Sampling of the diffraction field

Levent Onural

When optical signals, like diffraction patterns, are processed by digital means the choice of sampling density and geometry is important during analog-to-digital conversion. Continuous band-limited signals can be sampled and recovered from their samples in accord with the Nyquist sampling criteria. The specific form of the convolution kernel that describes the Fresnel diffraction allows another, alternative, full-reconstruction procedure of an object from the samples of its diffraction pattern when the object is space limited. This alternative procedure is applicable and yields full reconstruction even when the diffraction pattern is undersampled and the Nyquist criteria are severely violated. Application of the new procedure to practical diffraction-related phenomena, like in-line holography, improves the processing efficiency without creating any associated artifacts on the reconstructed-object pattern. © 2000 Optical Society of America

OCIS codes: 050.1940, 070.6020, 090.1760, 100.2000.

1. Introduction

Sampling of continuous signals, effects of undersampling, and related issues are very well understood, and the related signal-processing literature is abundant (see, for example, Ref. 1). Optical wave propagation, diffraction, and holography are also well-known and well-documented physical phenomena (see, for example, Ref. 2). Furthermore, there is an increasing tendency to apply, or combine, digital signal-processing techniques with optics and other wave-propagation-related fields.[3–5] However, because of the specific form of the convolution kernel that represents scalar wave propagation, the sampling of the Fresnel diffraction field and the reconstruction from those samples seem to be confusing; clarification of this fundamental issue is essential for digital processing of diffraction-based phenomena. The purpose of this paper is to provide that clarification.

What makes scalar wave propagation special is the quadratic-phase function

$$h_z(x, y) = \frac{1}{j\lambda z} \exp\left(j \frac{2\pi}{\lambda} z\right) \exp\left[j \frac{\pi}{\lambda z} (x^2 + y^2)\right], \quad (1)$$

The author (onural@bilkent.edu.tr; l.onural@ieee.org) is with the Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, TR-06533 Ankara, Turkey.

which has interesting properties. The Fresnel diffraction field, under coherent illumination and at a distance z that is due to a two-dimensional (2-D) object $f(x, y)$, is given by

$$\psi_z(x, y) = f(x, y) ** h_z(x, y), \quad (2)$$

where λ is the wavelength and the double asterisks denote 2-D convolution. For notational simplicity 2-D variables (x, y) are denoted as a vector: $\mathbf{x} = [x\ y]^\mathrm{T}$.

It is clear that, if $f(\mathbf{x})$ is band limited, $\psi_z(\mathbf{x})$ is also band limited and has the same band. Therefore both the diffraction field and the object can be recovered fully from the object's samples by use of the sinc interpolation (low-pass filtering) if the sampling rate is higher than the Nyquist rate. However, as a result of some unusual properties of the kernel, there is another simpler and potentially more useful reconstruction condition and procedure.

2. Sampling of the Diffraction Field

The Fourier transform of $h_z(x, y)$ can be found analytically as

$$H_z(u, v) = H_z(\mathbf{u})$$
$$= \exp\left(j \frac{2\pi}{\lambda} z\right) \exp\left[-j \frac{\lambda z}{4\pi} (u^2 + v^2)\right]$$
$$= \exp\left(j \frac{2\pi}{\lambda} z\right) \exp\left(-j \frac{\lambda z}{4\pi} \mathbf{u}^\mathrm{T}\mathbf{u}\right), \quad (3)$$

where $\mathbf{u} = [u \ v]^{\mathrm{T}}$, as usual. Therefore the Fourier transform of the diffraction field, if we assume a band-limited object $f(x, y)$, is

$$\Psi_z(u, v) =$$
$$\begin{cases} F(u, v)\exp\left(j\dfrac{2\pi}{\lambda}z\right)\exp\left[-j\dfrac{\lambda z}{4\pi}(u^2 + v^2)\right] & \text{if } (u, v) \in \mathbf{B}, \\ 0 & \text{else} \end{cases}$$
$$(4)$$

where $\mathbf{B}$ is the band occupied by $f(x, y)$. Let us denote the sampled version of $\psi_z(x, y)$ as $\psi_{zs}(\mathbf{x}) = \psi_{zs}(x, y)$ for which the periodic-sampling geometry is indicated by the sampling matrix $\mathbf{V}$ as

