nature photonics

Supplementary information https://doi.org/10.1038/s41566-026-02016-9

Volumetrically continuous 3D dynamic holography via cascaded light-field propagation

In the format provided by the authors and unedited

Contents

Supplementary Note 1: Multi-plane Fresnel holography

Multiplane holography typically reconstructs 3D scenes by superimposing multiple 2D planes. Traditional iterative optimization methods, when dealing with a large number of planes, incur significant computational costs, while non-iterative approaches (such as point cloud methods, look-up tables, and wavefront recording) speed up computation but sacrifice accuracy. Fresnel diffraction-based multiplane holography strikes a good balance between computational efficiency and reconstruction accuracy, making it widely applied in 3D holographic displays.

We note that when applied to multiplane holography, other numerical propagation models, such as the Angular Spectrum Method, do not alter the underlying layer-wise independent design paradigm. Consequently, similar to Fresnel holography, they remain subject to the inter-plane high-crosstalk limitations inherent to the layer-wise independent design approach.

As shown in Fig.S1, an image is projected at a position z_i along the propagation direction, and this image gradually defocuses as it moves away from z_i. The range of defocus represents the depth of field (DOF) of the imaging plane. DOF can serve as an indicator of the axial resolution for a single plane. In multiplane projection, overlapping DOFs between adjacent planes lead to crosstalk, degrading imaging quality

For Fresnel diffraction-based multiplane holography, the DOF of the imaging plane iii can be expressed by the following equation[1]:

$$DOF \propto \lambda \left[\frac{1}{n_h}\frac{z_i}{d\xi}\right]^2, \qquad (1)$$

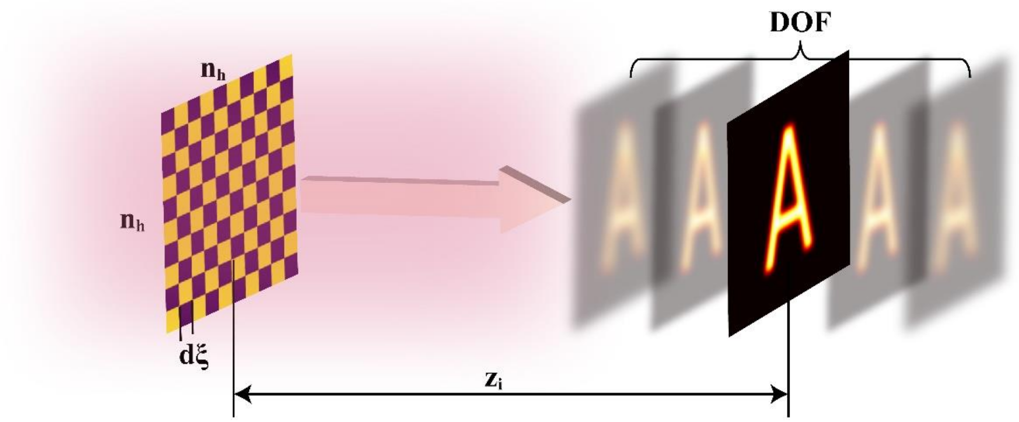

Fig.S1 | Depth of Field of a single projection plane. Projection is performed at a specific distance z_i along the propagation direction. As the distance deviates from z_i in both directions, the projected image gradually becomes defocused. The defocus range corresponds to the DOF of the imaging plane.

Here, λ is the wavelength of the incident light, $n_h \times n_h$ is the pixel number of the Fresnel hologram, z_i is the distance from the projection plane to the hologram plane, $d\xi$ is the pixel size of the hologram. The product of n_h and $d\xi$ corresponds to the aperture size. In a given imaging system, the ratio of the focal length to the aperture diameter is termed the f-number, defined as $f = \frac{z_i}{D}$. From this relationship, it is evident that reducing the aperture size increases the DOF and f-number. The f-number is inversely proportional to the numerical aperture (NA), as $f \approx \frac{1}{2 \cdot NA}$.

Considering that both the hologram's numerical aperture (NA) and physical parameters n_h and $d\xi$ are finite, Fresnel diffraction-based holography, including random-phase-based RV-CGH[1], exhibits an increasing DOF with the increasing distance z_i from the projection plane (Fig.S2a). To avoid crosstalk due to the overlap of adjacent plane DOFs, the spacing between planes away from the hologram must be increased, leading to non-uniform axial sampling and thus affecting depth perception.

As discussed, optimizing the imaging system can reduce the DOF of the projection planes, such as by introducing scattering media[2] to increase the effective numerical aperture, reducing the pixel spacing of the hologram, or increasing the number of pixels on the Spatial Light Modulator (SLM).

As shown in Fig. S2, in conventional layer-wise independent hologram design, the slices between projection planes are primarily composed of unstructured speckle, necessitating the discrete stacking of multiple independent 2D planes to reconstruct a 3D scene, thereby giving rise to intrinsic discontinuities.

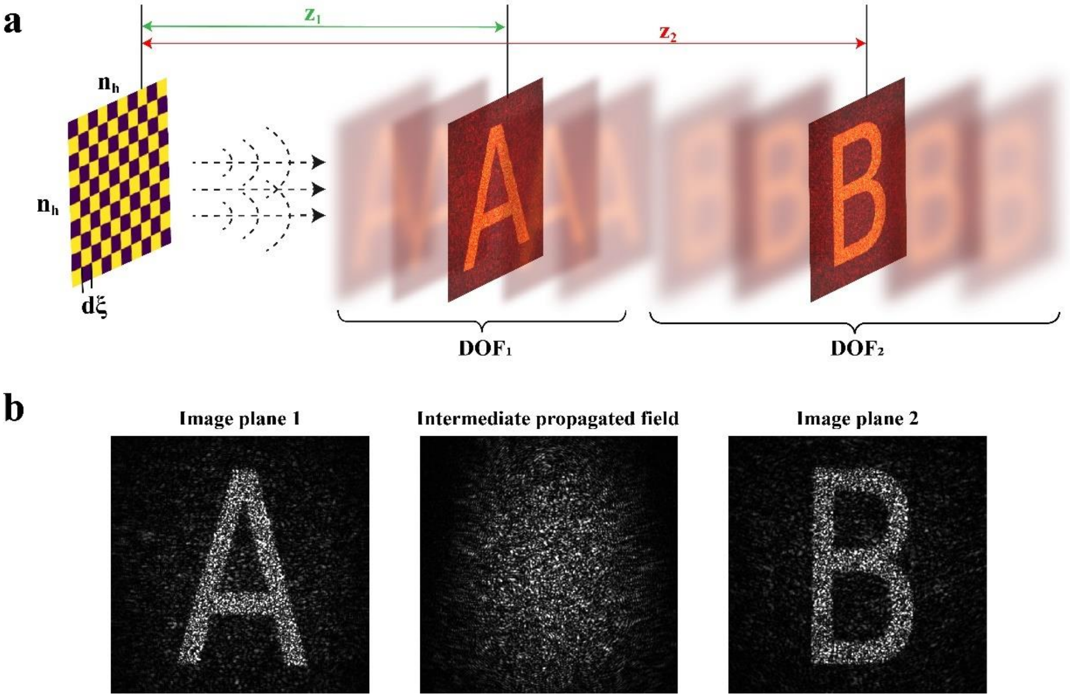

Fig. S2 | Illustration of DOF and crosstalk in conventional multiplane projection. a, The depth of field (DOF) of each projection plane depends on distance, increasing with propagation distance. Images gradually defocus as they move away from the projection plane. **b**, Experimental measurements of slices at different positions(represented by RV-CGH). To avoid crosstalk caused by overlapping DOFs, planes are spaced apart. Although the target images (“A” and “B”) are correctly reconstructed on their respective planes, the intermediate propagation fields are dominated by a speckle-like background.

Supplementary Note 2: Comparison with Fresnel holography and RV-CGH

This section presents a comparative study of the axial light field evolution characteristics in dual-plane projections using traditional Fresnel holographic CGH, RV-CGH, and the proposed method. To quantitatively analyze the distribution of the light field along the propagation direction, we extracted the XZ cross-sectional light field distributions at the same spatial location from the simulation results of the three methods (as shown by the green cross-sections in Fig.S3a-c). Fig.S3d-f display the corresponding light intensity distributions for the three methods, where the horizontal axis represents the propagation direction, and the green dashed lines precisely mark the designed depths of projection planes A and B on the Z-axis.

For traditional Fresnel holographic CGH (Fig.S3a), significant crosstalk exists between the projected images at a given depth interval. As shown in Fig.S3d, the light rays propagating from the A plane continue to the B plane, resulting in crosstalk and visible interference. Similarly, the light rays from the B plane cause significant crosstalk on the A plane (Fig.S3g). In RV-CGH (Fig.S3b), introducing high-dimensional random phases on the projection planes decorrelates the image fields between the two planes. After completing the imaging at the A plane, the light rays continue propagating, scattering into stray light that reduces crosstalk. However, the stray light continues to propagate to the B plane, as shown by the white dashed circle in Fig.S3e. Likewise, stray light from the B plane affects the A plane, as indicated by the white solid circle in Fig.S3e. While this method suppresses inter-plane crosstalk to some extent, the additional scattered light decreases the contrast of the target plane's image, as shown in Fig.S3h. Moreover, it is evident that in both methods, the total incident light energy must be distributed across the target planes. After imaging on any target plane, the light rays will become crosstalk or stray light for other planes, resulting in low efficiency in effective image formation.

The simulation results for the proposed method, shown in Fig.S3c, demonstrate ultra-low crosstalk and high-contrast dual-plane projections. From Fig.S3f, it can be seen that the optimization algorithm ensures that the output light from the A plane, after imaging, also serves as the input light for the B plane. This approach can be extended to more target planes, where the complex amplitude of each target plane is simultaneously optimized according to preset constraints, ultimately leading to the optimized design of the CGH. As shown in Fig.S3i, the image reproduced on the target plane exhibits high clarity and contrast.

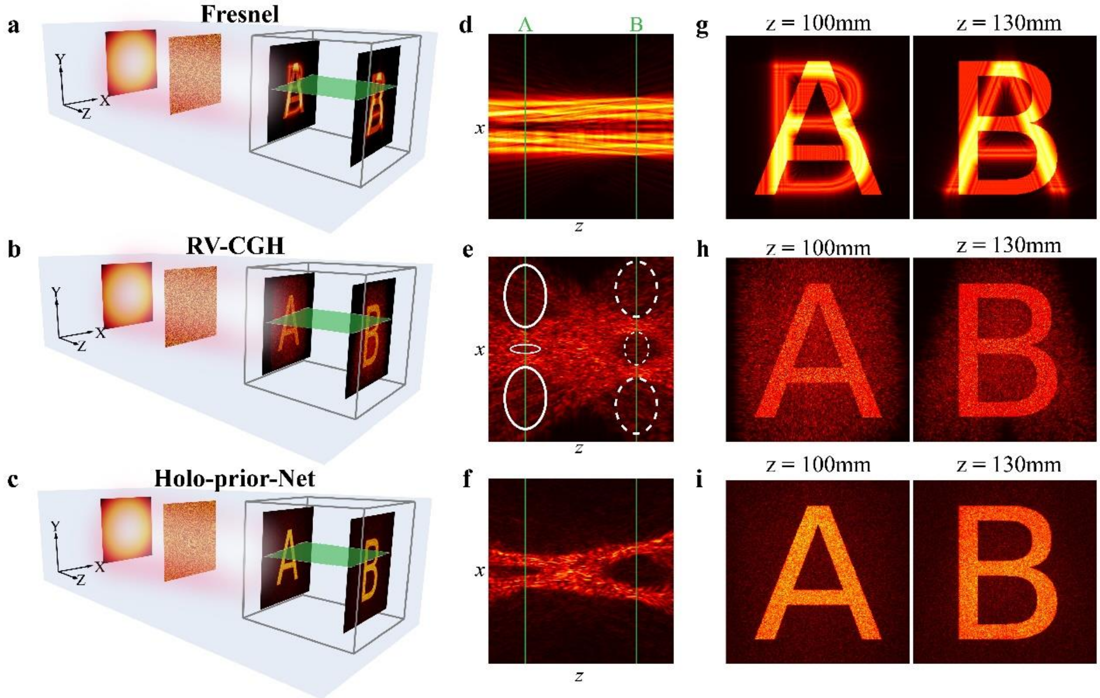

Fig. S3 | Axial evolution of light field during dual-plane projection using conventional 3D Fresnel CGH, RV-CGH, and the proposed method. a and d show the dual-plane projection and light field distribution on the X-Z cross-section using Fresnel CGH. **b** and **e** show the dual-plane projection and light field distribution on the X-Z cross-section using RV-CGH. **c** and **f** show the dual-plane projection and light field distribution on the X-Z cross-section using the proposed method. **g, h,** and **i** are the corresponding light intensity distributions on the projection planes.

As shown in Fig.S4, we further analyzed the Signal-to-Noise Ratio (SNR) comparison between two methods (RV-CGH and the proposed method) under the same input light power in dual-plane projections. When the depth interval between the two planes is large (e.g., z = 100 mm and z = 150 mm), the RV-CGH method clearly projects letter images, with SNR values of 15.7dB and 16.2dB, respectively. In contrast, the SNR for the proposed method reaches 21.2dB and 21.0dB. The images show that the background in the darker regions of the projected images using the proposed method is much darker, resulting in higher contrast. By fully optimizing the design, the proposed method utilizes the emitted light from each projection plane more efficiently, achieving higher energy utilization compared to RV-CGH's scattering suppression method.

Fig.S5 compares the simulation results for projecting three letter images using RV-CGH and the proposed method, with a target plane separation of 50 mm, under the same light source power. Assuming an ideal case where the total incident energy for the three target planes is 100%, the RV-CGH projections have energies of 48%, 60%, and 39.8% of the initial incident light, respectively. The sum of the energies for the three letters exceeds 100% because, in addition to the light directly imaged at plane A, part of the scattered light from letters B and C contributes to the energy at plane A. Similarly, scattered light also exists at planes B and C, resulting in a significant amount of speckle noise in the background and lowering the contrast. In contrast, the proposed method achieves energy efficiencies of 83.7%, 88.2%, and 78% at the three projection planes, respectively. At plane A, 83.7% of the incident light is used for effective imaging, and 88% of the energy continues to participate in imaging at plane B. The light from plane A efficiently contributes to the imaging at plane B, and the efficiency at plane C reaches 78%. The energy utilization for each projection plane is higher than that of the random phase method, indicating that the proposed method does not discard the emitted light from each projection plane but successfully uses it as an effective input for the next layer, improving energy utilization while maintaining high contrast.

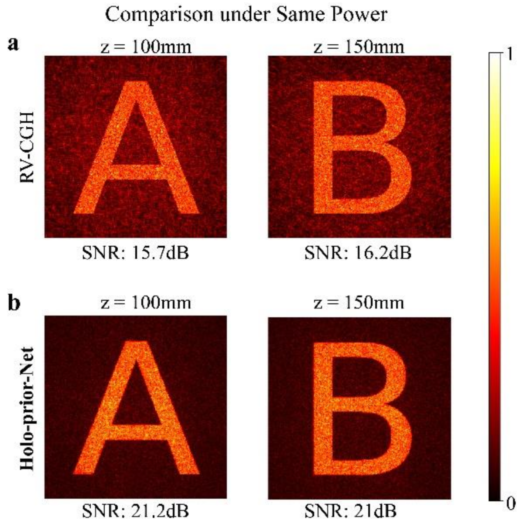

Fig.S4 | Comparison of crosstalk suppression effect between RV-CGH with random phase and the proposed method, using the same input laser power.

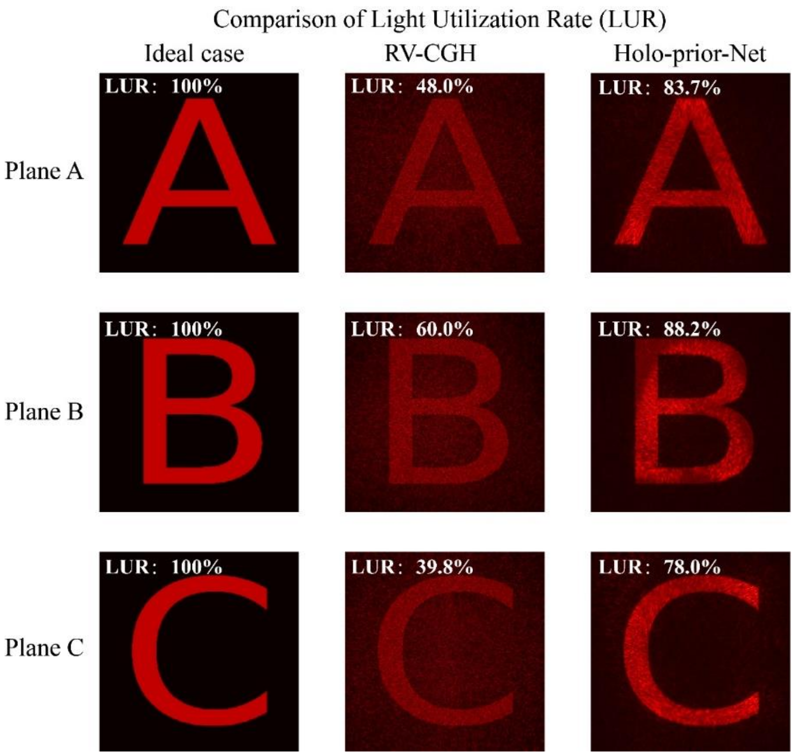

Fig.S5 | Comparison of light energy utilization between RV-CGH with random phase and the proposed method, with the same input laser power. Projection of letters A, B, and C on three planes spaced 50mm apart, calculating the energy fraction of effective imaging light on each plane compared to incident light.

Supplementary Note 3: Axial continuity and reduced lateral spacing

This section not only demonstrates the effectiveness of the proposed axial continuous depth projection but also further analyzes the imaging quality after reducing the lateral spacing. In light sheet holography-related studies[3], to avoid lateral crosstalk induced by Bessel beams, the minimum spacing between four axial continuous projection planes needs to be 3mm. To assess the lateral spacing and crosstalk characteristics of the proposed method, Fig. 4b in the main text presents the results of simultaneously projecting four axial continuous images and compares the simulation and experimental imaging quality with plane spacings of 600μm and 300μm. The smaller the plane spacing, the higher the accuracy required for experimental measurements.

As illustrated in Fig.S6a, we simultaneously project four images along the direction perpendicular to the display and analyze the variation in imaging quality under smaller spacings. Supplementary Fig.S6b-e demonstrates the impact of projection plane crosstalk for spacings of 300μm, 200μm, 100μm, and 50μm. As the spacing decreases, the crosstalk gradually increases: at 200μm, the imaging quality remains relatively good; however, as the spacing is further reduced to 100μm and 50μm, scattered light crosstalk gradually intensifies. This is mainly due to the trade-off between image complexity and the number of available holographic pixels — as the image complexity increases, more pixels are required to provide additional capability or finer pixel spacing for more precise control to avoid crosstalk.

