

One-dimensional beam shaping

Boaz Salik, Joseph Rosen, and Amnon Yariv

California Institute of Technology, Pasadena, California 91125

Received November 21, 1994; revised manuscript received February 21, 1995; accepted February 22, 1995

A general method is presented of optimizing one-dimensional holograms to realize an arbitrary diffracted field distribution over the propagation volume. A computationally simpler approach is developed for the case in which only the transverse distribution in one plane and the longitudinal distribution on one axis are of interest. Scaling effects in the optimized beams are studied and compared with those of canonical beams.

1. INTRODUCTION

Computer-generated holograms have been employed to construct desired transverse¹ and longitudinal² two-dimensional images. In particular, we have recently solved the problem of shaping the axial intensity distribution of a circularly symmetric beam.² For certain applications, e.g., photolithography³ and time-domain pulse shaping,⁴ the transverse image obeys one-dimensional diffraction laws, which cannot be treated as a special case of two-dimensional diffraction. In another paper⁵ we presented an analytical solution to the pseudo-nondiffracting (PND) one-dimensional beam. Here we consider the broader problem of diffraction in only one transverse coordinate (x) and present a general numeric method for controlling the field (amplitude and phase) over a finite spatial domain. We then develop a simplified approach to the specific problem of shaping a beam's axial intensity as an arbitrary pattern over a finite propagation distance and demonstrate it experimentally. We also consider scaling effects in the optimized PND beam.

2. BACKGROUND

Given a one-dimensional complex field distribution $g(x, z = 0) = g(x_1)$, the field at any subsequent point, in the Fresnel approximation,⁶ is

$$u(x, z) = \frac{\exp(ikz)}{\sqrt{i\lambda z}} \int_{-\infty}^{\infty} g(x_1) \exp\left[\frac{ik}{2z}(x - x_1)^2\right] dx_1, \quad (1)$$

where λ is the wavelength and $k = 2\pi/\lambda$ is the wave number.

Now we wish to impose certain conditions on $u(x, z)$ for some subspace $R_0 \subseteq R$, $R = \{(x, z) | z > 0\}$. In general, we can represent these conditions as an error function E to be minimized over R_0 :

$$E[g(x_1)] = \iint_{(x,z) \in R_0} e[u(x, z)] dx dz, \quad (2)$$

where e is the error density at point (x, z) .

For example, if we want the intensity $|u(x, z)|^2$ to come close to some $I_0(x, z)$ for $(x, z) \in R_0$, we might employ a Euclidean distance-square error function

$$\begin{aligned} E[g(x_1)] &= \iint_{(x,z) \in R_0} [I_0(x, z) - |u(x, z)|^2]^2 dx dz \\ &= \iint_{(x,z) \in R_0} \left\{ I_0(x, z) - \frac{1}{\lambda z} \left| \int_{-\infty}^{\infty} g(x_1) \right. \right. \\ &\quad \left. \left. \times \exp\left[\frac{ik}{2z}(x - x_1)^2\right] dx_1 \right|^2 \right\} dx dz. \quad (3) \end{aligned}$$

In general, then, the only way to find the $g(x_1)$ that minimizes E is to try all possible complex functions—usually an unattractive option. If, however, we write $g(x_1) = A(x_1)\exp[i\phi(x_1)]$, where A and ϕ are functions in a one-dimensional function space, so that $E[g(x_1)] = E(A, \phi)$, then, assuming that E is smooth in A and ϕ , we have

$$dE = \nabla_A E \cdot dA + \nabla_\phi E \cdot d\phi, \quad (4)$$

where $\nabla_A E$ and $\nabla_\phi E$ are energy gradients with respect to amplitude and phase, respectively. Hence we see that sufficient (though not necessary) conditions for $dE = 0$ are that $\nabla_A E = 0$ and $\nabla_\phi E = 0$. Now we can optimize A and ϕ independently using any convenient algorithm; although, in general, it may take more than one iteration of each to attain a local minimum for both, this is still faster in most cases than treating the full $(N \times N)$ -dimensional problem, which is computationally N^2 as complex as the N -dimensional problem (N being the length of vectors A and ϕ after quantization). Note that the problem of two-dimensional hologram design is also of current interest.⁷ Our method can also be used to optimize two-dimensional masks, realizing arbitrary three-dimensional diffraction patterns, although, in practice, separating the mask into amplitude and phase does not sufficiently shorten the computation time.

