

Smoothing of inter-layer edge artifacts in depth-map computer-generated holograms

SUNGJAE PARK, JONGHYUN LEE,  WONWOO CHOI,  SHIN-WOONG PARK,  AND HWI KIM* 

Department of Electronics and Information Engineering, College of Science and Technology, Korea University, Sejong-Campus, 2511 Sejong-ro, Sejong 30019, Republic of Korea

*Corresponding author: hwikim@korea.ac.kr

Received 8 September 2022; revised 8 November 2022; accepted 14 November 2022; posted 17 November 2022; published 12 December 2022

In the depth-map computer-generated hologram (CGH), inter-layer edge artifacts are observed in the discontinuous edges of section-wise depth-map objects. CGH synthesis, utilizing the hybrid smoothing method of silhouette masking and edge-apodization, alleviates unwanted inter-layer edge artifacts. The proposed method achieves improved de-artifact filtering that generates holographic images closer to the ground truth image of the depth-map object unattainable by the conventional CGH synthesis method. © 2022 Optica Publishing Group

<https://doi.org/10.1364/OL.475282>

Holographic displays are based on a wave-optic technology that uses the diffraction and interference of light to generate a coherent light image field in a viewer's retina plane. In terms of the quality of the holographic three-dimensional (3D) image, the best result shows real-world 3D scenes as they are. The essential feature of real-world 3D scenes is that the light field of the image is an incoherent superposition of spherical light waves from a 3D point cloud object. In particular, the 3D scene under incoherent illumination represents the ground truth image and preserves the accommodation cues that should be targeted by holographic 3D imagery. However, since a conventional holographic display uses a coherent light source, the reproduced holographic images have deficiency from the incoherent ground truth image.

Computer-generated holograms (CGHs) can be represented in several ways. CGH synthesis algorithms are generally classified into point cloud [1], triangle mesh [2], and depth-map models [3,4] according to the way they model 3D scenes. The depth-map CGH model based on fast Fourier transform (FFT) has been widely used due to its modeling simplicity and computational efficiency. Depth maps follow a standard 2.5-dimensional 3D content format, which is popularly distributed in the 3D broadcasting industry.

Recently, for phase-only CGH optimization, the stochastic gradient descent (SGD) method [5] and the deep neural network technique have been developed to generate photorealistic holographic images, the object model of which is based on the depth-map 3D image content [6]. The SGD method can be combined with techniques such as camera-in-the-loop training and deep learning to obtain almost ideal high-quality reconstruction results [5,7]. However, while most existing research has focused

on two-dimensional (2D) objects, depth-map 3D objects [8] still present a challenge. Figure 1 presents a traditional depth-map CGH, where edge artifacts are observed due to edge diffraction at the boundaries between adjacent objects of different depths. The term edge artifact is ascribed to inter-layer gaps, but emphasizes the artifacts of coherent light images, which do not appear in the natural incoherent light image generation process. As the depth gap of adjacent layers increases, so inter-layer edge artifacts become more severe.

In this Letter, we describe a method for generating a ground truth model with incoherent illumination and address a CGH synthesis method for enhancing the similarity of the CGH image to the incoherent ground truth. Results from both the proposed method and conventional depth-map CGH methods are compared to the ground truth image and their visual characteristics discussed.

The ground truth image of a 3D scene can be generated based on wave optics theory. Assume that there is an axially distant dual-layer opaque object (an opaque circle and triangle) as shown in Fig. 2(a). Then, under natural incoherent illumination, the viewer sees the incoherent sum of the light waves reflected by the dense point clouds comprising the target 3D object. The ground truth images of the target dual-layer object are presented according to viewers' pupil size in Fig. 2(b). Previous studies [3,4] show coherent results because all light coherence is considered in the cascaded Fresnel transform (CdFr) and the inverse cascaded Fresnel transform (ICdFr):

$$F_L = \text{CdFr}\{P_L; d, f_d, \rho\} \quad (1)$$

$$F' = \text{ICdFr}\{F_L; d, f_d\}, \quad (2)$$

where F_L and P_L are the complex light field distributions at the retina plane and the object plane, respectively, f_d is the eye focal length given by $f_d = d \times d_{eye} / (d + d_{eye})$, and ρ is the eye pupil radius. The observation of the incoherent image light field assumes that the object is a set of incoherent point emitted sources $P_L(x_m, y_n)$. The total field P_L is decomposed by $P_L = \sum_m \sum_n P_L(x_m, y_n)$. The observed light image pattern G_L in the retina plane (x_2, y_2) is represented by the modified incoherent sum of spherical waves emitted from $P_L(x_m, y_n)$ as

$$G_L(x_2, y_2) = \sum_{m=1}^M \sum_{n=1}^N |\text{CdFr}\{P_L(x_m, y_n); d, f_d, \rho\}|^2. \quad (3)$$

