

ARTICLE

Open Access

Spatial incoherence-driven optical reconstruction of holograms with observer shift-invariance

Yeo Ju Sohn¹ and Daeho Yang²✉

Abstract

Coherence preserves phase consistency between wavefields, enabling accurate recording and reconstruction in holography. Although recent advances in computational optics have realized holographic data acquisition using incoherent light by computationally retrieving information, optical reconstruction still requires partially coherent light sources. We demonstrate a hologram that reconstructs 3-dimensional distribution utilizing incoherence. By decomposing incoherent light into infinitesimal coherent lights and calculating their propagations, the incoherent sum is optimized to resemble the desired 3-dimensional scene, whereas individual coherent lights reconstruct completely different intensities. Incoherence provides high image quality and a wide eyebox, with the reconstructed intensity remaining shift-invariant under pupil displacement, allowing a 1000-fold expansion of the eyebox. We confirm the shift-invariance through a proof-of-concept experiment and demonstrate real-time synthesis of incoherent holograms using a neural network, significantly reducing computational costs. Our method could inspire new approaches in photonics using incoherent light and be practically adopted in holographic displays.

Introduction

Holography enables recording and reconstructing the wavefield of an object through coherent wavefield interference¹. The coherent interference of holography allows the recovery of phase information, making it irreplaceable for applications in non-destructive testing², data storage³, and biomedicine⁴. Moreover, 3-dimensional (3D) images can be reconstructed through phase information⁵, making holography valuable in microscopy⁶, biological 3D imaging⁷, and 3D object recognition⁸. Since holography relies on interference, coherent light sources are traditionally considered essential for obtaining accurate reconstructions.

However, advanced computational methods have enabled the acquisition of holographic data using incoherent light^{9–13}. Unlike traditional coherent holography, which records the wavefront of light, incoherent holography computationally reconstructs 3D location information, reducing speckle noise and minimizing sample damage. For instance, Fresnel incoherent correlation

holography (FINCH) adopts self-interference between wavefields with a common optical path to capture interference patterns¹¹, while optical scanning holography (OSH) uses active optical heterodyne scanning to record holograms¹². Furthermore, the complex-valued hologram is extracted without interference by modulating and recording multiple intensities in interferenceless coded aperture correlation holography (I-COACH)¹³.

Incoherent light sources have been extensively explored not only in holographic data acquisition but also in optical 3D reconstruction to address issues associated with coherent light sources^{14–16}. A superluminescent LED is utilized as a partially coherent light source to mitigate unwanted interference and noise¹⁴. Coherence of the light source is optimized¹⁵, or multiple sources are employed¹⁶, to reduce speckle noise and enhance resolution. However, optical reconstruction is limited to partially coherent regimes, as increased incoherence hinders 3D reconstruction by stochastically modulating wavefield¹⁷ (Fig. 1a).

Here, we demonstrate the optical reconstruction of a hologram based on incoherent light, requiring incoherence instead of coherence. Our method, Coherently-Reconstructed Incoherent Sum (CRIS), optically reconstructs the hologram via the incoherent sum of coherent lights. More specifically,

Correspondence: Daeho Yang (sheepys5@gachon.ac.kr)

¹Department of Family Medicine, Ewha Womans University College of Medicine, Seoul 07804, South Korea

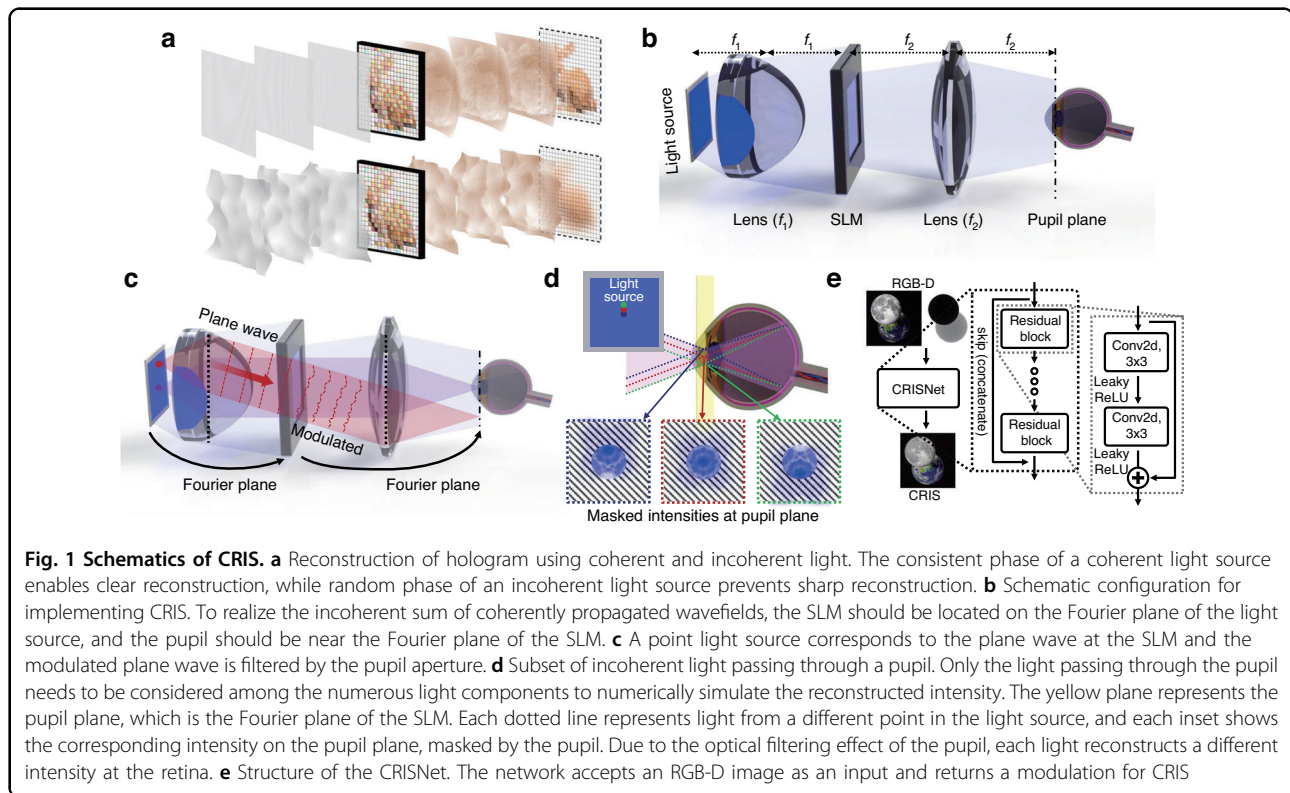
²Department of Physics, Gachon University, Seongnam, Gyeonggi-do 13120, South Korea

© The Author(s) 2025



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.





the intensity at the retina is calculated by propagating and accumulating the subset of incoherent light passing through the pupil, while the simulated intensity is optimized to resemble the target 3D scene. Spatial incoherence allows the scene to be viewed from various angles, achieving 3D reconstruction with wide eyebox—characteristics previously thought to be mutually exclusive in holographic and conventional displays^{18,19}. We theoretically demonstrate that CRIS remains shift-invariant under pupil displacement and experimentally confirm the shift-invariance. Theoretically, the shift-invariance allows for a 1000-fold expansion of the eyebox, though our proof-of-concept experiment achieved a 32-fold expansion. At the same time, we trained a neural network for real-time synthesis because simulating and optimizing an incoherent sum requires millions of times more computational resources than conventional holograms, making its use in holographic displays impractical. We anticipate our method will enable the widespread adoption of holographic displays by addressing key challenges—high image quality, real-time synthesis, and wide eyebox—without additional equipment, while also inspiring new approaches with incoherent light.

Results

For coherent light incident on a spatial light modulator (SLM), the scattering angle is determined by its pixel pitch. Consequently, the light is visible only within a small area, known as the eyebox, and even minor eye

movements can cause the reconstructed image to disappear. On the other hand, illuminating the SLM with incoherent light, such as that from a large-area LED, introduces a broad range of incident angles. Due to the large incident angle exceeding the diffraction angle of the SLM, the modulated light spread over various output directions, enabling a wide viewing angle regardless of the diffraction angle of the SLM, as seen in conventional displays. However, such displays are not capable of optically reconstructing arbitrary 3D scenes due to the incoherence of the light source (Fig. 1a).

