A practical inverse-problem approach to digital holographic reconstruction

Aurélien Bourquard,[1*] **Nicolas Pavillon,**[2,3] **Emrah Bostan,**[1] **Christian Depeursinge,**[2] **and Michael Unser**[1]

[1]*Biomedical Imaging Group, École polytechnique fédérale de Lausanne, Switzerland*
[2]*MicroVision and MicroDiagnostics Group, École polytechnique fédérale de Lausanne, Switzerland*
[3]*Current address: Biophotonics Laboratory, Immunology Frontier Research Center (IFReC), Osaka University, Suita, Osaka 565-0871, Japan*

*[*aurelien.bourquard@epfl.ch](mailto:aurelien.bourquard@epfl.ch)*

Abstract: In this paper, we propose a new technique for high-quality reconstruction from single digital holographic acquisitions. The unknown complex object field is found as the solution of a nonlinear inverse problem that consists in the minimization of an energy functional. The latter includes total-variation (TV) regularization terms that constrain the spatial amplitude and phase distributions of the reconstructed data. The algorithm that we derive tolerates downsampling, which allows to acquire substantially fewer measurements for reconstruction compared to the state of the art. We demonstrate the effectiveness of our method through several experiments on simulated and real off-axis holograms.

© 2013 Optical Society of America

OCIS codes: (090.1995) Digital holography; (100.2000) Digital image processing; (100.3010) Image reconstruction techniques; (100.3190) Inverse problems; (100.5070) Phase retrieval.

References and links

1. E. Cuche, P. Marquet, and C. Depeursinge, "Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms," Appl. Opt. **38**(34), 6994–7001 (1999).
2. I. Yamaguchi and T. Zhang, "Phase-shifting digital holography," Opt. Lett. **22**(16), 1268–1270 (1997).
3. G. Popescu, T. Ikeda, R. Dasari, and M. Feld, "Diffraction phase microscopy for quantifying cell structure and dynamics," Opt. Lett. **31**(6), 775–777 (2006).
4. Y. Awatsuji, T. Tahara, A. Kaneko, T. Koyama, K. Nishio, S. Ura, T. Kubota, and O. Matoba, "Parallel two-step phase-shifting digital holography," Appl. Opt. **47**(19), D183–D189 (2008).
5. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," Optik **35**(2), 227–246 (1972).
6. V. Katkovnik, A. Migukin, and J. Astola, "3D wave field reconstruction from intensity-only data: variational inverse imaging techniques," in *9th Euro-American Workshop on Information Optics* (Helsinki, Finland, 2010).
7. M. Cetin, W. Karl, and A. Willsky, "Edge-preserving image reconstruction for coherent imaging applications," in *Proceedings of ICIP 2002 International Conference on Image Processing* (Rochester, NY, USA, 2002).
8. X. Zhang and E. Lam, "Edge-preserving sectional image reconstruction in optical scanning holography," J. Opt. Soc. Am. A **27**(7), 1630–1637 (2010).
9. D. Brady, K. Choi, D. Marks, R. Horisaki, and S. Lim, "Compressive holography," Opt. Express **17**(15), 13,040–13,049 (2009).
10. M. Marim, M. Atlan, E. Angelini, and J. Olivo-Marin, "Compressed sensing with off-axis frequency-shifting holography," Opt. Lett. **35**(6), 871–873 (2010).
11. Y. Rivenson, A. Stern, and B. Javidi, "Compressive fresnel holography," J. Disp. Technol. **6**(10), 506–509 (2010).
12. S. Sotthivirat and J. Fessler, "Penalized-likelihood image reconstruction for digital holography," J. Opt. Soc. Am. A **21**(5), 737–750 (2004).
13. M. Marim, E. Angelini, J.-C. Olivo-Marin, and M. Atlan, "Off-axis compressed holographic microscopy in low-light conditions," Opt. Lett. **36**(1), 79–81 (2011).

14. N. Pavillon, C. Seelamantula, J. Kuhn, M. Unser, and C. Depeursinge, "Suppression of the zero-order term in off-axis digital holography through nonlinear filtering," Appl. Opt. **48**(34), H186–H195 (2009).
15. T. Colomb, J. Kühn, F. Charrière, C. Depeursinge, P. Marquet, and N. Aspert, "Total aberrations compensation in digital holographic microscopy with a reference conjugated hologram," Opt. Express **14**(10), 4300–4306 (2006).
16. M. Bigas, E. Cabruja, J. Forest, and J. Salvi, "Review of CMOS image sensors," Microelectron. J. **37**(5), 433–451 (2006).
17. L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Phys. D **60**(1–4), 259–268 (1992).
18. T. Chan and P. Mulet, "On the convergence of the lagged diffusivity fixed point method in total variation image restoration," SIAM J. Numer. Anal. **36**(2), 354–367 (1999).
19. O. Scherzer, ed., *Energy Minimization Methods* (Springer, 2011).
20. J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw Hills Companies, Inc., 1996).
21. M. Guizar-Sicairos, A. Diaz, M. Holler, M. Lucas, A. Menzel, R. Wepf, and O. Bunk, "Phase tomography from x-ray coherent diffractive imaging projections," Opt. Express **19**(22), 21,345–21,357 (2011).
22. K. Choi, R. Horisaki, J. Hahn, S. Lim, D. L. Marks, T. J. Schulz, and D. J. Brady, "Compressive holography of diffuse objects," Appl. Opt. **49**(34), H1–H10 (2010).

1. Introduction

Given an intensity hologram that is acquired in a proper configuration, it is generally possible to reconstruct the complex field, commonly through demodulation, either in the temporal domain with the so-called phase-shifting approach [1], or in the spatial domain in an off-axis configuration, where complex information can be retrieved through Fourier filtering [2]. Several solutions have been proposed for reconstruction following these approaches [3, 4]. The corresponding algorithms are presently used in practice and incorporated in digital holographic microscopy for instance. Their common characteristic is that the spectral information that is actually used for reconstruction reduces to the first diffraction orders.

Besides direct Fourier-based reconstruction, iterative procedures that are based on more implicit formulations have been proposed for the recovery of complex fields from intensity data. This includes the well known Gerchberg-Saxton algorithm used for phase retrieval in its original form [5] as well as more evolved strategies [6]. The recent trend in the literature is to formulate the task of digital holographic reconstruction as an inverse problem. For instance, Cetin *et al.* [7] have applied total-variation-type regularization to reconstruct amplitude profiles based on a linear physical model. Zhang and Lam have followed a similar approach for optical scanning holography [8]. In particular, they provided reconstruction results using model-matched phantom experiments. In their work, Bardy *et al.* [9] and Marim *et al.* [10] have demonstrated the possibility of reconstructing amplitude profiles from undersampled holograms in the context of Gabor and off-axis holography, respectively. Along the same line of research, Rivenson *et al.* [11] have investigated distinct undersampling strategies. Note that, although the aforementioned methods yield promising results compared to standard reconstruction techniques, they either require the acquisition of two or more holographic planes or assume a linearized model for reconstruction. Meanwhile, Sotthivirat and Fessler [12] have developed an optimization-transfer method for reconstruction that involves a nonlinear model and that also handles the case of a single acquisition. Focusing on image-plane holography, their work has shown substantial improvements on simulated data compared to the conventional approaches used for that particular setting.

In this paper, we propose a new variational-reconstruction approach that can be applied to experimentally-acquired holographic data. Our technique accurately models the measurement system based on a nonlinear forward model and allows to recover the complex object field from one single off-axis intensity hologram. Compared to the state of the art, the main contribution of our reconstruction algorithm is to employ side information apart from the measured data to improve the quality of reconstruction. Specifically, we make prior assumptions on some shared object properties, and also use the knowledge of a reference correction hologram (RCH)

to simplify the task of reconstruction and thus find suitable solutions. Another asset of our method is to exploit the complete image information at hand. The assumptions on the solution are used as a mean to regularize the amplitude and phase components of the complex field, which allows to perform reconstructions from spatially undersampled measurements. Note that an undersampling scheme was initially proposed in [13] for a particular off-axis protocol in low-level imaging conditions where the forward model can be linearized. However, unlike in our case, the protocol of [13] requires the complex amplitudes of the field to be first determined at the hologram plane through the combination of several intensity acquisitions. This is a step that can be avoided with our new nonlinear-model-based method presented in the sequel.

