https://doi.org/10.37188/lam.2026.123

News & Views Open Access

Encoding intelligence into light: a passive diffractive decoder for super-resolved, extended-depth projection

Xuhan Guo[1,*], Tao Lin[2] and Yikai Su[1]

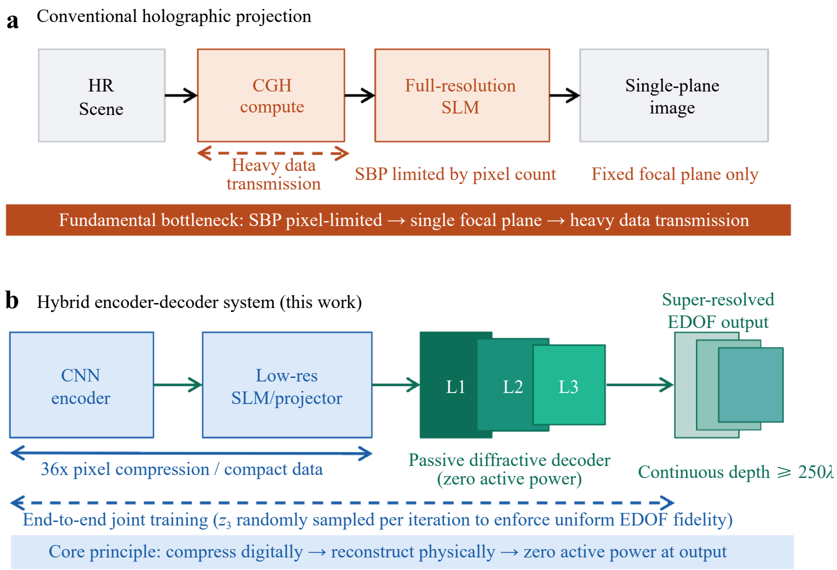
Abstract

High-resolution holographic projection necessitates a high pixel count, wide depth of field, and compact data transmission—requirements that conventional spatial light modulators cannot simultaneously satisfy owing to inherent trade-offs. This challenge is mitigated by the proposed hybrid system that combines a neural-network digital encoder with a passive all-optical diffractive decoder. The convolutional neural network compresses inputs 36-fold into compact phase patterns, while a jointly trained stack of passive diffractive layers reconstructs them optically into super-resolved images across a continuous axial volume. This architecture achieved up to a approximately 16-fold space-bandwidth product improvement and a continuous depth of field of ⩾ 250λ, with zero active power consumption at the output stage. It establishes a scalable framework toward realising energy-efficient, high-fidelity holographic projections for displays, metrology, and microscopy.

Keywords: Diffractive optical decoder, Pixel super-resolution, Extended depth of field, Hybrid digital-optical projection

Near-eye displays for augmented and virtual reality (AR/VR) must simultaneously deliver a high spatial resolution, large depth of field (DOF), and efficient data handling[1–4]. These performance metrics conflict with an each other: improving the spatial resolution requires finer spatial light modulator (SLM) pixels, thereby proportionally increasing the volume of phase data to be computed and transmitted; extending the DOF demands wider angular coverage, while the finite aperture of the SLM cannot be provided without sacrificing the space-bandwidth product (SBP); and compressing the hologram data deteriorates the high-frequency spatial information essential for ensuring image quality. Neural approaches to expand the SBP, such as the neural étendue expander[5], have successfully addressed individual facets of the problem; however, they have not resolved all three constraints simultaneously. The existing strategies address some of these challenges. Conventional CGH pipelines (Fig. 1a) drive a full-resolution SLM directly from the input scene, thus rendering the SBP intrinsically pixel-limited and fixing the focal plane. The neural étendue expander[5] widens the field of view through a learned diffractive surface, thereby expanding the SBP to approximately 6×, but operates on a single focal plane and requires active hardware at the output. The prior diffractive super-resolution (SR) decoder[6] is fully passive and achieves approximately 4× SBP improvement; however, it remains confined to a single projection plane and operates without the digital encoder needed to compress input data.

