Reconstruction guarantees for compressive tomographic holography

Yair Rivenson¹, Adrian Stern², and Joseph Rosen¹

¹Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, P.O. Box 653, Beer-Sheva 84105, Israel
²Department of Electro-Optics Engineering, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel

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

Received May 12, 2013; revised Month X, XXXX; accepted Month X, XXXX;
posted Month X, XXXX (Doc. ID XXXXX); published Month X, XXXX

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 has been applied for this task. While many applications of compressive sensing for tomography from digital holograms have been demonstrated in the last few years, the fundamental limits involved have not yet been addressed. In this Letter, 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, 070.0070, 100.3190, 110.1758, 100.6890, 100.6950, 100.1830

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 wavefield 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 tomography 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 works 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 we may 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 using resolution constraints. For sake of completeness, we first give a short background on CS before deriving the main result. Generally, the object sensing mechanism is expressed using a matrix-vector multiplication:

\[ \mathbf{g} = \mathbf{Φf}, \]

where \( \mathbf{f} \in \mathbb{C}^{N \times 1} \) is the sparse object, \( \mathbf{Φ} \in \mathbb{C}^{M \times N} \) is the sensing operator and \( \mathbf{g} \in \mathbb{C}^{M \times 1} \) represents the measurements, where the number of measurements \( M \) is smaller than the number of objects pixels, \( N \). In order to guarantee accurate reconstruction, CS theory requires the object to have a sparse representation in some transform domain, i.e., the vector \( \mathbf{f} \in \mathbb{C}^{N \times 1} \) needs to have only \( S \) meaningful entries (\( S \)-sparse signal). The second essential requirement is on the sensing mechanism, and it can be quantified using the coherence parameter, given by:

\[ \mu = \max_{k \neq \ell} \left( \frac{\langle \mathbf{Φ}_k, \mathbf{Φ}_\ell \rangle}{\| \mathbf{Φ}_k \|_2 \| \mathbf{Φ}_\ell \|_2} \right), \]

where \( \mathbf{Φ}_k \) denotes the \( k \)-th column vector of \( \mathbf{Φ} \) and \( \langle \cdot, \cdot \rangle \) is the vectors inner product. The coherence parameter is bounded by:

\[ \sqrt{(N - M) / (M(N - 1))} \leq \mu \leq 1. \]

The CS theory guarantees accurate object recovery, by evoking a convex optimization procedure, with the number of accurately reconstructed non-zero terms obeying [5]:

\[ S \leq 0.5(1 + 1 / \mu). \]

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 \( \lambda \), illuminates a 3D object volume. The object volume is discretized into \( N_{\text{obj}} \times N_x \times N_z \) voxels, with voxel dimensions of \( \Delta x \times \Delta y \times \Delta z \), where the object length is \( L = N_x \times \Delta z \). The 3D object wavefield can be recorded by standard holography methods on a CCD [12].

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

\[ g(u \Delta x, v \Delta y) = \sum_{i=1}^{N_z} F_{2D}^{-1} \left( e^{-j2\pi/u \Delta x \vert \Delta i \vert} e^{-j2\pi/v \Delta y \vert \Delta i \vert} \right) e^{j2\pi/u \Delta x i} \]

\[ \times F_{2D} \left( f * (p \Delta x, q \Delta y; r \Delta z) \right), \]

where \( F_{2D} \) is the discrete 2D Fourier transform. For notation convenience, we further assume that \( \Delta x = \Delta y = \Delta z \).
and \( N_z = N_z' = \sqrt{M} \), therefore the discrete indices are
\( 0 \leq m,n,p,q,u,v \leq \sqrt{M} - 1 \), and \( \Delta u_j = \Delta v_j = 1/(\sqrt{M} \Delta z) \).

Equation (4) can be written as a matrix-vector multiplication by concatenation of \( N_z \) matrices, \( \mathbf{H}_{k\alpha z} \),
each one of size \( \sqrt{M} \times \sqrt{M} \): 
\[
\mathbf{g} = \left[ \mathbf{H}_{z\alpha} ; \ldots ; \mathbf{H}_{N_z\alpha} \right] \left[ \mathbf{f}_{z\alpha} ; \ldots ; \mathbf{f}_{N_z\alpha} \right] = \Phi \mathbf{f},
\]
where the matrix \( \mathbf{Q}_{k\alpha z} \) is a diagonal matrix which accounts for the quadratic phase terms of Eq. (4), and \( \left[ \mathbf{f}_{z\alpha} ; \ldots ; \mathbf{f}_{N_z\alpha} \right] \) 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 \( \mathbf{G} = \Phi^T \Phi \), \( \Phi \in \mathbb{C}^{(N_z \times M) \times (N_z \times M)} \) (where \( \Phi \) is column normalized, see Eq. (1)), \( \Phi^T \Phi \) is Hermitian conjugate of \( \Phi \), and \( \mu \) is the maximal off-diagonal absolute value [5].

Due to construction of \( \Phi \) 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 case where the volume is divided into \( N_z = 3 \) equally spaced planes.