To address the limitation and utilize coherence among incoherent light, we used spatial coherence associated with an individual point light source. A single point light source exhibits spatial coherence, while a spatially incoherent light source can be considered as an ensemble of multiple uncorrelated point sources. By placing the SLM in the Fourier plane of the light source, light from a point forms a plane wave with a specific incident angle at the SLM plane (Fig. 1b). When considering only the light with the specific incident angle at the SLM plane, spatial coherence exists, and the light modulated by the SLM can reconstruct an arbitrary 3D wavefield.

After each modulated plane wave passes through another lens, the modulated light is Fourier transformed at the pupil plane (Fig. 1c). Although each plane wave is modulated by the same SLM pattern, the modulated light passing through the pupil differs due to the filtering effect



of the pupil. Since the pupil plane corresponds to the Fourier plane of the SLM, each modulated wave is represented as a Fourier-domain wave on the pupil plane. Each Fourier-domain wave, shifted proportionally to its plane wave angle, is masked by a pupil-defined Fourier filter (Fig. 1d). As a result, each modulated plane wave has different intensities at the retina, and the total intensity can be calculated as the incoherent sum of the individual intensities of the filtered waves. By integrating a subset of incident light imaged by the pupil aperture, the total intensity of CRIS can be expressed as,

$$I_{\text{CRIS}}(\vec{r}, z) = \frac{1}{\mathcal{N}} \sum_{\vec{k}_g} \left| \int_{\mathcal{A}} \overline{U}(\vec{k} + \vec{k}_g) e^{iz\sqrt{(2\pi/\lambda)^2 - |\vec{k}|^2}} e^{i\vec{k} \cdot \vec{r}} d\vec{k} \right|^2 \quad (1)$$

where, $I_{\text{CRIS}}(\vec{r}, z)$ is the reconstructed intensity, z is the distance along the longitudinal direction, $\mathcal{N}$ is a normalization constant, $\vec{k}_g = (k_{gx}, k_{gy})$ is the radial wavevector of the collimated incident light, $\overline{U}(\vec{k})$ is the Fourier transformed function of the SLM modulation $U(\vec{r})$, $\vec{k} = (k_x, k_y)$ is the radial wavevector of the modulated field, λ is the wavelength of the light, and $\mathcal{A}$ is the area of the pupil in the Fourier plane. Since the pupil blocks different spatial frequencies of the modulation based on $\vec{k}_g$, each filtered wavefield represents only a partial frequency region of the modulation. (Fig. 1d). In summary, the reconstructed total intensity for a specific depth can be calculated by accurately accounting for all coherent propagation of these partial modulations.

By introducing the pupil location vector from the center of the eyebox, (p_x, p_y, p_z) , it is possible to derive the shift-invariant formula (see “Methods”),

$$I_{\text{CRIS}}(\vec{r}, z, p_x, p_y, p_z) = I_{\text{CRIS}}(\vec{r}, z, 0, 0, 0) \quad (2)$$

Equation (2) is valid under the Fresnel approximation, and thus the reconstructed intensity is invariant even if the pupil shifts along all axes within the region where the Fresnel approximation holds. Therefore, if we find a proper Fourier transformed modulation $\overline{U}(\vec{k})$ satisfying $I_{\text{CRIS}}(\vec{r}, z) = I_{\text{target}}$, then we can reconstruct a 3D scene which can be seen everywhere by displaying the modulation $U(\vec{r})$. The shift-invariance can be understood as the cancellation between the grating phase of the incident angle and the shift in the Fourier domain.

When the pupil shifts in the Fourier plane, we can find the specific incident angle at which the light passes through the center of the pupil (Fig. 1c). Under the Fresnel approximation, the corresponding light has the same wavefield as the light with a zero incident angle, except for the oscillating phase induced by the incident angle. The same rationale can be applied to marginal

light; therefore, the incoherent sum would be the same as the case when the pupil is at the center. A more rigorous derivation of the shift-invariance of the CRIS, including the shift along the z -axis, can be found in the Methods.

To find the modulation satisfying Eq. (1), gradient descent methods can be used. After calculating the right-hand side of Eq. (1) for an arbitrary initial modulation, the difference between the incoherent sum and the target intensity is calculated. The modulation is updated to minimize the difference, and the process is repeated until the predefined number of iterations is reached. To numerically simulate incoherent light propagation, coherent propagations are calculated for hundreds of different incident angles, and all the coherent propagation results are incoherently added (see Supplementary Material). As a result, millions of coherent propagations are required to synthesize a single hologram, making the algorithm impractical for real-time applications.

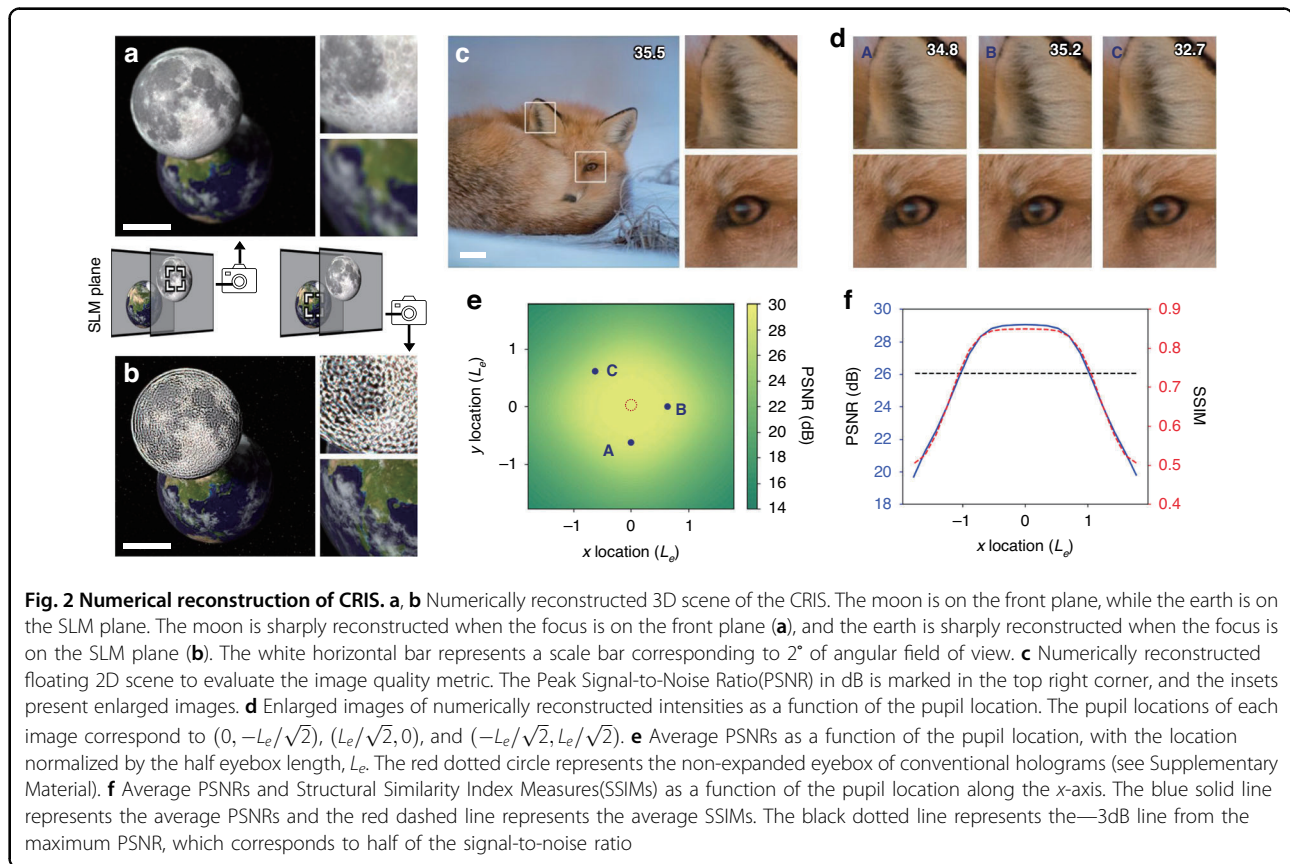
To overcome the limitation, we trained a neural network (CRISNet) where the model is composed of residual blocks with 3×3 convolutional layers (Fig. 1e). To train multi-depth holograms, the incoherent sum is numerically calculated for each depth and compared with the depth-dependent target image. The depth-dependent target image is synthesized by multiplying an all-in-focus image with a mask in which only the pixels at each corresponding depth are nonzero. As a result, the total cost function is given as

$$\mathcal{L} = \sum_n M(\vec{r}, d_n) [I_{\text{CRIS}}(\vec{r}, d_n) - I_{\text{target}}(\vec{r})] \quad (3)$$

where, d_n is distance between the n th layer and the SLM, $M(\vec{r}, d_n)$ is the mask for pixels in the n th layer, and $I_{\text{target}}(\vec{r})$ is the target intensity.