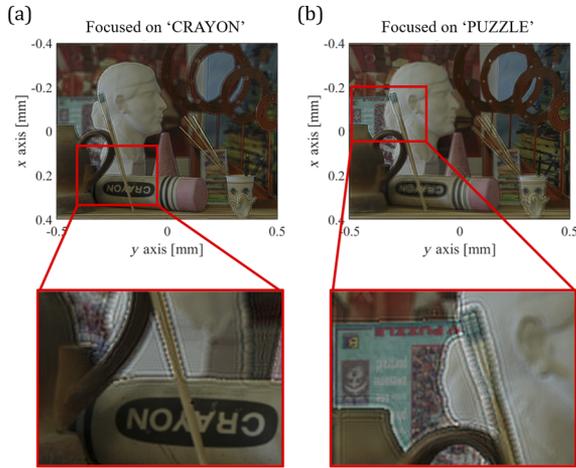


Fig. 1. Inter-layer edge-diffraction artifacts observed in the reconstructed image of a traditional depth-map CGH with focus on (a) ‘CRAYON’, and (b) ‘PUZZLE’.

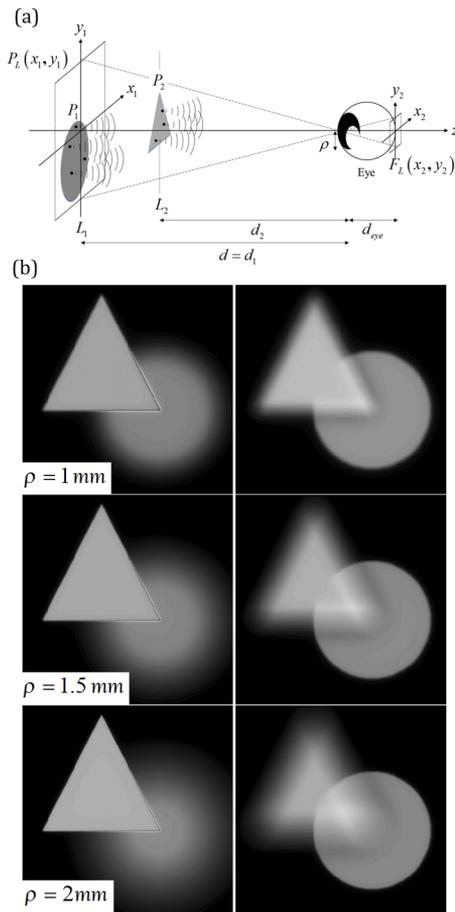


Fig. 2. (a) Schematic for incoherent wave-optic simulation, and (b) ground truth accommodation effect according to pupil radius: focusing on the triangle (the left panel G_{L_2}) and circle (the right panel G_{L_1}).

The method of generating the ground truth G_{L_2} focused on the front object (triangle) P_{L_2} of L_2 located in front of the observer’s reference is explained below. After propagating the L_1 object (circle) P_{L_1} located behind the observer’s reference to the L_2

position, the defocused light distribution field of the circle is calculated by applying the triangle’s silhouette mask M_2 . The conventional silhouette mask M_L is defined by

$$M_L(x_2, y_2) = \begin{cases} 0 & P_L(x_1, y_1) > 0 \\ 1 & P_L(x_1, y_1) = 0 \end{cases}. \quad (4)$$

The final expression obtained using Eq. (3) to reflect the incoherent properties of light in all these processes is as follows:

$$G_{L_2} = \sum_{m=1}^M \sum_{n=1}^N |F_{m,n}^{1,2} \times M_2|^2 + \sum_{m=1}^M \sum_{n=1}^N |F_{m,n}^{2,2}|^2, \quad (5)$$

where $F_{m,n}^{1,2} = CdFr\{P_{L_1}(x_m, y_n); d_1, f_{d_2}, \rho\}$,

$$F_{m,n}^{2,2} = CdFr\{P_{L_2}(x_m, y_n); d_2, f_{d_2}, \rho\},$$

and P_{L_1} and P_{L_2} are the circle and the triangle located on the L_1 and L_2 planes, respectively.