3. LONGITUDINAL BEAM SHAPING

Next we turn to the specific problem of shaping the intensity distribution along the propagation axis, $I_1(z)$, while maintaining some transverse constraint, $|u(x, z = z_0)|^2 \in S_x$. Here S_x is the space of allowed transverse intensities. Following Rosen and Yariv,² we assume for computational convenience (and without loss of generality) that

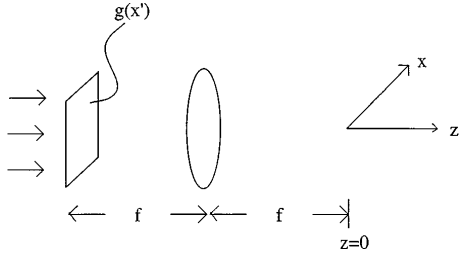


Fig. 1. System for realizing one transverse and one axial constraint.

$z_0 = 0$ and there is a lens with focal length f at $z = -f$ (Fig. 1). If we place our transparency $g(x')$ at $z = -2f$, the one-dimensional Fresnel approximation yields⁸

$$u(x, z) = \frac{\exp[ik(z + 2f)]}{\sqrt{i\lambda f}} \int_{-\infty}^{\infty} g(x') \times \exp\left[-\frac{ik(zx'^2 + 2fxx')}{2f^2}\right] dx', \quad (5)$$

and there is a convenient Fourier transform relationship between $g(x')$ and $u_t(x) = u(x, z = 0)$, easing the computational burden of the transverse constraint. Unfortunately, the axial profile is not as easy to compute. Setting $x = 0$ in Eq. (5), we have

$$u_1(z) = u(x = 0, z) = \frac{\exp[ik(z + 2f)]}{\sqrt{i\lambda f}} \int_{-\infty}^{\infty} g(x') \exp\left(-\frac{ikzx'^2}{2f^2}\right) dx'. \quad (6)$$

Although this can be treated as a Fourier transform in the variable x'^2 , doing so introduces significant quantization error, which is avoided if we integrate directly. Now we can express $g(x')$ in terms of $u_t(x, z = 0)$ because of the Fourier transform relation:

$$g(x') = \sqrt{\frac{i}{\lambda f}} \exp(-2ikf) \int_{-\infty}^{\infty} u_t(x) \exp\left(\frac{ikxx'}{f}\right) dx. \quad (7)$$

Thus we can define a transformation between the x - and z -axis distributions behind the lens:

$$T[u_t(x)] = u_1(z) = \frac{\exp(ikz)}{\lambda f} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_t(x) \times \exp\left[\frac{ik(2fxx' - zx'^2)}{2f^2}\right] dx dx'. \quad (8)$$

Now we can bypass any mask-plane functions by optimizing $u_t(x)$ subject to the constraint $u_t(x) \in S_x$ to yield $|u_1(z)|^2$ as close as possible to our desired $I_1(z)$. To optimize $u_t(x)$, we can still decompose it into amplitude and phase, as in Eq. (4). After arriving at the optimal $u_t(x)$, we simply use Eq. (7) to obtain our mask, $g(x')$.

In practice, our optical system has a finite aperture, placing limits on the validity of the Fresnel approximation. To quantify this, we add the third term of the binomial expansion in the expression for the longitudinal field distribution behind the lens⁶:

$$u_1(z) = u(0, z) = \frac{\exp[ik(z + 2f)]}{i\lambda f} \int_{-\infty}^{\infty} g(\tilde{x}) \times \exp\left\{-ik\left[\frac{\tilde{x}^2}{2f} - \frac{\tilde{x}^2}{2(z+f)} + \frac{\tilde{x}^4}{8(z+f)^3}\right]\right\} d\tilde{x}, \quad (9)$$

where $g(\tilde{x})$ is the field before the lens. There are two points of stationary phase in Eq. (9): $\tilde{x}_a = 0$ and $\tilde{x}_b = (z+f)\sqrt{-2z/f}$, where the root \tilde{x}_b is introduced by the fourth-order phase term. In the optical regime ($k \gg 1$), in order to ignore the contribution of this root in the integral, we must satisfy the condition $(z+f)\sqrt{-2z/f} > D/2$, where D is the diameter of the lens. From this condition we conclude that the Fresnel approximation is valid for $z > -f + D/\sqrt{8}$. In the conventional optics that we use (where $f \gg D$) our beams do not exist in the region $z < -f + D/\sqrt{8}$, and the Fresnel approximation holds.