The paper is organized as follows. In Sect. 2, we describe the nonlinear model that corresponds to our own holographic setup, and introduce the relevant notations. In Sect. 3, we introduce our formulation of the reconstruction problem, and discuss the relevance of this new approach for obtaining high-quality solutions in practice. We then derive our reconstruction algorithm in Sect. 4. In Sect. 5, we conduct experiments on synthetic and real holograms, and compare our method with the state of the art from both qualitative and quantitative standpoints. We conclude our work in Sect. 6.

2. Forward model

The general holographic model that we consider is standard in the literature. It involves a coherent object field $o(\mathbf{x};z)$ and a reference field $r(\mathbf{x};z)$ that interfere to form the measured intensities

$$I(\mathbf{x}) = |o(\mathbf{x};0) + r(\mathbf{x};0)|^2. \tag{1}$$

The coordinate vector $\mathbf{x} = (x_1,x_2)$ corresponds to the spatial position parallel to the hologram plane, while the coordinate z denotes the signed distance from the hologram plane in the direction of light propagation. Note that the squared norm appearing in Eq. (1) makes the intensities I depend nontrivially on the field o. This will cause our own reconstruction problem to be non-convex, as further discussed in Sect. 3.

Following the MKS system of units, the physical parameters of the acquisition system consist in the wavelength λ, the camera pixel size Δ_x, the detector resolution N, the numerical aperture NA of the microscope objective, the object-space immersion medium n_i, and the magnification factor M. Assuming a diffraction-limited configuration, the spatial bandwidth of the object corresponds to $\Delta\omega_0 = 2\mathrm{NA}\Delta_x N/(Mn_i\lambda)$ [14]. In the case of an off-axis hologram, a demodulated and non-aberrated object field can potentially be retrieved by employing an estimate $\tilde{r}$ of the reference wave $r(\mathbf{x};0)$ employed for recording, up to a multiplicative factor α. One common way to get $\tilde{r}$ is to use a calibration measurement with an empty field of view which accounts for the phase terms induced by the optical system, denoted RCH [15].

We assume in this work that the holographic device measures an infinitely thin object profile. The field o is thus expected to be focused at some position $z = -d$, the corresponding planar profile being denoted as Ψ_d. While the in-focus propagation distance d is classically independent from the actual object-field reconstruction, it is relevant to our method because we shall introduce some prior knowledge on Ψ_d in Sect. 3. Accordingly, we consider Ψ_d as our unknown, and reformulate Eq. (1) as

$$I(\mathbf{x}) = |(\mathscr{P}_d\mathscr{H}\Psi_d)(\mathbf{x}) + r(\mathbf{x};0)|^2. \tag{2}$$

The linear operators $\mathscr{P}_d$ and $\mathscr{H}$ denote digital propagation with distance d and spatial-frequency cut-off associated to the microscope objective (MO), respectively. These operators behave as a function of the parameters of the holographic setup, and their mathematical expressions are provided in Appendix A.

Finally, according to the pixel size Δ_x of the sensor array, the intensities $I(\mathbf{x})$ are sampled to

$$I[\mathbf{m}] = |(\mathscr{P}_d \mathscr{H} \Psi_d)(\mathbf{x}) + r(\mathbf{x};0)|^2|_{\mathbf{x}=\mathbf{m}\Delta_x}, \tag{3}$$

and a binary mask w predefines what samples are actually measured according to some down-sampling factor D. Each sample $I[\mathbf{m}]$ is thus counted as a measurement if and only if the value $w[\mathbf{m}]$ is unity for the same $\mathbf{m}$. The mask follows the pseudorandom structure of [13], and consists in binary values that are independent, identically distributed, and nonzero with probability $1/D$. As shown in our experimental section, this generalized formulation successfully handles cases where the amount of data that is available for reconstruction is substantially reduced. Note that downsampled acquisitions potentially allow for faster imaging when using cameras where pixel values are addressed in random access [16].

3. Reconstruction problem

According to Eq. (3), the goal of holographic reconstruction is to produce the most precise estimate of the object wave Ψ_d at reconstruction distance d, given noisy and quantized versions of the intensity samples I. In this section, we propose an estimate $\tilde{\Psi}_d$ of Ψ_d that involves all the available measured data. Since the complex field Ψ_d corresponds to an unknown quantity for our reconstruction algorithm, it is treated as a variable in the sequel.

3.1. Consistent formulation

Following a variational framework, we define the estimated complex field $\tilde{\Psi}_d$ as the solution of an inverse problem. This solution must be matched to the available measurements, and has to be regularized so as to make the problem well posed.

According to Eq. (3), we define the class of suitable solutions as those yielding measurements that are compatible with the image data after reintroduction into our forward model. Accordingly, given the samples of I acquired by the sensor array, one can express a so-called *consistency constraint* on Ψ_d as

$$I(\mathbf{x}) = |(\mathscr{P}_d \mathscr{H} \Psi_d)(\mathbf{x}) + \alpha \bar{r}(\mathbf{x})|^2, \forall \mathbf{x} \in \mathbf{m}\Delta_x \text{ with } w[\mathbf{m}] = 1. \tag{4}$$

The function $\bar{r}$ corresponds to a normalized version of $\tilde{r}$ whose phases are left untouched and whose amplitudes are replaced by their mean pointwise. This allows the actual reference amplitudes to be fully determined by the factor α obtained through the reconstruction process. Due to the presence of noise, we relax the strict data-fidelity constraint of Eq. (4) and rewrite it as

$$\mathscr{D}(\Psi_d, \alpha) = \sum_{\mathbf{m} \in \mathbb{Z}^2} w[\mathbf{m}] \left(|(\mathscr{P}_d \mathscr{H} \Psi_d)(\mathbf{x}) + \alpha \bar{r}(\mathbf{x})|^2 - I(\mathbf{x}) \right)^2 |_{\mathbf{x}=\mathbf{m}\Delta_x} < \varepsilon. \tag{5}$$

Besides Ψ_d, the multiplicative factor α influences the value of $\mathscr{D}$ in Eq. (5); both quantities shall be alternately optimized in our algorithm, as described in Sect. 4. Finally, the positive scalar ε determines to what extent the reconstructed solution can depart from the available measurements, depending on the level of noise that affect them. Based on the sampled intensities, this constant can be deduced from the signal-to-noise ratio (SNR) of the acquisition device ε_{SNR} through the relation

$$\varepsilon = \exp(-\varepsilon_{\text{SNR}} \ln(10)/10) \sum_{\mathbf{m} \in \mathbb{Z}^2} w[\mathbf{m}] I^2(\mathbf{x})|_{\mathbf{x}=\mathbf{m}\Delta_x}, \tag{6}$$

as shown in Appendix B. Under the sole data-fidelity constraint of Eq. (5), the solution $\tilde{\Psi}_d$ is underdetermined. In order to make the problem well posed, we regularize Ψ_d by minimizing the

total variation (TV) [17] of its phase and amplitude components separately. From a qualitative standpoint, our regularization approach is derived from the assumption that the unwrapped-phase and amplitude maps of the focused hologram are both well approximated by piecewise-constant functions as they are in-focus. This assumption shall prove fruitful for improving the reconstruction quality, as shown in the experimental section. Note also that our reconstruction algorithm proposed in Sect. 4 does not require any phase unwrapping to regularize the solution. Denoting the phase and amplitude TV values by $\text{TV}^\phi(\Psi_d)$ and $\text{TV}^\rho(\Psi_d)$, respectively, we write

$$\text{TV}^\phi(\Psi_d) = \int_{\mathbb{R}^2} \|\nabla(\arg_u \Psi_d)(\mathbf{x})\| d\mathbf{x}, \tag{7}$$

$$\text{TV}^\rho(\Psi_d) = \int_{\mathbb{R}^2} \|\nabla|\Psi_d|(\mathbf{x})\| d\mathbf{x}, \tag{8}$$

where $\arg_u \Psi_d$ denotes the unwrapped phase of Ψ_d. The above thus corresponds to the constrained minimization problem

$$\tilde{\Psi}_d = \arg\min_{\Psi_d} \left[\text{TV}^\phi(\Psi_d) + \gamma \text{TV}^\rho(\Psi_d) \right] \quad \text{s.t.} \quad \tilde{\mathscr{D}}(\Psi_d) < \varepsilon, \tag{9}$$

where γ is a positive scalar weight that balances the respective influence of phase and amplitude regularization, and where

$$\tilde{\mathscr{D}}(\Psi_d) = \arg\min_{\alpha} \mathscr{D}(\Psi_d, \alpha). \tag{10}$$

Given the structure of Eq. (1), the expression of $\mathscr{D}(\Psi_d, \alpha)$ in Eq. (5) contains two imbricated squaring operations. This implies that $\tilde{\mathscr{D}}(\Psi_d)$ is a non-convex functional of the field Ψ_d. The problem of Eq. (9) can thus have several local optima; its solution is non-unique in general.