$$\psi_{zs}(\mathbf{x}) = \sum_{\mathbf{n}} \psi_z(\mathbf{Vn})\delta(\mathbf{x} - \mathbf{Vn}), \qquad (5)$$

where $\mathbf{n} = [n_1 \ n_2]^{\mathrm{T}}$ and n_1 and n_2 are integers. Therefore the Fourier transform of ψ_{zs} can be written as

$$\Psi_{zs}(\mathbf{u}) = \mathscr{F}\{\psi_{zs}(\mathbf{x})\}$$
$$= \frac{1}{|\det \mathbf{V}|} \sum_{\mathbf{k}} \Psi_z(\mathbf{u} - \mathbf{Uk}), \qquad (6)$$

where $\mathbf{U} = 2\pi\mathbf{V}^{-\mathrm{T}}$, $\mathbf{k} = [k_1 \ k_2]^{\mathrm{T}}$, and k_1, and k_2 are integers. As expected, the sampling of $\psi_z(\mathbf{x})$ by the matrix $\mathbf{V}$ results in superposed, shifted replicas of the original Fourier transform in which the replicas are located at the $\mathbf{Uk}$.

The reconstruction of $f(x, y)$ from its diffraction $\psi_z(x, y)$ is simple: Just convolve $\psi_z(x, y)$ by $h_{-z}(x, y)$, which represents reverse propagation. In other words, in the Fourier domain

$$\Psi_z(u, v)H_{-z}(u, v) = F(u, v)H_z(u, v)H_{-z}(u, v)$$
$$= F(u, v) \qquad (7)$$

because $H_z(u, v)H_{-z}(u, v) = 1$.

It is interesting to see the result of the same backpropagation as applied to the sampled diffraction. In the Fourier domain, we have

$$\Psi_{zs}(\mathbf{u})H_{-z}(\mathbf{u}) = \frac{1}{|\det \mathbf{V}|} \sum_{\mathbf{k}} \Psi_z(\mathbf{u} - \mathbf{Uk})H_{-z}(\mathbf{u})$$
$$= \frac{1}{|\det \mathbf{V}|} \sum_{\mathbf{k}} F(\mathbf{u} - \mathbf{Uk})H_z(\mathbf{u} - \mathbf{Uk})H_{-z}(\mathbf{u}), \quad (8)$$

but

$$H_z(\mathbf{u} - \mathbf{Uk})H_{-z}(\mathbf{u}) = \exp\left[j\frac{\lambda z}{4\pi}(2\mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{u} - \mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{Uk})\right]. \qquad (9)$$

Therefore from Eqs. (8) and (9), we have

$$\Psi_{zs}(\mathbf{u})H_{-z}(\mathbf{u}) = \frac{1}{|\det \mathbf{V}|} \sum_{\mathbf{k}} F(\mathbf{u} - \mathbf{Uk})$$
$$\times \exp\left(-j\frac{\lambda z}{4\pi}\mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{Uk}\right)\exp\left(j\frac{\lambda z}{2\pi}\mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{u}\right). \quad (10)$$

Taking the inverse Fourier transform to find the result of backpropagation, we find

$$\mathscr{F}^{-1}\{\Psi_{zs}(\mathbf{u})H_{-z}(\mathbf{u})\}$$
$$= \sum_{\mathbf{k}} c_{\mathbf{k}}f\left(\mathbf{x} - \frac{\lambda z}{2\pi}\mathbf{Uk}\right)\exp(j\mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{x}), \quad (11)$$

where $c_{\mathbf{k}}$ is the complex constant

$$[1/|\det \mathbf{V}|]\exp[-j(\lambda z/4\pi)\mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{Uk}].$$

Therefore, as a consequence of the sampling of the diffraction by the sampling matrix $\mathbf{V}$, we get superposed, shifted replicas of the original object function $f(\mathbf{x})$ for which the replicas are the envelopes of complex sinusoids. For each $\mathbf{k}$, the location is $(\lambda z/2\pi)\mathbf{Uk}$, and the modulation carrier is given by $\exp(j\mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{x})$.