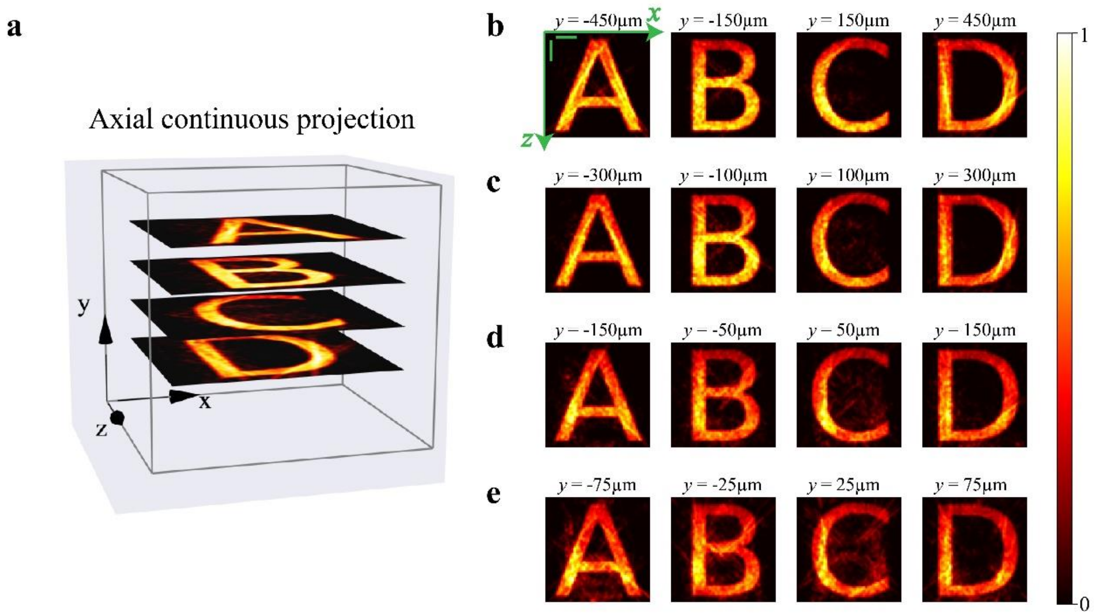

Fig.S6 | a, Schematic diagram of simultaneous projection of multiple axially continuous images. **b,** Image quality under different lateral spacings. (Horizontal scale bar: 1mm, vertical scale bar: 20mm.)

Supplementary Note 4: Enhancement of 3D holographic projection image quality

The complexity of the reconstructed 3D scene is constrained by the information capacity of the digital hologram, which can be quantified by the scalar bandwidth product (SBP). A hologram with a higher SBP generally provides projections with larger volumes and finer details. In some cases, however, due to the highly complex and densely packed projection images, the quality of the reconstructed 3D hologram may not be sufficiently clear (e.g., Fig. 3d in the main text). This limitation is related to the available holographic information capacity and system parameters rather than being specific to the proposed approach. The reconstruction quality is influenced by parameters including the hologram pixels number and pixel spacing, and these can be adjusted to further improve the image quality. In our experiments, the selection of hologram parameters was primarily influenced by the specifications of the commercial SLM (Meadowlark Optics: 1920 × 1152 pixels, pixel pitch 9.2μm). Future improvements in device resolution and aperture size may provide opportunities to increase the effective SBP, for example, by employing holograms with larger pixel arrays, smaller pixel pitches, thereby potentially enhancing projection quality.

Pixel Number and Projection Quality

We investigated the influence of hologram pixel counts on 3D scene reconstruction by considering an example of an axially continuous grayscale projection along the XZ plane (as shown in Fig.S7a), comparing the imaging quality before and after adjusting the hologram's pixel count. Fig. S7b, c, and d present the projection results using holograms with 512×512, 1024×1024, and 2048×2048 pixels, respectively. With a limited hologram pixel count (Fig.S7b), the reconstruction image of a potted succulent plant exhibits lower contrast, reduced grayscale fidelity, and lost fine spatial features. As the pixel count increases, Fig.S7c demonstrates improved grayscale graduation and clearer visual details. Fig.S7d (2048×2048, exceeding the SLM pixel count) provides additional improvement in projection quality, with enhanced smoothing and finer structural resolution, such as the details of the potted succulent plant. These observations suggest that higher pixel counts contribute to enhanced visual contrast, improved grayscale dynamic range, and better retention of fine structural features within the projected image.

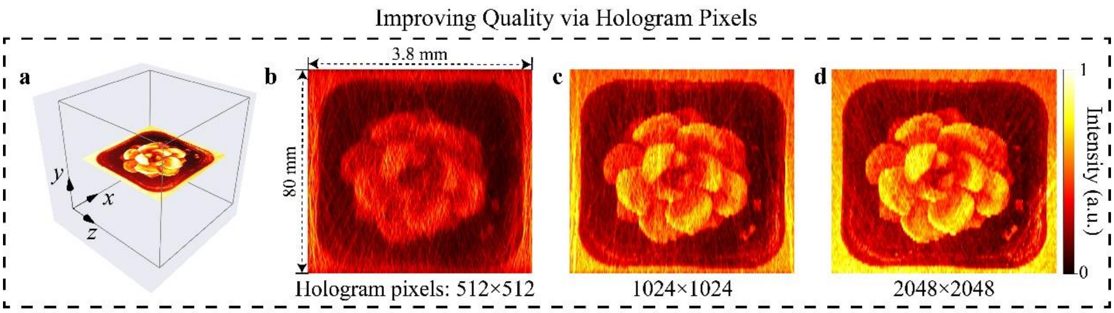

Fig.S7 | Effect of increasing hologram pixel size on projection image quality. **a,** Schematic diagram of axial projection. **b, c,** and **d** show the effects of using single-layer holograms with 512×512, 1024×1024, and 2048×2048 pixels, respectively, for axial continuous projection of a potted succulent plant.

Multi-layer Diffraction Layer Structure and Projection Quality

In this system, hologram design is based on the untrained neural network framework, where the phase distribution corresponds to the phase information of the network's diffraction layers. Increasing the number of pixels or diffraction layers provides additional degrees of freedom for optimization, which may improve the reconstruction fidelity for complex projection targets. In principle, multi-layer architectures, implemented through additional optical elements or diffraction layers, could further expand the available design space for hologram generation.

To investigate the influence of diffraction layers, we fixed the pixel count of each diffraction layer at 512×512, ensuring that the differences in control effects between different diffraction layers were not masked by the pixel count of a single layer. Fig.S8a and b show schematic diagrams of the network structure with one and three diffraction layers, respectively, for 3D scene hologram design. Fig.S8c illustrates the reconstructed image plane at a single propagation depth. Fig.S8d, e, and f show the projection results using 1, 2, and 3 diffraction layers.

The single-layer structure (Fig.S8d) reconstructs the overall structure of the potted succulent plant image but exhibits reduced image contrast and residual background intensity. With two diffraction layers (Fig. S8e), the reconstructed image shows improved contrast and reduced background noise, leading to clearer recovery of image details. Further increasing the number of diffraction layers to three (Fig. S8f) provides additional improvement in background suppression and grayscale consistency, resulting in a reconstruction with enhanced visual fidelity.

These observations suggest that, under a limited hologram pixel capacity, increasing the number of diffraction layers introduces additional optimization degrees of freedom, which may facilitate improved reconstruction fidelity for complex grayscale images.

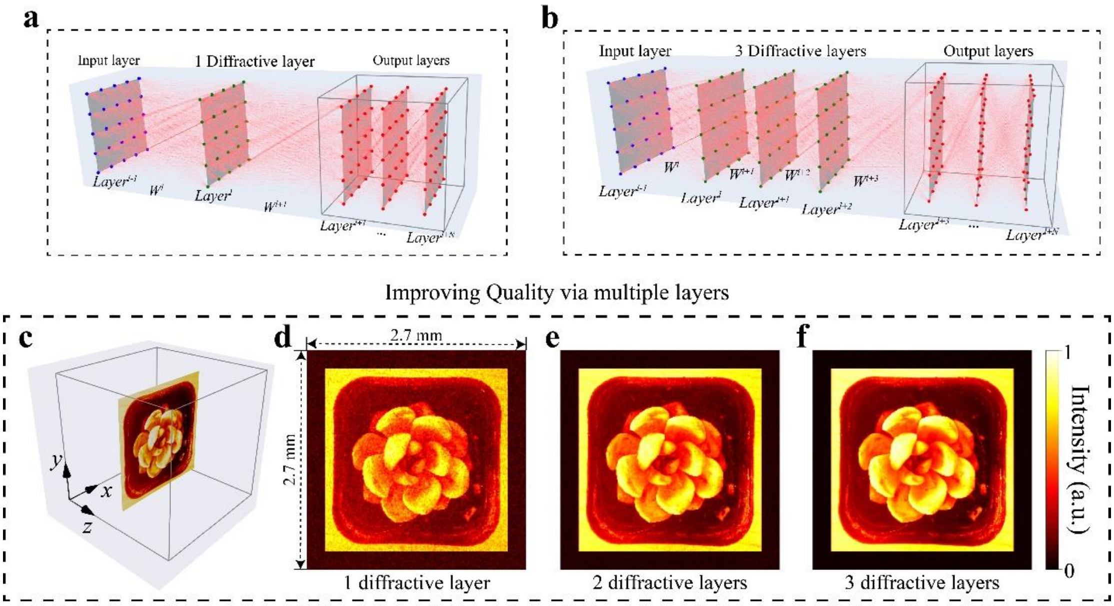

Fig.S8 | Effect of diffraction layer number on reconstruction quality at a specific depth. a-b, Network frameworks with single and three diffraction layers. **c,** illustrates of image reconstruction at a selected depth plane. **d-f,** Reconstruction results obtained using single-layer, two-layer, and three-layer holograms, each with 512×512 pixels, for a potted succulent plant target.

Pixel Pitch and Projection Quality

We compared the axial projection quality of different pixel pitches under the same hologram area. Fig.S9a, b, and c show the results with pixel pitches of 9.2μm (512×512), 9.2μm (512×512), and 2.3μm (2048×2048), respectively.

Fig.S9a and b show projections of holograms with the same pixel pitch and pixel count in different axial ranges of 80mm and 24mm. The reconstruction over the larger range (Fig.S9a) preserves more details than that over the shorter range (Fig.S9b), suggesting that the selected projection range influences the distribution of encoded information and reconstruction fidelity.

Fig.S9b and c compare projections with the same hologram area but different pixel pitches. When the axial range is reduced to 24 mm, the hologram with a pixel pitch of 9.2μm (Fig.S9b) only shows blurred outlines, while the hologram with a pixel pitch of 2.3μm (Fig.S9c) more clearly reproduces details, even within the same axial range. This indicates that a smaller pixel pitch can project finer details and improve the axial resolution by projecting over shorter axial distances.

We further investigated the effects of using smaller pixel pitches in holograms with the same pixel count. As shown in Fig.S10a and b, holograms with 2048×2048 pixels and pixel pitches of 9.2μm and 1μm were used for axial continuous projection. Despite the significant difference in hologram area, both setups successfully reconstructed the image, but the projection with a 1μm pixel pitch showed more grayscale variation, higher contrast, and smoother images.

These results indicate that reducing hologram pixel pitch can provide additional sampling capability and may facilitate improved reconstruction fidelity for complex grayscale images. The pixel pitch influences the spatial frequency range that can be encoded in the hologram, while the ultimate reconstruction performance depends on the combined effects of hologram sampling, pixel count, optical parameters, and system implementation. Future holographic platforms with higher pixel densities and smaller pixel pitches, including metasurface platforms[4,5], may provide additional opportunities for improving the reconstruction of complex 3D scenes.

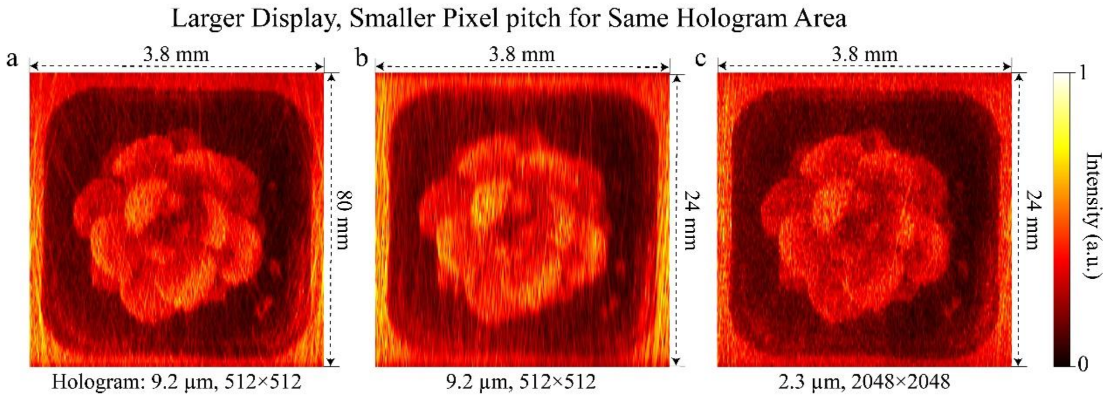

Fig.S9 | Effect of larger projection area and smaller pixel pitch on projection image quality under the condition of the same hologram area. The holograms have feature sizes of 9.2μm, 9.2μm, and 2.3μm, with pixel counts of 512×512, 512×512, and 2048×2048, respectively. **a** and **b** show projection results in axial ranges of 80 mm and 24 mm for holograms with the same parameters. **b** and **c** show projection results in a shorter axial range using holograms with different pixel pitches, while maintaining the same hologram area.

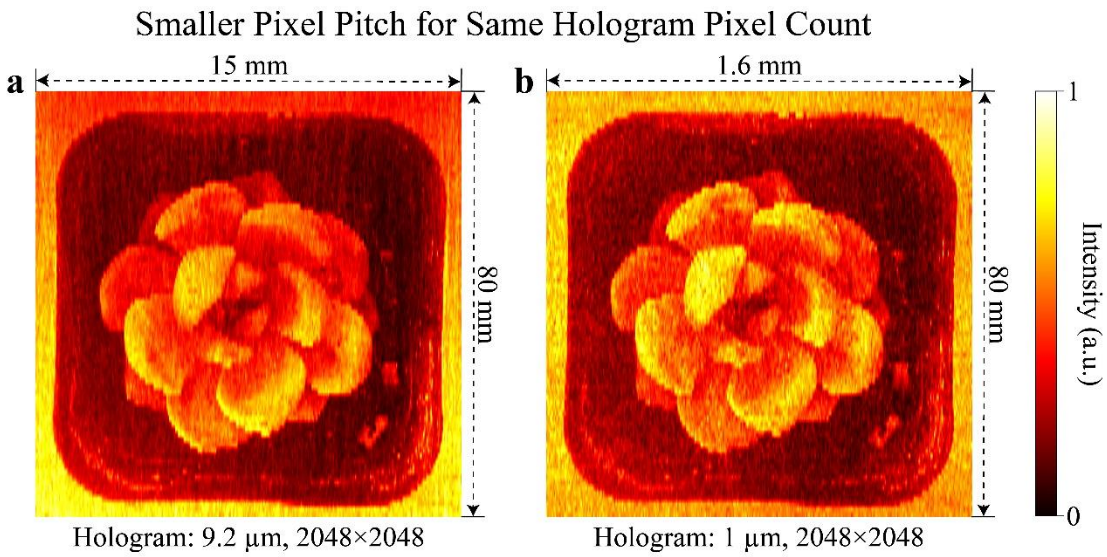

Fig.S10 | Effect of smaller pixel pitch on projection image quality with the same hologram pixel counts. The holograms have feature sizes of 9.2μm and 1μm, with pixel counts of 2048×2048.

Supplementary Note 5: Fully continuous 3D hollow sphere

To achieve realistic 3D holographic display, precise axial and lateral continuous reconstruction is crucial. We designed a 3D hollow sphere and validated its lateral and axial continuous reconstruction effects. Fig.4c in the main text demonstrates the experimental measurement of the reconstructed hollow sphere's 3D display and its overall contour from different angles.

To further verify the continuity of the 3D holographic projection, we performed slice analysis of the hollow sphere along different directions. Fig.S11 shows the projection results of slicing along the XY, YZ, and XZ directions at relatively large intervals (as depicted in the 3D schematic diagram). Each row corresponds to a slicing direction, with each slice being a ring whose radius gradually increases and then decreases, in line with the slicing characteristics of a 3D sphere. Unlike light-sheet holography, which simulates the 3D sphere by stacking several parallel rings along the axial direction, our method presents rings in the XY, YZ, and XZ planes at any direction, thus more accurately representing a 3D sphere.

To verify the fully continuous projection effect, we further performed closely spaced slice analysis on the 3D sphere. As shown in Fig.S12, we performed tight slicing along the XY, YZ, and XZ directions with very small intervals (as shown in the 3D schematic diagram). Despite the extremely small sampling interval, the experimental reconstruction results show slices of similar-sized rings in each row. This indicates that the projected 3D scene is not made up of sparsely spaced planes. Combining Fig.S11 and Fig.S12, whether with large or small interval slicing, our results consistently show complete rings, demonstrating that the projected 3D hollow sphere achieves fully continuous reconstruction both axially and laterally, enabling a more realistic fully continuous 3D projection.

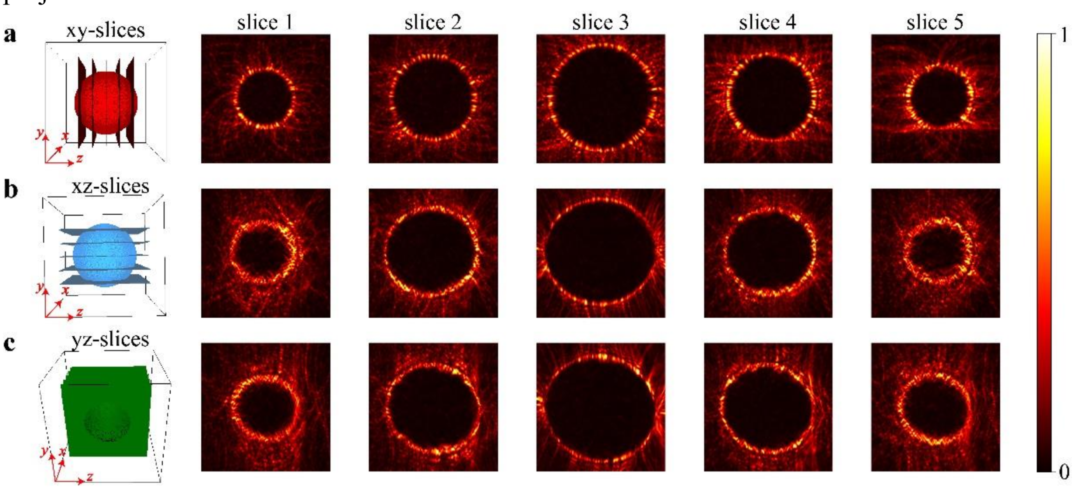

Fig.S11 | Large-spacing slice examination of the experimentally reconstructed hollow sphere along the XY, YZ, and XZ directions, demonstrating that each slice is a ring with a radius that first increases and then decreases, consistent with the slicing effect of a 3D sphere. **a,** Light field distribution on the XY plane sampled at large axial intervals. **b,** Light field distribution on the XZ plane sampled at large axial intervals. **c,** Light field distribution on the YZ plane sampled at large axial intervals.