3.2. Discretization

Using vector notation, we discretize the unknowns and the aforementioned operators according to the pixel size Δ_x. Each vector contains the components of the corresponding discretized map in lexicographic order. Accordingly, the discrete solution is defined as

$$\tilde{\mathbf{\Psi}}_d = \arg\min_{\mathbf{\Psi}_d} \left[\text{TV}^\phi(\mathbf{\Psi}_d) + \gamma \text{TV}^\rho(\mathbf{\Psi}_d) \right] \quad \text{s.t.} \quad \tilde{\mathscr{D}}(\mathbf{\Psi}_d) < \varepsilon. \tag{11}$$

The corresponding data term $\mathscr{D}$ takes the form

$$\mathscr{D}(\mathbf{\Psi}_d, \alpha) = \sum_i \mathbf{w}_i \left(|(\mathbf{A}\mathbf{\Psi}_d)_i + \alpha \bar{\mathbf{r}}_i|^2 - \mathbf{y}_i \right)^2, \tag{12}$$

where the vectors $\mathbf{w}$, $\mathbf{\Psi}_d$, $\bar{\mathbf{r}}$, and $\mathbf{y}$ correspond to the sampled versions of w, Ψ_d, $\bar{r}$, and I, respectively. The matrix $\mathbf{A}$ corresponds to the successive application of the linear operators $\mathscr{H}$ and $\mathscr{P}_d$ in discretized form. The discretized regularization terms read

$$\text{TV}^\phi(\mathbf{\Psi}_d) = \sum_i \left[(\mathbf{R}_1(\arg_u \mathbf{\Psi}_d))_i^2 + (\mathbf{R}_2(\arg_u \mathbf{\Psi}_d))_i^2 + \nu \right]^{1/2}, \tag{13}$$

$$\text{TV}^\rho(\mathbf{\Psi}_d) = \sum_i \left[(\mathbf{R}_1|\{\mathbf{\Psi}_d\}|)_i^2 + (\mathbf{R}_2|\{\mathbf{\Psi}_d\}|)_i^2 + \nu \right]^{1/2}, \tag{14}$$

where the $\mathbf{R}_{1,2}$ are convolution matrices implementing the discretized first derivatives in the horizontal and vertical directions, respectively, and where ν is an additive parameter that ensures

the differentiability of both TV functionals [18]. The horizontal and vertical filters associated with $\mathbf{R}_{1,2}$ are each one-dimensional, and correspond to $(z_{1,2}^{-1} - 1)$ in the $\mathscr{Z}$-transform domain. The symbol $|\{\cdot\}|$ denotes the absolute value applied to a vector pointwise. Similarly, $\mathbf{arg}_u$ takes the unwrapped argument of the vector components pointwise.

4. Reconstruction algorithm

We solve the problem of Eq. (11) based on a gradient-descent approach that we combine with a specifically devised line-search strategy. Specifically, we propose to apply the following iterative procedure:

1. Initialize the solution as $\mathbf{\Psi}_d = \mathbf{0}$.

2. Find the optimal reference factor $\tilde{\alpha} = \arg\min_\alpha \mathscr{D}(\mathbf{\Psi}_d, \alpha)$ to determine $\tilde{\mathscr{D}}(\mathbf{\Psi}_d)$.

3. If $\tilde{\mathscr{D}}(\mathbf{\Psi}_d) > \varepsilon$:

    (a) Compute the data-term gradient $\nabla\tilde{\mathscr{D}} = \partial\tilde{\mathscr{D}}(\mathbf{\Psi}_d)/\partial\mathbf{\Psi}_d$.
    (b) Compute the step $\tilde{\tau} \in \mathbb{R}_+$ s.t. $\tilde{\mathscr{D}}(\mathbf{\Psi}_d - \tilde{\tau}\nabla\tilde{\mathscr{D}}) = \varepsilon$ if it exists;
        optimize $\tilde{\tau} = \arg\min_\tau \tilde{\mathscr{D}}(\mathbf{\Psi}_d - \tau\nabla\tilde{\mathscr{D}})$ otherwise.
    (c) Update the solution as $\mathbf{\Psi}_d \leftarrow \mathbf{\Psi}_d - \tilde{\tau}\nabla\tilde{\mathscr{D}}$.

4. If $\tilde{\mathscr{D}}(\mathbf{\Psi}_d) \leq \varepsilon$:

    (a) Compute the phase-regularization gradient $\nabla\mathrm{TV}^\phi = \partial\mathrm{TV}^\phi(\mathbf{\Psi}_d)/\partial\mathbf{\Psi}_d$.
    (b) Compute the amplitude-regularization gradient $\nabla\mathrm{TV}^\rho = \partial\mathrm{TV}^\rho(\mathbf{\Psi}_d)/\partial\mathbf{\Psi}_d$.
    (c) Update the solution as $\mathbf{\Psi}_d \leftarrow \mathbf{\Psi}_d - \tau_{\mathrm{reg}}\left(\nabla\mathrm{TV}^\phi + \gamma\nabla\mathrm{TV}^\rho\right)$.

5. Return to Step 2 until V iterations are reached.

The positive scalar τ_{reg} is a step parameter that determines the strength of the regularization flow. The corresponding regularization gradients $\nabla\mathrm{TV}^\phi, \nabla\mathrm{TV}^\rho$ used in Step 4 and the data-term gradient $\nabla\tilde{\mathscr{D}}$ used in Step 3 are complex-valued with $\nabla\cdot = \partial\cdot/\partial\mathrm{Re}(\mathbf{\Psi}_d) + \mathrm{i}(\partial\cdot/\partial\mathrm{Im}(\mathbf{\Psi}_d))$. The explicit forms of the update terms are then obtainable from Eqs. (12), (13), and (14) using differential calculus. Our regularization technique does not require any explicit phase unwrapping, as explained in Appendix C. Finally, determining either of the two scalars $\tilde{\alpha}$ and $\tilde{\tau}$ in Steps 2 and 3(b) amounts to solving an elementary line-search problem. This issue is addressed in more detail in Appendix D.

Since the problem of Eq. (11) is non-convex, the solution obtained with our algorithm potentially depends on the initial estimate. In practice, however, no significant changes were observed, hence our generic choice of zero initialization in Step 1. While rigorous convergence analysis is nontrivial in non-convex settings [19], experimental evidence suggests that our algorithm converges to a fixed point. Indeed, the solution and the corresponding value of the regularization functional in Eq. (11) have been observed to stabilize as the iterations proceed. Each iteration of our algorithm is comparable to the standard technique in terms of computational cost.

5. Experiments

Based on the above developments, we conduct experiments in simulated and practical off-axis configurations. The holographic reconstructions are performed with our method as well as

with the conventional approach consisting in demodulation followed by back-propagation. Besides the standard reconstruction paradigm, we investigate cases where the amount of available measurements is reduced through random subsampling. In the sequel, a first set of experiments on simulated holograms evaluates the algorithms quantitatively in each case, using the oracle complex fields for comparison. The experiments that follow allow to determine the relevance and the competitiveness of our approach when real data are involved.

