

# Machine-learning-based high-speed lensless large-field holographic projection using double-sampling Fresnel diffraction method

Chentianfei Shen (沈陈天飞), Tong Shen (申桐), Qi Chen (陈祺), Qinghan Zhang (张馨瀚), and Jihong Zheng (郑继红)\*  
School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200433, China

\*Corresponding author: [jihongzheng@usst.edu.cn](mailto:jihongzheng@usst.edu.cn)

Received December 14, 2021 | Accepted March 1, 2022 | Posted Online March 25, 2022

Machine learning can effectively accelerate the runtime of a computer-generated hologram. However, the angular spectrum method and single fast Fresnel transform-based machine learning acceleration algorithms are still limited in the field-of-view angle of projection. In this paper, we propose an efficient method for the fast generation of large field-of-view holograms combining stochastic gradient descent (SGD), neural networks, and double-sampling Fresnel diffraction (DSFD). Compared with the traditional Gerchberg-Saxton (GS) algorithm, the DSFD-SGD algorithm has better reconstruction quality. Our neural network can be automatically trained in an unsupervised manner with a training set of target images without labels, and its combination with the DSFD can improve the optimization speed significantly. The proposed DSFD-Net method can generate 2000-resolution holograms in 0.05 s. The feasibility of the proposed method is demonstrated with simulations and experiments.

**Keywords:** computer-generated hologram; holographic display; machine learning.

**DOI:** [10.3788/COL202220.050502](https://doi.org/10.3788/COL202220.050502)

## 1. Introduction

Focus cues, image quality, field of view (FOV), and eye box are key issues in near-eye displays such as virtual reality and augmented reality<sup>[1,2]</sup>. Various optical schemes are adopted and applied in near-eye display realization to design more portable and compact near-eye display devices, such as coaxial prisms, planar displays based on transparent film arrays, freeform surfaces, multiplayer displays with directional backlighting, and stereoscopic see-through retinal projection<sup>[1-4]</sup>. The holographic near-eye display has achieved great progress in recent years<sup>[4-10]</sup>. The holographic near-eye display is a competitive way to realize near-eye displays because of its ability to meet compact structures and to reproduce three-dimensional (3D) images for a realistic and comfortable viewing experience<sup>[11]</sup>.

The Nyquist criterion limits the size of the reconstructed image when employing a standard diffraction method in computational holography, either the Fresnel diffraction algorithm or the Fraunhofer diffraction algorithm<sup>[12]</sup>. In recent years, researchers have tried to investigate computational holographic diffraction algorithms for large FOVs, such as the curved spatial light modulator (SLM) array method<sup>[13]</sup>, curved hologram method<sup>[14]</sup>, and time-multiplexing method<sup>[15]</sup>. By using spherical beam lighting, Chang *et al.*<sup>[16]</sup> and Qu *et al.*<sup>[17]</sup> suggested an

image amplified lensless holographic projection. However, the evaluation of their computer-generated hologram (CGH) requires a sophisticated iterative algorithm that is unfortunately time-consuming. In each iteration of the calculation, several Fourier transforms and inverse Fourier transforms have to be implemented. As the resolution of the image increases, the computation time increases accordingly. In order to accelerate the computing speed, machine learning technology has been used in the field of optical information processing<sup>[18-20]</sup>. Peng *et al.*<sup>[7]</sup> and Wu *et al.*<sup>[8]</sup> proposed to compute the CGH by machine learning methods. With the help of machine learning methods, the time for calculating the CGH is reduced to less than 0.2 s. However, their networks can only encode the CGH based on the single fast Fourier transform (S-FFT) and the angular spectrum diffraction algorithms and thus cannot obtain large FOV angles. Namely, it is not large enough for binocular observation for near-eye display.

