Imaging from the zero locations of far-field-intensity data

C.-W. Liao and M. A. Fiddy

Department of Electrical Engineering, University of Massachusetts Lowell, Lowell, Massachusetts 01854

C. L. Byrne

Department of Mathematics, University of Massachusetts Lowell, Lowell, Massachusetts 01854

Received May 22, 1997; accepted July 8, 1997; revised manuscript received July 17, 1997

A new method is presented that allows the recovery of an image from those coordinates at which its power spectrum is zero. The coordinates of these zeros are points that are common to both the intensity data and the complex amplitude. The imaging method is based on a spectral estimation technique known as the prior discrete Fourier transform (PDFT, DFT using prior knowledge), which incorporates information about the image, such as a support constraint. © 1997 Optical Society of America [S0740-3232(97)02512-X]

1. INTRODUCTION

The Fourier transform of a one-dimensional signal or image of compact support is well known to be an entire function of exponential type. It can be represented in terms of its real and complex zero locations by means of an infinite product of factors (the Hadamard product), each factor encoding a zero's coordinate in the complex spectral plane.[1] It is also well known that if only power spectral data are measured, the complex zeros of the spectrum and its complex conjugate are located symmetrically about the real axis, making it impossible to distinguish between those belonging to the spectrum and those of its complex conjugate, in the absence of any *a priori* knowledge. If there are N complex zeros located, then one can generate 2^{N-1} distinct complex functions, consistent with the measured data, by zero flipping.

In more than one dimension, an entire function is generally not factorizable into an infinite product of terms but is irreducible.[1,2] From this, one expects a unique complex spectrum to be associated with the power spectrum. This association has provided an incentive to find methods that compute the missing phase values from the measured intensity data. The irreducible function representing the complex spectrum encodes a zero structure that occupies the associated two-dimensional (2D) complex space, of which the real plane is the measurement domain. Approximations to these irreducible zero factors have been found by numerically estimating the analytic continuation of the power spectral function off the real plane (i.e., measured) data into the complex domain and tracking where its modulus is zero. Attempting to locate the zero contour of the spectrum, as distinct from its complex conjugate,[3–6] has proved difficult.

The approach adopted in this paper is based on earlier work that recognized that the zero crossings of 2D band-limited functions typically are at isolated points on the real plane[1,7] rather than on closed or open curves. A practical approach to image recovery is therefore to try to estimate the 2D complex function from these point zeros,[8–10] provided that there is a sufficient number of them to represent the function and that the representation adopted can be justified. Having found the zero locations, we use these to estimate the complex valued spectrum.

2. TWO-DIMENSIONAL BAND-LIMITED FUNCTIONS

It is well known[11] that in two dimensions there tend to be a number of point zeros on the real plane approximately equal to the number of Shannon samples required to represent the function. This has been observed under different circumstances and reported over the years. An entire function of exponential type on the 2D real plane is the limit of a function analytic in the associated four-dimensional space. The analyticity of the function allows us to model the function as a simple polynomial, locally. From solutions of the parabolic wave equation, we expect many points of the field $F(a, b)$ to be zero, and we can represent the field in the vicinity of these zeros[7] by a first-order approximation. Thus in the immediate neighborhood of a zero at a_0, b_0 we can approximate the function by the simple linear term:

$$F(a, b) = A(a - a_0) + iB(b - b_0), \tag{1}$$

where A and B can be assumed to be complex constants in a sufficiently small neighborhood. The increment or decrement of phase moving around the zero is equal to 2π and is well represented locally by the phase of Eq. (1).

We shall assume throughout that the object function to be reconstructed is the real-valued function $f(x, y)$ possessing nonzero values only within the compact support region $|x| \leq s$, $|y| \leq s$. The two-dimensional Fourier transform of $f(x, y)$ is the complex-valued function $F(a, b)$ given by

0740-3232/97/123155-07$10.00

$$F(a, b) = \iint f(x, y)\exp[-i(xa + yb)]\mathrm{d}x\mathrm{d}y/4\pi^2. \quad (2)$$

From the Fourier inversion theorem we have

$$f(x, y) = \iint F(a, b)\exp[i(xa + yb)]\mathrm{d}a\mathrm{d}b. \quad (3)$$

Since $f(x, y)$ is compactly supported, it can be represented by Fourier series in an infinite number of terms. For any $\Delta > 0$ such that $\pi/\Delta \geqslant s$, we can write

$$f(x, y) = \Delta^2 \sum_m \sum_n F(m\Delta, n\Delta)\exp[i(xm\Delta + yn\Delta)], \quad (4)$$

for all $|x| \leqslant \pi/\Delta$ and $|y| \leqslant \pi/\Delta$. So, in particular, representation (4) holds for all (x, y) within the support of the object. If the square $|x| \leqslant s$, $|y| \leqslant s$ is the smallest square containing the support of f, then $\Delta = \pi/s$ is called the Nyquist sampling rate. From Eq. (4) we see that for any $\Delta \leqslant \pi/s$ the object function f can be completely recovered from the (infinitely many) samples of F on the lattice $\{(m\Delta, n\Delta)\}$. In practice we usually do not know s exactly, so we choose Δ sufficiently small that s is almost certainly less than π/Δ; indeed, there is sometimes good reason to oversample, that is, to take Δ even smaller than our estimate of π/s. In any situation involving actual measurements, we can have only finitely many values of F, so the representation given by Eq. (4) must be replaced by something we can compute from the finite data.