The standard and proposed algorithms are implemented in MATLAB, using the angular-spectrum approximation to model light propagation [20]. The reconstructed objects are shown with the same resolution and field of view as those of the measured holograms. In practical configurations, an additive linear-phase field can appear in the reconstruction when tilt coming from sample positioning occurs between acquisition of the RCH and the one of the object. This additive field is removed to simplify the visual interpretation of the results. Note that our algorithm is set to regularize the reconstructions on a slightly larger field of view than the one displayed in order to minimize the effects of border conditions during digital propagation. As shown in the sequel, this improves the quality of reconstruction at the boundaries compared to the standard technique that uses apodization.

5.1. Synthetic data

In this set of experiments, the original complex fields are available, their amplitudes and phases at the focus plane being defined from the two-dimensional spatial maps of Table 1. The *Pentagon* picture originates from the BM3D database at http://www.cs.tut.fi/ foi/GCF-BM3D/ and the others from the USC-SIPI database at http://sipi.usc.edu/database/. After normalization of its spatial-domain amplitude and phase values to $[0,1]$ and $[0,\pi]$, respectively, each complex field is used to generate a distinct 512×512 hologram according to our forward model. These holograms were first obtained from larger objects and then restricted to a limited field of view—centered and with 512×512 samples—in analogy with a physical setup. The acquisition parameters that are used for hologram generation correspond to an off-axis configuration, choosing the wavelength as corresponding to the He-Ne laser line, *i.e.*, $\lambda = 633 \cdot 10^{-9}$, and choosing $\Delta_x = 6.45 \cdot 10^{-6}$, $N = 1,024$, NA $= 0.25$, $n_i = 1$, $M = 10$, $d = 4 \cdot 10^{-2}$, and $\tilde{r}(\mathbf{x}) = \exp(2\pi\mathrm{j}(k_1 + k_2)/5\Delta_x)$. The oracle maps mentioned in Table 1 are shown along with the corresponding holograms in Fig. 1. Note that the scales at which our object maps, holograms, and reconstructions are displayed are normalized in each figure for convenience.

Given the synthetic holograms and their corresponding parameters, our goal is to determine how accurately the oracle complex fields can be reconstructed by the standard and proposed methods. For each hologram, we consider the classical setting where all samples are available for reconstruction, as well as the setting $D = 2$ where only 50% of the data are kept through random downsampling. As reconstruction parameters, our algorithm uses $v = 10^{-10}$, $\tau_{\mathrm{reg}} = 1.5 \cdot 10^{-2}$, $\varepsilon_{\mathrm{SNR}} = 35$, and $V = 2,000$ in all synthetic experiments. The regularization weight is set to $\gamma = 0.5$. In order to neglect the possible influence of boundary artifacts, the SNR values are evaluated on fields of view that are centered and whose size is reduced by 200 pixels in each dimension compared to those of the corresponding holograms.

The numerical results that are obtained are reported in Tables 1 and 2. As an example, the reconstruction is shown for the phase-only hologram #5 in Fig. 2; the phase values $[0,\pi]$ are mapped to the grayscale range $[\mathrm{black}, \mathrm{white}]$. In the classical paradigm where all samples are available, the values of Table 1 indicate an overall improvement in terms of reconstruction quality when using our approach. Our visual results also demonstrate that less artifacts are produced by the proposed method at the object boundaries. Furthermore, besides backward compatibility with the standard approach, our reconstruction algorithm is able to successfully recover object profiles when the amount of sampled data is reduced. In particular, the corresponding SNR
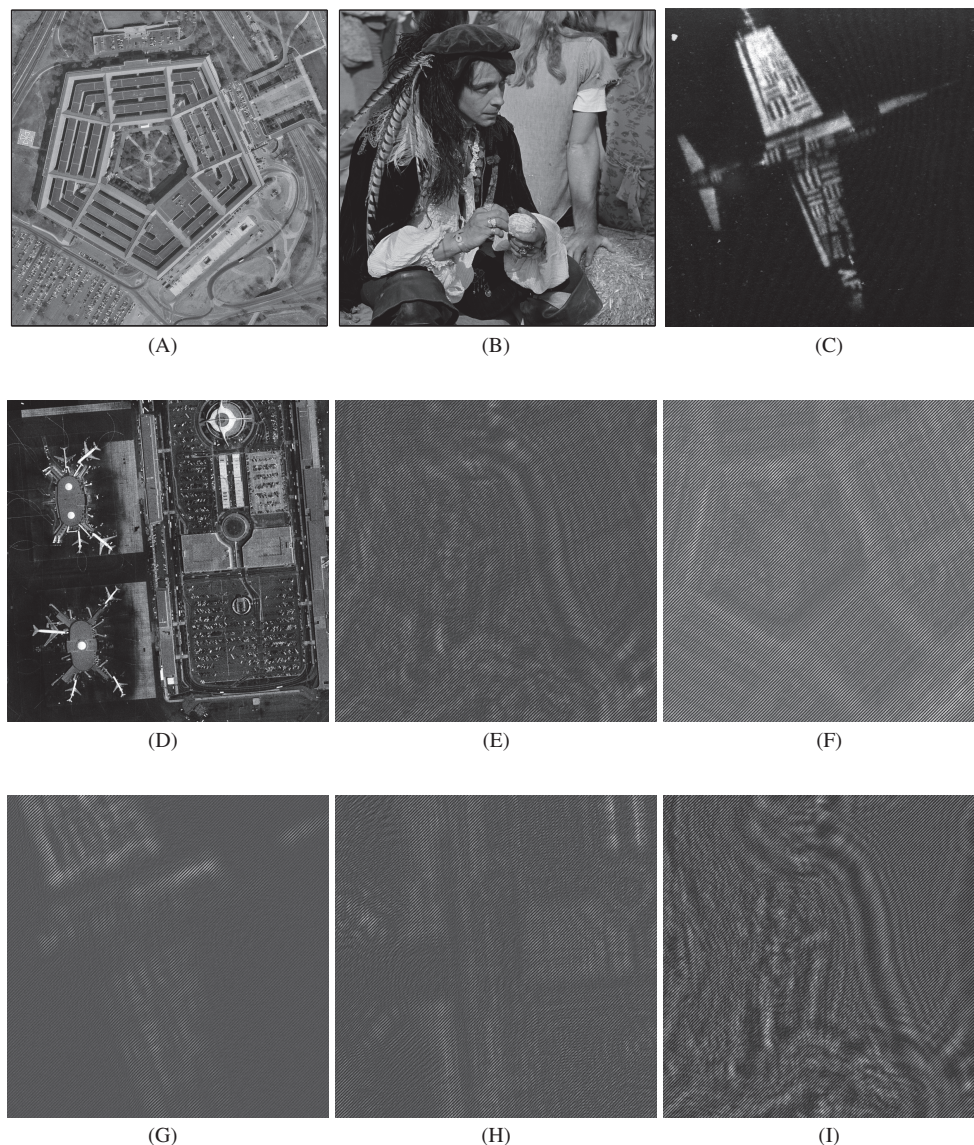
Fig. 1. Full-sized oracle maps and corresponding hologram acquisitions used in the synthetic experiments. From (A) to (D): *Pentagon*, *Man*, *Airplane*, and *Airport* maps defining the objects of Table 1. From (E) to (I): intensity holograms of size 512×512 associated with the objects no. 1, 2, 3, 4, and 5.

values of Table 1 remain stable as compared to the fully sampled configurations. By contrast, the SNR values that correspond to the standard reconstruction approach decrease dramatically when downsampling is used. The strong aliasing effects that are due to downsampling can indeed not be handled properly in that case since no regularization is used.

In Table 2, the proposed *field-of-view (FOV) ratio* measures the relative area of the solution where the reconstruction quality remains highest with respect to the oracle, considering the downsampled case only for our method. This area is determined through local MSE computa-

tions performed at every distance from the boundaries of the given object. It is obtained as the centered region where this local MSE does not exceed twice the global MSE. The FOV results further document the improvement of our method over the standard one. Our algorithm yields data-dependent values because its regularization strategy is itself data-adaptive.

