Phase retrieval of arbitrary complex-valued fields through aperture-plane modulation

Fucai Zhang,* Giancarlo Pedrini, and Wolfgang Osten
Institut für Technische Optik, Universität Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany
(Received 18 December 2006; published 4 April 2007)

We propose and experimentally demonstrate a simple and robust approach for retrieving arbitrary complex-valued fields from three or more diffraction intensity recordings. We need no *a priori* knowledge about the object field. The convergence rate is rapid. We obtained good results using experimental data with only 80 iterations (160 fast Fourier transforms). The method does not suffer any stagnation or ambiguity problem, and it also exhibits a high immunity to noise. The technique exhibits great potential in lensless phase-contrast imaging, wave-front sensing, and metrology for a wide spectral range.

DOI: 10.1103/PhysRevA.75.043805 PACS number(s): 42.30.Rx, 42.30.Kq, 42.30.Va

The phase of radiation scattered from an object carries important information about the object surface or its inner structure. Its measurement has been a key means of investigation in many fields, as in materials and biological sciences. Because only the magnitude of radiation can be measured directly over a large spectral range, the phase problem arises. Solutions to the phase problem fall into two categories: interferometric approaches [1–3] and beam-propagation-based approaches [4–6]. With the rapid progress in computing techniques, as well in digital imaging devices, beam-propagation-based phase recovery approaches have been receiving increasing interest. Much significant progress has been made, especially in x-ray diffraction imaging [7–9], in recent years. Several reasons that underlie this trend are the simple experimental setup, robustness to external influence, and suitability for various wavelengths.

The method proposed by Gerchberg and Saxton in 1972 [4] is the first widely accepted beam-propagation-based phase recovery method. The idea is that the missing phases can be recovered by iteratively applying the magnitude constraints in object and Fourier-transformed space. Fienup modified the Gerchberg-Saxton algorithm in 1978 by using finite support and non-negativity constraints in object space instead of the magnitude of the object [5]. In subsequent years, Fienup's algorithms and its variations have been successfully demonstrated for various applications, in particular for real and non-negative objects. In 1998, Miao *et al.* clarified the sampling requirement for unique phase retrieval and pointed out that oversampling of diffraction patterns usually guarantees the convergence [6]. In some important fields, such as electron microscopy and x-ray diffraction imaging, the object fields are complex valued. For instance, multiple scattering of electrons or a spatial variation in the anomalous scattering of x rays can both give rise to a complex object field. Although it has been shown that reconstruction of a complex object is possible if a strong support is available [10] or if a low-resolution image can also be measured [11], it is a common perception that, because of the loss of the non-negative constraint, phase retrieval in general for complex-valued objects is much more difficult than from real and non-negative objects. Strong support constraints include certain special shapes or separated supports, and the support needs also to be sharp and sufficiently tight (true boundary). One solution is then to develop experimental strategies. In recent work, Faulkner and Rodenburg [12] suggested moving an aperture to record a set of patterns diffracted from different part of the object. The object field in the overlapped area can be better determined because its diffraction would appear in several recordings, thus reducing the possibility of stagnation. In recent experimental work [13,14], Rodenburg *et al.* have successful demonstrated this technique with laser and hard x-ray sources using several hundred recordings. Another experimental strategy is to add several known phase curvatures into the incident beam when recording the diffraction patterns [15]. One theoretical solution proposed by McBride *et al.* is to introduce a difference map to specify the basic iteration to overcome the stagnation problem [16].

Wave-front sensing is another class of noninterferometric approach for phase recovery. The three wave-front sensors most commonly used are the Shack-Hartmann, the curvature, and the pyramid sensors. The essence of these sensors is the introduction of a known phase perturbation in the recording of intensities. From the intensities, the slope (or curvature) of the phase of the wave front can be estimated. These sensors usually can only work for smooth wave fronts.

In this paper, we show that one can significantly improve the convergence and robustness of the iterative method by adopting the essence of the wave-front sensing technique. A strong phase modulation is intentionally introduced into the object field. Thus, in the recorded diffraction pattern, any sampling would have contributions from all points in the object. Distributing information of one object point to all pixels of the sensor minimizes the effect of sensor noise and also eliminates the stagnation.

