Penalized-Likelihood Image Reconstruction for Digital Holography

Saowapak Sotthivirat

National Electronics and Computer Development Center
Ministry of Science and Technology, Klong Luang, Pathumthani 12120, THAILAND

Jeffrey A. Fessler

Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48109-2122

saowapak.sotthivirat@nectec.or.th, fessler@umich.edu

Conventional numerical reconstruction for digital holography using a filter applied in the spatial frequency domain to extract the primary image may yield suboptimal image quality because of the loss in high-frequency components and interference from other undesirable terms of a hologram. In this paper, we propose a new numerical reconstruction approach using a statistical technique. This approach reconstructs the complex field of the object from the real-valued hologram intensity data. Because holographic image reconstruction is an ill-posed problem, our statistical technique is based on penalized-likelihood estimation. We develop a Poisson statistical model for this problem and derive an optimization transfer algorithm that monotonically decreases the cost function each iteration. Simulation results show that our statistical technique has the potential to improve image quality in digital holography relative to conventional reconstruction techniques. © 2004 Optical Society of America

OCIS codes: 090.0090,090.1760,100.0100,100.2000,100.3010,100.3190

1. Introduction

Holography is a technique for recording and reconstructing both the amplitude and phase of a wave field.[1] Analog conventional holography consists of two optical processes: recording and reconstructing. First, the interference pattern between an object wavefront and a reference beam are recorded using photographic material. The recorded pattern is called a hologram. Then, one reconstructs the object's complex wavefront by illuminating the recording medium with a wave that is similar to the original reference beam. This process generates the zero-order image and the twin images called the primary (virtual) and conjugate (real) images. In in-line holography invented by Gabor,[2] the reconstructed holographic image suffers from an overlap of these three images, degrading resolution and contrast. To separate the primary image from other terms, Leith and Upatnieks[3] invented off-axis holography by introducing the reference beam at an angle with respect to the object beam. With

this technique, the zero-order, primary and conjugate images appear at different locations, so each image can be observed separately.

Because the processes of optical recording on photographic film and optical reconstruction preparation in analog conventional holography are time consuming and lack of real-time imaging ability, digital recording of a hologram on a digital detector (such as a charged-coupled-device (CCD) camera) and a numerical reconstruction of a complex object field on a computer become attractive alternatives and have been very useful in many applications.[4,5,6,7,8] In digital off-axis holography, the most common approach for extracting the primary image in numerical reconstruction is to perform a digital "spatial filter" that selects appropriate spatial frequencies in the Fourier domain of a hologram.[9] Limitations of that approach include a loss of high frequency components and interference from other terms in the hologram, which degrade the reconstructed holographic image quality. Phase-shifting or phase modulation[7,10,11] methods were proposed to suppress the zero-order and conjugate images, but they require at least three holograms to reconstruct one holographic image. The clever approach proposed in Ref. 12 estimates the complex object beam by solving a small system of equations; however, no noise model was considered.

To overcome the drawbacks of existing approaches, in this paper we propose a new numerical holographic reconstruction approach based on a statistical model for the measurements and a physical model of the optical system. Statistical image formation techniques have succeeded in many applications.[13,14,15,16] Statistical image reconstruction for digital holography can be formulated as an inverse problem in which we try to obtain a complex reconstructed holographic image (primary image) from hologram intensity data that are real.

Çetin *et al.*[17] proposed a statistical technique for Fourier holography and other coherent imaging applications. Their method was based on a Gaussian noise model and used a least-squares approach. Considering the digital recording process of a hologram, our statistical model follows a Poisson distribution having the mean associated with a squared magnitude of the interference between the object and reference beams. Due to the ill-posed nature of image reconstruction, our statistical technique uses penalized-likelihood (PL) estimation.[18,19,20] This optimization problem is very challenging because its negative log-likelihood function contains multiple global minima. Therefore, regularization is necessary to improve the problem conditioning and to reduce non-uniqueness. Moreover, we show that using two measured holograms can improve the results when reconstructing a complex holographic image with the same number of pixels as the recording device. (The use of two data sets to help estimate complex quantities has been applied in other optical problems.[21,22])

In PL estimation, the unknown parameter vector, which represents the complex object field, is estimated by minimizing a cost function. Since closed-form solutions are unavailable, we need an iterative algorithm to solve the problem. However, for the Poisson model, conventional gradient-based minimization is difficult. To simplify the optimization problem and to assure monotonic decreases in the cost function at each iteration, our proposed reconstruction approach is based on the use of optimization transfer and convexity techniques by finding a "surrogate function" that lies above the original cost function at each iteration.[14,23] Instead of minimizing the original cost function, we minimize the surrogate function using an iterative algorithm, such as the separable-paraboloidal-surrogate[14,23] or

2

conjugate gradient algorithm.

In this study, we demonstrate our holographic reconstruction method in the specific context of image-plane holography,[24, 25] which is a new imaging technique for forming a 3-D image of a "thick," partially transparent object without dissection. This technique has the same optical sectioning property as in confocal scanning microscopy but requires no xy scanning.[24, 25] The proposed reconstruction method can also be applied to phase retrieval problems, and to Fourier holography when the system matrix represents the Fourier transform, and to Fresnel holography when the system matrix represents the Fresnel transform.

This paper is organized as follows. Section 2 describes the measurement model of a digitally recorded hologram. Section 3 reviews conventional numerical reconstruction using a spatial filtering technique applied in the frequency domain, and introduces iterative reconstruction techniques. Section 4 proposes a statistical model for digital holography and introduces a new statistical holographic reconstruction technique based on PL estimation. Section 5 applies optimization transfer and convexity techniques to derive the surrogate functions and the iterative algorithm for holographic image reconstruction. Section 6 compares different holographic reconstructed images using our statistical approach with the conventional reconstruction approach. Finally, conclusions are given in Section 7.

2. Measurement Model for Digital Holography

For digital holography, a computer performs numerical reconstruction of the object from a hologram that is recorded by a digital detector. Figure 1 illustrates the recording process in digital holography.

We assume a linear relationship between the object beam u_o at the hologram (detector) plane and the object field f, as described by the following superposition integral:

$$u_o(\boldsymbol{r}) = \int h(\boldsymbol{r}; \boldsymbol{r}') f^{\mathrm{true}}(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r}', \tag{1}$$

where $h(\boldsymbol{r}; \boldsymbol{r}')$ denotes the continuous-space point spread function (PSF) of the optical system, $f^{\mathrm{true}}(\boldsymbol{r})$ denotes the true object, and $\boldsymbol{r}$ denotes 2-D spatial coordinates on the recording plane. For a planar object, $\boldsymbol{r}'$ denotes the 2-D coordinates within the object plane. For a 3-D object, $\boldsymbol{r}'$ denotes the 3-D coordinates within the object volume. For 3-D reconstruction in image-plane holography, one must scan the object (or the focal plane) along the optical axis and record a set of holograms. For simplicity, in this paper we focus on the 2-D case. In practice, the recording medium has finite thickness, but we ignore this effect here for simplicity.

For image-plane holography, the PSF h represents the characteristics of the imaging optics. For Fourier holography, h includes a Fourier transform, and for Fresnel holography, h includes a Fresnel transform (an approximation of the diffraction integral).[1] In all cases, the information about the object f of interest is embedded in the object beam u_o.

For analog conventional holography, the interference between the object and reference beams at the recording plane has the following continuous-space intensity:[1]

$$
\begin{aligned}
I(\boldsymbol{r}) &= |u_o(\boldsymbol{r}) + u_{\mathrm{ref}}(\boldsymbol{r})|^2 \\
&= |u_o(\boldsymbol{r})|^2 + |u_{\mathrm{ref}}(\boldsymbol{r})|^2 + u_o(\boldsymbol{r}) u_{\mathrm{ref}}^*(\boldsymbol{r}) + u_o^*(\boldsymbol{r}) u_{\mathrm{ref}}(\boldsymbol{r}),
\end{aligned}
\tag{2}
$$

where u_ref denotes the (known) field of the reference beam and $*$ represents the complex conjugate. The first two terms in Eq. (2) constitute the zero-order image; the third term, which is proportional to u_o, leads to the formation of the primary image; and the fourth term, which is proportional to u_o^*, leads to the formation of the conjugate image.[1] For off-axis holography, the reference beam is oriented at some angle resulting in a known spatial carrier frequency denoted by $\boldsymbol{\alpha}$. An example of such a reference beam is a plane wave that is tilted by an angle $\boldsymbol{\theta} = (\theta_x, \theta_y)$ in the x and y directions with respect to the optical axis, *i.e.*,

$$u_\mathrm{ref}(\boldsymbol{r}) = U_\mathrm{ref} \exp(-\imath 2\pi \boldsymbol{r} \cdot \boldsymbol{\alpha}) \tag{3}$$

and

$$\boldsymbol{\alpha} = \left(\frac{\sin \theta_x}{\lambda}, \frac{\sin \theta_y}{\lambda} \right), \tag{4}$$

where U_ref denotes the amplitude of the reference wave and λ is the wavelength.