The basic problem that we address in this paper is the reconstruction of the real object function $f(x, y)$ from finitely many samples $\{|F(m\Delta, n\Delta)|\}$, with Δ significantly less that our estimate of π/s; here $|F(m\Delta, n\Delta)|$ denotes the magnitude of the complex quantity $F(m\Delta, n\Delta)$. This problem is often called the phase-retrieval problem, although even if we had the phases as well, there would still be the problem of reconstructing from finite data. Our solution to this phase-retrieval problem will make use of a method called the prior discrete Fourier transform[12,13] (PDFT) for reconstructing f from finitely many oversampled values of the complex function F, using an estimate of s. We discuss first the PDFT and then the application of the PDFT to the solution of the phase-retrieval problem.

3. PRIOR DISCRETE FOURIER TRANSFORM FOR RECONSTRUCTION FROM FINITE DATA

Let us assume now that we have finitely many complex data values $\{F(m\Delta, n\Delta), |m| \leqslant M, |n| \leqslant N\}$. Since f is

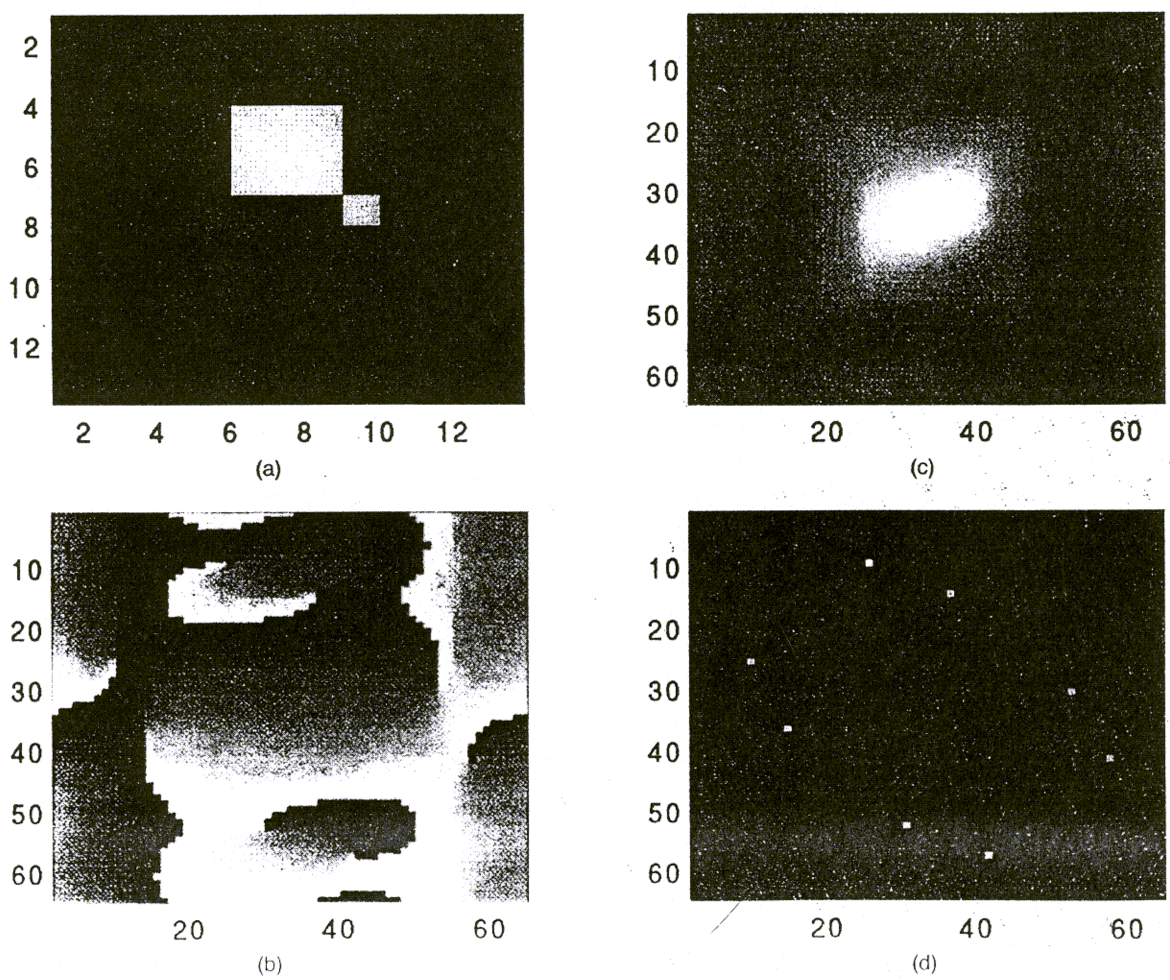

Fig. 1. (a) Object, (b) correct and wrapped phase function; (c) power spectrum; (d) real point zero locations.

taken to be real valued, the function F is conjugate symmetric, so that $F(-m\Delta, -n\Delta) = F(m\Delta, n\Delta)^*$, where * denotes complex conjugate. It is natural, therefore, to take our sampling points within a rectangle centered at the origin. Let us also assume that we have an estimate of s and that Δ has been chosen so that $s < \pi/\Delta$, so the data is oversampled.