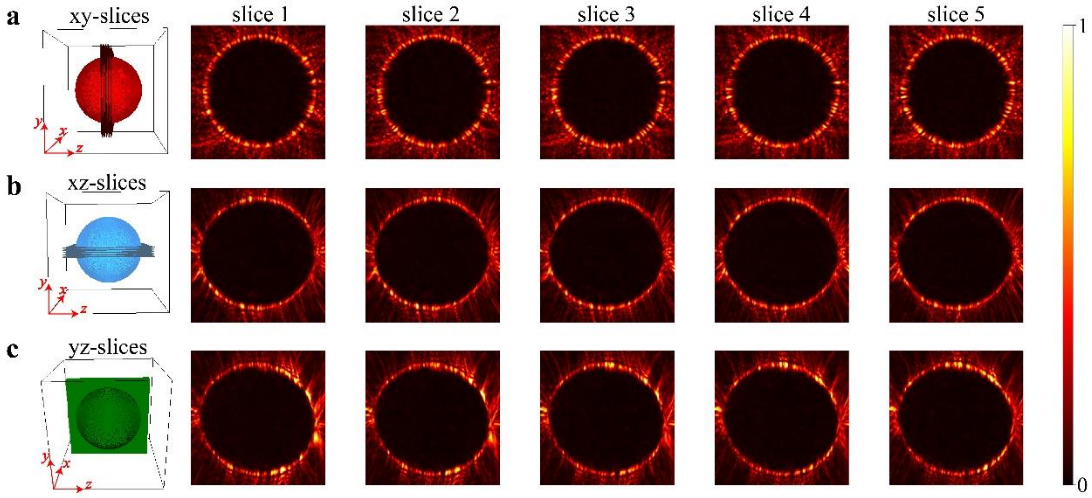

Fig.S12 | Small-spacing slice examination of the experimentally reconstructed hollow sphere along the XY, YZ, and XZ directions to verify continuity. **a,** Light field distribution on the XY plane sampled at small axial intervals. **b,** Light field distribution on the XZ plane sampled at small axial intervals. **c,** Light field distribution on the YZ plane sampled at small axial intervals.

Supplementary Note 6: Dynamic projection experiments

To validate the proposed method's capability to rapidly design dynamic target projections, we conducted experiments in which the target light field patterns were changed. The corresponding holograms were automatically generated by the proposed method based on the changing patterns and subsequently loaded onto the SLM. The light intensity distribution at the target position was then recorded using a camera. It is important to emphasize that the system implements dynamic hologram design rather than sequentially loading pre-computed holograms.

In the experiment, we designed a continuous switching of 14 target light field frames to simulate a dynamic scene where the snake moves according to the apple position, with both the snake and the apple represented as light spots. The results are shown in Supplementary Material Fig.S13, which records the 14 frames of the actual light field changes. The entire process, including scene switching and light field modulation, took less than 7 seconds. On average, the time to switch from one target light field to the phase generated by the neural network was about 0.43 seconds per frame, covering the entire process from phase generation, saving, and loading to the SLM. A dynamic demonstration video is provided in Supplementary Video 5.

The hardware configuration used in the experiment was as follows: NVIDIA GeForce RTX 3090 GPU and Intel® Xeon® Silver 4210R CPU @ 2.40 GHz with 32 GB RAM. To highlight the computational speed of the neural network, the experiment was conducted using only 256×256 pixels on the SLM. The 14-frame scene switching was completed within 7 seconds, including hologram generation and loading. With parallel computing or more efficient GPUs, the processing speed can be further accelerated.

In Supplementary Note 13, we provide a more detailed characterization of speed and scaling for more complex 3D projection results.

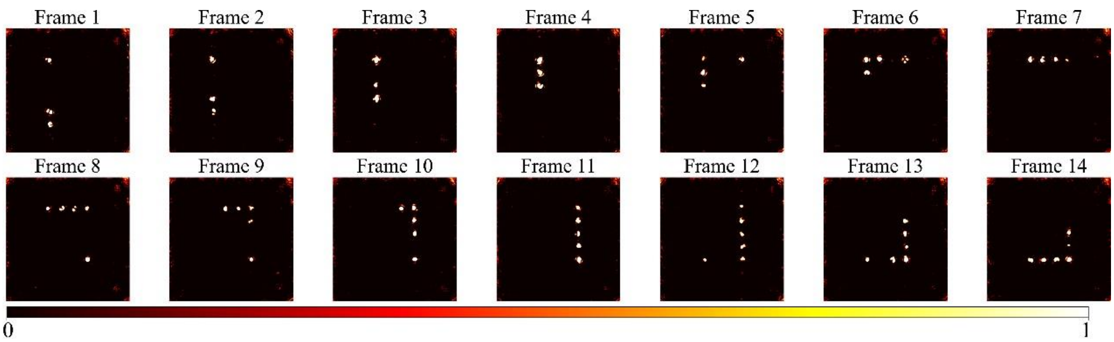

Fig.S13 | Dynamic projection. Fourteen frames of target light field sequences are designed to simulate the movement of a snake based on the apple's position. The experimental measurement results of the fast hologram generation and reconstruction are shown.

Supplementary Note 7: Video captions

Video.1: Simulation results of Fresnel holography, RV-CGH, and Proposed method. The light field variations at different depth positions within a 150mm axial range are shown. The projection planes are located at 50mm and 100mm. The propagation process from Letter A to Letter B illustrates how outgoing light from the preceding plane is utilized, rather than discarded.

Video.2: 3D volumetric rendering of experimentally reconstructed light fields for a continuous DNA double helix across different viewing angles. The 3D hologram was optically reconstructed and projected within an 80mm axial range. The lateral spatial dimension of the helix is 0.6mm, with a line width of approximately 50μm.

Video.3: 3D volumetric rendering of experimentally reconstructed light fields for a continuous hollow cylinder across different viewing angles. The 3D hologram was optically reconstructed and projected within a 40mm axial range, with the hollow cylinder having a diameter of 5mm and a wall thickness of 100μm.

Video.4: 3D volumetric rendering of experimentally reconstructed light fields for a continuous hollow light sphere across different viewing angles. The sphere has a lateral range of 1mm and an axial range of 40mm.

Video.5: Rapid dynamic hologram design. A sequence of 14 target light field distributions is dynamically designed, and the corresponding holograms are loaded onto the SLM. The experimental results were captured using a CCD camera, with a response time of less than 0.5 seconds per frame.

Video.6： Free-space volumetric visualization of a dynamically rotating 3D continuous double-helix light field in scattering smoke. Twenty 3D holograms corresponding to different spatial orientations of the double-helix structure were sequentially and continuously loaded on the SLM, producing a volumetric display of the dynamically rotating 3D double-helix light field.

Video.7: Projection of a continuous 3D volumetric light field, directly observable by the naked eye and viewable from multiple angles, without relying on post-processing or screen rendering.

Video.8： Axial continuous scanning experimental videos comparing different design paradigms. The translation stage moved over a 50 mm range with 0.1 mm steps, producing 500 raw camera frames. The letters “A” and “B” were projected with a 30 mm axial separation. The videos compare the layer-wise independent design paradigm (represented by RV-CGH) and the proposed cascaded propagation design paradigm. As the reconstruction planes move axially, the layer-wise design shows defocused output light and inter-plane crosstalk, whereas in the cascaded propagation design, the output light from one plane is constructively used as the imaging light for the next plane rather than being scattered into random speckles.

Video.9： Axial continuous scanning of the 3D double-helix structure. The translation stage moved over a 140 mm range with 0.2 mm steps, generating 700 raw camera frames. The scan includes intermediate depths not explicitly sampled during the design process, confirming the continuity of the projected light field.

Supplementary Note 8: Distinct Projection Paradigms and Axial Field Evolution

Axial continuous scanning further highlights the fundamental distinction between the conventional layer-wise independent design paradigm, exemplified by RV-CGH, and the proposed cascaded propagation approach. In the layer-wise independent scheme, the light field reconstructed on a target plane is deliberately engineered to rapidly defocus at non-target axial positions, decorrelating contributions from other planes into a speckle-like background. While this strategy visually suppresses inter-plane interference, it effectively discards useful optical energy, leading to reduced brightness, contrast, and SNR, and necessitating relatively large inter-plane spacing, as illustrated in the upper row of Fig. S14.

By contrast, in the proposed cascaded propagation scheme, the output field from each reconstructed plane is treated not as unwanted defocus noise but as a constructive input for subsequent planes. As shown in the lower row of Fig. S14, the light field evolves smoothly and controllably between adjacent planes, enabling a continuous transformation of the reconstructed intensity from plane A to plane B without relying on scattering into a speckled background. This controlled axial evolution allows higher brightness and cleaner backgrounds at each target plane for the same optical power, while preserving axial continuity of the 3D light field. Crucially, these observations confirm that crosstalk suppression in the proposed approach does not rely on defocus-induced random speckle, but emerges from a cascaded forward design in which the output of each plane constructively contributes to subsequent planes, transforming conventional crosstalk light into useful imaging light and preventing inter-plane interference at the design stage.

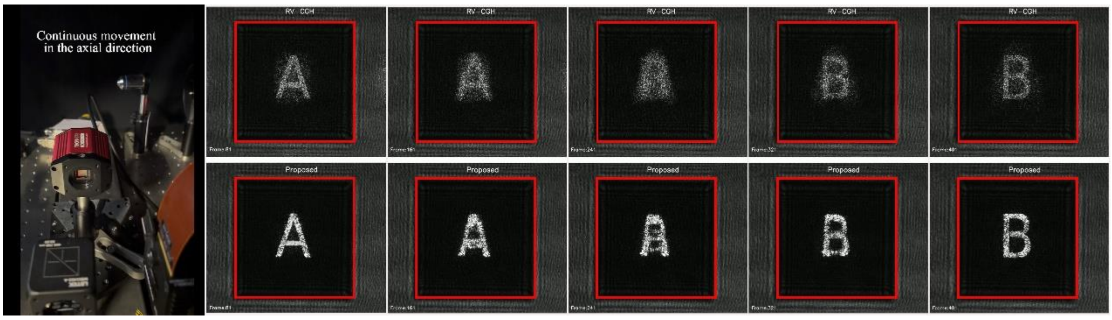

Fig.S14 | Axial continuous scanning under RV-CGH and proposed cascade propagation paradigm. Schematic of the axial scanning experiment is shown on the left. The upper row presents axial scan results obtained using RV-CGH. As the reconstruction plane moves along the axial direction, the emitted light from the projected image defocuses and is discarded, manifesting as crosstalk with adjacent planes. The lower row shows the continuous evolution of the light field at different axial positions under the proposed cascaded propagation scheme, using the same inter-plane spacing and optical conditions, where light is not scattered into random speckle. The red boxes indicate the holographically modulated region on the SLM, while areas outside the red boxes correspond to the SLM-reflected background light. Reflection of background light in both rows confirms that the two approaches are compared under identical illumination power. (See Supplementary Video 8 for the axial scanning experiment.)

To quantitatively assess inter-plane crosstalk suppression, the contrast-to-noise ratio (CNR) and crosstalk ratio were measured along the propagation direction for a two-plane letter projection (A and B, separated by 10 mm), as defined in Supplementary Note 14. Experimental axial CNR curves and corresponding reconstructed slices obtained from raw camera frames are shown in Fig. S15. In the RV-CGH scheme (Fig. S15a,b), CNR peaks remain below 3, with blurred target planes and intermediate slices dominated by speckle, indicating substantial inter-plane crosstalk. In contrast, for the proposed method (Fig. S15c,d), CNR curves exhibit two well-defined peaks at planes A and B, and intermediate slices reveal smooth axial evolution of the light field rather than random scattering. The reconstructed images show cleaner backgrounds and higher brightness for both letters, demonstrating effective crosstalk suppression and continuous axial evolution of the 3D light field. Corresponding axial crosstalk ratio distributions (Fig. S16) further indicate that the proposed method significantly reduces crosstalk at the target planes, minimizing energy leakage outside the intended projection areas. These results align with physical expectations, as the cascaded propagation scheme avoids random scattering.

Taken together, these qualitative and quantitative comparisons demonstrate that the cascaded propagation design paradigm inherently avoids speckle noise and energy loss characteristic of conventional approaches, resulting in higher energy efficiency, contrast, and SNR. We note that these improvements are not the primary objective but rather a direct consequence of the cascaded propagation design principle.

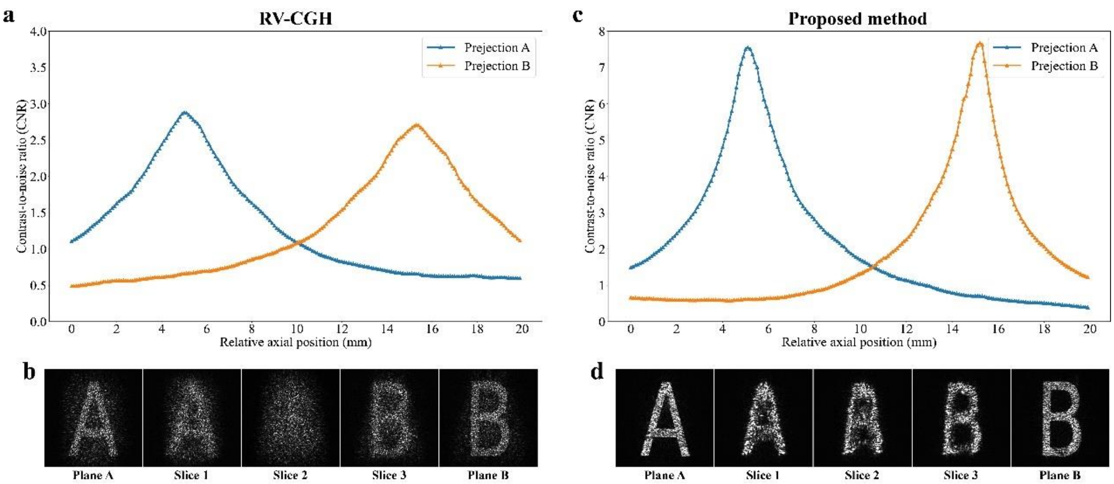

Fig. S15 | Experimental axial CNR profiles for double-plane letter projections (A and B, 10 mm separation). **a,c** Axial CNR profiles of RV-CGH and proposed method. **b,d** Reconstructed slices. Plane A and Plane B are the target projection planes; Slices 1–3 are intermediate sampling positions between them.

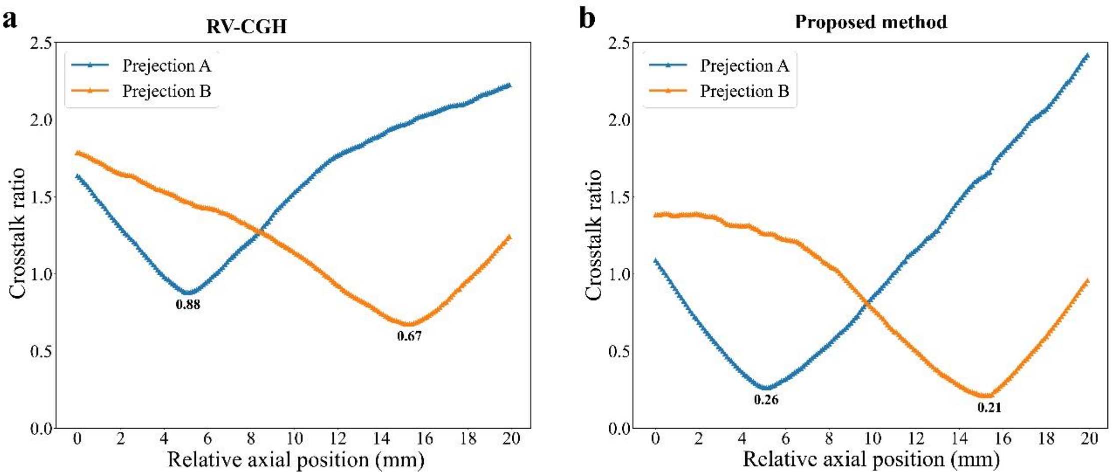

Fig. S16 | Axial crosstalk ratio for dual-plane letter projections (10 mm spacing) measured experimentally. a, RV-CGH exhibits substantially higher crosstalk (A: 0.88, B: 0.67), indicating significant energy leakage and inter-plane interference. **b**, The proposed cascaded propagation method achieves markedly lower crosstalk on the target planes (A: 0.26, B: 0.21), with reduced energy leakage.

Supplementary Note 9: Constructive Reuse of Complex Fields in Cascaded Propagation

The aim of these experiments is not to display visually complex amplitude or phase patterns, but to illustrate the physical characteristics of the proposed cascaded propagation framework in controlling complex-valued light fields. Specifically, this framework enables the coordinated optimization of amplitude and phase across multiple axial planes, allowing for controlled evolution of the complex field along the propagation direction.

In conventional layer-wise independent design approaches, such as Fresnel holography, RV-CGH, or light-sheet holography, each plane is typically constructed independently around its target intensity distribution, without coordinated phase control between planes. The output field from a preceding plane is often discarded and rapidly defocused at other axial positions, reducing its correlation. While this strategy can suppress inter-plane interference to some extent, it inherently limits the capability to represent three-dimensional complex field structures globally.

By contrast, in the proposed cascaded propagation framework, the phase at each projection plane is jointly optimized across all planes rather than relying on embedded random phase. This additional phase degree of freedom enables the output from a preceding plane to be constructively guided as input for subsequent planes, rather than being treated as uncorrelated background. This structured control is expected, from a physical standpoint, to improve the overall quality of three-dimensional light field projections by fully exploiting the light emerging from each plane.