Let $\boldsymbol{Y} = (Y_1, \ldots, Y_N)$ denote the (noisy) hologram measurement data recorded by a digital detector, where N denotes the number of measurement elements. We treat the measurement recorded by the ith element of the digital detector as a random variable whose mean is modeled as follows:

$$\mathrm{E}[Y_i] = |u_o(\boldsymbol{r}_i) + u_\mathrm{ref}(\boldsymbol{r}_i)|^2 + b_i, \qquad i = 1, \ldots, N, \tag{5}$$

where b_i denotes offsets due to effects such as dark current, and $\boldsymbol{r}_i$ denotes the center of the ith detector element. For simplicity, we treat the CCD camera response as a Dirac impulse at the center of each element. One could generalize Eq. (5) to include convolution with a CCD point response function. The reconstruction goal is to estimate the object f from the measurements $\boldsymbol{Y}$ using the model (5).

3. Numerical Holographic Reconstruction Methods

This section first reviews conventional numerical reconstruction using a filtering method and then introduces iterative reconstruction techniques.

A. *Conventional Filtering Approach*

The first step in conventional numerical holographic reconstruction is to apply a "filter" in the frequency domain to extract either the primary or conjugate image corresponding to one of the last two terms of Eq. (2). This conventional approach assumes that the reference beam is planar, as in Eq. (3). By substituting Eq. (3) into Eq. (2) and taking the Fourier transform, the spatial-frequency spectrum of the recorded interference pattern is converted into an angular spectrum of diffracted waves:

$$\mathcal{I}(\boldsymbol{\nu}) = \mathcal{I}_o(\boldsymbol{\nu}) + |U_\mathrm{ref}|^2 \delta(\boldsymbol{\nu}) + U_\mathrm{ref} \mathcal{U}_o(\boldsymbol{\nu} - \boldsymbol{\alpha}) + U_\mathrm{ref} \mathcal{U}_o^*(-\boldsymbol{\nu} - \boldsymbol{\alpha}), \tag{6}$$

where $\boldsymbol{\nu}$ denotes 2-D spatial frequencies, $\mathcal{I}_o$ denotes the Fourier transform of the intensity of the object beam, $\mathcal{U}_o$ denotes the Fourier transform of the field of the object beam, and $\boldsymbol{\alpha}$ was defined in Eq. (4). The zero-order spectrum, consisting of the first two terms of Eq. (6), must be eliminated to avoid having a background bias in the reconstructed image. The two first-order spectra, the last two terms of Eq. (6), lead to the primary and conjugate images,

respectively. Figure 2 shows a simulated recorded hologram for a planar reference beam tilted along the x direction, the magnitude of the Fourier transform of the hologram, and the reconstructed image using the conventional technique.

Considering the Fourier transform of the hologram in Fig. 2, one can extract either the primary or conjugate image by using a mask to select only one of the first-order spectra and then taking the inverse Fourier transform. Because Eq. (6) consists of four terms yet only one term is extracted, one must consider the appropriate dimensions of the reconstructed image. The usual choice is to reconstruct a holographic image that is the same size as the CCD array; this requires zero padding in the high spatial frequency regions surrounding the extracted frequency components. This conventional approach yields a reconstructed image with poorer resolution than the intrinsic recorder resolution due to the discarded high frequency components, and also suffers from degradation by interference from the residual frequency components of other undesirable terms.

For image-plane holography, no further processing of the extracted image is required after computing the inverse Fourier transform. However, for digital Fresnel holography, an additional processing step is required in which the (discretized) Fresnel transform of the image is computed. (Similarly for Fourier holography.) In contrast, the statistical approach described next requires no such post processing since the effects of the Fresnel or Fourier transform are incorporated into the PSF h in Eq. (1), and hence into the system matrix $\boldsymbol{A}$ defined in Eq. (9) below.

B. Iterative Reconstruction

For iterative holographic reconstruction, one must parameterize the continuous-space object in Eq. (1). We approximate the true object $f^{\text{true}}(\boldsymbol{r})$ using a linear combination of basis functions as follows:[26]

$$f^{\text{true}}(\boldsymbol{r}) \approx f(\boldsymbol{r}) = \sum_{j=1}^{P} x_j \chi_j(\boldsymbol{r}) \tag{7}$$

where $f(\boldsymbol{r})$ is an approximation of the true object in the continuous space, x_j is the unknown complex coefficient of the jth basis function, P is the total number of parameters (*e.g.*, pixels) to be estimated, and $\chi_j(\boldsymbol{r})$ is a basis function. For example, $\chi_j(\boldsymbol{r})$ is usually chosen to be the indicator function over the jth "pixel" in the object plane, and we adopt this choice in our numerical experiments. However, the formulation is general to any choice of basis functions. Combining Eqs. (1) and (7), we write the object beam $u_o(\boldsymbol{r}_i)$ in the following discrete form:

$$u_o(\boldsymbol{r}_i) = \int h(\boldsymbol{r}_i; \boldsymbol{r}') \sum_{i=1}^{P} x_j \chi_j(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r}' = \sum_{j=1}^{P} a_{ij} x_j = [\boldsymbol{A}\boldsymbol{x}]_i, \tag{8}$$

where a_{ij} is the ijth element of the system matrix $\boldsymbol{A}$, and $\boldsymbol{x} = (x_1, \ldots, x_P)$. Each a_{ij} denotes the contribution of the jth basis function to the the object beam $u_o(\boldsymbol{r}_i)$ at the center of the ith recorder element, and can be expressed in terms of the system PSF and the basis functions as follows:

$$a_{ij} = \int h(\boldsymbol{r}_i; \boldsymbol{r}') \chi_j(\boldsymbol{r}') \, \mathrm{d}\boldsymbol{r}'. \tag{9}$$

Combining these expressions with the definition $u_i \triangleq u_{\mathrm{ref}}(\boldsymbol{r}_i)$ leads to the following model for the measurement means:

$$\mathrm{E}[Y_i] = |[\boldsymbol{A}\boldsymbol{x}]_i + u_i|^2 + b_i, \qquad i = 1, \ldots, N. \tag{10}$$

We use this model for statistical holographic reconstruction. The goal is to estimate $\boldsymbol{x}$ from the measured Y_i's, since $\boldsymbol{x}$ parameterizes the unknown object f of interest. In the special case of space-invariant systems, the system matrix is Toeplitz and $\boldsymbol{A}\boldsymbol{x}$ corresponds to convolution between the image represented by $\boldsymbol{x}$ and the discrete-space system PSF.

Using Eq. (10), iterative techniques can estimate the complex object field from the measurement data (recorded holograms). Unlike the conventional filtering method, iterative techniques can use all of the information in the model (2), rather than discarding all but one of the four terms. Furthermore, iterative methods need not assume that the reference beam is a planar. Therefore, iterative methods could work for both in-line and off-axis holography. Since the recorded hologram is real, whereas the unknown object field is complex, if one attempts to estimate a complex holographic image having the same number of pixels as CCD elements, *i.e.*, $P = N$, then the problem will be under-determined since each x_j in Eq. (7) is complex valued, so consists of two unknown numbers. Regularization may help reduce this problem. Alternatively, one should choose the number of parameters P (and/or the number of recorded hologram samples N) such that $N \geq P$ so that the problem is not intrinsically under-determined.

Digital hologram measurements are degraded by noise, so before describing an iterative algorithm for holographic reconstruction in Section 5, we first formulate a statistical model for the noise.

4. Statistical Model

Statistical techniques for inverse problems require a model for the measurement statistics. In digital holography, the two major noise sources are light quanta statistics characterized by a Poisson distribution, and electronic readout noise characterized by a Gaussian distribution. Since Poisson distributed photon noise is inherent to optical imaging, whereas readout noise depends on detector design, we focus on the a Poisson component.[27] (Readout noise variance can be included in the b_i term below if needed.[27]) In particular, we model the noisy hologram measurements as having independent Poisson distributions with means described in Eq. (10):

$$Y_i \sim \mathrm{Poisson}\{|[\boldsymbol{A}\boldsymbol{x}]_i + u_i|^2 + b_i\}, \qquad i = 1, \ldots, N. \tag{11}$$

All terms in this expression are assumed known except for the unknown image vector $\boldsymbol{x}$. Because $\boldsymbol{x}$ is complex, the problem will be under-determined if the size of the data vector $\boldsymbol{Y}$ is less than twice the number of elements of $\boldsymbol{x}$, *i.e.*, if $N < 2P$.