By the discrete Fourier transform (DFT) estimate of $f(x, y)$ we mean the function of continuous x and y given by the truncated version of Eq. (4):

$$\begin{aligned}\mathrm{DFT}(x, y) &= \Delta^2 \sum_M \sum_N F(m\Delta, n\Delta) \\ &\quad \times \exp[i(xm\Delta + yn\Delta)], \\ &\qquad \text{for } |x|,|y| \leq s, \\ &= 0 \quad \text{otherwise}, \end{aligned} \tag{5}$$

where Σ_M indicates summation over $|m| \leq M$. If s is significantly less than π/Δ, this estimate may not be accurate. One reason for that is the following: we have made use of our knowledge about the support s when we set the DFT equal to zero for $|x|$, $|y| > s$, but one effect of this is that the DFT estimator so defined is no longer consistent with the measured data. In other words, if we take the Fourier transform of this estimator and evaluate it at the lattice points where we had data, we do not get back the values $F(m\Delta, n\Delta)$; to get back the data we must let the DFT exist all the way out to $|x|$, $|y| \leq \pi/\Delta$. A way out of this dilemma is to seek an estimator similar in form to the DFT but adjusted so as to be consistent with both the support and the data constraints; this is the PDFT.[12,13]

Let $p(x, y) = 1$ for $|x|$, $|y| \leq s$, and $p(x, y) = 0$ otherwise. The PDFT estimator of $f(x, y)$ is given by the function

$$\begin{aligned}\mathrm{PDFT}(x, y) &= p(x, y) \sum_M \sum_N \alpha(m, n) \\ &\quad \times \exp[i(xm\Delta + yn\Delta)], \end{aligned} \tag{6}$$

where the coefficients $\alpha(m, n)$ are determined by requiring that the Fourier transform of the PDFT(x, y) take on the values $F(m\Delta, n\Delta)$ for $|m| \leq M$, $|n| \leq N$.

The linear equations that result are

$$\begin{aligned}F(j\Delta, k\Delta) &= \sum_M \sum_N \alpha(m, n) \\ &\quad \times P[(j - m)\Delta, \quad (k - n)\Delta], \\ &\qquad |j| \leq M, |k| \leq N, \end{aligned} \tag{7}$$

where

$$P(a, b) = \iint p(x, y)\exp[-i(xa + yb)]\mathrm{d}x\mathrm{d}y/4\pi^2. \tag{8}$$

From our definition of $p(x, y)$ it follows that

$$P(a, b) = \sin(sa)\sin(sb)/\pi^2 ab. \tag{9}$$

Therefore system (7) becomes

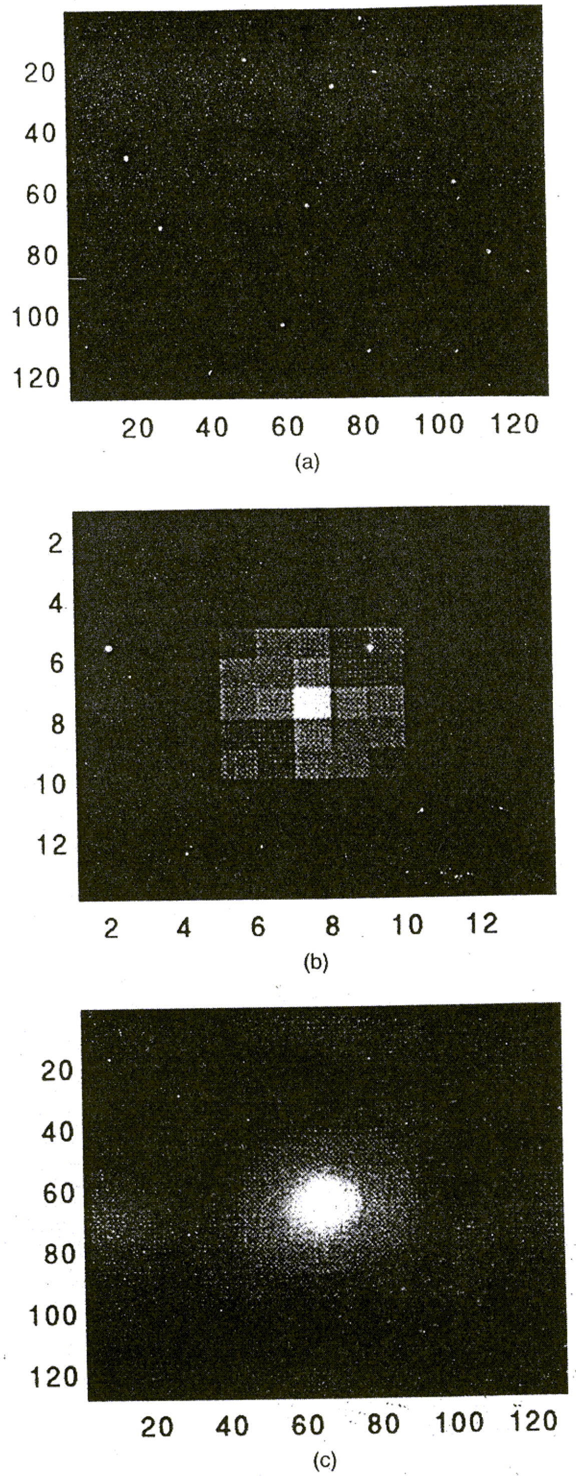

Fig. 2. (a) Nine data points used to estimate $r(x, y)'$; (b) $r(x, y)'$; (c) zero contours of $R(a, b)$.