To overcome the aforementioned FOV and time-consuming issues, we present a method combining the machine-learning-based technique with the lensless holographic projection. This method is based on the double-sampling Fresnel diffraction (DSFD) algorithm<sup>[17]</sup> and machine learning method for calculating the hologram with a phase-only SLM. Both the simulation and experiments are performed, and the evidence demonstrates

that the proposed method is able to provide an enlarged and qualified holographic image. The hologram can be real-time generated with high resolution. Therefore, the proposed method may hopefully find its application in near-eye display devices.

## 2. Methods

In a holographic display system, finding the phase value  $\phi$  on the SLM plane that best approximates the target image can be formulated as solving an optimization problem of the form

$$\hat{\phi} = \operatorname{argmin}\{L[s \cdot f(\phi), Y]\}, \quad (1)$$

where the function  $f(\phi)$  describes the light propagation,  $Y$  is the amplitude of the target image,  $L$  represents the loss function, and  $s$  is the scale factor of the system. The optical system not only determines the imaging quality of the projection but also is one of the most critical factors that affects the numerical calculation of diffraction propagation.

We discuss an intuitive reason for the use of the DSFD algorithm as the light propagation function. In a near-eye display system, a large FOV is a vital factor and is determined from the holographic image projection. We evaluate several lensless light propagation algorithms, including the S-FFT, the angular spectrum method (ASM), the DSFD, and the three-step diffraction<sup>[21]</sup>. Among these methods, both the S-FFT and ASM algorithms employ a plane wave as the light source, and hence the size of output images is limited. Although the three-step diffraction can obtain the largest field in the image reconstruction, the zeroth-order output and the first-order output overlap each other<sup>[21]</sup>. Therefore, we use the dispersive spherical wave as the light source of the system and implement the DSFD for calculating the light propagation. A schematic illustration of the holographic image projections with plane wave illumination and a diverging light source is presented in Fig. 1.

Figure 1(a) shows the scheme of CGH calculation for the S-FFT algorithm. The complex amplitude distribution  $U_i(x_i)$  on the image plane is given by

$$U_i(x_i) = \exp\left(-\frac{ikx_i^2}{2z}\right) F\left[U_{\text{SLM}}(x_{\text{SLM}}) \exp\left(-\frac{ikx_{\text{SLM}}^2}{2z}\right)\right], \quad (2)$$

in which  $U_{\text{SLM}}(x_{\text{SLM}})$  represents the complex amplitude of the SLM, and  $z$  is the distance between the CGH plane and the image plane.

The image size of S-FFT algorithm  $L_i$  is thus given by

$$L_i = \lambda z / \Delta x_{\text{SLM}}. \quad (3)$$

Here,  $\lambda$  is the wavelength, and  $\Delta x_{\text{SLM}}$  presents the pixel size of the SLM.

As depicted in Fig. 1(b), the diverging spherical wave propagates through two planes, the first plane is the hologram plane, and the second one is the image plane. According to the Fourier optics, the propagation of the light wave can be regarded as two

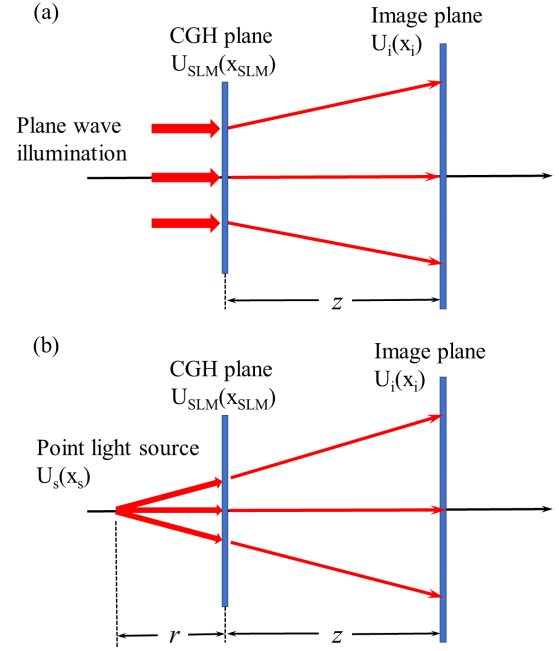


Fig. 1. Principle of S-FFT and DSFD algorithms for lensless holographic projection. (a) S-FFT algorithm with plane wave illumination; the maximum projecting image size is limited by the Nyquist criterion. (b) DSFD algorithm with diverging point light source; the image size is larger.