As illustrated in Fig. S17, each projection plane is assigned a deliberately optimized phase rather than a random phase, providing an extra degree of freedom to direct the evolution of the field from preceding planes into constructive contributions at subsequent planes. The cascaded propagation mechanism produces a continuous and structured evolution of the complex field along the depth direction, with energy transfer between planes no longer relying on defocusing for suppression.

These results demonstrate that coordinated optimization of complex amplitude across multiple projection planes enables controllable intensity formation and axial evolution, while avoiding the loss of output light as random scattered background or inter-plane crosstalk. This capacity for structured control constitutes a key physical feature distinguishing the cascaded propagation framework from traditional layer-wise independent design paradigms.

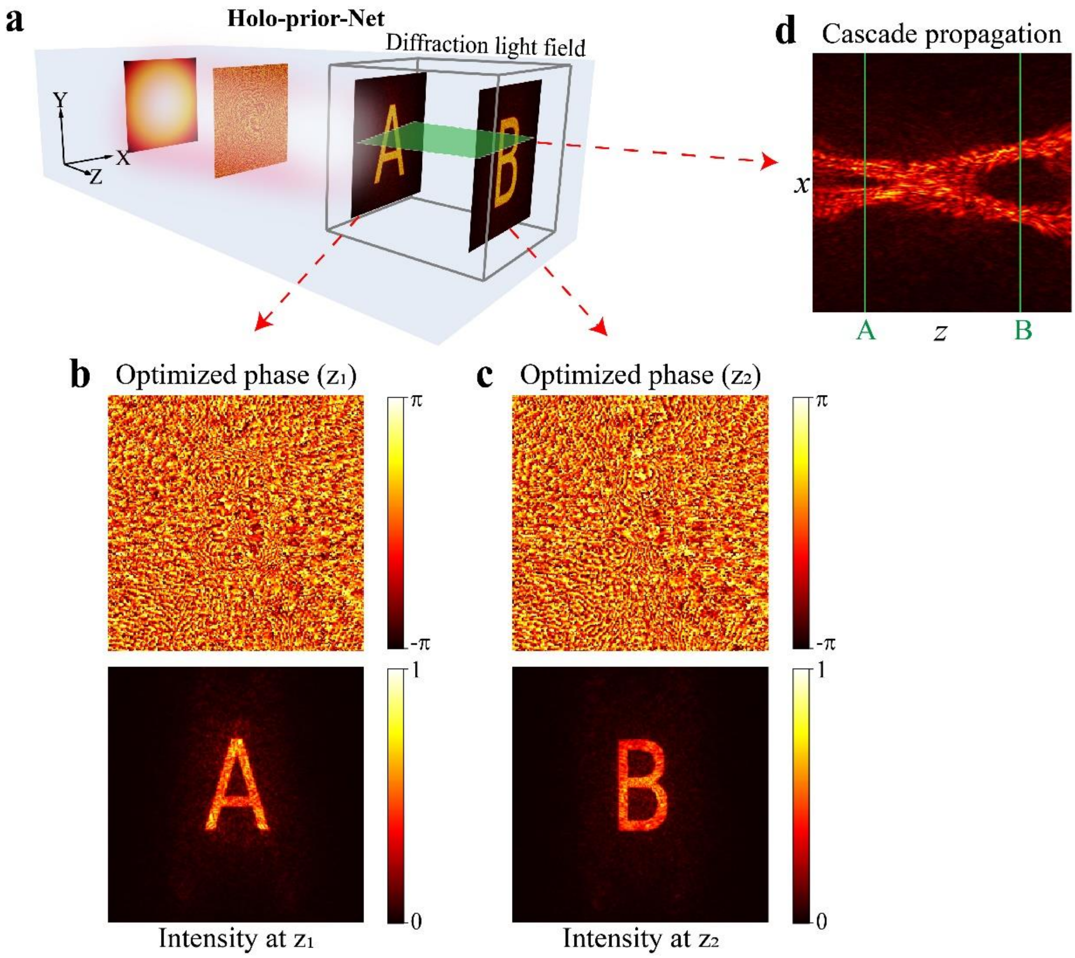

Fig. S17 | Cascade propagation enables constructive reuse of complex optical fields across depth. a, Schematic of the proposed cascade propagation framework, the complex amplitude of 3D diffraction light field is jointly optimized across double depth planes. **b-c**, Simulated intensity distributions and corresponding optimized phase maps at two representative depth planes. Each projection plane is assigned a deliberately optimized phase rather than a random phase, providing an additional degree of freedom to control how the output from preceding planes evolves constructively into subsequent planes. **d,** Illustration of the light field in an XZ cross-section, demonstrating that cascade propagation enables controlled intensity formation, in contrast to conventional layer-wise independent designs where output fields are treated as scattered background or discarded as crosstalk.

Supplementary Note 10: Experimental validation of volumetric continuity

To quantitatively evaluate the volumetric continuity of the projected 3D light field, we first performed an axial scanning experiment using a 3D double-helix structure as an example. The hologram was designed with an axial sampling interval of 1 mm and a total length of 120 mm. In the experiment, the light field was measured along the axial direction with a finer step of 0.2 mm, yielding 700 raw camera frames over a 140 mm range (slightly exceeding the designed 120 mm spiral structure). Notably, this dataset includes both the explicitly designed sampling planes and numerous intermediate slices that were not explicitly sampled during the optimization process.

Axial continuity was quantified using the metric C_{axial} (see definitions in Supplementary Note 14). For each axial slice, the two bright spots forming the double helix were automatically identified, and their region of interest (ROI) intensities were measured. The axial gap fraction was defined as the proportion of slices where the peak intensity fell below a preset threshold. Using a stringent threshold $\tau = 0.5$ (normalized peak intensity), we obtained $C_{axial} = 1$, indicating the absence of intensity gaps across all slices and quantitatively confirming the axial continuity. This result is consistent with physical expectations, as the cascade propagation mechanism ensures smooth evolution of the light field along the axial direction rather than producing gaps due to discrete sampling.

As shown in Fig.S18, the intensity distribution within the ROI for all 700 axial slices. In each slice, the intensity exceeds the threshold, visually confirming the absence of energy gaps along the depth axis. These quantitative findings are consistent with the axial scanning video provided in Supplementary Video 9, which demonstrates the smooth and continuous evolution of the light field in both axial and lateral directions.

It is important to note that the observed continuity does not rely solely on camera reconstruction. To provide direct physical evidence of volumetric 3D continuity, we further demonstrated free-space 3D structures within a scattering medium (Supplementary Videos 6-7), allowing direct, multi-viewpoint observation of continuous 3D light fields with the naked eye.

Taken together, these qualitative and quantitative results confirm that the projected 3D light field exhibits genuine axial and lateral continuity in physical space, rather than arising from discrete plane stacking or camera-based reconstruction. This continuity originates from the cascade propagation design principle, which enables structured and continuous evolution of the light field along the depth direction, rather than relying on discrete layers or numerical interpolation.

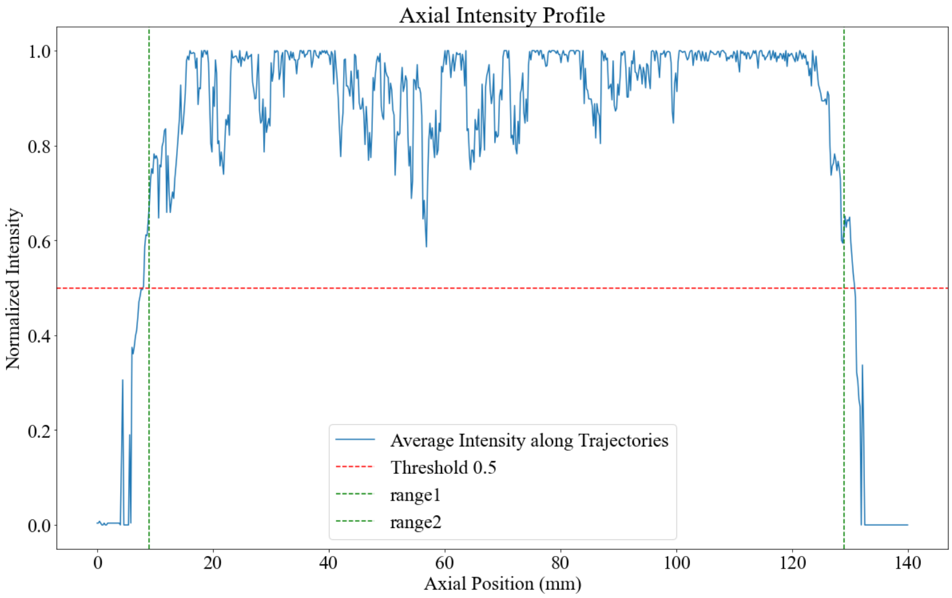

Fig. S18 | Quantification of axial continuity. Within the range of the double-helix structure (10 mm to 130 mm, indicated by the two green lines), the axial scanning measurements yield a continuity metric of $C_{axial} = 1$, At all axial positions, the regional light intensity exceeds the strict threshold of 0.5 (red dashed line), corresponding to zero slices with intensity gaps $N_{gap} = 0$.These results indicate that the experimentally projected light field evolves smoothly and continuously along the axial direction, with no observable discontinuities.

Supplementary Note 11: 4F system for projection result scaling

In our system, collimated light is incident on the hologram loaded onto the SLM, and the projected light field is obtained by direct free-space propagation without the use of optical lenses. Therefore, the projection's aspect ratio of 3D objects can be flexibly adjusted by setting the magnification of the 4F lens system placed after the SLM, enabling more realistic geometric proportions of the reconstructed sphere. We experimentally validated the scaling effects of different magnification rates using the 4F lens system on the projection image. Specifically, we tested effects of uniform scaling, shrinking, and enlarging.

Fig.S19a shows the results with lenses of focal length 100mm, where letters A and B were projected on two planes separated by a spatial depth interval of 10mm. Fig.S19b displays the setup with two lenses having focal lengths of 100mm and 50mm, respectively. In this case, the lateral size of the letters is reduced to half its original size, and the axial depth interval is reduced to 2.5mm, scaled down by a factor of 1/4 after passing through the 4F system. Fig.S19c corresponds to the case where the focal lengths of the two lenses are 50mm and 100mm. In this setup, the lateral size of the projected letters is enlarged to twice the original size, and the axial depth interval is increased to 40mm, four times the original depth interval of 10mm.

Therefore, the spatial dimensions of the 3D objects projected by our system can be flexibly scaled through subsequent optical imaging systems (such as selecting different magnification rates in the 4F system).

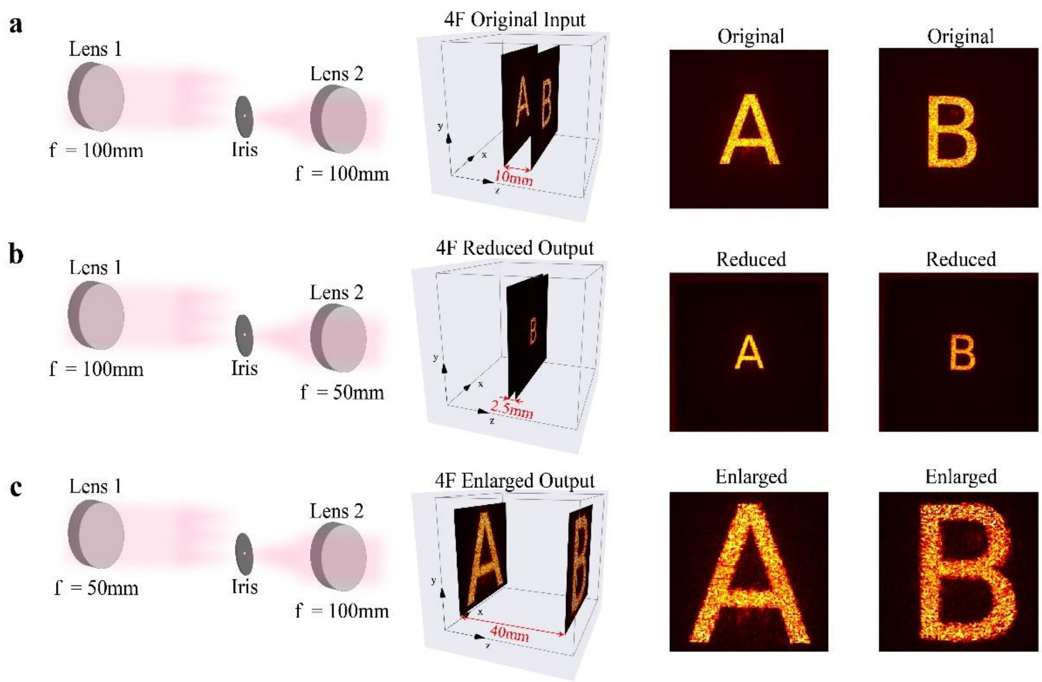

Fig.S19 | Experimental measurement results of projection imaging scaled by a 4F system. a, Equal-scale projection imaging through the 4F system. **b,** Projection imaging scaled down by the 4F system. **c,** Projection imaging scaled up by the 4F system.

Supplementary Note 12: Experimental and Computational Details

Experimental setups

As shown in Fig. S20, the optical layout of the experimental setup, including the sequence of components and system configuration, is presented. The spacing between adjacent holes on the optical platform is 25 mm, which can be used to reference the relative positions of the optical elements. The laser beam first passes through an attenuator (located ~175 mm from the laser output) to adjust the incident intensity on the camera and avoid overexposure. The beam then enters the beam-expanding system: Lens 1 (focal length 4.5 mm) is mounted on a support ~300 mm from the laser output, and Lens 2 (focal length 180 mm) is positioned on a support ~240 mm from Lens 1, with their separation approximately equal to the sum of their focal lengths (184.5 mm), achieving collimated beam expansion. An aperture is placed ~110 mm downstream of Lens 2 to control the beam size entering the subsequent optical path.

The distances between successive elements are as follows: aperture to Mirror 1 ~250 mm; Mirror 1 to Mirror 2 ~400 mm; Mirror 2 to Mirror 3 ~320 mm; Mirror 3 to half-wave plate ~225 mm; half-wave plate to beamsplitter (BS) ~250 mm; BS to SLM modulation plane ~60 mm (adjustable along the z-axis via a three-axis translation stage). The SLM-reflected beam propagates through the BS to the camera, with this segment adjustable according to the design requirements of the projected light field. The camera is mounted on a translation stage (150 mm travel) to allow scanning and reconstruction of the light field.

The Beam splitter directs the expanded beam to illuminate the SLM perpendicularly, and the modulated beam is delivered either to the camera on the translation stage or to the 3D display region. In the current implementation, the projected light field is taken directly from the zero-order diffraction without spatial filtering. Although a conventional 4f system could be used to select the first-order diffraction, we found it unnecessary in our setup. Experimental verification under the current system numerical aperture and sampling conditions showed that first-order light extracted via 4f filtering produces results consistent with those from the zero-order. Therefore, for system simplicity, the zero-order beam is used for projection. If desired, a standard 4f system can be easily incorporated for filtering.

Additionally, Fig. S21 shows close-up images of the experimental setup for camera-based reconstruction and 3D visualization, highlighting the key optical components and beam path configuration.

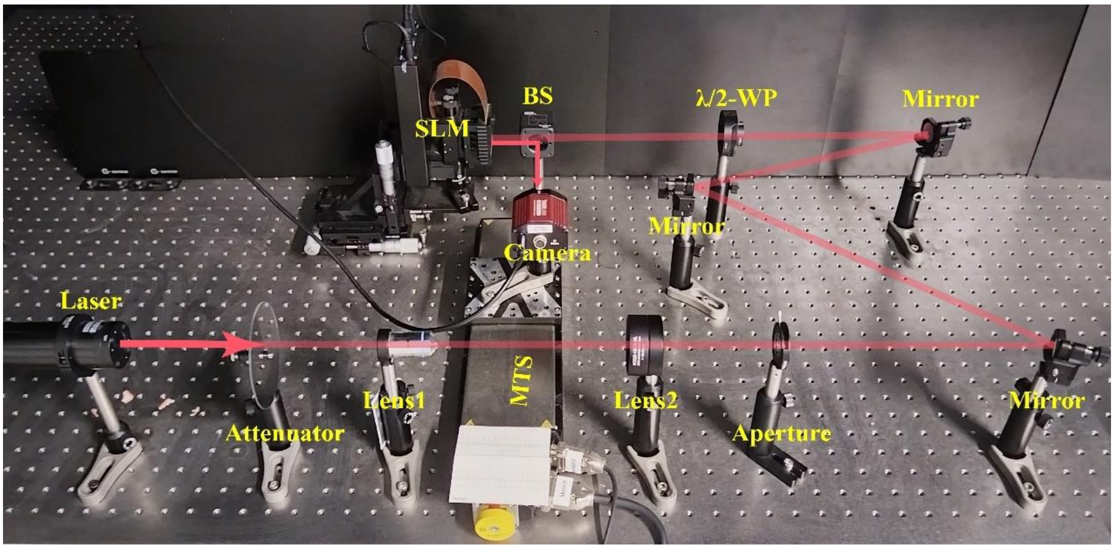

Fig.S20 | Experimental setup. HeNe laser (633nm; Thorlabs: HNL050L). Attenuator, Lens pairs with focal lengths of 4.5mm and 180mm, respectively. Beam splitter (BS, Thorlabs: CCM1-BS013). Half-wave plate(λ/2-WP), Spatial Light Modulator (SLM, Meadowlark Optics: SN5436, 1920 × 1152 pixels, pixel pitch 9.2μm). CMOS camera (Thorlabs: CS505MU). motorized stage (MTS, Oriental Motor, travel range: 150mm).

a **b**

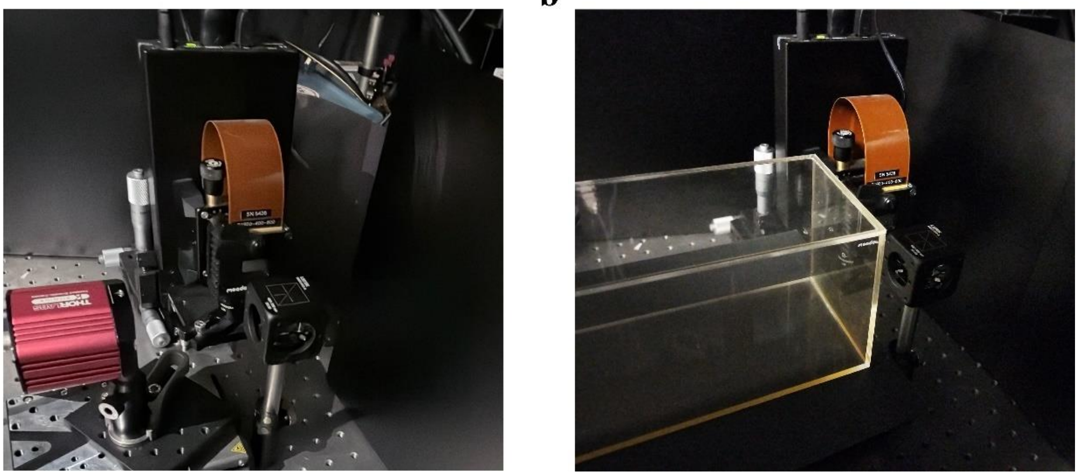

Fig. S21 | Experimental setup for reconstruction and 3D display. a, The light field is sequentially captured using a CMOS camera to reconstruct the 3D projection. **b**, 3D visualization of the projected light field in a transparent acrylic box containing a small amount of scattering smoke, enabling multi-angle observation of the free-space 3D structure (see Supplementary Videos 6-7).

