

Compressive holography of diffuse objects

Kerkil Choi,¹ Ryoichi Horisaki,² Joonku Hahn,¹ Sehoon Lim,¹ Daniel L. Marks,¹
Timothy J. Schulz,³ and David J. Brady^{1,*}

¹The Fitzpatrick Institute for Photonics, Department of Electrical and Computer Engineering,
Duke University, Durham, North Carolina 27708, USA

²Department of Information and Physical Sciences, Graduate School of Information Science and Technology,
Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

³Department of Electrical and Computer Engineering, Michigan Technological University,
Houghton, Michigan 49931, USA

*Corresponding author: david.brady@duke.edu

Received 4 March 2010; accepted 13 May 2010;
posted 28 May 2010 (Doc. ID 124904); published 23 June 2010

We propose an estimation-theoretic approach to the inference of an incoherent 3D scattering density from 2D scattered speckle field measurements. The object density is derived from the covariance of the speckle field. The inference is performed by a constrained optimization technique inspired by compressive sensing theory. Experimental results demonstrate and verify the performance of our estimates. © 2010 Optical Society of America

OCIS codes: 090.1995, 070.0070.

1. Introduction

Speckle imaging has been extensively studied in various areas. Speckle interferometry has been investigated in astronomical imaging [1,2]. Maximum-likelihood estimation approaches have been developed for radar imaging for structured covariance estimation based on the statistical model of speckle [3,4]. In medical imaging, several methods for speckle reduction have been proposed [5,6].

In [7], we demonstrated 3D tomographic reconstruction of a scattering object from its 2D digital hologram. The 3D scattering density was related to the 2D hologram via the first-order Born approximation. We successfully reconstructed a 3D tomographic volume of an object consisting of two dandelion parachutes whose number of voxels is 10 times larger than that of the measurements. This demonstration illustrated the fact that 2D holographic measurements naturally encode compressive 3D density information.

In [7], the scattering density of an object is assumed to be sparse in some basis. This assumption, however, is not valid for the field scattered from a diffuse object because its scattered field is random. The random field manifests itself as speckle [8]. This prevents us from applying the decompressive inference developed in [7] directly to the estimation of the scattering density of a diffuse object.

Nonetheless, we may assume that the *incoherent* scattering density of a diffuse object is sparse in some basis. From the theory of random processes, it is straightforward to see that the incoherent density of an object is proportional to the variance of the scattered speckle field of the object [8]. Based on this theory, this paper considers reconstruction of the incoherent density (i.e., intensity) of a diffuse object from its scattered fields collected in the same manner as in compressive holography.

The field scattered from a diffuse object can be modeled as complex circular Gaussian. This allows for building a statistical model between the incoherent scattering density of a diffuse object and the scattered field detected on the CCD, by integrating the statistical model into the propagation system as

described in Section 2. This model forms a hypothetical forward mapping that underpins our estimation. We show that the mapping often forms an ill-conditioned linear system and, thus, effectively results in measurement incompleteness when small singular values are ignored to produce numerically stable solutions. We resolve such measurement incompleteness by a constrained optimization technique enforcing sparsity constraints, which is inspired by compressive sensing theory.

Our inference is based on the estimation of the diagonal covariance [3] of the random process of the scattered speckle field. In particular, the diagonal “image” is constrained to form a sparse signal in the variational domain. This sparsity constraint is incorporated by enforcing minimum total variation (TV) of the estimate of the image.

Our main result is that one can reconstruct the 3D incoherent density of a diffuse object from its 2D speckle fields. We confirm this result experimentally in reconstructions with μm scale transverse resolution and mm scale axial resolution.

The paper is organized as follows. Our estimation problem is formulated in Section 2, and the estimation theory is developed in Subsection 2.D. Section 3 demonstrates and discusses our results with experimental data.

2. Problem Formulation

A. Continuous Forward Model: First Born Approximation

To retrieve a complex field, an interference pattern between a plane wave A and a 3D object with scattering density $\eta(x', y', z')$ is formed. Then a 2D detector array records the irradiance

$$\begin{aligned} I(x, y) &= |Ae^{j\alpha x} + E(x, y)|^2 \\ &= |A|^2 + |E(x, y)|^2 + Ae^{-j\alpha x}E(x, y) \\ &\quad + Ae^{j\alpha x}E^*(x, y), \end{aligned} \quad (1)$$

where the linear phase $\exp(j\alpha x)$ is introduced to shift the signal terms (e.g., E , E^* , and $|E|^2$ in the spatial frequency domain. Note that the introduction of a linear phase term in the y direction is omitted for brevity of analysis. The scattered field E is defined under the first Born approximation as

$$E(x, y) = \iiint dx' dy' dz' \eta(x', y', z') h(x - x', y - y', z - z'), \quad (2)$$

where h is the inverse Fourier transform of the propagation transfer function $\exp(i(z - z')\sqrt{k^2 - k_x^2 - k_y^2})$ [see Ref. [9] Eqs. (3–74)]. In Eq. (1), the DC term $|A|^2$, the autocorrelation term $|E(x, y)|^2$, and the twin-image term $Ae^{j\alpha x}E^*(x, y)$ are extracted and removed from the interference irradiance to isolate the complex field $A^*E(x, y)$ by utilizing the Leith–Upatnieks off-axis holography approach [10]. Then we may pro-

ceed with $Ae^{-j\alpha x}E(x, y)$. Furthermore, we may neglect the global effect of the reference $Ae^{j\alpha x}$ in the following analysis without loss of generality.

B. Discrete Forward Model

Let a 3D object (i.e., the scattering potential) be denoted by $\eta(x', y', z')$ with the convention that $z' = 0$ at the detector plane. When the object being illuminated is diffusive, as in this study, the scattering potential represents reflectance [3,8]. Let the sample spacings be $\Delta_x = \Delta_y = \Delta$. Also, let Δ_z be the sampling pitch in the z axis. Let the number of pixels along each dimension of the detector be N .

The sampled field on the detector with the given sample spacings can be written as [7]

$$\begin{aligned} E_{n_1 n_2} &= E(n_1 \Delta, n_2 \Delta) = \frac{1}{(2\pi)^2} \int \cdots \int dz' dx dy dk_x dk_y \\ &\quad \times \iint dx' dy' \eta(x', y', z') e^{-i(k_x x' + k_y y')} \\ &\quad \times \sum_{m'_1} \sum_{m'_2} \delta(x' - m'_1 \Delta) \delta(y' - m'_2 \Delta) \\ &\quad \times e^{iz' \sqrt{k^2 - k_x^2 - k_y^2}} \delta(z - z') e^{-i(k_x x + k_y y)} \\ &\quad \times \delta(x - n_1 \Delta) \delta(y - n_2 \Delta) \sum_{m'_1} \sum_{m'_2} \delta(k_x - m'_1 \Delta_k) \\ &\quad \times \delta(k_y - m'_2 \Delta_k) \sum_l \delta(z - l \Delta_z) \\ &= \frac{1}{N^2} \sum_l \sum_{m_1} \sum_{m_2} \left[\sum_{m'_1} \sum_{m'_2} \eta_{m'_1 m'_2 l} e^{-i2\pi \frac{m_1 m'_1 + m_2 m'_2}{N}} \right] \\ &\quad \times e^{il \Delta_z \sqrt{k^2 - m_1^2 \Delta_k^2 - m_2^2 \Delta_k^2}} e^{-i2\pi \frac{n_1 m_1 + n_2 m_2}{N}}, \end{aligned} \quad (3)$$

where $\eta_{m'_1 m'_2 l} = \eta(m'_1 \Delta, m'_2 \Delta, l \Delta_z)$ and $N = 1/(\Delta)$. This equation can simply be rewritten as

$$E_{n_1 n_2} = \mathcal{F}_{2D}^{-1} \left\{ \sum_l \hat{\eta}_{m_1 m_2 l} e^{il \Delta_z \sqrt{k^2 - m_1^2 \Delta_k^2 - m_2^2 \Delta_k^2}} \right\}, \quad (4)$$

where $\hat{\eta}$ denotes the Fourier transform of η and \mathcal{F}^{-1} denotes the inverse Fourier transform operator. Considering the linearity of Eq. (4), the isolated complex field measurement may be algebraically written as

