

Deep Learning-Based Artificially Focused CGH Method with Real-World Objects Using Eye Tracking for Holographic Near-Eye Displays

Tuvshinjargal Amgalan¹, Munkh-Uchral Erdenebat¹, Nyamsuren Darkhanbaatar¹,
Ki-Chul Kwon¹ and Nam Kim^{1*}

¹School of Information and Communication Engineering, Chungbuk National University, Cheongju, Chungbuk 361-763, South Korea

Abstract

We propose a novel computer-generated holography algorithm based on a deep neural network, incorporating an eye-tracking-based artificial focusing control system for near-eye holographic displays. By leveraging eye tracking, the system accurately identifies the focus point on a real object and applies blur to areas outside the focal point using depth information, thereby enhancing visual fidelity; and enhances depth perception, image clarity while effectively minimizing crosstalk and reducing the computational complexity typically associated with deep-neural network.

Author Keywords

Holographic display; Near-eye display; Computer-generated hologram (CGH); Synthetic-depth map; Eye-tracking.

1. Introduction

Holographic displays represent a promising approach to achieving compact near-eye reconstructions with a comfortable viewing experience, making them highly suitable for augmented and virtual reality applications [1]. Recently, deep learning-based computer-generated holography (CGH) methods have been rapidly developed to address several inherent challenges in near-eye displays, including focus detection, form factor, aberration correction, and vision correction [2]. However, for 3D reconstruction, multi-depth layer methods are often used due to the complexities involved in training, computation, and the requirements for extensive input data. These methods are also prone to artifacts caused by overlapping and occlusion [3]. In holographic reconstruction, evaluating multi-depth errors between reconstructed images at various diffraction distances and the target scene is critical. Therefore, the number of layers involved in the hologram generation process has a direct impact on the quality, speed, and distortion of the reconstruction. As a result, the number of layers involved in the hologram creation process has a direct influence on the quality, processing speed,

and level of distortion in the reconstruction. To minimize multi-depth discrepancies using an eye-gaze tracking system, the deep neural network can be trained to generate clearer, more accurate reconstructions that better represent the target scene at various depths, thereby enhancing the visual fidelity of near-eye holographic displays. Eye-tracking technology further enhances these systems by dynamically adjusting the focus and applying artificial blur to peripheral regions, mimicking natural human vision [4,5]. This approach reduces computational demands while improving the immersive quality of the display. By leveraging eye tracking, the system accurately identifies the focus point on a real object and applies blur to areas outside the focal point using depth information, thereby enhancing visual fidelity.

In this research, we propose an artificial focusing CGH method based on deep learning and eye-tracking system for holographic near-eye displays. By concentrating computational resources on the real objects in the display that correspond to the gaze point of attention; while blurring peripheral regions, this approach creates a more natural viewing experience. It reduces the complexity of rendering full 3D holograms across the entire field of view, thereby enhancing performance and reducing computational load without compromising the immersive experience. The proposed system's effectiveness is validated through simulations and experimental results, demonstrating its potential for near-eye holographic displays.

2. Proposed method

The proposed holographic near-eye display system integrates advanced deep learning techniques, synthetic depth mapping, and eye-tracking technology to enhance visual fidelity and computational efficiency. The workflow is designed to dynamically focus on the observer's gaze point while generating high-quality reconstructions of holographic near eye display. Figure 1 illustrates the schematic diagram featuring the eye-tracking system and the CNN based CGH generation process.

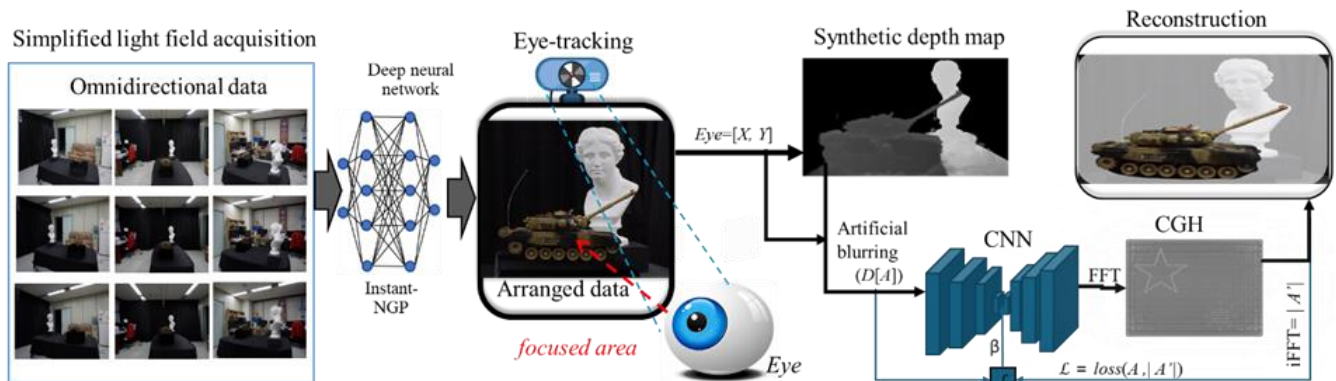


Fig 1. The schematic diagram of the proposed system, with artificial focused reconstruction based on the integration of eye-tracking technology and the conventional neural network-based CGH generation process.

