Journal of the Optical Society of America A
OPTICS, IMAGE SCIENCE, AND VISION

# Demonstration of single-shot digital holography using a Bayesian framework

Casey J. Pellizzari,[1,*] Matthias T. Banet,[2] Mark F. Spencer,[3] and Charles A. Bouman[1]

[1]School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA
[2]MZA Associates Corporation, Albuquerque, New Mexico 87106, USA
[3]Air Force Research Laboratory, Directed Energy Directorate, Kirtland AFB, New Mexico 87117, USA
*Corresponding author: casey.pellizzari@gmail.com

Received 11 September 2017; revised 16 November 2017; accepted 17 November 2017; posted 22 November 2017 (Doc. ID 306095); published 14 December 2017

**In this paper, we present experimental results for image reconstruction, with isoplanatic phase-error correction, from single-shot digital holography data. We demonstrate the utility of using a model-based iterative reconstruction (MBIR) algorithm to jointly compute the maximum *a posteriori* estimates of the phase errors and the real-valued object reflectance function. Specifically, we show that the MBIR algorithm is robust to noise and phase errors over a range of conditions.** 

***OCIS codes:*** (100.3190) Inverse problems; (100.3020) Image reconstruction-restoration; (010.7350) Wave-front sensing.

https://doi.org/10.1364/JOSAA.35.000103

## 1. INTRODUCTION

Digital holography (DH) can be used to sense both the amplitude and phase information returning from an actively illuminated object [1]. In practice, DH systems are sensitive to phase errors caused by index-of-refraction perturbations in the atmosphere or optical systems. These errors can often be estimated directly from the DH data. For wavefront sensing applications, the estimate of the phase errors is the desired sensor output [2]. Alternatively, for imaging applications, the phase errors must be estimated and corrected to form focused images.

Conventional techniques used to estimate phase errors from DH data involve maximizing an image sharpness metric [3–5]. These image-sharpening (IS) techniques reconstruct the complex-valued reflection coefficient, $g$, given by the complex-valued ratio of the reflected field to the incident field. For surfaces that are rough relative to the illumination wavelength, this leads to images with high spatial-frequency variations known as speckle. IS algorithms are sensitive to speckle variations and require incoherent averaging of multiple data realizations to estimate the phase errors with low error [3].

Recently, we developed a model-based iterative reconstruction (MBIR) algorithm for jointly computing the maximum *a posteriori* (MAP) estimates of the phase errors, $\phi$, and the real-valued reflectance, $r$, from single-shot DH data [6]. The reflectance is given by the real-valued ratio of the reflected power to the incident power. Furthermore, the reflectance can be expressed as $r = E[|g|^2]$, where $E[\cdot]$ indicates the expected value. The reflectance, in general, is smoother and has higher spatial correlation as compared to $g$. We are accustomed to seeing $r$ in conventional images, and it is of greater interest for many imaging applications. Additionally, by reconstructing $r$, we can leverage its higher spatial correlation to better constrain the estimation process and produce more accurate estimates of the phase errors with less data and less signal. In [6], the MBIR algorithm was shown to be robust to high noise and strong phase errors when tested on synthetic data.

In this paper, we analyze the performance of the MBIR algorithm using nonsynthetic, experimental data. First, we describe the laboratory setup used to generate DH data over a range of signal-to-noise ratios (SNRs) and atmospheric turbulence strengths. We then compare reconstructions from the MBIR algorithm to those from an IS algorithm.