SLM calibration

Before all experiments, the phase response of the SLM was calibrated to ensure accurate pure-phase modulation. The calibration followed the vendor-recommended diffraction efficiency method: by loading grating patterns covering gray levels 0-255, the intensity of the first-order diffracted light was measured to characterize the SLM's phase-drive response. The calibration optical setup is shown in Fig. S22. A HeNe laser, after beam expansion, passed through a half-wave plate to adjust the polarization and was incident on the SLM at a small oblique angle. The SLM-reflected beam was focused by a lens, with a pinhole placed near the focal plane to isolate the +1 diffraction order, and a power meter was used to record the diffracted intensity for each gray level.

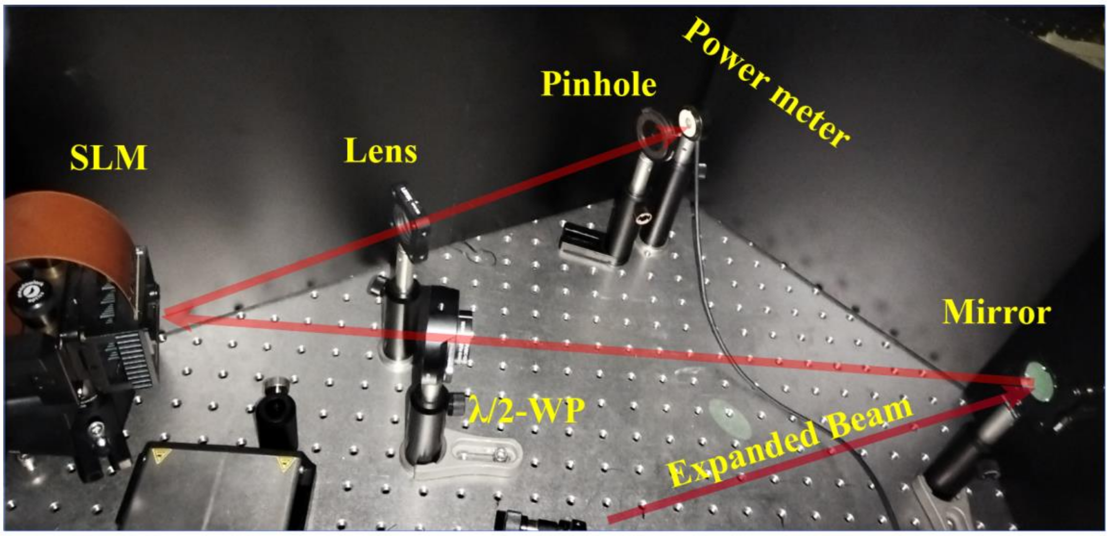

Fig.S22 | Optical layout for SLM phase calibration.

To ensure measurement stability, the experiments were conducted in a darkroom to avoid ambient light interference. Using the vendor-provided Blink software (Meadowlark Optics), a linear initial LUT (black.bmp wavefront calibration file) was loaded. As illustrated in Fig. S23, within the Pattern Generation module, stripe patterns were selected (Stripe mode, 8 stripes) to optimize diffraction spot spacing. For the SLM used in this work (1920 × 1152 pixels, Meadowlark Optics SN5436), patterns with gray level differences from 0 to 255 (8-bit) were sequentially loaded, yielding 256 intensity measurements, which were recorded in a spreadsheet (first column: gray level, second column: corresponding intensity).

After data acquisition, the measurements were exported in CSV format and imported into the vendor-provided DiffractiveLUT software. The software converted the measured diffraction efficiencies into phase responses, generating the phase − drive curve and producing a linearized look-up table (LUT). This LUT defines the mapping between the image gray levels and the corresponding phase modulation. A properly linearized LUT ensures that gray levels 0-255 (8-bit) correspond to phase modulation from 0 to 2π. As shown in Fig. S24, the measured phase response and the LUT-corrected linearized mapping. The generated LUT was applied to all holographic phase patterns in this work to compensate for the SLM's intrinsic nonlinear response.

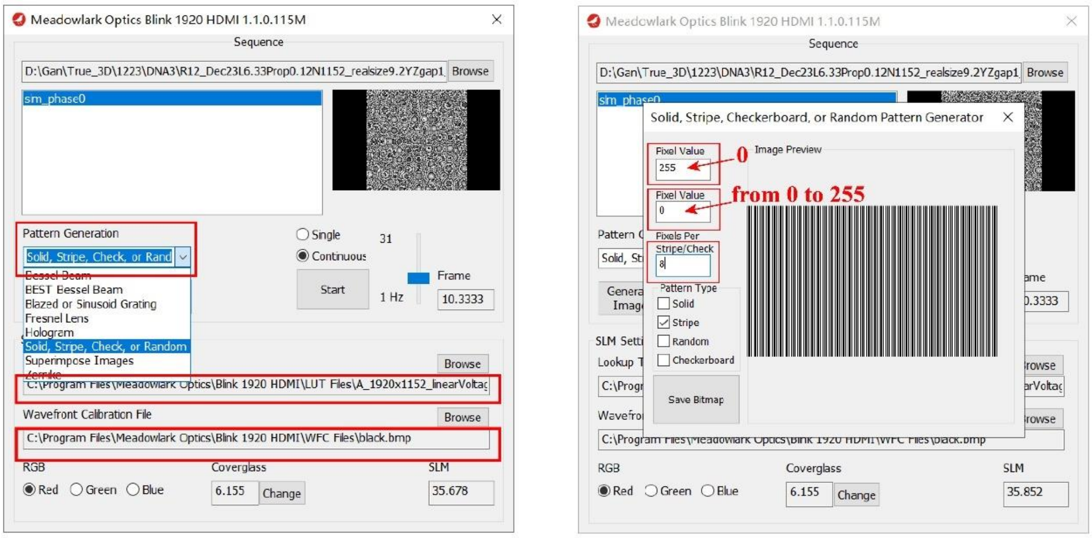

Fig.S23 | **a**, Blink software interface and SLM settings. **b**, Generation of calibration patterns using Blink for the SLM used in this work (Meadowlark Optics, SN5436; 1920 × 1152 pixels).

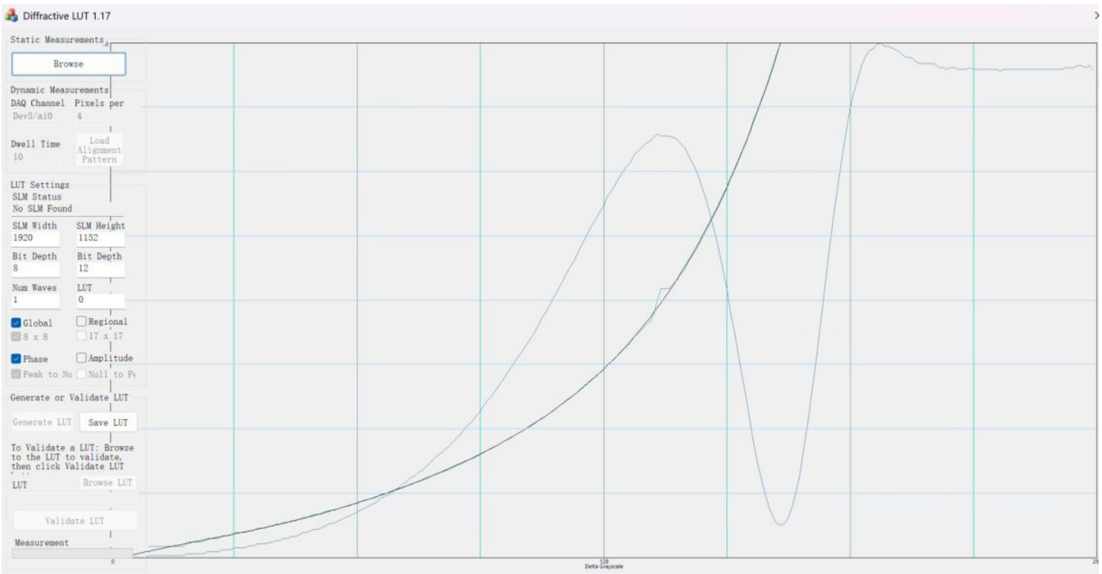

Fig.S24 | Phase calibration and LUT generation for the SLM, obtained from measured intensity variations of the grating patterns.

Algorithm pseudocode

Algorithm: Realization of Cascaded Propagation Design Paradigm

```
----------------------------------------------Hyperparameters----------------------------------------------
    - Wavelength:                       λ
    - Pixel pitch(SLM):                 Δp
    - Input of diffractive system:      Fixed illumination
    - Target 3D volume:                 Target_volume_matrix (Nx,Ny,Nz)
    - Axial sampling interval:          ΔZ
   - Number of diffractive layers:      num_layer
   - Projection Distance:               distance
    - learning_rate:                    lr = 1
    - epochs = 100

---------------------------------------------- Holo-Prior-Net ----------------------------------------------
# Initialize untrained network parameters
          ϕ(x,y) ← random phase pattern [-π, π]

# Architecture of Holo-Prior-Net ← Cascaded propagation
function Holo_Prior_Net( Nx, Ny,Nz, ΔZ, num_layer):

    #Input layer
    E_in ← Fixed Illumination(type = plane_or_gaussian, beam_waist ≥ SLM_aperture)

    #Diffractive layer(s)
    E ← E_in
    ϕ(x,y)=tf.Variable(tf.random.uniform(shape=[N1,N2],minval=-tf.constant(np.pi),
                          maxval=tf.constant(np.pi), dtype=tf.float32))
    num_layer =1 # Corresponding to a single SLM
     For ℓ = 1 to num_layer do
        # Apply layer phase modulation (w_i^l · exp (jϕ_i^l) · E_input_i^l  in Eq.(5) )
        E ← ApplyDiffractionLayer(E, ϕ(x,y))

        # Propagate field to next layer
        E ← FreeSpacePropagate(E, distance, λ, Δp)
      (FreeSpacePropagate denotes scalar diffraction propagation consistent with the Rayleigh-
       Sommerfeld formulation)

     #Output layers (mapping to 3D projection)
    # Axial cascaded propagation generates continuous 3D volume
    V_pred ← []
    E_projection ← E
     for k = 1 to Nz do
    # Cascaded propagation: output of current layer is input to next
```

```
        E_projection ← FreeSpacePropagate(E_projection, ΔZ , λ, Δp)
            V_pred.append(E_projection)
    end for

    return V_pred
end function
```

--- Optimization ---

```
V_pred = Holo_Prior_Net( Nx, Ny,Nz, ΔZ, num_layer)
Output_projection_volume = | V_pred | **2

#Criterion
loss ← MeanSquaredError( Output_projection_volume , Target_volume_matrix)

optimizer = tf.train.AdamOptimizer(learning_rate=lr ))

#Backpropagation and parameters update
# Gradients are computed with respect to the diffractive layer phase
grads_and_vars ← ComputePhaseGradients(loss, ϕ(x,y))

# Update diffractive layer phase parameters
training_op = optimizer.apply_gradients(grads_and_vars)
# saved diffractive layer phase
weights_op = tf.get_collection(tf.GraphKeys.TRAINABLE_VARIABLES)

init = tf.global_variables_initializer()
with tf.Session() as sess:
    sess.run(init)
    for epoch in range(epochs):
        _ = sess.run(training_op)
    end for

    # learnable parameters of the Holo_Prior_Net
    Optimized_weights = sess.run(weights_op)

normalized_phase ← Optimized_weights
cv2.imwrite( 'SLM_img.bmp', normalized_phase / normalized_phase.max() * 255)
```

Supplementary Note 13: Speed and scaling characterization

We explicitly measured and distinguished the computation times for forward propagation, backward propagation, file I/O, and SLM loading. Table 1 lists the detailed times for designing holograms for 100 projection planes at different resolutions. Forward propagation includes the time for loss function evaluation, while backward propagation accounts for parameter updates. To illustrate the relative contributions of each component, a stacked bar chart is provided in Fig. S25. From Table 1 and Fig. S25, the total training time and the proportion of forward and backward propagation at different resolutions are clearly shown. Overall, forward propagation time increases faster than backward propagation with resolution, indicating that the computational overhead at high resolution is dominated by the diffraction calculations in forward propagation.

Resolution	Total Time (s)	Forward (s / %)	Backprop (s / %)
128×128	15.37	5.06 / 32.9	10.31 / 67.1
256×256	20.38	7.82 / 38.4	12.56 / 61.6
512×512	38.24	19.68 / 51.5	18.56 / 48.5
1152×1152	62.02	47.53 / 76.6	14.49 / 23.4
1920×1152	89.02	69.10 / 77.6	19.92 / 22.4

Table 1 | Runtime breakdown of the hologram optimization pipeline for a fixed number of 100 depth planes at different pixel resolutions. Forward pass time includes loss computation, and backpropagation time includes parameter update.

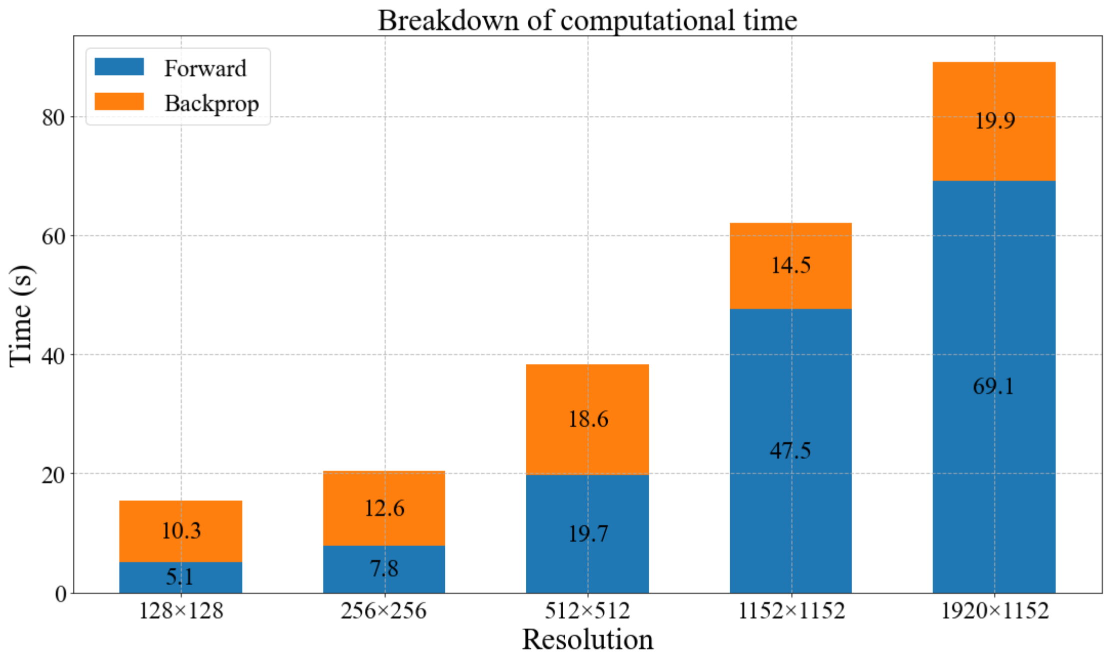

Fig.S25 | Breakdown of computation time for designing holograms of 100 projection planes at different resolutions. Forward propagation includes loss evaluation, while backward propagation includes parameter updates. Each bar represents the total computation time at a given resolution, with blue indicating forward propagation and orange indicating backward propagation. As the resolution increases, forward propagation dominates the total computation time.

In our experiments, all 3D projection examples were generated using 1152 × 1152-pixel holograms, for which the design of 100 sampled projection planes required ~62 s. For full 1920 × 1152 resolution, the time increased to ~89 s. File I/O time is not incurred at every optimization iteration but occurs once after design completion, averaging 0.018 s. SLM loading time is mainly determined by the SLM refresh rate (Meadowlark Optics, SN5436, 31 Hz), corresponding to ~0.032 s per loading.

We further systematically evaluated the scaling of computation time with the number of planes and pixel count up to the native SLM resolution. As shown in Figs. S26 and S27, computation time scales approximately linearly with both plane number and pixel count, with no observed superlinear growth, indicating the method remains computationally feasible for higher space-bandwidth products.

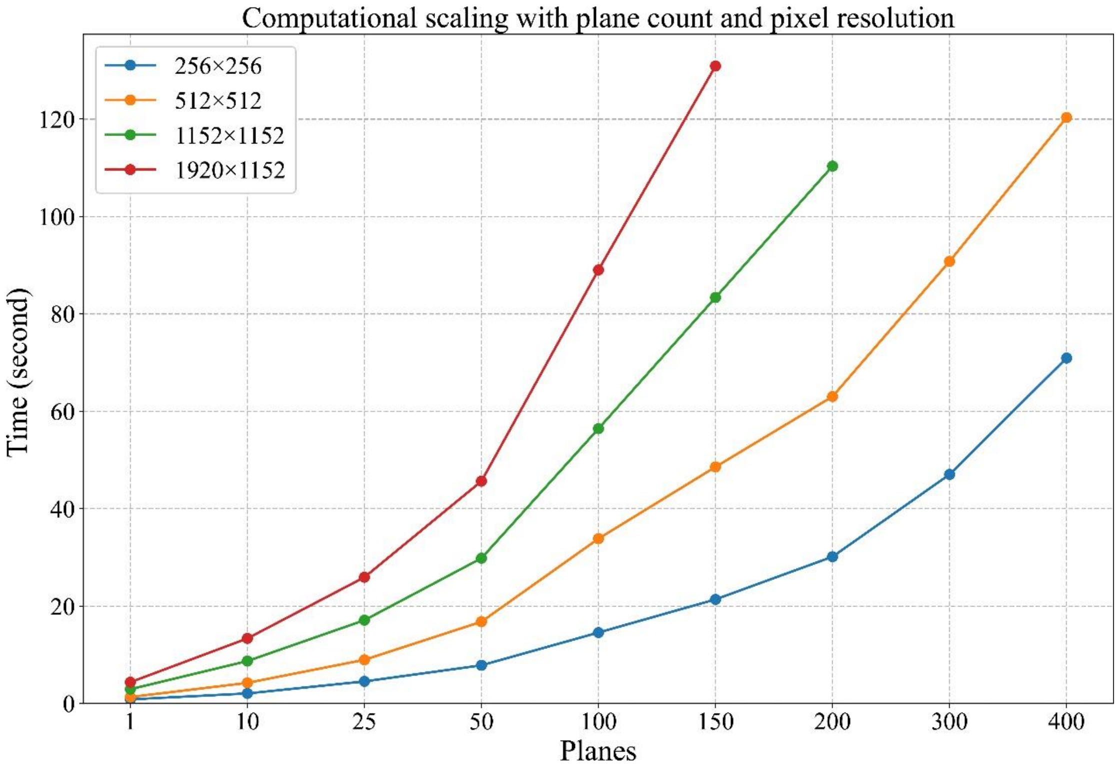

Fig.S26 | Coumputational scaling with plane count. Computation time for hologram optimization as a function of the number of depth planes, evaluated at four pixel resolutions: 256×256, 512×512, 1152 ×1152, and 1920×1152.