2.1. Data acquisition and generate 3D arranged data from real world objects based on deep neural networks

The process begins with the capture of omnidirectional perspectives of real-world objects using a simplified light field camera [6]. The captured data is organized and processed through an Instant-NGP model with hash-grid encoding, rapidly generating high quality 3D data of real-world objects within a virtual environment. The acquisition of 54 perspectives of real-world objects using a simplified light field acquisition system as shown in Fig 2 left side and generated arranged data, shown in the left panel of the lower section. This system captures six poses of the object, with each pose recorded three times as the object is shifted to a translation stage and acquired data as shown in Fig 2 right side. The automated nature of this setup ensures efficient and accurate data collection across multiple viewpoints.

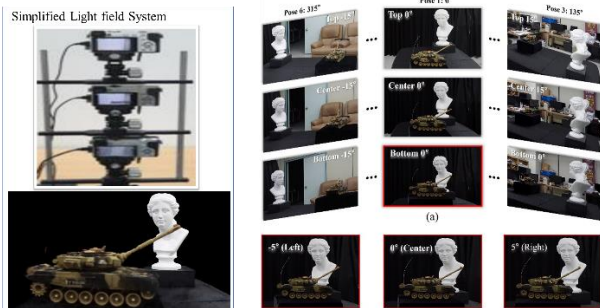


Fig 2. Data acquisition from real objects with 6 poses using Simplified light field system and generated arranged data.

2.2. Eye-Tracking and Artificial Blur

In the case of an eye-tracking system determines the observer's gaze point in real time, identifying the area of focus on the displayed content. Based on this information, artificial blur is applied to peripheral regions outside the focus point, using depth values from the synthetic depth map [7]. From the arranged 3D data, depth maps and position maps are rendered in real time. These two maps are then combined to generate a synthetic depth map, which provides precise depth information across the scene, including objects at varying distances. The process of obtaining the synthetic depth map is illustrated in Fig 3. The first position, on the right side of the object, is shown at the top and renders the near field at 200 mm. The second position, on the left side of the object, is shown at the bottom and renders the far field at 7000 mm. For each position, RGB, depth, and position map images are rendered from the arranged data. These images are then used to obtain the synthetic depth map, which provides more accurate depth information.

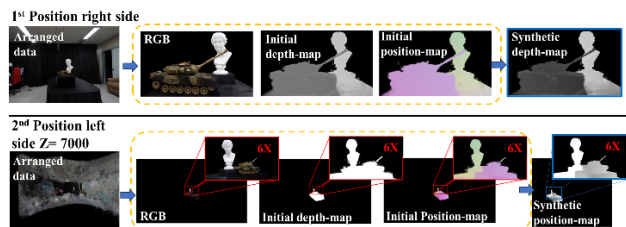


Fig 3. Process of generating the synthetic depth map, showing the rendering of near-field (200 mm) and far-field (7000 mm) from the arranged data, along with the corresponding RGB, depth, position and synthetic-depth maps for each position.

2.4 Holographic Display Implementation.

In this section, we describe the real-time rendering and integration of holographic data with the real-world environment on a near-eye holographic display. The generated holographic content, including synthetic depth maps and 3D reconstructions, is processed and visualized through a combination of eye-tracking technology and synthetic depth maps to improve the visual fidelity and enhance the viewer's immersive experience. First, the reconstructed holograms are combined with the real-world environment to create a seamless augmented reality experience. This involves synchronizing the synthetic depth map with the observer's gaze point, which is tracked using an eye-tracking system. The holographic optical elements (HOEs), recorded with a lens function, play a crucial role in projecting the holographic content onto the display while maintaining a natural integration with the real environment.

The display system utilizes a wide-angle, near-eye configuration, allowing for a large field of view (FOV) that enhances the immersion of the experience. Real-time rendering ensures that the displayed content adapts to the observer's movements and gaze direction, maintaining consistent visual fidelity. Additionally, the system leverages the synthetic depth map to apply artificial blur to peripheral areas outside the focal point, optimizing computational efficiency while ensuring that the viewer's attention remains focused on the central, clear 3D content. Finally, CNN based CGHs are displayed with high clarity, utilizing the depth cues provided by the synthetic depth maps, allowing for realistic 3D visualizations and significantly reducing computational load compared to traditional diffraction-based methods, as it does not require intensive real-time computation of light propagation across multiple layers. This is especially important for near-eye displays, where computational efficiency is crucial for maintaining high frame rates and visual fluidity.