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{w}, \quad (5)$$

where $\mathbf{g} \in \mathbb{C}^M$ and $\mathbf{f} \in \mathbb{C}^N$ are vectorized measurement and object elements and \mathbf{w} denotes additive noise.

A coherent optical system maps the object field f onto measurements g , according to Eq. (5). When an object is diffusive, the speckle phenomena are fundamental, as noted in studies performed in various fields [5,11,12]. Thus, both the object field and the measurements are modeled as random fields as described in the following section.

C. Statistical Model

Let g_k denote the k th 2D scattered speckle field detected on a 2D focal plane array (typically a CCD). Also, let f_k denote the k th 3D scattered speckle field of a diffuse object, induced by the 3D object density. Assume that $g_k \in \mathbb{C}^M$ contains $N_x(N_y)$ pixels in the $x(y)$ direction, respectively. Also, assume that $f_k \in \mathbb{C}^N$ contains $N_x(N_y)$ pixels in the $x(y)$ direction, respectively, and N_z pixels in the propagation (z) direction, such that $M = N_x \times N_y$, and $N = N_x \times N_y \times N_z$. Using Eq. (4), the k th measurements on the CCD can be rewritten as

$$g_k = G_{2D} Q B f_k + w_k = H f_k + w_k, \quad (6)$$

where $g_k \in \mathbb{C}^{M \times 1}$, w_k denotes independent additive Gaussian noise, and G_{2D} represents a 2D inverse DFT matrix. $B = \text{bldiag}(F_{2D}, F_{2D}, \dots, F_{2D})$, F_{2D} denotes a matrix representing the 2D DFT whose size is $M \times M$, and “bldiag” denotes the block-diagonal matrix. $Q = [P_1 P_2 \dots P_{N_z}]$ with $P_l = \text{diag}(\text{vec}(S_l))$. The spherical phase “image” S_l associated with the l th transverse slice is defined by

$$[S_l]_{m_1 m_2} = \exp(ik_0 l \Delta_z) \times \exp\left(il \Delta_z \sqrt{k_0^2 - m_1^2 \Delta_k^2 - m_2^2 \Delta_k^2}\right), \quad (7)$$

where k_0 is the wavenumber, $\text{vec}(A)$ represents a column vector formed by stacking columns of A [13], $\text{diag}(a)$ represents a diagonal matrix formed by placing the vector a on the diagonal, and Δ_k and Δ_z represent the transverse and axial sample spacings, respectively.

Let $f = [f_1 f_2 \dots f_N]^T$ with $N = N_x \times N_y \times N_z$. For fully developed speckles, the 3D scattered field f has the complex circular Gaussian probability density [14]:

$$p(f) = \frac{1}{\pi^N \det(R_f)} (-f^H R_f^{-1} f), \quad (8)$$

where $\mathbb{E}[f] = 0$, $R_f = \mathbb{E}[f f^H] = \text{diag}(\mathbb{E}[|f_1|^2], \mathbb{E}[|f_2|^2], \dots, \mathbb{E}[|f_N|^2]) = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_N) = \text{diag}(\alpha)$. Symbolically, we write $f \sim CN(0, R_f)$. Note that α_n is always nonnegative because $\alpha_n = [R_f]_{nn} = \mathbb{E}[|f_n|^2]$. Define the covariance of w as $\sigma^2 I$, with I denoting the $M \times M$ identity matrix. This statistical model implies that g also has a complex Gaussian probability density with zero mean and covariance $R_g: g \sim CN(0, R_g)$, with $R_g = H R_f H^H + \sigma^2 I = H(R_f + \sigma^2 I)H^H$. Note that $H H^H = I$, which implies $\text{rank}(H) = M$. Our objective is to estimate the incoherent scattering density α from several g_k .

D. Brief Notes on Challenges

As indicated by the models described above, we clearly have two challenges. The first challenge is the randomness of the fields. As discussed above, such

randomness causes fundamental phenomena of speckle. Speckle is considered undesirable because it degrades visual quality and, consequently, resolution of the reconstructed fields. Traditionally, the speckle artifacts are reduced by measuring multiple realizations of the fields with some diversity and by averaging the multiple fields on an intensity basis [8].

Another challenge is the underdeterminacy of the system in Eq. (6). In Eq. (6), the dimensions of the system matrix H clearly indicate that the inverse problem involved is underdetermined, despite the fact that multiple speckle fields are measured, which only helps resolve issues with randomness of the fields. This incompleteness of measurements makes it difficult to invert Eq. (6) in a stable and reliable way.

Motivated by these two challenges and inspired by the recent development of compressive sampling theory, we try to resolve the two challenges using an estimation-theoretic approach. We emphasize that our problem is “compressive” in the sense that the measurement matrix H is underdetermined, despite the multiple measurements. This compressive nature of the problem and the proposed approach is indicated by the title.

3. Estimation Methods

A. Incoherent Image Estimation

There can be several ways to estimate α , or equivalently R_f , from g_k . For example, one can estimate R_f from R_g by inverting $R_g = H R_f H^H + \sigma^2 I$ with the constraint that R_f is a diagonal matrix. Another approach to estimating R_f can be maximum-likelihood estimation, as in [3]. However, all these methods require an impractical amount of computational and time resources. Therefore, a mapping is considered to significantly reduce computational costs. The choice of mapping in this paper is inspired by spectrum estimation theory [15]. However, mappings other than that in Eq. (9) may also be considered.

