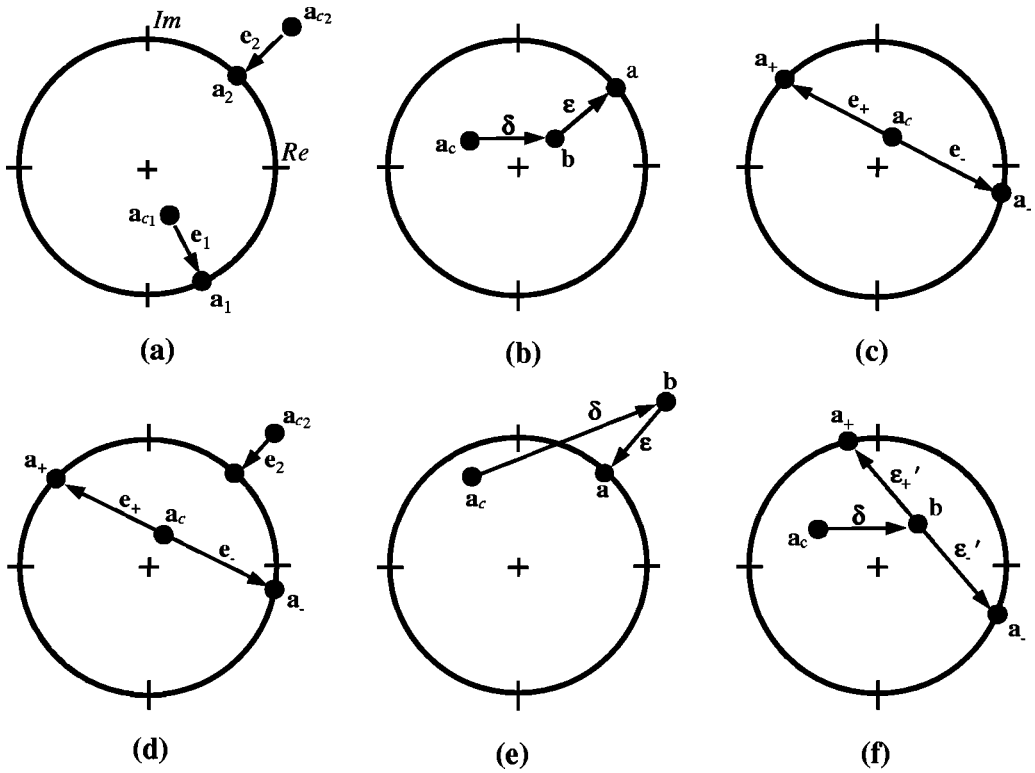

Fig. 1. Illustration of the encoding methods: (a) MDE, (b) ED, (c) PRE, (d) MD–PRE, (e) ED–PRE for $\mathbf{b}_{ij}$ outside the unit circle, (f) ED–PRE for $\mathbf{b}_{ij}$ inside the unit circle. The error diffused forward from previously encoded pixels is represented in the illustration by $\delta_{ij} \equiv (\varepsilon_{i-1,j} + \varepsilon_{i,j-1})/2$. In (f) the encoding error $\varepsilon' \equiv \varepsilon/\chi$, where ε is the amount of the encoding error that is diffused forward with use of Eq. (4).

C. Pseudorandom Encoding

For each desired complex value $\mathbf{a}_{c\,ij}$, PRE prescribes a random (complex-valued) variable $\mathbf{a}_{ij}$ such that the expected value $\langle \mathbf{a}_{ij} \rangle = \mathbf{a}_{c\,ij}$. One way that this condition can be met for phase-only SLM's is illustrated in Fig. 1(c). Here the encoded phase is $\psi_{ij} \equiv \arg(\mathbf{a}_{ij}) = \psi_{c\,ij} \pm \nu_{ij}/2$, where $\psi_{c\,ij} \equiv \arg(\mathbf{a}_{c\,ij})$ is the desired phase and $\pm\nu_{ij}/2$ is a phase offset. The positive or negative sign is randomly selected, each with a probability of 50%. If the uniformly distributed random variable $s_{ij} \in [-0.5, 0.5]$ is used, then the phase can be encoded as $\psi_{ij} = \psi_{c\,ij} + \mathrm{sgn}(s_{ij})\nu_{ij}/2$. For this phase random variable the expected value of complex transmittance is

$$\langle \mathbf{a}_{ij} \rangle = \cos(\nu_{ij}/2)\exp(\psi_{c\,ij}). \tag{6}$$

Values of the desired magnitude $|\mathbf{a}_{c\,ij}| \in [0, 1]$ are then encoded by choosing values of $\nu_{ij} \in [0, \pi]$ such that

$$\nu_{ij} = 2 \arccos(|\mathbf{a}_{c\,ij}|). \tag{7}$$

We will refer to this specific PRE algorithm as the inverse cosine algorithm.[5]

The error for encoding a single desired value is calculated by Eq. (2) with the encoded complex modulation $\mathbf{a}_{ij}$ as calculated in this subsection. This definition of encoding error differs from that in our previous papers.[1–8] In the earlier studies the *average* error contribution resulting from encoding a single pixel was evaluated rather than the *actual* contribution. This distinction is important because while several PRE algorithms have been developed for phase-only SLM's, they produce identical average errors and diffraction patterns of essentially identical performance.[7] However, when these various algorithms are blended with ED, (1) the actual error contributions are error diffused and (2) the performance of the diffraction patterns differs depending on which PRE algorithm is used.

For the particular desired functions $\mathbf{a}_{c\,ij}$ that we considered in this study, the inverse cosine PRE algorithm produces somewhat higher-fidelity reconstructions than the inverse sinc PRE method[1] and the phase reversal PRE method described in Sec. 3.C of Ref. 8. Therefore detailed simulations made with these alternate PRE algorithms are omitted because they provide little additional information over the results (presented in Section 4) found with the inverse cosine algorithm.