4. REALIZATION

In order to realize the complex holograms as real, positive masks, we used a standard off-axis implementation: Define the function

$$g'(x') = 1 + \text{Re}\left[g(x') \exp\left(i2\pi \frac{\sin \theta}{\lambda} x'\right)\right], \quad (10)$$

which is real and positive if $g(x')$ is properly normalized. The exponential shifts the z axis of our desired pattern a transverse distance of $f \sin \theta$ after the lens. To see this, we substitute $g(x') \exp\{i2\pi[(\sin \theta)/\lambda]x'\}$ into Eq. (5):

$$u(x, z) = \frac{\exp[ik(z + 2f)]}{\sqrt{i\lambda f}} \int_{-\infty}^{\infty} g(x') \exp[ik \sin(\theta)x'] \times \exp\left[-\frac{ik(zx'^2 + 2fxx')}{2f^2}\right] dx' = \frac{\exp[ik(z + 2f)]}{\sqrt{i\lambda f}} \int_{-\infty}^{\infty} g(x') \times \exp\left[-\frac{ik(zx'^2 + 2f\tilde{x}x')}{2f^2}\right] dx', \quad (11)$$

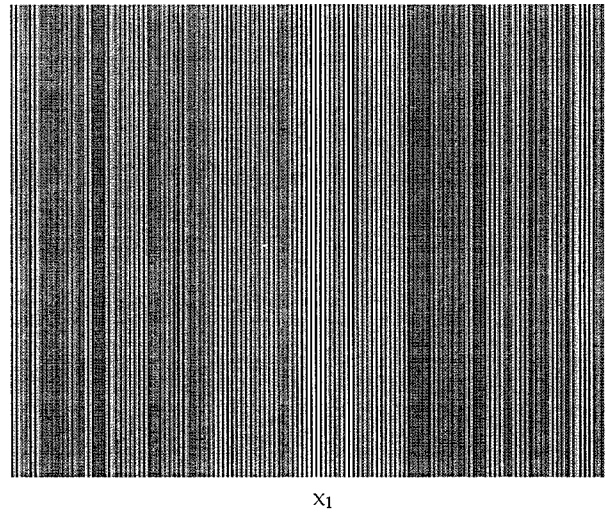


Fig. 2. Mask that generated the beam in Fig. 3(d) below. The mask was fabricated on tone-developed film for a cylindrical lens with focal length $f = 30$ cm, a HeNe laser at $\lambda = 633$ nm, and spatial frequency 206 cm^{-1} ($\theta = 13$ mrad, mask width of 1.24 cm).

