Phase retrieval algorithms: a personal tour [Invited]

James R. Fienup

The Institute of Optics, University of Rochester, Rochester, New York 14627, USA
(fienup@optics.rochester.edu)

Received 28 September 2012; accepted 15 October 2012;
posted 4 December 2012 (Doc. ID 181961); published 21 December 2012

This paper gives the reader a personal tour through the field of phase retrieval and related works that lead up to or cited the paper "Phase Retrieval Algorithms: a Comparison," [Appl. Opt. **21**, 2758 (1982)]. © 2012 Optical Society of America

OCIS codes: 100.5070, 100.3010, 010.7350, 110.0110, 030.0030, 070.0070.

1. Introduction

This is not your usual paper for Applied Optics, being neither just the latest results in, nor a comprehensive review of, a particular field. Solicited by Editor-in-Chief Joseph Mait, it is in part a retrospective paper about a personal journey in the field of phase retrieval, in the context of the 50th anniversary celebration of Applied Optics. He solicited the present paper because the paper "Phase Retrieval Algorithms: a Comparison" (henceforth: "the 1982 Applied Optics paper") [1] had, as of early 2012, received over 1,350 citations (Thompson Reuters' Web of Science), making it the fourth most cited paper in *Applied Optics*. Its annual rate of citation accelerated in the last decade (to over 100 per year) rather than the usual deceleration of citations that occurs for a paper of its vintage. Section 2 of the present paper describes the historical events that lead up to the 1982 Applied Optics paper, briefly summarizes the contents of that paper, and speculates on the reasons for its high citation rate. Section 3 describes some of the many fields within optics and other disciplines of the papers that cited the 1982 Applied Optics paper. Section 4 mentions some of the significant enhancements to the phase retrieval algorithm, including the first exposition of the "continuous" version of the hybrid input–output (HIO) algorithm.

1559-128X/13/010045-12$15.00/0

2. How It Came to Be

A. Computer Holography, Kinoforms, and Iterative Algorithms at Stanford

This is the most personal part of the story. As a graduate student in the Applied Physics Department of Stanford University, my thesis [2] advisor was Prof. Joseph W. Goodman in the Electrical Engineering Department, and he had a grant to study computer-generated holograms (CGHs) as a way to optically store data. Storing data optically in the Fourier domain rather than in the image domain or the bit domain, as is done now on the CDs and DVDs that we have all used for music and movies, has some distinct advantages. Slight motion of a Fourier transform hologram does not change the intensity of the image, making it tolerant to positioning the readout beam; small defects in a Fourier hologram cause added noise in the image rather than the complete dropping of a bit, making it error-tolerant; and the illumination of one hologram can create a whole two-dimensional (2D) array of spots in the image plane that can be read out simultaneously, making it a page-oriented as opposed to a bit-oriented memory. Since I was an amateur photographer and knew how things worked in the photographic darkroom, I was given the task of making some CGHs. At the time (early 1972) I was simply photographing pictures of holograms out of the book *Optical Holography* (see [3], Figs. 19.8 and 19.11), which included a Lohmann and Paris CGH [4] and a kinoform [5] (at that time Goodman's *Introduction to Fourier Optics*

[6] had no such pictures, although the present edition [7] does), trying a variety of photographic positive and negative materials. Bleaching the developed transparencies made phase holograms and kinoforms. I was intrigued by an accidental result: while processing a batch of Kodak 5254 color film, I accidentally exposed it to room light before development, and the resulting transparency made a reasonable kinoform without bleaching. That gave me the idea of using a red transparency film (for use with a He–Ne laser) for kinoforms instead of bleaching film, which was inconsistent and time-consuming. Trying different materials, I found that Kodachorome II gave a large phase effect, which one could see by eye by looking at its surface. Kodachrome II was a marvelous material, loved by photographers for its fine resolution and wonderful color reproduction, and celebrated in song [8]. To make a pure-phase CGH for illumination with, say, red He–Ne laser light, one would expose Kodachrome II to uniform, bright red light and to a desired pattern of blue/green light so that after processing it would be transparent to red light (with little or none of the red-absorbing dye) and have a desired thickness variation owing to the blue- and green-absorbing layers. An extra bonus was that there was a Kodak processing plant next to campus that could process the film overnight. After that discovery, I viewed Kodachrome II as a very easy way to produce kinoforms without the messy, smelly process of bleaching photographic film or plates.

Kinoforms [5] are CGHs that are phase-only transparencies that operate on-axis, i.e., without a carrier spatial frequency. They directly impose phase on a transmitted wavefront by virtue of their spatially varying thickness or bulk index of refraction (depending on the material). This makes them highly efficient: 100% of the light can go into the desired order of diffraction since there is only one order, the zeroth order. A problem with kinoforms, however, is that they can only impart a phase to the wavefront, whereas a field that will propagate to form an image of a desired object will vary in amplitude as well as phase. The amplitude of the Fourier transform of a real, nonnegative object is very bright on axis and very dim elsewhere, making the kinoform approximation very poor. This problem can be greatly reduced by assigning a quasi-random phase pattern to the object, which will result in a desired hologram field that is much more uniform, hence better, although still not uniform enough [9]. A way to adjust the quasi-random phase of the object to improve the uniformity of its Fourier transform was by an iterative algorithm invented by Gallagher and Liu [10], although they later found out [11] that it had been invented even earlier by the inventors of the kinoform [12]. What they developed was a way to perform spectrum shaping, i.e., how to assign a phase to an object or a field such that its spectrum (the intensity of its Fourier transform) would approach a desired distribution. This is an example of a synthesis problem. They also found out [11] that their algorithm was very similar to the Gerchberg–Saxton algorithm for solving a phase retrieval problem in electron microscopy [13]. These similar algorithms worked as follows. Suppose a function $f(x,y)$ has Fourier transform $F(u,v)$ (or they are related by some other propagation integral), and we are given their moduli (amplitudes) $|f(x,y)|$ and $|F(u,v)|$, where the moduli are either the square roots of measured field intensities or the square roots of desired image and hologram intensities. The algorithm is to start with a guess for the field in the object domain, typically assigning random numbers for the phase $\phi(x,y)$ of $g(x,y) = |f(x,y)| \exp[i\phi(x,y)]$, which has the measured or desired modulus. Then perform the following four steps: (1) transforming that field to the Fourier domain to give the field $G(u,v) = |G(u,v)| \exp[i\theta(u,v)]$, (2) changing that field to give it the measured or desired modulus in that domain, making it $G'(u,v) = |F(u,v)| \exp[i\theta(u,v)]$, (3) then inverse Fourier transforming that back to the object domain to give the image $g'(x,y) = |g'(x,y)| \times \exp[i\phi'(x,y)]$, and (4) then give it the measured or desired modulus, resulting in a new version of $g(x,y) = |f(x,y)| \exp[i\phi(x,y)]$ where the earlier version of $\phi(x,y)$ is replaced by $\phi'(x,y)$, completing one iteration. These four steps are repeated until no further progress is made or for a fixed number of iterations. This kind of iterative algorithm is widely referred to as the Gerchberg–Saxton algorithm, even though the patent [12] came first. (Publications in archival journals are more accessible than patents.)

