

DOI: [10.29026/oea.2024.230176](https://doi.org/10.29026/oea.2024.230176)

Advancing computer-generated holographic display thanks to diffraction model-driven deep nets

Vittorio Bianco and Pietro Ferraro*

Advancements are reported in computer-generated holography proofing RGB 4K display through a new strategy based on diffraction model-driven deep networks. In the new 4K-DMDNet, the network is not a “black box” anymore. Rather, the input-output relation must obey to the physics of wavefront propagation, which is embedded here as a constraint. Thus, a labelled dataset is not required, and the model shows superior generalization capabilities with respect to data-driven approaches. The method is promising for the new generation of RGB 4K holographic display, as well as augmented and virtual reality systems.

Bianco V, Ferraro P. Advancing computer-generated holographic display thanks to diffraction model-driven deep nets. *Opto-Electron Adv* 7, 230176 (2024).

Since its first introduction, holography has raised the fascinated attention of a broad audience, spanning from the research community to the general public. The promise to encode a complex wavefront into a 2D holographic pattern, and to recreate from them 3D image display is intriguing. In fact, important applications of such 3D holographic display can be foreseen, e. g. telepresence, cultural heritage fruition, surgery training, entertainment, as well as the development of next generation tools for gathering new insights in biology and medicine through augmented and Virtual Reality (VR)¹⁻⁴.

Within the framework of holographic display, Computer Generated Hologram (CGH)^{3,5} technology offers the possibility to skip the interferometric recording, since in theory any photograph of a real-world object can be converted into its hologram. In analogy to the classical holographic reconstructions of patterns registered onto a photographic plate, the CGH is sent to a Spatial Light Modulator (SLM) that manipulates the incoming

light according to the input and thus can optically reconstruct and display the image of the 3D object in sharp focus. Similarly, video sequences of images can be sent to the SLM in order to create optically displayed 3D videos.

However, limitations in the SLM and CGH technology have risked to hamper the enthusiasm around such applications. Existing SLMs can modulate the amplitude or the phase of the impinging light. Phase-Only Holograms (POHs) are typically sent to phase-only SLMs, being preferred to amplitude-only holograms due to the higher diffraction efficiency and elimination of the twin image. However, converting an image into the corresponding POHs is an ill-posed inverse problem. Existing algorithms to estimate the POH are based on the Gerchberg-Saxton (GS)⁶, Wirtinger⁷, and non-convex optimization schemes⁸, which are iterative, time consuming and typically converge only to local optima. As a result, the quality of optical display, in terms of artifacts, speckle noise, contrast, and reproduction times, is way too far

CNR- ISASI Institute of Applied Sciences & Intelligent Systems Viale Campi Flegrei, 34 80078 Pozzuoli (Na), Italy.

*Correspondence: P Ferraro, E-mail: pietro.ferraro@cnr.it

Received: 25 September 2023; Accepted: 28 September 2023; Published online: 16 January 2024



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License.

To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024. Published by Institute of Optics and Electronics, Chinese Academy of Sciences.

from the expectations of the general public for a wide-spread use.

Data-driven deep neural networks have been employed to solve the iterative POH estimation problem in real time and to achieve speckle-free holographic display². The network is trained to learn the non-linear mapping between the input images and the POH estimate using a labelled dataset. In this approach, the ground-truth is generated by employing the above-mentioned iterative solvers. The loss function is calculated between the network output and the iterative solution, which is used to update the weights of the network. Data-driven deep learning requires a large amount of labelled data and, above all, is intrinsically limited by the quality of the ground-truth. In other words, a network that learns the input-output link from the ground-truth estimate cannot perform better than the estimated data.

Researchers led by Prof. Liangcai Cao, propose a novel pathway to overcome the limitations associated with data-driven methods. Their approach relied on a model-driven deep learning strategy to achieve 4K real-time RGB holographic display with unprecedented quality⁹. Instead of using a labelled dataset, the physical diffraction model is enforced in the network as a constraint. Fresnel diffraction is used as a model for simulating the propagation of the light field from the hologram plane to the object best focus plane, a process that is fully embedded in the network and is carried out during the training. During the training stage, a set of RGB images is sent to the network, which generates for each of them the POHs corresponding to the R, G and B channels. The POHs are propagated using the Fresnel model to the object best focus plane to generate a guess of the input image. Input image and the guess at the network output are compared to calculate a loss function (Negative Pearson Correlation Coefficient, NPCC) and update the weights.

