Iterative image reconstruction algorithms using wave-front intensity and phase variation

H. M. Quiney and K. A. Nugent

School of Physics, University of Melbourne, Victoria 3010, Australia

A. G. Peele

Department of Physics, La Trobe University, Bundoora, Victoria 3086, Australia

Received November 9, 2004

Iterative algorithms that reconstruct images from far-field x-ray diffraction data are plagued with convergence difficulties. An iterative image reconstruction algorithm is described that ameliorates these convergence difficulties through the use of diffraction data obtained with illumination modulated in both intensity and phase. © 2005 Optical Society of America

OCIS codes: 100.3010, 100.3190, 100.5070.

The extraction of detailed structural information from x-ray diffraction data forms the basis of crystallography and drives the continuing development of new materials and pharmaceuticals. Since many biomolecules do not form mesoscopic crystals, increasing efforts have been made to determine structural information from nanoscale aperiodic samples using phase retrieval algorithms, and there have been recent suggestions that single molecules may be imaged with diffraction data obtained from free-electron laser sources.[1,2]

The iterative Gerchberg–Saxton (GS) scheme[3] and the hybrid input–output elaborations devised by Fienup[4,5] serve as prototypes of the phase retrieval algorithms used to extract structural information from diffraction data. Implicit in these algorithms is the assumption that the incident radiation is of uniform intensity and constant phase and that the target object is represented by a complex transmission function $\xi_o(\mathbf{x})$ that is wholly contained within some known support S. The characteristic function $\xi(\mathbf{x})$ that describes the diffraction of a weakly scattering planar object wholly enclosed within S under specified experimental conditions is $\xi(\mathbf{x})=\xi_s(\mathbf{x})\xi_o(\mathbf{x})$, where $\xi_s(\mathbf{x})$ is a unit step function within S. The function $\xi(\mathbf{x})$ is associated with a wave field $F(\mathbf{k})=\hat{\mathcal{F}}[\xi(\mathbf{x})]$, where $\hat{\mathcal{F}}$ denotes the Fourier transform operator; the reciprocal relation is $\xi(\mathbf{x})=\hat{\mathcal{F}}^{-1}[F(\mathbf{k})]$, where $\hat{\mathcal{F}}^{-1}$ denotes the inverse of $\hat{\mathcal{F}}$.

The iterative algorithms that are described as being projections onto convex sets can be defined within a common framework by use of an operator formalism. The iterative procedure is initiated with a trial characteristic function $\xi^{(0)}(\mathbf{x})$. The estimate of this function after k iterations, $\xi^{(k)}(\mathbf{x})$, is obtained with the recursion

$$\xi^{(k)}(\mathbf{x}) = \hat{\mathcal{T}}\xi^{(k-1)}(\mathbf{x}), \qquad \text{for } k = 1,2, \ldots , \tag{1}$$

where $\hat{\mathcal{T}}$ is an operator of the form $\hat{\mathcal{T}}=\hat{\mathcal{T}}_1\hat{\mathcal{T}}_2\cdots\hat{\mathcal{T}}_m$, with each operator $\{\hat{\mathcal{T}}_n : n=1,m\}$ imposing a constraint on $\xi(\mathbf{x})$. Common to each of these algorithms is the use of a subsidiary constraint operator $\hat{\mathcal{T}}_{\mathrm{FT}}$ that imposes the Fourier modulus constraint on $\xi(\mathbf{x})$ through $|F(\mathbf{k})|$, with the known values of the modulus of $F(\mathbf{k})$ obtained directly from intensity measurements $I(\mathbf{k})$. The iterative sequence is presumed to terminate when $\xi^{(k)}(\mathbf{x})$ satisfies predetermined acceptance criteria. In practice, the iteration may stagnate without satisfying the acceptance criteria, since the underlying assumptions regarding the convexity of the sets generated by one or more of the operators may not be justified. The rate of convergence of these algorithms is often slow, unpredictable, and sensitive to details of the form of $\xi(\mathbf{x})$.

Many iterative strategies of this type have been devised in attempts to overcome the convergence difficulties encountered in image reconstruction from far-field diffraction data by use of plane-wave illumination; reviews, formal developments, and numerical studies have been presented recently in, for example, Refs. 6–9.