$$\begin{aligned}F(j\Delta, k\Delta) &= \sum_M \sum_N \beta(m, n)\mathrm{sinc}[(j - m)\Delta] \\ &\quad \times \mathrm{sinc}[(k - n)\Delta], \\ &\qquad |j| \leq M, |k| \leq N, \end{aligned} \tag{10}$$

with $\beta(m, n) = (s/\pi)^2\alpha(m, n)$ and $\mathrm{sinc}(t) = \sin(t)/t$, $\mathrm{sinc}(0) = 1$.

The PDFT method consists of solving Eq. (10) for the $\{\beta(m, n)\}$ and then using the resulting $\{\alpha(m, n)\}$ in Eq. (6) as the estimate of $f(x, y)$. If Δ is nearly equal to π/s then the PDFT and the DFT are not greatly different; however, when Δ is half of π/s or smaller we begin to see a difference.

One note of caution: as the Δ becomes smaller, system (10) becomes increasingly ill-conditioned and hence sensitive to noise in the data. To combat this, we always multiply the values of P in Eq. (7) corresponding to $m = j$, $n = k$ by (for example) 1.001 or 1.0001 before solving Eq. (10). This regularization minimizes the ill-conditioning[13] by improving the condition number of the P matrix.

3. PRIOR DISCRETE FOURIER TRANSFORM FOR PHASE RETRIEVAL

We assume now that we have only the magnitude data $\{|F(m\Delta, n\Delta)|, |m| \leq M, |n| \leq N\}$, where again we have an estimate of the support limit s and have chosen Δ so that π/Δ is greater than our estimate of s. The complex-valued function $F(a, b)$ can be written in terms of its real and imaginary parts as follows:

$$F(a, b) = R(a, b) + iQ(a, b), \tag{11}$$

where both R and Q are real-valued functions. Letting $r(x, y)$ and $q(x, y)$ be the inverse Fourier transforms of R and Q, respectively, we have that

$$r(x, y) = [f(x, y) + f(-x, -y)]/2, \tag{12}$$

$$q(x, y) = [f(x, y) - f(-x, -y)]/2i. \tag{13}$$

The function $R(a, b)$ is real and symmetric; that is, $R(-a, -b) = R(a, b)$. Those data values $|F(m\Delta, n\Delta)|$ that are (nearly) zero correspond to (nearly) zero values of both $R(m\Delta, n\Delta)$ and $Q(m\Delta, n\Delta)$. We begin by using (some of) these point zeros, along with our support information, to reconstruct $r(x, y)$ by means of the PDFT.

To reconstruct $r(x, y)$ from zeros of $R(a, b)$ by means of the PDFT, we must also have at least one nonzero value of R; otherwise, the PDFT estimate of r will be identically zero. We know that since $F(-a, -b) = F(a, b)$, $F(0, 0)$ is real, therefore $R(0, 0) = |F(0, 0)|$ or $R(0, 0) = -|F(0, 0)|$; we take $R(0, 0) = |F(0, 0)|$, arguing that $f(x, y)$ is often nonnegative in practice. The PDFT estimator of $r(x, y)$ then takes the form

$$r(x, y)' = \mathrm{PDFT}_r(x, y) = p(x, y) \sum_z c(m, n) \times \exp[i(xm\Delta + yn\Delta)], \tag{14}$$