At this stage, these observations confirm the relevance of our approach for holographic reconstruction in both classical and downsampled configurations. They remain to be corroborated in the real-data experiments that follow.

Table 1. Reconstruction Quality in Synthetic Experiments

			SNR for Amplitude Map / SNR for Phase Map			
	Original Object		$D = 1$		$D = 2$	
#	Amplit.	Phases	Standard	Variational	Standard	Variational
1	*Pentagon*	*Man*	21.46 / 21.25	**24.20 / 22.38**	6.60 / 1.55	**23.47 / 21.84**
2	*Pentagon*	-	24.08 / -	**24.53** / -	6.64 / -	**23.91** / -
3	*Airplane*	*Airplane*	**18.29** / 17.71	18.07 / **17.75**	4.61 / −3.77	**17.77 / 17.37**
4	*Airport*	*Airport*	19.14 / 18.66	**20.23 / 20.17**	5.87 / 0.46	**19.59 / 19.75**
5	-	*Man*	- / 20.74	- / **22.98**	- / 2.51	- / **22.54**

Table 2. Effective Field of View in Synthetic Experiments

			FOV Ratio		
	Original Object		$D = 1$		$D = 2$
#	Amplit.	Phases	Standard	Variational	Variational
1	*Pentagon*	*Man*	0.84	**0.88**	**0.87**
2	*Pentagon*	-	0.84	**0.86**	**0.86**
3	*Airplane*	*Airplane*	0.84	**0.95**	**0.95**
4	*Airport*	*Airport*	0.84	**0.86**	**0.86**
5	-	*Man*	0.84	**0.90**	**0.89**

5.2. Real data

In this second experimental part, we consider holograms that are acquired practically from distinct physical objects. The $1{,}024 \times 1{,}024$ hologram *Neuron* and the 512×512 hologram *Epithelial* consist in acquisitions of neural and epithelial cell samples, respectively. The 512×512 holograms *USAF 5-4* and *USAF 9-8* correspond to USAF targets with mixed phase and amplitude information. As in the synthetic case of Sect. 5.1, the acquisition settings that have been used in the optical setup are off-axis. The parameters that are common to all acquisitions are $\Delta_x = 6.45 \cdot 10^{-6}$ and $n_i = 1$. The other scalar parameters are reported in Table 3, while the approach followed to generate the reference waves is described in [14].

The proposed algorithm is parameterized as above, except $\tau_{\text{reg}} = 3 \cdot 10^{-4}$, $\gamma = 1$ for the hologram *USAF 9-8*, and except ε_{SNR}. In this practical setting, the latter quantity depends nontrivially on the acquisition conditions and is determined experimentally. The values of ε_{SNR} specific to each hologram are thus reported in Table 3 along with the acquisition parameters. Moreover, since the amplitude of the reference wave tends to vary slowly in space due to the non-ideality of the holographic device, we propose to determine α in Eq. (5) in a spatially adaptive manner.

(A) Standard algorithm with $D = 1$ (B) Proposed algorithm with $D = 1$

(C) Standard algorithm with $D = 2$ (D) Proposed algorithm with $D = 2$

Fig. 2. Reconstruction of the phase-only *Man* object from the synthetic hologram #5 without downsampling ($D = 1$) and with downsampling ($D = 2$).

Specifically, we perform one distinct optimization for each 8×8-pixel block of the solution. Note that each optimized value of α is still estimated based on the principles of Appendix D, although the sums involved in the computations are now restricted to the corresponding block.

The reconstructions from *Neuron*, *Epithelial*, and *USAF 5-4* are shown in Figs. 3, 4, 5, and 6. While the standard technique only handles the fully sampled scenario, distinct downsampling factors $D = 1, 2, 4$ are considered with our algorithm. When all samples are available, the visual results show similar advantages of our algorithm over the classical method as in the synthetic case. For instance, the field of view of the reconstructed objects is extended when using our iterative approach, unlike in the standard approach where the signal must be attenuated at borders in order to properly handle the boundary conditions. These results thus demonstrate that the boundary issues arising in digital propagation can be suitably addressed when using an im-

Table 3. Object-Dependent Parameters Used for Optical Acquisition

Hologram		Acquisition					Algorithm
Name	Type	λ	N	NA	M	d	ε_{SNR}
Neuron	*Phase-only*	$680 \cdot 10^{-9}$	1024	0.45	20	$3 \cdot 10^{-2}$	25.0 dB
Epithelial	*Phase-only*	$680 \cdot 10^{-9}$	512	0.25	10	$4 \cdot 10^{-2}$	32.0 dB
USAF 5-4	*Mixed*	$661 \cdot 10^{-9}$	512	0.4	20	$5 \cdot 10^{-2}$	30.0 dB
USAF 9-8	*Mixed*	$661 \cdot 10^{-9}$	512	0.4	20	$6 \cdot 10^{-2}$	18.5 dB

plicit formulation of the reconstruction problem such as Eq. (11). In this set of real experiments, our method is also observed to reduce the amount of noise in the recovered phase and amplitude profiles. In particular, the continuity of their background is improved. When increasing D, the reconstruction quality remains acceptable even though some fine details are inevitably lost due to the associated decrease in resolution. This decrease is associated to small patterns of "pointillism" type appearing in the holograms.

The contrast and the sharpness of object features are also well preserved when using our algorithm. In particular, the results on *USAF 5-4* shown in Figs. 5 and 6 demonstrate that sharp transitions and large regular structures are reconstructed accurately. The resolution performance of the standard and proposed methods can be compared in Figure 7 where the reconstruction from *USAF 9-8* is shown in the fully sampled setting. In this example, the resolution capability of both methods is similar. While the use of TV regularization produces a slight smearing effect in some places, the level of noise is strongly reduced in the reconstruction obtained with our method. This set of practical experiments thus corroborates the observations made in the synthetic case, and validates our approach as applied to real off-axis holograms subject to noise.

6. Conclusions

We have devised an algorithm for off-axis holographic reconstruction that is based on a consistent problem formulation. Based on suitable regularity assumptions, our technique has allowed to reconstruct complex-valued object profiles satisfactorily, including in the case where the amount of measured samples is decreased. Compared to the standard technique, the proposed method avoids the presence of boundary artifacts in the solutions, and reduces the perceived level of noise in practical configurations. From a general perspective, the obtained results further illustrate the interest of inverse-problem approaches in holography.

A. Definition of $\mathcal{H}$ and $\mathcal{P}_d$

The operator $\mathcal{H}$ corresponds to an amplitude filter defined in the Fourier domain as

$$\hat{H}(\boldsymbol{\omega}) = \begin{cases} 1, & \|\boldsymbol{\omega}\| \leq \Delta\omega_o/2 \\ 0, & \|\boldsymbol{\omega}\| > \Delta\omega_o/2. \end{cases} \tag{15}$$

We remind that this frequency limitation fundamentally comes from the NA of the objective. Given the distance d and the wavenumber $k = 2\pi/\lambda$, $\mathcal{P}_d$ corresponds to a phase filter in the Fourier domain that is defined as

$$\hat{P}_d(\boldsymbol{\omega}) = \exp\left(ikd \left(1 - \lambda^2 \|\boldsymbol{\omega}\|^2 \right)^{1/2} \right). \tag{16}$$

This propagation model is based on the angular-spectrum method [20].