steps. In the first step, since the SLM is illuminated by a diverging light wave, the procedure from the hologram plane to the point light source can be taken as the inverse Fraunhofer diffraction. The complex amplitude distribution  $U_s$  on the source plane is given by

$$U_s(x_s) = \exp\left(\frac{ikx_s^2}{-2r}\right) F^{-1}[U_{\text{SLM}}(x_{\text{SLM}})], \quad (4)$$

where  $F^{-1}$  represents the inverse Fourier transform, and  $k = 2\pi/\lambda$  is the wave number. In the second step, the procedure from the point light source to the image plane can be regarded as the Fresnel diffraction. The complex amplitude distribution  $U_i$  on the image plane is thus expressed as

$$U_i(x_i) = \exp\left[-\frac{ikx_i^2}{2(z+r)}\right] F\left\{U_s(x_s) \exp\left[-\frac{ikx_s^2}{2(z+r)}\right]\right\}. \quad (5)$$

Since the Fourier transform is used in the above diffraction procedures, the sampling interval should be calculated according to

$$\Delta x_s = \frac{\lambda r}{L_{\text{SLM}}}, \quad \Delta x_i = \frac{\lambda(r+z)}{N\Delta x_s}. \quad (6)$$

The maximal size of the diffraction image is determined by

$$L_i = \frac{z+r}{r} L_{\text{SLM}}, \quad (7)$$

where  $L_{SLM}$  is the size of the SLM itself, and  $r$  is the distance between the point source and the SLM.

The use of the diverging wave as the light source may enlarge the image sizes with a ratio depending on the relation of  $r$  and  $z$ . According to Eqs. (3) and (7), Fig. 2 shows the comparison between the image size for the S-FFT algorithm based on plane wave illumination and that of the DSFD algorithm based on a diverging point light source under the same diffraction distance. It can be clearly observed that the DSFD algorithm can be used to enlarge the image size.

The procedure for calculating the resulting light field on the image plane is shown in Eq. (5). Evidently, the complex amplitude representation on the SLM plane is required to display such holograms. Usually, SLMs are classified into phase-only, amplitude-only, and complex amplitude modulation types. Aside from availability, phase-only SLMs are often preferred because of their high light efficiency. Notably, light is only steered but not attenuated. Nevertheless, calculating holograms that function with phase-only SLMs is one of the main challenges in developing holographic displays. The common method to encode the complex amplitude CGH into phase-only CGH is an iterative phase optimization. The iterative Gerchberg–Saxton (GS) algorithm is the standard way to solve the problem of phase retrieval of a field on two separate planes, as shown in Fig. 3.

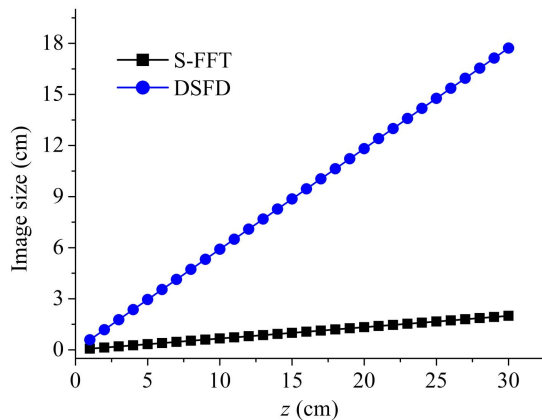


Fig. 2. Projection image size of the S-FFT algorithm and DSFD algorithm, where the wavelength is 532 nm, pixel size is 8  $\mu\text{m}$ , the number of pixels is 1920, and the distance between the point light source and the SLM plane is 2.6 cm.