In the experiment, the simplified light field image acquisition system consists of three Sony a6000 cameras captured 54 perspectives with 6000x4000 px resolution. From the captured perspectives, the instant-NGP model regenerated high-quality 3D arranged data with a total of 13,623,184 parameters in 70 seconds with a loss of 0.000747. The eye-tracking system utilized a

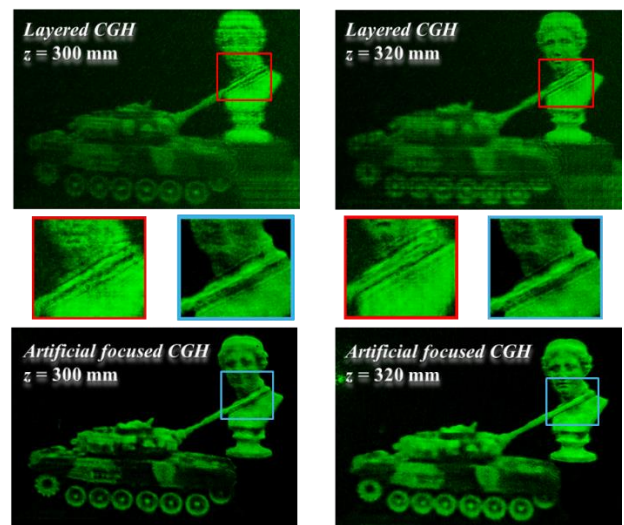


Fig 4. Reconstructed optical images. Top row: Layered CGHs at z=300 mm (tank) and z=320 mm (Venus). Bottom row: Artificially focused CGHs at the same focus distances.

standard 1920×1080 2D Microsoft LifeCam™ camera, while the graphic user interface (GUI) was developed based on OpenCv architecture applications. The CNN total epoch was 50 times, and the learning rate is 0.001 with Adam optimization and total training loss is 0.00183, validation loss is 0.00198. Figure 4 shows the reconstructed images from optical experiments. The first row displays the layered CGHs, with the focus at $z=300$ mm (tank focused) and $z=320$ mm (Venus focused). The second row illustrates the artificial-focused CGHs, with separate focus points at $z=300$ mm (tank focused) and $z=320$ mm (Venus focused).

3. Conclusion

This study demonstrates an innovative approach to holographic near-eye displays by integrating advanced eye-tracking systems, deep learning algorithms, and holographic optical elements. The proposed method efficiently addresses key challenges such as multi-depth discrepancies, computational complexity, and visual fidelity. By leveraging synthetic depth maps, artificial focusing, and real-time blur application based on gaze tracking, the system achieves natural and immersive viewing experiences. Experimental results validate the effectiveness of the approach, showcasing high-quality reconstructions seamlessly integrated with real-world environments on wide-angle, near-eye displays.

These findings contribute to the advancement of holographic display technologies, paving the way for more practical and efficient implementations in augmented and virtual reality applications. Future work will focus on refining the computational framework, further optimizing hardware integration, and exploring broader applications for this system.

Acknowledgements

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2023-2020-0-01846) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

4. References

1. J. Carmigniani, B. Furht, M. Anisetti, P. Ceravolo, E. Damiani, M. Ivkovic. "Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*." 2010 Dec 14;51(1):341–77.
2. L. Chen, R. Zhu, H. Zhang. "Speckle-free compact holographic near-eye display using camera-in-the-loop optimization with phase constraint." *Optic Express*. 2022;30(26):46649.
3. L. Shi, B. Li, W. Matusik. "End-to-end learning of 3D phase-only holograms for holographic display." *Light Sci Appl* 11, 247 (2022).
4. A. Patney, M. Salvi, J. Kim, A. Kaplanyan, C. Wyman, N. Benty, et al. "Towards Foveated Rendering for Gaze-Tracked Virtual Reality." *ACM Transactions on Graphics (TOG)*,35(6):179, 2016.
5. T. Amgalan, M-U. Erdenebat, A. Khuderchuluun, J-R. Jung, S-K. Gil, and Nam-Kim, "Artificial focusing method of 3D CGH based on eye-tracking controller for holographic display," in 2024 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR), Technical Digest Series (Optica Publishing Group, 2024), paper Fr1G_3.
6. M.-U. Erdenebat, T. Amgalan, A. Khuderchuluun, O.-S. Nam, Jeon S.-H, Kwon K.-C, Kim N, "Comprehensive high-quality three-dimensional display system based on a simplified light-field image acquisition method and a full-connected deep neural network," *Sensors* 23(14), 6245 (2023).
7. Amgalan T, Wu H.-Y, Erdenebat M.-U, Khuderchuluun A, Kwon K.-C, Dashdondov K, et al. "Advanced depth-layering using a synthetic-depth map and deep neural network for real-world object-based holographic displays," *Opt. Express* 32, 47011-47025 (2024)