Because $H H^H = I$, the minimum-norm solution [13] to the inverse problem in Eq. (6) is given by

$$\hat{f} = H^H (H H^H)^{-1} g = H^H g = H^H H f + H^H w, \quad (9)$$

from which a quadratic mapping can be formed from K independent measurements as

$$\hat{s}_n = \frac{1}{K} \sum_{k=1}^K |\hat{f}_{nk}|^2. \quad (10)$$

This process is analogous to the periodogram estimation for spectrum estimation [15], motivating us to consider the expected value of the estimate \hat{s} . Let $H = [h_1 h_2 \dots h_N]$, where h_n denotes the n th column of H . Then, $[H^H H]_{mn} = \langle h_m, h_n \rangle = h_m^H h_n$ with $\langle a, b \rangle$ denoting the inner product between a and b . Hence, the expected value of \hat{s} can be expressed as

$$\begin{aligned}
E[\hat{s}] &= [E[\hat{s}_1]E[\hat{s}_2] \cdots E[\hat{s}_N]]^T \\
&= \frac{1}{K} \sum_{k=1}^K \text{Diag}(E[\hat{f}_k \hat{f}_k^H]) = \text{Diag}(E[\hat{f} \hat{f}^H]) \\
&= \text{Diag}(E[(\mathbf{H}^H \mathbf{H} \mathbf{f} + \mathbf{H}^H \mathbf{w})(\mathbf{H}^H \mathbf{H} \mathbf{f} + \mathbf{H}^H \mathbf{w})^H]) \quad (11) \\
&= \text{Diag}(\mathbf{H}^H \mathbf{H} E[\mathbf{f} \mathbf{f}^H] \mathbf{H}^H \mathbf{H} + \mathbf{H}^H E[\mathbf{w} \mathbf{w}^H] \mathbf{H}) \\
&= \text{Diag}(\mathbf{H}^H \mathbf{H} \mathbf{R}_f \mathbf{H}^H \mathbf{H} + \sigma^2 \mathbf{H}^H \mathbf{H}) \\
&= \text{Diag}(\mathbf{H}^H \mathbf{H} \mathbf{R}_f \mathbf{H}^H \mathbf{H}) + \sigma^2 \text{Diag}(\mathbf{H}^H \mathbf{H}),
\end{aligned}$$

where the third equality holds because the probability density of \hat{f}_k is the same for all k . In particular, the last equality in Eq. (11) indicates that

$$\begin{aligned}
E[\hat{s}_n] &= \sum_{m=1}^N |\langle \mathbf{h}_n, \mathbf{h}_m \rangle|^2 \alpha_m + \sigma^2 \langle \mathbf{h}_n, \mathbf{h}_n \rangle \\
&= \sum_{m=1}^N |\mathbf{h}_n^H \mathbf{h}_m|^2 \alpha_m + \sigma^2 \|\mathbf{h}_n\|^2, \quad (12)
\end{aligned}$$

where $\|\mathbf{h}\|$ denotes the Euclidean norm of \mathbf{h} . Equation (12) implies that

$$\mathbf{d} = E[\hat{s}] = \mathcal{B} \boldsymbol{\alpha} + \sigma^2 \boldsymbol{\omega}, \quad (13)$$

where \mathbf{d} are defined as *synthetic measurements*, the elements of $\mathcal{B} \in \mathbb{R}^{N \times N}$ are formed by squaring the magnitudes of the corresponding elements of $\mathbf{H}^H \mathbf{H}$: $\mathcal{B}_{nm} = |\langle \mathbf{h}_n, \mathbf{h}_m \rangle|^2$, and $\boldsymbol{\omega} = [\|\mathbf{h}_1\|^2, \|\mathbf{h}_2\|^2, \dots, \|\mathbf{h}_N\|^2]^T$.

Equation (13) suggests that an estimate of $\boldsymbol{\alpha}$ may be obtained by solving a constrained convex optimization problem defined by

$$\begin{aligned}
\boldsymbol{\alpha}^* &= \arg \min_{\boldsymbol{\alpha}} \Phi(\boldsymbol{\alpha}), \\
\text{subject to } &\|\mathbf{d} - \mathcal{B} \boldsymbol{\alpha}\|_2^2 < \epsilon(\sigma^2), \quad (14)
\end{aligned}$$

for some small $\epsilon(\sigma^2)$ that depends on noise level σ . Alternatively, we may solve the Lagrangian unconstrained formulation of the constrained optimization problem in Eq. (14):

$$\boldsymbol{\alpha}^* = \arg \min_{\boldsymbol{\alpha}} \frac{1}{2} \|\mathbf{d} - \mathcal{B} \boldsymbol{\alpha}\|_2^2 + \beta \Phi(\boldsymbol{\alpha}). \quad (15)$$

The functional $\Phi(\boldsymbol{\alpha})$ imposes constraints on $\boldsymbol{\alpha}$. The parameter β controls the relative strength of the constraints to the data fidelity term. In particular, for decompressive inference, Φ enforces the sparsity constraints. Typical choices for a sparse basis include a particular wavelet basis. When an orthogonal sparse basis Ψ is chosen, the sparsity is enforced by minimizing $\Phi(\boldsymbol{\theta}) = \|\boldsymbol{\theta}\|_1 = \sum_n |\theta_n|$, where $\boldsymbol{\theta} = \Psi \boldsymbol{\alpha}$. Another choice for the sparsity constraints is the TV. When the incoherent scattering density is expected to have a smooth surface, the density may be assumed to be sparse in the variational (e.g., gradient)

domain. This sparsity can be incorporated by minimizing TV of the estimate. We define TV for our problem as $\Phi(\boldsymbol{\alpha}) = \|\boldsymbol{\alpha}\|_{\text{TV}} = \sum_{n_x} \sum_{n_y} \sum_{n_z} |\nabla(\boldsymbol{\alpha}_{n_z})_{n_x, n_y}|$, where $\boldsymbol{\alpha}_{n_z}$ denotes the n_z th 2D transverse slice of the 3D $\boldsymbol{\alpha}$, and $|\nabla(\boldsymbol{\alpha}_{n_z})_{n_x, n_y}|$ denotes the magnitude of the 2D gradient vector at location (n_x, n_y) in the n_z th transverse slice.

We adapt the two-step iterative shrinkage/thresholding algorithm to solve the optimization problem in Eq. (15). It is worth mentioning that a choice of the sparsity constraints may have a significant impact on the accuracy of estimates [16,17]. As compressive sensing theory indicates, Φ should be chosen such that the mutual coherence is small enough to ensure the accurate reconstructions for the given number of measurements. Also, Φ should produce a representation that is as sparse as possible to ensure the optimal accuracy of the estimate. Developing or choosing such sparse bases is a challenging problem and is an active research area [18].

B. On Some Properties of \mathcal{B}

For \mathbf{H} , the number of rows M is typically several hundreds of thousands, and the number of columns N is also several hundreds of thousands. Hence, the size of \mathcal{B} is as large as hundreds of thousands times hundreds of thousands. In general, this causes a storage problem. Even when large memory is available, the operations involving the matrix and the associated vectors are computationally extensive and, thus, consume impractical amount of time resource. Fortunately, \mathcal{B} has useful properties, which allow for efficient computations by using fast Fourier transforms (FFTs).

Recall that

$$\mathbf{H} = \mathbf{G}_{2D} \mathbf{Q} \mathbf{B} = [\mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_{N_c}]. \quad (16)$$

\mathbf{H}_k is a Toeplitz-block-Toeplitz matrix that is approximated by a circulant-block-circulant (CBC) matrix [13]. Thus, it can be decomposed as $\mathbf{H}_k = \mathbf{G}_{2D} \mathbf{P}_k \mathbf{F}_{2D} = \mathbf{F}_{2D}^H \mathbf{P}_k \mathbf{F}_{2D}$. Note that \mathbf{P}_k is a diagonal matrix. Therefore, we need to store only the diagonal elements of \mathbf{P}_k . \mathbf{H}_k can be readily retrieved from the diagonal elements by using FFTs. \mathcal{B} also has a similar, useful property. The following lemma and theorem are useful for clarifying the property.

Lemma 3.1. Given a matrix \mathbf{A} , define \mathbf{B} as a matrix that is constructed by squaring the absolute value of each element of \mathbf{A} : $[\mathbf{B}]_{mn} = |[A]_{mn}|^2$. If \mathbf{A} is CBC, then so is \mathbf{B} .