Figure 2 presents the numerical reconstruction results of the CRIS. The CRIS is synthesized to reconstruct the scene, with the moon located at the floating plane in front of the SLM and the earth located at the SLM plane. To numerically simulate CRIS, the virtual image of the Earth is assumed to be located at infinity, corresponding to 0 diopters, while the virtual image of the Moon is assumed to be located 67 cm from the pupil, corresponding to 1.5 diopters. The numerical reconstruction presents that the earth appears blurred while the moon remains sharp when the focus is on the front plane, implying 3D reconstruction (Fig. 2a). In contrast, the earth becomes sharp when the focus is on the SLM plane, although the defocus blur of the moon is not natural (Fig. 2b). The defocus blur can be modified by adopting additional loss functions in sake of reducing image quality^{20,21} (See Supplementary Material).





Given that most image quality metrics are designed for 2D images, we utilized a 2D image dataset, the DIV2K validation dataset²², to quantitatively analyze the image quality of the CRIS. We assumed that the virtual images were positioned 1.5 diopters from the SLM plane to simulate floating objects, where 1.5 diopters correspond to a typical interaction length of 67 cm, approximately the length of a human arm. One of the numerically reconstructed floating images, for the case where the pupil is at the center of the eyebox, is presented in Fig. 2c. Additional reconstruction intensities for different target images are shown in Fig. 7.

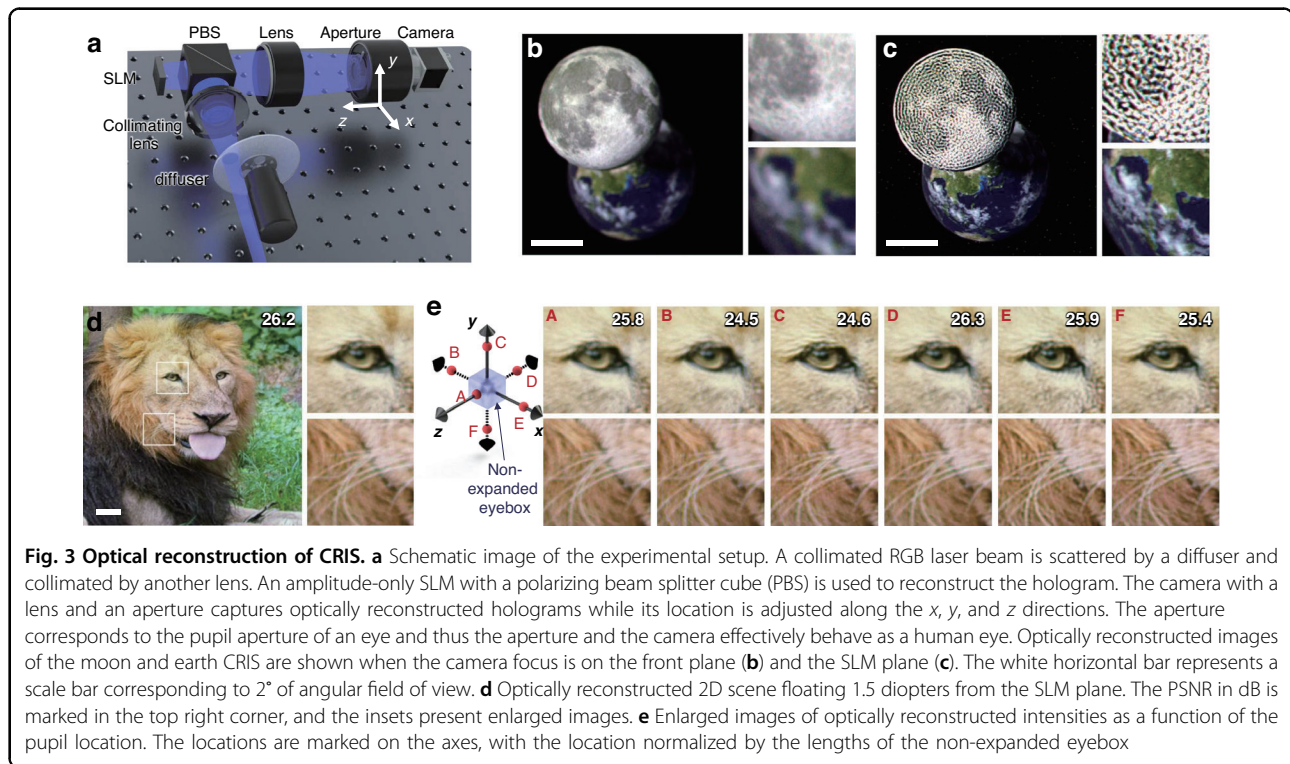
To demonstrate the shift-invariance of the CRIS, the same image is numerically reconstructed at different locations of the eyebox (Fig. 2d). According to the theoretically derived formula (see “Methods”), the half length of the eyebox along the x - and y -axes is given as $L_e = (8\lambda f^2/D)^{1/4}$, where D is the diopter. Each point in Fig. 2d corresponds to the points $(0, -L_e/\sqrt{2})$, $(L_e/\sqrt{2}, 0)$, and $(-L_e/\sqrt{2}, L_e/\sqrt{2})$. To visualize the distribution of image quality, the CRIS is synthesized for all images in the dataset and the PSNR values are averaged for each pupil location through the numerical reconstruction. Figure 2e presents the average PSNR as a function of the x and y location in a unit of L_e . As the shift-invariance does not

hold beyond the valid range of the Fresnel approximation, the average PSNR decreases near the distance L_e . A cross-sectional graph confirms that the full width at half maximum (FWHM) of the PSNR is $2L_e$, indicating agreement between the numerical simulation results and the theoretically derived formula within 1.6% (Fig. 2f). The average PSNR for the pupil at the center of the eyebox is 29 dB and that for the pupil at the distance L_e is 26.3 dB.

One of the notable points of the eyebox length is that the eyebox length does not depend on the pixel pitch of an SLM. Therefore, by increasing the focal length of the lens, it is possible to increase the eyebox while sustaining the field of view by using an SLM with a larger physical size. Another important point is that the eyebox length decreases as the diopter of the hologram increases. However, increasing the diopter by a factor of two only results in a 15% decrease in the eyebox length, as it is inversely related to the $1/4$ power of the diopter. (See “Methods”)

Addressing the etendue or eyebox limitations of holographic displays has long been studied. For instance, to increase the etendue, a scattering mask is placed while the scattering noise is reduced by optimizing the wavefield²³ and adopting a neural etendue expander²⁴. Moreover, a photon sieve is utilized to enhance the viewing angle²⁵.





However, supporting high image quality for every pupil location remains challenging, whereas CRIS can achieve the wide eyebox and high image quality with a single hologram.

To confirm the shift-invariance of the CRIS, optical reconstruction (i.e., experimental reconstruction) is performed (Fig. 3). Although spatial coherence should be suppressed in realizing the CRIS, sustaining temporal coherence is ideal for providing high image quality. To create the light source assumed in the theoretical calculation, red, blue, and green lasers are coupled to a single-mode fiber, and a rotating diffuser is employed to suppress spatial coherence. Even though temporally coherent light was used in the experiment, image quality is anticipated to be sustained without temporal coherence (see Supplementary Material). An amplitude-only SLM is used to optically reconstruct the CRIS whereas the CRIS is synthesized to be represented solely by amplitude. Contrary to conventional holograms, the Fourier plane is filled with light from various incident angles, making the use of a Fourier filter impossible. Although not explicitly stated in the numerical reconstruction, all the reconstructed images and benchmarks are based on the amplitude-only representation.

