Reconstruction guarantees for compressive tomographic holography

Yair Rivenson,[1],* Adrian Stern,[2] and Joseph Rosen[1]

[1]*Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel*

[2]*Department of Electro-Optics Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel*

Corresponding author: rivenson@ee.bgu.ac.il

Three-dimensional (3D) object tomography from a two-dimensional recorded hologram is a process of high-dimensional data inference from undersampled data. As such, recently, techniques developed in the field of compressive sensing and sparse representation have been applied for this task. While many applications of compressive sensing for tomography from digital holograms have been demonstrated in the past few years, the fundamental limits involved have not yet been addressed. We formulate the guarantees for compressive sensing-based recovery of 3D objects and show their relation to the physical attributes of the recording setup.   © 2013 Optical Society of America

OCIS codes: (090.1995) Digital holography; (070.0070) Fourier optics and signal processing; (100.3190) Inverse problems; (100.6890) Three-dimensional image processing; (100.6950) Tomographic image processing; (100.1830) Deconvolution.

http://dx.doi.org/10.1364/OL.38.002509

Numerical reconstruction obtained by digitally focusing on different 3D object depth planes is one of the biggest advantages of digital holography, enabling a single-shot object acquisition rather than physically focusing on each plane separately. The hologram records a 2D wave field of the 3D object, making the reconstruction of the 3D object data from its 2D projection inherently ill posed. Consequently, often the reconstructed in-focus plane image is distorted. In recent years, compressive sensing was successfully combined with digital holography (see [1] and references therein). One of the applications is reconstruction of a 3D object tomograph from its single 2D hologram, initially demonstrated in 2009 [2,3]. This problem was approached by formulating the object reconstruction as a compressive sensing [4,5] inverse problem. The word "compressive" in this sense refers to the fact that the holographic sensing process encodes and compresses 3D object information into 2D holographic measurements.

Despite vast applicative efforts in recent years, in publications, such as [1–3,6–11], to name a few, there has been little theoretical work investigating the fundamental limits of 3D object reconstruction from its 2D hologram. The important question is, how accurately may we infer the 3D object points from our set of 2D measurements, regardless of the specific recovery method being used? In this Letter, reconstruction accuracy is formulated by using resolution constraints. For the sake of completeness, we first give a short background on compressive sensing before deriving the main result. Generally, the object sensing mechanism is expressed using a matrix-vector multiplication:

$$\mathbf{g} = \Phi\mathbf{f}, \qquad (1)$$

where $\mathbf{f} \in \mathbb{C}^{N\times1}$ is the sparse object, $\Phi \in \mathbb{C}^{M\times N}$ is the sensing operator, and $\mathbf{g} \in \mathbb{C}^{M\times1}$ represents the measurements, where the number of measurements, M, is smaller than the number of object pixels, N. In order to guarantee accurate reconstruction, compressive sensing theory requires the object to have a sparse representation in some transform domain, i.e., the vector $\mathbf{f} \in \mathbb{C}^{N\times1}$ needs to have only S meaningful entries (S-sparse signal). The second essential requirement applies to the sensing mechanism, and it can be quantified by using the coherence parameter, given by

$$\mu = \max_{k\neq l}|\langle\boldsymbol{\phi}_k,\boldsymbol{\phi}_l\rangle|/\{\|\boldsymbol{\phi}_k\|_2\|\boldsymbol{\phi}_l\|_2\}, \qquad (2)$$

where $\boldsymbol{\phi}_l$ denotes the lth column vector of Φ and $\langle\cdot\rangle$ is the vectors' inner product. The coherence parameter is bounded by $\sqrt{(N-M)/[M(N-1)]} \le \mu \le 1$. Compressive sensing theory guarantees accurate object recovery by evoking a convex optimization procedure, with the number of accurately reconstructed nonzero terms obeying [5]

$$S \le 0.5(1 + 1/\mu). \qquad (3)$$

In order to evaluate the performance of the compressive holography framework, we wish to quantify its coherence parameter, which is derived from the system's forward model. We assume this model obeys the Born approximation [2]. A possible model of the acquisition geometry is illustrated in Fig. 1. A monochromatic plane wave, with wavelength λ, illuminates a 3D object volume. The object volume is discretized into $N_{\text{object}} = N_x \times N_y \times N_z$ voxels, with voxel dimensions of $\Delta x \times \Delta y \times \Delta z$, where the object length is $L_z = N_z \times \Delta z$. The 3D object wave field can be recorded by standard holography methods on a CCD [12].