Correspondence: Xuhan Guo (guoxuhan@sjtu.edu.cn)
[1]State Key Laboratory of Photonics and Communications, School of Information and Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[2]North Ocean Photonics Co., Ltd., Shanghai 201306, China



c Performance comparison with prior approaches

Method	SBP gain	DOF range	Data compression	Decoder power	Generalization
Conventional CGH (SLM-limited)	1×	Single plane	None	Active (compute)	—
Neural étendue expander[5]	~6×	Single plane	Partial	Active	Limited
Diffractive SR decoder[6]	~4×	Single plane	None	Passive	Limited
Hybrid encoder-decoder system, Ozcan et al.[7]	~16×	$\geqslant 250\lambda$	36×	Passive	Yes

λ = illumination wavelength; SBP = space-bandwidth product;
DOF = depth of field; SR = super-resolution; EDOF = extended depth of field

Fig. 1 Architecture and comparison of the diffractive projection systems. **a** Conventional holographic projection. **b** Hybrid encoder-decoder system. **c** Performance comparison with three representative approaches[5–7].

As shown in the performance comparison with the representative approaches in Fig. 1c, no prior approach has simultaneously achieved high SBP gains, extended DOFs, data compression, and passive outputs.

Ozcan et al.[7] have demonstrated a hybrid system that resolved the projection problem (Fig. 1b). A lightweight convolutional neural network encoder compresses high-resolution input images 36-fold into compact phase patterns, which are then fed to a low-resolution SLM or projector. A jointly trained stack of passive diffractive layers (L1, L2, L3) reconstructs the wavefront downstream all-optically, thus synthesising super-resolved images across a continuous axial volume. The diffractive decoder—conceptually derived from the D^{2}NN framework[8] and the preceding single-plane SR decoder[6]—gains both digital front-end and volumetric depth coverage. Once fabricated, this hybrid system can operate as a permanently embedded analogue processor that does not require active power. The core training innovation is the randomised depth-sampling strategy, where the target reconstruction plane z_3 is drawn uniformly from the full target DOF range at each iteration. This approach compels the system to maintain high fidelity across the entire axial volume instead of optimising a single focal plane through a learned generalisation—a common limitation in classical wavefront coding. The resulting system achieved an approximately

16-fold SBP improvement and a continuous extended DOF (EDOF) spanning ⩾ 250λ in the simulation. Proof-of-concept experiments validated this approach at two wavelengths: a three-dimensionally printed single-layer decoder in the terahertz band (λ ≈ 750 nm) demonstrated a DOF of 50 mm (≈ 67λ), and an SLM-emulated layer at λ = 635 nm achieved 30 mm (≈ 47λ). External generalisations on 5,000 Quickdraw doodles spanning 100 unseen categories yielded a mean PSNR of 19.13±1.69 dB and an SSIM of 0.724±0.070, thereby confirming that the encoder-decoder combination learned a broadly applicable SR mapping instead of merely memorising the training distribution. Translating these numerical results into fabricated hardware requires robustness against two practical challenges: interlayer mechanical misalignment and finite-phase quantisation. The authors addressed both through an extended 'vaccination' training strategy originally introduced for misalignment-resilient diffractive networks[9]. By injecting controlled random displacements of the diffractive layers during training, the decoder acquired intrinsic tolerance to positional errors up to ±0.5λ laterally with less than 1 dB PSNR penalty. The same co-design philosophy applied to phase quantisation recovered the performance degradation to within 0.09 dB of the ideal even at 3-bit resolution, thereby encoding the fabrication constraints directly into the optimisation objective.

Whereas these results are compelling, the current implementation presents trade-offs that must be addressed in future studies. First, the experimental demonstrations used input images of only 24 × 24 pixels. Scaling to megapixel natural content would require substantially larger diffractive apertures and a corresponding expansion of training data. Second, the experimental DOF (47–67λ) is significantly smaller than the simulated DOF of ⩾ 250λ. This limitation stemmed from using the single-layer architecture selected for fabrication simplicity and can be mitigated by transitioning to multilayer lithography. Third, visible-wavelength operations demand sub-micrometre feature sizes over centimetre-scale apertures. Although the wafer-scale fabrication of multilayer diffractive processors has been demonstrated[10], the interlayer alignment, yield, and cost at display-relevant scales can be further improved. Finally, the system currently operates under monochromatic illumination; a polychromatic extension would require either an achromatic diffractive design or wavelength-multiplexed training, which may introduce additional complexity.