The Gallagher and Liu paper (the only one of the iterative algorithm papers of which I was aware in late 1973) inspired me to try out their algorithm for kinoforms, to seek improvements in the algorithm, to control quantization errors in other CGHs, and to use it for general spectrum shaping [14]. For reducing quantization errors in kinoforms and CGHs, however, I found that, while it would reduce the standard deviation of the error in the image of a bit pattern, the maximum error for any single bit could remain undesirably large. To keep the bit error rate small, it was important to reduce the errors of those outliers. For what follows, we refer to the function $g(x,y)$ above as the input to an iteration of the algorithm and $g'(x,y)$ after the inverse Fourier transform as the output. The key to reducing the error of the outliers is the observation that if the input $g(x,y)$ gives rise to the output $g'(x,y)$, then if we make a small change in the input, replacing it by $g(x,y) + \Delta g(x,y)$, where $\sum_{x,y} |\Delta g(x,y)|^2 \ll \sum_{x,y} |g(x,y)|^2$, then the expected value of the new output image is $g'(x,y) + \alpha \Delta g(x,y)$ plus some additional noise [14], where α is a constant that depends on the statistics of $|F(u,v)|$ and $|G(u,v)|$. Suppose that an image pixel that was supposed to be zero was a positive number. Then to drive it toward zero, one could subtract from the current input α^{-1} times the positive number, and the new output would be closer to zero at that pixel. This is the input–output point of view. In it, we think of the three operations—Fourier transforming, giving the result the desired Fourier modulus, and then inverse transforming—as a nonlinear system. Small enough changes to the input to this nonlinear system result in closely related small changes to the output of the system. That is like saying that a nonlinear function can be approximated as linear one over a small enough range of input values. Note that, when this is done, the new input no longer has the desired modulus at that pixel like it does for the other iterative algorithms, so the new input image is no longer an estimate of the object.

During the time that this work on iterative algorithms was being done (1973–1974), Prof. Goodman was at the Institut d'Optique in Orsay, France, on a year-long sabbatical. After returning, he recommended to me: see if an iterative algorithm like this can also solve the phase retrieval problem. Reconstructing astronomical images from intensity interferometry data [15] or from stellar speckle interferometry data [16] was of particular interest.

The phase retrieval problem, as found in x-ray crystallography, astronomical imaging, Fourier transform spectroscopy and some other fields, is different from that found in electron microscopy or spectrum shaping. One typically has $|F(u,v)|$ from measurements, but instead of knowing $|f(x,y)|$, one typically has much weaker information, often knowing, for example, that f is real-valued and nonnegative, $f \geq 0$, and knowing something about the support of f, i.e., knowing that f must be zero outside of some region, the support constraint. Reconstructing f from $|F(u,v)|$ (and some constraints on f, such as support and nonnegativity) is known as the phase retrieval problem. It is so named because if the true Fourier transform is $F(u,v) = |F(u,v)| \exp[i\psi(u,v)]$, then reconstructing f is equivalent to retrieving the phase ψ from $|F|$ (and some constraints on f), since once one has ψ, one can simply inverse Fourier transform $F(u,v) = |F(u,v)| \exp[i\psi(u,v)]$ in a computer to get $f(x,y)$. The phase retrieval problem had been studied for many years, but no practical algorithm for solving it had been invented. Furthermore, the prospects seemed dim because it was understood that in the case of one-dimensional (1D) functions there were ordinarily a large number of other functions having the same Fourier modulus; that is, the solution would not be unique, as studied for spectroscopy [17] and astronomy (although just in one dimension) [18].

I attempted using the iterative algorithm, including using the input–output approach, for solving the phase retrieval problem. Unfortunately, at that same time sufficient funds to use Stanford University's mainframe computer were running thin, so it was suggested that I instead use a new Interdata minicomputer that was available in the Electrical Engineering Department for no extra cost. That computer, however, was extremely slow and had very little memory as compared with the mainframe. For that reason the phase retrieval experiments were limited to 1D functions. Constraints used were a noise-free Fourier modulus, a support constraint (the width of the object was assumed known), and a nonnegativity constraint. The results are shown in Fig. 2 of [19]. While there was considerable variation in the various reconstructed images, the general shape of the object could be discerned, despite the known lack of uniqueness for this 1D phase problem. Presumably the nonnegativity constraint helped to prevent the solutions from being wildly different. These 1D results were not of sufficient quality for journal publication, but were included in a subsubsection of my dissertation [2], which was mostly about the Referenceless On-Axis Complex Hologram (ROACH) made with Kodachrome II film, which allowed on-axis amplitude control as well as the phase control as from a kinoform [20].

B. Early Phase Retrieval at ERIM

My first job after graduating was at what was then known as the Environmental Research Institute of Michigan (ERIM), which was an independent, not-for-profit research institute, which not long before had been the Willow Run Laboratories of the University of Michigan. It was a wonderful hotbed of advanced research in various forms of reconnaissance (radar and optical), holography, and other fields. (It was bought by Veridian Systems by the time I departed for academia 27 years later, and has since become part of General Dynamics—Advanced Information Systems.) While initially working on projects having to do with such things as automatic focusing for synthetic-aperture radar, optical processing for radio interferometer data, and diffractive optics, I applied for internal research and development (IR&D) funding to continue the phase retrieval work I had started at Stanford. Such funding was precious and my proposal was turned down; the topic was deemed not sufficiently important. Then in the fall of the following year I attended the OSA Annual Meeting in Tucson, where a talk by Frieden and Currie [21] on a phase retrieval algorithm, even though it only worked for special cases, created a great deal of excitement in the audience. That event emboldened me to be more forceful in my request for IR&D funds, which were granted the next year.

At ERIM I repeated the digital experiments performed at Stanford, but with 2D objects, which our PDP 11/45 computer could handle, with the help of a special processor for performing fast Fourier transforms (FFTs). I tried several variations on the input–output algorithm applicable to the image reconstruction problem in astronomical interferometric imaging that used different ways of modifying the input image (the fourth step of the iterative algorithm) based on the current output image and the previous input image. From the Gerchberg–Saxton point of view, the next input image would be the output image, but modified to satisfy the object-domain constraints of support and nonnegativity, i.e., set the next input image to the output image where the output image satisfies the constraints, and set it to zero where the output image violates the constraints. This is referred to as the "error-reduction" approach. The "input–output" approach would be to set the next input image to the previous input image where the output image satisfies the constraints, and set it to the previous input image minus a constant times the output image where the output image violates the constraints. Another approach was based on the interesting property of this nonlinear system, that if one takes an output of the system and uses that as a new input, then the new output is the same as the new input: the system leaves it unchanged because the new input already satisfies the Fourier-domain constraint. Hence, no matter what input produces a given output, one can pretend that the output resulted from itself as an input. The "output–output" approach, then, would be to set the next input equal to the output where the output satisfies the constraints, and set it to the output minus a constant times the output where the output violates the constraints. I also tried other variations, and tried mixing and matching different operations from different approaches to handling the values where the output image either satisfies or violates the constraints. This was not the beautiful mathematics of an Einstein that predicted what would happen long before an experiment was performed; this was the trial and error approach that Edison used to invent a practical light bulb: keep trying different things (guided by physics, mathematics, and intuition) until you find something that works; and then refine that. There *is* beautiful mathematics surrounding the phase retrieval problem, and it is centered around the zeros of the Fourier transform analytically extended to the complex plane [17,18,22,23], for example]; but that beautiful mathematics had yielded no practical phase retrieval algorithms. More than one of the variations I tried worked to a degree; but the best one, both in the 1D experiments at Stanford and the 2D experiments at ERIM, was the "HIO" algorithm. The HIO approach is to set the next input equal to the output where the output satisfies the constraints, and set it to the input minus a constant times the output where the output violates the constraints; it is a hybrid between the output–output (first line) and input–output (second line) approaches. It is given by Eq. (44) in Ref. [1], and was described in words in terms of Eq. (2) and Eq. (5) in Ref. [19],