Since the system matrix $\boldsymbol{A}$ is usually ill-conditioned in imaging problems, we use penalized-likelihood estimation to reconstruct the holographic image by finding the minimizer $\hat{\boldsymbol{x}}$ of a cost function of the following form:

$$\Phi(\boldsymbol{x}) = L(\boldsymbol{x}) + V(\boldsymbol{x}), \tag{12}$$

where L denotes the negative log-likelihood function corresponding to the statistical model, and V denotes a roughness penalty function. Our focus is image-plane holography, which

uses incoherent illumination. For imaging with highly coherent illumination, one may need to consider the effects of speckle when designing the roughness penalty.[28, 29]

Ignoring irrelevant constants independent of $\boldsymbol{x}$, the negative log-likelihood function corresponding to the statistical model (11) is given by:

$$L(\boldsymbol{x}) = \sum_{i=1}^{N} h_i([\boldsymbol{Ax}]_i), \tag{13}$$

where h_i is a real function of a complex scalar argument defined as follows:

$$h_i(l) \triangleq -y_i \log(|l + u_i|^2 + b_i) + (|l + u_i|^2 + b_i), \tag{14}$$

where l denotes a general complex argument. It will be convenient to also write h_i as an explicit function of the real and imaginary components of its argument, $i.e.$, $h_i(l) = h_i(l^R, l^I)$, where $l = l^R + \imath l^I$, and where the superscripts R and I denote the real and imaginary parts, respectively. Each y_i denotes a (real-valued) realization of the Poisson random variable Y_i.

We consider penalty functions that discourage differences between neighboring object pixels having the following form:[23]

$$V(\boldsymbol{x}) = \beta \sum_{i=1}^{r} \psi([\boldsymbol{Cx}]_i), \tag{15}$$

where ψ is a potential function that determines the behavior of the penalty function, $\boldsymbol{C}$ is a penalty matrix that defines adjacent neighboring pairs of pixels, β is a regularization parameter that controls the degree of smoothness in the reconstructed image, and r is the number of pairs of neighboring object pixels. Examples of potential functions include the quadratic potential function $\psi(t) = t^2/2$, and nonquadratic potential functions, such as:[30]

$$\psi(t) = \delta^2 \left[\left|\frac{t}{\delta}\right| - \log\left(1 + \left|\frac{t}{\delta}\right|\right) \right], \tag{16}$$

where δ is a parameter that controls the degree of edge preservation.[30, 31, 32] The smaller δ, the stronger the degree of edge preservation. For simplicity, the penalty matrix we used in this paper consists of horizontal and vertical adjacent neighbors, which is called a first-order neighborhood. An example of the matrix $\boldsymbol{C}$ for a 2×2 image is:

$$\boldsymbol{Cx} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ x_4 - x_3 \\ x_3 - x_1 \\ x_4 - x_2 \end{bmatrix}.$$

Our goal is to estimate $\boldsymbol{x}$ by finding the minimizer of the cost function:

$$\hat{\boldsymbol{x}} \triangleq \arg\min_{\boldsymbol{x}} \Phi(\boldsymbol{x}).$$

Since closed-form solutions for the minimizer are unavailable, iterative algorithms are needed.

5. The Algorithm

In this section, we approach the minimization problem by using optimization transfer and convexity techniques. These lead to an iterative algorithm that monotonically decreases the cost function.

A. Optimization Transfer

Directly minimizing the cost function in Eq. (12) is difficult for nonquadratic h_i's. To simplify the optimization problem and to assure monotonic decreases in the cost function at each iteration, one can apply an optimization transfer approach by finding a "surrogate" function ϕ that lies above the cost function.[14,23,31,33] Each iteration, we obtain the next estimate by finding the minimizer of the surrogate function:

$$\boldsymbol{x}^{n+1} \triangleq \arg\min_{\boldsymbol{x}} \phi(\boldsymbol{x}; \boldsymbol{x}^n) \tag{17}$$

where $\boldsymbol{x}^n$ denotes the estimate at the nth iteration.

Choosing a surrogate function ϕ that satisfies the following monotonicity condition ensures that the iterates $\boldsymbol{x}^n$ will monotonically decrease the cost function Φ:[14,31,33]

$$\Phi(\boldsymbol{x}^n) - \Phi(\boldsymbol{x}) \geq \phi(\boldsymbol{x}^n; \boldsymbol{x}^n) - \phi(\boldsymbol{x}; \boldsymbol{x}^n), \qquad \forall \boldsymbol{x} \geq \boldsymbol{0}. \tag{18}$$

Instead of using the condition above, we choose a surrogate function $\phi(\boldsymbol{x}; \boldsymbol{x}^n)$ that satisfies the following sufficient conditions:

1. $\phi(\boldsymbol{x}^n; \boldsymbol{x}^n) = \Phi(\boldsymbol{x}^n)$
2. $\phi(\boldsymbol{x}; \boldsymbol{x}^n) \geq \Phi(\boldsymbol{x}), \qquad \forall \boldsymbol{x} \in \mathbb{C}^P$
3. $\left. \dfrac{\partial}{\partial x_j} \phi(\boldsymbol{x}; \boldsymbol{x}^n) \right|_{\boldsymbol{x}=\boldsymbol{x}^n} = \left. \dfrac{\partial}{\partial x_j} \Phi(\boldsymbol{x}) \right|_{\boldsymbol{x}=\boldsymbol{x}^n}, \qquad \forall j,$ $\qquad(19)$

where $\mathbb{C}^P$ defines a P-dimensional complex space.

The next section presents the surrogate functions for the cost function given in Eq. (12).

B. Paraboloidal-Surrogate Functions

We first focus on the likelihood part. Since quadratic choices for the surrogate ϕ are particularly easy to minimize, our goal now is to find a parabola that lies above the negative log-likelihood function. Figure 3 illustrates the one-dimensional plot of the marginal cost function $h_i(l^R, 0)$. In this plot the marginal cost function has two optimal minima. However, a 2-D plot of $h_i(l^R, l^I)$ consists of multiple solutions that lie on a circle in the complex plane. Therefore, it is a challenging problem to find the correct optimal solution.

Since l and u_i are complex, we adapt De Pierro's multiplicative trick[34] to rewrite $h_i(l)$ in Eq. (14) as follows:

$$
\begin{aligned}
h_i(l^R, l^I) &= -y_i \log\left(\alpha_i^{R,n} \left[\frac{(l^R + u_i^R)^2 + b_i/2}{\alpha_i^{R,n}} \right] + \alpha_i^{I,n} \left[\frac{(l^I + u_i^I)^2 + b_i/2}{\alpha_i^{I,n}} \right] \right) \\
&\quad + [(l^R + u_i^R)^2 + b_i/2] + [(l^I + u_i^I)^2 + b_i/2], \tag{20}
\end{aligned}
$$

8

where we define

$$\alpha_i^{R,n} = \frac{(l^{R,n} + u_i^R)^2 + b_i/2}{k_i^n}$$

$$\alpha_i^{I,n} = \frac{(l^{I,n} + u_i^I)^2 + b_i/2}{k_i^n}$$

$$k_i^n = |l_i^n + u_i|^2 + b_i$$

$$l_i^n = [\boldsymbol{Ax}^n]_i.$$

Combining the fact that $\alpha_i^{R,n} + \alpha_i^{I,n} = 1$ with the convexity of the negative log function in Eq. (20) leads to our first surrogate function in which we separate the real and imaginary parts:

$$h_i(l) = h_i(l^R, l^I) \leq h_i^{R,n}(l^R) + h_i^{I,n}(l^I), \tag{21}$$

where

$$h_i^{R,n}(l^R) \triangleq -y_i \alpha_i^{R,n} \log\left(\frac{(l^R + u_i^R)^2 + b_i/2}{\alpha_i^{R,n}}\right) + (l^R + u_i^R)^2 + b_i/2 \tag{22}$$

$$h_i^{I,n}(l^I) \triangleq -y_i \alpha_i^{I,n} \log\left(\frac{(l^I + u_i^I)^2 + b_i/2}{\alpha_i^{I,n}}\right) + (l^I + u_i^I)^2 + b_i/2. \tag{23}$$

In 1-D, the surrogates $h_i^{R,n}$ and $h_i^{I,n}$ each have two minima (see Fig. 3) and are symmetric about the line $l = -u_i$. To facilitate the minimization in Eq. (17), we want to find parabolic surrogates that lie above these curves. A parabolic surrogate function for the real part has the following form:

$$q_i^{R,n}(l^R) = h_i^{R,n}(l_i^{R,n}) + \dot{h}_i^{R,n}(l_i^{R,n})(l^R - l_i^{R,n}) + \frac{1}{2}c_i^{R,n}(l^R - l_i^{R,n})^2 \tag{24}$$

where $\dot{h}_i^{R,n}$ is the first derivative of $h_i^{R,n}$ and $c_i^{R,n}$ is the curvature of the parabola $q_i^{R,n}$. This parabola has the same value as $h_i^{R,n}$ at the current estimate $l^R = l_i^{R,n}$, and the same first derivatives at that point. A parabolic surrogate function for the imaginary part is denoted by $q_i^{I,n}(l^I)$ and has an analogous form.