Figure 1 illustrates the experimental arrangement. A pixelated phase plate, mounted on a linear stage, is positioned downstream of an extended object. The object can be of transmission or reflection type. The distances from the plate to the object and the plate to the sensor [charge-coupled device (CCD)] are Z_1 and Z_2, respectively. The phase of the plate is known and distributed uniformly in the range of 0 to 2π. The wave front at the object plane U_{obj} and that before the plate are related by the Sommerfeld-Rayleigh diffraction integral. The wave front after the plate and that on the CCD are related by the Fresnel integral, since Z_2 is selected to be large enough. Thus, the Fresnel algorithm [17] can be ap-

*Electronic mail: fucai@ito.uni-stuttgart.de

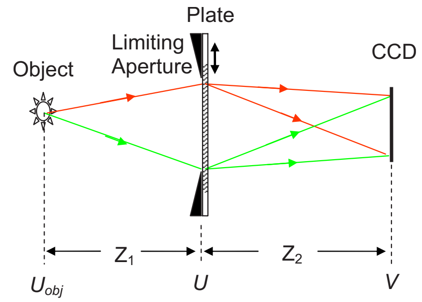

FIG. 1. (Color online) Experimental arrangement of wave-front-modulation-based phase retrieval approach.

plied for a fast calculation. Supposing that the pixel pitch of the CCD is Δx, the number of pixels of the CCD in one dimension is N, and λ is the wavelength of light, according to the sampling interval relationship in the Fresnel algorithm, the pixel pitch of the plate needs to be $\Delta\xi = \lambda Z_2/(N\Delta x)$. In order for the CCD to resolve the diffraction pattern, closely before the plate is placed a square aperture with a side length $\lambda Z_2/\Delta x$. Notice that the recorded diffraction pattern is not oversampled.

M diffraction patterns are collected, as the plate is shifted transversely by a multiple of the plate pixel pitch. The new phase retrieval algorithm, as illustrated in Fig. 2, starts from a random estimate of U, the wave front before the plate, and proceeds with the following steps in an iterative manner. (1) Modulate the current estimate of U_n with the plate phase $\varphi_m = \Psi(\xi - m\Delta\xi)$, yielding a modulated wave front $\tilde{U}_n = U_n \exp(j\varphi_m)$, where n is the iteration number, m is the current plate position ($m = 0, 1, \ldots, M$), and the function $\Psi(\xi)$ represents the phase distribution of the plate. (2) Propagate the wave front $\tilde{U}_n$ using the Fresnel algorithm to the CCD to obtain an estimate V_n of the diffraction field. (3) Replace the magnitude of V_n with the recorded one. (4) Propagate the corrected wave front $\tilde{V}_n$ back to the plate plane. (5) Remove the plate phase modulated in step 1 to obtain a new and improved estimate of the wave front before the plate. (6) Repeat steps 1–5 for the next plate position, i.e., $m = m+1$. If the last plate position is reached, then use the first one. This process ends when the change of amplitude between two successive retrieved wave fronts before the plate is sufficiently small. After the object field on the plate plane has been obtained, further propagation to the object plane gives the original object field. Under the Fresnel approximation, the object resolution in reconstruction is approximately given by

$$\delta x = \lambda Z_1/S = (Z_1/Z_2)\Delta x, \tag{1}$$

where S, $= \lambda Z_2/\Delta x$, represents the size of the limiting aperture before the plate.

In this algorithm, the wave front can be smooth or totally random in both its amplitude and phase. The only information needed to know is the phase retardance introduced by the plate. The use of a plate with a random phase results in more uncorrelated diffraction patterns, allowing for retrieving more information from each recording, and thus eliminating the stagnation problem. The solution given by the approach is unique since the random phase of the plate effectively breaks any symmetry that may exist in the object field. Therefore, even trivial ambiguities, such as the object field shift and the twin image as appearing in the Fienup algorithms, do not exist in the proposed approach. Furthermore, in the recording of the diffraction pattern, the approach avoids the over exposure problem. Here, the nondiffracted light will be diffracted by the phase plate. By selecting the phase of the plate and the size of its pixel, one can make the diffraction pattern have a uniform intensity distribution over the whole sensing area of the CCD.