$$g_{k+1}(x,y) = \begin{cases} g'_k(x,y), & (x,y) \notin \gamma \\ g_k(x,y) - \beta g'_k(x,y), & (x,y) \in \gamma \end{cases}, \tag{1}$$

where $g_k(x,y)$ is the input at the kth iteration, $g'_k(x,y)$ is the output at the kth iteration, β is the feedback constant, and γ is the set of points for which the output violates the object-domain constraints. For a nonnegativity and support constraint, the condition $(x,y) \notin \gamma$ can be expressed as $(x,y) \in S \,\&\, g'_k(x,y) \geq 0$, where S is the set of points inside the support constraint. A value of β somewhere between 0.5 and 1.0, say 0.7, usually works well. I think of the value of β as being like the force with which one depresses the accelerator while driving a car up a narrow mountain road, seeking the top. Press too gently (e.g., β less than 0.5), and the car will make steady progress, but too slowly; press too forcefully (e.g., β greater than 1.0), and the car will progress much more quickly, but is likely to veer off a cliff and crash.

The 2D results, using support and nonnegativity constraints in the object domain, were dramatic. The reconstructed images, while not perfect, were very good representations of the object, much better than the 1D case. This implied that the 2D phase problem was unique—or else I would have been very lucky for the algorithm to have stumbled onto the solution that matched the object I had started with rather than one of the other solutions—in contrast to the nonuniqueness typical of the 1D case. Here we consider the solution to be unique if it is the same to within a translation or a 180° rotation, since those operations do not change the Fourier modulus, and one still gets the same picture of the object. These results created substantial interest at the 1977 OSA Annual Meeting in Toronto [24] and were published in Optics Letters in 1978 [19], which is the seminal paper on the approach, although its citation rate is less than half that of the 1982 Applied Optics paper.

With good image reconstruction results in hand, searching for external funding immediately commenced. One potential sponsor, skeptical of the veracity of the results, arranged for B. L. McGlamery to simulate Fourier modulus data from stellar speckle interferometry of a photograph of a satellite and send it to me in a blind test, which was a success, as described in [25,26], the latter of which won the SPIE's Rudolph Kingslake Medal and Prize for best paper in Optical Engineering that year. Papers on the effect of noise [27], results with real stellar speckle interferometry data [28], understanding the uniqueness of phase retrieval [29], estimating the support of the object from its autocorrelation function [30], and on a wide variety of applications of the iterative transform algorithm [14] were published along with summary papers on the phase retrieval algorithm [31,32], all before the 1982 Applied Optics paper. So what was new in that paper, and why is that paper so highly cited?

The 1982 Applied Optics paper [1] was an extended version of a presentation at the 1981 Annual Meeting of the OSA [33], whose main new point was that the error-reduction approach (satisfying constraints in each domain) was equivalent to a steepest-descent gradient search algorithm with a particular line-search step size, minimizing an error metric, which was the sum of squared differences between the computed Fourier modulus and the measured one. As pointed out in [1], "The relationship between a gradient search method and the error-reduction algorithm for a problem in digital filter design is discussed in Ref. 13," i.e., in [34]. [1] also gives a proof of convergence in the weak sense that the error could only decrease or stay the same at each iteration. Liu and Gallagher [11] showed this for the case of intensities in image and Fourier domains, but in [1] it was proven for the general problem of arbitrary data and constraints. It also gave the name "HIO" to the version of the most successful iterative transform algorithm, but it was the same algorithm already described in [19], which did not go into that detail on account of the necessary brevity of *Optics Letters* (one researcher using HIO admitted to me that he did not realize that it was same algorithm as described in [19]). It further showed that a conjugate gradient method was superior to a steepest-descent method. It also gave an example with helpful tips to making the algorithm work successfully. While this was all good, it probably does not deserve to have more citations than the seminal paper [19]. Some possible explanations for this are as follows. Perhaps the word "Comparison" in the title made some people who did not read it thoroughly think that it was a review paper on phase retrieval algorithms, which it was not meant to be; it was meant to compare iterative transform algorithms with gradient search algorithms. On the other hand, since the competing phase retrieval algorithms at the time were not effective for general objects, it was in effect comparing all the dominant phase retrieval algorithms. Perhaps it is because it had a great deal of mathematical detail on the algorithms, although [14] had quite a bit of detail as well. Perhaps it is because it discusses both the image reconstruction and the wavefront sensing applications, although [14] did this as well. Perhaps a few authors of papers, for these or some other reasons, referenced [1] but not [19], and subsequent authors reading these papers just referenced [1] without reading it in detail to find out that it was just one of several earlier papers [14,19,26,27,28,31,32] in which I had written about the algorithm.

That this group of papers is so highly cited partly arises from the generality of the Gerchberg–Saxton algorithm and these derivative algorithms, as was the main point of [14]:

> "This paper discusses an iterative computer method that can be used to solve a number of problems in optics. This method can be applied to two types of problems: (1) synthesis of a Fourier transform pair having desirable properties in both domains, and (2) reconstruction of an object when only partial information is available in any one domain.
>
> …
>
> Both the synthesis and the reconstruction problems can be expressed as follows:
>
> Given a set of constraints placed on an object and another set of constraints placed on its Fourier transform, find a Fourier transform pair (i.e., an object and its Fourier transform) that satisfies both sets of constraints."

This paper [14] was not just a showing of results for a few different applications, but was also a call to arms to the optics community to apply related algorithms to many different problems in different fields; yet it has received only 1/5 the citations of the 1982 Applied Optics paper.

What follows next are brief discussions of some of the application areas explored in the papers that cite [1], with what is currently the most important such application (as measured by citations) saved for last.

3. Application Areas Citing the 1982 Applied Optics Paper

In this section we mention some of the more popular applications areas resulting in citations to the 1982 Applied Optics paper. Some of them were already mentioned above: reconstruction of phase from the intensity distribution of electron beams in two planes; image reconstruction for astronomical imaging including from stellar speckle interferometry data, amplitude interferometry data, and intensity interferometry data; and various forms of spectrum shaping for CGHs, and for the reduction of quantization noise. In what follows are some additional areas. This is not meant to be a comprehensive review or even listing of all these areas, but is merely designed to give the reader an appreciation for the wide range of possibilities. The particular references cited are meant only to be examples with which I was familiar and which cited the 1982 Applied Optics paper, rather than necessarily being the first or most noteworthy papers on the topics.

A. Synthesis Problems

1. *Beam Shaping*

The iterative algorithm has been used by many for other applications of spectrum shaping, also known as beam shaping. Typically one wants a phase-only transparency or transparencies, to avoid absorption of light, to produce a desired intensity pattern of light. An example of beam shaping is the design of a phase plate that has smoothly varying phase that produces a high-power laser beam having a uniform far-field speckle distribution for use in inertial confinement fusion [35]. Three-dimensional (3D) beam shaping has also been performed [36].

2. *Optical Encryption*

One can design a diffractive optical element that allows for the encryption of data for security applications [37]. The problem involves jointly designing two diffractive elements having quasi-random phases, with finite pupils, that together reconstruct an image although neither diffractive element by itself gives any hints as to what is in the image.

3. *Optical Communication*

By designing the temporal phase of a light beam and transmitting it through a fiber, one can compensate for the dispersion of the fiber and temporally concentrate the energy at the output to give a desired bit pattern [38].