In contrast with the plane-wave formulation, a representative model of a focused beam illuminating an object consists of a Gaussian intensity profile and a spherical wave front, defining a characteristic radiation function

$$\xi_{\mathrm{rad}}(\mathbf{x}) = \exp(-\mu r^2), \tag{2}$$

where $\mu=\mu_r+i\mu_i$ and r is the perpendicular distance from the symmetry axis of the beam. The precise mathematical form of this function, which depends on the experimental arrangements is unimportant for the development of the argument that follows, provided that the variation in both intensity and phase is of comparable smoothness with $\xi_{\mathrm{rad}}(\mathbf{x})$ and that it can be characterized accurately over S. Iterative phase retrieval algorithms may readily be adapted to determine directly the details of the wave field of a focused beam since the wave field in the focal plane behaves as if it were a target whose diffracted image is obtained in the detector plane. The aberration functions required to correct the optics of the Hubble Space Telescope[10] play roles that are equivalent to the characteristic illumination functions we employ here and that were obtained with conventional iterative phase retrieval algorithms.[5] It was also noted in the numerical studies reported in Ref. 10 that the accuracy obtained from iterative phase retrieval techniques was significantly improved with out-of-focus images or, equivalently, quadratically phase-modified images. Here we explore a related approach in the context of noncrystalline phase retrieval.

The characteristic function that defines the mapping between the object plane and the diffraction intensity data is $\xi(\mathbf{x})=\xi_o(\mathbf{x})\xi_s(\mathbf{x})\xi_{\mathrm{rad}}(\mathbf{x})$. A new partition of the available information may be made by defining the multiplicative curvature operators

$$\hat{\mathcal{T}}_\mu = A\exp(-\mu_r r^2)\exp(-i\mu_i r^2), \tag{3}$$

$$\hat{\mathcal{T}}_\mu^{-1} = A^{-1}\exp(\mu_r r^2)\exp(i\mu_i r^2), \tag{4}$$

where $\hat{\mathcal{T}}_\mu$ applies a Gaussian intensity profile of maximum amplitude A and spherical curvature, and $\hat{\mathcal{T}}_\mu$ performs the inverse operation.

We can use this information about the illumination to assist in the extraction of the only information that is of immediate interest, which is $\xi_o(\mathbf{x})$. We achieve this by defining a new Fourier modulus constraint operator $\hat{\mathcal{T}}_{\mathrm{FT},\mu}$ of the form

$$\hat{\mathcal{T}}_{\mathrm{FT},\mu} = \hat{\mathcal{T}}_\mu^{-1}\hat{\mathcal{T}}_{\mathrm{FT}}\hat{\mathcal{T}}_\mu, \tag{5}$$

which replaces $\hat{\mathcal{T}}_{\mathrm{FT}}$ wherever it appears in the GS algorithm; similarly $\xi_o^{(n)}(\mathbf{x})$ replaces $\xi^{(n)}(\mathbf{x})$ in Eq. (1).

This operator imposes the Fourier modulus constraint with data obtained with known incident intensity and phase variation but applies any other constraints to the current estimate of the unmodified characteristic object function $\xi_o^{(n)}(\mathbf{x})$. The influence of applied intensity and curvature variation on the determination of $\xi_o(\mathbf{x})$ may be understood by considering parameterized autocorrelation function

$$g_\mu(\mathbf{r}) = \exp(-\mu^* r^2)\int \exp(-2\mu_r x^2) \times\exp(-2\mu^*\mathbf{r}\cdot\mathbf{x})\xi_o(\mathbf{x})\xi_o(\mathbf{x}+\mathbf{r})\mathrm{d}\mathbf{x}, \tag{6}$$

which reduces to $g_0(\mathbf{r})$, the autocorrelation function of $\xi_o(\mathbf{x})$, in the limits $\mu_r \to 0_+$ (constant intensity) and $\mu_i \to 0$ (constant phase).

For illumination that has constant phase perpendicular to the propagation direction and a Gaussian intensity distribution, $\mu_r > 0$ and $\mu_i = 0$, any variation in intensity may be wholly absorbed into a redefinition of $\xi_o(\mathbf{x})$, so that the assumption that $\xi_o(\mathbf{x})$ is real requires that $g_{\mu=\mu_r}(\mathbf{r})$ is also real, and no new information can be obtained. Nevertheless, we have incorporated this feature into the algorithm because variation of intensity over the dimension of $\xi_o(\mathbf{x})$ is a likely consequence of the presence of significant phase curvature.