3. Recovery from Samples

After the observations given in Section 2, we can state that the recoverability conditions of a continuous object $f(\mathbf{x})$ from the samples of its Fresnel diffraction pattern by a sampling matrix $\mathbf{V}$ are

(1) If $f(\mathbf{x})$ is band limited, say, within a band $\mathbf{u} \in \mathbf{B}$ and if the matrix $\mathbf{V}$ is chosen to satisfy nonoverlapping replicas in the Fourier domain [no aliasing, see Eq. (6)], the $f(\mathbf{x})$ can be fully recovered by low-pass-filtering (the band is $\mathbf{B}$) the result of the backpropagated sampled (discrete) Fresnel diffraction field. This process is simply the 2-D Nyquist sampling and reconstruction case. If $f(\mathbf{x})$ is recovered its continuous diffraction field is also known and vice versa. Mathematically, we have

$$f(\mathbf{x}) = \mathscr{F}^{-1}\{\Psi_{zs}(\mathbf{u})H_{-z}(\mathbf{u})H_{\mathrm{LP}}(\mathbf{u})\}, \qquad (12)$$

where

$$H_{\mathrm{LP}}(\mathbf{u}) = \begin{cases} |\det \mathbf{V}| & \text{if } \mathbf{u} \in \mathbf{B} \\ 0 & \text{else} \end{cases} \qquad (13)$$

and the subscript LP denotes the low-pass filtering.

(2) If $f(\mathbf{x})$ is space limited, say, within a region $\mathbf{x} \in \mathbf{R}$ and if the matrix $\mathbf{V}$ is chosen to satisfy nonoverlapping replicas of $f(\mathbf{x})$ after backpropagating the sampled (discrete) Fresnel diffraction field [see Eq. (11)], $f(\mathbf{x})$ can be recovered fully by the windowing of the results of backpropagation

$$f(\mathbf{x}) = W_{\mathbf{R}}(\mathbf{x})\left[\sum_{\mathbf{k}} c_{\mathbf{k}} f\left(\mathbf{x} - \frac{\lambda z}{2\pi}\mathbf{U}\mathbf{k}\right)\exp(j\mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{x})\right],$$

$$(14)$$

where

$$W_{\mathbf{R}}(\mathbf{x}) = \begin{cases} \dfrac{1}{c_{\mathbf{k}}} & \text{if } \mathbf{x} \in \mathbf{R} \\ 0 & \text{else} \end{cases} \qquad (15)$$

and $\mathbf{R}$ is the window support.

Note that the above two requirements are not the same; indeed, they are alternative requirements because a band-limited signal cannot be space limited at the same time. Also note that either condition 1 or condition 2 is sufficient for complete recovery of the image. Indeed, this is the reason that makes this sampling–recovery problem an interesting one. In straightforward digital signal-processing applications it is quite common to perform the sampling–recovery in accord with the requirement 1, which is a general and a very-well-known result for any kind of convolution kernel. However, the rather obscure second result might have a wider application in typical optical environments and could provide great savings during digital signal processing because it is quite common to have rather small objects in typical applications. A small object can still be recovered fully even if the sampling strategy significantly violates the first condition (i.e., from severe aliasing) because sampling can still easily satisfy the second condition.

In optical applications space limitedness is a consequence of finite-sized objects. Band limitedness, on the other hand, is a consequence of the imaging device and the imaging environment. Knowing the above two recoverability conditions and having the freedom to choose either one depending on the application would definitely give the signal-processing engineer an advantage.