Proof: see Appendix A.

Note that $[\mathcal{B}]_{mn} = |[\mathbf{H}^H \mathbf{H}]_{mn}|^2$.

Theorem 3.2. Each block $\mathbf{H}_1^H \mathbf{H}_2$ of $\mathbf{H}^H \mathbf{H}$ is a CBC matrix. Hence, $\mathbf{H}^H \mathbf{H}$ is a circulant-block-circulant-block (CBCB) matrix, and thus, \mathcal{B} is a CBCB matrix as well.

Proof: see Appendix B.

This theorem implies that each block of \mathcal{B} can also be decomposed as $\mathcal{B}_{l_1 l_2} = \mathbf{F}_{2D}^H \boldsymbol{\Upsilon}_{l_1 l_2} \mathbf{F}_{2D}$, where $\boldsymbol{\Upsilon}_{l_1 l_2} = \text{diag}(v_1^{l_1 l_2}, v_2^{l_1 l_2}, \dots, v_M^{l_1 l_2})$. Therefore, if

$v_1^{l_1 l_2}, v_2^{l_1 l_2}, \dots, v_M^{l_1 l_2}$ can be computed efficiently and fast, then \mathcal{B} can also be efficiently stored and retrieved. In addition, computations pertaining to \mathcal{B} can be efficiently performed using FFTs. The following corollary shows how the eigenvalues can be computed using FFTs.

Corollary 3.3. Let $Y_{l_1 l_2}$ denote an $N_x \times N_y$ 2D “image” obtained by multiplying the two quadratic phase images S_{l_1} and S_{l_2} defined in Eq. (7):

$$\begin{aligned} [Y_{l_1 l_2}]_{m_1 m_2} &= [S_{l_1}^* S_{l_2}]_{m_1 m_2} \\ &= \exp(ik(l_2 - l_1)\Delta_z) \exp(i(l_2 - l_1)\Delta_z \sqrt{k^2 - m_1^2 \Delta_k^2 - m_2^2 \Delta_k^2}). \end{aligned} \quad (17)$$

Define $W_{l_1 l_2}$ as the 2D inverse DFT of $Y_{l_1 l_2}$: $W_{l_1 l_2} = \mathcal{F}_{2D}^{-1}(Y_{l_1 l_2})$ with \mathcal{F}_{2D}^{-1} denoting a 2D inverse Fourier transform operator. Let $[Z_{l_1 l_2}]_{m_1 m_2} = |[W_{l_1 l_2}]_{m_1 m_2}|^2$. Then, the eigenvalues, $v_1^{l_1 l_2}, v_2^{l_1 l_2}, \dots, v_M^{l_1 l_2}$, of $\mathcal{B}_{l_1 l_2}$ can be obtained by taking a 2D DFT of $Z_{l_1 l_2}$.

Proof: see Appendix C.

Thus, these diagonal blocks of \mathcal{B} can be efficiently computed using 2D FFTs prior to the iterative data processing step.

Another useful property of $\mathcal{B}_{l_1 l_2}$ is that $\mathcal{B}_{l_1 l_2}^H$ is also a CBC matrix because $\mathcal{B}_{l_1 l_2}^H = (F_{2D}^H \Upsilon_{l_1 l_2} F_{2D})^H = F_{2D}^H \Upsilon_{l_1 l_2}^H F_{2D}$, which shows that $\mathcal{B}_{l_1 l_2}^H$ has the same form of decomposition and, therefore, is also a CBC matrix.

The iterative algorithm to solve Eq. (15) requires many evaluations of $\mathcal{B}x$ and $\mathcal{B}^H y$, for some $N \times 1$ vectors x and y . These multiplications are the most time-consuming computational tasks, required at every iteration. The properties of \mathcal{B} and \mathcal{B}^H discussed above enable efficient and practically affordable computations of $\mathcal{B}x$ and $\mathcal{B}^H y$ via the use of 2D FFTs.

We roughly estimate the computational cost required by our proposed method. Let $N_{\max} = \max(N_x, N_y)$. Assume that the eigenvalues of all $\mathcal{B}_{l_1 l_2}$ are computed and stored prior to the execution of the iterative algorithms. The computations of $\mathcal{B}_{l_1 l_2} a$ and $\mathcal{B}_{l_1 l_2}^H b$ require $O(N_{\max}^2 \log(N_{\max}))$ opera-

tions. Also, there are N_z^2 of such matrix-vector multiplications. Therefore, the computational complexity for each iteration is approximately $N_z^2 \times O(N_{\max}^2 \log(N_{\max}))$. In particular, if $N_z \ll N_{\max}$, as in our problem, the computational complexity per iteration is roughly $O(N_{\max}^2 \log(N_{\max}))$.

Another important property of \mathcal{B} is its rank. The rank is a quantity that characterizes the invertibility of the system or the stability of its inversion. As described in Eq. (16), H is rank deficient, and so is $H^H H$: the ranks of both H and $H^H H$ are M . However, the matrix \mathcal{B} can have a higher rank than $H^H H$ does because of the nonlinear mapping between \mathcal{B} and $H^H H$. Namely, the nonlinear operation that squares the absolute value of each element of $H^H H$ creates a new matrix, \mathcal{B} , that has a higher rank. This numerical phenomenon regarding the rank is generally matrix dependent.

To illustrate this rank phenomenon, we create some example matrices in Fig. 1. In Fig. 1(a), we define $H = A \in \mathbb{R}^{100 \times 1000}$ as a matrix whose elements are independent identically distributed Gaussian random variables with zero mean and unit variance. The rank of A is 100. Its Gram matrix $A^H A \in \mathbb{R}^{1000 \times 1000}$ also has the rank 100. Finally, the matrix $\mathcal{B} = B \in \mathbb{R}^{1000 \times 1000}$, obtained by taking the absolute value of each element of A and by squaring the absolute values, has full rank: $\text{rank}(\mathcal{B}) = 1000$. This rank relationship is illustrated by comparing their associated singular value spectra and *effective condition numbers* denoted by $\text{econd}(M)$ and defined as

$$\text{econd}(M) = \frac{\max\{\sigma(M) | \sigma(M) > \epsilon\}}{\min\{\sigma(M) | \sigma(M) > \epsilon\}}, \quad (18)$$

where $\max\{\sigma(M) | \sigma(M) > \epsilon\}$ and $\min\{\sigma(M) | \sigma(M) > \epsilon\}$ represent the maximum and minimum *numerically nonzero* singular values of the matrix M , respectively. ϵ was chosen to be 10^{-8} .