In the case of uniform intensity and nonvanishing phase variation, $\mu_r = 0$ and $\mu_i \neq 0$, the autocorrelation function is

$$g_{\mu=i\mu_i}(\mathbf{r}) = \exp(i\mu_i r^2)\int \exp(i2\mu_i\mathbf{r}\cdot\mathbf{x})\xi_o(\mathbf{x})\xi_o(\mathbf{x}+\mathbf{r})\mathrm{d}\mathbf{x}. \tag{7}$$

The multiplicative factor $\exp(i\mu_i r^2)$ conveys no new information about $\xi_o(\mathbf{x})$ since this function simply isolates the known variation of incident phase across the object. The complex weighting term $\exp(2i\mu_i\mathbf{r}\cdot\mathbf{x})$ in the integrand, however, conveys a great deal of discriminatory information about the object that may be used to assist in image reconstruction. In the case of real $\xi_o(\mathbf{r})$ a real function $g_0(\mathbf{r})$ is transformed into a complex function $g_\mu(\mathbf{r})$, where the real and imaginary parts encode independent information about the relative orientation of vectors $\mathbf{r}$ and $\mathbf{x}$ in Eq. (6) through scalar product $\mathbf{r}\cdot\mathbf{x}$.

In the tests that follow we used the images in Fig. 1 to examine the rate of convergence of the modified algorithm. To set a length scale for μ, these images are defined to occupy a two-dimensional space $-1/2 \leqslant x \leqslant 1/2$ and $-1/2 \leqslant y \leqslant 1/2$. Since we established that $\mu_r > 0$ yields no information that is useful in image reconstruction and that it is most likely to degrade the rate of image reconstruction, we set this beam intensity parameter at $\mu_r = 4$ in all the tests, which multiplies the test image by unity at (0,0) and by a factor of 1.4×10^{-1} at $(\pm 1/2, \pm 1/2)$.

To demonstrate the efficiency of this new scheme without specific reference to physical dimensions, we characterize the phase curvature by the Fresnel number of the 128×128 pixel support, $N_F^{(s)}$, rather than μ_i. For our present purposes we define $N_F^{(s)}$ to be equal to half the number of phase oscillations across S. In terms of the scaling we adopted, the numerical values of the two quantities are related for this system by $\mu_i = 32\pi N_F^{(s)}$.

The mean-square error in the image reconstruction, $\delta^{(k)}$, of an $N\times N$ pixel image at GS iteration k is defined to be

$$\delta = \frac{1}{N^2}\left\{\sum_{i=1}^{N}\sum_{j=1}^{N}[\xi_o(\mathbf{x}_{ij}) - \xi_o^{(k)}(\mathbf{x}_{ij})]^2\right\}^{1/2}. \tag{8}$$

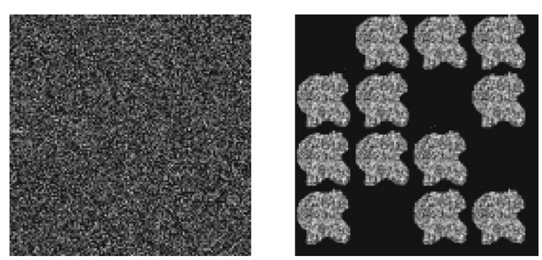

Fig. 1. Test images used in numerical studies. Both images are 128×128 pixel gray-scale objects, which in practice are zero padded to 256×256 pixels to satisfy the oversampling requirement. The left-hand image was generated by assigning each pixel a random number, n on the interval $0 < n < 1$, and is denoted $\xi_{\mathrm{random}}(\mathbf{x})$. The right-hand image consists of a number of identical, irregularly shaped cells, the interiors of which are also generated by random pixel assignment. The cells are arranged on a regular but incompletely filled grid. This image is denoted $\xi_{\text{quasi-periodic}}(\mathbf{x})$.

Table 1. Dependence of Number of Iterations m Required to Achieve an Image Reconstruction with a Mean-Square Error of 1.0×10^{-6} per Pixel on Fresnel Number $N_F^{(s)}$

$N_F^{(s)}$	0	1	2	3	4	5	6	7	8	9	10	11	12
$m_{\text{random}}{}^{a,b}$	—	1227	798	422	271	190	125	101	79	66	60	56	55
$m_{\text{quasi-periodic}}{}^{b,c}$	—	477	270	255	157	118	96	88	90	78	78	81	81

[a] Reconstruction of ξ_{random}.
[b] The conventional GS scheme, corresponding to uniform phase limit $N_F^{(s)}=0$, leads to a stagnated iterative scheme without successful image reconstruction.
[c] Reconstruction of $\xi_{\text{quasi-periodic}}$.