## 2. EXPERIMENTAL SETUP

Figures 1 and 2 describe our experimental setup. For our master oscillator (MO) laser, we used an Oxxius laser source with 300 mW of continuous-wave power and approximately 50 m of coherence length at 532 nm. We used a half-wave plate and polarized beam-splitting cube to create two optical legs. In the first optical leg, we coupled the beam from the MO laser into a single-mode, polarization-maintaining (PM) fiber through a variable neutral-density (ND) filter and a half-wave plate. We then used a collimating lens and 20× optical beam expander to flood-illuminate the object. In the second optical leg, we coupled the beam from the MO laser into a separate single-mode, PM fiber through a half-wave plate to create an off-axis local oscillator (LO). This off-axis LO created a quasi-uniform reference beam with the appropriate tilt for digital holographic detection in the off-axis image plane recording

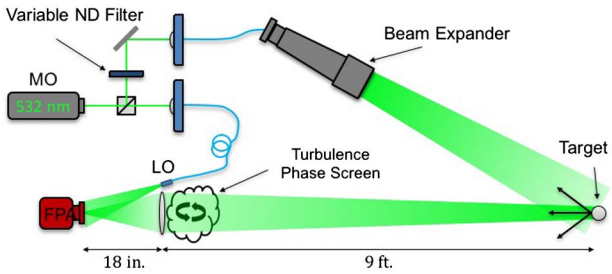


**Fig. 1.** Simplified diagram of our experimental setup. The output from a 532 nm laser was split into two paths. One path was attenuated by a variable ND filter and sent through a beam expander to illuminate the object. The other path was interfered with the scattered signal using an off-axis IPRG. A phase screen was placed in front of the imaging lens to simulate isoplanatic atmospheric turbulence.

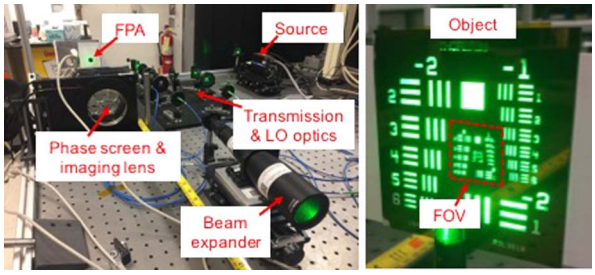


**Fig. 2.** Images of the transmitter and receiver optics (left) and the object (right) used in our experiment. The red-dashed box on the object shows the approximate field of view (FOV) for the imaging system.

geometry (IPRG) [2]. For detection, we used a single lens to image the scattered light from the object onto the focal-plane array (FPA) of a Guppy PRO f-125 FireWire camera. We saved the recorded digital holograms in a 964 × 964, 16-bit TIFF file.

We positioned an Air Force resolution chart, backed by cardstock, 9 ft away from the imaging lens and ensured that it was uniformly illuminated. In addition, we positioned the beam expander to minimize specular reflections from the object. After setting up the imaging system with an image distance of 18″, we simulated both weak and strong isoplanatic turbulence by placing different phase screens directly in front of the imaging lens. A clear piece of plastic was used to simulate the weak turbulence, and a Lexitek phase screen was used for the strong turbulence. We recorded digital holograms at four different signal levels, which we varied with the adjustable ND filter.

The digital hologram, $h$, must be demodulated and filtered to obtain the data, $y$, used for image reconstruction. Figure 3 shows the steps used to obtain $y$ for both the weak-turbulence (top row) and strong-turbulence (bottom row) cases. In looking at the spectrum of the strong-turbulence data, we see an irregularity towards the top of the pupil-plane image, which may have been caused by a reflection from the Lexitek phase screen. To avoid this irregularity, we used a smaller subset of the pupil-plane image samples, as indicated by the green dashed square. The resulting, smaller pupil-plane image data, $y$, was 200 × 200 pixels compared to 338 × 338 pixels for the weak-turbulence data.