In the numerical near field Fresnel approximation [1], $\Delta x_{\text{CCD}} = \Delta x$, $\Delta y_{\text{CCD}} = \Delta y$, and the number of pixels is $N_{\text{holo}} = N_x \times N_y$. Let us look at the following discrete forward model, relating the 3D object, f, to its recorded 2D field, g:

$$g(u\Delta x, v\Delta y) = \sum_{r=1}^{N_z}\mathcal{F}_{2D}^{-1}\left\{e^{-j\pi\lambda r\Delta z[(\Delta v_x m)^2+(\Delta v_y n)^2]}e^{j\frac{2\pi}{\lambda}r\Delta z}\right.$$
$$\left.\times \mathcal{F}_{2D}[f(p\Delta x, q\Delta y; r\Delta z)]\right\}, \qquad (4)$$
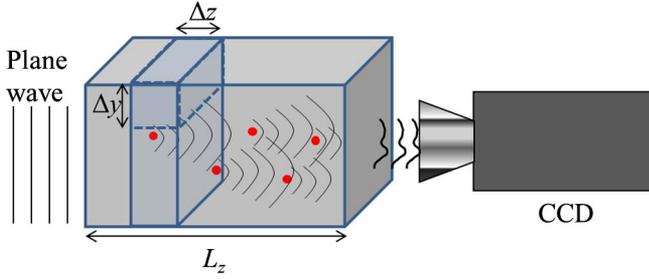
Fig. 1.   Schematic description of the framework. A plane wave illuminating a volume (shaded gray). The wavefront scattered from the different particles is holographically recorded on a CCD.
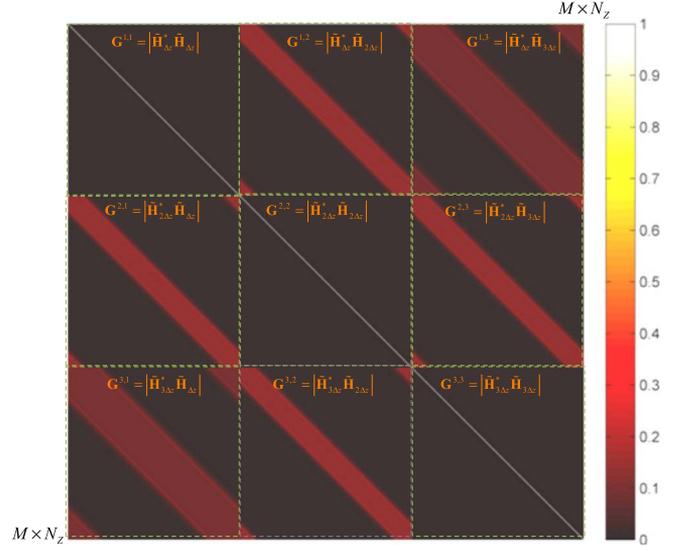

Fig. 2.   Gram matrix for the 3D–2D forward Fresnel sensing operator. The partition to nine submatrices, marked by dashed lines, shows the coherence between the point spread function of two different planes.

where $\mathcal{F}_{2D}$ is the discrete 2D Fourier transform. For notation convenience, we further assume that $\Delta x = \Delta y = \Delta$ and $N_x = N_y = \sqrt{M}$; therefore the discrete indices are $0 \le m, n, p, q, u, v \le \sqrt{M} - 1$, and $\Delta v_x = \Delta v_y = 1/(\sqrt{M}\Delta)$.