It is also interesting to show that, as the depth approaches zero, Eq. (11) tends to converge to simple sampling of the object:

$$\lim_{z \to 0} \sum_{\mathbf{k}} c_{\mathbf{k}} f\left(\mathbf{x} - \frac{\lambda z}{2\pi}\mathbf{U}\mathbf{k}\right)\exp(j\mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{x})$$

$$= f(\mathbf{x}) \sum_{\mathbf{k}} \exp(j\mathbf{k}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{x})$$

$$= f(\mathbf{x}) \sum_{\mathbf{k}} \delta(\mathbf{x} - \mathbf{V}\mathbf{k}). \qquad (16)$$

4. Simulated Examples

The simple one-dimensional (1-D) example shown in Fig. 1 gives further insight. A space-limited object (a slit) is shown in Fig. 1(a), and the real part of its diffraction pattern at a particular z is shown in Fig. 1(b). The diffraction pattern is sampled at a rate that is significantly lower than the Nyquist rate. Because this is a 1-D example, the sampling matrix is
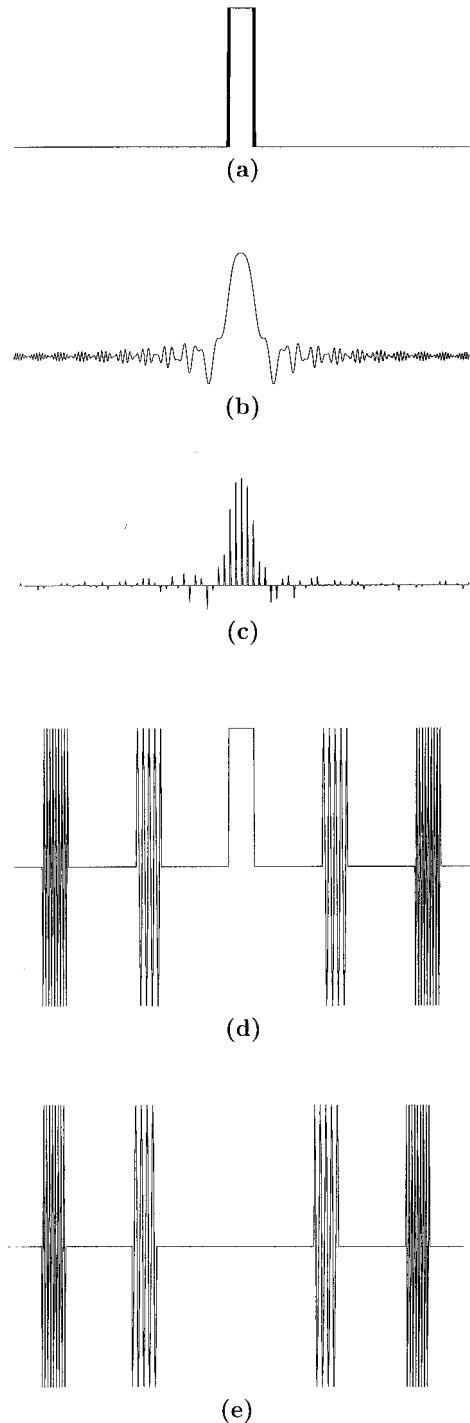

Fig. 1. Reconstruction of a 1-D object from its undersampled diffraction pattern: (a) a 1-D object (slit), (b) its diffraction pattern, (c) the sampled diffraction pattern (the sampling rate is below the Nyquist rate), (d) the reconstruction of the real part from the undersampled diffraction pattern, and (e) the reconstruction of the imaginary part from the undersampled diffraction pattern.

a scalar: $\mathbf{V} = T$; the result of the sampling is shown in Fig. 1(c). Figures 1(d) and 1(e) display the real and the imaginary parts, respectively, of the reconstruction from the sampled diffraction pattern. No physical units are presented because the dimensions

were normalized in discrete simulations that yielded these results.

The original object and its shifted and modulated replicas are clearly seen from Fig. 1. What we mean by modulation is multiplication by a complex sinusoidal function; therefore the modulation alters only the phase. As is expected from Eq. (11), the modulation frequency increases as the replica moves away from the origin. Full reconstruction can easily be achieved from Fig. 1(d) by use of a window that keeps the desired center pulse and eliminates the others. Furthermore, as a consequence of the modulation of the shifted pulses, we can also explain why low-pass filtering works for the purpose of full reconstruction in the case of a band-limited object. The sampling rate can be reduced, and the result will be more closely spaced, modulated replicas. This outcome can be seen from Eq. (11), where $\mathbf{U} = 2\pi/T$ for this 1-D example. The frequency of the carrier signal is $k(2\pi/T)$, where k represents the index associated with the modulated replicas. Interestingly, the effect of aliasing for this particular sampling problem has a convenient form, as is shown in Figs. 1(d) and 1(e). In most practical applications most of the signal energy of a band-limited object is still concentrated over a limited space; therefore either one of the reconstruction techniques would be satisfactory. The choice should depend on the convenience of the associated signal-processing implementation.