It can be seen that the Gram matrix is built up of \( N_z \times N_z \) sub-blocks, each one representing the correlation between the point spread functions of two depth planes. Since the normalized \( H_{k\alpha} \) are orthonormal, the corresponding diagonal sub-blocks \( \mathbf{G}_{k\alpha} = \mathbb{H}_{k\alpha}^T \mathbb{H}_{k\alpha} = I \), i.e., all being zero beside the diagonal. This means that in order to find the coherence parameter we should look for it in any off-diagonal block, \( \mathbf{G}_{k\alpha} = \mathbb{H}_{k\alpha}^T \mathbb{H}_{k\alpha} \), which is equal to (for \( k \neq \ell \)):
\[
\mathbb{H}_{k\alpha} \mathbb{H}_{\ell\alpha} = \mathcal{F}_2^2 D^2 Q_{k\alpha} \mathcal{F}_2^2 D^2 Q_{\ell\alpha} = \mathcal{F}_2^2 D^2 Q_{(k-\ell)\alpha}.
\]

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

\[
\mu = \max_{k \neq \ell} \left| \mathbb{H}_{k\alpha} \mathbb{H}_{\ell\alpha} \right| = \max_{k \neq \ell} \left| \mathcal{F}_2^2 D^2 Q_{(k-\ell)\alpha} \right| \approx \max_{k \neq \ell} \left( \frac{\Delta^2}{\lambda (1-k) \Delta z} \right) = \frac{\Delta^2}{\lambda \Delta z}.
\]

Substituting Eq. (7) to Eq. (3) we find that the number of object features that can be accurately reconstructed is bounded by:
\[
S \leq 0.5 \left( 1 + \frac{\lambda \Delta z}{\Delta^2} \right).
\]

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 / 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: \( \mu \approx N_z \Delta^2 / (\lambda \Delta z) \). This expression settles 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 \( \mu \) gets, 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 \( (MN - M) / (M(N - 1)) = 1/\sqrt{M} \), 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 \( L = 316 \mu m \), randomly populated with \( S \) identical point particles (particle size \( \Delta = 5 \mu m \)) is illuminated using a plane wave with wavelength of \( \lambda = 633 n m \). The detector plane is positioned 1 mm from the volume and composed of 64-64 pixels with pitch of \( \Delta = 5 \mu m \). The object’s volume 2D wavefield is acquired, using phase shifting holography. The 3D object is then reconstructed using the two-step iterative shrinkage/thresholding (TwIST) solver [15] for object reconstruction, solving:
\[
\hat{f} = \arg \min \| \mathbf{g} - \Phi \mathbf{f} \|_\infty + \tau \| \mathbf{f} \|_1.
\]

The first term in Eq. (9) represents the data fidelity, \( \| \mathbf{g} - \Phi \mathbf{f} \|_\infty \), the \( p \)-norm, \( \tau \) 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 \( N_z = [2:10] \) 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 bigger than \( 10^{-6} \). The simulation was repeated for 10 times and the result shown in Fig. 3 is the mean of the different simulation instances. The shaded area denotes region of accurate reconstruction (MSE<10\(^{-6}\)) 3D object tomography reconstruction from its 2D hologram. From Fig. 3, we notice that as the axial resolution becomes finer, less 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 wavefield. For 3D objects, the number of recoverable particles is inverse proportional to the number of depth slices.

The results from the numerical experiment can be also interpreted as a super-resolving capability of the framework, meaning that if we have a small number of features, \( S \), we can reconstruct an object with a longitudinal resolution \( \Delta z \) which is improved compared with the classical resolution limit \( \Delta z \approx 4 \Delta^2 / \lambda = 158 \mu m \) [2]. When working in the near field Fresnel approximation, i.e., the entire diffraction pattern is essentially captured by the detector, the longitudinal resolution, \( \Delta z \), depends on the object’s feature size, \( \Delta \) [12]. 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 object’s sparsity. Hence, the simulation demonstrates that this
limitation can be practically relaxed when applying the CS framework. In fact, even the pessimistic estimation of the coherence parameter \[ \Delta z = 3\Delta^2 / \lambda \], 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 longitudinal accuracy of \( L_z/10 = 31.6\mu m \), which is 5 times finer than the fundamental limitations.

To conclude, we have formulated exact guarantees for 3D object tomography recovery from its 2D captured diffraction field when using the compressive sensing framework. The result shows that a digital hologram can indeed “compress” a 3D object with 5 degrees of freedom, where the number of degrees of freedom depends on the sensor’s resolution, axial resolution of the object and illuminating wavelength.

Acknowledgements

The authors wish to thank Israeli Ministry of Science and Technology (MOST) for supporting this research. AS and JR wishes to thank the Israel Science Foundation (grants No.1039/09 and 439/12) for supporting this work.

References


Fig. 1. (color online) Schematic description of the framework: A plane wave illuminating a volume (denoted in shaded gray). The wavefront scattered from the different particles is holographically recorded on a CCD.

Fig. 2. (color online) Gram matrix for the 3D-2D forward Fresnel sensing operator. The partition to 9 sub-matrices, marked by dashed lines, shows the coherence between the point spread function of two different planes.

Fig. 3. (color online) Simulation results showing the normalized number of reconstructed 3D object’s particles (MSEx10^-9) 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.
Informative Complete reference list