Single-shot and lensless complex-amplitude imaging with incoherent light based on machine learning

Ryoichi Horisaki1,2 · Kazuki Fujii1 · Jun Tanida1

Received: 8 April 2018 / Accepted: 25 July 2018 / Published online: 31 July 2018 © The Author(s) 2018

Abstract

We present a single-shot incoherent light imaging method for simultaneously observing both amplitude and phase without any imaging optics, based on machine learning. In the proposed method, an object with a complex-amplitude field is illuminated with incoherent light and is captured by an image sensor with or without a coded aperture. The complex-amplitude field of the object is reconstructed from a single captured image using a state-of-the-art deep convolutional neural network, which is trained with a large number of input and output pairs. In experimental demonstrations, the proposed method was verified with a handwritten character database, and the effect of a coded aperture printed on an overhead projector film in the reconstruction was examined. Our method has advantages over conventional wavefront sensing techniques using incoherent light, namely simplification of the optical hardware and improved measurement speed. This study shows the importance and practical impact of machine learning techniques in various fields of optical sensing.

Keywords Phase retrieval · Digital holography · Incoherent holography · Coded aperture · Deep learning · Convolutional neural network

1 Introduction

An optical field is expressed as a complex-amplitude, which describes both the amplitude and phase of a light wave, in wave optics [1]. Conventional cameras just observe amplitude information only. Phase information is important for visualizing transparent objects, e.g., biological cells in biology, and compensating for aberrations/scattering with adaptive optics, e.g., deep tissue imaging in biomedicine and telescopic imaging through atmospheric turbulence in astronomy.

An established approach for complex-amplitude imaging, in which the transmittance/reflectance and phase delay of an object are observed, is digital holography (DH) [2–5]. DH basically assumes coherent illumination, such as a laser, but several incoherent DH methods have been proposed [6–10].

One issue with these incoherent DH methods is a tradeoff between the number of shots and the space–bandwidth product due to the use of a spatial or temporal reference carrier. Another approach for phase imaging with incoherent light is to use a Shack–Hartmann wavefront sensor [11]. Single-shot complex-amplitude/phase imaging methods based on the Shack–Hartmann wavefront sensor have utilized a lens array or a holographic optical element to observe a stereo or plenoptic image, and this results in a tradeoff between the spatial and angular resolutions [12–17]. Thus, these methods involve a compromise between the space–bandwidth product and the simplicity of the optical setup.

In this paper, to solve the above issues, we propose a method for single-shot complex-amplitude imaging with incoherent light and no imaging optics. We verify the effectiveness of a deep convolutional neural network as the reconstruction algorithm and the use of a coded aperture (CA) to improve the reconstruction fidelity. This approach extends the range of applications of complex-amplitude imaging and shows the importance of machine learning techniques in optical sensing.
2 Method

In the proposed method, a complex-amplitude object field illuminated with incoherent light is captured by an image sensor with or without a coded aperture (CA), as shown in Fig. 1. A CA has been used to improve the reconstruction conditions because its Fourier spectrum is significantly extended compared with that of a conventional non-coded aperture. Using CAs, single-shot amplitude imaging with incoherent light and single-shot reference-free DH with coherent light have been demonstrated [18–23].

The imaging process of the proposed method is written in general form as

\[ g = H[f] \]  
\[ = H[f, \exp(i\phi)\]. \]
where \( g \in \mathbb{R}^{N\times 1} \) is the captured intensity image and \( H[\cdot] \) is the forward operator. \( f \in \mathbb{C}^{N\times 1}, f_r \in \mathbb{R}^{N\times 1}, \) and \( f_\phi \in \mathbb{R}^{N\times 1} \) are the complex-amplitude, amplitude, and phase of the object field, respectively, and \( f = f_r, \exp(i\phi) \).

Here, \( N \) is the number of pixels along one spatial direction, \( \mathbb{R}^{P\times Q} \) is a \( P \times Q \) matrix of real numbers, and \( \mathbb{C}^{P\times Q} \) is a \( P \times Q \) matrix of complex numbers.

The inversions for reconstructing the amplitude and the phase are separable, as indicated by Eq. (2). In this paper, for simplicity, we solve the following two inversions independently:

\[ \hat{f}_r = H_r^{-1}[g], \]  
\[ \hat{f}_\phi = H_\phi^{-1}[g], \]
where \( H_r^{-1}[\cdot] \) and \( H_\phi^{-1}[\cdot] \) are the inverse operators for the amplitude and the phase, respectively.