4. Antireflection Coatings

One can iteratively design antireflection coatings, where one wants certain features in a desired spectral response but has constraints on the number of layers and the indices of refraction of the multilayer structure [39]. This can be extended to designing other types of multilayer coatings as well.

B. Reconstruction Problems

1. Wavefront Sensing for Radio Antennas

By making a simple power measurement from a point source on a satellite, one can measure the far-field pattern of the telescope. That, combined with knowledge of the transverse shape of the antenna, allows the phase (i.e., longitudinal shape of the antenna) to be reconstructed with a phase retrieval algorithm [40], thereby diagnosing the deformations of the radio dish.

2. Wavefront Sensing for Optics

Similar to the problem for radio dishes, measuring a point-spread function (PSF) of an optical system, together with knowledge of the shape of the exit pupil of the optical system, allows one to determine the aberrations of the optical system with a phase retrieval algorithm. This was done for determining the aberrations due to atmospheric turbulence [41], for determining the aberrations of the flawed Hubble Space Telescope [42], and has been developed for phasing up the segments of the future James Webb Space Telescope [43], for optical metrology in the manufacture of optical surfaces [44], and for measuring x-ray focused beams [45]. With sufficient forms of measurement diversity, it is possible to reconstruct the amplitude of the pupil as well as the phase, i.e., without any prior knowledge of the pupil [46]. There are many forms of diversity, mostly modifying the phase in the exit pupil making the total phase the unknown phase plus a known phase perturbation. This makes the phase retrieval algorithms more robust to noise, to algorithm stagnation, and to the possibility of nonuniqueness and is important for applications, such as wavefront sensing for space telescopes or for optical metrology, where one usually needs a very accurate, reliable solution. The most common type of measurement diversity is focus diversity, as will be used for the James Webb Space Telescope [43,47]. Other forms of diversity include wavelength diversity, piston diversity (for segmented systems), actuator poking diversity, and transverse translation diversity (of a structure or an illumination beam relative to the object) [48,49] and diversity of field position to assess misalignments [50]. Diverse measurements are important under stressing conditions, such as when dealing with broadband or undersampled data [50,51] or when reconstructing amplitude as well as phase [46]. When one measures two planes of intensity that are very near to one another, such that one can estimate the partial derivative of the 3D intensity with respect to the axial coordinate, then one can also use what is commonly called the transport-of-intensity approach [52].

3. FROG

Frequency-resolved optical gating (FROG) determines the temporal characteristics of fast (femtosecond) laser pulses [53,54]. FROG traces, as a function of frequency and time delay, are the squared magnitude of the Fourier transform of a 2D signal, hence solvable by a phase retrieval algorithm.

4. Blind Deconvolution

When one or more blurred images have unknown PSFs, then straightforward image restoration by deconvolution is not possible. Given constraints on both the object, such as support and nonnegativity and the PSF (support and nonnegativity again), then one can use an algorithm of the iterative transform type or a gradient search algorithm, but with an additional Wiener filtering step [55].

5. Tomography

Tomographic imaging can suffer from missing projections [56], or unknown phases in optical refraction tomography [57] and diffraction tomography [58], which iterative algorithms can fill in.

6. Miscellaneous

There are also a number of citing papers in areas outside my knowledge base that find phase retrieval algorithms pertinent, for example for determining complex Ginzburg–Landau equations [59] and current distributions in Josephson junctions [60].

The most fun is Elser's application of an HIO-like algorithm to solving the Sudoku problems [61] that one can find in many daily newspapers. In its most common form, one must find numbers 1 through 9 to place in a 9×9 grid of squares, separated into 3×3 boxes of size 3×3 squares each, such that all the numbers 1 through 9 appear exactly once in each length-9 row, once in each length-9 column, and once in each 3×3 box. Although there are no Fourier transforms or phases, the problem does involve solving for systems of equations, with the constraints mentioned above as well as the constraints given by the starting numbers given in some of the squares.

7. Lensless Imaging

Holography is an approach to imaging without lenses, relying on the interference of a coherent field, reflected from or transmitted through an object and propagated to the detector plane, with a reference beam. If there is no reference beam, however, then it is still possible to reconstruct an image from a single intensity pattern of the propagated field if there are strong enough constraints on the object, such as a sharply defined support constraint [62], which can be natural to the object (a bright object on a dark background) or by virtue of an illumination pattern [63], or having a low-resolution image [64]. Alternatively, by correlating the measured speckle intensity, by the same process as Hanbury–Brown and Twiss intensity interferometry one can estimate the Fourier magnitude of the underlying intensity reflectivity of the object and with a phase retrieval algorithm reconstruct an incoherent image of the coherently illuminated object [65]. Emmett Leith once told me, in a friendly, joking way, that he called these approaches "anti-holography," since they did away with the need for holography to reconstruct complex-valued fields in some instances.

Most papers have an initial surge in citations, followed by a decay. The 1982 Applied Optics paper, in contrast, had a relatively steady citation rate (15 to 30 per year) for many years, but then starting in 2002 experienced a surge that continues to this day. This surge was a result of the final application area described next.

8. X-ray Coherent Diffractive Imaging (CDI)

A particular kind of lensless imaging is to illuminate a microscopic object with a coherent beam of x rays, measure the intensity of its far-field diffraction pattern, and reconstruct an image from that data using a phase retrieval algorithm (employing support and possibly nonnegativity constraints on the object). The phase problem in x-ray diffraction has a long history, and multiple Nobel prizes have been awarded to those who solved it for crystals. For many years it was limited to crystals, in which case the x-ray illumination beam needed to be spatially coherent only over a modest number of unit cells in the crystal. The 1982 Applied Optics paper mentioned x-ray crystallography as an application of the phase retrieval algorithm. Sayre [66] pointed out the need to sample the diffraction pattern at intervals twice as fine as the usual reciprocal-lattice (Bragg reflection) points, which by themselves give undersampled data. The ability to avoid the undersampling is trivial when a small object is not a periodic crystal structure, since there is data everywhere in the far field, rather than just at reciprocal-space points. Eventually with brighter, more-coherent x-ray beams from synchrotron radiation and more recently x-ray free-electron lasers and even "table-top" x-ray sources, it is possible to obtain bright beams with spatial coherence widths greater than a micrometer, enabling x-ray diffraction experiments to be performed on noncrystalline samples, and the field of x-ray CDI grew [67]. It was obvious that images could be reconstructed from this data, since it had been done already with longer optical wavelengths, both in computer simulation experiments [62] and with data gathered in the laboratory, even for complex-valued objects [68]. Nevertheless, it required a successful experiment with x rays [69] to galvanize the field. Reference [70] gives an excellent account of the phase retrieval approach to x-ray CDI.

4. Some Other Algorithm Developments

A number of developments in the understanding of and in potential improvements in the iterative algorithms have taken place.

Why the HIO algorithm could seemingly climb out of local minima while the error-reduction algorithm is often doomed to be trapped in the same local minima was only partly understood at the beginning. Later we were able to demonstrate conclusively that HIO could climb out of such local minima [71], although it was not guaranteed to do so [72]. Takajo and Takahashi shed light on the convergence properties of HIO [73,74,75].

Only briefly explored in the 1982 Applied Optics paper, the conjugate gradient nonlinear optimization algorithm was shown by Lane to be effective for image reconstruction [76]. Gradient-search nonlinear optimization techniques were generally found to be superior to iterative transform algorithms for the application of wavefront sensing [42,77,78], presumably because the bulk of the phase can be represented by a tens to a few hundreds of basis-function coefficients, a much smaller search space than the thousands to hundreds of thousands of unknowns represented by a point-by-point phase map. The expression of a phase function by a basis-function expansion with a moderate number of terms also serves as a regularization of the phase retrieval problem and avoids many local minima associated with nonphysical phases [42].