With the same 3D scene presented in the numerical reconstruction, the optical reconstruction shows the sharp moon and the blurred earth when the camera focus is on the front plane, indicating successful reconstruction

of floating objects (Fig. 3b). When the camera focus is on the SLM plane, the earth becomes sharp (Fig. 3c). The suppressed spatial coherence prevents speckle noise induced by laser coherence and dust, which are the major sources of noise in conventional holograms, resulting in high image quality. Additional 3D scenes for different pupil locations are shown in Fig. 8.

To assess image quality metrics based on the aperture location, we optically reconstruct a 2D floating image and evaluate the PSNR as the aperture location varies. Figure 4d shows the optically reconstructed image with enlarged insets when the aperture is at the center of the eyebox. Even when the aperture is moved beyond the boundaries of the non-expanded eyebox, the image quality remains similar (Fig. 4e). Furthermore, shifting the aperture along the z -axis does not affect image quality or field of view, contrary to conventional holograms. In other words, we experimentally confirmed the shift-invariance, which theoretically allows a 1000-fold expansion of the eyebox, where the theoretical eyebox expansion ratio is calculated by comparing $2L_e$ and the conventional eyebox (see “Methods”). Since the restricted numerical aperture of our setup limits the incident angle of light, the measured eyebox volume is $4.5 \times 4.5 \times 12 \text{ mm}^3$, representing a 32-fold expansion. Nonetheless, the experimental demonstration of the shift-invariance suggests that the eyebox can be further increased by adopting optics with a larger numerical aperture (See “Methods”).

Although the CRIS can resolve the eyebox problem, there still remain limitations related to depth representation, including the absence of motion parallax and unnatural defocus blur. Since the CRIS provides the same 3D scene regardless of pupil locations, motion parallax does not exist in the CRIS. The absence of motion parallax could be an issue in flat-panel holographic displays; however, it is not problematic for near-eye displays with head motion-induced parallax. Considering that holographic displays for near-eye displays have been extensively studied recently^{24,26,27} due to the augmented, virtual, and mixed reality applications, CRIS offers a promising solution for these contexts, particularly for supporting accommodation effect.

Discussion

Holographic displays have long been considered to be the future of displays due to the advantage of providing photorealistic 3D scenes without vergence-accommodation conflicts²⁸. However, to deliver immersive experiences, holographic displays should match conventional 2D displays in every aspect while also reconstructing 3D scenes. Although recent advancements have addressed issues related to high computational costs and low image quality, a small eyebox problem still remains. Despite various proposals for eyebox expansion, every method has its own challenges, including low image quality, double vision, or low responsiveness, leading to less immersive experiences compared to conventional displays^{29,30}. Our method addresses the limited eyebox problem while providing coherent-artifact-free reconstruction like conventional displays, offering immersive experiences.

Recently, ordinary displays have been utilized to reconstruct holograms without relying on coherent light sources³¹. While both the ordinary display approach and our approach utilize incoherent light sources, the two methods have distinct goals. The former approach focuses on utilizing ordinary displays under general conditions without constraints; however, a wide eyebox is not expected because the optimized intensity is limited to a fixed camera. In contrast, CRIS prioritizes achieving eyebox expansion by applying specific constraints to the positions of the light source and SLM, while also utilizing the filtering effect of the pupil aperture.

Unnatural defocus blur is a common issue in holographic displays, manifesting as both weak defocus blur and defocus blur with speckles. In holograms with smooth phase, suppressed light scattering provides high image quality but leads to weak defocus blur³². Conversely, in holograms with non-smooth phase, increased light scattering induces high-contrast speckles in out-of-focus areas, disrupting the correct perception of depth cues³³. Although the CRIS provides natural defocus blur for objects on the SLM plane, the

defocus blur for objects at other planes remains unnatural. Various methods can address the unnatural defocus blur^{20,34}, and employing these methods can relieve the issue (See Supplementary Material).

In summary, we proposed a new hologram, CRIS, and theoretically derived the shift-invariance property. The proof-of-concept experiment demonstrates that the shift-invariance holds for all axes and reconstructing a 3D scene is possible. The CRIS not only inspires new approaches using incoherent light but also resolves key challenges of holographic displays such as eyebox limitations, image quality, real-time synthesis, and responsiveness. We expect the CRIS to promote the widespread use of holographic displays by providing 3D reconstruction with the enhanced properties comparable to those of 2D displays. Moreover, the incoherent optical reconstruction will enhance the understanding of incoherent light in holography and enable various applications based on its accurate consideration.

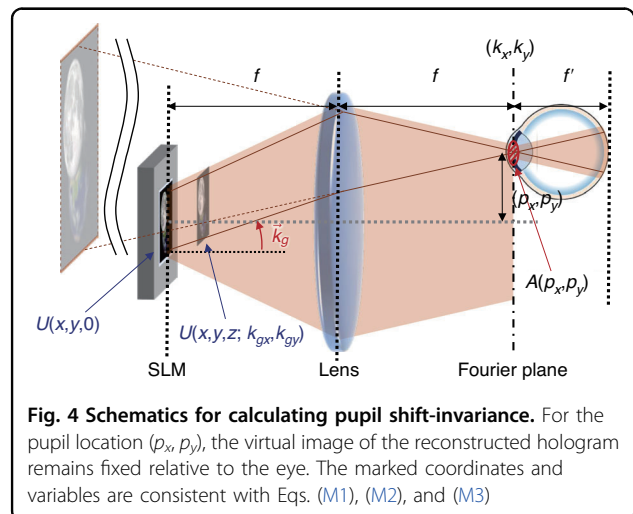
Materials and methods

Shift-invariance property under Fresnel approximation

According to the Angular Spectrum Method (ASM), the Fourier-transformed wavefield $\bar{U}$ can be expressed as $\bar{U}(k_x, k_y, z) = \bar{U}(k_x, k_y, 0)e^{iz\sqrt{k^2 - k_x^2 - k_y^2}} \approx \bar{U}(k_x, k_y, 0)e^{ikz\left(1 - \frac{k_x^2 + k_y^2}{2k^2}\right)}$, where we used the Fresnel approximation. The propagated wavefield $U(x, y, z)$ can be expressed as,

$$U(x, y, z) = \iint \bar{U}(k_x, k_y, 0) \times e^{ikz\left(1 - \frac{k_x^2 + k_y^2}{2k^2}\right)} e^{ik_x x + ik_y y} dk_x dk_y \quad (\text{M1})$$

By considering spatially incoherent light as incoherent sum of numerous coherent light components, Eq. (M1)



can be used to express a hologram reconstructed by coherent light. To consider coherent light with different incident angles, the grating phase, $e^{-ik_{gx}x - ik_{gy}y}$, should be applied to the wavefield before propagation, $U(x, y, 0)$. As a result, Eq. (M1) can be modified to consider the propagated wavefield with different propagation angle, (Fig. 4)

$$U(x, y, z; k_{gx}, k_{gy}) = \int \int \overline{U}(k_x + k_{gx}, k_y + k_{gy}, 0) \times e^{ikz \left(1 - \frac{k_x^2 + k_y^2}{2k^2}\right)} e^{ik_x x + ik_y y} dk_x dk_y \quad (\text{M2})$$

where the grating phase, $e^{-ik_{gx}x - ik_{gy}y}$, is removed by shifting the Fourier-transformed wavefield to $\overline{U}(k_x + k_{gx}, k_y + k_{gy}, 0)$.

A pupil of an eye behaves as an aperture, so only a limited area, $\mathcal{A}(p_x, p_y)$, should be integrated in Eq. (M2) to numerically reconstruct a visible intensity. By integrating the holograms reconstructed by coherent light components with different incident angles, the incoherent sum can be expressed as

$$I(x, y, z, p_x, p_y) = \frac{1}{\mathcal{N}} \sum_{k_{gx}, k_{gy}} \left| \int \int_{\mathcal{A}(p_x, p_y)} \overline{U}(k_x + k_{gx}, k_y + k_{gy}, 0) \times e^{ikz \left(1 - \frac{k_x^2 + k_y^2}{2k^2}\right)} e^{ik_x x + ik_y y} dk_x dk_y \right|^2 \quad (\text{M3})$$

where $\mathcal{A}(p_x, p_y)$ denotes the area depending on the position of the pupil. Here we assumed the pupil as a circle with the center (p_x, p_y) and radius p_r .