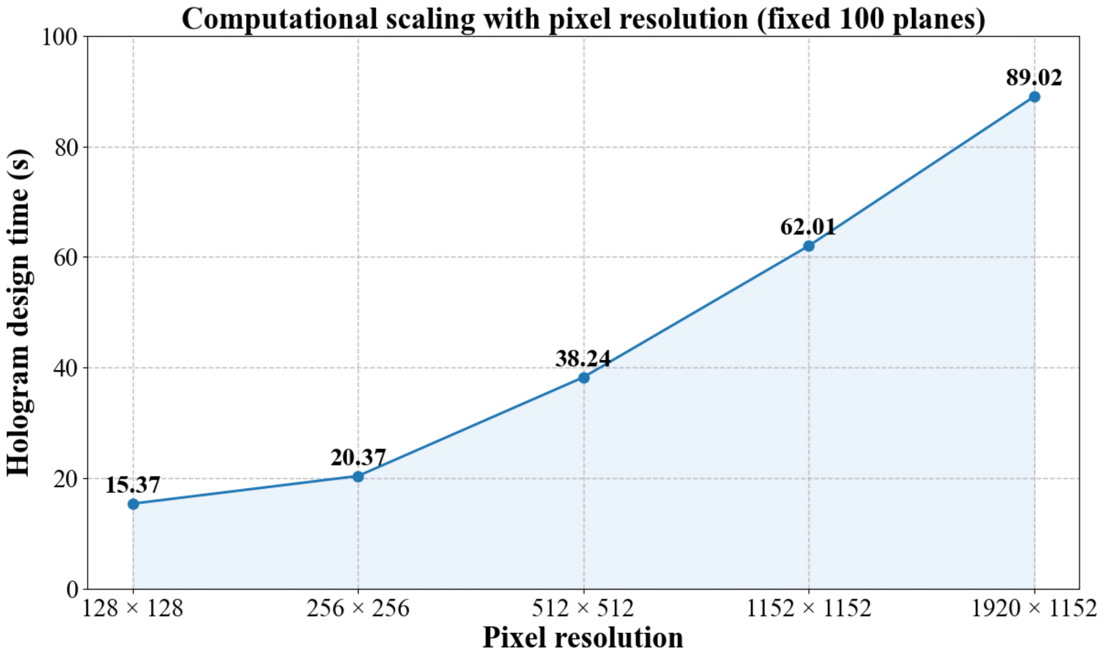

Fig.S27 | Coumputational scaling with pixel resolution. Computation time for hologram optimization as a function of the number of pixel resolution up to the native SLM resolution, evaluated at 100 planes.

We also quantified the peak GPU memory usage as a function of spatial resolution for a fixed number of depth planes. As shown in Fig. S28, memory usage increases monotonically with pixel count, as expected, without any explosive growth, demonstrating predictable and manageable scalability. All computations were performed on a single GPU (NVIDIA GeForce RTX 3090), and higher-performance GPUs are expected to further improve computational efficiency.

To illustrate the iterative behavior and convergence of the optimization, a representative loss curve as a function of iteration number is shown in Fig. S29. The proposed 3D loss function typically stabilizes after 80–100 iterations. Accordingly, a 100-iteration early-stop strategy was employed for the 3D hologram designs presented in this work. Under these conditions, the 3D hologram design time is ~62 s per example.

The reported computation speed does not rely on any precomputation or warm-start initialization. The learnable phase on the SLM is always initialized with random values, without reference to previous holograms, pretraining, caching, or other warm-start strategies. Therefore, the optimization time reflects the true cost of each 3D hologram design.

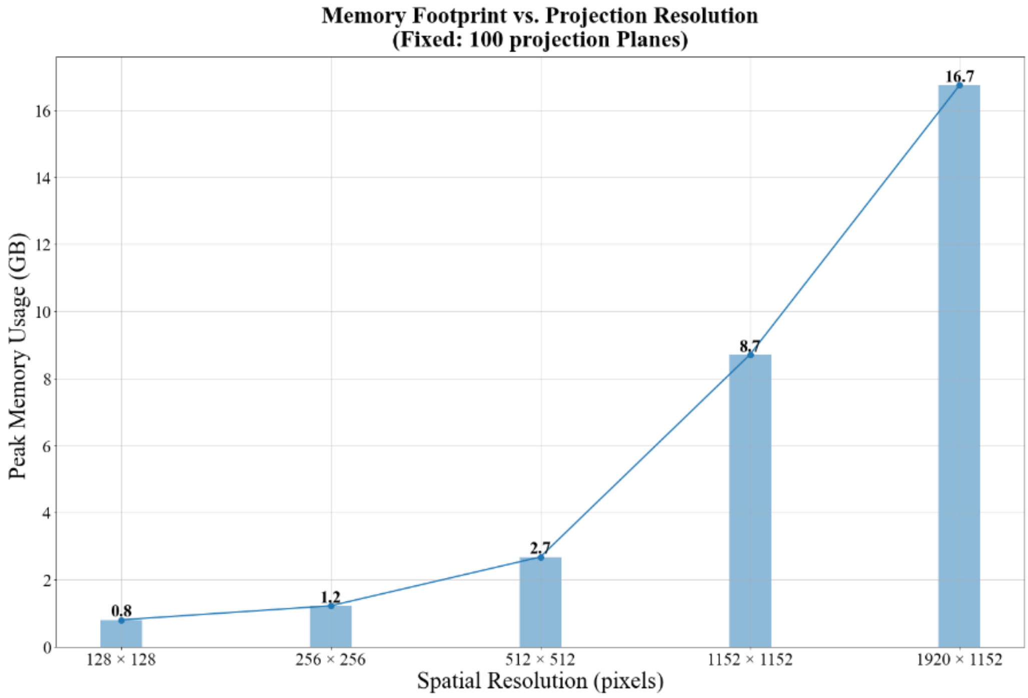

Fig.S28 | Memory footprint scaling with pixel resolution.

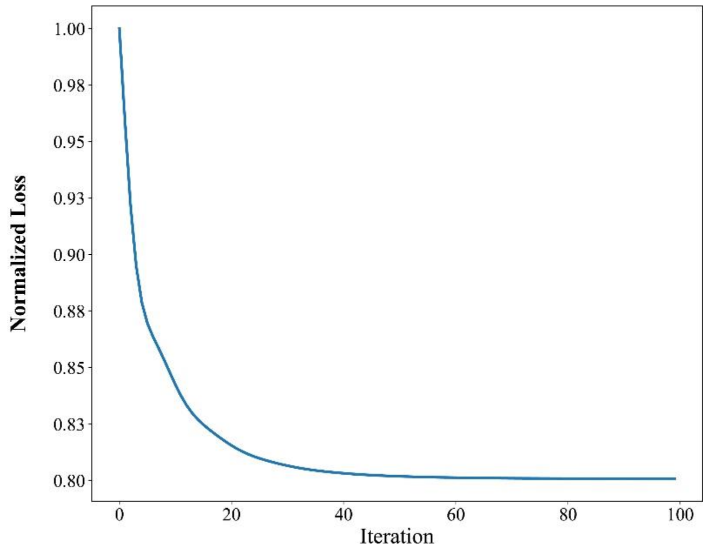

Fig. S29 | Convergence behavior. The loss decreases smoothly and saturates after 80 to100 iterations.

Supplementary Note 14: Quantitative Metrics

In this work, the continuity refers to the continuous three-dimensional physical light field generated in the actual projection, rather than numerical sampling continuity in hologram computation. The continuity of the proposed method originates from a fundamentally different design paradigm. In conventional layer-wise independent design hologram (e.g., Fig. S2), the slices between projection planes consist primarily of unstructured speckles, requiring stacking of independent layers with spacing, which inherently produces discontinuities. In contrast, our method treats the light field of each layer as a physically meaningful input to subsequent layers, achieving a physically continuous volumetric light field rather than a stack of independent planes, thus avoiding discontinuities inherent to conventional design principles.

Axial continuity is defined as the absence of intensity gaps above a threshold across consecutive slices within a region. The intensity gap is defined as $I(z_i) < \tau \cdot \max(I)$, indicating regions where the intensity falls below the threshold τ. Axial continuity is quantified via the gap fraction, where N_{gap} is the number of slices containing gaps, and N is the total number of slices. C_{axial} closer to 1 indicates stronger axial continuity. The axial continuity quantification index is as follows:

$$N_{gap} = \sum_{k=1}^{N} 1 \times [I(z_k) < \tau \cdot \max(I)] \tag{2}$$

$$C_{axial} = 1 - \frac{N_{gap}}{N} \tag{3}$$

Crosstalk is defined by the crosstalk ratio as the ratio of leaked optical energy in non-target regions ($E_{background}$) to the energy in the target region(E_{signal}):

$$Crosstalk\ \text{ratio} = \frac{E_{background}}{E_{signal}} \tag{4}$$

We further define the contrast-to-noise ratio (CNR) as:

$$CNR = \frac{\mu_{signal} - \mu_{background}}{\sigma_{background}} \tag{5}$$

where μ_{signal} and $\mu_{background}$ are the mean intensities of the signal and background regions, respectively, $\sigma_{background}$ is the standard deviation of the background. CNR quantifies the visual distinguishability of the reconstructed signal against background fluctuations caused by noise and crosstalk.

For all metric calculations, the ROI refers to the target pattern region in the ideal image, and the mask is its binary representation, used to extract signal and noise pixels. Both ROI and mask are automatically generated from the ideal target image, ensuring objectivity. The noise region comprises all pixels outside the ROI. In simulations, dynamic ranges are normalized to 0-255, while experimental data use the camera's native dynamic range (0-4095, 12-bit). All metrics are computed without subtracting stray-light floors, ensuring conservative and reproducible evaluation.

Contrast calculations are based on the measured intensity I from the CMOS camera, defined as follows:

$$Contrast = 10 \log_{10} \frac{I_{min}}{I_{max}} \quad (6)$$

For SNR, the signal power is the mean square intensity within the target region, and noise is defined as the pixel-wise difference between the measured image and the target image in the background region. Noise power is computed as the mean square of these differences:

$$SNR = 10 \log_{10} \frac{P_{signal}}{P_{noise}} \quad (7)$$

Supplementary Note 15: Robustness and Sensitivity Analyses

sensitivity to z-registration errors

To experimentally assess the robustness of the proposed method to z-registration errors, we measured the standard contrast-to-noise ratio (CNR) of two projection images as a function of axial position. As shown in Fig. S30, within a 20 mm measurement range, the CNR of two projection planes (A and B) was evaluated across 200 axial scan positions. Distinct CNR peaks were observed at the preset depths of A and B (separated by 10 mm), while CNR gradually decreased away from the target planes. These results indicate that the proposed method exhibits tolerance to axial misalignment.

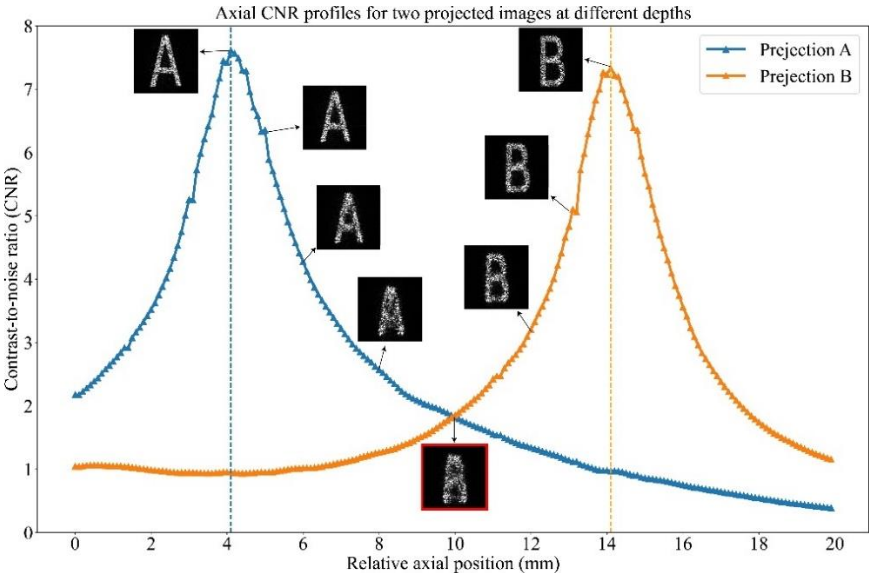

Fig. S30 | Axial CNR curves under axial position misalignment. Distinct CNR peaks are observed at the preset depths of A and B (separated by 10 mm), while CNR decays smoothly away from the target planes. These results indicate robust axial selectivity and tolerance to axial plane misalignment in the proposed method.

sensitivity to SLM phase noise:

We further evaluated robustness by explicitly introducing random phase noise on the SLM. Uniformly distributed phase noise with amplitudes of 0.1π, 0.2π, 0.3π, 0.4π, and 0.5π was added to the optimized holograms. For each noise level, we measured the reconstructed light fields at representative axial positions corresponding to depth planes A and B. As shown in Fig. S31a, the reconstructed geometry remained intact across all tested noise levels, with no noticeable structural distortion or loss of target features, consistent with the physical expectation that random phase perturbations primarily reduce interference efficiency. Increased phase noise led to a gradual reduction in reconstruction brightness. To quantify this effect, we introduced the normalized total intensity (NTI), defined as the ratio of total intensity in the target region to that of the ideal reference image. For both depth planes (Fig. S31b), NTI decreased monotonically with increasing phase noise, reflecting progressive energy loss due to reduced phase coherence. These results demonstrate that the proposed method is robust to moderate phase noise.

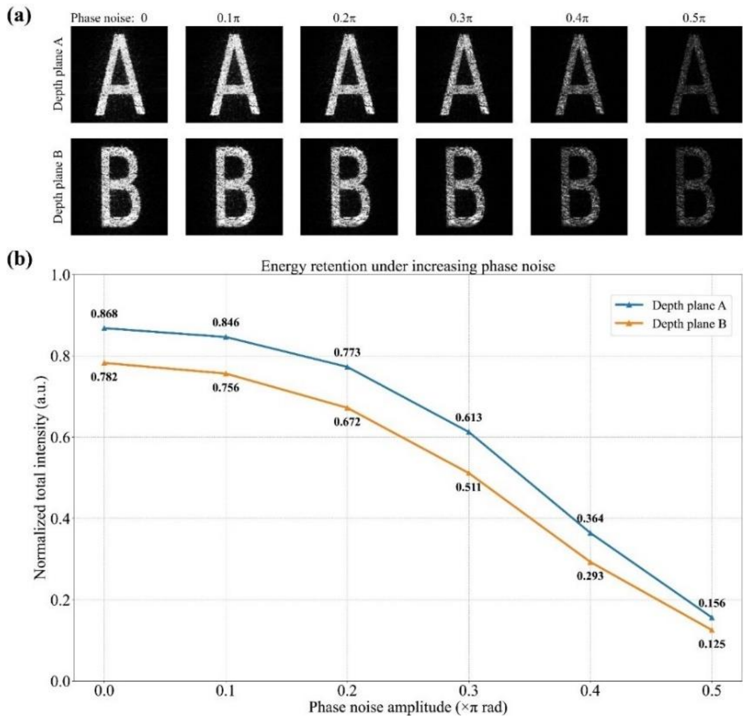

Fig. S31 | Robustness to SLM phase noise. (a) Experimental reconstructions at two representative depth planes (A and B) under increasing random phase noise amplitudes from 0 to 0.5π. The geometric structure of the reconstructed images remains well preserved across all noise levels. (b) Normalized energy retention within the target region as a function of phase noise amplitude, showing a monotonic decrease due to reduced coherent interference efficiency.

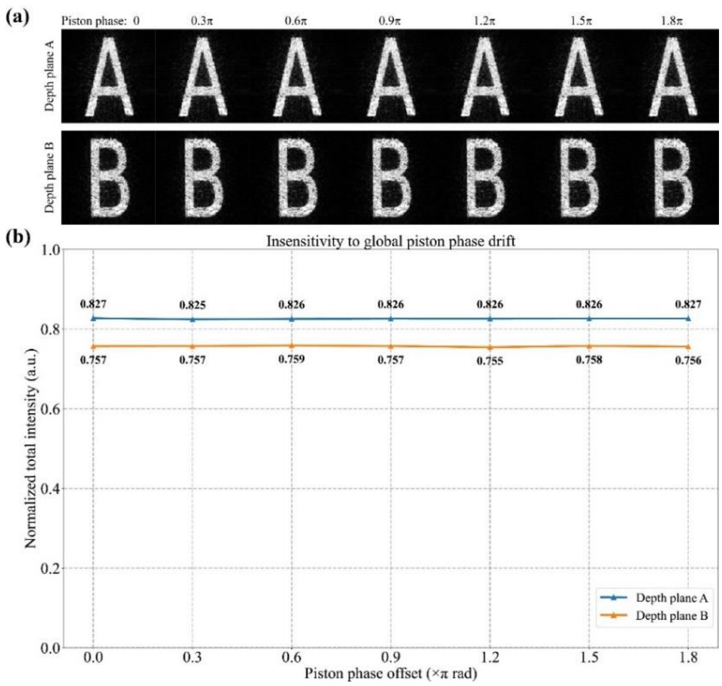

Fig. S32 | Robustness against piston phase drift. (a) Experimental measurement of intensity distributions at two representative depths (A and B) under increasing piston phase drift applied to the SLM. (b) Normalized total intensity (NTI) within the target regions as a function of piston phase drift.

sensitivity to Piston drift:

We also introduced global piston phase shifts on the SLM. As shown in Fig. S32, despite significant global phase offsets, the reconstructed image structures remained visually unchanged. Quantitative analysis of NTI within the target region showed negligible variation across tested piston drift levels. This behavior is expected, as a global piston corresponds to a spatially uniform phase shift, which does not alter the phase gradient or spatial frequency content of the wavefront and therefore does not affect free-space propagation or intensity formation. These results confirm that the proposed method is inherently insensitive to global piston phase errors.

sensitivity to wavelength detuning:

To evaluate robustness against wavelength detuning, holograms were designed at different wavelengths (627-639 nm), while the reconstruction was always performed with a fixed 633 nm laser. This experiment probes the effect of wavelength mismatch between design and reconstruction. As shown in Fig. S33, reconstructions at both representative depths (A and B) remained recognizable under all tested detuning conditions, although image quality showed slight blurring. CNR was computed for each wavelength offset (Δλ = 0、±2、±4、±6 nm). As expected, CNR decreased gradually with increasing wavelength mismatch, indicating the method's tolerance to detuning without catastrophic structural degradation. The gradual degradation of CNR with increasing wavelength mismatch quantitatively characterizes the tolerance of the proposed method to wavelength detuning.