Figure 1(b) shows the singular value spectra of \mathcal{B} associated with two examples of H , defined in Section 2, for small M and N : $M = 225$ and $N = 900$. In particular, two examples B_1 and B_2 of \mathcal{B} are created to compare the condition behavior of \mathcal{B} for different choices of the axial sample spacing. Specifically, the axial sample spacing (i.e., the distance)

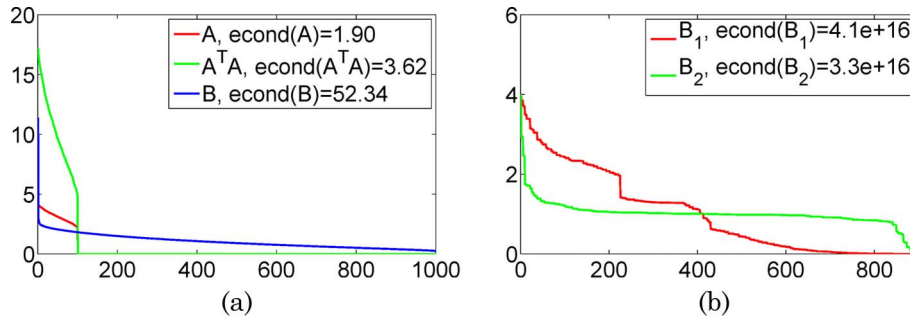


Fig. 1. (Color online) (a) Singular value spectra of a 100×1000 Gaussian random matrix A , its Gram matrix $A^H A$, and its associated B matrix and (b) singular value spectra of two examples of \mathcal{B} , denoted by B_1 and B_2 , associated with two examples of H created with two different axial sample spacings.

between two transverse planes for \mathbf{B}_1 is smaller than that for \mathbf{B}_2 : ($\Delta_z(\mathbf{B}_1) < \Delta_z(\mathbf{B}_2)$). Supposedly, \mathbf{B}_2 is much better conditioned than \mathbf{B}_1 . Despite the large effective condition numbers, 10^{16} , for both matrices, \mathbf{B}_1 has more small singular values than \mathbf{B}_2 does. The number of the singular values greater than 0.5 of \mathbf{B}_2 is around 850, whereas that of \mathbf{B}_1 is only around 500. Hence, \mathbf{B}_1 is more ill conditioned than \mathbf{B}_2 . This tendency pertaining to the axial sample spacing and the condition number of the matrices remains the same even when the matrix sizes (significantly) increase.

C. Preconditioning of System Equation

The sparsity-constrained estimates are known to be near optimal, provided that the columns of the system matrix form approximately an orthonormal basis [16,19]. To transform our system matrix to such an orthonormal basis, we apply a preconditioning method and convert $\mathbf{d} = \mathcal{B}\alpha + \sigma^2\omega$ to $\tilde{\mathbf{d}} = \mathcal{P}\mathbf{d} = \mathcal{P}\mathcal{B}\alpha + \sigma^2\mathcal{P}\omega = \tilde{\mathcal{B}}\alpha + \sigma^2\tilde{\omega}$. Therefore, we alternatively solve

$$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \|\tilde{\mathbf{d}} - \tilde{\mathcal{B}}\alpha\|_2^2 + \beta\Phi(\alpha), \quad (19)$$

instead of Eq. (15).

If \mathcal{B} has full rank, the preconditioner \mathcal{P} may be chosen such that $\tilde{\mathcal{B}} = \mathcal{P}\mathcal{B}$ is as unitary as possible. An immediate choice of unitary $\tilde{\mathcal{B}}$ would be the identity matrix: $\tilde{\mathcal{B}} = I$. In this case, the preconditioner \mathcal{P} is merely the inverse of \mathcal{B} : $\mathcal{P} = \mathcal{B}^{-1}$, and the estimation problem becomes a denoising problem: $\tilde{\mathbf{d}} = \mathcal{P}\mathcal{B}\alpha + \sigma^2\tilde{\omega} = \alpha + \sigma^2\tilde{\omega}$. However, the matrix could often be rank deficient or ill conditioned. In such cases, some nonzero singular values of \mathcal{B} are so small that they either excessively amplify the noise or cause numerical instability when $\mathcal{P} = \mathcal{B}^{-1}$ is applied. To mitigate this problem, we choose \mathcal{P} in the Tikhonov regularized sense. That is $\mathcal{P} = (\mathcal{B}^H\mathcal{B} + \lambda_t I)^{-1}\mathcal{B}^H$. λ_t is a Tikhonov regularization parameter. Note that λ_t suppresses or ignores a small number of singular vectors associated with numerically unstable or zero singular values.

Recall that the size of \mathcal{B} is typically several hundreds or thousands times several hundreds of thousands. In general, computing $(\mathcal{B}^H\mathcal{B} + \lambda_t I)^{-1}$, for the size of \mathcal{B} , is a computationally formidable task. Fortunately, the CBCB property of \mathcal{B} , discussed in theorem 3.2, enables efficient and fast computation of the inverse. Several efficient computation approaches have been proposed for similar inversion problems [20,21]. We adapt the method described in [21] to find the Tikhonov pseudoinverse. We sketch major steps here, and refer readers to Ref. [21] for details. Note that $\mathbf{K} = (\mathcal{B}^H\mathcal{B} + \lambda_t I)$ can be readily proved to be CBCB using theorem 3.2. Therefore, a diagonal-block matrix whose blocks are all 2D DFT matrices can diagonalize all blocks of \mathbf{K} . The resulting diagonal-block matrix can then be permuted to a block-diagonal matrix. The size of each block of this permuted matrix is relatively small. In fact, the size of such a block for \mathbf{K} is only $N_z \times N_z$. Therefore, each block can be separately inverted efficiently and fast. The resulting inverse is then permuted back to the original shape and is multiplied by the 2D DFT block diagonal matrix to result in \mathbf{K}^{-1} . \mathcal{B} is then computed using $\tilde{\mathcal{B}} = (\mathcal{B}^H\mathcal{B} + \lambda_t I)^{-1}\mathcal{B}^H\mathcal{B}$. $\tilde{\mathcal{B}}$ resembles the identity matrix: it has diagonal elements close to 1, with suitable normalization and has comparatively small off-diagonal elements. The detailed steps are described in lemmas 2, 3, and 4 in [21]. Note that $\tilde{\mathbf{d}} = \tilde{\mathcal{B}}\alpha + \sigma^2\tilde{\omega}$ implicitly ignores a small number of singular vectors associated with “small” singular values because of the effect of λ_t . The constrained optimization technique enforcing the sparsity constraints is adapted to resolve such incompleteness of $\tilde{\mathbf{d}}$.

While λ_t may be systematically chosen, we chose it by trial and error. Our choice for λ_t is 10^{-8} in the following reconstruction results, unless otherwise stated.

4. Experimental Results and Discussions

A. Descriptions on Experiments

Figures 2(a) and 2(b) show photographs of the experimental setup and a 3D diffuse object “DISP,” respectively. The scattered light in the object arm interferes

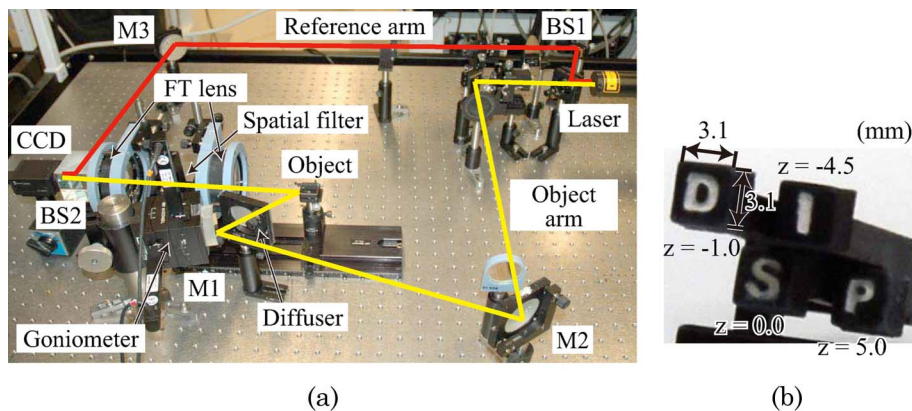


Fig. 2. (Color online) (a) Photograph of experimental setup: M, mirror; BS, beam splitter; and FT lens, Fourier transform lens. (b) Photograph of a 3D diffuse object “DISP.”