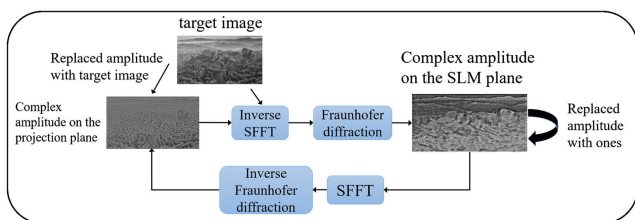


Fig. 3. GS algorithm workflow for computing a phase-only CGH from target image.

Unfortunately, the GS algorithm inevitably requires long computation time and leads to serious speckle noise in the image reconstruction.

In order to further improve the image quality and reduce the calculating time of the CGH, the method based on the combination of the DSFD algorithm and machine learning is proposed. Gradient descent is a way to minimize an objective function parameterized by a model’s parameters by updating the parameters in the opposite direction of the gradient of the objective function to the parameters. We first implement stochastic gradient descent (SGD) to optimize the loss function in Eq. (1); see Fig. 4. We give an initial random phase on the SLM plane and calculate the complex field on the image plane with the DSFD algorithm. Then, we calculate the loss between the target image and the simulated projection image. Finally, we backpropagate the error between the target images and simulate reconstruction with a stochastic descent optimization algorithm to update the phase-only holograms. Since the iterative procedure of the SGD does not perform the inverse calculation of the light propagation and only needs to calculate the diffraction once, time consumption is only half of the traditional GS method for each epoch. In addition, by adjusting the learning rate, the SGD algorithm converges much faster than the traditional GS method.

Despite the fact that the SGD algorithm can reduce the computation time of holograms by more than half, we still need a fast way to obtain the phase-only hologram. We then combine the above DSFD algorithm with a neural network to form our DSFD-Net model. The DSFD-Net model can be trained in an unsupervised learning method of the mapping from the target image to the hologram without labels. The generation and reconstruction of phase-only CGH can be depicted as the encoding and decoding process of target images. Our neural network works as the encoder part in the system and translates the target image to the corresponding phase-only CGH. The output of the network is the input of our decoder. The decoding part is the fixed DSFD model that has been described above. The architecture of our training procedure and the U-Net is shown in Fig. 5. As an unsupervised learning model, the data sets and validation sets do not need to be labeled.

The U-Net model uses a down-sampling and then up-sampling structure. The use of a skip connection at the same stage ensures that the final CGH output incorporates more low-level

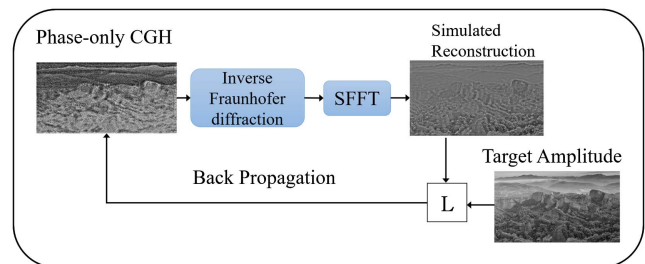
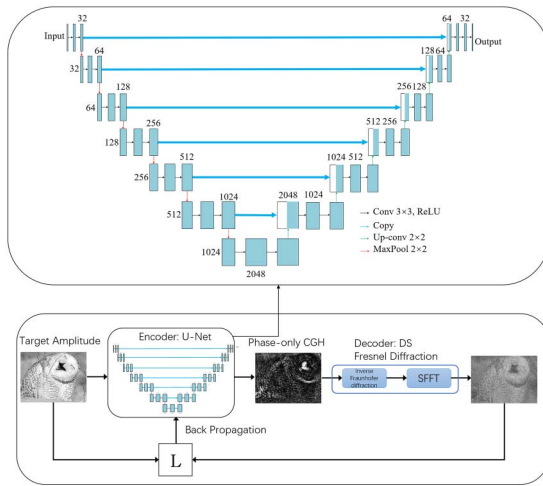


Fig. 4. SGD algorithm workflow for computing a phase-only CGH from target image.