The results in Table 1 record the number of iterations of the phase-modified GS algorithm that are required to achieve a reconstruction of the test images with $\delta=1.0\times10^{-6}$, for which the reconstructions are indistinguishable to the eye from the original images. A striking feature of the results reported in Table 1 is that the introduction of phase curvature is sufficient to transform the classic GS scheme from one that frequently stagnates into one that usually converges rapidly. Apart from the modification of the Fourier modulus constraint operator within the GS algorithm to be of the form of Eq. (5), no further change is required.

In cases in which convergence of the modified algorithm is rapid, the phase curvature overlay on the object, typically involving $N_F^{(s)}>5$, becomes the dominant feature, imposing variable-range oscillatory structures on the object that are included exactly within the reconstruction algorithm through the use of Eq. (5). This additional information is sufficient to give the apparently random structures within each image a unique identity, fixing both the relationship between pixels and the absolute position of pixels with respect to the underlying phase and intensity perturbations. This procedure rapidly filters from the large number of symmetry-equivalent structures a single function $\xi_o(\mathbf{x})$ that is consistent both with the data $I(\mathbf{k})$ and the known phase and intensity variation characterized by μ.

Assuming the validity of Gaussian optics,[11] representative estimates can be made of critical parameters that describe illumination of a sample by a focused synchrotron beam. For a Gaussian beam of waist width w_0 and wavelength λ we define $x=z/z_R$, where z is the distance from the focal plane in the axial direction of the beam and $z_R=\pi w_0^2/\lambda$ is the Rayleigh length. The width of the beam in the plane perpendicular to the beam axis at z is $w(z)=w_0[1+(z/z_R)^2]^{1/2}$, and phase variation $\Delta\varphi$ is given by $\Delta\varphi=2x$. It is reasonable to assume that $w_0=10$ nm will be achievable at a wavelength $\lambda=0.1$ nm within a few years, with $w_0=50$ nm available at present. A Fresnel number $N_F=5$ corresponds to $\Delta\varphi=10\pi$ rad and fixes the object diffraction plane at $z\simeq49\ \mu$m. Within these parameters, $z_R\simeq3.1\ \mu$m and the effective width of the illumination is $w(5\pi z_R)\simeq150$ nm. These parameters define a focused Gaussian beam that will illuminate a sample within a spatial support of approximately 150 nm diameter. This provides sufficient phase curvature over the dimension of the sample for the significant benefits of the phase retrieval algorithm described in this Letter to become readily apparent, and it represents an achievable experimental configuration.

This study suggests strongly that for characteristic functions that are not intrinsically more complicated than the ones considered here, accurate, rapid, and unambiguous reconstruction of two-dimensional nanoscale objects may be possible using intense, focused beams of radiation and a single diffraction data set. Although the success of this approach depends on the availability of diffraction data that are not overwhelmed by noise and on the ability to characterize accurately both the intensity and the phase variation of the incident radiation over the sample, several technological opportunities immediately present themselves as warranting further careful investigation of this approach. These include the development of nanocrystallography in which two-dimensional films replace three-dimensional crystal samples that are difficult or impossible to obtain in many important instances.

The Australian Research Council is acknowledged for grant support and for the award of an Australian Research Council Federation Fellowship (K. A. Nugent) and a Queen Elizabeth II Fellowship (A. G. Peele). H. M. Quiney's e-mail address is quiney@unimelb.edu.au.

References

1. J. Miao, H. N. Chapman, J. Kirz, D. Sayre, and K. O. Hodgson, Annu. Rev. Biophys. Biomol. Struct. **33**, 157 (2004).
2. J. Miao, K. O. Hodgson, and D. Sayre, Proc. Natl. Acad. Sci. U.S.A. **98**, 6641 (2001).
3. R. Gerchberg and W. Saxton, Optik (Stuttgart) **35**, 237 (1972).
4. J. R. Fienup, Appl. Opt. **21**, 2758 (1982).
5. J. R. Fienup, Appl. Opt. **32**, 1737 (1993).
6. V. Elser, J. Opt. Soc. Am. A **20**, 40 (2003).
7. H. H. Bauschke, P. L. Combettes, and D. R. Luke, J. Opt. Soc. Am. A **19**, 1334 (2002).
8. S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. H. Spence, Phys. Rev. B **68**, 140101 (2003).
9. K. A. Nugent, A. G. Peele, H. N. Chapman, and A. P. Mancuso, Phys. Rev. Lett. **91**, 203902 (2003).
10. J. R. Fienup, J. C. Marron, T. J. Schulz, and J. H. Seldin, Appl. Opt. **32**, 1747 (1993).
11. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (Wiley, 1991), pp. 81–89.