The problem solved by the network is summarized by the authors in the elegant formula:

$$\begin{aligned} & \text{find } \mathbf{H} \\ \text{s.t. } & |\text{Prop}(\mathbf{H})|^2 = \mathbf{I}, \end{aligned}$$

where \mathbf{H} is the POH to be estimated, \mathbf{I} is the input image, and Prop denotes the Fresnel propagation from the hologram to the object best-focus plane. After training, the weights are “frozen” and the network is able to predict the POHs from the intensity images.

In this way, they surpass the common perception of networks as “black boxes” by achieving a larger generalization capability, since the network learns from physical constraints rather than data examples.

In their experiments, the three R, G, and B POHs are sent to a phase-only SLM (3840×2160 pixels, with 3.74 μm pixel pitch) in turn, and accordingly lasers at wavelengths 638 nm, 520 nm, 450 nm are sequentially switched on, so that the three colour channels are rapidly optically displayed. The switching period can be made shorter than the human eye integration time, so that the time multiplexing process is perceived by the human eyes as an RGB optical display.

Previous works addressed the POH generation issue using model-driven deep learning¹⁰⁻¹². Camera in the loop (CITL) strategies¹⁰ and Phase Dual Resolution Network (PDRNet)¹² are good examples in this sense. Prof. Cao’s group also developed in a previous work the Holo-Encoder¹³, a model-driven network able to provide single wavelength POHs with very fast inference time. All the model-driven networks proposed so far exhibited limited phase convergence though.

One of the key-enabling ideas behind the research of Cao’s group is that limited convergence should be traced back to insufficient constraints of the inverse problem⁹. In their work, the constraint of the reconstructed images is empowered in the frequency domain, in particular by zero-padding the spectrum^{9,14} and thus oversampling the phase in the spatial domain. The oversampling operation is directly added in the network pipeline that emulates the process of Fresnel propagation⁹.

Another important distinctive feature of their work is the introduction of the sub-pixel convolution method as a way to extend the number of learnable parameters, N , which is a sort of meter to identify the learning capability of network models (Fig.1). The conventional way to augment N is to increase the network depth at the cost of slower convergence, higher computational burden, longer training times, and more demanding hardware required. The proposed network model is a U-Net CNN consisting of a downsampling path and an upsampling path. Sub-pixel convolution extends N of the network upsampling path of a fourfold factor without extending the network depth, which is an important breakthrough to retrieve high-fidelity POH reconstructions with fast inference. Authors demonstrated that the new sub-pixel convolution outperforms the conventional upsampling methods, i.e. transposed convolution and nearest neighbor resize convolution, in terms of Peak Signal-to-Noise Ratio (PSNR), obtaining the remarkable PSNR=19.27 dB⁹.

Authors benchmarked the new 4K-DMDNet using full-color numerical simulations and testing the optical

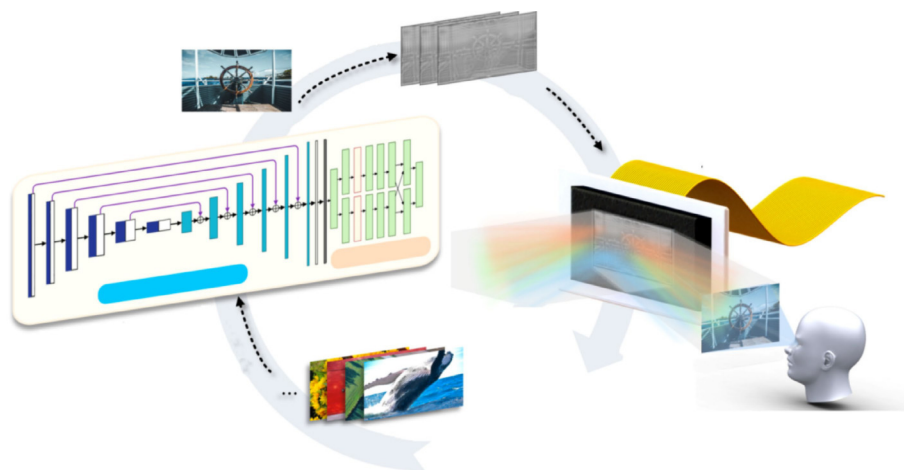


Fig. 1 | Conceptual scheme of the generation and reconstruction process of 4K POHs by the 4KDMDNet.