To explicitly represent the dependency of integration area on the pupil location, we rewrite Eq. (M3) by shifting k_x and k_y . Considering f is the focal length of the lens used to make the spatial Fourier transform, we can define the shifted wavevectors as, $k_x \equiv k'_x + k_{px}$, $k_y \equiv k'_y + k_{py}$, $k_{px} \equiv kp_x/f$, $k_{py} \equiv kp_y/f$. The re-written equation is given as

$$I(x, y, z, p_x, p_y) = \frac{1}{\mathcal{N}} \sum_{k_{gx}, k_{gy}} \left| \int \int_{\mathcal{A}(0,0)} \overline{U}(k'_x + k_{px} + k_{gx}, k'_y + k_{py} + k_{gy}, 0) \times e^{ikz \left(1 - \frac{(k'_x + k_{px})^2 + (k'_y + k_{py})^2}{2k^2}\right)} e^{i(k'_x + k_{px})x + i(k'_y + k_{py})y} dk'_x dk'_y \right|^2 \quad (\text{M4})$$

Since the summation range of $k_{gx}(k_{gy})$ is from negative infinity to positive infinity, we can arbitrarily shift $k_{gx}(k_{gy})$, and thus $k_{gx}(k_{gy})$ can be replaced with $k_{gx} - k_{px}(k_{gy} - k_{py})$. After exchanging $k_{gx}(k_{gy})$ and performing simple

calculation yields the following equation,

$$I(x, y, z, p_x, p_y) = \frac{1}{\mathcal{N}} \sum_{k_{gx}, k_{gy}} \left| \int \int_{\mathcal{A}(0,0)} \overline{U}(k'_x + k_{gx}, k'_y + k_{gy}, 0) \times e^{ikz \left(1 - \frac{k_x'^2 + k_y'^2}{2k^2}\right)} e^{ik'_x \left(x - \frac{zk_{px}}{k}\right) + ik'_y \left(y - \frac{zk_{py}}{k}\right)} dk'_x dk'_y \right|^2 \quad (\text{M5})$$

During the calculation, the term $e^{-ikz \left(\frac{k_{px}^2}{2k^2} + \frac{k_{py}^2}{2k^2}\right)} e^{ik_{px}x + ik_{py}y}$ is removed because it does not include k'_x or k'_y and the pure phase term can be removed in the calculation of the absolute square.

To make the result clearer, $x(y)$ is exchanged to $x + \frac{zk_{px}}{k}$ ($y + \frac{zk_{py}}{k}$),

$$\begin{aligned} I\left(x + \frac{p_x}{f}z, y + \frac{p_y}{f}z, z, p_x, p_y\right) &= \frac{1}{\mathcal{N}} \sum_{k_{gx}, k_{gy}} \left| \int \int_{\mathcal{A}(0,0)} \overline{U}(k'_x + k_{gx}, k'_y + k_{gy}, 0) \times e^{ikz \left(1 - \frac{k_x'^2 + k_y'^2}{2k^2}\right)} e^{ik'_x x + ik'_y y} dk'_x dk'_y \right|^2 \\ &= I(x, y, z, 0, 0) \end{aligned} \quad (\text{M6})$$

The right-hand side of Eq. (M6), $I(x, y, z, 0, 0)$, does not depend on p_x and p_y while the left-hand side term depends on them. Equation (M6) can be rewritten as a form of

$$I\left(x, y, z, p_x, p_y\right) = I\left(x - \frac{p_x}{f}z, y - \frac{p_y}{f}z, z, 0, 0\right) \quad (\text{M7})$$

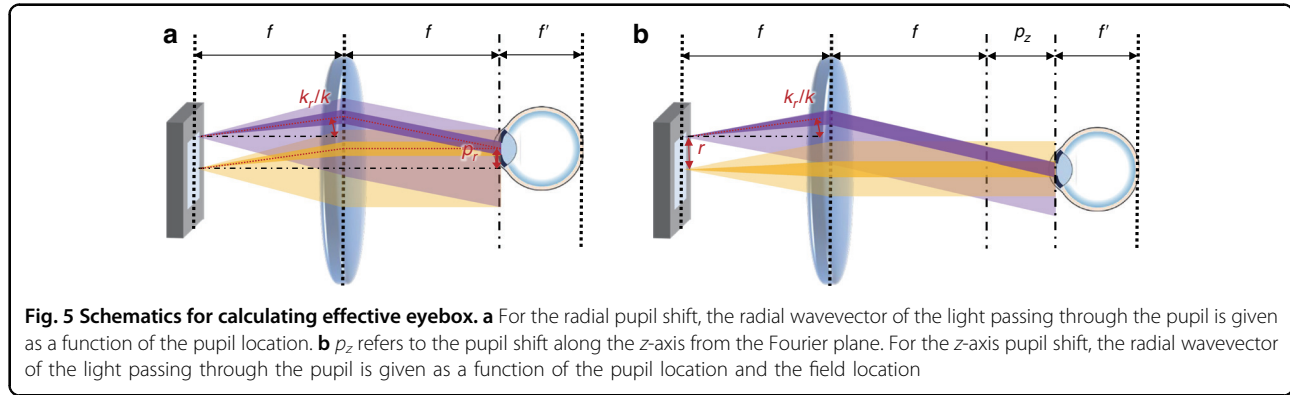
Although the reconstructed intensity is shifted by $(\frac{p_x}{f}z, \frac{p_y}{f}z)$ according to Eq. (M7), the intensity seen by an eye is not shifted because the eye is also shifted. Such a condition can be easily found by drawing the virtual image optically reconstructed by the lens (Fig. 4). In other words, the observed intensity of a hologram reconstructed by a spatially incoherent light does not depend on the pupil location and thus the hologram can be seen everywhere.

Generalization of shift-invariance along z-axis

To account for the pupil shift along the z-axis, we need to calculate the complete optical propagation from the SLM to the retina. The propagation can be calculated using the following procedure (Fig. 5b).

1. Propagate for a distance f
2. Phase altered by a lens with the focal length f
3. Propagate for a distance $f + p_z$





4. Phase altered by an eye lens with the focal length f'
5. Propagate for a distance f'

After simple calculations under Fresnel approximation, the wavefield at the retina is given as,

$$U(x, x, f + f + p_z + f')$$

$$= \frac{1}{N_0} \iint e^{-ik_x x \left(\frac{f}{f'}\right) - ik_y y \left(\frac{f}{f'}\right)} \bar{U}(k_x, k_y, 0) dk_x dk_y \quad (\text{M8})$$

where, N_0 is the normalization constant. The derivation of Eq. (M8) is provided in the Supplementary Material. The right-hand side of Eq. (M8) remains unaffected by the shift of the pupil along the z-axis, p_z , thus resulting in shift-invariant intensity formed at the retina. The only alteration is the magnification determined by the ratio of focal lengths, $\frac{f}{f'}$. In summary, by combining the result of the previous subsection and Eq. (M8), we can get the following shift-invariant property,

$$I(x, y, z, p_x, p_y, p_z) = I(x, y, z, 0, 0, 0) \quad (\text{M9})$$

Although we rigorously derived the shift-invariance of CRIS, the shift-invariance along the x , y , and z axes is not unique property of our optical system. For example, a lens with an object placed at its focal length also exhibits shift-invariance, as the virtual image forms at infinity and remains unchanged regardless of eye movement due to negligible parallax. Our study, however, demonstrates that a 3D scene can still be reconstructed even when the virtual image of the SLM is at infinity, enabled by wavefield optimization through the filtering effect of the pupil.

3-dimensional eyebox size

According to Eq. (M9), the reconstructed intensity does not differ depending on the location of the pupil. However, the fundamental limitation of the CRIS is the Fresnel approximation, and the eyebox is limited by the valid area of

the Fresnel approximation. If we consider a plane wave with a radial wavevector, (k_x, k_y) , at the SLM, then the condition for the Fresnel approximation can be expressed as

$$\frac{d}{8\lambda} \left(\frac{k_r}{k}\right)^4 \ll 1 \quad (\text{M10})$$

where, $k_r = \sqrt{k_x^2 + k_y^2}$, d is the propagation distance, λ is the wavelength of the light, and k is the magnitude of the wavevector. More specifically, the expansion of the eye box of the CRIS originates from the cancellation of the grating phase from the incident angle by the shift in the Fourier domain. However, the large incident angle breaks the approximation $d\sqrt{k^2 - k_r^2} \approx kd\left(1 - \frac{k_r^2}{2k^2}\right)$, and the cancellation becomes imperfect.