(A) Standard algorithm with $D = 1$

(B) Proposed algorithm with $D = 1$

(C) Proposed algorithm with $D = 2$
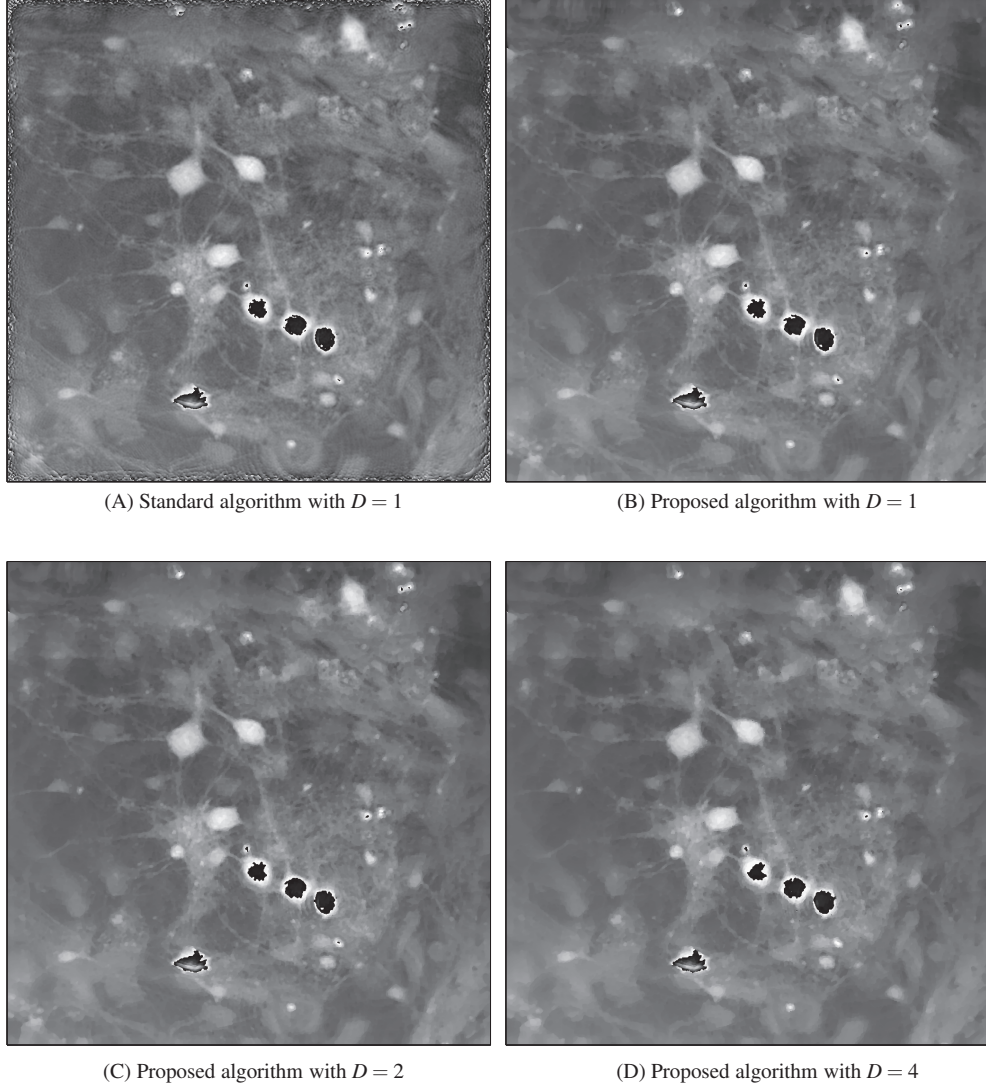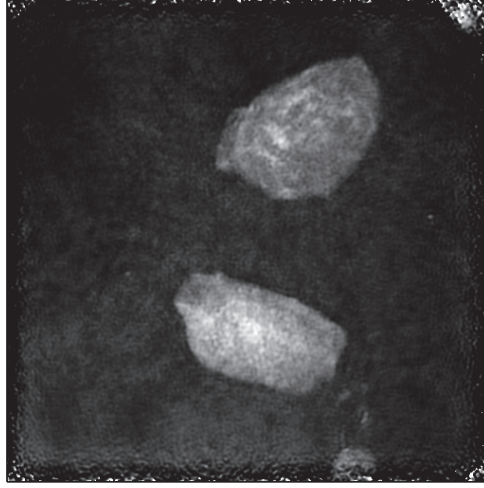
(D) Proposed algorithm with $D = 4$

Fig. 3. Reconstruction from the phase-only hologram *Neuron*.

B. Data-fidelity constant

As mentioned in Sect. 3, we constrain suitable estimates of the unknown complex field to be consistent with the measured intensity samples. According to our forward model, a given solution Ψ_d should thus yield intensities of the form

$$\tilde{I}(\mathbf{x}) = |(\mathscr{P}_d \mathscr{H} \Psi_d)(\mathbf{x}) + \alpha \bar{r}(\mathbf{x})|^2 \tag{17}$$
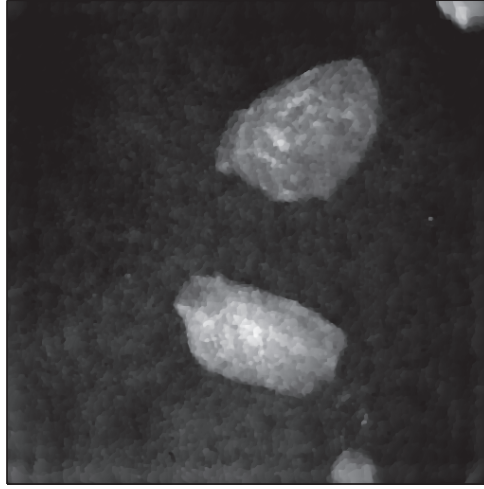
that are close to $I(\mathbf{x})$, $\forall \mathbf{x} \in \mathbf{m}\Delta_x$ with $w[\mathbf{m}] = 1$. Assuming that the noise level affecting the samples of I corresponds to a SNR of ε_{SNR} in dB, we constrain the similarity between $\tilde{I}$ and I to match this value as a lower limit. Taking only the non-masked samples into account, we obtain
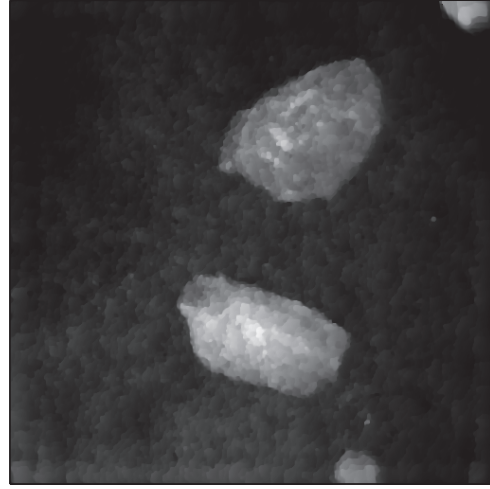
(A) Standard algorithm with $D = 1$ (B) Proposed algorithm with $D = 1$
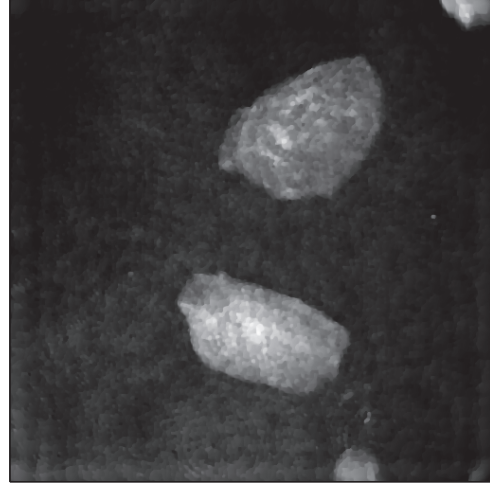
(C) Proposed algorithm with $D = 2$ (D) Proposed algorithm with $D = 4$

Fig. 4. Reconstruction from the phase-only hologram *Epithelial*.

$$\varepsilon_{SNR} < 10\log_{10}\left(\frac{\sum_{\mathbf{m}\in\mathbb{Z}^2} w[\mathbf{m}]I(\mathbf{x})^2|_{\mathbf{x}=\mathbf{m}\Delta_x}}{\sum_{\mathbf{m}\in\mathbb{Z}^2} w[\mathbf{m}](\tilde{I}(\mathbf{x})-I(\mathbf{x}))^2|_{\mathbf{x}=\mathbf{m}\Delta_x}}\right), \tag{18}$$

which expands as

$$\sum_{\mathbf{m}\in\mathbb{Z}^2} w[\mathbf{m}](\tilde{I}(\mathbf{x})-I(\mathbf{x}))^2|_{\mathbf{x}=\mathbf{m}\Delta_x} < \exp\left(-\varepsilon_{SNR}\ln(10)/10\right)\sum_{\mathbf{m}\in\mathbb{Z}^2} w[\mathbf{m}]I^2(\mathbf{x})|_{\mathbf{x}=\mathbf{m}\Delta_x}. \tag{19}$$