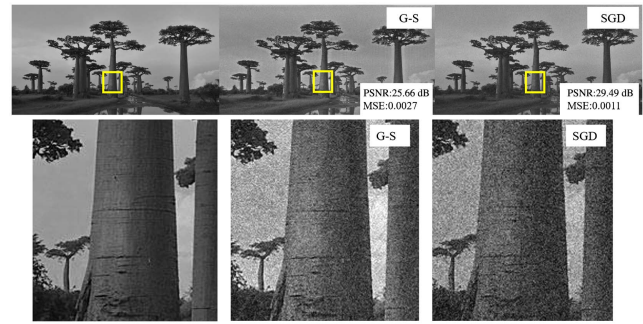


**Fig. 5.** Illustration of our wave propagation model. A target image is first converted to an amplitude value, and then is passed to a phase-encoder network (i.e., the U-Net). At the SLM plane, we display the CGH and propagate the light field to the target plane. During the training phase, the loss between the projection image and the target amplitude can be calculated and is then propagated back to train the phase-encoder network.

features and retains all the information in the image. This advantage is very suitable for CGH computation. The length and width of the image tensor are reduced by half after each down-sampling in our U-Net, and the geometric feature extraction of the input image is realized after down-sampling is repeated six times. When the up-sampling of next six times is implemented, the reconstructed original size image tensor is obtained. In order to avoid the disappearance of the gradient during the network training, the residual connection is employed to realize the cross-layer transfer of the gradient. After each convolution, batch normalization is performed to avoid overfitting. In the U-Net training procedure, we use the amplitude of the  $1920 \times 1080$  image as the training input. The U-Net outputs the corresponding CGH. We simulate the physical diffraction processing with the CGH generated by our U-Net.

### 3. Simulation

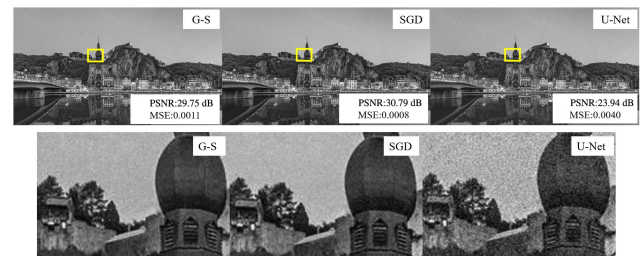
We simulate the light propagation on Google Colab with PyTorch, which is essentially based on python and compute unified device architecture (CUDA), to demonstrate the performance of different algorithms with graphics processing units (GPUs). The GPU is NVIDIA Tesla P100 with 16 GB memory. To keep consistent with the experimental situation, the pixel pitch of the CGH is set as  $8 \mu\text{m}$ , and the resolution is  $1920 \times 1080$ . The wavelength of the laser is 532 nm. The distance between the diverging point light source and the hologram is 2.6 cm, and the propagation distance is 26 cm. Figure 6 shows the simulated results of the SGD method and GS method. We use the mean square error (MSE) and peak signal-to-noise ratio (PSNR) to quantify the quality of the reconstructed images.