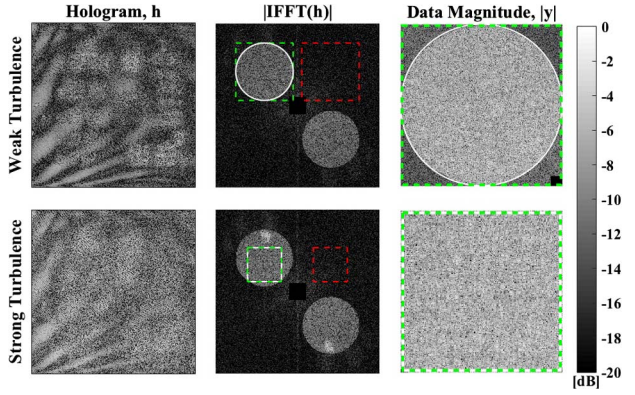


**Fig. 3.** Example of weak-turbulence (top row) and strong-turbulence (bottom row) data. The left column shows the raw 964 × 964 digital holograms, $h$. The second column shows the spectrum magnitude of $h$ along with the regions of interest. The green dashed line shows the subspace used for the complex pupil image, $y$, having a magnitude shown in the third column. The white solid line shows the boundary of the binary aperture transmission function, $a$, used for processing, and the red dashed line shows a region of the data, $y_n$, which contains primarily measurement noise used for computing the SNR. Note that the center 100 pixels × 100 pixels of the spectrum have been masked for plotting purposes only.

The solid white lines in Fig. 3 indicate the boundary of the binary aperture transmission function, $a$, used for image processing (i.e., $a = 1$ inside the white line and zero elsewhere). We used a circular aperture for the weak-turbulence data that matched the pupil function of our imaging system. For the strong-turbulence data, we used a square function that filled the entire data window. The red dashed line in Fig. 3 shows a region of the data, $y_n$, which contains primarily measurement noise. To quantify the SNR, we define the samples in the solid white region, which contains the signal, as $y_s$ and compute SNR according to

$$\mathrm{SNR} = \frac{s^2(y_s) - s^2(y_n)}{s^2(y_n)}, \tag{1}$$

where $s^2(\cdot)$ computes the sample variance of the argument.

## 3. ESTIMATION FRAMEWORK

The output data, indicated by a green dashed line in Fig. 3, is a complex-valued pupil-plane image and can be represented by a vector, $y \in \mathbb{C}^N$. In [6] we showed that $y$ can be formulated using an additive noise model given by

$$y = Af + w. \tag{2}$$

Here, $f \in \mathbb{C}^N$ is the field in the object plane given by $f = \Gamma g$, where $g \in \mathbb{C}^N$ is the object's complex-valued reflection coefficient and $\Gamma \in \mathbb{C}^{N\times N}$ is a diagonal matrix that applies the object-plane, quadratic-phase factor from the Fresnel propagation integral [7]. Given a reflectance function, $r$, the reflection coefficient, $g$, can be modeled using a complex normal distribution given by $p(g|r) \sim CN(0, \mathcal{D}(r))$, where $\mathcal{D}(\cdot)$ denotes an operator that produces a diagonal matrix from its vector argument. The vector $w \in \mathbb{C}^N$ in Eq. (2) is the measurement noise that has a complex normal distribution, $p(w) \sim CN(0, \sigma_w^2 I)$,

where $\sigma_w^2$ is the noise variance, and $I$ is the identity matrix. Finally, the matrix $A$ accounts for the propagation and measurement geometry and can be decomposed as

$$A = \mathcal{D}(a)\mathcal{D}(\exp\{j\phi\})F. \tag{3}$$

In Eq. (3), $a \in \mathbb{R}^N$ is the entrance-pupil transmission function and $\phi$ is the phase-error function. Finally, we choose the reconstruction parameters such that $F \in \mathbb{C}^{N\times N}$ is a two-dimensional discrete Fourier transform (DFT) matrix scaled so that $F^H F = I$, where the superscript, $H$, indicates the Hermitian transpose.