According to the sufficient conditions in Eq. (19), we must choose the parabola $q_i^{R,n}$ to satisfy the following conditions:

1. $q_i^{R,n}(l_i^{R,n}) = h_i^{R,n}(l_i^{R,n})$

2. $q_i^{R,n}(l^R) \geq h_i^{R,n}(l^R n), \quad \forall l^R$

3. $\dot{q}_i^{R,n}(l_i^{R,n}) = \dot{h}_i^{R,n}(l_i^{R,n}).$

Similar conditions apply to $q_i^{I,n}$. The first and third conditions are satisfied by the construction of $q_i^{R,n}$ and $q_i^{I,n}$. The only remaining problem is to find curvatures $c_i^{R,n}$ and $c_i^{I,n}$ that satisfy the second conditions. For the fastest convergence rate,[14] ideally we would choose the *smallest* curvature for which the second condition is satisfied. However, a closed-form

solution for this optimal choice has proven elusive. Instead, we have chosen the curvatures using the following general expression:

$$c_i^{R,n} = \max_{l^R \in \mathbb{R}} \frac{\dot{h}_i^{R,n}(l^R) - \dot{h}_i^{R,n}(l_i^{R,n})}{l^R - l_i^{R,n}}, \tag{25}$$

and likewise for $c_i^{I,n}$. Although this curvature is not optimal, Appendix A shows that this choice leads to a parabolic surrogate that is guaranteed to lie above the cost function. For the specific model in Eq. (11), the curvatures in Eq. (25) have the following closed-form solution:

$$c_i^{R,n} = \frac{2y_i[b_i^2 + 2b_i(l_i^{R,n} + u_i^R)^2]^{1/2}(l_i^{R,n} + u_i^R)^2}{k_i^n \left[b_i^2 + b_i\left(2(l_i^{R,n} + u_i^R)^2 + [b_i^2 + 2b_i(l_i^{R,n} + u_i^R)^2]^{1/2}\right)\right]} + 2. \tag{26}$$

Appendix B derives this expression. Computation time per iteration could be reduced by using precomputed curvatures.[35]

The following inequalities summarize the construction of the surrogate functions:

$$
\begin{aligned}
L(\boldsymbol{x}) &= \sum_{i=1}^{N} h_i([\boldsymbol{Ax}]_i) \\
&\leq \sum_{i=1}^{N} h_i^{R,n}([\boldsymbol{Ax}]_i^R) + h_i^{I,n}([\boldsymbol{Ax}]_i^I) \\
&\leq \sum_{i=1}^{N} q_i^{R,n}([\boldsymbol{Ax}]_i^R) + q_i^{I,n}([\boldsymbol{Ax}]_i^I) \\
&\triangleq Q(\boldsymbol{x}; \boldsymbol{x}^n).
\end{aligned} \tag{27}
$$

Since the likelihood surrogate function Q is quadratic, many algorithms could find its minimizer. Minimizing Q is simpler than minimizing the original cost function Φ. In this paper, we apply the separable-paraboloidal-surrogate (SPS) algorithm[14,23] for this problem. The conjugate gradient (CG) method could also be applied easily because nonnegativity constraints are not required.

C. The Separable-Paraboloidal-Surrogate Algorithm

In this section, we derive the SPS algorithm for Eq. (27). To apply the SPS approach, we separate pixels using the additive convexity technique developed by De Pierro,[34] permitting simultaneous updates. Using the convexity of $q_i^{R,n}$, we have:

$$
\begin{aligned}
q_i^{R,n}([\boldsymbol{Ax}]_i^R) &= q_i^{R,n}\left(\sum_{j=1}^{P} p_{ij}\left[\frac{[a_{ij}(x_j - x_j^n)]^R}{p_{ij}} + l_i^{R,n}\right]\right) \\
&\leq \sum_{j=1}^{P} p_{ij} q_i^{R,n}\left(\frac{[a_{ij}(x_j - x_j^n)]^R}{p_{ij}} + l_i^{R,n}\right),
\end{aligned} \tag{28}
$$

where we choose p_{ij}'s that satisfy $p_{ij} \geq 0$ and $\sum_{j=1}^{P} p_{ij} = 1$. As in previous work,[23] we use $p_{ij} = \frac{|a_{ij}|}{\sum_{j'=1}^{P} |a_{ij'}|}$ in our simulation study. An analogous inequality applies to $q_i^{I,n}$. Combining the real and imaginary components leads to the following surrogate function:

$$Q(\boldsymbol{x}; \boldsymbol{x}^n) \leq \sum_{j=1}^{P} Q_j^n(x_j) \tag{29}$$

$$Q_j^n(x_j) \triangleq Q_j^{R,n}(x_j) + Q_j^{I,n}(x_j)$$

$$Q_j^{R,n}(x_j) \triangleq \sum_{i=1}^{N} p_{ij} q_i^{R,n} \left(\frac{[a_{ij}(x_j - x_j^n)]^R}{p_{ij}} + l_i^{R,n} \right), \tag{30}$$

and we define $Q_j^{I,n}$ similarly.

To obtain the update at each iteration, we minimize $Q_j^n(x_j)$. When no penalty is used, we obtain the maximum-likelihood estimate as is derived in the next section.

D. Maximum-Likelihood Estimation

Since $Q_j^n(x_j) = Q_j^n(x_j^R + \imath x_j^I)$ is a quadratic function of two real variables, x_j^R and x_j^I, we minimize Q_j^n using one step of Newton's method, which involves a 2×2 matrix-vector multiplication for each pixel as follows:

$$x_j^{n+1} \triangleq \arg\min_{x_j} Q_j^n(x_j)$$

$$= x_j^n - \boldsymbol{H}_j^{-1} \nabla Q_j^n(x_j^n), \qquad j = 1, \ldots, P, \tag{31}$$

where the gradient of Q_j^n is defined by

$$\nabla Q_j^n(x_j^n) \triangleq \left[\begin{array}{c} \frac{\partial}{\partial x_j^R} Q_j^n(x_j) \\ \frac{\partial}{\partial x_j^I} Q_j^n(x_j) \end{array} \right] \Bigg|_{x_j = x_j^n}$$

$$= \left[\begin{array}{c} \sum_{i=1}^{N} a_{ij}^R \dot{h}_i^{R,n}(l_i^{R,n}) + a_{ij}^I \dot{h}_i^{I,n}(l_i^{I,n}) \\ \sum_{i=1}^{N} -a_{ij}^I \dot{h}_i^{R,n}(l_i^{R,n}) + a_{ij}^R \dot{h}_i^{I,n}(l_i^{I,n}) \end{array} \right]$$

$$= \left[\begin{array}{c} \frac{\partial}{\partial x_j^R} L(\boldsymbol{x}) \\ \frac{\partial}{\partial x_j^I} L(\boldsymbol{x}) \end{array} \right] \Bigg|_{\boldsymbol{x} = \boldsymbol{x}^n} \triangleq \left[\begin{array}{c} \dot{L}_j^{R,n} \\ \dot{L}_j^{I,n} \end{array} \right], \tag{32}$$

and the jth 2×2 Hessian matrix, $\boldsymbol{H}_j$, is

$$\boldsymbol{H}_j \triangleq \left[\begin{array}{cc} d_j^{RR} & d_j^{RI} \\ d_j^{IR} & d_j^{II} \end{array} \right] = \left[\begin{array}{cc} \frac{\partial^2}{\partial (x_j^R)^2} Q_j^n(x_j) & \frac{\partial^2}{\partial x_j^R \partial x_j^I} Q_j^n(x_j) \\ \frac{\partial^2}{\partial x_j^I \partial x_j^R} Q_j^n(x_j) & \frac{\partial^2}{\partial (x_j^I)^2} Q_j^n(x_j) \end{array} \right] \Bigg|_{x_j = x_j^n}$$

$$= \left[\begin{array}{cc} \sum_{i=1}^{N} \frac{1}{p_{ij}} [(a_{ij}^R)^2 c_i^{R,n} + (a_{ij}^I)^2 c_i^{I,n}] & \sum_{i=1}^{N} \frac{a_{ij}^R a_{ij}^I}{p_{ij}} (-c_i^{R,n} + c_i^{I,n}) \\ \sum_{i=1}^{N} \frac{a_{ij}^R a_{ij}^I}{p_{ij}} (-c_i^{R,n} + c_i^{I,n}) & \sum_{i=1}^{N} \frac{1}{p_{ij}} [(a_{ij}^I)^2 c_i^{R,n} + (a_{ij}^R)^2 c_i^{I,n}] \end{array} \right] . \tag{33}$$