D. Blended Minimum-Distance Pseudorandom Encoding

MD–PRE was first introduced in Ref. 2 and has been further developed in Refs. 3 and 8. This method trades off desirable performance properties of the MDE algorithm with those of the PRE algorithm. MDE applied to a phase-only SLM is known to produce the highest possible diffraction efficiency for a given function of any encoding algorithm.[22] However, MDE is quite susceptible to intermodulation distortion[23] and (especially for spot array generators) can produce large sidelobes at sum and difference frequencies of the desired diffraction pattern.[24] PRE results in lower diffraction efficiency, and the errors between the desired and the resulting design are due to interference with background speckle, which is a necessary by-product of the PRE method. Various simulations and

experiments have shown that blending leads to overall better performance.[2,3,8] Specifically, the background speckle intensity is reduced and the efficiency is increased over PRE alone, and the distortion and sidelobe levels are reduced over MDE alone.

The MD–PRE algorithm can be expressed as follows:

If $|\mathbf{a}_{c\,ij}| > 1$ encode by MDE algorithm [Eq. (1)] as illustrated in Fig. 1(a).

Otherwise encode by PRE algorithm in Subsection 2.C as illustrated in Fig. 1(c).

The combined algorithm is illustrated in Fig. 1(d).

The performance of the algorithm depends on a single free parameter γ. This parameter is the maximum magnitude of the desired complex values $\mathbf{a}_{c\,ij}$. With PRE it is possible to encode only values of magnitude less than or equal to unity.[7] Therefore for PRE alone we normally scale the complex values of the desired function so that γ equals unity. PRE also permits encoding for values of γ that are less than zero, but this produces increased levels of speckle noise and lower diffraction efficiency.[4] However, MDE can encode values of any magnitude. For MD–PRE we have always found a particular value of γ greater than unity that minimizes the approximation errors and another value that minimizes the maximum-intensity noise sidelobe in the diffraction pattern. Currently there is no method that provides an *a priori* estimate of the optimal value of γ. Instead, the best value of γ is found by repetitive simulations of the encoding algorithm. We will show in Section 4 that the free parameter γ also controls the performance of ED and ED–PRE algorithms.

E. Blended Error-Diffusion Pseudorandom Encoding

The blending of ED with PRE is similar in philosophy to the blending of MDE with PRE. However, there are multiple possible ways that this might be accomplished. Among the various blending approaches we have considered are the following:

I. Apply ED [Eq. (3)] to values of $\mathbf{b}_{ij}$ outside the unit circle and apply PRE to values of $\mathbf{b}_{ij}$ inside the unit circle. The error from Eq. (5) for values encoded by ED is diffused forward by use of Eq. (4). The error from Eq. (2) for values encoded by PRE is not diffused forward; i.e., it is treated as zero in Eq. (4). The rationale for this is that the average error produced by PRE is automatically diffracted into speckle background and does not also need to be diffused into adjacent pixels.

II. Same as I except that the average error from PRE is diffused forward by Eq. (4).

III. Same as I except that the actual error from PRE is diffused forward by Eq. (4).

IV. Same as I except that a fraction $\chi \in [0,1]$ of the actual error from each PRE encoded value is diffused forward by Eq. (4). We will define the amount of error that is diffused from a PRE encoded pixel (rather than the total error) as

$$\varepsilon_{ij} = \chi(\mathbf{b}_{ij} - \mathbf{a}_{ij}). \tag{8}$$

We empirically found through various simulations and experimentation that method IV (for optimized values of χ) produces significantly better performance than methods I and II and somewhat better performance than method III. The sensitivity of method IV with respect to values of $\chi \in [0.5,1]$ is apparent though not dramatic (as is illustrated in Section 4.) Therefore the performance of method IV depends on the two free parameters γ and χ. Since PRE, MD–PRE, and ED–PRE all use random variables, the performance of each algorithm also depends on the particular random sequence used. In Section 4 we also present the variations in performance for an ensemble of random sequences. We are not recommending that ED–PRE be optimized in terms of all three variables (γ, χ, and which sample of an ensemble of random sequences is selected) for our envisioned real-time and near-real-time applications, but rather we present these analyses to provide insight into the performance of the algorithms.

In the remainder of this paper, method IV will be referred to as the ED–PRE algorithm. It is implemented as follows:

Given the $N = nm$ desired values $\mathbf{a}_{c\,ij}$, N uniformly distributed random numbers $s_{ij} \in [-0.5, 0.5]$, and specific values for γ and χ,

1. Normalize all $N\mathbf{a}_{c\,ij}$ so that the maximum of the values $|\mathbf{a}_{c\,ij}|$ equals γ.
2. For $i = 1$ to n and $j = 1$ to m.
3. Calculate $\mathbf{b}_{ij}$ using Eq. (4).
4. If $|\mathbf{b}_{ij}| > 1$ Encode $\mathbf{b}_{ij}$ to $\mathbf{a}_{ij}$ using Eq. (3) Calculate encoding error ε_{ij} using Eq. (5). Otherwise Encode

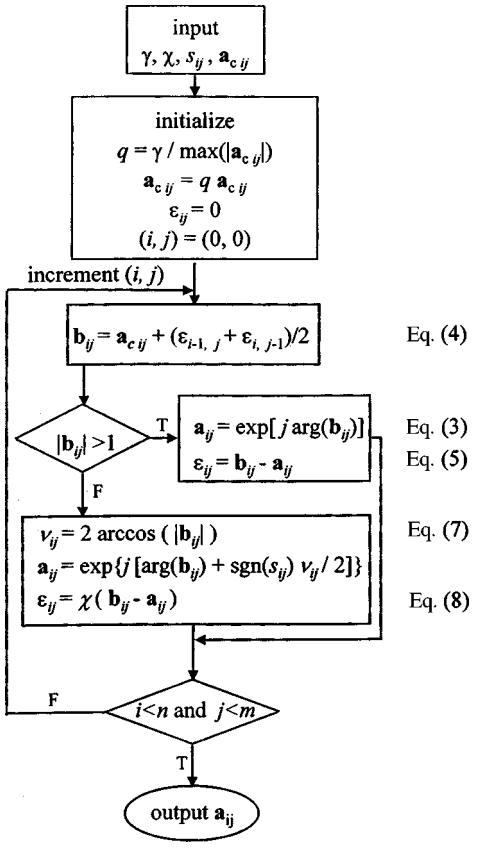

Fig. 2. Flow chart for the ED–PRE algorithm.

$\mathbf{b}_{ij}$ to $\mathbf{a}_{ij}$ using the PRE method in Subsection 2.C and Calculate ε_{ij}, the encoding error to diffuse forward, using Eq. (8).