A 2-D simulated example is given in Fig. 2. Figure 2(a) shows a simple circular hole. The diffraction pattern at a particular distance is shown in Fig. 2(b). The diffraction pattern of Fig. 2(b) is sampled with a hexagonal geometry. Therefore the associated sampling matrix is

$$\mathbf{V} = \begin{bmatrix} T & T \\ T & -T \end{bmatrix}.$$

Again, to show the effect of the undersampling, we chose T to be rather large, and therefore significant aliasing occurs. In this example the corresponding $\mathbf{U}$ is

$$\mathbf{U} = \begin{bmatrix} \dfrac{\pi}{T} & \dfrac{\pi}{T} \\ \dfrac{\pi}{T} & -\dfrac{\pi}{T} \end{bmatrix}.$$

The reconstructed field is shown in Fig. 2(c). As is expected, as a consequence of the discussions of Section 2 and of Eq. (11), the form of the aliasing is very convenient: the aliasing generates modulated replicas of the reconstructed original, which is seen as the plain white spot at the center of the image. Full reconstruction, even when there is severe aliasing, is possible just by the windowing of the desired object at the center and the elimination of the modulated replicas. The reconstruction obtained in this way does not suffer any losses, and there is no blurring or other artifacts on the object.
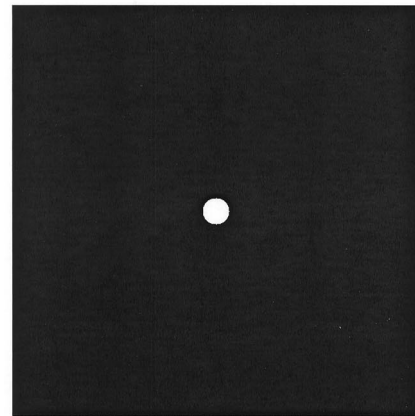
(a)
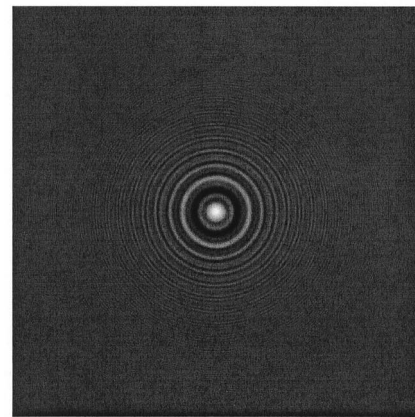
(b)
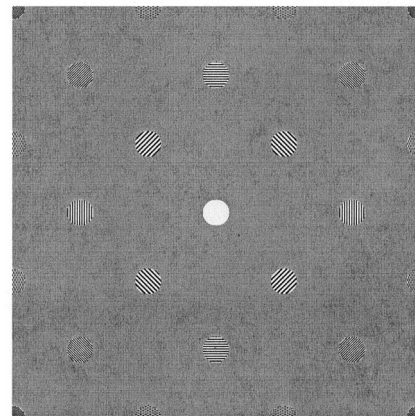
(c)

Fig. 2. Simulated data used in the reconstruction of a 2-D object from its undersampled diffraction pattern: (a) a 2-D object (a transparent circular hole on an opaque background), (b) its diffraction pattern, and (c) reconstruction from the undersampled diffraction field (the black background is shifted to a gray value to permit the observation of negative field values, as well). The darkest values represent the most negative values, whereas the lightest tones correspond to the highest (most positive) values.