We use a regression algorithm for the inversions of Eqs. (3) and (4) in the reconstruction process. Regressions, including deep learning, are typical techniques in machine learning, and have applied regressions to optical sensing and control [24–29]. In addition, several studies for phase retrieval, ghost imaging, and superresolution imaging with deep learning have been reported [30–32]. In this paper, we modify a deep convolutional residual network (ResNet [33]) for each of the inversions. Residual learning in ResNet prevents stagnation during the learning process and optimizes deep layers efficiently. ResNet has been used for phase retrieval and computer-generated holography and showed favorable results [29, 30].

The network architecture for the reconstruction is shown in Fig. 2. The network is composed of multi-scale ResNets, as shown in Fig. 2a. \( K \) is the number of filters at each of the convolutional layers. “D” is a downsampling block, as shown in Fig. 2b. “U” is an upsampling block, as shown in Fig. 2c. “R” is a residual convolutional block, as shown in Fig. 2d. “S” is a skip convolutional block, as shown in Fig. 2e. The definitions of the layers are as follows [34]: “\( \text{Conv}(S, L) \)” is a two-dimensional convolutional layer with a filter size \( S \) and a stride \( L \). “\( \text{TConv}(S, L) \)” is a two-dimensional transposed convolutional layer with a filter size \( S \) and a stride \( L \). “\( \text{BatchNorm} \)” is a batch normalization layer. “\( \text{ReLU} \)” is a rectified linear unit layer.

The network architectures for reconstructing the amplitude and phase are the same, as shown in Fig. 2a, but they are trained with different datasets. The amplitude datasets are composed of pairs of amplitude images \( f_r \) and their captured intensity images \( g \). On the other hand, the phase datasets are composed of pairs of phase images \( f_\phi \) and their captured intensity images \( g \).

3 Experimental demonstration

The proposed method was demonstrated experimentally. The complex-amplitude object was implemented with two spatial light modulators (SLMs: LC 2012 manufactured by Holoeye, pixel pitch: 36 µm, pixel count: \( 768 \times 1024 \)); one of them was operated in the phase mode, and the other one was operated in the amplitude mode. The phase and amplitude SLMs were located at positions 5 and 7 cm from an image sensor (PL-B953 manufactured by PixelLink, pixel pitch: 4.65 µm, pixel count: \( 768 \times 1024 \)), respectively. Those two SLMs were illuminated with an incoherent green light-emitting diode (M565L3 manufactured by Thorlabs, nominal wavelength: 565 nm, full width at half maximum of spectrum: 103 nm) without any spatial or spectral filter. Our method does not suppose amplitude and phase layers at different distances. A single complex-amplitude layer is readily trainable in the same manner by implementing such an object. In the case with the CA, the CA was printed on an overhead projector film (OHP film: VF-1421N manufactured by Kokuyo) by a laser printer (ApeosPort-V C2276 manufactured by Fuji Xerox, resolution: 1200 dpi) and it

![Fig. 1 Schematic diagram of single-shot, lensless complex-amplitude imaging with/without a CA](image-url)
was located at a position 1 cm from the image sensor. The CA was a random binary pattern with the maximum printable resolution.

The amplitude and phase images simultaneously displayed on the two SLM were different handwritten digits randomly selected from the EMNIST database, where the pixel count of the images was $28 \times 28$ [35]. The captured intensity images were reduced to $28 \times 28$ pixels, and then they were expanded to $32 \times 32(=N^2)$ with zero padding for adjusting the captured image size to the input one of the network, where the input and output image sizes are a power of two. These resized captured intensity images and the original amplitude and phase images, which were also expanded to $32 \times 32(=N^2)$ with zero padding, were provided to the network for reconstructing the amplitude and phase images with zero padding. The output images from the network were cropped to remove the zero padding area, and the final amplitude and phase images with a pixel count of $28 \times 28$ were produced. The number of convolutional filters was $64(=K)$. An algorithm called “Adam” was used to optimize the network with an initial learning rate of 0.0001 for the step width in the optimization, a mini-batch size of 50, and a maximum number of epochs of 50 [36]. The number of training pairs was 100,000 in both the amplitude and phase reconstructions.

The experimental results without the CA are shown in Figs. 3 and 4. The trained amplitude and phase networks were respectively tested with 1000 amplitude images and 1000 phase images that had not been used for the training process. Ten examples of the amplitude and phase test images are shown in Fig. 3a and b, respectively. Their captured intensity images from the amplitude and phase images in Fig. 3 are shown in Fig. 4a. The amplitude and phase images reconstructed from Fig. 4a are shown in Fig. 4b and c. The peak signal-to-noise ratios (PSNRs) of the 1000 reconstructed amplitude images and 1000 reconstructed phase images were 20.9 and 14.3 dB, respectively. These results demonstrated the effectiveness of
the deep network for complex-amplitude imaging with incoherent light.