A detailed flow chart of the complete ED–PRE algorithm is presented in Fig. 2. Figures 1(e) and 1(f) graphically illustrate the blended encoding process for values of $\mathbf{b}_{ij}$ both inside and outside the unit circle.

3. SIMULATION PROCEDURES

This section presents simulated diffraction patterns for, and evaluates the performance of, the various encoding algorithms for two specific test patterns, and the procedures used in those experiments, are described.

The ED, MD–PRE, and ED–PRE, as described in Section 2, are implemented for values of $\gamma \in [1, 2.5]$ in increments of 0.1. PRE corresponds to MD–PRE with $\gamma = 1$. Also, the performance of MDE is reported. It corresponds to MD–PRE with $\gamma = \infty$. It was observed for the functions encoded in this study for $\gamma \geqslant 2.5$ that the performance of MD–PRE is nearly identical to the performance of MDE. ED–PRE is implemented for values of $\chi \in [0, 1]$ in increments of 0.1. One set of simulations is performed, with the identical set of random numbers s_{ij} used for each encoding and for each value of γ and χ. These simulations are used to make direct comparisons of the performance of the algorithms. A second set of simulations is performed to determine the statistical variations in an ensemble of runs. For this evaluation the simulation is repeated an additional 20 times, each time with a different N sample sequence s_{ij}. The best-case and worst-case performance is reported for the 21 trials of each algorithm over a range of values of γ and for fixed values of χ.

Two test functions are selected for encoding. For continuity with our previous studies we use the same $N = 128 \times 128$ pixel test function $\mathbf{a}_{c\,ij}$, which produces a 7×7 spot array.[6,8] This function reconstructs off axis. The function was selected because it is typical of current diffractive optic designs that have diffraction efficiencies close to the theoretical maximum.[22] Specifically, when the fully complex function is encoded by MDE, the diffraction efficiency η is found to be 96% (see Table 1 and 2 below). The fully complex function itself has a diffraction efficiency of ~44%. The efficiency indicates how much amplitude information is in the function. Since diffraction efficiency of the fully complex function $\mathbf{a}_{c\,ij}$ can be shown to be identical to the average intensity of the modulation,[4] the square root of η gives the root-mean-square amplitude of the complex function of 0.66. This shows that a significant part of the fully complex function is not phase-only and requires some type of encoding. For comparison, the diffraction efficiency of a fully complex function used to generate a 49-spot array can be as low as 2% (when the phases of all diffracted spots are identical), and the root-mean-square amplitude is then 0.14. We know that the encoding errors [see Eqs. (2), (5), and (8)] would be greater and the performance would decrease for the lower-efficiency complex function;[1,4] however, further consideration of this point is beyond the scope of this paper.

The second test function is identical to the first except that a linear phase ramp has been removed from the first test function, so that its diffraction pattern reconstructs centered on the optical axis. Diffraction patterns centered at multiple locations are evaluated to give a more complete appreciation of the performance of the various encoding algorithms. (However, in practice one should anticipate the presence of an on-axis order because of practical limitations in perfectly controlling the fabrication of a diffractive optic or the phase settings of a SLM.)

The diffraction pattern is simulated by performing a fast Fourier transform (FFT) of the test function. The modulation value of an SLM pixel is represented by a single complex number. The 128×128 array of numbers is zero padded to form a 512×512 array that is transformed by the FFT. The padding samples and interpolates the diffraction pattern at 1/4 the diffraction limit, thus producing a realistic-looking diffraction pattern.

Diffraction patterns simulated in this way are evaluated to determine signal-to-noise ratio (SNR), signal-to-peak-noise ratio (SPR), and nonuniformity (NU). NU measures the relative deviation of the spot array from perfectly uniform. The peak intensities of the 49 spots is measured. The average and standard deviation of the intensities are calculated, and NU is the ratio of the standard deviation to the average intensity.

The average intensity of the 49 spots is also used in SNR and SPR calculations. For SNR the average spot intensity is divided by the average noise intensity. For SPR the average spot intensity is divided by the peak-noise sidelobe. The noise intensities are calculated with use of the entire 512×512-sample diffraction pattern region excluding a 128×128 window that just surrounds the spot array.

In previous studies we have used SPR and NU as our key measures of fidelity.[3–8] We also have reported SNR to provide comparisons with the work of other authors. However, since ED does not uniformly distribute noise of the full NRB, we also calculate a reduced-bandwidth SPR. For the modified SPR (SPR_m) the peak-noise intensity is found by using the intensity pattern that occupies the central 256×256 of the 512×512-sample image of the diffraction pattern.

In addition, the diffraction efficiency η is evaluated. However, rather than using the 512×512 FFT, we calculate the FFT of the 128×128 array directly without the zero padding. Efficiency is calculated as the sum of the intensities of 49 spots divided by the total energy in the 128×128-point diffraction pattern. This analysis does not take into account device-specific pixel aperture effects that determine the amount of energy diffracted into higher-order replicas that are outside the NRB of the diffraction pattern.

Gray-scale images of the intensity patterns are presented for several encodings to provide additional information on the generation of background noise. To bring out the background noise, we saturate the gray-scale level in each image so that full white corresponds to 3% of the average peak intensity. In each case, images of the entire NRB (i.e., all 512×512 samples) are shown.

4. COMPARISONS OF THE ENCODING ALGORITHMS

The results of encoding by ED, MDE, MD–PRE, and ED–PRE are presented in this section. Results for both on-axis and off-axis test functions are given. For each function the discussion first focuses on results derived with the single random sequence s_{ij} and preferred values of χ. Then the results are presented of the sensitivity analyses as a function of χ and for the ensemble of 21 random sequences.

For better appreciation of the property improvements possible with blended algorithms, we first present some reference designs using ED and MDE. Figure 3 shows how the background noise differs as a function of γ for the ED algorithm. For the ED algorithm as described in Ref. 17 the value of γ is presumably 1. The resulting diffraction pattern is shown in Fig. 3(a). The most pronounced noise appears at the corners of the 512×512-sample image. Faint noise peaks from the noise cloud extend out into the upper-left and lower-right corners of the (256×256-sample) reduced-bandwidth window. For $\gamma = 1.2$ the peak noise at the corners of the image [Fig. 3(b)] is reduced, and faint noise spikes appear over a larger extent of both the full NRB and the reduced-bandwidth window. An even more uniform distribution of noise spikes is seen in Fig. 3(c) for $\gamma = 1.3$. For γ near 1.6 the image [Fig. 3(d)] closely resembles the MDE design [Fig. 3(e)]. This result was unexpected. However, if γ is increased further, the diffraction pattern becomes quite distorted and no longer resembles the MDE design [Fig. 3(f)].