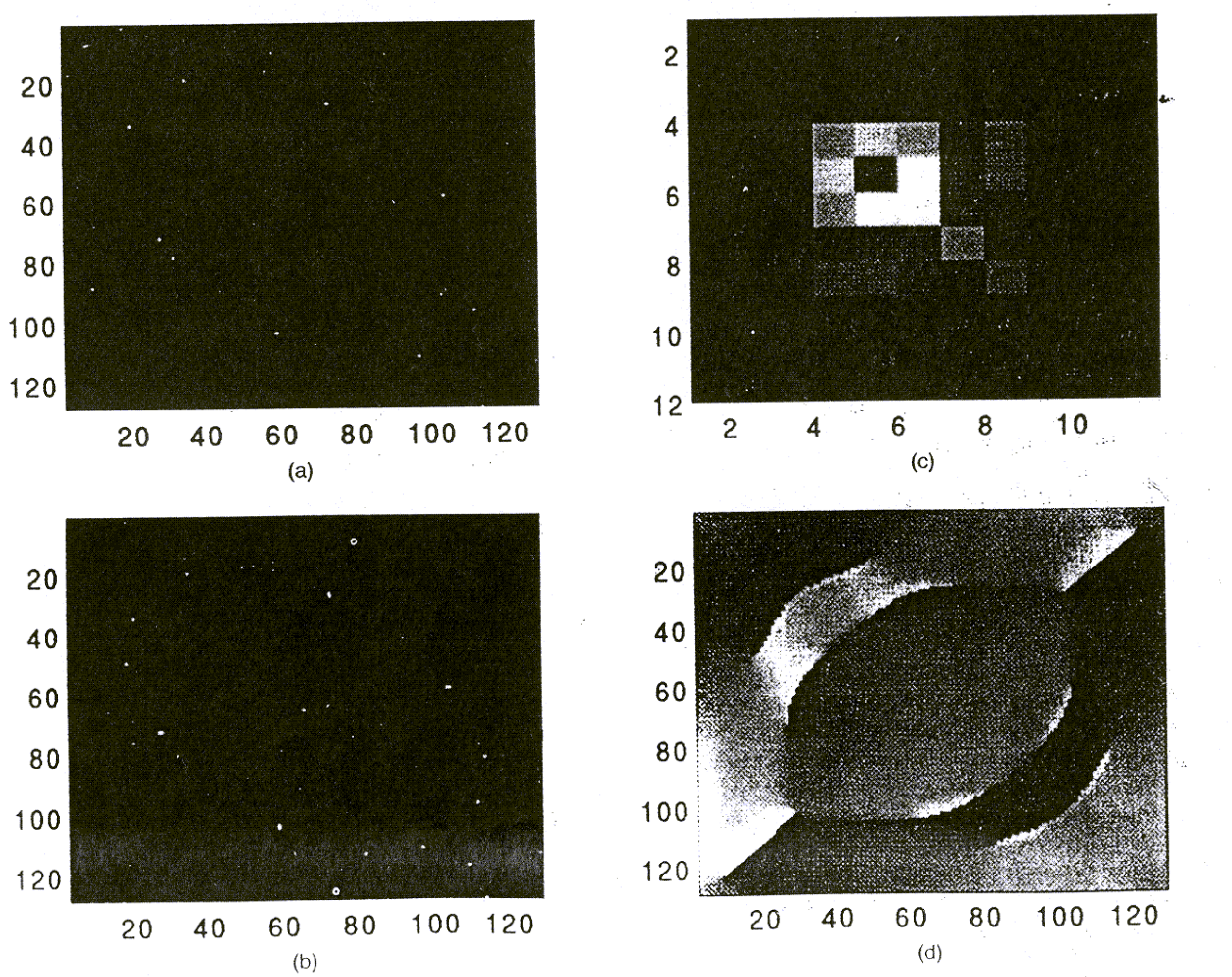

Fig. 3. (a) Twelve locations at which values of $Q(a, b)$ are defined; (b) total number of complex data points available for PDFT estimator; (c) reconstruction of object; (d) estimated Fourier phase.