Our goal is to jointly compute the MAP estimates of $r$ and $\phi$ from $y$, which are given by

$$(\hat{r}, \hat{\phi}) = \underset{(r,\phi)\in\Omega}{\operatorname{argmin}}\{-\log p(y|r,\phi)p(r)\log p(\phi)\}, \tag{4}$$

**Initialize:** $r' \leftarrow |A^H y|^{\circ 2}$, $\phi' \leftarrow \mathbf{0}$
**Repeat{**
**E-step:**
$Q(r,\phi;r',\phi') = \mathbf{E}[-\log p(y|f,\phi)\, p(f|r)\, p(r) \log p(\phi)\,|y,r',\phi']$
**M-step:**
$\hat{r} \leftarrow \underset{r}{\operatorname{argmin}}\, Q(r,\phi';r',\phi')$
$\hat{\phi} \leftarrow \underset{\phi}{\operatorname{argmin}}\, Q(r',\phi;r',\phi')$
**Updates Variables:** $r' \leftarrow \hat{r}$, $\phi' \leftarrow \hat{\phi}$
**}**

**Fig. 4.** EM algorithm for joint MAP estimation of $r$ and $\phi$. Here, $|\cdot|^{\circ 2}$ indicates the element-wise magnitude square of a vector.

where $\Omega$ represents the jointly feasible set. Direct optimization of Eq. (4) is not practical, because it requires computing the determinate and inverse of a dense matrix. Instead, we use the expectation maximization (EM) algorithm to replace the cost function with a surrogate function given by

$$Q(r,\phi;r',\phi') = E[-\log p(y|f,\phi)p(f|r)p(r)\log p(\phi)|y,r',\phi'], \tag{5}$$

where $r'$ and $\phi'$ are the current values of $r$ and $\phi$, respectively, and the expectation is taken over the random vector, $f$, to form the $Q$ function [6]. Evaluation of the expectation in Eq. (5) constitutes the E-step of the EM algorithm.

Figure 4 shows the alternating minimization approach used for implementing the M-step. We used iterative coordinate descent (ICD) to minimize $Q$ with respect to $r$ and $\phi$ [8]. Appendix A provides details on the exact form of $Q$ as well as the prior models and parameters used for this experiment.

We use the iterative initialization process for $\phi$ described in [6]. The EM algorithm is run for $N_K$ iterations. We then restart the process by using our last estimate of $\phi$ as the new initial estimate, but reinitializing $r$ as $r' \leftarrow |A^H y|^{\circ 2}$. This process is repeated $N_L$ times. On the $N_L$th time, we allow the algorithm to run until $\|r^{(k)} - r^{(k-1)}\|_1 / \|r^{(k-1)}\|_1 \le 1 \times 10^{-4}$, where $k$ is the iteration index and $\|\cdot\|_1$ indicates the L1 norm of a vector. For this work, we set $N_K = 10$ and $N_L = 50$ for the weak-turbulence reconstructions and $N_K = 100$ for the strong-turbulence reconstructions.