with the light in the reference arm to form the interference intensity distribution on the CCD. The CCD has 1624×1224 resolution with $4.4\ \mu\text{m}$ pixel pitch with 16 bit quantization. The Mach–Zehnder geometry is adapted to conduct off-axis holography. The beams are created with an He–Ne laser with wavelength $632.8\ \text{nm}$. The spatial filter is applied in the Fourier domain via $4f$ optics. One of the two Fourier transform lenses is used to Fourier transform the scattered object beam, and the other is used to inverse Fourier transform the filtered scattered object beam. The Fourier transform lenses have the same focal length of $75\ \text{mm}$. The spatial filter is inserted between the two Fourier transform lenses to ensure complete separation of the object Fourier spectra, the autocorrelation Fourier spectra, and the twin-image Fourier spectra. The separation is applied only along the vertical axis to maximize the utilizable Fourier spectra. In other words, the filter height is chosen such that the size of the object Fourier spectra is nearly a quarter of that of the whole object Fourier spectra only along the vertical direction.

Note that the beam in the object arm illuminates the object through a stationary diffuser. The purposes of the diffuser are to fully develop the speckles in each speckle field and also to minimize the correlations between the speckle fields. We induce a random phase by introducing the effect of a stationary diffuser into each speckle field. To change the random phase patterns in each speckle field, a goniometer is introduced to illuminate the diffuser from different angles for each speckle field measurement. The diffuser pro-

duces multiple stationary random phases through the angular diversity of the illumination. Note that, for each speckle field, the random phase needs to be stationary. Therefore, while each hologram is being taken, both the goniometer and the diffuser are fixed. Readers may be referred to [8] for more details on how a diffuser can be used to create uncorrelated speckle measurements.

B. Reconstructions with Experimental Data

The size of the interference intensity image on the CCD is 1624×1224 pixels. This image is Fourier transformed and filtered to separate the object Fourier spectra from the Fourier spectra of the autocorrelation and the twin image. Consequently, the size of the separated speckle field, associated with only the object, is 1582×222 pixels.

A conventional approach to speckle reduction is to average the intensities of multiple backpropagated fields. *Backpropagation* is defined as finding f from a single g by computing Eq. (9). Conceptually, this computation is equivalent to propagating a 2D field in the direction backward to the propagating direction (i.e., backward propagation or backpropagation). The reconstructions from experimental data in this section are compared to the results obtained by applying the conventional approach.

Figure 3(a) shows a backpropagation reconstruction with a single speckle field. Figure 3(b) shows the average of the backpropagation reconstruction intensities of 50 speckle fields. As expected, the reconstruction is smoother and has a better contrast

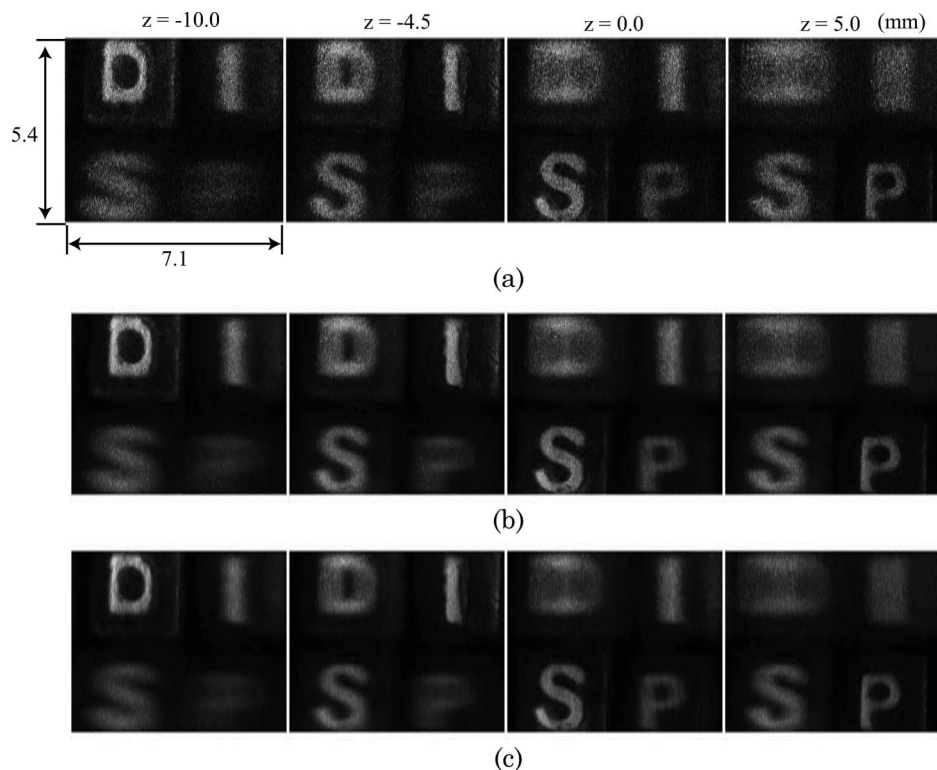


Fig. 3. (a) Intensity of the backpropagation with a single speckle field, (b) the average of the backpropagation intensities of 50 speckle fields [see Eq. (10)], and (c) an image obtained by applying a 5×5 median filter to the image in (b).

when multiple intensities are averaged compared to the backpropagation reconstruction intensity created from a single speckle field. Nonetheless, speckle artifacts still manifest as rough surface features. Moreover, the intensity average reconstruction shows poor axial resolution, as is clear from blurred features in the axial direction. For example, although the letter “S” becomes in focus and presents clear features on the third plane, the estimated intensity of the “S” is strongly blurred into the second and fourth planes. For reference, Fig. 3(c) shows a speckle reduction result obtained by applying a 5×5 median filter to the estimate in Fig. 3(b).

Figure 4(a) shows an estimate obtained by solving Eq. (15) when \vec{d} is constructed with 50 speckle fields. Note that speckle artifacts have been significantly reduced, as indicated by the smooth surfaces of the letters “D,” “I,” “S,” and “P.” In addition, the estimate shows improved axial resolution, by suppressing the letters not in focus. However, the blurred letters are still undesirable. Figure 4(b) shows a Tikhonov pseudoinverse estimate from 50 speckle fields: $\vec{d} = \tilde{\mathcal{B}}\alpha + \sigma^2\tilde{\omega}$. In this reconstruction, although the axial resolution is improved by suppressing the blurred letters in the out-of-focus planes, the speckle artifacts are still manifest as the rough surfaces of the letters in focus. Figure 4(d) shows an estimate obtained by solving Eq. (19) from the same \vec{d} . The estimate shows remarkably suppressed speckle artifacts and improved axial resolution. For com-