To gain some insight into the similarity between the results for MDE and ED with $\gamma = 1.6$, we produced histograms of the deviations $|\mathbf{a}_{c\,ij} - \mathbf{a}_{m\,ij}|$ between the encoded values produced by ED and by MDE, where the subscripts e and m designate values encoded by ED and MDE, respectively. Since the encoded values are on the unit circle, the largest deviation possible would be 2. For $\gamma = 1$ more than 50% of the deviations of the 16,384 pixels are larger than 0.6. For $\gamma = 1.1$ the histogram is nearly uniformly distributed over the full range from 0 to 2. However, as γ is increased further, the histograms have an increasing number of deviations near zero. For $\gamma = 1.5$ the deviations are less than 0.25 for 87% of the pixels, and there are deviations in excess of 0.6 for less than 4% of the pixels. For $\gamma = 1.6$ only 82% of the deviations are less than 0.25; but an even smaller amount, 1% of the deviations, exceed 0.6. The ED gray-scale image appears most similar to the MDE image when the number of large deviations are minimized. As γ is increased further, the percentage of large deviations increases to the point that the histogram becomes nearly uniformly distributed between 0 and 2. Thus for the encoding of the particular test function, the ED design tends to converge to, and produce a somewhat close approximation to, the MDE diffraction pattern for γ in the range 1.4–1.7.

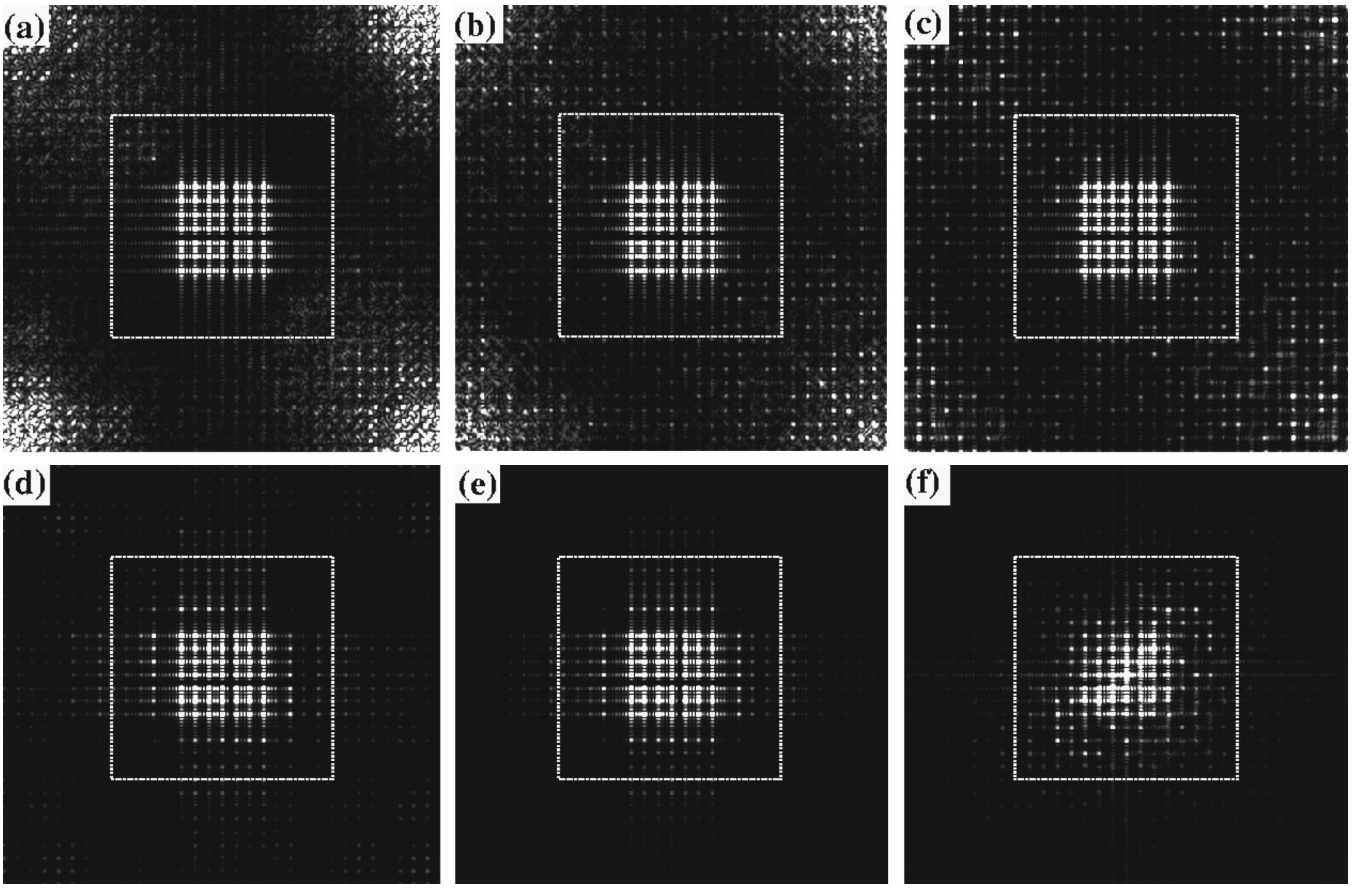

Fig. 3. On-axis diffraction patterns for encoding by (a)–(d) and (f) ED and (e) MDE. The value of γ used for ED is (a) 1.0, (b) 1.2, (c) 1.3, (d) 1.6, and (f) 5.5. The dotted square encloses the area used to calculate SPR_m. The gray-scale intensities are scaled so that full white corresponds to 3% of the average peak intensities of the 49 desired spots.