Having a good support constraint greatly improves the convergence of the algorithms. Approaches to reconstructing upper bounds on the support of the object from the support of its autocorrelation function (which can be computed from the given Fourier modulus data) were improved [30,79] and the "shrink-wrap" technique for refining the support constraint during the iterations was developed [70].

To avoid algorithm stagnation (being caught in a local minimum), it helps to "sneak up on the solution" when solving the phase retrieval problem. For example, in the "expanding Fourier modulus" method [64] the algorithm converges more successfully if one first reconstructs a low-resolution image from the lower spatial frequencies, then gradually adds higher spatial frequencies reconstructing a successively finer-resolution image. For wavefront sensing applications, it helps to retrieve the large, low-order Zernike coefficients first and then work up to the higher-order Zernike coefficients.

The error-reduction algorithm was known to converge in a weak sense, as discussed earlier. Later, Youla [80] set forth a formalism for a more general set of problems and the popular method of projections onto convex sets was developed [81], including a proof of convergence in a strong sense. The error-reduction algorithm is indeed a projection onto sets algorithm, but it was not described in this Hilbert-space formalism. For phase retrieval, the set of function having the same Fourier modulus is highly nonconvex and the error-reduction/projections-onto-sets algorithm works poorly for it, as described earlier.

Additional variations on the HIO algorithm were developed. It was shown that HIO could be understood in terms of the projections onto sets theory and was a special case of the Douglas–Rachford algorithm [82]. The hybrid projection–reflection (HPR) algorithm [83] is a generalization of the HIO algorithm with comparable performance. Elser [84] developed the "difference map" algorithm, which also has HIO as a special case.

Another variation is a continuous version of the HIO algorithm (CHIO). It was documented only briefly in a 1-page conference paper [85] and never discussed in an archival journal, so a fuller account is given here. CHIO was designed to overcome a drawback of HIO. With successive iterations, HIO has a tendency for the values at a given pixel to oscillate somewhat with increasing iteration number. The conjecture was that the oscillations had to do with the fact that the input image at the next iteration is a discontinuous function of the output image. Figure 1 shows the relationship between the next input value, $g_{k+1}(x)$ as a function of the previous input $g_k(x)$ and it output $g'_k(x)$ for three different versions of the iterative algorithm when employing a nonnegativity constraint for a point (x,y) inside the support constraint. In Fig. 1(a) we see that for the error-reduction algorithm, the next input is zero when the output value is negative and equal to the output value where it is nonnegative. In Fig. 1(b) we see the relationship described in Eq. (1) for HIO. There is a discontinuous relationship between the value of the next input $g_{k+1}(x,y)$ as a function of the current output $g'_k(x,y)$, depending on whether the current output value is greater than or less than zero. This discontinuity might make the algorithm more violent than what is optimal. The fourth step of the CHIO algorithm, the update of the input, is illustrated in Fig. 1(c) and is given by

$$g_{k+1}(x,y)=\begin{cases} g'_k(x,y), (x,y)\in S\,\&\,\alpha g_k(x,y)\le g'_k(x,y) \\ g_k(x,y)-\left(\frac{1-\alpha}{\alpha}\right)g'_k(x,y), 0\le g'_k(x,y)\le \alpha g_k(x,y) \\ g_k(x,y)-\beta g'_k(x,y), \text{otherwise} \end{cases}. \tag{2}$$

Other forms are possible; for example, by employing terms of higher order in $g'_k(x,y)$, one can get curves that have continuous values and continuous derivatives everywhere. The equation above has continuous values but not continuous derivatives.

To test the CHIO algorithm, the satellite model shown in Fig. 2 was used as the object. It was embedded in an array of greater than twice its width and height to ensure that the Fourier intensity is sampled better than Nyquist. Figure 3 shows the magnitude of its Fourier transform, the data from which we will try to reconstruct the object. Figure 4 shows the support constraint used for this experiment. A better support constraint could be derived from the support of the autocorrelation of the object [79], but this much looser support constraint was chosen purposely to make the reconstruction more difficult for the algorithms (otherwise the HIO algorithm reconstructed it too easily). Figure 5 shows images reconstructed from (a) HIO, using $\beta = 0.7$, and (b) CHIO, using $\beta = 0.7$ and $\alpha = 0.4$ after 160 iterations. Note that in both cases the reconstructed image is the twin image (rotated 108°), which is

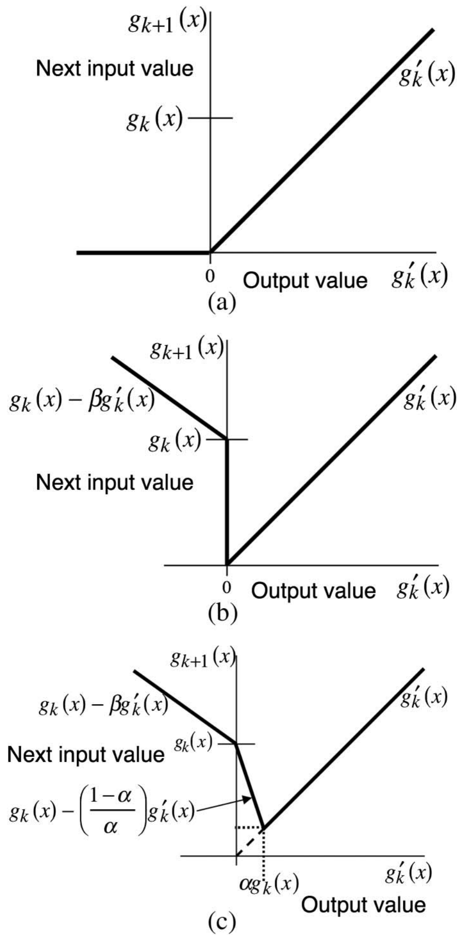


Fig. 1. Next input value as a function of the output value for (a) error-reduction, (b) HIO, and (c) CHIO algorithms.

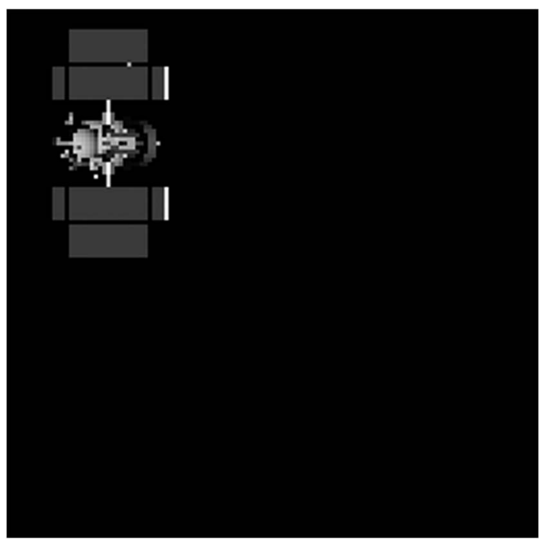


Fig. 2. Object.