Simulations have been done for different kinds of object fields, ranging from an isolated square amplitude plus a spherical phase to a Gaussian random amplitude plus a uniform random phase. In all cases the iteration converges rapidly when the number of recordings is larger than 4. Figure 3 shows a typical simulation result. The test object magnitude is a fractal pattern (varying from 0 to 1), and the object phase contains two singular points (varying from 0 to 2π). Both the irregular and the soft edges in the magnitude and the singularity in the phase have been thought to be difficult for existing phase retrieval algorithms. In Fig. 3, one can see that the recovered amplitude is almost indistinguishable from the original one, while the recovered phase differs from the original one only up to a constant offset. In this simulation, four diffraction patterns quantized to 256 levels are used. To qualify the accuracy of this method, the signal-to-noise ratio (SNR), defined as

$$\sum_{m,n} |U(m,n)|^2 \Bigg/ \sum_{m,n} [|\hat{U}(m,n)| - |U(m,n)|]^2, \tag{2}$$

is used as a measure, where $\hat{U}(m,n)$ and $U(m,n)$ represent the original wave front and the recovered wave front before the plate, respectively. From the simulation, it is found that the convergence curve consists of three phases. In the first

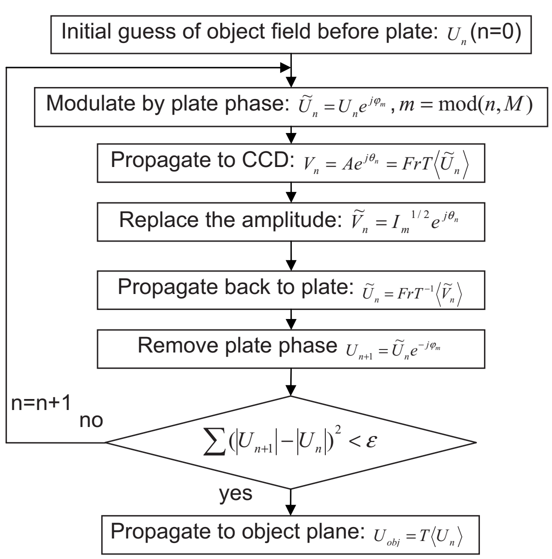

FIG. 2. Procedure for the wave-front-modulation-based phase retrieval algorithm. n is the iteration number; M the total number of recordings; I_m the mth recorded intensity pattern; FrT: the Fresnel propagation operator; T: the beam propagation operator.

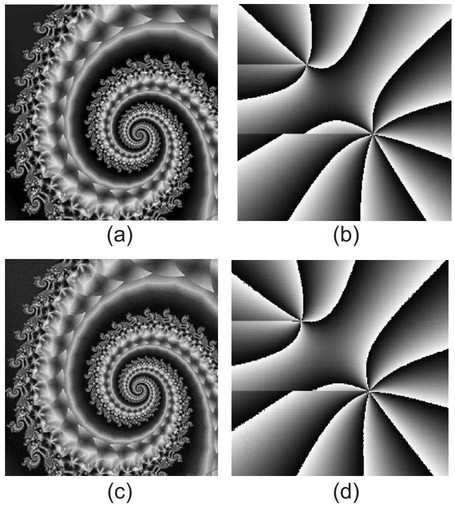

FIG. 3. Simulation results of the wave-front-modulation-based phase retrieval method for a complex-valued field; (a) and (b) are original amplitude and phase; (c) and (d) are retrieved amplitude and phase.

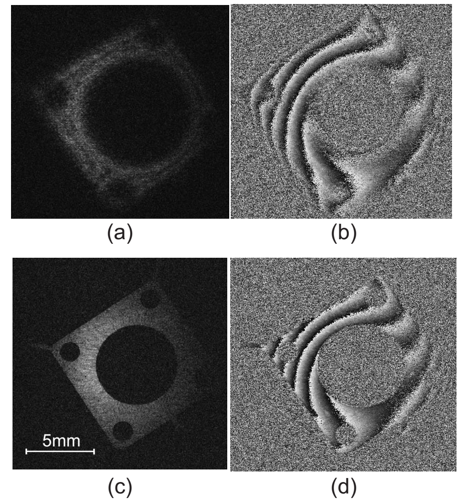

FIG. 5. Experiment results of the wave-front-modulation-based phase retrieval algorithm. (a) and (b) are the retrieved amplitude and phase at the plate plane; (c) and (d) are the retrieved object amplitude and object phase.