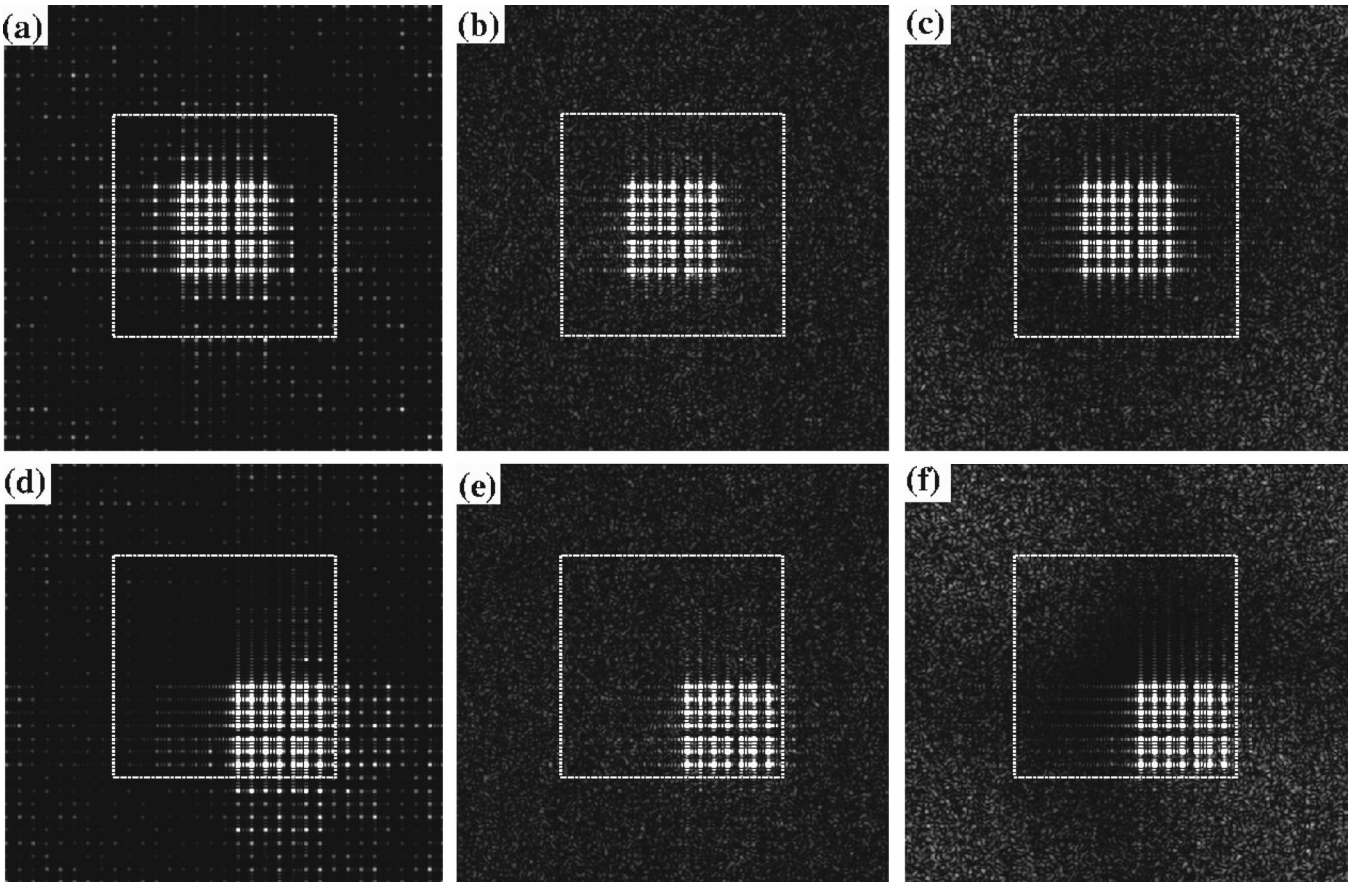

Fig. 4. (a)–(c) On-axis and (d)–(f) off-axis diffraction patterns for the maximum SPR design by (a) ED, $\gamma = 1.5$; (b) MD–PRE, $\gamma = 1.4$; (c) ED–PRE, $\gamma = 1.3$, $\chi = 0.6$; (d) ED, $\gamma = 1.5$; (e) MD–PRE, $\gamma = 1.4$; (f) ED–PRE, $\gamma = 1.3$, $\chi = 1$. The gray-scale normalization and the dotted square are identical to those used in Fig. 3.

A. Comparisons of On-Axis Designs

Figures 4(a)–4(c) show the background noise produced by ED, MD–PRE, and ED–PRE designs that have been tuned by γ and χ to produce the maximum value of SPR. Several performance metrics for each of these designs are reported in Table 1. For the ED algorithm with $\gamma = 1.5$ [Fig. 4(a)] the noise peaks appear to be even more uniformly distributed than for any of the designs in Fig. 3. For the MD–PRE algorithm [Fig. 4(b)] the noise background appears to be the most uniformly distributed of any result in Fig. 4. For the ED–PRE algorithm [Fig. 4(c)] the noise pattern demonstrates good features of both: of ED in that the noise is most pronounced in the corners of the image, but of MD–PRE in that the noise is more uniformly distributed over the corner regions than with ED. Thus ED–PRE appears to randomize the noise background, which reduces the noise peaks.

The background noise can be compared quantitatively in terms of SPR and SNR. Results for various encodings are presented in Fig. 5 and Table 1. Figure 5 presents both SPR and SPR_m (see Section 3) for ED, MD–PRE, and ED–PRE. For the ED design SPR_m is maximum for $\gamma = 1$. For MD–PRE at $\gamma = 1$ (i.e., the PRE algorithm) SPR_m is lower than for the ED design. However for $\gamma = 1.4$, SPR_m for MD–PRE is larger than for any ED design. The ED–PRE design results in an even larger value of this fidelity metric. For the full NRB the largest values of the SPR metric for ED–PRE and MD–PRE are essentially identical to those for SPR_m. In fact the entire SPR curve for MD–PRE is identical to the SPR_m curve. This is due to the presence of MDE-type sidelobes [e.g., in Fig. 3(b)] that are at the level of the random background noise. The SPR of the ED design is severely reduced (to a level even less than for MDE) by the inclusion of the noise peaks in the corner of the diffraction pattern.