The experimental results with the CA are shown in Fig. 5. The network was trained and tested with the same amplitude and phase images as the previous experiment but with different captured intensity images through the CA. The captured intensity images with the CA from the amplitude and phase images in Fig. 3 are shown in Fig. 5a. The amplitude and phase images were also reconstructed from the captured intensity images, as shown in Fig. 5b and c. The PSNRs of the amplitude and phase images were 23.5 and 15.3 dB, respectively. These results showed that the amplitude and phase reconstructions were improved in the case with the CA.

In the experiments, the calculation time for reconstructing the test images was 0.5 s when using a computer with an Intel Dual Xeon processor with a clock rate of 2.1 GHz and a memory size of 128 GB, and NVIDIA Quadro P4000 graphics processing unit with a memory size of 8 GB. The training time was 18 h.

4 Conclusion

We presented a method for single-shot, lensless complex-amplitude imaging with incoherent light and no imaging optics. A deep convolutional neural network was used for reconstructing both the amplitude and phase from a single intensity image. We demonstrated the proposed method experimentally with a CA implemented on an OHP film using handwritten datasets. The results showed the effectiveness of the deep network and the CA for complex-amplitude imaging with incoherent light.

Our method reduces the hardware size/cost and the measurement time in incoherent complex-amplitude imaging, while maintaining the image quality. It is readily extendable to a reflective setup and multi-dimensional complex-amplitude imaging, such as three-dimensional spatial imaging and spectral imaging. The CA is not limited to only amplitude modulation but can also be applied to phase or complex-amplitude modulation [22]. Our method is applicable to non-visible spectral regions, such
as X-rays, infrared light, and terahertz radiation. Therefore, the method is promising in various fields, including, biology, security, and astronomy.

**Funding** This work was supported by JSPS KAKENHI Grant Numbers JP17H02799 and JP17K00233, and JST PRESTO Grant Number JPMJPR17PB.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**References**

1. Goodman, J.W.: Introduction to Fourier Optics. McGraw-Hill, New York (1996)
2. Nehmetallah, G., Banerjee, P.P.: Applications of digital and analog holography in three-dimensional imaging. Adv. Opt. Photonics 4, 472–553 (2012)
3. Osten, W., Faridian, A., Gao, P., Körner, K., Naik, D., Pedrini, G., Singh, A.K., Takeda, M., Wilke, M.: Recent advances in digital holography. Appl. Opt. 53, G44–G63 (2014)
4. Bhaduri, B., Edwards, C., Pham, H., Zhou, R., Nguyen, T.H., Goddard, L.L., Popescu, G.: Diffraction phase microscopy: principles and applications in materials and life sciences. Adv. Opt. Photonics 6, 57–119 (2014)
5. Memmolo, P., Miccio, L., Paturzo, M., Caprio, G.D., Coppola, G., Netti, P.A., Ferraro, P.: Recent advances in holographic 3D particle tracking. Adv. Opt. Photonics 7, 713–755 (2015)
6. Rosen, J., Brooker, G.: Non-scanning motionless fluorescence three-dimensional holographic microscopy. Nat. Photonics 2, 190–195 (2008)
7. Rosen, J., Brooker, G.: Fresnel incoherent correlation holography (FINCH): a review of research. Adv. Opt. Technol. 1, 151–169 (2012)
8. Vrijayakmar, A., Kashter, Y., Kelner, R., Rosen, J.: Coded aperture correlation holography new type of incoherent digital holograms. Opt. Express 24, 12430–12441 (2016)
9. Quan, X., Matoba, O., Awatsuji, Y.: Single-shot incoherent digital holography using a dual-focusing lens with diffraction gratings. Opt. Lett. 42, 383–386 (2017)
10. Liu, J.-P., Tahara, T., Hayasaki, Y., Poon, T.-C.: Incoherent digital holography: a review. Appl. Sci. 8, 143 (2018)
11. Platt, B.C., Shack, R.: History and principles of Shack–Hartmann wavefront sensing. J. Refract. Surg. 17, S573–7 (1995)
12. Parhasarathy, A.B., Chu, K.K., Ford, T.N., Merz, J.: Quantitative phase imaging using a partitioned detection aperture. Opt. Lett. 37, 4062–4064 (2012)
13. Saita, Y., Shinto, H., Nomura, T.: Holographic Shack–Hartmann wavefront sensor based on the correlation peak displacement detection method for wavefront sensing with large dynamic range. Optica 2, 411–415 (2015)
14. Gong, H., Agbana, T.E., Pozzi, P., Soloviev, O., Verhaegen, M., Vdovin, G.: Optical path difference microscopy with a Shack–Hartmann wavefront sensor. Opt. Lett. 42, 2122–2125 (2017)