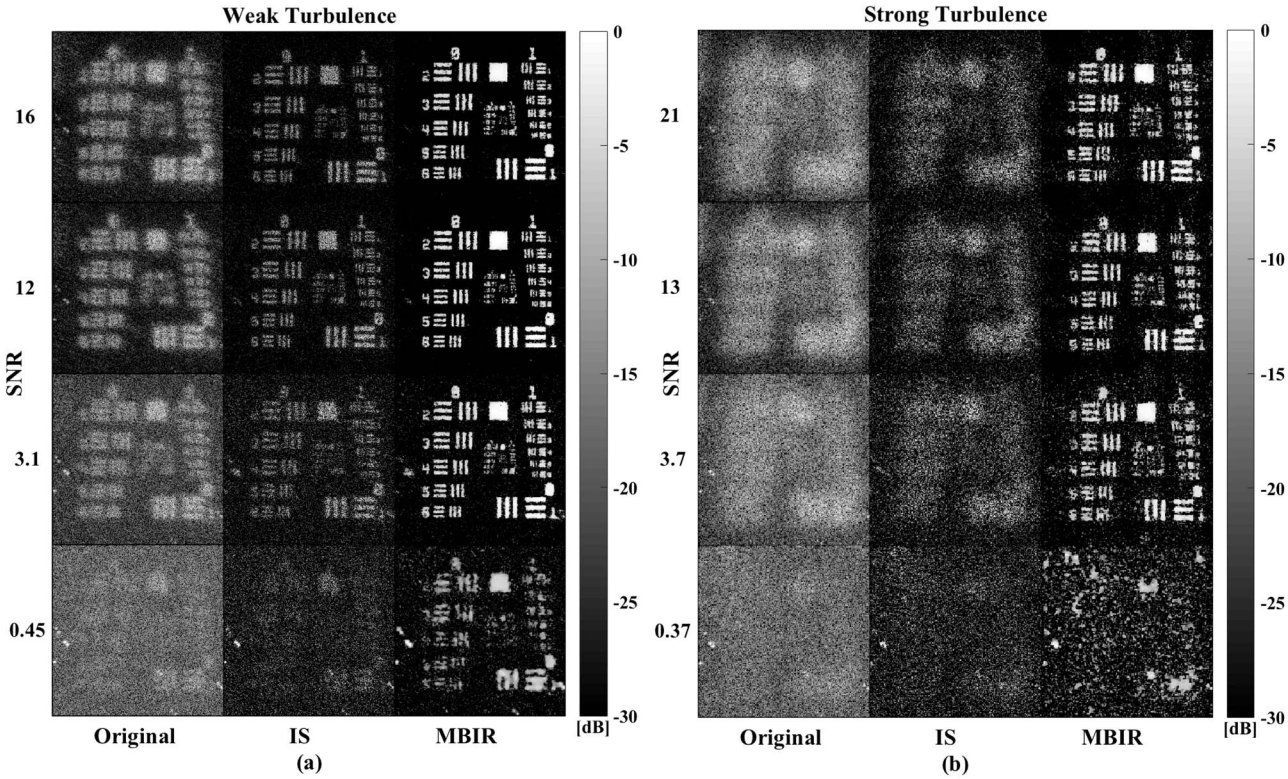


**Fig. 5.** (a) 338 × 338 weak-turbulence reconstructions, and (b) 200 × 200 strong-turbulence reconstructions. For both subplots, the left column shows the original blurry image, the middle column shows the image-sharpening results, and the third column shows the MBIR results. It is important to note that these results are shown using a log-based decibel scale.

## 4. RESULTS AND CONCLUSIONS

We chose to compare the MBIR algorithm to the point-by-point IS approach presented in [3] using the $M_2$ sharpness metric. The algorithm computes the phase-error estimate according to

$$\hat{\phi} = \underset{\phi}{\operatorname{argmax}}\{-\|(|F^H\mathcal{D}(\exp\{j\phi\})^H y|^{\circ 2})^{\circ 0.5}\|_1\}, \tag{6}$$

where $^\circ$ indicates the application of an exponent to each vector element. Following the process described in [3], we used 20 iterations of conjugate gradient to optimize Eq. (6), and the algorithm was initialized using a 15th-order Zernike polynomial estimate obtained using an iterative method to estimate only up to the 3rd-order terms, then up to the 4th, and so on, continuing up to 15th order.

Figure 5 shows reconstructions for both the IS and MBIR algorithms along with the original blurry images for four different SNRs and two turbulence strengths. To compress the large dynamic range that occurs in coherent images, we present the images using a log-based decibel scale given by $r_{\mathrm{dB}} = 10\log_{10}(\tilde{r})$, where $\tilde{r} \in [0, 1]$ is the normalized reflectance function. Figure 6 shows the phase-error estimates corresponding to the reconstructions in Fig. 5.

The weak-turbulence results show that the IS algorithm is able to correct most of the phase errors; however, the resultant images are speckled, they have low contrast, and there is residual blurring. On the other hand, the MBIR algorithm produces images that have high contrast, with most of the object near the peak image value and the background more than 30 dB below that. Furthermore, the MBIR algorithm produces images that have much less speckle variation and residual blurring compared with the IS algorithm. At the lowest SNR, the MBIR algorithm makes it easier to distinguish the object from the background.