This expression is equivalent to Eq. (5). The left-hand side of Eq. (19) corresponds to the data term $\mathscr{D}(\Psi_d, \alpha)$, while its right-hand side is the definition of the data-fidelity constant ε in Eq. (6). Note that this derived result is specific to the quadratic form of Eq. (5) that is matched to

(A) Standard algorithm with $D = 1$

(B) Proposed algorithm with $D = 1$

(C) Proposed algorithm with $D = 2$

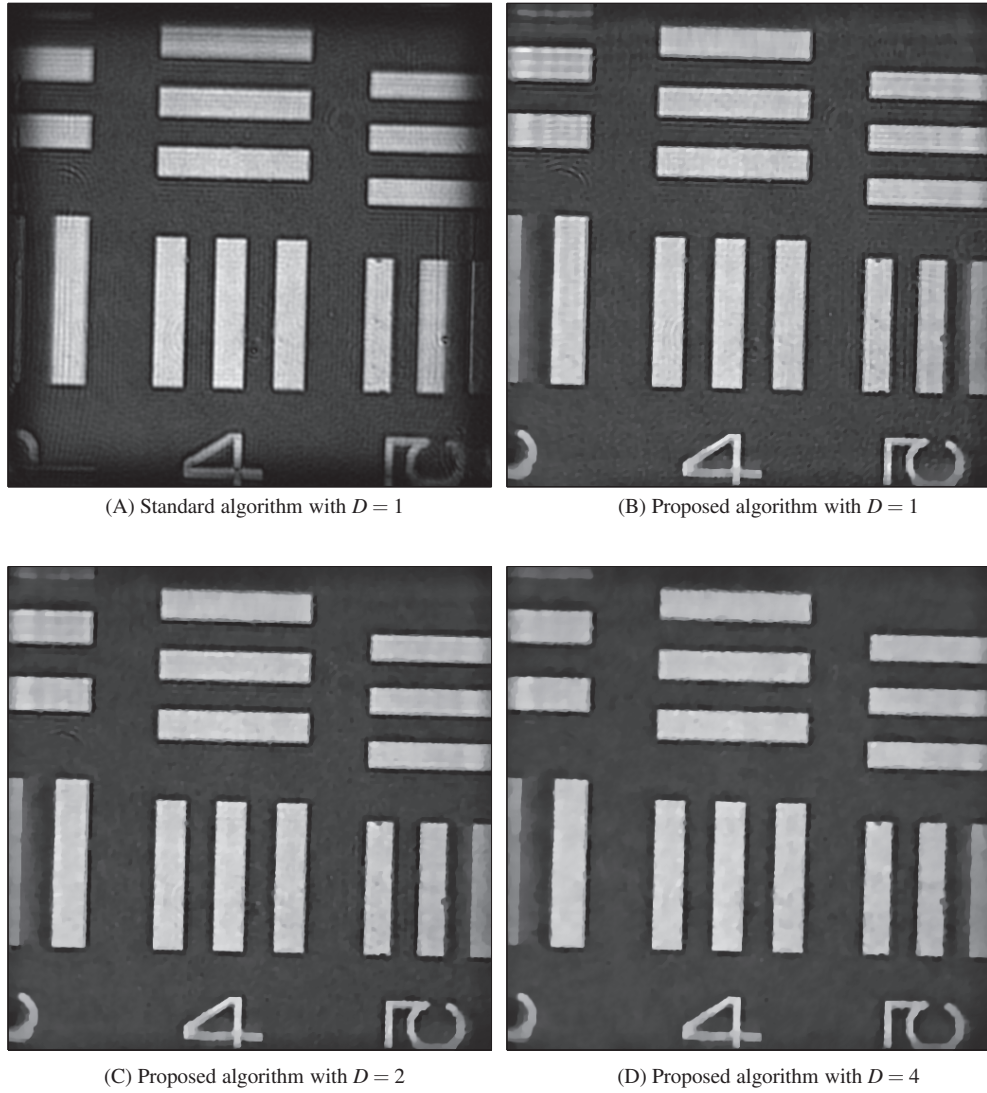(D) Proposed algorithm with $D = 4$

Fig. 5. Reconstructed amplitudes from the hologram *USAF 5-4*.

a Gaussian noise model. Another noise model (*e.g.*, Poisson) would correspond to a distinct formula for ε.

C.   Spatial derivatives of unwrapped phase

The phase derivative can be estimated independently of wrapping [21]. More specifically, since our finite-difference filters only consider two adjacent phases at the same time, only the relative phase differences are relevant. Assuming that the phase jumps cannot exceed π between two pixels, and defining $v_i = ((\mathbf{R}_{1,2}(\mathbf{arg}\Psi_d))_i \bmod 2\pi)$, we have the relation

$$(\mathbf{R}_{1,2}(\mathbf{arg}_u\Psi_d))_i = \begin{cases} v_i, & v_i < \pi \\ v_i - 2\pi, & v_i \geq \pi. \end{cases} \tag{20}$$

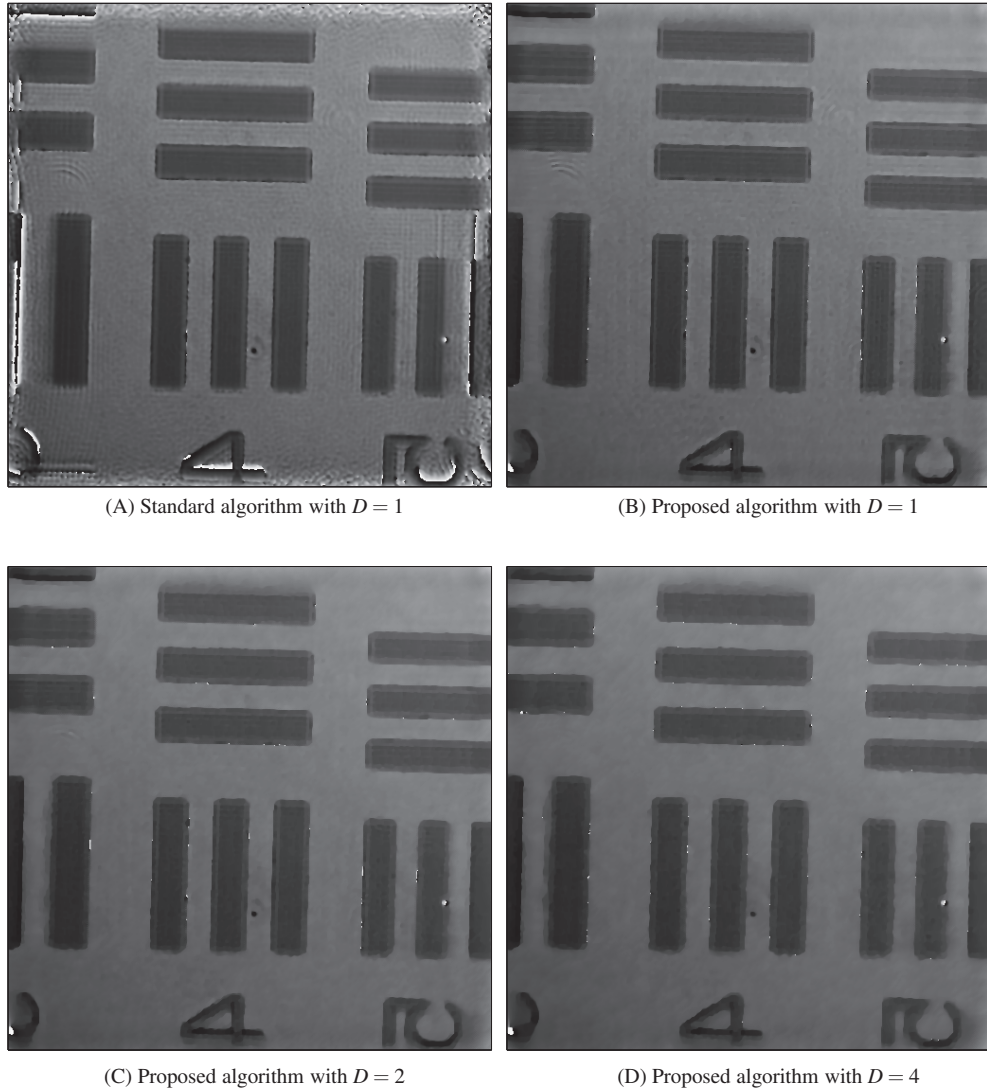|	
(A) Standard algorithm with $D = 1$	(B) Proposed algorithm with $D = 1$
(C) Proposed algorithm with $D = 2$	(D) Proposed algorithm with $D = 4$

Fig. 6. Reconstructed phases from the hologram *USAF 5-4*.