phase, the algorithm converges slowly, and it may last from tens to thousands of iterations depending on the number of recordings used. When an improved estimate has been found, i.e., the SNR is greater than 5, the algorithm enters the second phase. In this phase, the algorithm converges rapidly to a final status within a few tens of iterations. In the final phase, the SNR fluctuates slightly with a period equal to the number of recordings. The number of iterations in the first phase can be reduced dramatically by using the recorded intensity directly as the constraint. This is understandable because the intensity has sharper transition areas than the amplitude. A sharp transition on the CCD plane implies more high-frequency components on the plate plane. Therefore, it is more likely to push the algorithm into the "right convergence route." However, more investigation is required to clarify this observation. The convergence curves for different numbers of recordings are shown in Fig. 4. The object is the same as in Fig. 3. The number of iterations in which the intensity is used as constraint is 30, 95, and 1800, respectively, when the number of recordings is 5, 4, and 3. For the rest of the

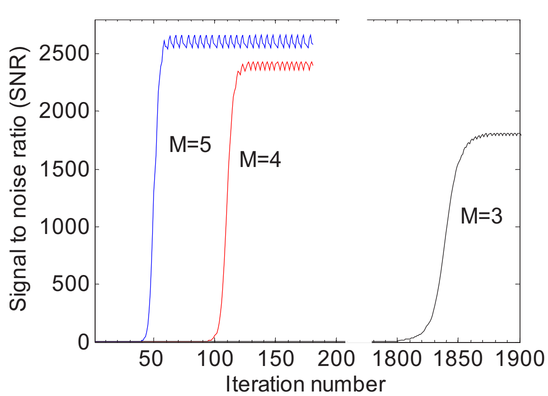

FIG. 4. (Color online) Convergence performance of wave-front-modulation-based phase retrieval algorithm. M is the number of recordings.

iterations, the modulus is used instead. If only the modulus were used as the constraint from the beginning, the iteration number to get a SNR of 100 would be 94, 600, and 5043 for 6, 5, and 4 recordings, respectively. The larger the number of recordings the algorithm uses, the more robust and faster it will be. When the number of recordings is larger than 3 ($M \geqslant 3$), it is always possible to recover the phase. The observation agrees with the theoretical expectation because the number of unknowns in a complex object field is twice larger than the number of samplings in one recording [4]. Although it is possible in theory to do phase retrieval using only two intensities, it is not feasible because any distortions in the recordings or errors in the calculation will make the iteration stagnate. Therefore, three recordings is the minimum requirement for the proposed approach to work. In fact, the number of recordings here has the same role as the oversampling factor in Ref. [6]. The computational complexity is $4N^2 \log(N)$ for an iteration in the proposed approach, where N is the effective number of samplings of the retrieved object fields. In comparison, it is $4\rho N^2 \log(N\rho^{1/2})$ in Fienup algorithms for the same number of effective object samplings, where ρ is the oversampling factor. Generally, the proposed approach is ρ times faster than the Fienup algorithm for an iteration.

An experiment has also been conducted to verify this approach. A fabricated phase plate was used in the experiment. The phase plate consisted of 1200×1400 pixels with a size of 8 μm^2. The light source was a He-Ne laser with a wavelength of 632.8 nm. The 8-bit, uncooled camera had 1300 $\times 1030$ pixels with a size 6.7 μm^2. The recorded diffraction patterns were cropped to 1024×1024 pixels in calculation, and thus the distance Z_2 was calculated to be 86.7 mm. The object field was generated by illuminating a binary photographic slide with a spherical wave front emerging from a single-mode fiber. Figure 5 shows the recovered amplitude and phase, at the plate plane and at the object plane, after 80 iterations from five recordings. After 60 iterations, the amplitude pattern started to become recognizable and no obvious improvement was observed after 75 iterations. Different initial guesses (constant or random) have also been tested. The convergence rate differs slightly. Measurement of the plate profile using a confocal microscope showed that the plate had slow and irregular transition areas between pixels. The diffraction from those areas could be one source for the residual noise in the reconstruction. Integrating a pinhole array with the plate could eliminate this effect. Other reasons for the residual noise could be an imperfect alignment of the plate with respect to the CCD as well as noise from the laser and the camera.