5. Applications to In-Line Holography

The results described in Section 4 can be applied easily to in-line holography, which is related intrinsicly to diffraction.[5] The coherent illumination in in-line holography comes directly from the back.
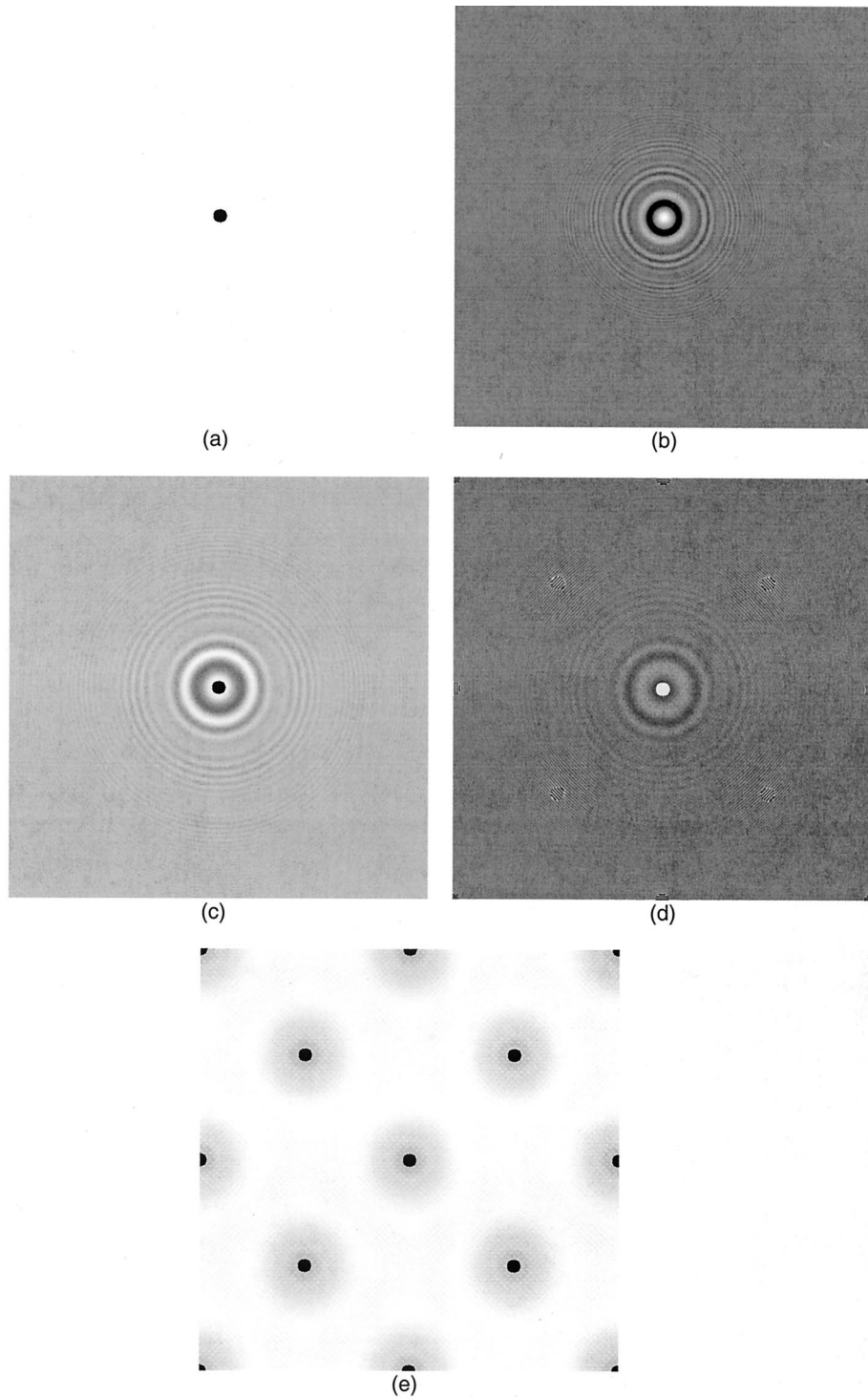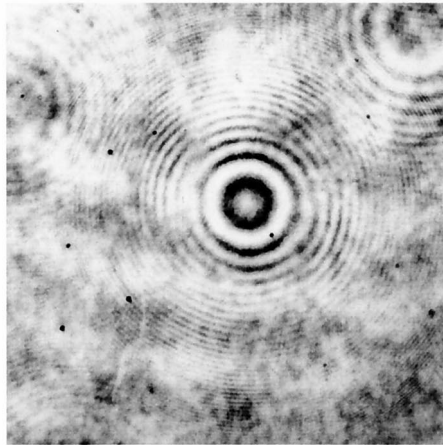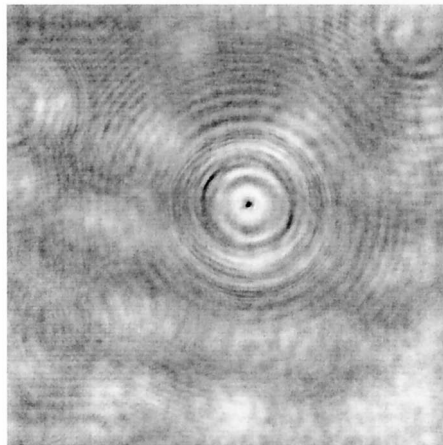
Fig. 3. Simulated data used to show the application of in-line holography: (a) a small 2-D opaque object on a transparent background, (b) its in-line hologram, (c) the conventional reconstruction of an object from its in-line hologram, i.e., the intensity, (d) the reconstructions from the undersampled in-line hologram's field, and (e) the reconstruction from the undersampled hologram's intensity.