The method of generating the ground truth G_{L_1} focused on the rear object (circle) in L_1 located behind the observer’s reference follows. After propagating P_{L_1} located in L_1 to the position of L_2 , apply the silhouette mask M_2 of P_{L_2} and backpropagate to the position of L_1 once more. The ground truth G_{L_1} is generated as follows by linearly combining this result with the result of backpropagation of P_{L_2} located in L_2 to the position of L_1 :

$$G_{L_1} = \sum_{m=1}^M \sum_{n=1}^N |F_{m,n}^{1,1}|^2 + \sum_{m=1}^M \sum_{n=1}^N |F_{m,n}^{2,1}|^2, \quad (6)$$

where $F_{m,n}^{1,1}$ is the observed field of the $(m, n)^{\text{th}}$ point at the L_1 plane occluded by the silhouette mask M_2 . $F_{m,n}^{1,1}$ is represented by the observation of the occluded field projected to the L_1 plane, P' , $F_{m,n}^{1,1} = CdFr\{P'; d_1, f_{d_1}, \rho\}$, where P' is expressed as

$$P' = ICdFr\{CdFr\{P_{L_1}(x_m, y_n); d_1, f_{d_2}, \rho\} \times M_2; d_1, f_{d_2}\}.$$

The observation of the front object (triangle) is expressed by

$$F_{m,n}^{2,1} = CdFr\{P_{L_2}(x_m, y_n); d_2, f_{d_1}, \rho\}. \quad (7)$$

The right and left panels of Fig. 2(b) present the ground truth images of G_1 and G_2 with varied pupil radii. As the pupil radius

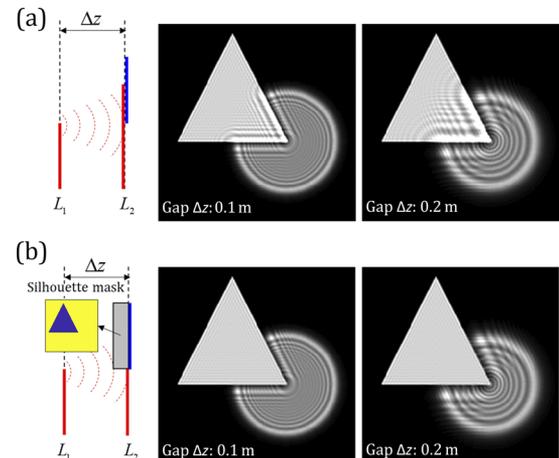


Fig. 3. (a) Principle of edge diffraction in traditional depth-map CGH and reconstructed image according to the gap Δz (0.1 m and 0.2 m) between two objects; (b) schematic diagram of silhouette mask method and reconstructed images according to gap Δz (0.1 m and 0.2 m).

is reduced, the degree of defocus blurring in the observed image decreases. Actually, those images are the ultimate 3D image which we want to obtain through holographic display.

However, conventional CGH using a coherent light source inevitably produces the optical artifacts induced by the

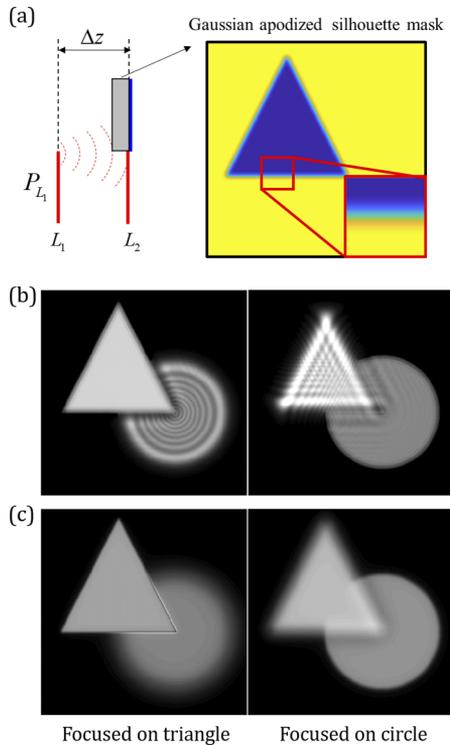


Fig. 4. (a) Example of a Gaussian apodized silhouette mask using GAF, (b) reconstructed images of depth-map CGH using GAF, and (c) ground truth with pupil radius 1 mm.