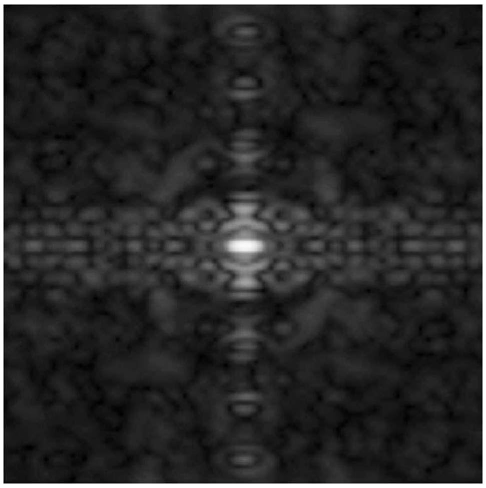
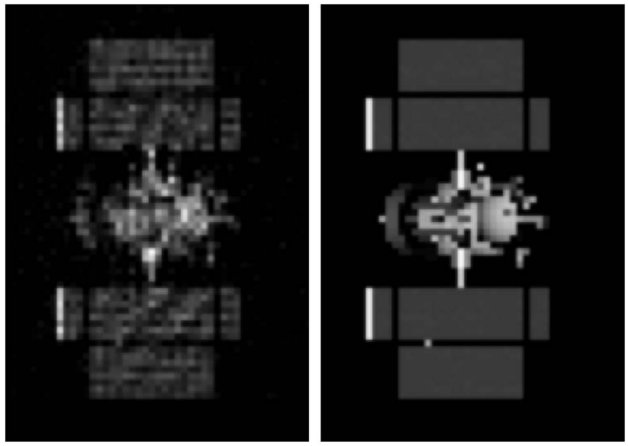


Fig. 3. Magnitude of the Fourier transform of the object.



(a) (b)

Fig. 5. Reconstructed images after 160 iterations by (a) HIO and (b) CHIO.

considered to be an acceptable solution. For this number of iterations the image from HIO is considerably worse than from CHIO, suggesting the superiority of CHIO over HIO, although with more iterations HIO eventually converged as well.

Figure 6 shows the convergence of the two algorithms. The top curves show the object-domain error metric,

$$E_o^2 = \frac{\sum_{(x,y)\in\gamma} |g'_k(x,y)|^2}{\sum\limits_{(x,y)} |g'_k(x,y)|^2}, \tag{3}$$

where γ is the set of points where the object-domain constraints (of support and nonnegativity) are violated; this is how well the reconstructed image satisfies the data and the constraints, and can be computed in real-world scenarios as the algorithm proceeds. The bottom curves show the absolute error [86],

$$E_{abs}^2 = \min_{(x_o,y_o)} \frac{\sum_{(x,y)} |g'_k(x-x_o, y-y_o) - f(x,y)|^2}{\sum\limits_{(x,y)} |f(x,y)|^2}, \tag{4}$$

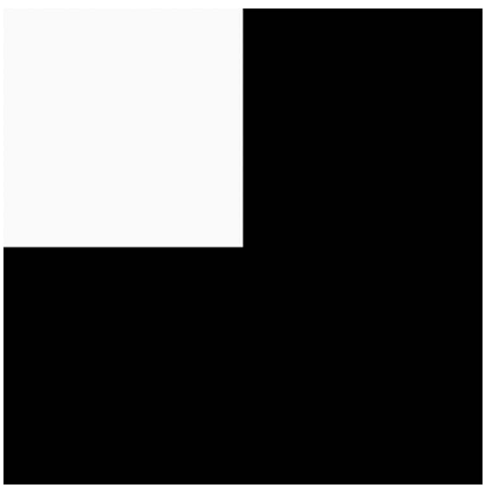


Fig. 4. Support constraint.

which is the normalized mean-squared difference between the true object, f, and the reconstructed image. This metric is computed for both the reconstructed image and its twin, and the smaller of the two numbers is reported. The minimization over translations is done to subpixel accuracy [87]. This metric is known only for computer-simulation experiments for which the true object is known. From these curves we can see that CHIO converged much more rapidly than HIO for this particular case; HIO did eventually converge to the solution after more iterations.

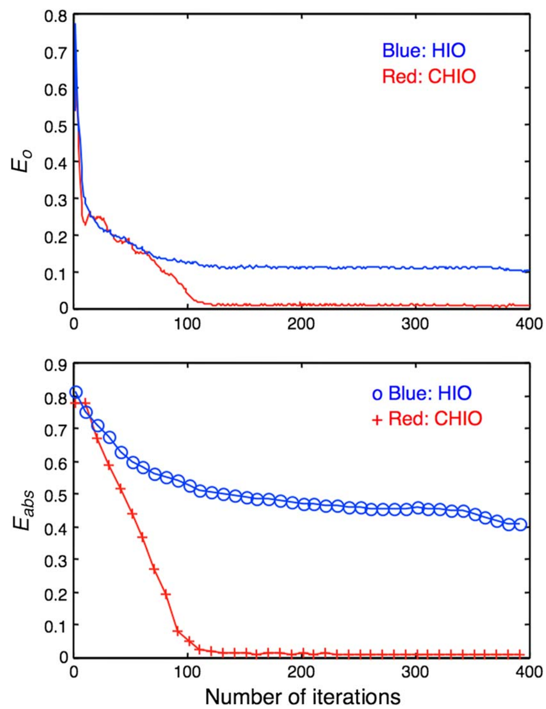


Fig. 6. (Color online) Convergence as a function of iteration number for HIO and CHIO. Top: object-domain error metric, bottom: absolute error (with respect to the true object).

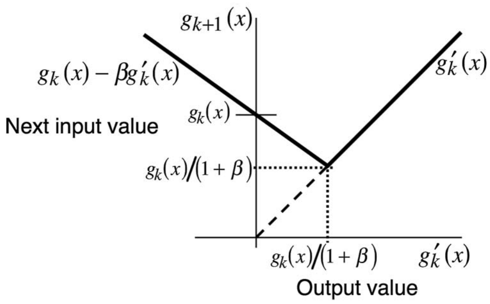


Fig. 7. Next input value as a function of the output value for the HPR algorithm.

The fourth step of the HPR algorithm [83, Eq. (21)] can be written

$$g_{k+1}(x,y) = \begin{cases} g_k'(x,y), \ (x,y) \in S \ \& \ g_k(x,y)/(1+\beta) \leq g_k'(x,y) \\ g_k(x,y) - \beta g_k'(x,y), \ \text{otherwise} \end{cases}, \quad (5)$$

and is illustrated in Fig. 7. From this we see that HPR is a special case of CHIO with $\alpha = 1/(1+\beta)$.

5. Concluding Remarks

The iterative algorithms invented by Gerchberg and Saxton and those working in computer holography were adapted and improved for image reconstruction and wavefront sensing, including the development of the HIO algorithm and gradient-based nonlinear optimization algorithms, and were expanded to numerous different application areas because of the ubiquity of Fourier transforms in physics and engineering and the generality and simplicity of the algorithms, allowing them to effectively solve a multitude of reconstruction and synthesis problems. These areas and algorithms continue to be actively developed, with the application to x-ray coherent diffractive imaging as the current area of most rapid growth. For more detailed reviews of phase retrieval, the reader is referred to [88,89] for astronomical imaging, [90] for electron diffraction, [91] for crystallography, and [92] for coherent diffractive imaging.

I would like to thank my many collaborators at ERIM/Veridian Systems and my students and postdocs at the University of Rochester for making this such a memorable journey.