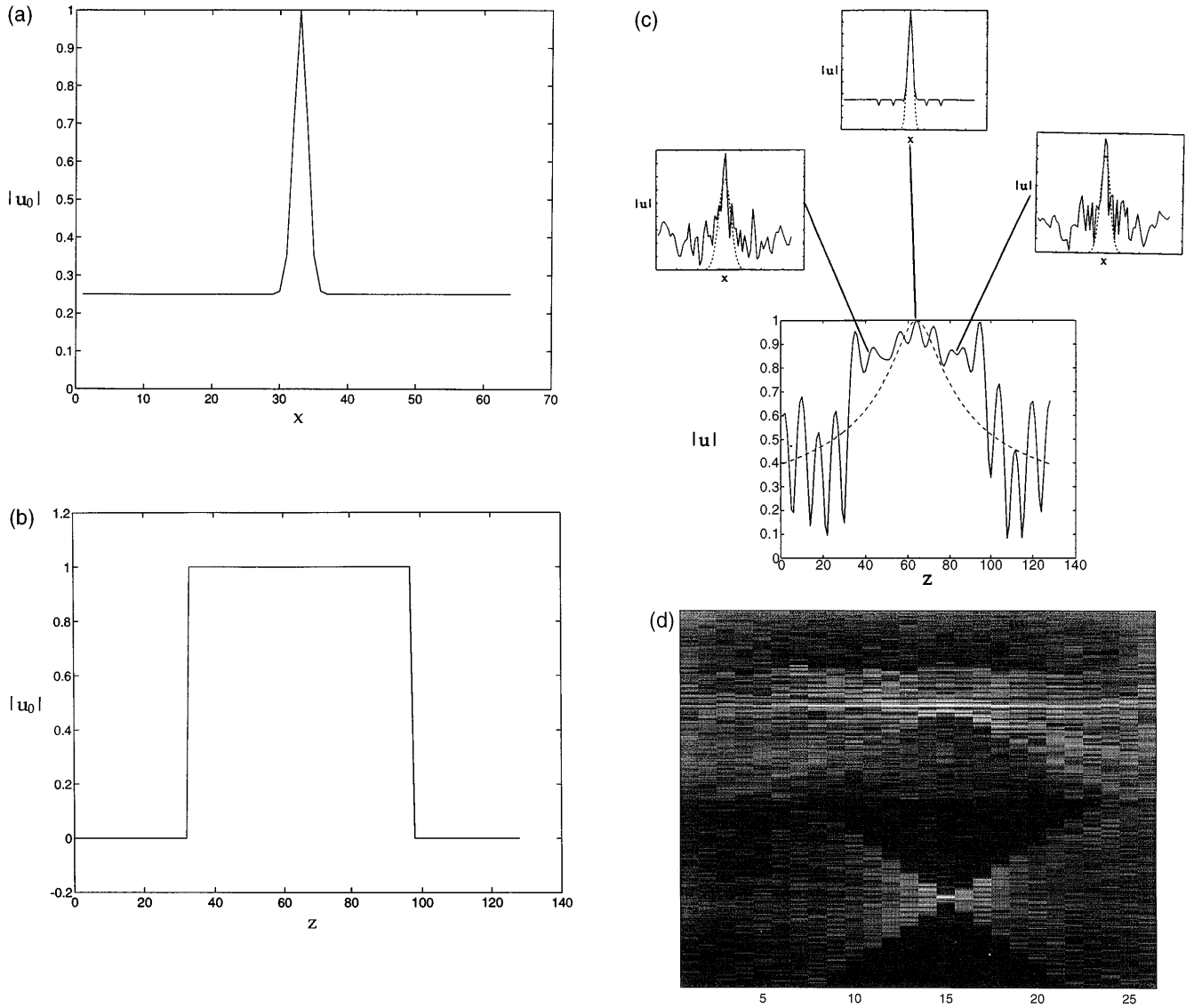


Fig. 3. (a) Transverse constraint for PND beam, (b) axial constraint, (c) simulated PND beam axial intensity and three transverse profiles, and (d) experimentally obtained PND beam (top) compared with an ordinary focused beam over $\Delta z = 15$ cm. The mask used, along with relevant parameters, is shown in Fig. 2.

where $\tilde{x} = x - f \sin \theta$.

Now taking the real part adds the inverted pattern around $(x, z) = (-f \sin \theta, 0)$, and the dc term simply contributes a focused spot at the origin, which does not affect our shifted pattern if f and θ are large enough. A typical mask is shown in Fig. 2.

As examples, we demonstrate this method for two cases. The first is a one-dimensional PND beam, where we forced $u_t(x)$ to assume a transverse Gaussian pulse width of $\sigma = 1_{\text{pixel}}$ and background level $\leq 25\%$ of the peak amplitude [Fig. 3(a)]. $I_1(z)$ was simply set to a constant over $\Delta z = 65_{\text{pixels}}$, with $2f = 128_{\text{pixels}}$ as shown in Fig. 3(b); outside this region there were no constraints on $I_1(z)$. This is a natural application, requiring only one axial and one transverse constraint. We used the error function

$$E[u_t(x)] = \int_{z \in \Delta z} [|u_1(z)| - \sqrt{I_1(z)}]^2 dz, \quad (12)$$

where $u_1(z)$ is given by Eq. (8). This is the Euclidean

distance square in field amplitude, and a simple gradient descent⁹ was used to alternately optimize $|u_t(x)|$ and Phase $[u_t(x)]$ (the algorithm always converged within eight iterations for vectors of 64 pixels). Figure 3(c) shows the simulated field amplitude along the z axis (solid curves), compared with a Gaussian of the same width (dashed curves), and three transverse cross sections along the beam. Figure 3(d) shows the two intensities experimentally, over the portion of the x - z plane close to the focus ($\Delta z = 15$ cm, $f = 30$ cm).