projection rendering. As for the numerical simulations, the 4K-DMDNet greatly outperformed the outcomes of both traditional GS and the Holo-Encoder. It was able to suppress artifacts and reject speckle noise, providing POHs with PSNR=20.49 dB in only 0.26 s. It also outperformed GS and Holo-Encoder in optical projections, showing excellent quality RGB 4K projection onto a camera at 0.3 m distance. The model-driven unsupervised approach intrinsically allowed better generalization, thus authors proved the successful optical projection of a binary pattern very different from the typical training images. And last but not least, the ability to reconstruct 3D scenes with objects at different focus planes is demonstrated by the authors. In all cases projections look natural, artifact-free and speckle-free, which is an essential requirement to unlock holographic display for use by a broad audience.

For the future, we can expect the use of this novel technology for head-up displays in portable augmented and virtual reality apparatus, new generation 3D holographic display that surpasses the current state of the art¹, and also for helping lab technicians in metalenses design and additive manufacturing, as claimed by the authors. Further interesting developments are foreseeable, e.g. replacing the U-Net with updated models such as generative adversarial networks or graph neural networks.

References

- Zhang CL, Zhang DF, Bian ZP. Dynamic full-color digital holographic 3D display on single DMD. *Opto-Electron Adv* 4, 200049 (2021).
- Shi L, Li BC, Kim C et al. Towards real-time photorealistic 3D holography with deep neural networks. *Nature* 591, 234–239 (2021).
- He ZH, Sui XM, Jin GF et al. Progress in virtual reality and augmented reality based on holographic display. *Appl Opt* 58, A74–A81 (2019).
- Bianco V, D'Agostino M, Pirone D et al. Label-free intracellular multi-specificity in yeast cells by phase-contrast tomographic flow cytometry. *Small Methods* 7 (2023).
- Sahin E, Stoykova E, Mäkinen J et al. Computer-generated holograms for 3D imaging: a survey. *ACM Comput Surv* 53, 32 (2021).
- Gerchberg RW. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik* 35, 237–246 (1972).
- Chakravarthula P, Peng YF, Kollin J et al. Wirtinger holography for near-eye displays. *ACM Trans Graph* 38, 213 (2019).
- Zhang JZ, Pégard N, Zhong JS et al. 3D computer-generated holography by non-convex optimization. *Optica* 4, 1306–1313 (2017).
- Liu KX, Wu JC, He ZH et al. 4K-DMDNet: diffraction model-driven network for 4K computer-generated holography. *Opto-Electron Adv* 6, 220135 (2023).
- Peng YF, Choi S, Kim J et al. Speckle-free holography with partially coherent light sources and camera-in-the-loop calibration. *Sci Adv* 7, eabg5040 (2021).
- Ishii Y, Shimobaba T, Blinder D et al. Optimization of phase-only holograms calculated with scaled diffraction calculation through deep neural networks. *Appl Phys B* 128, 22 (2022).
- Yu T, Zhang SJ, Chen W et al. Phase dual-resolution networks for a computer-generated hologram. *Opt Express* 30, 2378–2389 (2022).
- Wu JC, Liu KX, Sui XM et al. High-speed computer-generated holography using an autoencoder-based deep neural network. *Opt Lett* 46, 2908–2911 (2021).
- Ferraro P, De Nicola S, Coppola G et al. Controlling image size as a function of distance and wavelength in Fresnel-transform reconstruction of digital holograms. *Opt Lett* 29, 854–856 (2004).



Scan for Article PDF