References

1. J. R. Fienup, "Phase retrieval algorithms: a comparison," Appl. Opt. **21**, 2758–2769 (1982).
2. J. R. Fienup, "Improved synthesis and computational methods for computer-generated holograms," Ph.D. Thesis (Stanford University, 1975) (ProQuest, UMI No. 7525523, or at http://www.optics.rochester.edu/workgroups/fienup/Professor.htm).
3. R. J. Collier, C. B. Burkhardt, and L. H. Lin, *Optical Holography* (Academic, 1971).
4. A. W. Lohmann and D. P. Paris, "Binary Fraunhofer holograms, generated by computer," Appl. Opt. **6**, 1739–1748 (1967).
5. L. B. Lesem, P. M. Hirsch, and J. A. Jordan, Jr., "The kinoform: a new wavefront reconstruction device," IBM J. Res. Devel. **13**, 150–155 (1969).
6. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, 1968).
7. J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts, 2005).
8. Paul Simon, "Kodachrome".
9. D. Kermisch, "Image reconstruction from phase information only," J. Opt. Soc. Am. **60**, 15–16 (1970).
10. N. C. Gallagher and B. Liu, "Method for computing kinoforms that reduces image reconstruction error," Appl. Opt. **l2**, 2328–2335 (1973).
11. B. Liu and N. C. Gallagher, "Convergence of a spectrum shaping algorithm," Appl. Opt. **l3**, 2470–2471 (1974).
12. P. M. Hirsch, J. A. Jordan, Jr., and L. B. Lesem, "Method of making an object-dependent diffuser," U.S. patent 3,619,022 (9 Nov. 1971).
13. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," Optik **35**, 237–246 (1972).
14. J. R. Fienup, "Iterative method applied to image reconstruction and to computer-generated holograms," Opt. Eng. **l9**, 297–305 (1980).
15. R. Hanbury Brown and R. Q. Twiss, "A test of a new type of stellar interferometer on Sirius," Nature **178**, 1046–1048 (1956).
16. A. Labeyrie, "Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images," Astron. Astrophys. **6**, 85–87 (1970).
17. E. Wolf, "Is a complete determination of the energy spectrum of light possible from measurements of the degree of coherence," Proc. Phys. Soc. London **80**, 1269–1272 (1962).
18. A. Walther, "The question of phase retrieval in optics," Opt. Acta **l0**, 41–49 (1963).
19. J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," Opt. Lett. **3**, 27–29 (1978).
20. D. C. Chu, J. R. Fienup, and J. W. Goodman, "Multiemulsion on-axis computer generated hologram," Appl. Opt. **12**, 1386–1388 (1973).
21. B. R. Frieden and D. G. Currie, "On unfolding the autocorrelation function," J. Opt. Soc. Am. **66**, 1111 (1976) (Abstract).
22. R. H. T. Bates, "Contributions to the theory of intensity interferometry," Mon. Not. R. Astron. Soc. **142**, 413–428 (1969).
23. P. J. Napier and R. H. T. Bates, "Inferring phase information from modulus information in two-dimensional aperture synthesis," Astron. Astrophys. Suppl. **l5**, 427–430 (1974).
24. J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," presented at the Annual Meeting of the OSA, Toronto, Ontario, October 1977, Abstract: J. Opt. Soc. Am. **67**, 1389 (1977).
25. J. R. Fienup, "Space object imaging through the turbulent atmosphere," Proc. SPIE **149**, 71–81 (1978).
26. J. R. Fienup, "Space object imaging through the turbulent atmosphere," Opt. Eng. **18**, 529–534 (1979).
27. G. B. Feldkamp and J. R. Fienup, "Noise properties of images reconstructed from Fourier modulus," Proc. SPIE **231**, 84–93 (1980).
28. J. R. Fienup and G. B. Feldkamp, "Astronomical imaging by processing stellar speckle interferometry data," Proc. SPIE **243**, 95–102 (1980).
29. T. R. Crimmins and J. R. Fienup, "Ambiguity of phase retrieval for functions with disconnected support," J. Opt. Soc. Am. **71**, 1026–1028 (1981).
30. J. R. Fienup, T. R. Crimmins, and W. Holsztynski, "Reconstruction of the support of an object from the support of its autocorrelation," J. Opt. Soc. Am. **72**, 610–624 (1982).
31. J. R. Fienup, "Phase retrieval in astronomy," in *Signal Recovery and Synthesis with Incomplete Information and Partial Constraints*, digest of papers (Optical Society of America, 1983), paper ThA8-1-5.
32. J. R. Fienup, "Image reconstruction for stellar interferometry," in *Current Trends in Optics*, F. T. Arecchi and F. R. Aussenegg, eds. (Taylor and Francis and Halsted, 1981), Invited Papers