The second example is a sequence of two PND beams separated by a dark region [Fig. 4(a)]. Since this requires three transverse constraints, one for each beam and one for the dark region, we optimized using integral (5) over the region of interest. The error function used was again Euclidean in amplitude:

$$E[g(x_1)] = \iint_{(x,z) \in R_0} [\sqrt{I_0(x,z)} - |u(x,z)|]^2 dx dz, \quad (13)$$

and again we used a gradient descent to optimize $|g(x_1)|$ and Phase $[g(x_1)]$. The simulated and actual beam intensities over the $x-z$ plane are shown in Figs. 4(b) and 4(c), respectively, over a distance $\Delta z = 8$ cm.

5. SCALING

Finally, we tested the dependence of the minimum z -axis error in optimized PND beams on the size of the z constraint, Δz , and the allowed pulse width σ . In all finite-power canonical beams the region of near-constant

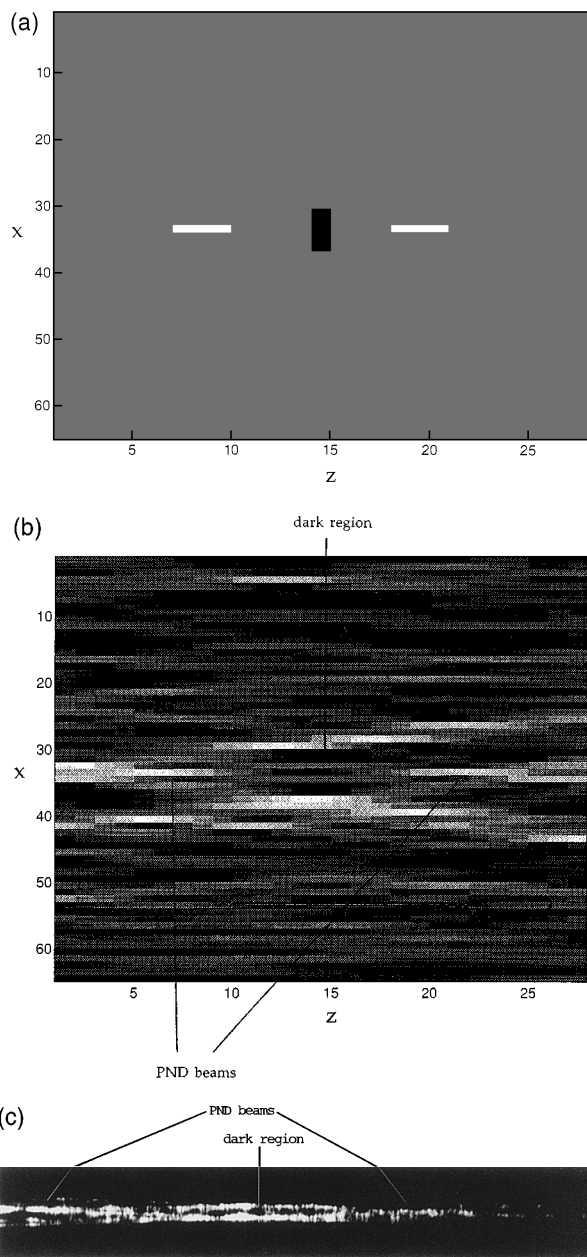


Fig. 4. (a) Two-dimensional constraint for the second example discussed in Section 4; white denotes the field maximum, black denotes the field minimum, and gray denotes unconstrained areas. (b) Simulated intensity distribution. (c) Observed intensity distribution over $\Delta z = 8$ cm. The lens has $f = 30$ cm, $\lambda = 633$ nm, and mask spatial frequency of 206 cm^{-1} (mask width of 1.24 cm).