**Fig. 6.** Performance evaluation of the GS algorithm and the SGD algorithm.

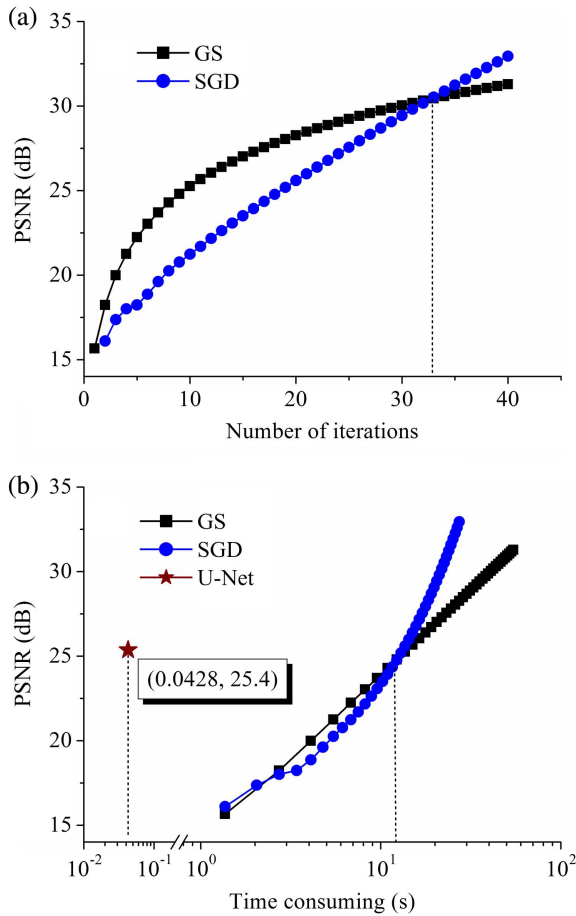
A gray-level image in Fig. 7 is employed to demonstrate the effectiveness of our U-Net. A comparison between the U-Net and iterative methods demonstrates that the proposed U-Net method produces the reconstructed images with acceptable quality. The PSNR is more than 23 dB.

The numerical reconstructions are presented in Fig. 8. We test 100 random images from the testing dataset, and Fig. 8(a) indicates that both the GS and SGD iterative algorithms can achieve a high quality ( $> 25$  dB) after sufficient iterations are performed (say 30 iterations). When the number of iterations is small and the time consumed is low, the PSNR of the GS method is higher than that of the SGD method. But, after the number of iterations exceeds 35, the situation is reversed, i.e., the PSNR of SGD is better than that of GS. We find that, for running the same iterations, the SGD method consumed half the time of the GS method, so the SGD method takes less than half the time compared to the traditional GS method to achieve high-quality reconstruction results (PSNR  $> 30$  dB). The SGD method is a better iterative method compared to the GS method when high-quality reconstruction of images is required. The results are achieved on the assumption that the wave propagation model used for optimizing the SLM phase pattern is the same for simulating the final image. The U-Net only takes an average of 0.05 s to generate  $1920 \times 1080$  holograms and achieve an average of 25 dB PSNR of the reconstruction image quality.



**Fig. 7.** Performance evaluation of our U-Net and the iterative methods. The PSNR and MSE values indicate the reconstruction image quality of the algorithm.



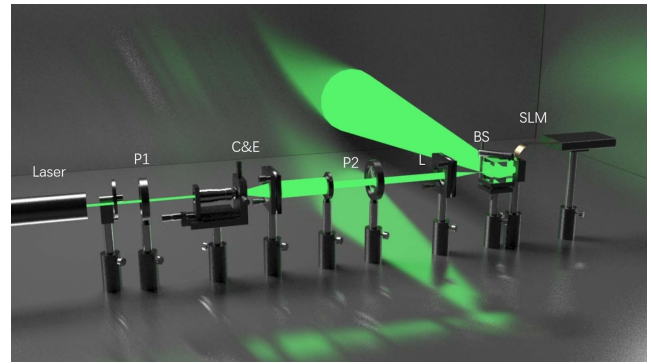


**Fig. 8.** Comparison of average calculating speed and image quality achieved by several CGH techniques. (a) Images are reconstructed with similar quality at the same number of iterations by GS and SGD algorithms. (b) The SGD algorithm requires less time than the GS algorithm for high-quality reconstruction. The U-Net takes less than 0.05 s, which is far less than that of iterative methods. The horizontal of Fig. 8(b) is in logarithmic scale.