11

After matrix multiplication, the explicit form for the SPS algorithm becomes simply:

$$
\begin{bmatrix} x_j^{R,n+1} \\ x_j^{I,n+1} \end{bmatrix} = \begin{bmatrix} x_j^{R,n} - \frac{1}{\det \boldsymbol{H}_j}\left(d_j^{II}\dot{L}_j^{R,n} - d_j^{RI}\dot{L}_j^{I,n}\right) \\ x_j^{I,n} - \frac{1}{\det \boldsymbol{H}_j}\left(-d_j^{RI}\dot{L}_j^{R,n} + d_j^{RR}\dot{L}_j^{I,n}\right) \end{bmatrix} \tag{34}
$$

where the determinant of the Hessian matrix, $\boldsymbol{H}_j$, is

$$
\det \boldsymbol{H}_j = d_j^{RR}d_j^{II} - (d_j^{RI})^2. \tag{35}
$$

The surrogate functions derived in this section do not include the penalty function. Without regularization, a noisy image might be obtained after several iterations. Therefore, in the next section we derive the surrogate function for the penalty term in PL estimation. The derivation extends our previous work[14,23] to the case of complex images.

E. Penalty Surrogate Function and Penalized-Likelihood Estimation

Lacking any prior information that would relate the real and imaginary parts of the unknown image $\boldsymbol{x}$, we employ separate penalty functions for the two parts. Since we separately penalize the real and imaginary parts, using different regularization parameters for the real and imaginary parts provides more flexibility than having only one regularization parameter for both. Thus, we use a penalty function of the following form:

$$
V(\boldsymbol{x}) = \beta^R \sum_{i=1}^{r} \psi([\boldsymbol{C}^R \boldsymbol{x}^R]_i) + \beta^I \sum_{i=1}^{r} \psi([\boldsymbol{C}^I \boldsymbol{x}^I]_i), \tag{36}
$$

where $\boldsymbol{C}^R$ and $\boldsymbol{C}^I$ are penalty matrices for the real and imaginary parts of the estimates, and β^R and β^I are the corresponding regularization parameters. To preserve edges, we used the nonquadratic potential function ψ in Eq. (16) in our simulations.

Similar to the nonquadratic likelihood function, we derive the following surrogate functions:

$$
V(\boldsymbol{x}) \le V'(\boldsymbol{x}; \boldsymbol{x}^n) \le \sum_{j=1}^{P} S_j^n(x_j), \tag{37}
$$

where V' is called the paraboloidal-surrogate function for the penalty function and S_j^n is called the separable-paraboloidal surrogate function for the penalty function. The first inequality is derived by forming a parabola that lies above the original penalty function and the second inequality is derived using the convexity of ψ. If a quadratic potential function is used instead, then the first parabola step is unnecessary. The paraboloidal-surrogate function $V'(\boldsymbol{x}; \boldsymbol{x}^n)$ has the following form:

$$
V'(\boldsymbol{x}; \boldsymbol{x}^n) = \beta^R \sum_{i=1}^{r} \varphi([\boldsymbol{C}^R \boldsymbol{x}^R]_i; [\boldsymbol{C}^R \boldsymbol{x}^{R,n}]_i) + \beta^I \sum_{i=1}^{r} \varphi([\boldsymbol{C}^I \boldsymbol{x}^I]_i; [\boldsymbol{C}^I \boldsymbol{x}^{I,n}]_i), \tag{38}
$$

where the function $\varphi(t; s)$ is a parabola in t for fixed s, where t and s denote real scalar arguments, and φ is defined by

$$
\varphi(t; s) = \psi(s) + \dot{\psi}(s)(t - s) + \frac{1}{2}\omega(s)(t - s)^2, \tag{39}
$$

12

and the curvature ω of the parabola surrogate[32] is given by

$$\omega(s) = \frac{\dot{\psi}(s)}{s}.$$

Since the paraboloidal-surrogate function V' is convex, we apply the additive convexity technique developed by De Pierro[34] to obtain the separable-paraboloidal-surrogate function that lies above V' as follows:

$$
\begin{aligned}
S_j^{R,n}(x_j^R) & \triangleq \beta^R \sum_{i=1}^{r} \gamma_{ij}^R \varphi\left(\frac{c_{ij}^R (x_j - x_j^n)^R}{\gamma_{ij}^R} + [\boldsymbol{C}^R \boldsymbol{x}^{R,n}]_i; [\boldsymbol{C}^R \boldsymbol{x}^{R,n}]_i \right) \\
S_j^{I,n}(x_j^R) & \triangleq \beta^I \sum_{i=1}^{r} \gamma_{ij}^I \varphi\left(\frac{c_{ij}^I (x_j - x_j^n)^I}{\gamma_{ij}^I} + [\boldsymbol{C}^I \boldsymbol{x}^{I,n}]_i; [\boldsymbol{C}^I \boldsymbol{x}^{I,n}]_i \right) \\
S_j^{n}(x_j) & = S_j^{R,n}(x_j^R) + S_j^{I,n}(x_j^I),
\end{aligned}
\tag{40}
$$

and we define $\gamma_{ij}^o = \frac{|c_{ij}^o|}{\sum_{j'=1}^{P} |c_{ij'}^o|}$, where o represents R or I. From Eq. (39), the first derivative of $\varphi(t; s)$ evaluated at $t = s$ is $\dot{\psi}(s)$, thus the gradient of S_j^n is

$$
\begin{aligned}
\nabla S_j^n(x_j^n) & = \left. \begin{bmatrix} \frac{\partial}{\partial x_j^R} S_j^{R,n}(x_j^R) \\ \frac{\partial}{\partial x_j^I} S_j^{I,n}(x_j^I) \end{bmatrix} \right|_{x_j = x_j^n} = \begin{bmatrix} \beta^R \sum_{i=1}^{r} c_{ij}^R \dot{\psi}([\boldsymbol{C}^R \boldsymbol{x}^{R,n}]_i) \\ \beta^I \sum_{i=1}^{r} c_{ij}^I \dot{\psi}([\boldsymbol{C}^I \boldsymbol{x}^{I,n}]_i) \end{bmatrix} \\
& = \left. \begin{bmatrix} \frac{\partial}{\partial x_j^R} V(\boldsymbol{x}) \\ \frac{\partial}{\partial x_j^I} V(\boldsymbol{x}) \end{bmatrix} \right|_{x = \boldsymbol{x}^n} \triangleq \begin{bmatrix} \dot{V}_j^{R,n} \\ \dot{V}_j^{I,n} \end{bmatrix}.
\end{aligned}
\tag{41}
$$

Because there is no coupling between x_j^R and x_j^I in the penalty or its surrogate functions, the Hessian matrix for S_j^n is diagonal:

$$
\nabla^2 S_j^n(x_j^n) = \begin{bmatrix} p_j^{R,n} & 0 \\ 0 & p_j^{I,n} \end{bmatrix},
\tag{42}
$$

where

$$
p_j^{o,n} = \left. \frac{\partial^2 S_j^{o,n}}{\partial (x_j^o)^2} \right|_{x_j = x_j^n} = \beta^o \sum_{i=1}^{r} \frac{(c_{ij}^o)^2}{\gamma_{ij}^o} \omega([\boldsymbol{C}^o \boldsymbol{x}^{o,n}]_i).
\tag{43}
$$