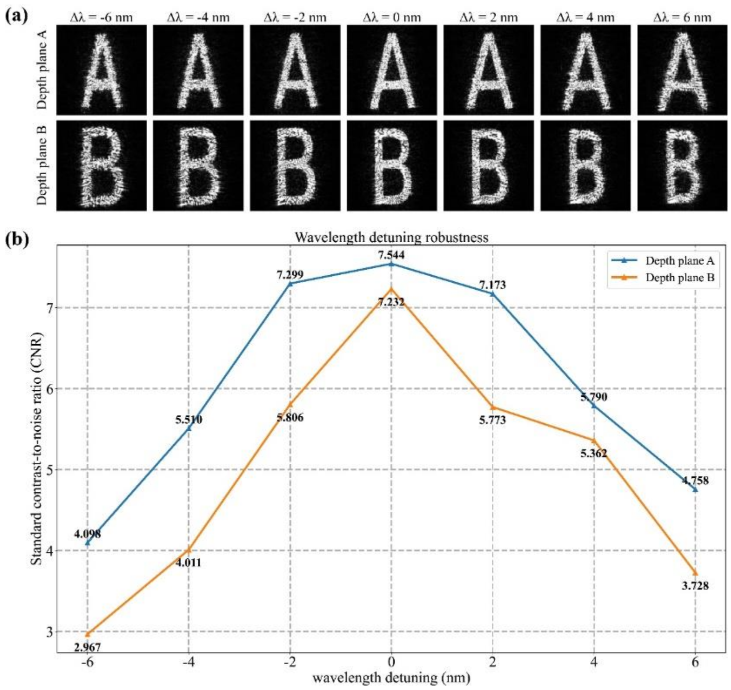

Fig. S33 | Robustness against wavelength detuning. (a) Experimental reconstructions at two representative depths (A and B) obtained using a fixed experimental laser wavelength of 633 nm, while the holograms are designed assuming different wavelengths (627~639 nm). (b) Standard contrast-to-noise ratio (CNR) as a function of wavelength detuning (Δλ = 0, ±2, ±4, ±6 nm).

coherence length:

In all demonstrated continuous 3D holographic projections, the interference and continuity of the light field remained stable throughout the experimentally probed volumetric region, with no discernible limitations attributable to coherence length. While the potential influence of coherence length at significantly larger volumes represents an interesting direction for future studies, such investigations are beyond the scope of the present work.

sensitivity to angular tilt misalignment:

Experimentally, we intentionally introduced controlled tilts along the y-axis on the SLM to evaluate angular robustness. Tilts were applied by rotating the SLM adjustment screws clockwise in six full turns, measuring images at different depths after each rotation, followed by reversing the screws counterclockwise to return to near-initial alignment. As shown in Fig. S34a, reconstructed structures at representative depths A and B remained intact under all tested tilt conditions, although the image positions shifted laterally. Fig. S34b quantifies the y-axis pixel displacement, which increased monotonically with tilt and exhibited clear depth dependence. Upon reversal of tilt, images returned to their original positions, demonstrating that the system tolerates angular misalignment without irreversible degradation. Simulations further confirmed this behavior: Fig. S35 shows reconstructed continuous 3D helical curves under tilts of 0°, -0.2°, and 0.3° from three viewpoints. Despite tilting, the helices remained continuous and structurally intact, with predictable geometric displacements and no structural degradation, consistent with experimental observations.

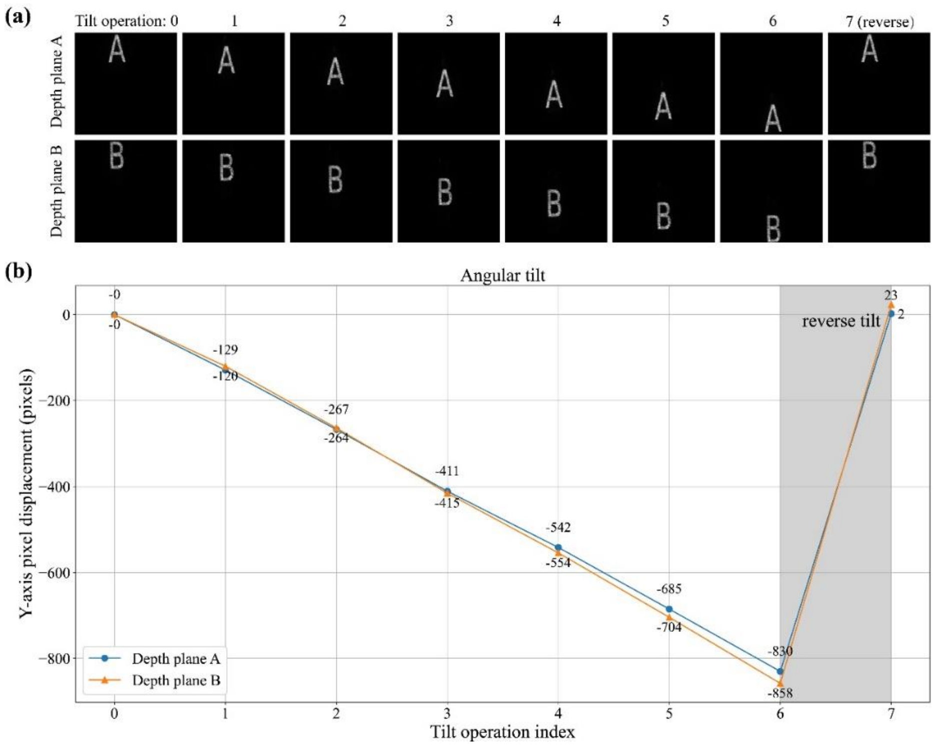

Fig. S34 | Robustness against angular tilt misalignment. a. Experimental reconstructions at two representative depths (A and B) under increasing angular tilt applied to the SLM along the y direction. Each tilt step（tilt operation index 1-6）corresponds to one full clockwise rotation of the SLM tilt adjustment screw. The final frame is obtained by reversing the tilt direction with six full counterclockwise rotations (tilt operation index 7). b. Measured y-axis pixel displacement of the reconstructed images at depths A and B as a function of the tilt operation index. While angular tilt induces a systematic lateral shift that depends on depth, the reconstructed image structures remain intact and recover to their original positions after reverse tilt, demonstrating strong robustness against angular misalignment.

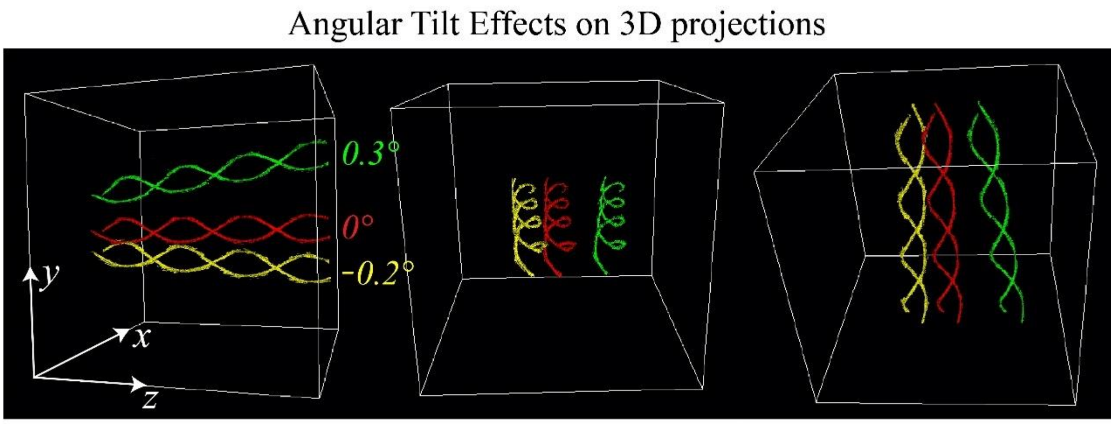

Fig. S35 | Simulated robustness of 3D helical structures under different tilt angles. Visualization of continuous helical holograms with applied tilts along the y-axis, shown from multiple viewpoints to highlight the effect of tilt on 3D projection.

Impact of pixel pitch

We investigated the effect of SLM pixel pitch (4μm-18.4μm) on projection quality while keeping hologram size and other optical parameters fixed. Simulated reconstructions of two depth planes (A and B) were analyzed using the crosstalk ratio and CNR. The crosstalk ratio is defined as the ratio of non-target energy in background regions to the target signal energy. The hologram's wavefront complexity can be quantified via the space-bandwidth product (SBP) [6]:

$$\mathrm{SBP} = \mathrm{N} \times \mathrm{N} = \left(\frac{L}{d}\right)^2 \tag{8}$$

where L is the hologram length,, d is the pixel pitch,, N is the number of pixels. Smaller pixel pitches (d) yield higher SBP, allowing finer image features and improved reconstruction quality.由

As shown in Fig. S36, smaller pixel pitches produced sharper images with cleaner backgrounds at both depth planes. Increasing pixel pitch (decreasing SBP) led to reduced axial resolution and increased energy leakage. Quantitative metrics showed monotonic trends: as pixel pitch increased from 4 μm to 18.4 μm, crosstalk ratio increased while CNR decreased, reflecting the loss of controllable spatial degrees of freedom. This monotonic behavior is consistent across qualitative reconstructions and quantitative metrics, demonstrating that pixel pitch is a fundamental and predictable factor affecting the robustness of continuous 3D holographic projection. These results provide practical guidance for selecting hologram pixel pitch under SBP constraints.

Taken together, these tests demonstrate that the proposed holistic 3D holographic design strategy exhibits predictable and physically interpretable robustness characteristics. They also clarify the practical limits and operational stability of the framework, indicating that it can be reliably deployed under realistic SBP constraints and alignment tolerances.

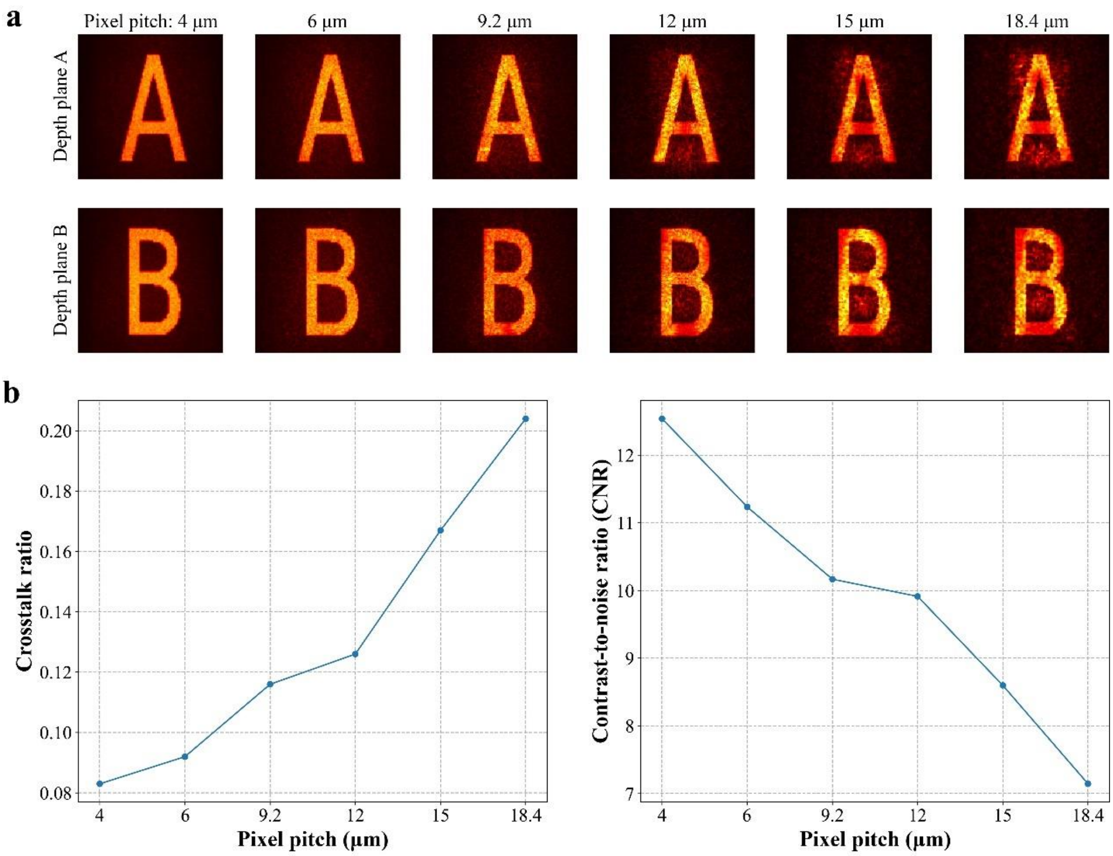

Fig. S36 | Impact of pixel pitch on crosstalk ratio and CNR. a, Simulated reconstructions of two depth planes (A and B) for pixel pitches ranging from 4 µm to 18.4 µm. b, Quantitative crosstalk ratio (left) and CNR (right) as a function of pixel pitch. Increasing pixel pitch leads to higher energy leakage and crosstalk, and reduced contrast.

Supplementary Note 16: Alternative Loss Functions for Reference

In this work, a single MSE loss was used successfully for all 3D projection examples presented. We note that the form of the loss function can be flexibly adapted for different tasks or specific reconstruction targets. For completeness, we provide below alternative loss function forms associated with particular reconstruction objectives; these are offered as reference and are not a required component of the method presented here.

We can use the mean squared error (MSE) loss function to compute the error between the overall 3D diffracted light field and the target 3D scene, introducing constraints on the 3D light field. Additionally, we incorporated constraint conditions to maximize the energy in the target region, ensuring that the total energy of each plane from z=2 to z=N closely matches the energy of the initial plane. The objective function is formulated as follows:

$$\mathrm{L}_{\mathrm{3D\_MSE}} = \frac{1}{X \times Y \times Z} \sum_{z=1}^{Z} \sum_{x=1}^{X} \sum_{y=1}^{Y} \left| \left| V_{pred}(x,y,z) \right|^2 - \left| V_{target}(x,y,z) \right|^2 \right|^2 \tag{9}$$

$$\mathrm{L}_{\mathrm{EMax}} = -\lambda_2 \sum_{z=1}^{Z} \sum_{(x,y)\in\Omega_z} \left| E_{pred}(x,y,z) \right|^2 \tag{10}$$

$$\mathrm{L}_{EConst} = \lambda_3 \sum_{z=2}^{Z} \left| \sum_{(x,y)\in\Omega_z} \left| E_{pred}(x,y,z) \right|^2 - \sum_{x=1}^{X} \sum_{y=1}^{Y} \left| E_{pred}(x,y,1) \right|^2 \right|^2 \tag{11}$$

$$\mathrm{Loss} = \lambda_1 \mathrm{L}_{\mathrm{3D\_MSE}} + \lambda_2 \mathrm{L}_{\mathrm{EMax}} + \lambda_3 \mathrm{L}_{EConst} \tag{12}$$

In this context, X、Y and Z represent the dimensions of the 3D light field. $\Omega_z$refers to the imaging region of the target image on the plane at z. $\mathrm{L}_{\mathrm{3D\_MSE}}$ calculates the MSE error between the overall 3D diffracted light field $V_{pred}(x,y,z)$ and the target 3D scene $V_{target}(x,y,z)$, ensuring that the network output approaches the target 3D light field. $\mathrm{L}_{\mathrm{EMax}}$ represents energy maximization within the target imaging region, constraining the light field energy to be distributed as uniformly as possible within the valid imaging region. L_{EConst} (Energy Constraint) is used to calculate the total energy across each projection plane, ensuring that the total energy of the light field in each plane's effective imaging region closely matches the total energy of the initial plane. This constraint prevents the scattering of light rays in a random manner, guiding the network framework to optimize the usage of emitted light from each projection plane as input for the subsequent layer's imaging. The error function used is composed of the three aforementioned terms, each with a weighting coefficient λ_1、λ_2 and λ_3, which can be adjusted to balance the influence of these loss functions. The weight coefficient λ_1 can be set to 1, while λ_2 and λ_3 are both set to 0.1.

Supplementary Note 17: Comparisons under smoke environment

In the comparative experiments, the camera position and focus plane were kept fixed, and holograms corresponding to the proposed method and RV-CGH were alternately displayed in video acquisition mode over short intervals. The imaging focus plane and incident light power were maintained identical, and the short-time alternation ensured consistent smoke conditions, allowing fair comparison between the two methods.

For the 2D XZ-plane projection experiment (Fig. S37a), the proposed method reconstructs beam contours clearly, preserving the front-to-back twisting features and axial separation of the double twisted curve structure. Background noise is low, and the overall structure exhibits high visibility and contrast. In contrast, Fig. S37b shows that RV-CGH produces more blurred contours. The front-end features are partially recognizable but with reduced brightness and contrast, and structural information in the rear region is largely lost due to strong background interference, making axial positions difficult to distinguish. These results demonstrate a clear difference between the two methods in 2D smoke visualization.

To further evaluate volumetric reconstruction and 3D visibility, we compared a 3D hollow curved surface under smoke conditions. This target consists of a ring-like structure that gradually contracts and expands along the propagation direction, forming a continuously varying 3D surface. As shown in Fig. S38a, the proposed method successfully reconstructs a 3D structure with clearly defined contours and features. Viewed from above the light field, the structure exhibits smooth size variation along the propagation direction, with narrower features in the middle region and overall high contrast and structural integrity. In contrast, Fig. S38b shows that RV-CGH exhibits significant degradation. Overall contrast is reduced, and fine structures in the middle region are completely unresolved, replaced by a larger, blurred light field.