By accepting these trade-offs, Ozcan et al.[7] have provided a new conceptual perspective to holographic projection: not as a monolithic computational task but as a deliberate partition between electronics and optics. State-of-the-art neural CGH networks can synthesise photorealistic 4K holograms in real time[11,12]; however, the output resolution remains restricted to the SLM pixel count. Integrating such a front end directly with a passive diffractive decoder would delegate resolution recovery and axial depth control entirely to the optics, thereby complementing previous large-étendue waveguide approaches[13] that address field-of-view limitations in the eye box. Recent advances in all-optical tensor computing[14] and diffractive generative modelling[15] show a broader class of hybrid platforms in which digital intelligence is compiled into static optical hardware during deployment and require no runtime energy for inference. In addition to displays, the encoder-decoder principle may be applied to computational metrology and lensless microscopy[16], where a passive extended-DOF projector can replace active axial scanning. This study highlights two critical areas for investigation: the scalability of end-to-end training when handling megapixel-scale content and the adaptability of the framework under broadband illumination.

Acknowledgements

This study was supported by the National Research and Development Program of China (2023YFB2804702), the National Natural Science Foundation of China (NSFC) (62550072, 62341508), the Shanghai Science and Technology Innovation Action Plan (25LN3201000, 25JD1405500, and 24JD1401500), and the Shanghai Municipal Science and Technology Major Project.

Author contributions

X.H.G and Y.K.S initiated and supervised the project.

Conflict of interest

The author declares no competing interests.

Received: 22 June 2026 Revised: 09 July 2026 Accepted: 13 July 2026

References

1. Blanche, P. A. Holography, and the future of 3D display. *Light: Advanced Manufacturing* **2**, 28 (2021).
2. Chang, C. L. et al. Toward the next-generation VR/AR optics: a review of holographic near-eye displays from a human-centric perspective. *Optica* **7**, 1563-1578 (2020).
3. Gopakumar, M. et al. Full-colour 3D holographic augmented-reality displays with metasurface waveguides. *Nature* **629**, 791-797 (2024).
4. Xiong, J. H. et al. Augmented reality and virtual reality displays: emerging technologies and future perspectives. *Light: Science & Applications* **10**, 216 (2021).
5. Tseng, E. et al. Neural étendue expander for ultra-wide-angle high-fidelity holographic display. *Nature Communications* **15**, 2907 (2024).
6. Işıl, Ç. et al. Super-resolution image display using diffractive decoders. *Science Advances* **8**, eadd3433 (2022).
7. Chen, H. L. et al. Super-resolution image projection over an extended depth of field using a diffractive decoder. *Light: Science & Applications* **15**, 236 (2026).
8. Lin, X. et al. All-optical machine learning using diffractive deep neural

networks. *Science* **361**, 1004-1008 (2018).
9. Mengu, D. et al. Misalignment resilient diffractive optical networks. *Nanophotonics* **9**, 4207-4219 (2020).
10. Shen, C. Y. et al. Broadband unidirectional visible imaging using wafer-scale nano-fabrication of multi-layer diffractive optical processors. *Light: Science & Applications* **14**, 267 (2025).
11. Shi, L. et al. Towards real-time photorealistic 3D holography with deep neural networks. *Nature* **591**, 234-239 (2021).
12. Liu, N. H. et al. Propagation-adaptive 4K computer-generated holography using physics-constrained spatial and Fourier neural operator. *Nature Communications* **16**, 7761 (2025).
13. Choi, S. et al. Synthetic aperture waveguide holography for compact mixed-reality displays with large étendue. *Nature Photonics* **19**, 854-863 (2025).
14. Zhang, Y. F. et al. Direct tensor processing with coherent light. *Nature Photonics* **20**, 102-108 (2026).
15. Chen, S. Q. et al. Optical generative models. *Nature* **644**, 903-911 (2025).
16. Zuo, C. et al. Deep learning in optical metrology: a review. *Light: Science & Applications* **11**, 39 (2022).