Therefore, if the object's opacity is denoted as $a(x, y)$, where 1 and 0 represent total opacity and total transparency, respectively, the in-line hologram is the intensity of the diffraction pattern of the mask, $1 - a(x, y)$, at a distance z. So that the desired object information is not lost because of the nonlinear cross term,
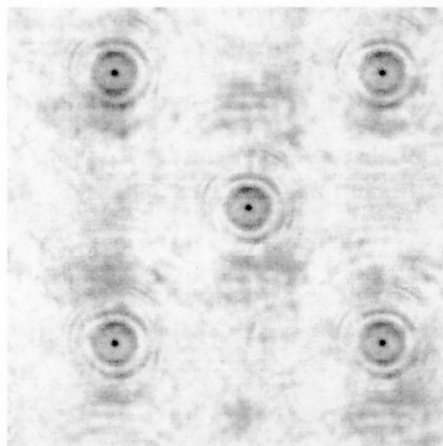
$a(x, y)$ must block only a small fraction of the background illumination. Reconstruction from an in-line hologram can be achieved by the illumination of the hologram by coherent light; in this case, the reconstructed pattern will be seen to be superposed on the hologram (this hologram is the hologram of the

(a)

(b)

(c)

Fig. 4. Real data used to show the application to in-line holography: (a) portion of a real optical in-line hologram of a dust particle on a glass substrate, (b) the conventional reconstruction by digital means, and (c) the reconstruction from the undersampled hologram.

original object at a distance of $2z$), which is called the twin image.

The effect of the undersampling of an in-line hologram on reconstruction is shown in Fig. 3 by simulation. Figure 3(a) shows the object-plane mask:

There is a small, opaque, circular object at the center of a transparent background. The in-line hologram of this object at a particular distance is shown in Fig. 3(b). Reconstruction from this hologram (with no undersampling) is displayed in Fig. 3(c). Figure 3(e) shows the reconstruction from the intensity from the same hologram when the hologram is severely undersampled; the replicas are as expected. To provide insight about the modulation of the replicas, we show the imaginary part of the field amplitude in Fig. 3(c). Please note that the information recovery from the undersampled diffraction pattern, as given by Eq. (14), is useful when the object has a finite size. In-line holography violates this condition because the illuminated background (theoretically) extends to infinity even if the object size is small. However, this is not a severe problem and can be overcome in various ways as long as the object size is still small. In the simulations given here this problem is solved by the elimination of the average illumination level (the dc part) of the hologram; this technique is straightforward in a digital signal-processing environment.

The same undersampling and reconstruction processes are also applied to an optical hologram, as shown in Fig. 4(a). The reconstruction from this hologram (with no undersampling) is shown in Fig. 4(b). Reconstruction is carried out by digital signal processing.[5] Reconstruction from the same hologram but after severe (hexagonal) undersampling is shown in Fig. 4(c). Again, the average illumination level of the hologram is eliminated before reconstruction.