Equation (4) can be written as a matrix-vector multiplication by concatenation of N_z matrices, $\mathbf{H}_{r\Delta z}$, $r \in [1, N_z]$, each one of size $\sqrt{M} \times \sqrt{M}$:

$$\mathbf{g} = [\mathbf{H}_{\Delta z}; \ldots; \mathbf{H}_{N_z\Delta z}][\mathbf{f}_{\Delta z}; \ldots; \mathbf{f}_{N_z\Delta z}] = \Phi \mathbf{f},$$
$$\mathbf{H}_{r\Delta z} = \mathcal{F}_{2D}^{-1}\mathbf{Q}_{r\Delta z}\mathcal{F}_{2D}. \tag{5}$$

where the matrix $\mathbf{Q}_{r\Delta z}$ is a diagonal matrix that accounts for the quadratic phase terms of Eq. (4), and $[\mathbf{f}_{\Delta z}; \ldots; \mathbf{f}_{N_z\Delta z}]$ is a lexicographical representation of the 3D object.

A standard way to obtain the coherence parameter from Eq. (2) is by calculating the Gram matrix $\mathbf{G} = \tilde{\Phi}^*\tilde{\Phi}$, $\mathbf{G} \in \mathbb{C}^{(N_z \times M) \times (N_z \times M)}$ (where $\tilde{\Phi}$ is the column normalization of Φ, and $\tilde{\Phi}^*$ is the Hermitian conjugate of $\tilde{\Phi}$); then μ is the maximal off-diagonal absolute value [5]. Because of the construction of $\tilde{\Phi}$ by concatenation [Eq. (5)], the Gram matrix has a structured form as demonstrated in Fig. 2. In Fig. 2, a Gram matrix is shown for the case where the volume is divided into $N_z = 3$ equally spaced planes. It can be seen that the Gram matrix is built of $N_z \times N_z = 9$ subblocks, each one representing the correlation between the point spread functions of two depth planes. Since the normalized $\tilde{H}_{r\Delta z}$ are orthonormal, the corresponding diagonal subblocks $\mathbf{G}^{k,k} = \tilde{\mathbf{H}}_{k\Delta z}^*\tilde{\mathbf{H}}_{k\Delta z} = \mathbf{I}$, i.e., all are zero except the diagonal. This means that, in order to find the coherence parameter, we should look for it in any off-diagonal block, $\mathbf{G}^{k,l} = \tilde{\mathbf{H}}_{k\Delta z}^*\tilde{\mathbf{H}}_{l\Delta z}$ which is equal to (for $k \ne l$)

$$\tilde{\mathbf{H}}_{k\Delta z}^*\tilde{\mathbf{H}}_{l\Delta z} = \mathcal{F}_{2D}^{-1}\mathbf{Q}_{k\Delta z}^*\mathcal{F}_{2D}\mathcal{F}_{2D}^{-1}\mathbf{Q}_{l\Delta z}\mathcal{F}_{2D} = \mathcal{F}_{2D}^{-1}\mathbf{Q}_{(l-k)\Delta z}\mathcal{F}_{2D}. \tag{6}$$

Therefore the coherence parameter is given by [12,13]

$$\mu = \max_{k \ne l}|\tilde{\mathbf{H}}_{k\Delta z}^*\tilde{\mathbf{H}}_{l\Delta z}| = \max_{k \ne l}|\mathcal{F}_{2D}^{-1}\mathbf{Q}_{(k-l)\Delta z}\mathcal{F}_{2D}|$$
$$\approx \max_{k \ne l}\left\{\frac{\Delta^2}{\lambda(l-k)\Delta z}\right\} = \frac{\Delta^2}{\lambda\Delta z}. \tag{7}$$