The plane wave with the radial wavevector, (k_x, k_y) , would be focused at the point $(fk_x/k, fk_y/k)$ on the Fourier plane and the light is filtered by the pupil (Fig. 5a). By assuming that the radius of the pupil is smaller than the distance between the optical axis and the center of the pupil, we can approximate all the radial wavevectors of the plane wave light passing through the pupil as (k_x, k_y) . It is simple to show that the radial wavevector magnitude of the light passing through the pupil is given as,

$$\frac{k_r}{k} = \frac{1}{f} \sqrt{p_x^2 + p_y^2} \quad (\text{M11})$$

where the pupil position is assumed as $(p_x, p_y, 0)$.

For the case where the pupil is shifted along the z-axis, k_r depends on a distance from the SLM center, r . (Fig. 5b) The wavevector magnitude can be calculated as $\frac{k_r}{k} = \frac{rp_z}{f^2}$, and the maximum wavevector magnitude is at the edge of the field of view (FoV). By rewriting in terms of FoV, the



maximum of the wavevector magnitude is given as

$$\frac{k_{r\text{MAX}}}{k} \approx \frac{1}{f} p_z \frac{\text{FoV}}{2} \quad (\text{M12})$$

By combining Eq. (M11) and Eq. (M12), the eyebox relation can be summarized as,

$$\frac{k_{r\text{MAX}}}{k} \approx \frac{1}{f} \sqrt{p_x^2 + p_y^2 + p_z^2 \left(\frac{\text{FoV}}{2} \right)^2} \quad (\text{M13})$$

According to Eq. (M10) and Eq. (M13), the eyebox can be written as,

$$\frac{1}{f} \sqrt{p_x^2 + p_y^2 + p_z^2 \left(\frac{\text{FoV}}{2} \right)^2} < \left(\frac{8\lambda}{d} \right)^{1/4} \approx \left(\frac{8\lambda}{Df^2} \right)^{1/4} \quad (\text{M14})$$

where, $D \approx d/f^2$ is the depth of the hologram in diopter and the condition “much smaller” is adjusted to “smaller” to account for the fourth power of k_r/k . The important points of Eq. (M14) are that the eyebox of the CRIS is given as an ellipsoid and its volume depends on the depth of the hologram.

By restricting pupil shift along the x -axis, i.e., $p_y = 0$, $p_z = 0$, the length of the eyebox along the x -axis can be calculated as $|p_x| < \left(\frac{8\lambda f^2}{D} \right)^{1/4}$. As a result, the half eyebox length along the x - and y -axes is given as,

$$L_e \equiv \left(\frac{8\lambda f^2}{D} \right)^{1/4} \quad (\text{M15})$$

Considering the parameters used in the numerical reconstruction, the eyebox length $2L_e$ is 12 mm, which is nearly the same as the FWHM of the PSNR, 12.2 mm, with an error of 1.6%. The slight difference can be attributed to the fact that the boundary of the approximation is not well-defined. Interestingly, the eyebox does not depend on the pixel pitch or the number of pixels of the display. Under typical conditions, the eyebox length of a glass-type display ($f = 20$ mm, $\lambda = 515$ nm, $D = 1.5$ diopters) is 11.5 mm, while that of a flat-panel-type display ($f = 1000$ mm, $\lambda = 515$ nm, $D = 1.5$ diopters) is 81.4 mm. With our numerical and experimental parameters, the theoretical eyebox volume calculated from Eq. (M14) is 3800 mm^3 , while that of conventional holograms is 3.97 mm^3 , resulting in an enhancement of ~ 1000 times.

According to Eq. (M15), the eyebox size of CRIS is independent of the pixel pitch of the SLM. Consequently, the eyebox expansion ratio decreases linearly as the pixel pitch decreases. Based on this relationship, it might be anticipated that CRIS would lose its practicality if an SLM with a pixel pitch approaching the wavelength were

developed. Nevertheless, CRIS remains advantageous in terms of accessibility and cost-effectiveness, even under such circumstances. This is because a reduction in pixel pitch, assuming a fixed display resolution, leads to a decrease in the physical size of the SLM and its field of view. To preserve the field of view and enable eyebox expansion, the resolution must increase correspondingly, significantly raising SLM costs and computational costs.

Experimental setup

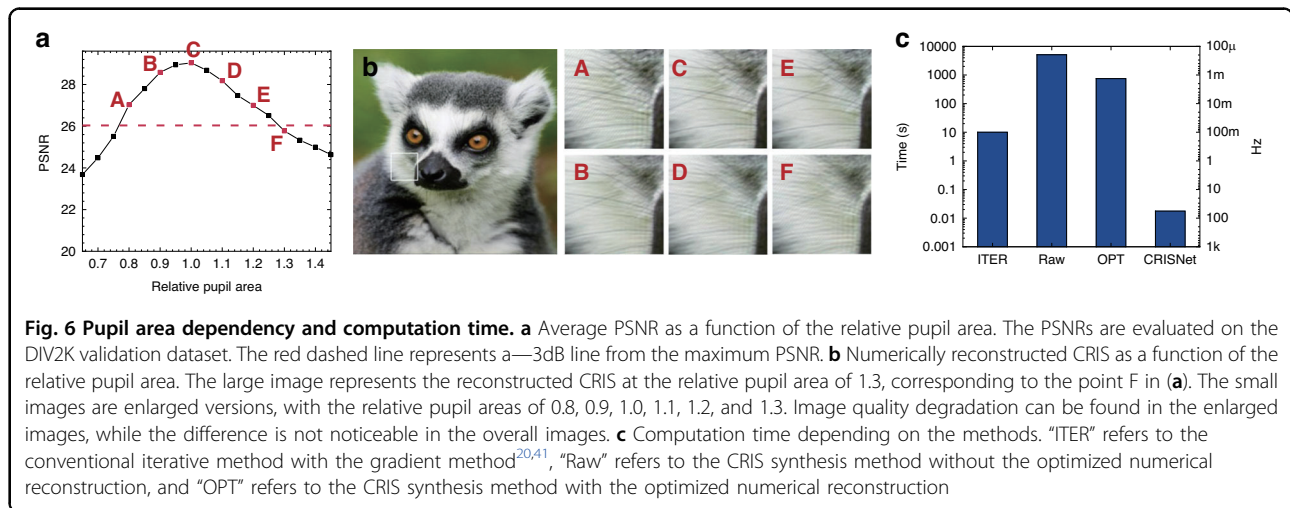
The lasers with wavelengths of 638 nm, 515 nm, and 450 nm are coupled by the single mode fiber and the lasers become spatially incoherent due to the rotating diffuser. The holographic diffuser with a divergence angle of $\pm 10^\circ$ and transmission efficiency of 85% is adopted. The spatially incoherent light collimated by the lens is illuminated through the PBS. Since an amplitude-only SLM rotates the polarization of light by an amount proportional to the modulation, only the modulated pattern of light passes through the PBS. The amplitude-only modulator, IRIS-F55 (MAY Inc.) with 1920×1080 resolution and $6.3 \mu\text{m}$ pixel pitch, is used to modulate the light, but only 1080×1080 area is used in the optical reconstruction to minimize the noise from the limited numerical aperture. It operates in grayscale at 180 Hz, which corresponds to 60 Hz for color display as it processes RGB sequentially. The RGB lasers are synchronized with the SLM, with each color illuminating during its corresponding frame.

We implemented CRIS using an amplitude-only SLM because phase-only SLMs require 4f-filtering to achieve high image quality. While holographic displays using phase-only SLMs can achieve high-quality images, their image quality becomes lower than that of amplitude-only SLMs due to the pixelated structures when 4f-filtering is not applied³⁵. Since CRIS uses numerous angles of incident light to expand the eyebox, blocking high-frequency components during 4f-filtering would also block most incident light and prohibit eyebox expansion.