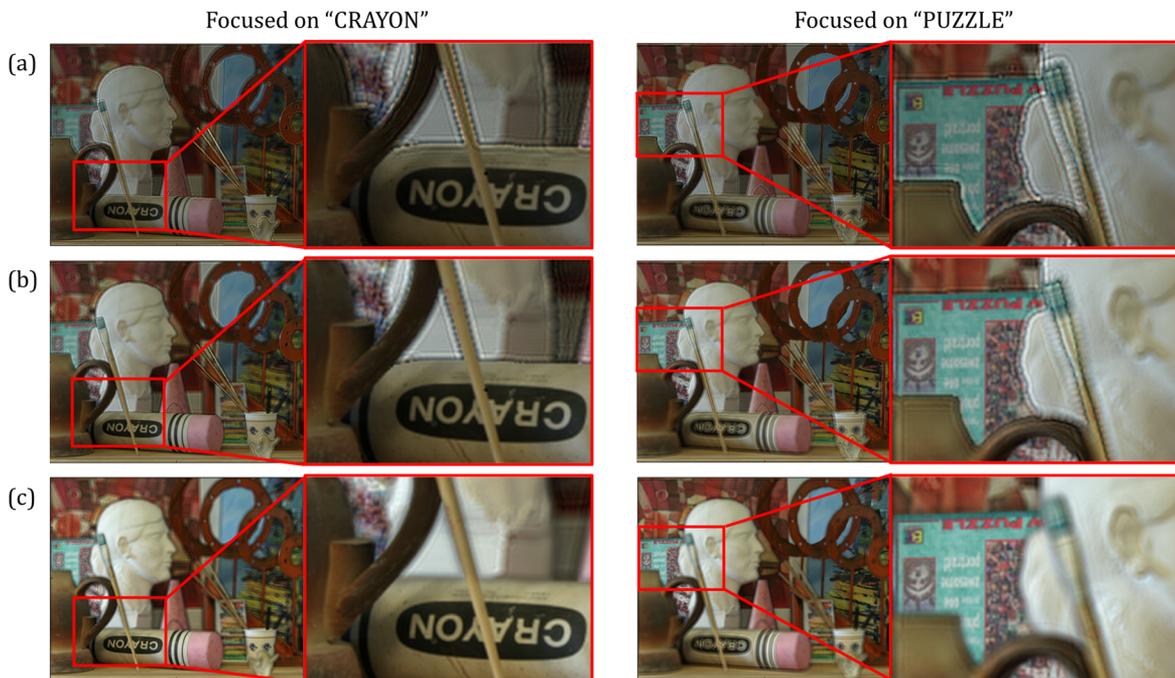


Fig. 5. Reconstructed image of full-color depth-map CGH [RGB-D data type: primary colors of red (633 nm), green (532 nm), and blue (473 nm)] using (a) traditional depth-map method, (b) silhouette masking method, and (c) GAF method. The image resolution, pixel pitch of CGH, and maximum depth are set to 1600×2000 , $20 \mu\text{m} \times 20 \mu\text{m}$, and 0.5 m, respectively. The distances from the CGH plane to the ‘PUZZLE’ plane, the plaster statue and brush plane, and the ‘CRAYON’ plane are 0.23 m, 0.31 m, and 0.39 m, respectively.

diffraction and interference of the light waves. Since the depth-map CGH calculated by the traditional method sequentially calculates and synthesizes the image information of each layer according to a stepwise model, inter-layer diffraction artifacts are generated around sharp edges of 3D objects (Fig. 3). The edge diffraction of the object at L_1 affects the object at L_2 through the axial distance Δz .

Figure 3(a) shows that as the gap between two objects increases, the area affected by edge diffraction broadens and its intensity appears to increase. The common technique to achieve an improved occlusion effect is the silhouette masking method, as illustrated in Fig. 3(b) [9]. The silhouette masking method prevents noise by blocking the edge diffraction that occurs in the rear layers. The effectiveness of silhouette masking is shown in Fig. 3(b). Note that the numerical simulation performed in Fig. 1 employed silhouette masking.

However, the unnatural defocus effect is still observed on the edge of the rear object. To improve the similarity of CGH reconstruction images to the incoherent ground truths, we propose a method using a gray-scale silhouette mask designed by convolving a Gaussian apodizing filter (GAF) G_δ to the conventional binary silhouette mask as presented in Fig. 4. The conventional silhouette mask in Fig. 3(b) is a binary mask, inducing strong edge-diffraction interference. The tailored edge of the gray-scale silhouette mask shown in Fig. 4(a) reduces edge diffraction [Fig. 3(b)]. The GAF tailors the discontinuous edges to smoothly apodized Gaussian profile edges. The GAF method proceeds the same way as the conventional depth-map CGH, except for the application of the Gaussian apodization filter to the layered images and the binary silhouette mask. Optimal design of the GAF significantly alleviates inter-layer edge diffraction in the depth-map CGH without losing the natural accommodation effect.