The strong-turbulence results show that the IS algorithm is not able to estimate strong phase errors from single-shot data. Conversely, the MBIR algorithm is able to produce highly focused images with reduced speckle variations and high contrast at all but the lowest SNR. These experimental results closely resemble the simulated results in [6].

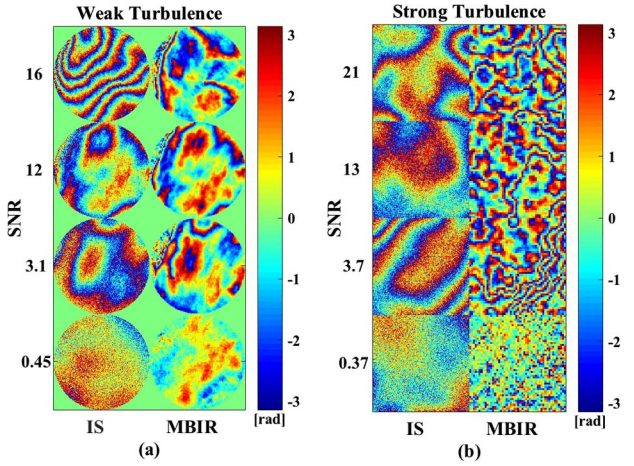


**Fig. 6.** IS and MBIR phase error estimates, $\hat{\phi}$, for (a) the weak-turbulence cases and (b) the strong-turbulence cases, corresponding to Fig. 5.

In summary, we have experimentally demonstrated the utility of the MBIR algorithm for image reconstruction from single-shot DH data. The MBIR algorithm was shown to outperform an IS algorithm over a range of SNRs and turbulence strengths. Overall, the MBIR algorithm was able to produce highly focused images with reduced speckle variation and high contrast for all but the lowest SNR cases.

## APPENDIX A: EM SURROGATE FUNCTION

Following [6], the EM surrogate function is given by

$$\begin{aligned} Q(r,\phi;r',\phi') = &-\frac{1}{\sigma_w^2}2\,\mathrm{Re}\{y^H A_\phi \mu\} + \log|\mathcal{D}(r)| \\ &+ \sum_{i=1}^{N}\frac{1}{r_i}(C_{i,i}+|\mu_i|^2) \\ &+ \sum_{\{i,j\}\in\mathcal{P}} b_{i,j}\frac{|\Delta_r|^p}{p\sigma_r^p}\left(\frac{\left|\frac{\Delta_r}{T\sigma_r}\right|^{q-p}}{1+\left|\frac{\Delta_r}{T\sigma_r}\right|^{q-p}}\right) \\ &+ \sum_{\{i,j\}\in\mathcal{P}} b_{i,j}\frac{|\Delta_{\overline{\phi}}|^2}{2\sigma_{\overline{\phi}}^2}. \end{aligned} \tag{A1}$$

Here, $\mathrm{Re}\{\cdot\}$ indicates the real part of the argument, the superscript, $H$, indicates the Hermitian transpose, and the subscript, $\phi$, indicates the dependence of $A$ on $\phi$. The variables $\mu$ and $C$ are the mean and covariance matrix for the complex normal posterior distribution, $p(f|y, r, \phi)$, and are given by

$$\mu = C\frac{1}{\sigma_w^2}A_{\phi'}^H y \tag{A2}$$

and

$$C = \left[\frac{1}{\sigma_w^2}A_{\phi'}^H A_{\phi'} + \mathcal{D}(r)^{-1}\right]^{-1} \approx \mathcal{D}\left(\frac{\sigma_w^2}{1+\frac{\sigma_w^2}{r}}\right). \tag{A3}$$

For circular apertures, we approximate the covariance matrix as shown in Eq. (A3), since it simplifies computations, and we have found that it works well. For square apertures, the approximation in Eq. (A3) is exact.