To obtain the update of the SPS algorithm in PL estimation, we combine the surrogates for the likelihood and penalty. Thus, the overall surrogate function to be minimized becomes

$$
\phi_j^n(x_j) = Q_j^n(x_j) + S_j^n(x_j).
\tag{44}
$$

So the update x_j is obtained by:

$$
x_j^{n+1} = \arg \min_{x_j} \phi_j^n(x_j).
\tag{45}
$$

Similar to Eq. (34), the update of the SPS algorithm with regularization becomes

$$
\begin{bmatrix} x_j^{R,n+1} \\ x_j^{I,n+1} \end{bmatrix} = \begin{bmatrix} x_j^{R,n} \\ x_j^{I,n} \end{bmatrix} - \frac{1}{\det \tilde{\boldsymbol{H}}_j} \begin{bmatrix} (d_j^{II} + p_j^{R,n})(\dot{L}_j^{R,n} + \dot{V}_j^{R,n}) - d_j^{RI}(\dot{L}_j^{I,n} + \dot{V}_j^{I,n}) \\ -d_j^{RI}(\dot{L}_j^{R,n} + \dot{V}_j^{R,n}) + (d_j^{RR} + p_j^{R,n})(\dot{L}_j^{I,n} + \dot{V}_j^{I,n}) \end{bmatrix},
$$
(46)

where the new Hessian matrix $\tilde{\boldsymbol{H}}_j$ is

$$
\tilde{\boldsymbol{H}}_j = \boldsymbol{H}_j + \begin{bmatrix} p_j^{R,n} & 0 \\ 0 & p_j^{I,n} \end{bmatrix},
$$
(47)

and its determinant is

$$
\det \tilde{\boldsymbol{H}}_j = (d_j^{RR} + p_j^{R,n})(d_j^{II} + p_j^{I,n}) - (d_j^{RI})^2.
$$
(48)

The SPS algorithm outline for holographic image reconstruction is shown below. For simplicity, we supress the "n" superscript.

$\hat{\boldsymbol{x}}$ = initial image
$a_i = \sum_{j=1}^{P} |a_{ij}|, \qquad i = 1, \ldots, N$
for $n = 1, \ldots,$ Niters
$\quad \hat{l} = \sum_{j=1}^{P} a_{ij} \hat{x}_j, \qquad i = 1, \ldots, N$
$\quad k_i^n = |\hat{l} + u_i|^2 + b_i, \qquad i = 1, \ldots, N$
$\quad h_i = \frac{-2y_i(\hat{l}+u_i)}{k_i^n} + 2(\hat{l} + u_i), \qquad i = 1, \ldots, N$
$\quad$ Compute c_i^R and c_i^I using Eq. (26).
$\quad$ for $j = 1, \ldots, P$
$\qquad \dot{L}_j = \sum_{i=1}^{N} a_{ij}^* h_i$
$\qquad \dot{L}_j^R = \text{Re}\{\dot{L}_j\}, \; \dot{L}_j^I = \text{Im}\{\dot{L}_j\}$
$\qquad d_j^{RR} = \sum_{i=1}^{N} \frac{a_i}{|a_{ij}|}[(a_{ij}^R)^2 c_i^R + (a_{ij}^I)^2 c_i^I]$
$\qquad d_j^{II} = \sum_{i=1}^{N} \frac{a_i}{|a_{ij}|}[(a_{ij}^I)^2 c_i^R + (a_{ij}^R)^2 c_i^I]$
$\qquad d_j^{RI} = d_j^{IR} = \sum_{i=1}^{N} \frac{a_i a_{ij}^R a_{ij}^I}{|a_{ij}|}(-c_i^R + c_i^I)$
$\qquad \dot{V}_j^R = \beta^R \sum_{i=1}^{r} c_{ij}^R \dot{\psi}([\boldsymbol{C}^R \hat{\boldsymbol{x}}^R]_i)$
$\qquad \dot{V}_j^I = \beta^I \sum_{i=1}^{r} c_{ij}^I \dot{\psi}([\boldsymbol{C}^I \hat{\boldsymbol{x}}^I]_i)$
$\qquad p_j^R = \beta^R \sum_{i=1}^{r} \frac{(c_{ij}^R)^2}{\gamma_{ij}^R} \omega([\boldsymbol{C}^R \hat{\boldsymbol{x}}^R]_i)$
$\qquad p_j^I = \beta^I \sum_{i=1}^{r} \frac{(c_{ij}^I)^2}{\gamma_{ij}^I} \omega([\boldsymbol{C}^I \hat{\boldsymbol{x}}^I]_i)$
$\qquad \det \boldsymbol{H}_j = (d_j^{RR} + p_j^R)(d_j^{II} + p_j^I) - (d_j^{RI})^2$
$\qquad \hat{x}_j^R = \hat{x}_j^R - \frac{1}{\det \boldsymbol{H}_j}[(d_j^{II} + p_j^R)(\dot{L}_j^R + \dot{V}_j^R) - d_j^{RI}(\dot{L}_j^I + \dot{V}_j^I)]$
$\qquad \hat{x}_j^I = \hat{x}_j^I - \frac{1}{\det \boldsymbol{H}_j}[-d_j^{RI}(\dot{L}_j^R + \dot{V}_j^R) + (d_j^{II} + p_j^R)(\dot{L}_j^I + \dot{V}_j^I)]$
$\quad$ end
end

F. Number of Holograms Used

In principle, our statistical technique can be applied to data with any number of measurement elements N and to a model with any number of image pixels P. Whereas N is fixed by the choice of the measurement device (*e.g.*, CCD camera pixels), the value of P can be selected by the algorithm designer. A natural choice for P would be the number of CCD elements, which is the size of a single hologram. However, when $N < 2P$ the problem is under-determined, so the regularization term will be particularly important. An alternative is to estimate half as many pixels as there are the CCD elements, *i.e.*, $(P = N/2)$, from a single hologram. However, this option requires interpolation and downsampling processes that might introduce some artifacts in the reconstructed image. To study the effect of the sizes of the reconstructed image relative to the amount of data, we considered the following three different cases in our simulations.

Case 1: Use one hologram to reconstruct a holographic image whose size is half the number of CCD elements (half-size reconstruction), *i.e.*, $P = N/2$.

Case 2: Use one hologram to reconstruct a holographic image whose size is the same as the number of CCD elements (full-size reconstruction), *i.e.*, $P = N$.

Case 3: Use two holograms to reconstruct a holographic image whose size is the same as the number of CCD elements (full-size reconstruction), *i.e.*, $N = 2P$.

6. Simulation Results

In this section, we compare the conventional numerical reconstruction technique with our statistical reconstruction for the three cases described above. Moreover, we examined the effect of possible *a priori* knowledge that the object f is real. We implemented this constraint by zeroing the imaginary part of x^n after each iteration.

A. Effect of Numbers of Data Sets

A 128×128 original image (Fig. 4a) that is complex was degraded by the PSF, interference pattern, and Poisson noise (Fig. 4b) as in Eq. (11). We assumed a space-invariant optical system with a PSF that is a 7×7 jinc function, $\frac{J_1(2\pi r)}{\pi r}$, where J_1 is a Bessel function of the first kind and r is a polar-coordinate parameter, with full width at half maximum (FWHM) of 3.5 pixels. We used the following two 2D reference beams:

$$
\begin{aligned}
u_{r1}(n_1, n_2) &= 200 \exp\left(-\imath \frac{2\pi}{3} n_1\right) \\
u_{r2}(n_1, n_2) &= 150 \exp\left(-\imath \frac{2\pi}{4} n_1\right),
\end{aligned}
\tag{49}
$$

where the pixel indices n_1 and n_2 range from 0 to 127. For experiments with only one hologram, we used the first reference beam. The offset b_i is assigned to be 5 and 10 for the first and second hologram, respectively. The Poisson noise has the peak signal-to-noise ratio (PSNR) of 29 dB and 24 dB in the first and second hologram (Fig. 4b), respectively.

The PSNR in the data is defined as follows:

$$\text{PSNR} \triangleq 10 \log_{10} \left(\frac{\max_i (y_i - b_i)^2}{\frac{1}{N} \sum_{i=1}^{N} (y_i - \text{E}[y_i])^2} \right). \tag{50}$$

Each simulated real-valued hologram has the same size (128×128 pixels) as the original complex-valued image.