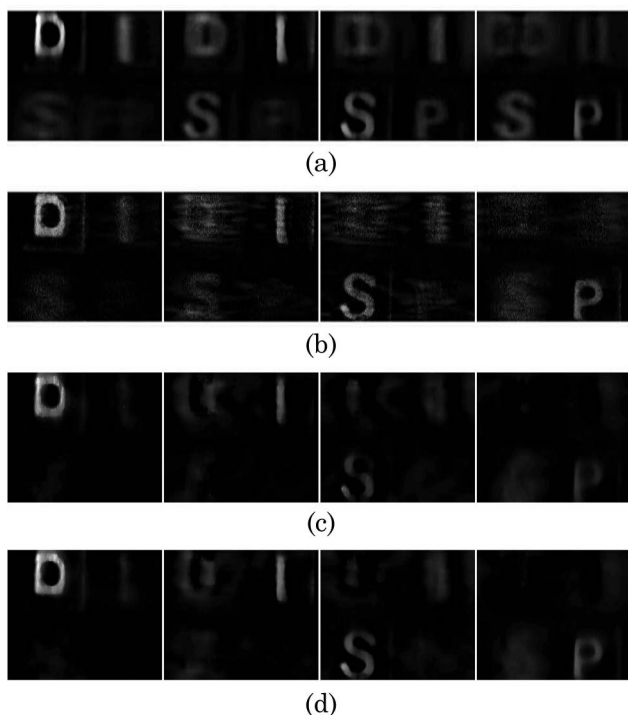


Fig. 4. (a) Estimate obtained by solving Eq. (15), (b) a Tikhonov-regularized pseudoinverse (i.e., $\vec{d} = \tilde{\mathcal{B}}\alpha + \sigma^2\tilde{\omega}$) with $\lambda_t = 10^{-8}$ using 50 speckle fields both for (a), (b) and (c), (d) reconstructions (with $\beta = 0.3$) obtained by solving the preconditioned formulation in Eq. (19) with (c) a single speckle field and (d) 50 speckle fields.

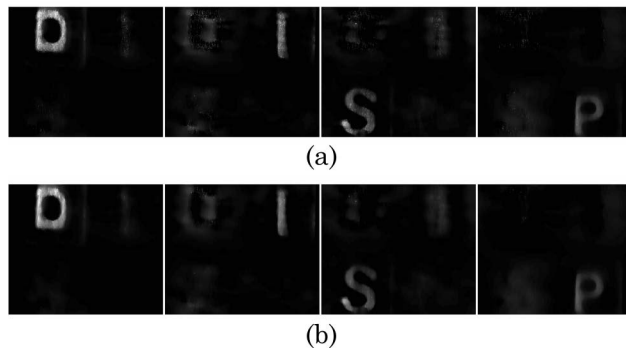


Fig. 5. Reconstructions with (a) $\beta = 0.1$ and (b) $\beta = 0.2$.

parison, Fig. 4(c) demonstrates an estimate obtained by solving Eq. (19) with a single speckle field. The axial resolution of the estimate appears comparable to the estimate created with 50 speckle fields. In contrast, the letters “S” and “P” in the estimate created with 50 speckle fields are clearer compared to the single estimate created with a single speckle field.

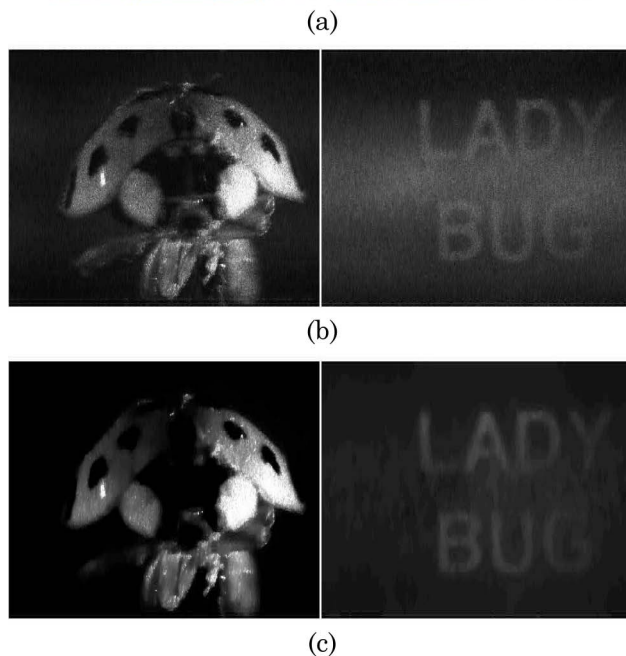
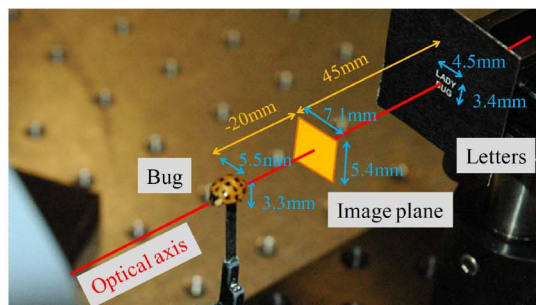


Fig. 6. (Color online) (a) Photograph of the objects: a real ladybug in front and the letters “LADYBUG” in the back with their dimensions; the yellow rectangle shows the location of the in-focus plane, (b) the average image of the backpropagation reconstruction intensities of 20 speckle fields, and (c) our reconstruction obtained by solving Eq. (19).

The regularization parameter β was chosen by trial and error, such that the estimate shows least speckle artifacts and “visually best” axial resolution. Our choice is $\beta = 0.3$ for the reconstructions in Fig. 4. Reconstructions obtained with other values of β are also shown in Fig. 5, to illustrate the effect of β on reconstructions.

Figure 6 shows another experiment. In the experiment, we placed the letters “LADYBUG” directly behind a real ladybug in the axial direction, as illustrated by Fig. 6(a). Figure 6(b) shows the average intensity of the backpropagation reconstructions created with 20 object speckle fields. The wings of the real ladybug show speckle artifacts, and the letters “LADYBUG” are disguised in the blurs of the real ladybug. Figure 6(c) shows an estimate obtained by solving Eq. (19). The surfaces of the wings become smoother. In particular, image contrast has been improved, as illustrated by the black spots on the wings. Also, the blurs placed on the letters have been reduced, while the letters become smoother. Readers may notice, however, that some of the features appear lost around the center of the real ladybug and in some part of the letters as well. We found that some scattered light was obscured by the object. The lack of this obscuring effect in the mathematical model in the algorithm causes the artifact in the estimate. A new model that considers the obscuring effect would be an interesting future work.

5. Conclusions

We have proposed an approach for estimating the incoherent scattering density of a diffuse object. Experimental results have successfully demonstrated the strength of our methods for producing the 3D reconstructions of diffuse objects, which substantiated suppressed speckle artifacts and the improved depth resolution. The suppression of speckle artifacts and the improvement in the axial resolution result from inverting the synthetic operator \mathbf{B} , which acts like a (generalized) 3D point spread function of the associated 3D incoherent image, and encouraging the sparsity of the estimates. Interesting future work might include the application of our method to medical imaging applications, such as optical coherence tomography or ultrasound imaging, where speckles are also fundamental but undesirable.