Substituting Eq. (7) into Eq. (3), we find that the number of object features that can be accurately reconstructed is bounded by

$$S \le 0.5(1 + \lambda\Delta z/\Delta^2). \tag{8}$$

While the coherence parameter used in Eq. (8) is based on a worst-case analysis and is considered rather pessimistic [14], it gives an evaluation and a trend for actual performance results, regardless of the reconstruction method or algorithm being used. Equation (8) is useful to reveal the dependence of successful object reconstruction on the system parameters. The result in Eq. (8) predicts that, by increasing the wavelength or using coarser axial resolution, the number of object features that can be reconstructed accurately is increased. Another way to interpret these results is noticing that $\mu \approx N_z\Delta^2/(\lambda L_z)$. This expression concurs with the mathematical intuition that, as the ratio of the number of unknowns to the number of equations, which is N_z, increases, the larger μ becomes, and fewer features, S, can be accurately reconstructed. From a physical perspective, the coherence parameter measures the maximum correlation between any two point spread functions. This corresponds to the correlation between two object points having the same lateral position, but located in two adjacent planes. We also note that the coherence parameter is lower bounded by $\sqrt{(MN_z - M)/[M(MN_z - 1)]} \approx 1/\sqrt{M}$, i.e., by the square root of the number of detector pixels.

By loosening the requirement on accurate reconstruction and allowing some degree of error, we obtain results that are more practical, while obeying the trend predicted from Eq. (8). This is demonstrated in the following numerical experiment: a volume of length $L_z = 316$ μm, randomly populated with S identical point particles (particle size $\Delta = 5$ μm) is illuminated by a plane wave with a wavelength of $\lambda = 633$ nm. The detector plane is positioned 1 mm from the volume and is
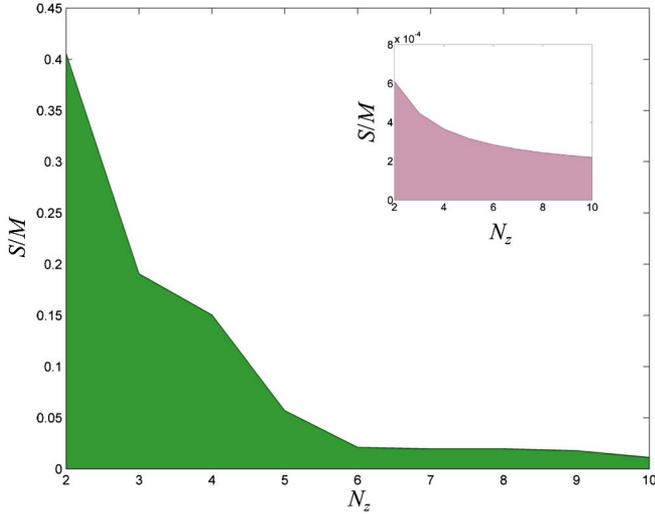
Fig. 3.   Simulation results showing the normalized number of reconstructed 3D object's particles (MSE $< 10^{-8}$) as a function of number of objects planes, for a constant volume length, L_z, given a sensor with M pixels. The theoretical exact reconstruction guarantee according to Eq. (8) is placed in the inset.

composed of 64×64 pixels with pitch of $\Delta = 5$ μm. The object's volume 2D wave field is acquired, using phase-shifting holography [12]. The 3D object is then reconstructed by using the two-step iterative shrinkage/thresholding (TwIST) solver [15], solving

$$\hat{\mathbf{f}} = \arg\min_{\mathbf{f}} \|\mathbf{g} - \Phi\mathbf{f}\|_2 + \tau\|\mathbf{f}\|_1. \qquad (9)$$

The first term in Eq. (9) represents the data fidelity, $\|\cdot\|_p$ is the ℓ_p-norm, and τ is a regularization parameter. We have limited the run of the TwIST solver to 1,000 iterations. For each simulation instance the volume was divided to $N_z \in [2:10]$ planes. For each volume division we increase the number of particles in the volume and solve Eq. (9). The number of particles continues to increase till the reconstruction accuracy, quantified by the mean-square error (MSE), becomes larger than 10^{-8}. The simulation was repeated 10 times, and the result, shown in Fig. 3, is the mean of the different simulation instances. The shaded area denotes region of accurate (MES $< 10^{-8}$) 3D object tomography reconstruction from its 2D hologram. From Fig. 3, we notice that, as the axial resolution becomes finer, fewer particles can be reconstructed, which is expected from the analysis. The quantity S/M in Fig. 3 represents the ratio between the number of recovered particles to the number of detector pixels. In the limit of $N_z = 1$ the ratio $S/M = 1$, meaning that the maximum recoverable particles equals the number of CCD pixels, as expected when a 2D object reconstruction is carried out from its 2D wave field. For 3D objects, the number of recoverable particles is inversely proportional to the number of depth slices, N_z.