To reduce the number of unknowns, we allow the phase-error function, $\phi$, to be modeled on a grid that has lower resolution than the measured data, denoted as $\overline{\phi}$. To scale $\overline{\phi}$ to the resolution of $\phi$, we use a nearest-neighbor interpolation scheme given by $\phi = P\overline{\phi}$, where $P$ is an $N \times N/n_b^2$ interpolation matrix with elements in the set $\{0, 1\}$, and $n_b$ is the factor of subsampling used in both dimensions. In this work, we used $n_b = 4$.

The last two terms in Eq. (A1) result from using Markov random field prior models for $r$ and $\overline{\phi}$. The variable $b_{i,j}$ is the weight between neighboring pixel pairs ($r_i$ and $r_j$, or $\phi_i$ and $\phi_j$), and $\mathcal{P}$ is the set of all pair-wise cliques falling within the same neighborhood. We used a $3 \times 3$ Gaussian kernel with standard deviation 0.1 pixels for $b$. $\Delta_r = r_i - r_j$ and $\Delta_\phi = \phi_i - \phi_j$ are the difference between pixel-pair values, $T$ is a unitless threshold value that controls the transition of the potential function from having the exponent, $q$, to having the exponent, $p$ [9]. In this work, $T = 0.1$, $q = 2$, and

$p = 1.1$. The variable, $\sigma_r$, controls the variation in $\hat{r}$, and $\sigma_{\overline{\phi}}$ controls the variation in $\hat{\overline{\phi}}$. We set $\sigma_{\overline{\phi}} = 0.5$ rad for weak turbulence and $\sigma_{\overline{\phi}} = 1.0$ rad for strong turbulence. The values of $\sigma_r$ and $\sigma_w$ were set according to $\sigma_r \leftarrow \gamma^{-1}[s^2(r^{(0)})]^{0.5}$ and $\sigma_w \leftarrow [N^{-1}y^H y]^{0.5}$, where $r^{(0)}$ is the initialized value of $r$, $\gamma$ is a unitless parameter introduced to tune the amount of regularization in $r$, and $s^2(\cdot)$ computes the sample variance of a vector's elements. We set $\gamma = 2$ for $N = 338^2$ and $\gamma = 1.5$ for $N = 200^2$.

## REFERENCES


1. T.-C. Poon and J.-P. Liu, *Introduction to Modern Digital Holography: With MATLAB* (Cambridge University, 2014).
2. M. F. Spencer, R. A. Raynor, M. T. Banet, and D. K. Marker, "Deep-turbulence wavefront sensing using digital-holographic detection in the off-axis image plane recording geometry," Opt. Eng. **56**, 031213 (2017).
3. S. T. Thurman and J. R. Fienup, "Phase-error correction in digital holography," J. Opt. Soc. Am. A **25**, 983–994 (2008).
4. J. C. Marron, R. L. Kendrick, N. Seldomridge, T. D. Grow, and T. A. Höft, "Atmospheric turbulence correction using digital holographic detection: experimental results," Opt. Express **17**, 11638–11651 (2009).
5. A. E. Tippie and J. R. Fienup, "Multiple-plane anisoplanatic phase correction in a laboratory digital holography experiment," Opt. Lett. **35**, 3291–3293 (2010).
6. C. Pellizzari, M. F. Spencer, and C. A. Bouman, "Phase-error estimation and image reconstruction from digital-holography data using a Bayesian framework," J. Opt. Soc. Am. A **34**, 1659–1669 (2017).
7. J. W. Goodman, *Introduction to Fourier Optics* (Roberts & Company, 2005).
8. C. A. Bouman, "Model based image processing," 2013, https://engineering.purdue.edu/~bouman/publications/pdf/MBIP-book.pdf.
9. J. B. Thibault, K. Sauer, C. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multi-slice helical CT," Med. Phys. **34**, 4526–4544 (2007).