The observed differences primarily arise from unwanted outgoing light generated by multiple projection planes in RV-CGH, which creates additional background interference and structural crosstalk, reducing contrast. In smoke, these defocused background components are further diffused and amplified, leading to substantial loss of structural detail and reduced visibility. By contrast, the proposed cascaded propagation framework mitigates these sources of crosstalk and energy loss, resulting in a more complete 3D light field under the same smoke conditions.

These comparative results demonstrate that, under scattering conditions, the proposed method preserves more complete 3D structures and higher visibility compared to RV-CGH, further highlighting its effectiveness and advantages.

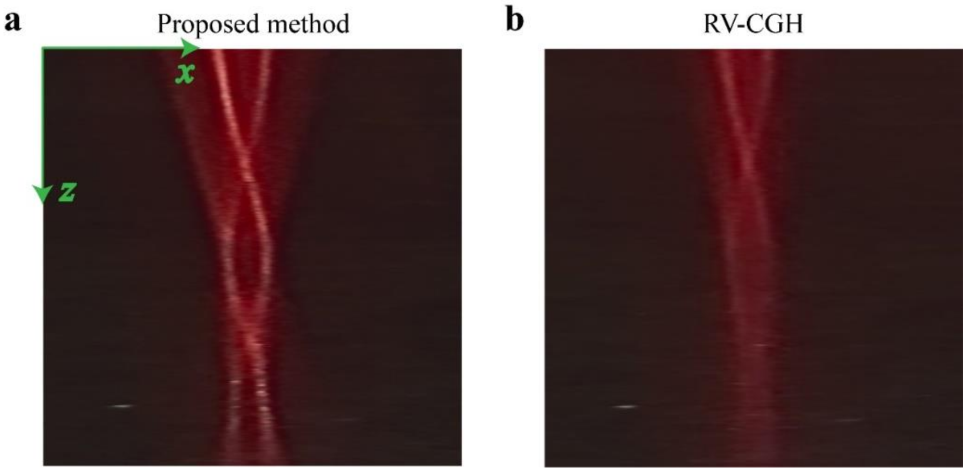

Fig. S37 | Comparison of XZ-plane reconstructed light fields between the proposed method and RV-CGH under smoke environment. a, Reconstruction result of the proposed method. **b,** Reconstruction result of RV-CGH. (XZ aspect ratio: 8:1). Evaluated under identical incident power, camera focal plane, and smoke density, proposed method exhibits clear axial (z-direction) structural preservation, where the double twisted curves remain distinctly separable and recognizable along the propagation direction. In contrast, the RV-CGH reconstruction shows reduced axial clarity and increased depth-direction blurring, leading to degraded volumetric structural visibility.

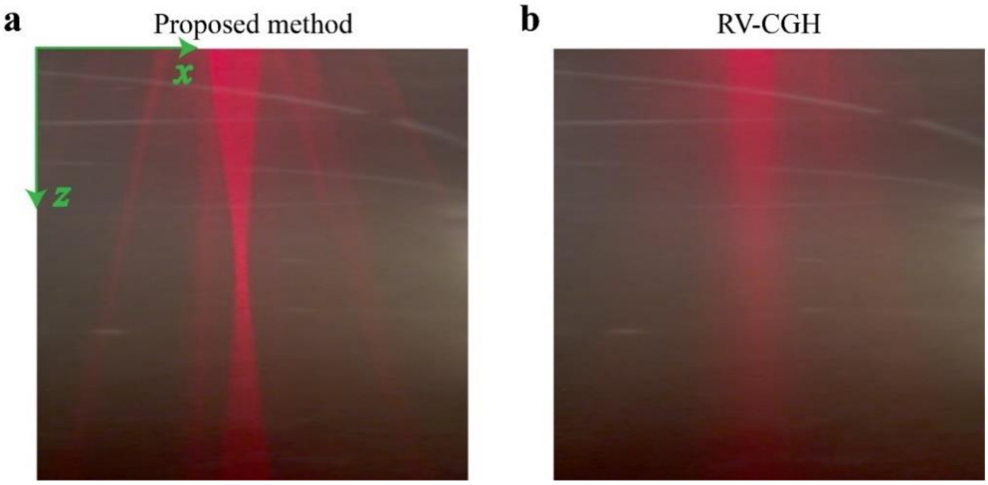

Fig. S38 | Comparison of reconstruction quality for a 3D hollow curved-surface light field under smoke environment. a, Reconstruction result of the proposed method. **b,** Reconstruction result of RV-CGH. (XZ aspect ratio: 4:1). Evaluated under identical incident power, camera focal plane, and smoke density, proposed method reconstructs a continuous hollow 3D curved surface with clearly defined geometric boundaries and well-preserved volumetric continuity. In contrast, the RV-CGH reconstruction exhibits pronounced axial blurring and severe structural degradation.

Supplementary Note 18: System Performance and Heterogeneous Scene Evaluation

3D resolution characterization

To further characterize the practical operating boundaries of the proposed method, we quantitatively evaluated the reliable resolvable feature spacing in both the lateral and axial directions under the current experimental configuration using a USAF resolution target and a dual-point axial spacing measurement approach.

For the lateral resolution assessment, a digital USAF-1951 resolution target was displayed via SLM. As shown in Fig. S39a, the test pattern was successfully reconstructed and captured by the camera. The full-field view (left) shows a clear and detailed projection, while the zoomed-in region (right) highlights the finest resolvable features within the area of interest. To quantitatively evaluate the lateral resolution, a horizontal intensity line profile was extracted from the region marked by the green box in Fig. S39a, as shown in Fig. S39b. The normalized intensity distribution exhibits three distinct peaks separated by well-defined minima, confirming their resolvability. The measured center-to-center spacing between adjacent peaks is approximately 24.15μm, which represents the reliably resolvable lateral feature spacing under the current system. The line-profile analysis provides quantitative evidence beyond visual inspection, with sufficient peak-to-valley contrast indicating that the three-line structure is clearly distinguishable.

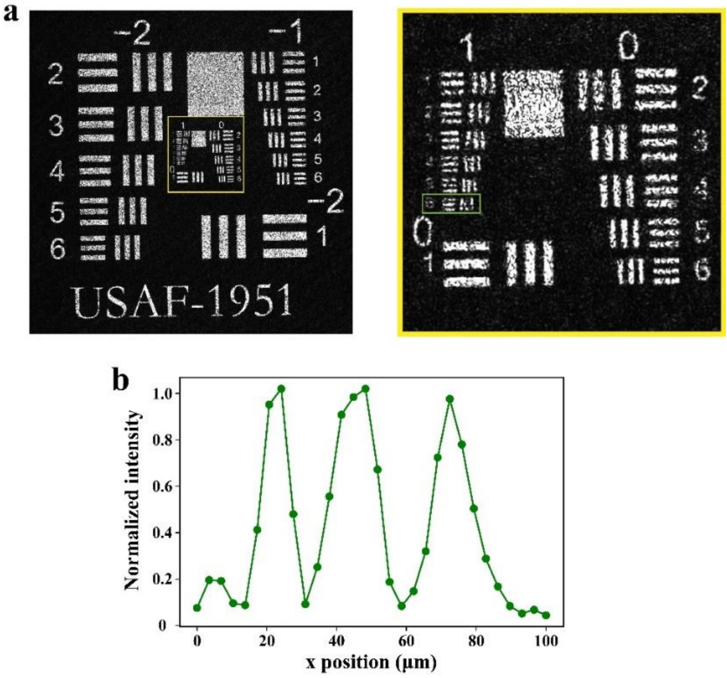

Fig. S39 | Lateral resolution evaluation of the holographic projection system using a projected USAF-1951 resolution target. a, Reconstructed USAF-1951 target captured by the camera (left, full field of view, 6900×6900μm). For clarity, the right panel shows a magnified view of the region highlighted by the yellow box in the left image (1656×1656μm). The resolvable features correspond to the three-line pattern in Group 1, Element 6 (green box). **b,** Normalized intensity profile extracted along the vertical direction across the three-line structure indicated by the green box. The peak-to-peak spacing between adjacent intensity maxima is used to quantify the resolvable feature size and evaluate the lateral resolution of the reconstructed projection. The measured center-to-center spacing between adjacent peaks is approximately 24.15μm.

To characterize the axial resolution, we further performed a dual-point axial spacing experiment. Two tightly focused points located at different depths were reconstructed, with progressively decreasing axial separation. The axial intensity distributions were extracted and analyzed to evaluate their resolvability. As shown in Fig. S40, the two focal spots exhibit a transition from clearly separable to partially overlapping peak structures as the axial spacing decreases. When the separation is 800μm, two distinct intensity maxima are clearly observed (Fig. S40b). When the spacing is reduced to 700μm, the valley between the peaks rises, indicating stronger coupling and increased crosstalk (Fig. S40d). By systematically varying the axial separation and analyzing the resolvability of the intensity profiles, we determine that the reliably resolvable axial spacing under the current imaging conditions is approximately 700μm.

These experimental results quantitatively characterize the 3D resolution capability of the proposed method and help define its practical operating boundaries.

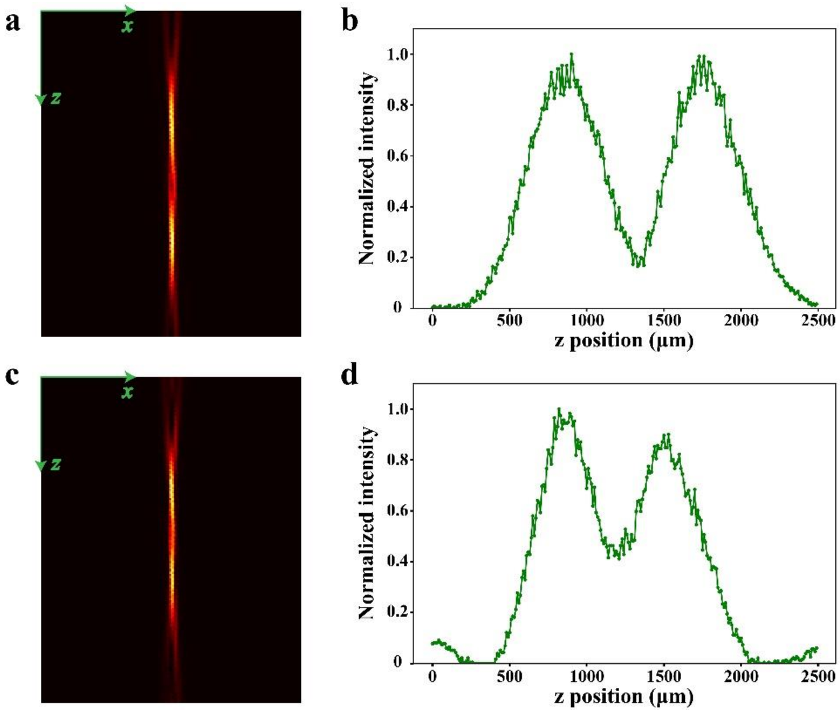

Fig. S40 | Axial resolution evaluation based on dual-focus axial separation measurements. a, c, Reconstructed intensity distributions of two focal spots located at different axial positions with small separation distances. The center-to-center axial separations are 800μm and 700μm, respectively. The xz-plane views illustrate the evolution of the axial intensity distribution. **b, d,** Corresponding normalized intensity profiles extracted along the z-direction, respectively. The peak-valley structures reflect the axial resolvability between adjacent focal features. A pronounced dip is observed between the two peaks, with the valley intensity below half of the peak value, indicating partial resolvability.

Heterogeneous scenes validation

To further evaluate the performance of the proposed framework under heterogeneous target conditions, we conducted additional stress tests involving the reconstruction of structurally diverse heterogeneous patterns at different spatial regions and depths.

First, heterogeneous targets were projected at different regions within the same image plane. As shown in Fig. S41, a USAF resolution target was placed in the left region, while a Siemens star pattern was placed in the right region. The USAF target contains dense high-frequency features, whereas the Siemens star exhibits radially varying spatial frequency components, jointly validating the ability of the system to preserve and control high-frequency details across different orientations and frequency distributions. The results demonstrate faithful reconstruction of both targets, with clear spatial selectivity and negligible inter-region interference.

In addition, heterogeneous targets at different axial depths were also experimentally validated. Specifically, a semantically structured "Nature Photonics" text pattern (instead of single letter) was reconstructed on one depth plane, while a USAF-1951 resolution target with dense high-spatial-frequency line features was reconstructed on another plane. These two targets were reconstructed at distinct axial positions and occupied separate spatial regions within the holographic volume. As shown in Fig. S42, despite significant structural differences between the heterogeneous targets, both reconstructions remain clearly distinguishable, with fine details well preserved.

Across these heterogeneous stress tests involving both spatially separated and axially separated targets, the proposed method consistently maintains clear feature separation and low crosstalk. These heterogeneous configurations introduce substantially increased spatial-frequency diversity, providing a more stringent validation and demonstrating the effectiveness and applicability of the proposed method for reconstructing heterogeneous scenes.

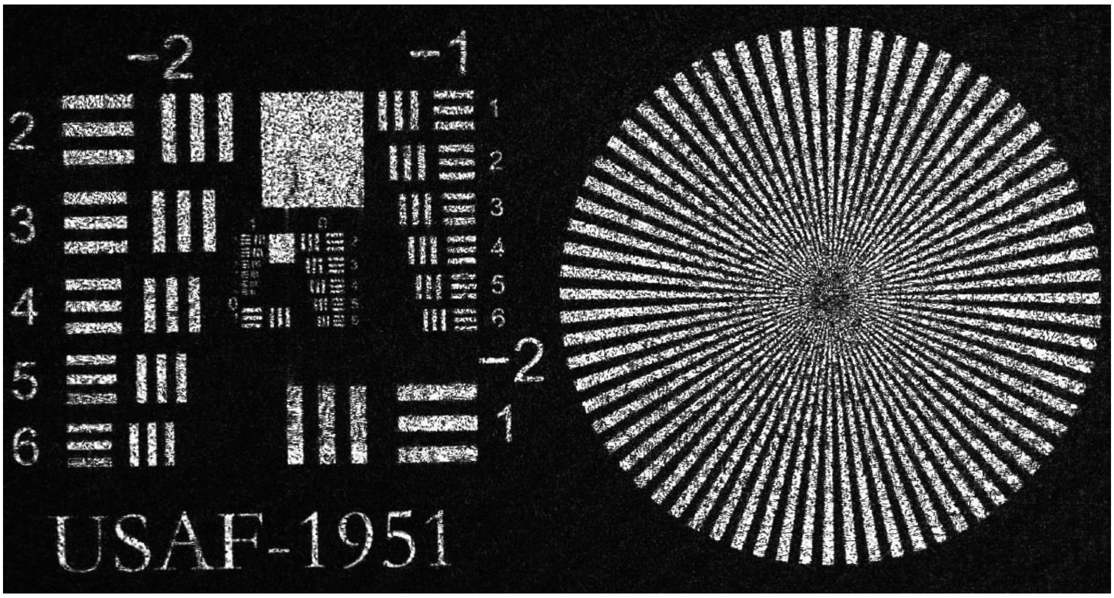

Fig. S41 | Experimental validation of heterogeneous targets placed in different holographic regions. The left region shows the USAF resolution target, corresponding to high-frequency features, while the right region shows a Siemens star with 64 spokes, exhibiting radially varying spatial frequencies. The overall physical size of the field is 7065×4492μm. Both targets are simultaneously reconstructed in spatially separated regions on the same projection plane, demonstrating good spatial selectivity and negligible inter-region interference.

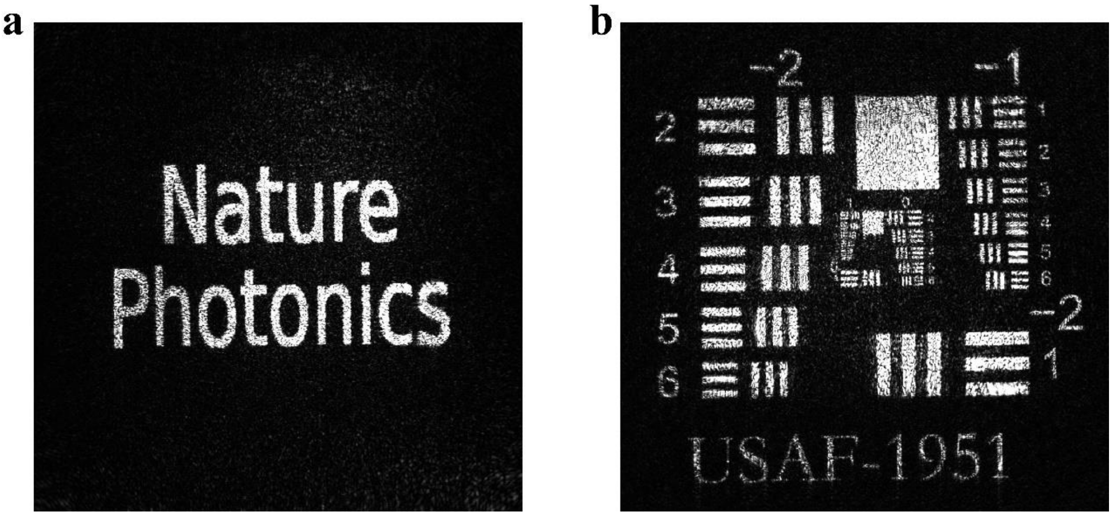

Fig. S42 | Experimental validation of heterogeneous targets placed at different holographic depths. **a,** A semantically structured target (“Nature Photonics” text) projected at one depth plane, rather than a single simple letter. **b,** A USAF-1951 resolution target, containing dense high-spatial-frequency features and fine line structures, projected at another depth plane. Both targets have a dimension of 5175×5175μm, with a 3cm axial separation. These heterogeneous targets are reconstructed at different depths, occupying distinct spatial regions within the holographic volume, demonstrating robust feature preservation and effective heterogeneous target projection.

Supplementary References

1. Makey, G. *et al.* Breaking crosstalk limits to dynamic holography using orthogonality of high-dimensional random vectors. *Nature photonics* **13**, 251-256 (2019).
2. Yu, P. *et al.* Ultrahigh-density 3D holographic projection by scattering-assisted dynamic holography. *Optica* **10**, doi:10.1364/optica.483057 (2023).
3. Dorrah, A. H. *et al.* Light sheets for continuous-depth holography and three-dimensional volumetric displays. *Nature Photonics* **17**, 427-434, doi:10.1038/s41566-023-01188-y (2023).
4. Genevet, P. & Capasso, F. Holographic optical metasurfaces: a review of current progress. *Reports on Progress in Physics* **78**, 024401 (2015).
5. Huang, L., Zhang, S. & Zentgraf, T. Metasurface holography: from fundamentals to applications. *Nanophotonics* **7**, 1169-1190 (2018).
6. Melde, K., Mark, A. G., Qiu, T. & Fischer, P. Holograms for acoustics. *Nature* **537**, 518-522 (2016).