The results from the numerical experiment can be also interpreted as a superresolving capability of the framework, meaning that if we have a small number of features, S, we can reconstruct an object with a longitudinal resolution Δz, which is improved compared with the classical resolution limit $\Delta z \approx 4\Delta^2/\lambda = 158$ μm [2]. When working in the near-field Fresnel approximation, i.e., when the entire diffraction pattern is essentially captured by the detector, the longitudinal resolution Δz depends on the object's feature size, Δ [2]. This longitudinal resolution is the minimal detail that can be recovered by any reconstruction algorithm that does not take into account *a priori* knowledge, such as the object's sparsity. Hence, the simulation demonstrates that this limitation can be relaxed in practice when applying the compressive sensing framework. In fact, even the pessimistic estimation of the coherence parameter [14] predicts that when we need to resolve only two $S = 2$ identical particles, lying in two adjacent planes, we get from Eq. (8) $\Delta z = 3\Delta^2/\lambda$, which is 33% superior to the classical limit. Our simulation has shown that a 3D object volume with about 50 identical particles can be reconstructed with a longitudinal accuracy of $L_z/10 = 31.6$ μm, which is 5 times finer the than the fundamental limitations.

To conclude, we have formulated exact guarantees for 3D object tomography recovery from its 2D captured diffracted field when using the compressive sensing framework. The result shows that a digital hologram can indeed "compress" a 3D object with S degrees of freedom, where the number of degrees of freedom depends on the sensor's resolution, the axial resolution of the object. and the illuminating wavelength.

References

1. Y. Rivenson, A. Stern, and B. Javidi, Appl. Opt. **52**, A423 (2013).
2. D. J. Brady, K. Choi, D. L. Marks, R. Horisaki, and S. Lim, Opt. Express **17**, 13040 (2009).
3. L. Denis, D. Lorenz, E. Thiébaut, C. Fournier, and D. Trede, Opt. Lett. **34**, 3475 (2009).
4. E. Candes and M. Wakin, IEEE Signal Process. Mag. **25**, 21 (2008).
5. A. M. Bruckstein, D. L. Donoho, and M. Elad, SIAM Rev. **51**, 34 (2009).
6. X. Zhang and E. Y. Lam, J. Opt. Soc. Am. A **27**, 1630 (2010).
7. A. F. Coskun, I. Sencan, T.-W. Su, and A. Ozcan, Opt. Express **18**, 10510 (2010).
8. Y. Rivenson, A. Stern, and J. Rosen, Opt. Express **19**, 6109 (2011).
9. S. Lim, D. Marks, and D. Brady, Appl. Opt. **50**, H75 (2011).
10. L. Tian, Y. Liu, and G. Barbastathis, in *Biomedical Optics and 3-D Imaging*, OSA Technical Digest (Optical Society of America, 2012), paper DW4C.3.
11. L. Williams, G. Nehmetallah, and P. Banerjee, Appl. Opt. **52**, 1702 (2013).
12. T. Kreis, *Handbook of Holographic Interferometry*, 1st ed. (Wiley-VCH, 2004), Chap. 3.
13. Y. Rivenson and A. Stern, Opt. Lett. **36**, 3365 (2011).
14. M. Elad, IEEE Trans. Signal Process. **55**, 5695 (2007).
15. J. M. Bioucas-Dias and M. A. T. Figueiredo, IEEE Trans. Image Process. **16**, 2992 (2007).