from ICO-12 Meeting, Graz, Austria, September 1981, pp. 95–102.
33. J. R. Fienup, "Comparison of phase retrieval algorithms," presented at the Annual Meeting of the O.S.A., Kissimmee, Florida, October 1981, Abstract: J. Opt. Soc. Am. **71**, 1641 (1981).
34. M. T. Manry and J. K. Aggarwal, "The design of multidimensional FIR digital filters by phase correction," IEEE Trans. Circuits Syst. **23**, 185–199 (1976).
35. J. A. Marozas, "Fourier transform-based continuous phase-plate design technique: a high-pass phase-plate design as an application for OMEGA and the National Ignition Facility," J. Opt. Soc. Am. A **24**, 74–83 (2007).
36. S. Buhling and F. Wyrowski, "Solving tolerancing and three-dimensional beam shaping problems by multifunctional wave optical design," Opt. Eng. **40**, 1590–1597 (2001).
37. E. G. Johnson and J. D. Brasher, "Phase encryption of biometrics in diffractive optical elements," Opt. Lett. **21**, 1271–1273 (1996).
38. Y. Ivankovski and D. Mendlovic, "High-rate-long-distance fiber-optic communication based on advanced modulation techniques," Appl. Opt. **38**, 5533–5540 (1999).
39. J. Skaar, "Iterative design of antireflection coatings based on the direct and inverse scattering transform," Opt. Commun. **232**, 45–48 (2004).
40. D. Morris, "Phase retrieval in the radio holography of reflector antennas and radio telescopes," IEEE Trans. Antennas Propag. **33**, 749–755 (1985).
41. J. N. Cederquist, J. R. Fienup, C. C. Wackerman, S. R. Robinson, and D. Kryskowski, "Wave-front phase estimation from Fourier intensity measurements," J. Opt. Soc. Am. A **6**, 1020–1026 (1989).
42. J. R. Fienup, J. C. Marron, T. J. Schulz, and J. H. Seldin, "Hubble Space Telescope characterized by using phase retrieval algorithms," Appl. Opt. **32**, 1747–1767 (1993).
43. B. H. Dean, D. L. Aronstein, J. S. Smith, R. Shiri, and D. S. Acton, "Phase retrieval algorithm for JWST flight and testbed telescope," Proc. SPIE **6265**, 626511 (2006).
44. G. R. Brady and J. R. Fienup, "Nonlinear optimization algorithm for retrieving the full complex pupil function," Opt. Express **14**, 474–486 (2006).
45. M. Guizar-Sicairos and J. R. Fienup, "Measurement of coherent x-ray focused beams by phase retrieval with transverse translation diversity," Opt. Express **17**, 2670–2685 (2009).
46. S. T. Thurman, R. T. DeRosa, and J. R. Fienup, "Amplitude metrics for field retrieval with hard-edge and uniformly illuminated apertures," J. Opt. Soc. Am. A **26**, 700–709 (2009).
47. B. H. Dean and C. W. Bowers, "Diversity selection for phase-diverse phase retrieval," J. Opt. Soc. Am. A **20**, 1490–1504 (2003).
48. H. M. L. Faulkner and J. M. Rodenburg, "Movable aperture lensless transmission microscopy: a novel phase retrieval algorithm," Phys. Rev. Lett. **93**, 023903 (2004).
49. M. Guizar-Sicairos and J. R. Fienup, "Phase retrieval with transverse translation diversity: a nonlinear optimization approach," Opt. Express **16**, 7264–7278 (2008).
50. J. S. Smith, D. L. Aronstein, B. H. Dean, and D. S. Acton, "Phase retrieval on broadband and under-sampled images for the JWST testbed telescope," Proc. SPIE **7436**, 7436D (2009).
51. G. R. Brady and J. R. Fienup, "Effect of broadband illumination on reconstruction error of phase retrieval in optical metrology," Proc. SPIE **6617**, 66170I (2007).
52. M. R. Teague, "Deterministic phase retrieval: a Green's function solution," J. Opt. Soc. Am. **73**, 1434–1441 (1983).
53. D. J. Kane and R. Trebino, "Characterization of arbitrary femtosecond pulses using frequency-resolved optical gating," IEEE J. Quantum Electron. **29**, 571–579 (1993).
54. R. Trebino and D. J. Kane, "Using phase retrieval to measure the intensity and phase of ultrashort pulses: frequency-resolved optical gating," J. Opt. Soc. Am. A **10**, 1101–1111 (1993).
55. R. G. Lane, "Blind deconvolution of speckle images," J. Opt. Soc. Am. A **9**, 1508–1514 (1992).
56. C. M. Vest and I. Prikryl, "Tomography by iterative convolution: empirical study and application to interferometry," Appl. Opt. **23**, 2433–2440 (1984).
57. G. K. Datta and R. M. Vasu, "Non-interferometric methods of phase estimation for application in optical tomography," J. Mod. Opt. **46**, 1377–1388 (1999).
58. M. H. Maleki and A. J. Devaney, "Phase-retrieval and intensity-only reconstruction algorithms for optical diffraction tomography," J. Opt. Soc. Am. A **10**, 1086–1092 (1993).
59. R. P. Yu, D. M. Paganin, and M. J. Morgan, "Inferring generalized time-dependent complex Ginzburg-Landau equations from modulus and gauge-field information," Phys. Rev. B. **77**, 134512 (2008).
60. M. Carmody, E. Landree, L. D. Marks, and K. L. Merkle, "Determination of the current density distribution in Josephson junctions," Phys. C **315**, 145–153 (1999).
61. V. Elser, I. Rankenburg, and P. Thibault, "Searching with iterated maps," Proc. Natl. Acad. Sci. **104**, 418–423 (2007).
62. J. R. Fienup, "Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint," J. Opt. Soc. Am. A **4**, 118–123 (1987).
63. J. R. Fienup, "Lensless coherent imaging by phase retrieval with an illumination pattern constraint," Opt. Express **14**, 498–508 (2006).
64. J. R. Fienup and A. M. Kowalczyk, "Phase retrieval for a complex-valued object by using a low-resolution image," J. Opt. Soc. Am. A **7**, 450–458 (1990).
65. P. S. Idell, J. R. Fienup, and R. S. Goodman, "Image synthesis from nonimaged laser speckle patterns," Opt. Lett. **12**, 858–860 (1987).
66. D. Sayre, "Some implications of a theorem due to Shannon," Acta Cryst. **5**, 843 (1952).
67. K. A. Nugent, D. Paganin, and T. E. Gureyev, "A phase odyssey," Phys. Today **54**, 27–32 (2001).
68. J. N. Cederquist, J. R. Fienup, J. C. Marron, and R. G. Paxman, "Phase retrieval from experimental far-field data," Opt. Lett. **13**, 619–621 (1988).
69. J. Miao, P. Charalambous, J. Kirz, and D. Sayre, "Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," Nature **400**, 342–344 (1999).
70. H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. H. Spence, U. Weierstall, T. Beetz, C. Jacobsen, and D. Shapiro, "High-resolution *ab initio* three-dimensional x-ray diffraction microscopy," J. Opt. Soc. Am. A **23**, 1179–2000 (2006).
71. J. H. Seldin and J. R. Fienup, "Numerical investigation of the uniqueness of phase retrieval," J. Opt. Soc. Am. A **7**, 412–427 (1990).
72. J. R. Fienup and C. C. Wackerman, "Phase retrieval stagnation problems and solutions," J. Opt. Soc. Am. A **3**, 1897–1907 (1986).
73. H. Takajo, T. Takahashi, H. Kawanami, and R. Ueda, "Numerical investigation of the iterative phase-retrieval stagnation problem: territories of convergence objects and holes in their boundaries," J. Opt. Soc. Am. A **14**, 3175–3187 (1997).
74. H. Takajo, T. Takahashi, R. Ueda, and M. Taninaka, "Study on the convergence property of the hybrid input output algorithm used for phase retrieval," J. Opt. Soc. Am. A **15**, 2849–2861 (1998).
75. H. Takajo, T. Takahashi, and T. Shizuma, "Further study on the convergence property of the hybrid input–output algorithm used for phase retrieval," J. Opt. Soc. Am. A **16**, 2163–2168 (1999).
76. R. G. Lane, "Phase retrieval using conjugate gradient minimization," J. Mod. Opt. **38**, 1797–1813 (1991).
77. J. R. Fienup, "Phase-retrieval algorithms for a complicated optical system," Appl. Opt. **32**, 1737–1746 (1993).
78. J. R. Fienup, "Wave front sensing by nonlinear optimization," in *Frontiers in Optics*, OSA Technical Digest (CD) (Optical Society of America, 2006), paper FML2.
79. T. R. Crimmins, J. R. Fienup, and B. J. Thelen, "Improved bounds on object support from autocorrelation support and application to phase retrieval," J. Opt. Soc. Am. A **7**, 3–13 (1990).

80. D. C. Youla, "Generalized image restoration by method of alternating orthogonal projections," IEEE Trans. Circuits Syst. **25**, 694–702 (1978).
81. D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1—theory," IEEE Trans. Med. Imag. **1**, 81–94 (1982).
82. H. H. Bauschke, P. L. Combettes, and D. R. Luke, "Phase retrieval, error reduction algorithms, and Fienup variants: a view from convex optimization," J. Opt. Soc. Am. A **19**, 1334–1345 (2002).
83. H. H. Bauschke, P. L. Combettes, and D. R. Luke, "A hybrid projection reflection method for phase retrieval," J. Opt. Soc. Am. A **20**, 1025–1034 (2003).
84. V. Elser, "Phase retrieval by iterated projections," J. Opt. Soc. Am A **20**, 40–55 (2003).
85. J. R. Fienup, "Phase retrieval with continuous version of hybrid input–output," in *Frontiers in Optics*, OSA Technical Digest (CD) (Optical Society of America, 2003), paper ThI3.
86. J. R. Fienup, "Invariant error metrics for image reconstruction," Appl. Opt. **36**, 8352–8357 (1997).
87. M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," Opt. Lett. **33**, 156–158 (2008).
88. R. H. T. Bates, "Astronomical speckle imaging," Phys. Rep. **90**, 203–297 (1982).
89. J. C. Dainty and J. R. Fienup, "Phase retrieval and image reconstruction for astronomy," *Image Recovery: Theory and Application*, H. Stark, ed. (Academic, 1987), pp. 231–275.
90. W. O. Saxton, *Computer Techniques for Image Processing in Electron Microscopy* (Academic, 1978).
91. R. P. Millane, "Phase retrieval in crystallography and optics," J. Opt. Soc. Am. A **7**, 394–411 (1990).
92. K. A. Nugent, "Coherent methods in the x-ray sciences," Adv. Phys. **59**, 1–99 (2010).