The above assumption on maximal phase jumps corresponds to the required properties for proper phase sampling [21]. In settings such as [22] where this assumption is violated, our method may still be applied, provided that only the solution amplitudes are regularized. While preliminary investigations indicate satisfactory results on objects whose phases are random and uniformly distributed in $[-\pi, \pi[$, this issue remains to be addressed in further research.

D. Line search

Each of the optimization problems that arise in our algorithm when determining $\tilde{\alpha}$ and $\tilde{\tau}$ can be recast as the resolution of a simple polynomial equation. According to Eq. (5), both data-term quantities appearing in Steps 2 and 3(b) are written in the generic form
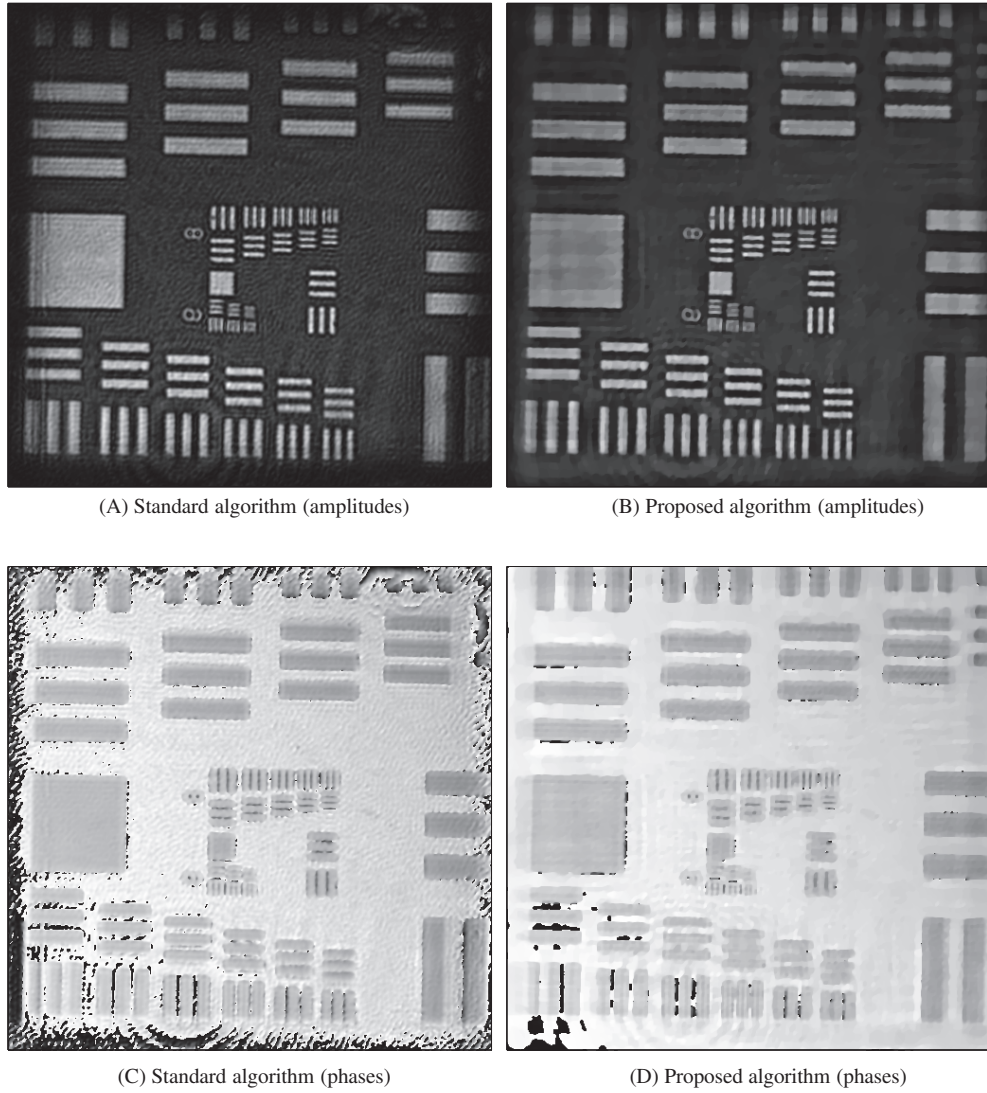
(A) Standard algorithm (amplitudes)   (B) Proposed algorithm (amplitudes)

(C) Standard algorithm (phases)   (D) Proposed algorithm (phases)

Fig. 7. Reconstructed amplitudes and phases from the fully sampled hologram *USAF 9-8*.

$$\mathscr{C}(\chi) = \sum_{\mathbf{m} \in \mathbb{Z}^2} \left(|a[\mathbf{m}] + \chi b[\mathbf{m}]|^2 - c[\mathbf{m}] \right)^2, \tag{21}$$

where a and b, and c are known constant sequences—the latter being real-valued and nonnegative—and where χ is the scalar variable of interest to optimize. Expanding the terms of Eq. (21), and factoring out χ, we obtain the simplified expression

$$\mathscr{C}(\chi) = A\chi^4 + B\chi^3 + C\chi^2 + D\chi + E, \tag{22}$$

where

$$
\begin{aligned}
A &= \sum_{\mathbf{m}\in\mathbb{Z}^2} |b[\mathbf{m}]|^4, \\
B &= \sum_{\mathbf{m}\in\mathbb{Z}^2} 4\mathrm{Re}\{a[\mathbf{m}]b^*[\mathbf{m}]\}|b[\mathbf{m}]|^2, \\
C &= \sum_{\mathbf{m}\in\mathbb{Z}^2} 2|b[\mathbf{m}]|^2(|a[\mathbf{m}]|^2 - c[\mathbf{m}]) + 4\left(\mathrm{Re}\{a[\mathbf{m}]b^*[\mathbf{m}]\}\right)^2, \\
D &= \sum_{\mathbf{m}\in\mathbb{Z}^2} 4\mathrm{Re}\{a[\mathbf{m}]b^*[\mathbf{m}]\}(|a[\mathbf{m}]|^2 - c[\mathbf{m}]), \\
E &= \sum_{\mathbf{m}\in\mathbb{Z}^2} (|a[\mathbf{m}]|^2 - c[\mathbf{m}])^2.
\end{aligned}
\tag{23}
$$

Given their definition, the coefficients $A, B, C, D,$ and E are real scalars that only depend on a, b, and c. Depending on the optimization problem, $\mathscr{C}$ has to be either equalized to ε, or minimized with respect to the scalar parameter. In the first case, the solution—if it exists—corresponds to the smallest real and nonnegative root χ_i of the fourth-degree polynomial

$$
A\chi^4 + B\chi^3 + C\chi^2 + D\chi + (E - \varepsilon).
\tag{24}
$$

In the second case, the argument minimizing $\mathscr{C}$ must cancel the derivative of $\mathscr{C}$ with respect to χ. Accordingly, the solution is the real and nonnegative root of the polynomial

$$
4A\chi^3 + 3B\chi^2 + 2C\chi + D
\tag{25}
$$

that yields the smallest value of $\mathscr{C}$.

Acknowledgments