Table 1. Best Encoding Performance for the On-Axis Function

Encoding Method	χ	γ	η(%)	SPR	SNR	NU(%)
Maximum SPR Design						
ED–PRE	0.6	1.3	72	53	784	3.3
ED	—	1.5	91	12	2660	10.2
MD–PRE	—	1.4	76	47	1000	5.8
Minimum NU Design						
ED–PRE	0.9	1.1	53	17	350	1.6
ED	—	1.2	58	4	421	5.2
MD–PRE	—	1.3	70	41	727	5.5
MDE[a]	—	$\propto$	96	17	5220	19.1
PRE[b]	—	1.0	44	24	258	7.9

[a] For MDE γ is not an adjustable parameter.
[b] PRE has both best SPR and NU for γ equal to unity.

Therefore the blended algorithms provide a way to reduce peak noise over the full NRB by distributing the noise more uniformly.

The other key fidelity metric is NU. Table 1 shows that the ED–PRE design having the highest SPR also has a lower value of NU than does ED, MDE, or MD–PRE. Thus it outperforms the MD–PRE design in both SPR and NU despite having a somewhat lower SNR and diffraction efficiency η.

Table 1 also reports the performance of designs that produce the lowest overall value of NU. The ED–PRE algorithm produces an exceptionally low value of NU; however, SPR is now even lower than for MDE, and η is much lower than for the maximum SPR design. For the minimum-NU MD–PRE design γ is lower only by 0.1 than γ for the maximum SPR design, and all the performance differences between the two designs as measured by the metrics in Table 1 are only slight. The NU and SPR curves in Fig. 5 also suggest that MD–PRE is less sensitive to γ than is ED–PRE.

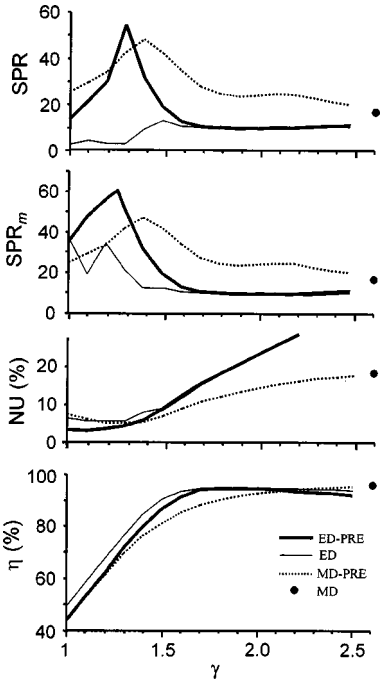

Fig. 5. Performance curves of the various encoding algorithms as a function of γ for the on-axis test function.[25] For ED–PRE the specific curve shown for NU is for $\chi = 0.9$, for SPR is for $\chi = 0.6$, and for SPR$_m$ is for $\chi = 0.7$. These curves achieve the best performance for ED–PRE as reported in Table 1.

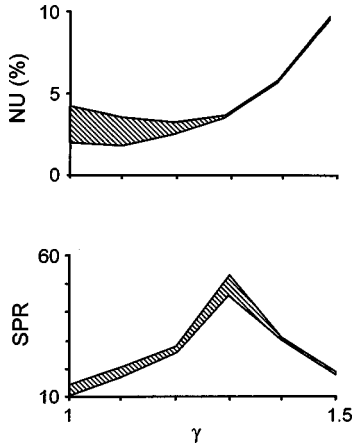

Fig. 6. Sensitivity of fidelity metrics of ED–PRE to the free parameter $\chi \in [0.5, 1]$ for the on-axis test function.

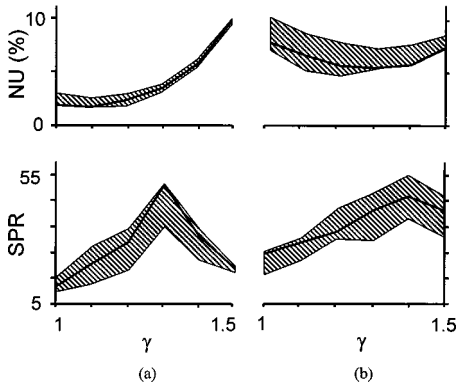

Fig. 7. Statistical variations of the fidelity metrics of (a) ED–PRE and (b) MD–PRE for an ensemble of encodings of the on-axis function. Shaded regions are bounded by the maximum and minimum values found for 21 random trials of each encoding algorithm. The respective curves from Fig. 5 are reproduced for comparison.

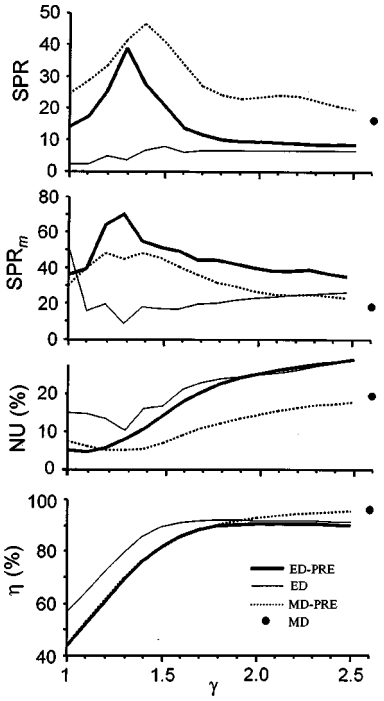

Fig. 8. Performance curves of the various encoding algorithms as a function of γ for the off-axis test function. For ED–PRE the specific curve shown for NU is for $\chi = 0.7$, for SPR is for $\chi = 1$, and for SPR$_m$ is for $\chi = 0.6$. These curves achieve the best performance for ED–PRE as reported in Table 2.

The results shown for ED–PRE in Fig. 5 are for the value of χ that produces either the highest value of SPR or the lowest value of NU. Figure 6 shows the range of variation of these performance metrics for values of $\chi \in [0.5, 1]$. The value of χ appears to have the greatest effect on NU, and the sensitivity decreases with increasing γ.

Figure 7 shows the range of variation of SPR and NU for an ensemble of the 21 random sequences of s_{ij}. The variation of ED–PRE is for the same values of χ as reported in Table 1. The ED–PRE curves from Fig. 5 are replotted for comparison. Whereas Fig. 5 and Table 1 report that ED–PRE produces higher SPR than MD–PRE, Fig. 7 shows for the ensemble that MD–PRE can produce a slightly higher value of SPR. A much larger ensemble is needed for us to be able to say whether MD–PRE outperforms ED–PRE in SPR on average and how frequently. Figure 7 also shows that ED–PRE can have substantially lower NU than MD–PRE and that NU for ED–PRE is much less sensitive to statistical variation than is MD–PRE.