Appendix A: Proof of Lemma 3.1

Proof. Because A is a CBC matrix, it can be expressed as

$$A = \text{circulant}(A_1, A_2, \dots, A_k)$$

$$= \begin{bmatrix} A_1 & A_2 & A_3 & \dots & A_k \\ A_k & A_1 & A_2 & \dots & A_{k-1} \\ A_{k-1} & A_k & A_1 & \dots & A_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_2 & A_3 & A_4 & \dots & A_1 \end{bmatrix}, \quad (\text{A1})$$

where each block A_k is circulant and can be written as

$$A_k = \text{circulant}(a_1^k, a_2^k, \dots, a_M^k), \quad (\text{A2})$$

with the superscript k denoting the block index. A circulant matrix can be expressed in matrix polynomial: $A_k = \sum_n a_n^k \Pi^n$, where $\Pi = \text{circulant}(0, 1, 0, 0, \dots, 0)$ [13]. Similarly, the CBC matrix can be written as $A = \sum_l \Pi^l \otimes A_l$ with \otimes denoting the Kronecker product [13]. Hence, $A = \sum_l \Pi^l \otimes (\sum_n a_n^l \Pi^n)$. Because \mathbf{B} is obtained by squaring A elementwise, it can be written as $\mathbf{B} = \sum_l \Pi^l \otimes (\sum_n |a_n^l|^2 \Pi^n)$. Therefore, $\mathbf{B} = \text{circulant}(\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_k)$ with $\mathbf{B}_k = \text{circulant}(|a_1^k|^2, |a_2^k|^2, \dots, |a_M^k|^2)$. This concludes that if A is a CBC matrix, then \mathbf{B} is also a CBC matrix.

Appendix B: Proof of Theorem 3.2

Proof. It is well known that a matrix C is a CBC matrix if and only if it can be written as $C = F_{2D}^H \Lambda F_{2D}$ for some diagonal matrix Λ [13,20]. First, we note that $H_{l_1}^H H_{l_2} = F_{2D}^H P_{l_1}^H P_{l_2} F_{2D} = F_{2D}^H \Lambda_{l_1, l_2} F_{2D}$. Because P_k is a diagonal matrix for all k , Λ_{l_1, l_2} is also a diagonal matrix. This proves that $H_{l_1}^H H_{l_2}$ is a CBC matrix. We also have

$$H^H H$$

$$= \begin{bmatrix} H_1^H H_1 & H_1^H H_2 & H_1^H H_3 & \dots & H_1^H H_{N_z} \\ H_1^H H_{N_z} & H_1^H H_1 & H_1^H H_2 & \dots & H_1^H H_{N_z-1} \\ H_1^H H_{N_z-1} & H_1^H H_{N_z} & H_1^H H_1 & \dots & H_1^H H_{N_z-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_1^H H_2 & H_1^H H_3 & H_1^H H_4 & \dots & H_1^H H_1 \end{bmatrix}. \quad (\text{B1})$$

Recall that \mathcal{B} is constructed by squaring $H^H H$ elementwise: $[\mathcal{B}]_{mn} = |[H^H H]_{mn}|^2$. We can show that each block \mathcal{B}_{l_1, l_2} , for which $[\mathcal{B}_{l_1, l_2}]_{mn} = |[H_{l_1}^H H_{l_2}]_{mn}|^2$, is a CBC matrix by letting $A = H_{l_1}^H H_{l_2}$ and $\mathbf{B} = \mathcal{B}_{l_1, l_2}$ in lemma 3.1. Hence, \mathcal{B} is a block matrix composed of multiple CBC matrices. This proves that \mathcal{B} is a CBCB matrix.

Appendix C: Proof of Corollary 3.3

Proof. Note that $P_{l_1}^H P_{l_2} = \text{diag}(Y_{l_1, l_2})$. Because $H_{l_1}^H H_{l_2} = F_{2D}^H P_{l_1} P_{l_2} F_{2D}$,

$$H_{l_1}^H H_{l_2} = \text{circulant}(\text{vec}(\mathbf{W}_{l_1, l_2})) = \text{circulant}(\text{vec}(\mathcal{F}_{2D}^{-1}(Y_{l_1, l_2}))).$$

Thus, $\mathcal{B}_{l_1, l_2} = \text{circulant}(\text{vec}(\mathbf{Z}_{l_1, l_2}))$, where $[\mathbf{Z}_{l_1, l_2}]_{m_1 m_2} = |[W_{l_1, l_2}]_{m_1 m_2}|^2$. Therefore, the eigenvalues $v_1^{l_1, l_2}, v_2^{l_1, l_2}, \dots, v_M^{l_1, l_2}$ of \mathcal{B}_{l_1, l_2} can be obtained by taking a 2D DFT of \mathbf{Z}_{l_1, l_2} .

This research was supported by the Defense Advanced Research Projects Agency (DARPA) under the Air Force Office of Scientific Research contract FA9550-06-1-0230.

References

1. R. H. T. Bates, "Astronomical speckle imaging," *Phys. Rep.* **90**, 203–297 (1982).
2. J. R. Fienup, R. G. Paxman, M. F. Reiley, and B. J. Thelen, "3-D imaging correlography and coherent image reconstruction," *Proc. SPIE* **3815**, 60–69 (1999).
3. T. Schulz, "Penalized maximum-likelihood estimation of covariance matrices with linear structure," *IEEE Trans. Signal Process.* **45**, 3027–3038 (1997).
4. A. D. Lanterman, "Statistical radar imaging of diffuse and specular targets using an expectation-maximization algorithm," *Proc. SPIE* **4053**, 20–31 (2000).
5. R. G. Dantas and E. T. Costa, "Ultrasound speckle reduction using modified gabor filters," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **54**, 530–538 (2007).
6. D. L. Marks, T. S. Ralston, and S. A. Boppert, "Speckle reduction by I-divergence regularization in optical coherence tomography," *J. Opt. Soc. Am. A* **22**, 2366–2371 (2005).
7. D. J. Brady, K. Choi, D. L. Marks, R. Horisaki, and S. Lim, "Compressive holography," *Opt. Express* **17**, 13040–13049 (2009).
8. J. W. Goodman, *Speckle Phenomena in Optics* (Roberts, 2006).
9. J. W. Goodman, *Introduction to Fourier Optics* (Roberts, 2005).
10. E. N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory," *J. Opt. Soc. Am.* **52**, 1123–1130 (1962).
11. Y. K. Park, W. Choi, Z. Yaqoob, R. Dasari, K. Badizadegan, and M. Feld, "Speckle-field digital holographic microscopy," *Opt. Express* **17**, 12285–12292 (2009).
12. Z. Yaqoob, D. Psaltis, M. Feld, and C. Yang, "Optical phase conjugation for turbidity suppression in biological samples," *Nat. Photon.* **2**, 110–115 (2008).
13. T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing* (Prentice-Hall, 2000).
14. J. W. Goodman, *Statistical Optics* (Wiley Interscience, 2000).
15. P. Stoica and R. L. Moses, *Spectral Analysis of Signals* (Prentice-Hall, 2005).
16. E. Candés, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.* **59**, 1207–1223 (2006).
17. R. G. Baraniuk, E. Candés, R. Nowak, and M. Vetterli, "Compressive sampling," *IEEE Signal Process. Mag.* **25**, 12–13 (2008).
18. R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Trans. Inf. Theory* **49**, 3320–3325 (2003).
19. J. Tropp, "Just relax: convex programming methods for identifying sparse signals in noise," *IEEE Trans. Inf. Theory* **52**, 1030–1051 (2006).
20. T. F. Chan and J. A. Olkin, "Circulant preconditioners for toeplitz-block matrices," *Numer. Algorithms* **6**, 89–101 (1994).
21. B. Fischer and J. Modersitzki, "Fast inversion of matrices arising in image processing," *Numer. Algorithms* **22**, 1–11 (1999).