Figure 4c shows the conventional reconstruction using an apodizing Gaussian mask. We applied a 41×41 Gaussian mask with FWHM of 27.2 pixels to the selected region in the frequency domain of the hologram. The magnitude and phase of the reconstructed image appear to be blurry while noise still remains. Owing to the effect of the filtering method, noise cannot be removed completely without oversmoothing edges. Figures 4d-f show our statistical holographic reconstruction based on PL estimation for the three different cases. Because of non-convextity of the cost function Φ, the reconstructed image can be quite sensitive to the initial image. We used the image from the conventional approach as the initial image for the iterations. We used the nonquadratic penalty function in Eq. (16) with regularization parameters $\beta^R = \beta^I = 10$ and edge-preserving parameters $\delta^R = \delta^I = 1$. As in most Baysian image reconstruction methods, these parameters are determined by trial and error. The SPS algorithm was run for 200 iterations and its most expensive computation at each iteration includes five 2-D convolutions or five FFT's. Although SPS might not be the fastest existing algorithm for this problem, the focus of the paper is to illustrate the potential of the penalized-likelihood reconstruction technique to digital holography. Unlike the conventional technique, the penalized-likelihood reconstruction technique with a nonquadratic penalty can reduce noise significantly while still preserving edges.

Figure 4d shows the half-size reconstruction using one hologram (Case 1). For display, the reconstructed image was linearly interpolated to match the size of the original image. Figures 4e-f show full-size reconstructions using one and two holograms, respectively (Cases 2 and 3). The half-size reconstruction has less noise but a little more artifacts than the full-size reconstruction using one hologram, perhaps as a result of interpolation and downsampling processes. Figures 5 and 6 show profiles of the magnitude and phase of the reconstructed images. As expected, full-size reconstruction with two holograms yields the best reconstructed image with the smallest normalized root mean-squared error (NRMSE). The NRMSE in percentage is defined as follows:

$$\text{NRMSE} = \frac{\|\hat{\boldsymbol{x}} - \boldsymbol{x}_{\text{true}}\|}{\|\boldsymbol{x}_{\text{true}}\|} \times 100\%, \tag{51}$$

where $\hat{\boldsymbol{x}}$ is the reconstructed image, $\boldsymbol{x}_{\text{true}}$ is the true image, and $\| \cdot \|$ denotes the Euclidean norm.

Figure 7 shows the contours of the marginal objective functions at one particular pixel for the cases of one and two holograms. For this illustration, we examined the noiseless and blurless case without regularization to clearly demonstrate how the statistical technique using two holograms can help reduce non-uniqueness. As shown in Fig. 7a, since there are multiple minimizers, the algorithm converges to an estimate that depends strongly on the initial guess. When two holograms are used, the solutions become more distinct, and thus the algorithm often converges to the desired solution as in Fig. 7b. However, even with two

16

holograms the algorithm can converge to a limit that depends on the initial estimate because the cost function is non-convex and not have local minima. In all cases, our statistical technique decreases the cost function monotonically, although this alone does not ensure convergence to a global minimizer for non-convex cost functions.

B. Real Object Constraint

If the object is known *a priori* to be real, then we can constrain the estimate $\hat{x}$ to be real by zeroing its imaginary part at each iteration. In this case, we may not need to have $N \geq 2P$ since effectively the number of unknowns is reduced by a factor of two. Thus we expect the statistical method to yield similar results for the three cases, except possibly for some artifacts caused by interpolation for the case of half-size reconstruction. A 128×128 real image (Fig. 8a) was degraded using the same parameters as in the previous section. The conventional numerical reconstruction in Fig. 8d is blurry as a result of the Gaussian filter. Figures 8e-g show 200 iterations of penalized-likelihood reconstruction with the non-quadratic penalty function ($\beta^R = \beta^I = 5, \delta^R = \delta^I = 5$). Although the NRMSEs for all three cases differ slightly, they all provide similar reconstructed holographic images with less blur than the conventional reconstruction. Because of the real object constraint, using one hologram appears to be adequate for reconstructing a good holographic image.

7. Conclusions

We have demonstrated the potential for reconstructing a digital holographic image using the proposed statistical technique. Because the method uses all the information in the recorded hologram rather than just one term, this approach can improve the quality of the image relative to the conventional numerical reconstruction technique that uses a spatial filter applied in the spatial frequency domain. Moreover, unlike the conventional approach, our statistical technique is not limited by the assumption of a planar reference beam. Because of the ill conditioning and non-uniqueness of the problem, our statistical holographic reconstruction is based on PL estimation. We constructed a statistical model for this system and developed a monotonic algorithm. Although the unique global minimum is not guaranteed due to the non-convexity of the negative log-likelihood function in this problem, one can partially overcome the problem of multiple minima by using regularization and multiple recorded holograms. For a real object, the realness constraint can be enforced at each iteration so that the algorithm can converge faster and one hologram should be sufficient to yield a good reconstructed image.

Appendix A

For simplicity, we consider only the real or imaginary part and ignore the subscript i in the following proofs.

Lemma 1: If $h(l)$ and $q(l)$ are differentiable and the following three conditions are satisfied:

(C1) $h(m) = q(m)$ for some m
(C2) $\dot{q}(l) \geq \dot{h}(l), \forall l \geq m$
(C3) $\dot{q}(l) \leq \dot{h}(l), \forall l \leq m$,
then it follows that $q(l) \geq h(l), \forall l$ and thus $q(l)$ is a surrogate for $h(l)$, *i.e.*, $q(l) \geq h(l), \forall l$.

Proof: For $l \geq m$, then

$$
\begin{aligned}
q(l) &= q(m) + \int_m^l \dot{q}(t)dt \\
&\geq h(m) + \int_m^l \dot{h}(t)dt = h(l)
\end{aligned}
\tag{A1}
$$

For $l \leq m$, then

$$
\begin{aligned}
q(l) &= h(m) + \int_l^m [-\dot{q}(t)]dt \\
&\geq h(m) + \int_l^m [-\dot{h}(t)]dt = h(l)
\end{aligned}
\tag{A2}
$$

Thus, $q(l) \geq h(l), \forall l$ under the above conditions. □

Lemma 2: If $h(l)$ is differentiable and the following maximum is finite and nonnegative,

$$
c(m) = \max_{l \neq m} \frac{\dot{h}(l) - \dot{h}(m)}{l - m},
$$

then

$$
q(l) = h(m) + \dot{h}(m)(l - m) + \frac{1}{2}c(m)(l - m)^2
\tag{A3}
$$

is a parabolic surrogate for h, *i.e.*, $q(l) \geq h(l), \forall l$.

Proof: Condition (C1) of Lemma 1 is clearly satisfied by q when $l = m$. To prove Condition (C2), for $l \geq m$, we differentiate Eq. (A3) with respect to l and substitute $c(m)$ with the proposed curvature as follows:

$$
\begin{aligned}
\dot{q}(l) &= \dot{h}(m) + c(m)(l - m) \\
&\geq \dot{h}(m) + \frac{\dot{h}(l) - \dot{h}(m)}{l - m}(l - m) = \dot{h}(l).
\end{aligned}
\tag{A4}
$$

Similarly, $\dot{q}(l) \leq \dot{h}(l)$, for $l \leq m$, so Condition (C3) is satisfied. Because all three conditions of Lemma 1 are satisfied, $q(l)$ is a parabolic surrogate for $h(l)$. □

Appendix B

The first derivative of h^o in Eq. (22) or (23) is

$$
\dot{h}^o(l; l^n) = \frac{-2y(l + u^o)[(l^{o,n} + u^o)^2 + b/2]}{k^n[(l + u^o)^2 + b/2]} + 2(l + u^o).
\tag{B1}
$$

Thus we define

$$
f(l) \triangleq \frac{\dot{h}^o(l; l^n) - \dot{h}^o(l^{o,n}; l^n)}{l - l^{o,n}} = \frac{2y}{k^n}\left[\frac{(l + u^o)(l^{o,n} + u^o) - b/2}{(l + u^o)^2 + b/2}\right] + 2.
\tag{B2}
$$

To obtain the maximum of the above continuous function, we equate the first derivative to zero:

$$\dot{f}(l) = \frac{2y}{k^n} \left[\frac{-(l^{o,n} + u^o)(l + u^o)^2 + b(l + u^o) + \frac{b}{2}(l^{o,n} + u^o)}{[(l + u^o)^2 + b/2]^2} \right] = 0. \qquad (B3)$$

Then the optimal l^* that yields the maximum is

$$l^* = \frac{b + [b^2 + 2b(l^{o,n} + u^o)^2]^{1/2}}{2(l^{o,n} + u^o)} - u^o \qquad (B4)$$

and

$$f(l^*) = \frac{2y[b^2 + 2b(l^{o,n} + u^o)^2]^{1/2}(l^{o,n} + u^o)^2}{k^n \left[b^2 + b \left(2(l^{o,n} + u^o)^2 + [b^2 + 2b(l^{o,n} + u^o)^2]^{1/2} \right) \right]} + 2 \qquad (B5)$$

is the curvature of the parabolic surrogate function.