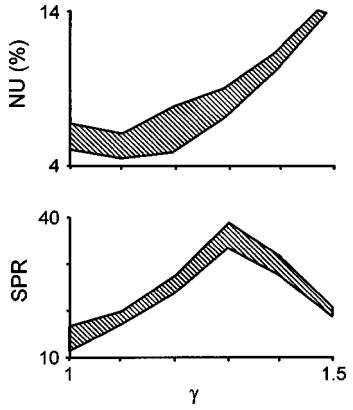

Fig. 9. Sensitivity of fidelity metrics of ED–PRE to the free parameter $\chi \in [0.5,1]$ for the off-axis test function.

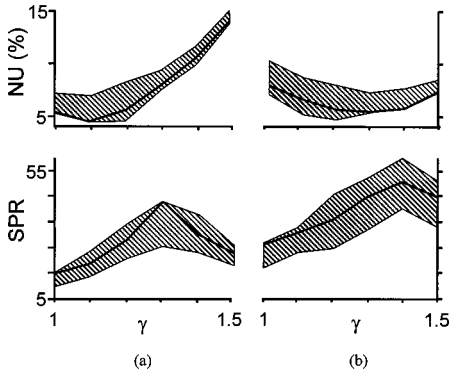

Fig. 10. Statistical variations of the fidelity metrics of (a) ED–PRE and (b) MD–PRE for an ensemble of encodings of the off-axis test function. Shaded regions are bounded by the maximum and minimum values found for 21 random trials of each encoding algorithm. The respective curves from Fig. 8 are reproduced for comparison.

Table 2. Best Encoding Performance for the Off-Axis Function

Encoding Method	χ	γ	η(%)	SPR	SNR	NU(%)
Maximum SPR Design						
ED–PRE	1.0	1.3	66	39	621	9
ED	—	1.5	59	8	2360	16.5
MD–PRE	—	1.4	76	47	1000	5.8
Minimum NU Design						
ED–PRE	0.7	1.1	52	18	340	4.5
ED	—	1.3	80	5	772	10.2
MD–PRE[a]	—	1.3	70	41	727	5.5

[a] MDE, PRE, and MD–PRE have identical performance for either off-axis or on-axis function. See Table 1 for MDE and PRE performance.

B. Comparisons of Off-Axis Designs

Figures 4(d)–4(f), Figs. 8–10, and Table 2 present the results for the off-axis design. These follow the same format as do Figs. 4(a)–4(c), Figs. 5–7, and Table 1. Figures 4(d), 4(e), and 4(f) presents the intensity diffraction patterns for ED, MD–PRE, and ED–PRE, respectively. The background noise pattern for each encoding method demonstrates textures (i.e., spiky for ED, white diffuse for MD–PRE, and colored diffuse for ED–PRE) similar to those shown in Figures 4(a)–4(c) for the on-axis design. Figure 8 and Table 2 compare the performance of the three algorithms. The major differences between the performance of the on-axis and the off-axis designs are that for the off-axis design (1) ED–PRE now clearly produces a much larger value maximum value of SPR_m than does MD–PRE; however, MD–PRE produces a much larger value of SPR than does ED–PRE; and (2) the minimum value of NU produced by ED–PRE is larger and is only slightly lower than the value of NU for MD–PRE. A major similarity is that the performance of off-axis and on-axis designs is nearly identical for MDE and MD–PRE. The only difference is in some values of SPR_m for MD–PRE. This is probably because in the off-axis design some of the MDE-type sidelobes lie outside the reduced-bandwidth region (which is the same area as was used for evaluating the on-axis designs).

These results point out one advantage of MD–PRE over ED–PRE: MD–PRE is less sensitive to where the desired pattern is centered. Certainly, other error diffusion kernels could be designed to center the noise reconstruction around the desired reconstruction. However this would further complicate the design process. Also, if the desired pattern is distributed over the full NRB, there may be no practical way to select ED weighting coefficients that spatially separate noise from the desired pattern. ED–PRE may provide little improvement in these cases.

For the sensitivity analyses (Fig. 9) it is found that NU and SPR for ED–PRE are both somewhat more sensitive to χ for the off-axis designs (Fig. 9) than for the on-axis designs (Fig. 6). In the statistical comparisons of ED–PRE with MD–PRE for the off-axis design, Fig. 10 shows that MD–PRE (for $\gamma = 1.4$) will almost always have a higher SPR than ED–PRE, while NU for MD–PRE is, at worst, only slightly higher than for ED–PRE. Slight differences between MD–PRE for on-axis and off-axis cases occurred because the ensemble of 21 random sequences used for calculating the curves in Fig. 7 was different from that used in Fig. 10.

5. CONCLUSIONS

We have introduced a new type of single-pixel encoding algorithm ED–PRE and have compared its performance

with that of existing algorithms. The design criteria of principle concern, NU and SPR, emphasize fidelity rather than energy efficiency. Low SPR is sought for the entire NRB so that successive designs can address and utilize the entire SBWP available to the SLM. Depending on the specific test function encoded, the blended ED–PRE algorithm performs nearly as well as and sometimes better than the MD–PRE, and these blended algorithms substantially outperform the nonblended MDE, PRE, and ED algorithms. Even when the test function is centered and the performance metrics are calculated over a reduced-bandwidth window (which was the original intended application of the ED algorithm used here), the ED–PRE algorithm produces better performance than ED alone. Apparently ED–PRE uses the properties of ED to filter the background noise and distribute it nonuniformly over the diffraction plane, and it uses the properties of PRE to diffuse and reduce the peak intensity of the ED noise spikes.

In this paper we have also delineated the differences between the properties of ED and PRE. As a final distinction we note that both methods use diffusion, but it is used in different ways. In PRE the encoding error is scattered or diffused over the entire diffraction plane, forming a laser speckle pattern. In ED the encoding errors from one pixel are diffused forward into neighboring pixels in the modulation plane. The resulting noise background for ED is spiky rather than diffuse or speckled. One might choose to model the error sequence from ED as a stochastic process.[20] This is valid only to the extent that the sequence is described as a random process. However, for PRE the error sequence is a random process to the extent that a random-number generator represents a random process.