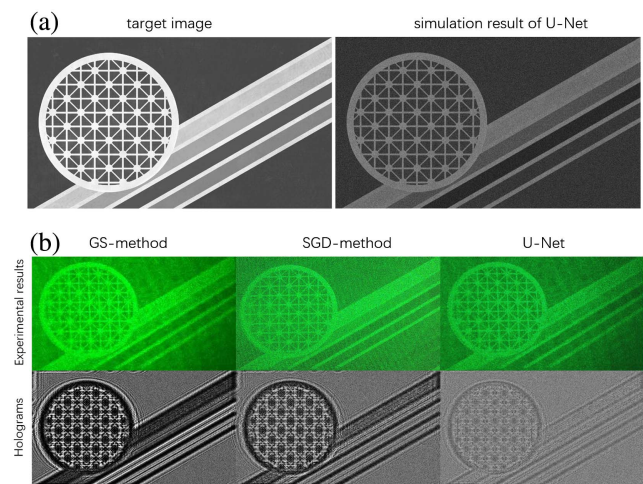
### 4. Experiments

We also build an actual phase-only holographic display prototype to validate our simulation results. All of the experiments are performed under the same condition. The schematic of the experimental setup is shown in Fig. 9. We load the CGH on a HOLOEYE PLUTO-2-VIS-014 reflective SLM. The pixel size of the SLM is 8 μm, and the pixel number is 1920 × 1080. The green laser with the wavelength of 532 nm is used. The patterns are projected on the wall and are captured by a camera. The parameters used in the experiments are consistent with those used in the simulations above.

Figure 10 demonstrates the effectiveness of our proposed method. The simulated image and the experimental result of our U-Net are shown in Fig. 10(a). The size of the projection image on the wall is 15.36 cm × 8.64 cm, which is consistent with the simulated result. We then compare the reconstruction quality of GS holography, SGD method, and our U-Net in Fig. 10(b). All of them are based on the DSFD algorithm.



**Fig. 9.** Schematic of the experimental setup [P1, P2, polarizers; C&E, collimator and expander; L, lens; BS, beam splitter].



**Fig. 10.** (a) Simulated optical image and the experimental result based on U-Net. (b) Comparison of reconstruction quality of different encoding methods.

Compared with the reconstruction results of the iterative method, the U-Net method can achieve qualified digital holography reconstruction.

The results show that our DSFD-Net model has great potential for designing a lensless holographic projection system with large FOV. Besides, current methods using machine learning for calculating the diffraction process are basically for simple diffraction algorithms, such as the S-FFT algorithm and the angular spectrum diffraction algorithm. For different tasks, the corresponding machine learning methods and systems have their own adaptations. Whether machine learning algorithms can be applied to more complex algorithms such as the DSFD algorithm is still unknown. In this paper, we use machine learning for DSFD algorithms with calculation of multiple diffraction processes and varying sampling frequencies to demonstrate that machine learning can be applied to different types of diffraction algorithms and that the corresponding CGHs can be calculated in real time. However, the proposed method still has some unsolved issues. For example, training a digital hologram of a

high-resolution image using a convolutional neural network requires a very large GPU memory size. At the current stage, it is difficult for the proposed method to further increase the image resolution with available GPUs. In the next work, we will try to compress the size of the neural network and find other network structures to adapt our method to higher resolution images. In the future, we will continue to study the algorithm of CGH, especially 3D digital holography based on machine learning.

## 5. Discussion

In this paper, the machine learning techniques are introduced to generate the hologram used in an image magnified lensless holographic projection system. Compared to the iterative method, neural network can compress computation time to the several milliseconds level. Meanwhile, the neural network can match various projection systems to meet the corresponding requirement of the near-eye display devices. The proposed method is applicable to augmented reality displays, virtual reality displays, and, hopefully, other real-time 3D display systems in the future.

## Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 61975122) and the National Key Research and Development Program of China (No. 2018YFA0701802).