In the above simulation and experiment, a pure phase plate is used as wave-front modulator. In general, a plate with a complex modulation property will also work. A pure phase plate, however, gives the best energy efficiency. Here, we recorded diffraction patterns by translating the plate to different positions. If a spatial light modulator (SLM) were adopted, several uncorrelated phase maps could also be used. A system using a SLM would be more compact and flexible. In our experiment, the plate has 16 phase levels. But simulation shows that a plate having a binary phase shift, 0 and π, works as well. A plate is required in the approach, but the requirement on its quality is not as strict as that for a Fresnel zone plate or a refractive lens. In Fresnel-plate-based transmission x-ray microscopy, the resolution is determined by the width of the outermost zone, and several tens of nanometers are required to get a good lateral resolution. This requirement is at the cutting edge of current fabrication technology. In the proposed approach, the plate pixel pitch could be set to a value available by current fabrication capability simply by selecting a larger distance between the CCD and the plate. In a high-resolution microscope, a lens having an aspherical surface is required for good image quality. This has proved to be difficult to fabricate and to calibrate. A phase plate, especially one with two phase levels, would be much easier to fabricate and to calibrate than a zone plate and aspherical lens.

In conclusion, we have proposed and demonstrated a technique for the phase retrieval of any complex-valued field by using wave-front modulation and the measurement of three or more diffraction patterns. The technique combines the essence of the iterative phase retrieval technique with that of the wave-front sensing technique. Meanwhile, it removes the smooth curvature requirement in wave-front sensing and solves the stagnation problem of current phase retrieval methods when they are used for complex objects. Experimental results confirm that this technique is a practical method for lensless microscopy. Potential future directions of this research include the optimization of the phase plate design, the evaluation of the system performance, and the development of a strategy for accelerating the convergence. It is believed that the simple and robust technique would greatly improve practical phase imaging and find applications in wave-front sensing, and metrology for a wide spectral range.

The authors thank Thomas Schoder for the plate fabrication and D. Wang, R. Groves, and Y. Fu for valuable discussions. This work is support by DFG (Grant No. OS 111/19-1).

[1] D. Gabor, Nature (London) **161**, 777 (1948).
[2] I. McNulty *et al.*, Science **256**, 1009 (1992).
[3] S. Eisebitt *et al.*, Nature (London) **432**, 885 (2004).
[4] R. W. Gerchberg and W. O. Saxton, Optik (Stuttgart) **35**, 237 (1972).
[5] J. R. Fienup, Opt. Lett. **3**, 27 (1978).
[6] J. Miao, D. Sayre, and H. N. Chapman, J. Opt. Soc. Am. A **15**, 1662 (1998).
[7] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, Nature (London) **400**, 342 (1999).
[8] J. Miao, C. C. Chen, C. Song, Y. Nishino, Y. Kohmura, T. Ishikawa, D. Ramunno-Johnson, T. K. Lee, and S. H. Risbud, Phys. Rev. Lett. **97**, 215503 (2006).
[9] M. A. Pfeifer, G. J. Williams, I. A. Vartanyants, R. Harder, and I. K. Robinson, Nature (London) **442**, 63 (2006).
[10] J. R. Fienup, J. Opt. Soc. Am. A **4**, 118 (1987).
[11] J. R. Fienup and A. M. Kowalczyk, J. Opt. Soc. Am. A **7**, 450 (1990).
[12] H. M. L. Faulkner and J. M. Rodenburg, Phys. Rev. Lett. **93**, 023903 (2004).
[13] J. M. Rodenburg, A. C. Hurst, and A. G. Cullis, Ultramicroscopy **107**, 227 (2007).
[14] J. M. Rodenburg, A. C. Hurst, A. G. Cullis, B. R. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson, Phys. Rev. Lett. **98**, 034801 (2007).
[15] K. A. Nugent, A. G. Peele, H. N. Chapman, and A. P. Mancuso, Phys. Rev. Lett. **91**, 203902 (2003).
[16] W. McBride, N. L. O'Leary, and L. J. Allen, Phys. Rev. Lett. **93**, 233902 (2004).
[17] F. Zhang, I. Yamaguchi, and L. P. Yaroslavsky, Opt. Lett. **29**, 1668 (2004).