6. Physical Interpretation of the Information Recovery from Samples

The mathematical foundations of the recovery of the original object field from the samples of its diffraction pattern are given in Section 2; the rules and the procedures for recovery are stated in Section 3. It is also interesting to add a physical interpretation to the recovery process.

The key analogy is to consider the sampling grid as a diffraction grating. The diffraction grating for the case of sampling is a periodic diffracting element that is opaque everywhere except at the sample locations, where it is totally transparent. Let us assume that the diffraction pattern $\psi_z(\mathbf{x})$ is recorded as a 2-D complex-valued (both amplitude and phase) mask. If this mask were illuminated by a reverse-propagating plane wave, the result would be the reconstruction of the original object at its original location, as given by Eq. (7). However, a plane wave passing through the diffraction grating (the sampling grid) would generate a number of plane waves that each propagate at a different angle; the directional cosines of these diffracted plane waves can be found easily from the sampling grid. Therefore if the recorded $\psi_z(\mathbf{x})$ is first multiplied by the sampling grid and then illuminated by a backpropagating plane wave the overall effect is equivalent to the illumination of $\psi_z(\mathbf{x})$ by a number of plane waves that each propagate backward at a different angle. Each such plane-wave component (usually called a diffraction

order) will result in a modulated term in which the carrier is a complex sinusoid and the envelope is a shifted replica of the original, as given by Eq. (11). The density and the pattern of the sampling grid will affect the angular separation of the plane-wave components; this, in turn, will affect the locations of the reconstructed replicas.

7. Conclusion

Contrary to the common belief that it is necessary to avoid undersampling (sampling below the Nyquist rate) for complete object recovery, objects can be recovered fully from their undersampled diffraction patterns even if the undersampling is severe. The condition for full recoverability from the undersampled diffraction patterns is to have finite-sized objects. There is full theoretical support for this result, as has been given in this paper. Simulations and applications to physical in-line holograms have shown that the presented procedure is feasible and useful. The ability to undersample the diffraction pattern without sacrificing the quality of the reconstructed object pattern gives computational (both storage and CPU time) savings when digital signal-processing techniques are employed in diffraction-related imaging applications. Further significant savings are possible when there is no need to keep the full image size during reconstruction. The results are valid for any arbitrary depth z except in the near field, where the Fresnel approximation is no longer valid.

The results that we have given in this paper are consequences of the special form of the convolution kernel that represents Fresnel diffraction. Because the fractional Fourier transform[6-8] is also based essentially on this kernel, the obtained results can be applied to the discretization of this transform, too.

References

1. D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing* (Prentice-Hall, New York, 1984), Sec. 1.4.
2. M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, New York, 1980), Chap. 8.
3. K. Nishihara, S. Hatano, and K. Nagayama, "New method of obtaining particle diameter by the fast Fourier transform pattern of the in-line hologram," Opt. Eng. **36,** 2429–2439 (1997).
4. S. Belaid, D. Lebrun, and C. Ozkul, "Application of two-dimensional wavelet transform to hologram analysis—visualization of glass fibers in a turbulent flame," Opt. Eng. **36,** 1947–1951 (1997).
5. L. Onural and P. D. Scott, "Digital decoding of in-line holograms," Opt. Eng. **26,** 1124–1132 (1987).
6. H. M. Ozaktas, O. Arikan, M. A. Kutay, and G. Bozdaği, "Digital computation of the fractional Fourier transform," IEEE Trans. Sig. Process. **44,** 2141–2150 (1996).
7. H. M. Ozaktas, M. A. Kutay, and D. Mendlovic, "Introduction to the fractional Fourier transform and its applications," in *Advances in Imaging and Electron Physics*, P. W. Hawkes, ed. (Academic, San Diego, Calif., 1999), Vol. 106, Chap. 4, pp. 239–291.
8. H. M. Ozaktas, B. Barshan, D. Mendlovic, and L. Onural, "Convolution, filtering, and multiplexing in fractional Fourier domains and their relation to chirp and wavelet transforms," J. Opt. Soc. Am. A **11,** 547–559 (1994).