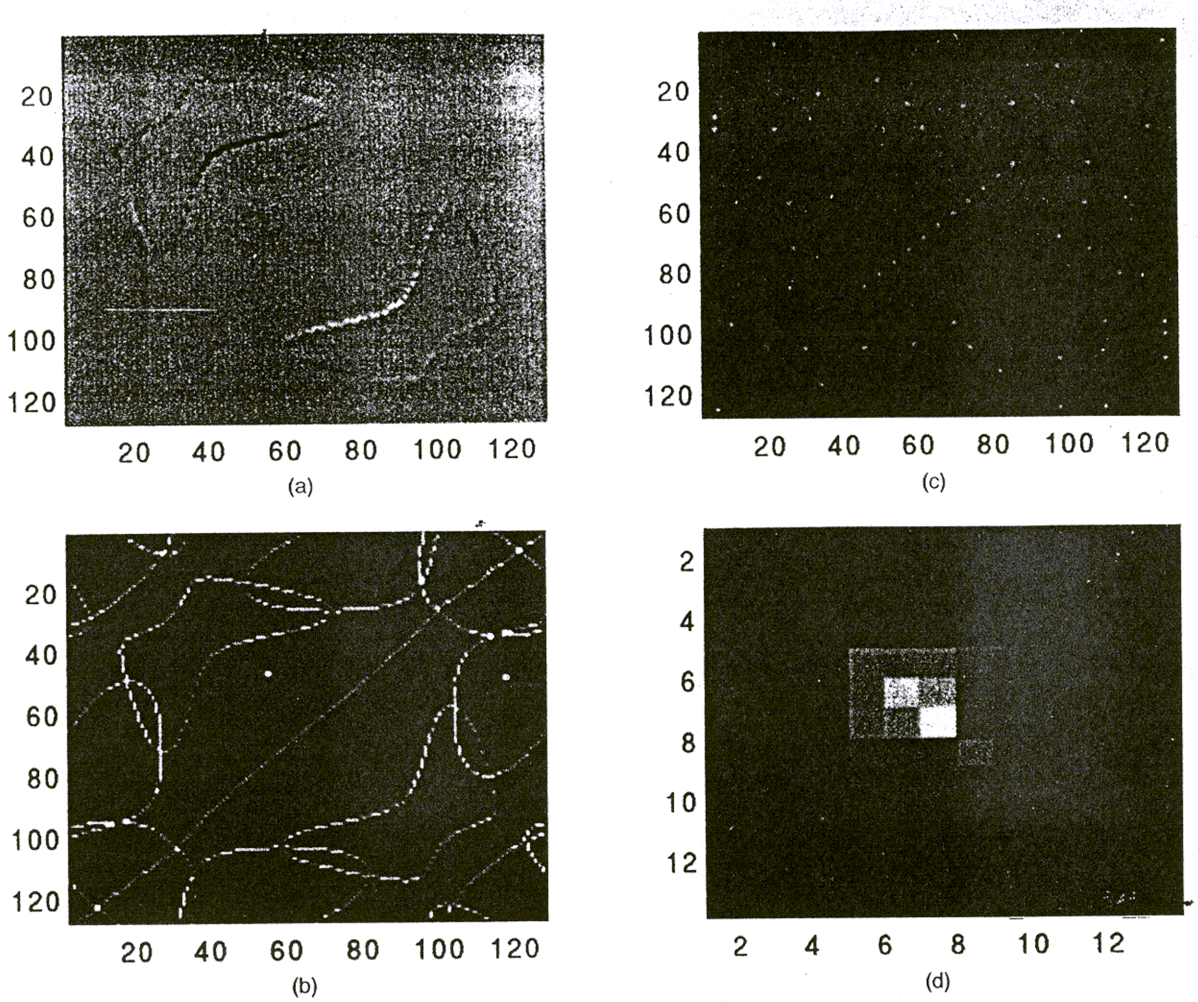

Fig. 4. (a) Zero contours of the magnitude of the estimated imaginary part of the spectrum; (b) zero contours of both the real and the imaginary parts, indicating the points at which they intersect; (c) total complex data points available for PDFT; (d) object estimate.

where Σ_z denotes summation over those pairs (m, n) for which the lattice point $(m\Delta, n\Delta)$ is one of the point zeros being used to generate the function, as well as over the pair (0, 0). Once again, we find the coefficients $c(m, n)$ by forcing the Fourier transform of Eq. (14) to agree with $|F(0, 0)|$ at the point (0, 0) and to be zero at the point zero locations being used in the sum. Once we solve the resulting system of linear equations for the $c(m, n)$, we put these values into Eq. (14); this is our estimate $r(x, y)'$ of $r(x, y)$.

Now we take the Fourier transform of this estimate $r(x, y)'$ to get $R(a, b)'$. From the values $R(m\Delta, n\Delta)'$ for $|m| \leq M$, $|n| \leq N$, we obtain an estimate of the curves along which $R(a, b) = 0$. Once we have estimated where $R(a, b) = 0$, we have locations at which we can claim to know $|Q(a, b)| = |F(a, b)|$. Along the zero curves of $R(a, b)$, the function $Q(a, b)$ can change sign only when a zero of $|F(a, b)|$ is encountered. Therefore we now know (estimates of) $Q(a, b)$, except for the sign, along a number of curves in the (a, b) plane.

The next step is to select a certain finite number of these locations and to assign the signs in an arbitrary way, consistent only with the antisymmetry of Q; that is, $Q(-a, -b) = -Q(a, b)$. With these real numbers as data now, we again use the PDFT and the prior knowledge of s to estimate the purely imaginary function $q(x, y)$. Once we have obtained our estimate $q(x, y)'$ of $q(x, y)$, we Fourier transform back to get $Q(a, b)'$, our estimate of $Q(a, b)$. From $Q(a, b)'$ we now have (estimates of) the zero curves of $Q(a, b)$, which must then be locations at which we know the values of $R(a, b)$, except for the sign.

We can continue indefinitely in this fashion, but at some point we stop and use some of the estimated values of $R(a, b)$ and $Q(a, b)$, along with our value of s, to do a final PDFT estimate of $f(x, y)$. We choose to estimate the symmetric $r(x, y)$ first because we cannot estimate an unsymmetric $f(x, y)$ from symmetric point zeros and a symmetric support function $p(x, y)$; we always get what looks like a twin image if we view it as an estimate of f itself. It is better to view the result of this first step as an estimate only of $r(x, y)$ and then to go back to (a, b) space to estimate the full zero structure of $R(a, b)$.

There are many opportunities for variation in what we have just described, and the results we present in this paper will illustrate some of these possibilities. What we have outlined here is the main idea behind the use of the PDFT for phase retrieval.

4. NUMERICAL RESULTS

A simple binary object as shown in Fig. 1(a) was used in these examples; it covers an area of only 4 × 4 pixels. This object is well known to lead to an irreducible spectrum[7] with a readily seen set of real point zeros. Figure 1(b) shows the correct and the wrapped phase function for the spectrum of this object and Fig. 1(c) its power spectrum. Although there is evidence of some point zeros being located toward the edges of the array, we use only those that are surrounding the central region. Those particular zeros are shown in Fig. 1(d). Knowing that the object is real and assuming that it is more positive than negative, we can define the central square root of the intensity value $|F(0, 0)|$ to be equal to $R(0, 0)$, thus giving nine data points from which to estimate $r(x, y)'$ with the PDFT, as shown in Fig. 2(a) where the support constraint used in the PDFT is a window of 5×5 pixels. This estimate is shown in Fig. 2(b) and the magnitude of the Fourier transform of this, $R(a, b)$ is shown in Fig. 2(c). Note that $R(a, b)$ has zero contours, and at these locations we can deduce values for $Q(a, b)$ from the intensity data.

Figure 3(a) shows just 12 locations at which values of $Q(a, b)$ have been defined. The sign of these values is chosen to reflect the known Hermitian character of the complex spectrum. Figure 3(b) shows the total number of complex data points now available to use with the PDFT; they include an estimate of the real part (when the imaginary part is zero), the imaginary part (where the real part is zero), and points where both the real and the imaginary parts are zero. The PDFT reconstruction using all of these 21 points is shown in Fig. 3(c). The correct object shape is beginning to emerge in the 5 × 5 reconstruction window. The Fourier phase associated with this estimate is still in error, however, as shown in Fig. 3(d).