The GAF method is described as follows. First, calculate $F_{L_1}^{1,2}$, the optical field at the retina plane, which is the observation

of P_{L_1} with the eye-focus on the L_2 plane. The GAF should be applied to the image information P_{L_1} at this step:

$$F_{L_1}^{1,2} = CdFr\{G_\delta * P_{L_1}; d_1, f_{d_2}, \rho\}. \quad (8)$$

Second, calculate $F_{L_1}^{1,2'}$ by applying a gray-scale silhouette mask, $G_\delta * M_2$ to $F_{L_1}^{1,2}$ as

$$F_{L_1}^{1,2'}(x_2, y_2) = (G_\delta * M_2) \times F_{L_1}^{1,2}. \quad (9)$$

Finally, the CGH is obtained by linearly combining $F_{L_1}^{1,2'}$ and F_{L_2} (the observed image of the object at L_2). The application of the GAF filter to F_{L_2} is adopted and the CGH pattern is expressed as

$$CGH = ICdFr\{F_{L_1}^{1,2'} + G_\delta * F_{L_2}; d_1, f_{d_2}\}. \quad (10)$$

The reconstruction image of the CGH at the retina plane is

$$Q_{1(2)} = CdFr\{CGH; d_{1(2)}, f_{1(2)}, \rho\}. \quad (11)$$

The CGHs reconstructed using the proposed method are presented in Fig. 4(b). The distance of the triangle and circle objects is set to 0.2 m under the same conditions as in Fig. 3. The effect of the GAF is a reduction of the inter-layer edge-diffraction artifact. This simulation result is comparable to the ground truth image with a pupil radius of 1 mm [Fig. 4(c)], and we have obtained a CGH with a more natural accommodation effect superior to the conventional depth-map CGH [Fig. (3)]. The GAF suppresses the high-frequency components generating unnatural edge-diffraction artifacts. It was found that the GAF needs to be applied to the layered object image light fields as well as the binary silhouette mask as described above. To further compare the proposed method and the conventional method, we measure the deviation of the CGH reconstructions. The deviation is evaluated at the region-of-interest (ROI) designating the area around the layer contour boundaries as represented by

$$\Delta = \sum_{n=1,2} \int_{ROI} \sqrt{(|G_{L_n}| - |Q_n|)^2 / |G_{L_n}|^2}. \quad (12)$$

The deviation of the traditional method [Fig. 3(a)] from the ground truth is calculated to be 0.401 and the deviation index of the silhouette mask method [Fig. 3(b)] to be 0.4008, while

for the proposed GAF method [Fig. 4(b)], the calculated value is reduced to 0.1702. This means that the GAF transforms the CGH image closer to the ground truth.

In Fig. 5, the advantages of the proposed method are prominently presented through the reconstructed image of depth-map CGH with full-color images in the art dataset [10]. In the traditional depth-map CGH [Fig. 5(a)], the inter-layer diffraction artifact is observed at the boundary between the adjacent objects. Although the binary silhouette mask reduces the edge diffraction artifact slightly [Fig. 5(b)], unnatural noise remains. The proposed GAF process generates the CGH observation results shown in Fig. 5(c). This numerical demonstration shows the GAF method generating a relatively natural and more realistic defocus effect by damping the inter-layer diffraction artifact around the contour boundary of depth-map objects.

Funding. Alchemist Project (MOTIE&KEIT Grant: 1415179744, 20019169); Samsung Research Funding & Incubation Center of Samsung Electronics (SRFC-TB1903-05).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. A. Symeonidou, D. Blinder, and P. Schelkens, *Opt. Express* **26**, 10282 (2018).
2. O. Kunieda and K. Matsushima, *Appl. Opt.* **58**, G104 (2019).
3. S. Park, J. Lee, S. Lim, M. Kim, S. Ahn, S. Hwang, S. Jeon, J. Jeong, J. Hahn, and H. Kim, *Opt. Express* **29**, 26793 (2021).
4. J. Roh, K. Kim, E. Moon, S. Kim, B. Yang, J. Hahn, and H. Kim, *Opt. Express* **25**, 14774 (2017).
5. S. Choi, J. Kim, Y. Peng, and G. Wetzstein, *Optica* **8**, 143 (2021).
6. L. Shi, B. Li, C. Kim, P. Kellnhofer, and W. Matusik, *Nature* **591**, 234 (2021).
7. S. Choi, M. Gopakumar, Y. Peng, J. Kim, and G. Wetzstein, *ACM Trans. Graph.* **40**, 240 (2021).
8. C. Chen, B. Lee, N. N. Li, M. Chae, D. Wang, Q. H. Wang, and B. Lee, *Opt. Express* **29**, 15089 (2021).
9. H. Zhang, L. Cao, and G. Jin, *Appl. Opt.* **56**, F138 (2017).
10. D. Scharstein and C. Pal, in *2007 IEEE Conference on Computer Vision and Pattern Recognition*, (IEEE, 2007), pp. 1–8.