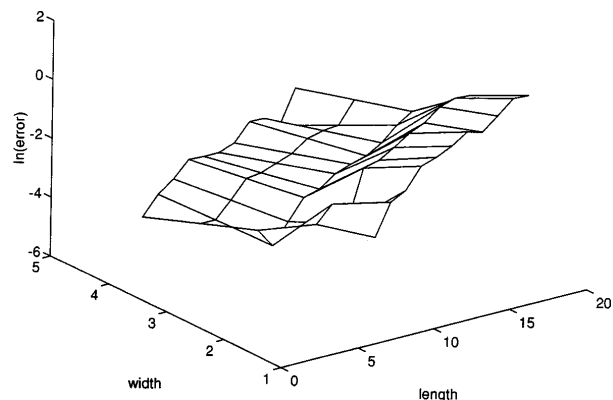


Fig. 5. Log(error) versus focal depth and beam width of optimized PND beam.

amplitude, Δz , increases with the beam width (e.g., for Gaussian beams we have $z_0 = n\pi\sigma^2/\lambda$). An appropriate analogy here is the region Δz over which constant intensity can be attained within a given error. To this end we optimized PND beams with varying Δz and plotted these parameters versus the minimum error attained (Fig. 5). Although there are deviations, the clear trend is an increase in error with larger Δz and smaller σ . Therefore, for the maintenance of constant error, any increase in focal depth (Δz) requires an offsetting increase in the beam width σ . As expected, optimization allows us to improve depth of focus but does not affect its dependence on fundamental physical parameters.

6. CONCLUSION

We have presented a general method for optimizing holograms to realize an arbitrary field distribution over the propagation region and developed a simplified method for shaping the axial field of one-dimensional holograms subject to one transverse constraint. We demonstrated the two methods experimentally and tested the error limitations of our algorithm as a function of the transverse and longitudinal constraint sizes, concluding that the optimized PND beams scale in the same general way as finite-power canonical beams.

An important application of this approach is controlling the behavior of temporal pulses in a dispersive medium. The analogy between temporal pulse propagation and one-dimensional Fresnel diffraction is treated in Ref. 4, and in Ref. 5 we present an analytical solution to the specific problem of extending the distance over which a temporal pulse remains narrow (its depth of time focusing). Here we have solved the general problem of controlling the exact shape of a temporal pulse over some propagation distance. Additionally, our optimization procedure for extending depth of focus generally yields longer focal depths than those of analytical beams given the same hologram resolution. For example, in Fig. 3 we demonstrate a tenfold increase in the depth of focus of a one-dimensional beam, whereas the analytical solution in Ref. 5 increases depth of focus by a factor of 2. This advantage of optimized beams is due to the fact that analytical beams do not, in general, lie at an error minimum.

ACKNOWLEDGMENT

This material is based on work supported under a National Science Foundation Graduate Research Fellowship and grants from the Advanced Research Projects Agency and the U.S. Army Research Office.

REFERENCES

1. W. H. Lee, in *Progress in Optics*, E. Wolf, ed. (North-Holland, Amsterdam, 1978), Vol. XVI, pp. 119–173.
2. J. Rosen and A. Yariv, “Synthesis of an arbitrary axial field profile by computer-generated holograms,” *Opt. Lett.* **19**, 843–845 (1994).
3. M. D. Levenson, “Phase-shifting mask strategies: isolated bright contacts,” *Microlithography World* (September–October 1992).
4. E. B. Treacy, “Optical pulse compression with diffraction gratings,” *IEEE J. Quantum Electron.* **QE-5**, 454–458 (1969).
5. J. Rosen, B. Salik, A. Yariv, and H. K. Liu, “Pseudo non-diffracting slitlike beam and its analogy to the pseudo nondispersing pulse,” *Opt. Lett.* **20**, 423–425 (1995).
6. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968), pp. 59–61.
7. R. Piestun, B. Spector, and J. Shamir, “Control of wavefront propagation with diffractive elements,” *Opt. Lett.* **19**, 771–773 (1994).
8. C. W. McCutchen, “Generalized aperture and the three-dimensional diffraction image,” *J. Opt. Soc. Am.* **54**, 240–244 (1964).
9. L. R. Foulds, *Optimization Techniques* (Springer-Verlag, New York, 1981), pp. 329–335.
10. A. Yariv, *Optical Electronics* (Saunders, Philadelphia, 1991), pp. 48–49.