From this estimate of the object, one can calculate $q(x, y)$ and identify its zero contours. At these locations new estimates for $R(a, b)$ can be inferred from the measured intensity data. As the number of estimated complex spectral values increases, the PDFT estimate using them and the prior knowledge that the object is confined to a 5 × 5 window improves. Indeed, from the overlap of zero contours of the real and the imaginary parts, one can more precisely sign the real and imaginary parts, since a sign change of either the real or the imaginary part can occur only when the part passes through a zero. Figure 4(a) shows the zero contours of the magnitude of the estimated imaginary part of the spectrum and Fig. 4(b) the zero contours of both the real and imaginary parts, indi-

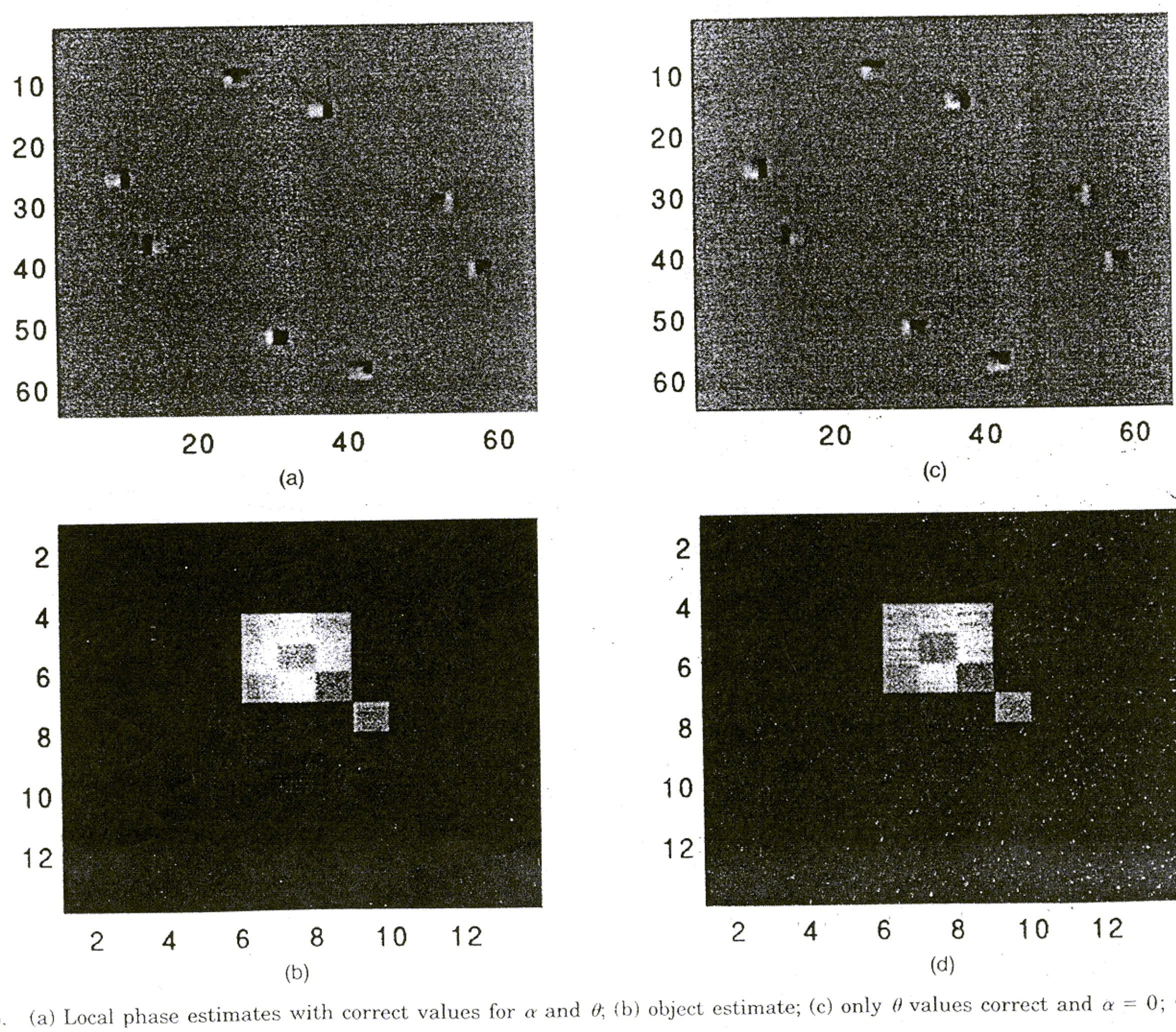

Fig. 5. (a) Local phase estimates with correct values for α and θ; (b) object estimate; (c) only θ values correct and $\alpha = 0$; (d) object estimate.

cating the points at which they intersect. From this information, additional complex spectral values are deduced, shown in Fig. 4(c), leading to the object estimate in Fig. 4(d).