Although we have presented encoding algorithms specifically designed for continuous phase-only SLM's, it should be clear that ED–PRE can also be developed both for quantized and for coupled modulation characteristics by blending the approaches presented in Refs. 5, 8, 17. We would expect the background noise to have an appearance and the performance metrics to show dependencies on the free parameters γ and χ similar to those observed for phase-only modulation. However, it would be interesting to find out whether the relative performance differences between ED, ED–PRE, and MD–PRE are maintained for different modulation characteristics. We have no reason to suspect that this either might or might not be the case.

In summary, ED–PRE can improve on the fidelity of ED alone. We suspect that ED–PRE (as illustrated by the two examples presented here) is competitive with MD–PRE and that in some cases it may even outperform MD–PRE.

ACKNOWLEDGMENTS

This study was supported by U.S. Office of Naval Research contract N00014-96-1-1296 and NASA cooperative agreement NCC5-222.

*When the work was performed the authors were with University of Louisville. M. Duelli's current address is Optical Coating Laboratory, Inc., 2789 Northpoint Parkway, MS 125-3, Santa Rosa, California 95407-7397.

REFERENCES AND NOTES

1. R. W. Cohn and M. Liang, "Approximating fully complex spatial modulation with pseudorandom phase-only modulation," Appl. Opt. **33**, 4406–4415 (1994).
2. L. G. Hassebrook, M. E. Lhamon, R. C. Daley, R. W. Cohn, and M. Liang, "Random phase encoding of composite fully complex filters," Opt. Lett. **21**, 272–274 (1996).
3. R. W. Cohn and W. Liu, "Pseudorandom encoding of fully complex modulation to bi-amplitude phase modulators," in *Diffractive Optics and Micro-Optics*, Vol. 5 of 1996 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1996), pp. 237–240.
4. R. W. Cohn and M. Liang, "Pseudorandom phase-only encoding of real-time spatial light modulators," Appl. Opt. **35**, 2488–2497 (1996).
5. R. W. Cohn, "Pseudorandom encoding of complex valued functions onto amplitude coupled phase modulators," J. Opt. Soc. Am. A **15**, 868–883 (1998).
6. R. W. Cohn and M. Duelli, "Ternary pseudorandom encoding of Fourier transform holograms," J. Opt. Soc. Am. A **16**, 71–84 (1999); Errata, 1089–1090 (1999).
7. R. W. Cohn, "Analyzing the encoding range of amplitude-phase coupled spatial light modulators," Opt. Eng. **38**, 361–367 (1999).
8. M. Duelli, M. Reece, and R. W. Cohn, "A modified minimum-distance criterion for blended random and nonrandom encoding," J. Opt. Soc. Am. A **16**, 2425–2438 (1999).
9. R. W. Cohn and L. G. Hassebrook, "Representations of fully complex functions on real-time spatial light modulators," in *Optical Information Processing*, F. T. S. Yu and S. Jutamulia, eds. (Cambridge U. Press, Cambridge, UK, 1998), Chap. 15, pp. 396–432.
10. B. R. Brown and A. W. Lohmann, "Complex spatial filter," Appl. Opt. **5**, 967–969 (1966).
11. W. J. Dallas, "Computer-generated holograms," in *The Computer in Optical Research*, B. R. Frieden, ed. (Springer-Verlag, Berlin, 1980), Chap. 6, pp. 291–366.
12. R. D. Juday, "Optimal realizable filters and the minimum Euclidean distance principle," Appl. Opt. **32**, 5100–5111 (1993).
13. L. B. Lesem, P. M. Hirsch, and J. A. Jordon, Jr., "The kinoform: a new wavefront reconstruction device," IBM J. Res. Dev. **13**, 150–155 (1969).
14. J. L. Horner and P. D. Gianino, "Phase-only matched filtering," Appl. Opt. **23**, 812–816 (1984).
15. R. W. Floyd and L. Steinberg, "An adaptive algorithm for spatial grayscale," Proc. Soc. Inf. Disp. **17**, 78–84 (1976).
16. R. Hauk and O. Bryngdahl, "Computer-generated holograms with pulse density modulation," J. Opt. Soc. Am. A **1**, 5–10 (1984).
17. S. Weissbach, F. Wyrowski, and O. Bryngdahl, "Digital phase holograms: coding and quantization with an error diffusion concept," Opt. Commun. **72**, 37–41 (1989).
18. S. Weissbach and F. Wyrowski, "Error diffusion procedure: theory and applications in optical signal processing," Appl. Opt. **31**, 2518–2534 (1992).
19. A. G. Kirk, A. K. Powell, and T. J. Hall, "The design of quasi-periodic Fourier plane array generators," in *Optical Information Technology*, S. D. Smith and R. F. Neale, eds. (Springer-Verlag, Berlin, 1993), pp. 47–56.
20. E. Barnard, "Optimal error diffusion for computer-generated holograms," J. Opt. Soc. Am. A **5**, 1803–1811 (1988).
21. M. Duelli and R. W. Cohn, "Pseudorandom encoding for real-valued ternary spatial light modulators," Appl. Opt. **38**, 3804–3809 (1999).
22. F. Wyrowski, "Upper bound of the diffraction efficiency of

diffractive phase elements," Opt. Lett. **16**, 1915–1917 (1991).
23. J. A. Davis and D. M. Cottrell, "Random mask encoding of multiplexed phase-only and binary phase-only filters," Opt. Lett. **19**, 496–498 (1994).
24. J. M. Goodman, "Effects of film nonlinearities," in *Introduction to Fourier Optics* (McGraw-Hill, San Francisco, Calif., 1968), Sec. 8-6, pp. 230–241.
25. To maintain as much consistency as possible in comparing all the curves and tables and to avoid excessive computation, we have calculated and reported all performance metrics for values of $\gamma = 1, 1.1, 1.2, \ldots$. This results in adequately smooth and sampled curves except in one case. For the SPR_m curve of ED–PRE in Fig. 5, finer sampling led to a significant increase in SPR_m , from 53 at $\gamma = 1.2$ and 1.3 to 60 at $\gamma = 1.26$. This additional point is included in the plot in Fig. 5. We also checked the maxima of other SPR and SPR_m performance curves, using finer sampling increments. However, since the change in appearance is minimal and the maximum values of the curves would change by no more than a few tenths, these additional findings are omitted.