Acknowledgment

References

1. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1996).
2. D. Gabor, "A New Microscope Principle," Nature **161**, 777–778 (1948).
3. E. N. Leith and J. Upatnieks, "Reconstructed Wavefronts and Communication Theory," J. Opt. Soc. Am. A **52**, 1123–1130 (1962).
4. H. Chen, M. Shih, E. Arons, E. Leith, J. Lopez, D. Dilworth, and P. C. Sun, "Electronic holographic imaging through living human tissue," Applied Optics **33**, 3630–3632 (1994).
5. E. Leith, H. Chen, Y. Chen, D. Dilworth, J. Lopez, R. Masri, J. Rudd, and J. Valdmanis, "Electronic holography and speckle methods for imaging through tissue using femtosecond gated pulses," Applied Optics **30**, 4204–4210 (1991).
6. Y. Takaki and H. Ohzu, "Fast numerical reconstruction technique for high-resolution hybrid holographic microscopy," Applied Optics **38**, 2204–2211 (1999).
7. Y. Takaki, H. Kawai, and H. Ohzu, "Hybrid holographic microscopy free of conjugate and zero-order images," Applied Optics **38**, 4990–4996 (1999).
8. E. Cuche, F. Bevilacqua, and C. Depeursinge, "Digital holography for quantitative phase-contrasting imaging," Optics Letters **24**, 291–293 (1999).
9. E. Cuche, P. Marquet, and C. Depeursinge, "Spatial filtering for zero-order and twin-image elimination in digital off-axis holography," Applied Optics **39**, 4070–4075 (2000).
10. I. Yamaguchi and T. Zhang, "Phase-shifting digital holography," Optics Letters **22**, 1268–1270 (1997).
11. S. De Nicola, P. Ferraro, A. Finizio, and G. Pierattini, "Wave front reconstruction of Fresnel off-axis holograms with compensation of aberrations by means of phase-shifting digital holography," Optics and Lasers in Engineering **37**, 331–340 (2002).

12. M. Liebling, T. Blu, E. Cuche, P. Marquet, C. Depeursinge, and M. Unser, "A novel non-diffractive reconstruction method for digital holographic microscopy," in *Proc. IEEE Int'l Symp. on Biomedical Imaging* (2002), pp. 625–628.

13. C. A. Bouman and K. Sauer, "A Unified Approach Statistical Tomography Using Co-ordinate Descent Optimization," IEEE Trans. Image Processing **5**, 480–492 (1996).

14. H. Erdoğan and J. A. Fessler, "Monotonic Algorithms for Transmission Tomography," IEEE Trans. Med. Imaging **18**, 801–814 (1999).

15. J. A. Conchello, "Superresolution and convergence properties of the expectation-maximization algorithm for maximum-likelihood deconvolution of incoherent images," J. Opt. Soc. Am. A **15**, 2609–2619 (1998).

16. T. J. Holmes, "Maximum-likelihood image restoration adapted for noncoherent optical imaging," J. Opt. Soc. Am. A **5**, 666–673 (1988).

17. M. Çetin, W. C. Karl, and A. S. Willsky, "Edge-preserving image reconstruction for coherent imaging applications," in *Proc. IEEE Int'l Conf. on Image Processing* (2002), vol. 2, pp. 481–484.

18. B. W. Silverman, *Density estimation for statistics and data analysis* (Chapman and Hall, New York, 1986).

19. J. A. Fessler and A. O. Hero, "Penalized Maximum-Likelihood Image Reconstruction Using Space-Alternating Generalized EM Algorithms," IEEE Trans. Image Processing **4**, 1417–1429 (1995).

20. H. L. Van Trees, *Detection, estimation, and modulation theory* (Wiley, New York, 1968).

21. J. L. Marroquin and M. Tapia, "Parallel algorithms for phase unwrapping based on Markov random field models," J. Opt. Soc. Am. A **12**, 2578–2585 (1995).

22. W. Kim and M. H. Hayes, "Phase retrieval using two Fourier-transform intensities," J. Opt. Soc. Am. A **7**, 441–449 (1990).

23. S. Sotthivirat and J. A. Fessler, "Image Recovery Using Partitioned-Separable Paraboloidal Surrogate Coordinate Ascent Algorithms," IEEE Trans. Image Processing **11**, 306–317 (2002).

24. E. Arons and E. Leith, "Coherence confocal-imaging system for enhanced depth discrimination in transmitted light," Applied Optics **35**, 2499–2506 (1996).

25. P.-C. Sun and E. N. Leith, "Broad-source image plane holography as a confocal imaging process," Applied Optics **33**, 597–602 (1994).

26. Y. Censor, "Finite series expansion reconstruction methods," Proc. IEEE **71**, 409–419 (1983).

27. D. L. Snyder, A. M. Hammoud, and R. L. White, "Image recovery from data acquired with a charge-coupled-device camera," J. Opt. Soc. Am. A **10**, 1014–1023 (1993).

28. M. Cetin, W. C. Karl, and A. S. Willsky, "Edge-preserving image reconstruction for coherent imaging applications," in *Proc. IEEE Intl. Conf. on Image Processing* (2002), vol. 2, pp. 481–4.

29. M. Çetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," IEEE Trans. Image Processing **10**, 623–631 (2001).

30. K. Lange, "Convergence of EM Image Reconstruction Algorithms with Gibbs Smooth-

ing," IEEE Trans. M. Imaging **9**, 439–446 (1990).

31. J. A. Fessler, "Grouped Coordinate Descent Algorithms for Robust Edge-Preserving Image Restoration," in *Image reconstruction and restoration II*, (ed. T. J. Schulz) (Proc. SPIE 3170, 1997), pp. 184–194.

32. P. J. Huber, *Robust Statistics* (Wiley, New York, 1981).

33. J. A. Fessler, "Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction," IEEE Trans. Med. Imaging **16**, 166–175 (1997).

34. A. R. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," IEEE Trans. Med. Imaging **14**, 132–137 (1995).

35. S. Sotthivirat, *Statistical Image Recovery Techniques for Optical Imaging Systems*, Ph.D. thesis, University of Michigan, Ann Arbor, MI 48109-2122 (2003).

List of Figure Captions

Fig. 1. Diagram of digital holography.

Fig. 2. Holographic reconstruction using a filtering method.

Fig. 3. Illustration of the marginal cost, $h_i(l^R, 0)$, and surrogate functions as a function of l^R. The solid line is the original marginal cost function. The other two lines lying above the cost function are the surrogate functions. The function with the dashed line is called the paraboloidal surrogate function which has the same first derivative and the same point as the original cost function at $l = l^n$.

Fig. 4. Holographic reconstruction of a complex object. The top image of each pair represents the magnitude of the image and the bottom image represents the phase of the image, except for the hologram data in (b). (a) Original image. (b) Two different holograms. (c) Conventional reconstruction using an apodizing Gaussian filter (NRMSE=40.0%). (d) Half-size penalized-likelihood reconstruction using one hologram (NRMSE=17.5%). Linear interpolation in the vertical direction to the same size as the original image is performed for display. (e) Full-size penalized-likelihood reconstruction using one hologram (NRMSE=17.3%). (f) Full-size penalized-likelihood reconstruction using two holograms (NRMSE=14.1%).

Fig. 5. Profiles of the magnitude of the numerical reconstructed images across the second row of circles.

Fig. 6. Profiles of the phase of the numerical reconstructed images across the second row of circles.

Fig. 7. Contours of the marginal objective functions at one pixel when (a) using one hologram and (b) using two holograms for full-size reconstruction. The "x" mark indicates the optimal solution at $20 + i110$ and the "o" marks indicate the updates of the estimates starting at $150 + i150$.

Fig. 8. Penalized-likelihood reconstruction of a real object using the real object constraint. (a) Original image. (b) and (c) Hologram data. (d) Conventional reconstruction using an apodizing Gaussian filter (NRMSE=43.8%). (e) Half-size penalized-likelihood reconstruction using one hologram (NRMSE=22.8%). (f) Full-size penalized-likelihood reconstruction using one hologram (NRMSE=21.1%). (g) Full-size penalized-likelihood reconstruction using two holograms (NRMSE=17.2%).

Fig. 1.

Recorded Hologram   Magnitude of Fourier Transform   Reconstructed Holographic Image

Fig. 2.

Fig. 3.

(a)   (b)   (c)

(d)   (e)   (f)

Fig. 4.

Fig. 5.

Fig. 6.

26

(a)

(b)

Fig. 7.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

Fig. 8.