The success of this approach depends on the accuracy with which one can locate the zeros of the intensity function, and this might require some care when noisy data are used. Also, the estimates for the real and the imaginary parts will be dependent on the choice of the prior function $p(x, y)$; if this is poor or contains incorrect information about the object, it will lead to errors in the final estimate.

An interesting and complementary approach to using this method is illustrated next. In previous publications,[9,10] an initial estimate for the phase of the spectrum was calculated with a product expansion incorporating the real zero locations as if they were the roots of a factorizable polynomial. This procedure met with some success but does not always provide a phase estimate good enough that an iterative procedure such as the error reduction technique can recover the true phase. In its more general form, Eq. (1) can be written as

$$F(a, b) = \exp(-i\alpha)\{A[(a - a_0)\cos\theta + (b - b_0)\sin\theta] + iB[(a - a_0)\sin\theta + (b - b_0)\cos\theta]\}. \quad (15)$$

Around each of these real point zeros, we do know that the phase of the function will behave very much like the phase of this function with corrections to the phase, α, for the offset of the zero from the x,y origin and with the orientation θ of the phase wrap. With this model, the phases, and hence the complex values, at the points immediately surrounding each zero can be estimated. The results are shown in Fig. 5(a). Based on these complex data, totaling 64 points in this case, the PDFT estimate is very good, as shown in Fig. 5(b). In Fig. 5(c) we assumed that the phase offset values for each zero are all zero, and so only θ is known. Under these conditions, the object estimate is again excellent [Fig. 5(d)]. This suggests that if one can make use of spectrum symmetries and if from the intensity map one can infer the orientation of the linear phase term around each zero, then the PDFT can provide a good object estimate directly.

5. CONCLUSIONS

There are many situations in which only Fourier magnitude data or intensity data can be measured. We have demonstrated a new and rigorous method for estimating a complex spectrum from its (finitely many) intensity data samples. The samples used are those locations at which the intensity data are zero. The method provides an optimal estimate for the object function in the sense that the PDFT is optimal. It also provides simultaneously an interpolated/extrapolated complex data set from which improved image resolution results. Numerical examples were shown indicating the success of the approach. A modification to the basic method was also presented, in which the phase behavior in the neighborhood of each intensity zero is also estimated and incorporated into the PDFT restoration procedure. This method can provide an improved estimate with just the zero points and their immediate (phased) neighbors, but only if the phase offsets and phase-wrap orientations can be reasonably well estimated *a priori*. Function symmetries can assist with this.

ACKNOWLEDGMENT

M. A. Fiddy acknowledges support of NATO travel grant 910274.

REFERENCES

1. M. A. Fiddy, "The role of analyticity in image recovery," *Image Recovery: Theory and Applications*, H. Stark, ed. (Academic, Orlando, Fla., 1987), pp. 499–529.
2. M. H. Hayes and J. H. McClellan, "Reducible polynomials in more than one variable," Proc. IEEE **70**, 197–198 (1982).
3. H. M. Berenyi, H. V. Deighton, and M. A. Fiddy, "The use of bivariate polynomial factorization algorithm in two dimensional phase problems," Opt. Acta **32**, 687–700 (1985).
4. M. Nieto-Vesperinas and J. C. Dainty, "Phase recovery for two dimensional digital objects by polynomial factorization," Opt. Commun. **58**, 83–88 (1986).
5. R. H. T. Bates, H. Jiang, and B. L. K. Davey, "Multidimensional system identification through blind deconvolution," Multidimensional Syst. Signal Process. **1**, 127–142 (1990).
6. R. H. T. Bates, B. K. Quek, and C. R. Parker, "Some implications of zero sheets for blind deconvolution and phase retrieval," J. Opt. Soc. Am. A **7**, 468–474 (1990).
7. M. S. Scivier and M. A. Fiddy, "Phase ambiguities and the zeros of multidimensional band-limited functions," J. Opt. Soc. Am. A **2**, 693–698 (1985).
8. P. Chen, M. A. Fiddy, A. H. Greenaway, and Y. Wang, "Zero estimation for blind deconvolution from noisy sampled data," in *Digital Image Recovery and Synthesis II*, Paul S. Edell, ed., Proc. SPIE **2029**, 16–24 (1993).
9. P. Chen and M. A. Fiddy, "Image reconstruction from power spectral data using point zero locations," J. Opt. Soc. Am. A **11**, 2210–2214 (1994).
10. P. Chen, M. A. Fiddy, C-W. Liao, and D. A. Pommet, "Blind deconvolution and phase retrieval using point zeros," J. Opt. Soc. Am. A **13**, 1524–1531 (1996).
11. N. B. Baranova, A. V. Mamaev, N. F. Pilipetsky, V. V. Shkunov, and B. Ya. Zel'dovich, "Wave-front dislocations," J. Opt. Soc. Am. **73**, 525–527 (1983).
12. C. L. Byrne and M. A. Fiddy, "Signal reconstruction as Wiener filter approximation," in *Photon Correlation Techniques and Applications*, J. B. Abbiss and A. E. Smart, eds., Vol. 1 of OSA Proceedings Series (Optical Society of America, Washington, D.C., 1988), pp. 18–23.
13. C. L. Byrne, R. M. Fitzgerald, M. A. Fiddy, T. J. Hall, and A. M. Darling, "Image restoration and resolution enhancement," J. Opt. Soc. Am. **73**, 1481–1487 (1983).