Due to the lens after the SLM, the virtual image of the SLM is located at infinity, corresponding to 0 diopters. In contrast, the virtual image of the front plane is located 67 cm away from the pupil, corresponding to 1.5 diopters. The lens after the SLM forms the Fourier plane of the hologram and the aperture mounted on the motorized stage moves near the Fourier plane. As a result, the aperture emulates a moving iris of a human.

The calibration of the SLM is performed by assigning one value to the entire SLM and measuring the modulated intensity. After measuring the intensity, a fitted function of output-input is applied to the input value to correct the modulation. Although the LCoS modulation of the wavefield depends on the incident angle, the difference is negligible if the angle difference is not much larger than 10° ³⁶. To expand the viewing angle to larger than 10° ,





digital micromirror devices can be adopted to minimize the dependency of the modulation on the incident angle. The overall optical efficiency of the experimental setup is $\sim 4\%$, limited by the oversized illumination area, scattering from the diffuser and other optical components, and the unused light at zero-value pixels on the amplitude-only SLM.

The main reason for observing only a 32-fold expansion experimentally, instead of the theoretical 1000-fold, is the limited numerical aperture of the optical system. To observe the 1000-fold eyebox expansion, the lens after the SLM must accommodate all the incident angles reflected by the SLM. In the experimental setup, we used an optical system with a numerical aperture of 0.3, which allowed only a small portion of the light to pass through the lens. By using an optical system with a numerical aperture of larger than 0.8, it would be possible to achieve the theoretical maximum.

Pupil area dependency

During the synthesis of the CRIS, only the propagated light passing through the pupil area $\mathcal{A}$ contributes to the reconstruction. Variations in the pupil area result in changes in the reconstructed intensity, leading to degradation of image quality (Fig. 6a). Considering that the pupil area depends on the brightness, the pupil area used in the synthesis can be calibrated by detecting ambient brightness. However, even under fixed ambient brightness, emotional arousal can cause variations in pupil area of up to $\pm 20\%$ for a single individual^{37,38}. According to the results of the numerical simulation, the variation in pupil area of $\pm 20\%$ can reduce the PSNR by 2 dB, which is not significant in most cases, given the transient nature of the variation. By estimating a 3 dB degradation in the PSNR, the FWHM of the relative pupil area is $\sim 51\%$. In addition, interference and

speckle-like patterns are not observed in the reconstructions with mismatched pupil areas because additional and subtractive light components are incoherent (Fig. 6b).

Details of CRISNet

Real-time synthesis of the CRIS was achieved using the CRISNet and unsupervised training was adopted for the optimal image quality³⁹. However, the numerical reconstruction of the CRIS requires 500 times more computational time than coherent propagation, making it an obstacle during training (Fig. 6c). The major computational challenge arises from the requirement to compute coherent propagations for a large number of different incident angles, with an expanded Fourier domain to accommodate off-axis propagations. Through optimization of the numerical reconstruction, such as employing a random selection of 100–300 incident angles instead of 400, along with cached tensors, the computation time can be reduced to 14.7%. CRISNet was trained by adopting the optimized reconstruction method, and the computation time was further reduced to $1/40000$, enabling operation at 57 Hz on an NVIDIA RTX4090 without image quality degradation. The network was trained with a learning rate of 0.0001 for 100 epochs using the ADAM optimizer. The DIV2K training dataset²² was used to train the CRISNet, and each image was randomly cropped to a size of 1024×1024 to reduce the risk of overfitting. Moreover, during training, we observed a continuous increase in numerically reconstructed PSNRs synthesized from the DIV2K validation dataset, distinct from the DIV2K training dataset, indicating effective generalization and avoidance of overfitting. Since the DIV2K training dataset contains only 2D images, we generated random depth maps to synthesize multi-depth scenes. Each depth map is divided into an 11×11 grid, with each sub-patch in the



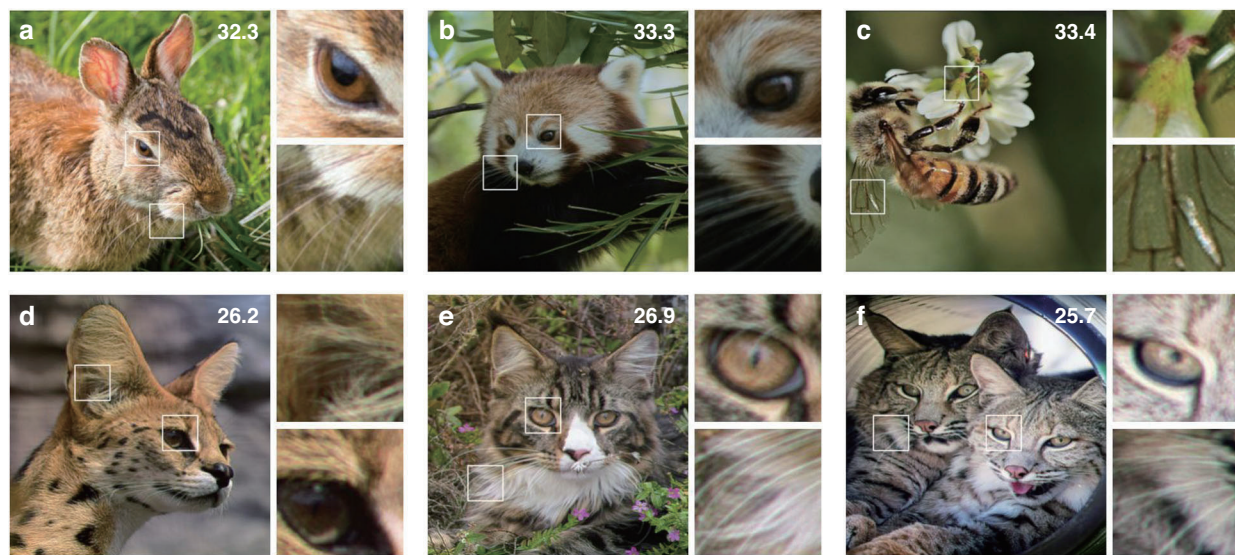


Fig. 7 Reconstructed images of CRIS. Images in the Flickr2K dataset are numerically reconstructed (a–c) and optically reconstructed (d–f) when the location of the pupil is at the center of the eyebox. The PSNR in dB is marked in the top right corner, and the insets present enlarged images

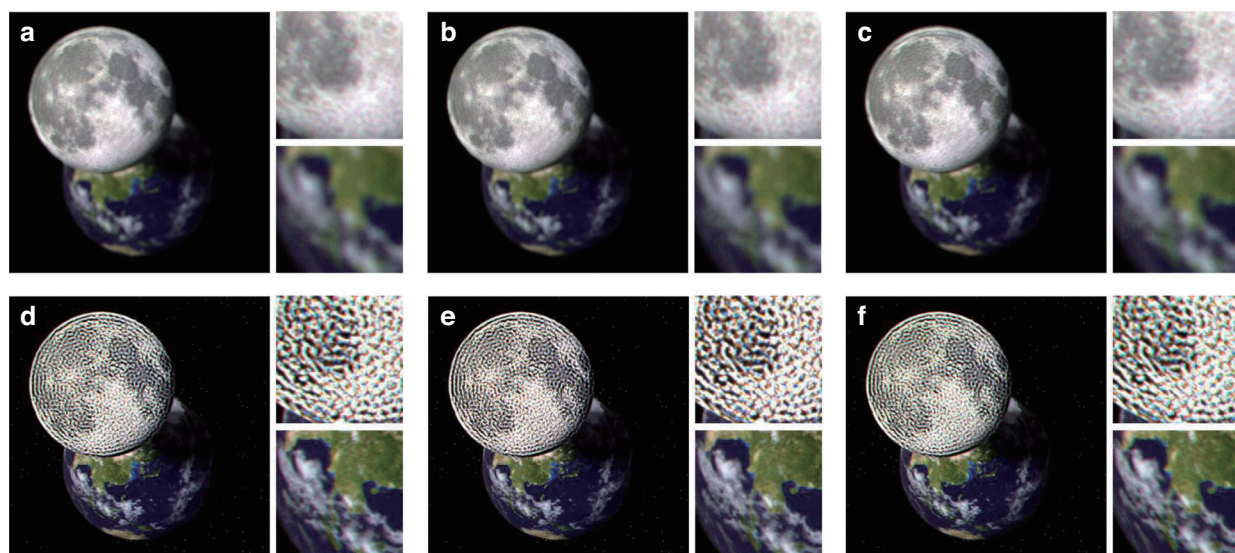


Fig. 8 3D reconstruction at different pupil locations. The 3D scene is optically reconstructed for different pupil locations when the camera focus is on the front plane (a–c) and the SLM plane (d–f). Each pupil location of (a–c) corresponds to the points marked with the letters A, B, and C in Fig. 4e

grid assigned a uniform value randomly selected from a predefined range⁴⁰.