## References

1. T. Zhan, K. Yin, J. Xiong, Z. He, and S. T. Wu, "Augmented reality and virtual reality displays: perspectives and challenges," *iScience* **23**, 101397 (2020).
2. K. Yin, Z. He, J. Xiong, J. Zou, K. Li, and S.-T. Wu, "Virtual reality and augmented reality displays: advances and future perspectives," *J. Phys. Photonics* **3**, 022010 (2021).
3. B. C. Kress and P. Schelkens, "Optical waveguide combiners for AR headsets: features and limitations," *Proc. SPIE* **11062**, 110620J (2019).
4. C. Chang, K. Bang, G. Wetzstein, B. Lee, and L. Gao, "Toward the next-generation VR/AR optics: a review of holographic near-eye displays from a human-centric perspective," *Optica* **7**, 1563 (2020).
5. A. Maimone, A. Georgiou, and J. S. Kollin, "Holographic near-eye displays for virtual and augmented reality," *ACM Trans. Graph.* **36**, 85 (2017).
6. P. Sun, S. Chang, S. Liu, X. Tao, C. Wang, and Z. Zheng, "Holographic near-eye display system based on double-convergence light Gerchberg-Saxton algorithm," *Opt. Express* **26**, 10140 (2018).
7. Y. Peng, S. Choi, N. Padmanaban, and G. Wetzstein, "Neural holography with camera-in-the-loop training," *ACM Trans. Graph.* **39**, 185 (2020).
8. J. Wu, K. Liu, X. Sui, and L. Cao, "High-speed computer-generated holography using an autoencoder-based deep neural network," *Opt. Lett.* **46**, 2908 (2021).
9. R. Horisaki, R. Takagi, and J. Tanida, "Deep-learning-generated holography," *Appl. Opt.* **57**, 3859 (2018).
10. Y. Zhao, L. Cao, H. Zhang, D. Kong, and G. Jin, "Accurate calculation of computer-generated holograms using angular-spectrum layer-oriented method," *Opt. Express* **23**, 25440 (2015).
11. Z. He, X. Sui, G. Jin, and L. Cao, "Progress in virtual reality and augmented reality based on holographic display," *Appl. Opt.* **58**, A74 (2019).
12. S. J. Liu, D. Xiao, X. W. Li, and Q. H. Wang, "Computer-generated hologram generation method to increase the field of view of the reconstructed image," *Appl. Opt.* **57**, A86 (2018).
13. F. Yaras, H. Kang, and L. Onural, "Circular holographic video display system," *Opt. Express* **19**, 9147 (2011).
14. R. Kang, J. Liu, G. Xue, X. Li, D. Pi, and Y. Wang, "Curved multiplexing computer-generated hologram for 3D holographic display," *Opt. Express* **27**, 14369 (2019).
15. S. J. Liu, N. T. Ma, F. X. Zhai, N. N. Liu, P. P. Li, Y. Q. Hao, and D. Wang, "Large field-of-view holographic display method with speckle noise suppression based on time multiplexing," *J. Soc. Inf. Disp.* **29**, 758 (2021).
16. C. Chang, Y. Qi, J. Wu, J. Xia, and S. Nie, "Image magnified lensless holographic projection by convergent spherical beam illumination," *Chin. Opt. Lett.* **16**, 100901 (2018).
17. W. Qu, H. Gu, H. Zhang, and Q. Tan, "Image magnification in lensless holographic projection using double-sampling Fresnel diffraction," *Appl. Opt.* **54**, 10018 (2015).
18. K. Wang, Q. Kema, J. Di, and J. Zhao, "Y4-Net: a deep learning solution to one-shot dual-wavelength digital holographic reconstruction," *Opt. Lett.* **45**, 4220 (2020).
19. F. Niknam, H. Qazvini, and H. Latifi, "Holographic optical field recovery using a regularized untrained deep decoder network," *Sci. Rep.* **11**, 10903 (2021).
20. J. Tang, J. Wu, K. Wang, Z. Ren, X. Wu, L. Hu, J. Di, G. Liu, and J. Zhao, "RestoreNet-Plus: image restoration via deep learning in optical synthetic aperture imaging system," *Opt. Lasers Eng.* **146**, 106707 (2021).
21. H. Pang, A. Cao, W. Liu, L. Shi, and Q. Deng, "Effective method for further magnifying the image in holographic projection under divergent light illumination," *Appl. Opt.* **58**, 8713 (2019).