CRISNet consists of 20 residual blocks, each containing two Conv2D layers, two leaky ReLU layers, and a shortcut connection. Each Conv2D layer has 32 input channels and 32 output channels, except for the first and last layers. The first layer has 4 input channels, and the last layer has 3 output channels, corresponding to the input and output channels of the network (Figs. 7 and 8).

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. RS-2023-00245184 and RS-2024-00341142). The main calculations were performed by using the supercomputing resource of the Korea Institute of Science and Technology Information (KISTI).

Author contributions

Y.S. and D.Y. conceived the idea and wrote the manuscript. D.Y. performed the numerical simulation and experiment.



Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Conflict of interest

The authors declare no competing interests.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41377-025-01823-z>.

Received: 25 September 2024 Revised: 23 February 2025 Accepted: 9 March 2025

Published online: 12 May 2025

References

- Gabor, D. A new microscopic principle. *Nature* **161**, 777–778 (1948).
- Erf, R. *Holographic Nondestructive Testing* (Elsevier, 2012).
- Ashley, J. et al. Holographic data storage technology. *IBM J. Res. Dev.* **44**, 341–368 (2000).
- Park, Y., Depeursinge, C. & Popescu, G. Quantitative phase imaging in biomedicine. *Nat. Photonics* **12**, 578–589 (2018).
- Poon, T.-C. *Digital Holography and Three-Dimensional Display: Principles And Applications* (Springer Science & Business Media, 2006).
- Kim, M. K. & Kim, M. K. *Digital Holographic Microscopy* (Springer, 2011).
- Xu, W., Jericho, M., Meinertzhagen, I. & Kreuzer, H. Digital in-line holography for biological applications. *Proc. Natl Acad. Sci. USA* **98**, 11301–11305 (2001).
- Javidi, B. & Tajahuerce, E. Three-dimensional object recognition by use of digital holography. *Opt. Lett.* **25**, 610–612 (2000).
- Liu, J.-P., Tahara, T., Hayasaki, Y. & Poon, T.-C. Incoherent digital holography: a review. *Appl. Sci.* **8**, 143 (2018).
- Tahara, T. Review of incoherent digital holography: applications to multi-dimensional incoherent digital holographic microscopy and palm-sized digital holographic recorder-holosensor. *Front. Photonics* **2**, 829139 (2022).
- Rosen, J. & Brooker, G. Digital spatially incoherent fresnel holography. *Opt. Lett.* **32**, 912–914 (2007).
- Poon, T.-C. Optical scanning holography—a review of recent progress. *J. Opt. Soc. Korea* **13**, 406–415 (2009).
- Vijayakumar, A. & Rosen, J. Interferenceless coded aperture correlation holography—a new technique for recording incoherent digital holograms without two-wave interference. *Opt. Express* **25**, 13883–13896 (2017).
- Peng, Y., Choi, S., Kim, J. & Wetzstein, G. Speckle-free holography with partially coherent light sources and camera-in-the-loop calibration. *Sci. Adv.* **7**, eabg5040 (2021).
- Lee, S. et al. Light source optimization for partially coherent holographic displays with consideration of speckle contrast, resolution, and depth of field. *Sci. Rep.* **10**, 18832 (2020).
- Kuo, G., Schiffer, F., Lanman, D., Cossairt, O. & Matsuda, N. Multisource holography. *ACM Trans. Graph.* **42**, 1–14 (2023).
- Ghoushchi, V. P., Aas, M., Ulusoy, E. & Ürey, H. Effect of spatial coherence of led sources on image resolution in holographic displays. in *Advances in Display Technologies VII* Vol. 10126, 63–70 (SPIE, 2017).
- Blanche, P.-A. Holography, and the future of 3d display. *Light. Adv. Manuf.* **2**, 446–459 (2021).
- Chang, C., Bang, K., Wetzstein, G., Lee, B. & Gao, L. Toward the next-generation vr/ar optics: a review of holographic near-eye displays from a human-centric perspective. *Optica* **7**, 1563–1578 (2020).
- Yang, D. et al. Diffraction-engineered holography: beyond the depth representation limit of holographic displays. *Nat. Commun.* **13**, 1–11 (2022).
- Kavaklı, K., Itoh, Y., Ürey, H. & Aksit, K. Realistic defocus blur for multiplane computer-generated holography. In *Proc. IEEE Conference Virtual Reality and 3D User Interfaces (VR)* 418–426 (IEEE, 2023).
- Agustsson, E. & Timofte, R. Ntire 2017 challenge on single image super-resolution: dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* (IEEE, 2017).
- Kuo, G., Waller, L., Ng, R. & Maimone, A. High resolution étendue expansion for holographic displays. *ACM Trans. Graph.* **39**, 66–1 (2020).
- Tseng, E. et al. Neural étendue expander for ultra-wide-angle high-fidelity holographic display. *Nat. Commun.* **15**, 2907 (2024).
- Park, J., Lee, K. & Park, Y. Ultrathin wide-angle large-area digital 3d holographic display using a non-periodic photon sieve. *Nat. Commun.* **10**, 1–8 (2019).
- Gopakumar, M. et al. Full-colour 3d holographic augmented-reality displays with metasurface waveguides. *Nature* **629**, 791–797 (2024).
- Kim, J. et al. Foveated ar: dynamically-foveated augmented reality display. *ACM Trans. Graph.* **38**, 99–1 (2019).
- Yarag, F., Kang, H. & Onural, L. State of the art in holographic displays: a survey. *J. Disp. Technol.* **6**, 443–454 (2010).
- Liu, T. et al. Research progress of large viewing-area 3d holographic near-eye display. *Laser Photonics Rev.* **18**, 2300641 (2023).
- Stein, N. et al. A comparison of eye tracking latencies among several commercial head-mounted displays. *i-Percept.* **12**, 204166952098338 (2021).
- Shigematsu, O., Naruse, M. & Horisaki, R. Computer-generated holography with ordinary display. *Opt. Lett.* **49**, 1876–1879 (2024).
- Maimone, A., Georgiou, A. & Kollin, J. S. Holographic near-eye displays for virtual and augmented reality. *ACM Trans. Graph.* **36**, 1–16 (2017).
- Zhao, Y., Cao, L., Zhang, H., Kong, D. & Jin, G. Accurate calculation of computer-generated holograms using angular-spectrum layer-oriented method. *Opt. Express* **23**, 25440–25449 (2015).
- Lee, B., Kim, D., Lee, S., Chen, C. & Lee, B. High-contrast, speckle-free, true 3d holography via binary CGH optimization. *Sci. Rep.* **12**, 1–12 (2022).
- Yang, D. & Lee, H. -S. Direct amplitude-only hologram realized by broken symmetry. *Sci. Adv.* **10**, eadp1205 (2024).
- Lizana, A. et al. Influence of the incident angle in the performance of liquid crystal on silicon displays. *Opt. Express* **17**, 8491–8505 (2009).
- Partala, T. & Surakka, V. Pupil size variation as an indication of affective processing. *Int. J. Hum. Comput. Stud.* **59**, 185–198 (2003).
- Oliva, M. & Anikin, A. Pupil dilation reflects the time course of emotion recognition in human vocalizations. *Sci. Rep.* **8**, 4871 (2018).
- Shimobaba, T. et al. Deep-learning computational holography: a review. *Front. Photonics* **3**, 8 (2022).
- Yu, H. et al. Deep learning-based incoherent holographic camera enabling acquisition of real-world holograms for holographic streaming system. *Nat. Commun.* **14**, 3534 (2023).
- Chakravarthula, P., Peng, Y., Kollin, J., Fuchs, H. & Heide, F. Wirtinger holography for near-eye displays. *ACM Trans. Graph.* **38**, 1–13 (2019).

