Content-adaptive targeting scheme for holographic displays

Ziyang Chen[a], Dongyeon Kim[b], Rafal Mantiuk[b], and Kaan Akşit[a]

[a]University College London, London, United Kingdom
[b]University of Cambridge, Cambridge, United Kingdom

ABSTRACT

Computer-Generated Holography (CGH) algorithms identify Three-Dimensional (3D) holograms by utilizing focal stacks as their target images. However, periodically slicing depth planes to generate a focal stack can lead to artifacts and reduced contrast in visually salient parts on a 3D scene. We introduce a content-adaptive targeting scheme that prioritizes contrast and optimizes slicing of the depth planes. This way, our new targeting scheme reduces the slicing artifacts in reconstructed 3D holographic images.

1 INTRODUCTION

A holographic display modulates a complex-valued wave field using a Spatial Light Modulator (SLM) to reconstruct desired 3D images. Focal stacks reconstructed on multiple planes are conventionally employed as ground truth assets for 3D CGH acquisition. Several studies have investigated focal stack representation methods to determine the most suitable approach as ground truth for CGH optimization. Some works[1,2] on 3D holograms investigated paths to render realistic blur with a new targeting scheme and loss function to reduce the ringing effect. The results showed improved image quality but exhibited a reduced effect of depth-dependent blur, likely due to incomplete utilization of the display system's pupil. Additionally, Lee et al.[3] proposed the focal stack generation with incoherent propagation kernels, defined as the auto-correlation of the coherent propagation kernel. This was achieved using emerging kHz-range fast SLMs with low-bit depth, enabling time-multiplexing to suppress speckle noise and allowing the optimized CGH to operate in the incoherent regime. However, CGH algorithms for fast SLMs are computationally expensive as they have to deal with more frames. These works provide promising results focusing more on the realization of the defocus blur but overlooked the slicing artifacts that arise from the axial propagation of holographic displays. In detail, the axial intensities in holography correlate with each other. Consequently, slicing artifacts are inherently present in RGB-D-based focal stacks, and these artifacts are more noticeable in the contents with large depth discontinuities.

In this work, we propose a new focal stack targeting scheme for CGH acquisition that takes into account the axial holographic projection and detours the depth-slicing problem by leveraging the human visual perception factors. The simulated results maintain the image qualities at the salient areas while reducing the slicing artifact.

2 METHODOLOGY AND EVALUATION

Our method regularizes contrast and slicing boundaries in the process of generating a focal stack by solving the following minimization problem:

$$\hat{d} \leftarrow \underset{d}{\operatorname{argmin}} \left(\underbrace{-\lambda_1 \log(\mathrm{C}(\mathsf{F}))}_{\text{Contrast Loss}} + \underbrace{\lambda_2 \sum_{i=1}^{N} |\mathcal{S}(\mathsf{F}_i)| * |\mathcal{S}(\mathsf{F}_{i+1})| * \mathfrak{s}(I)}_{\text{Slicing Boundaries Loss}} \right), \tag{1}$$

where, $d \in \mathbb{R}^{1\times N}$ denotes the depth plane distribution of the focal stack $\mathsf{F} \in \mathbb{R}^{N\times H\times W}$ with a height of H and a width of W. $C(\mathsf{F})$ denotes our formulated contrast calculation, $\mathcal{S}$ denotes the Sobel edge detector, and $\mathfrak{s}$ is the saliency predictor for the RGB image I. To constrain the d in the correct order and range, we add another two terms: $\mathcal{L}_{order} = \lambda_3 \sum_{i=1}^{N} \mathrm{ReLU}(d_i - 1) + \mathrm{ReLU}(-d_i)$ and $\mathcal{L}_{range} = \lambda_4 \sum_{i=1}^{N-1} \mathrm{ReLU}(d_i - d_{i+1})$. Our focal stack generation process follows the incoherent propagation method in the work of Lee et al.[3] which each of the focal planes is the summation of all the planes considering their occlusions mask. We replace the masking process with a differentiable approximation of the actual discretization process:

$$M_{(\mathcal{D},d)} = \mathrm{softmax}\left(-\frac{|\mathcal{D} - \frac{d_i + d_{i+1}}{2}|}{\tau} \right), \quad i \in \{1,...,N-1\}, \tag{2}$$

where, $\mathcal{D}$ is the depth map of the RGB image, τ denotes the temperature to make sure we get the sharp boundaries between depth map masks. For the perceived contrast estimation, a focal stack is decomposed into multiple levels of Laplacian and Gaussian pyramids. The contrast level defined by each spatial frequency band is defined by the ratio between the Laplacian pyramid and the bilinearly upsampled Gaussian in the next pyramid level.[4] Then, the contrast

Further author information: (Send correspondence to Ziyang Chen)
Ziyang Chen: E-mail: ziyang.chen.22@ucl.ac.uk

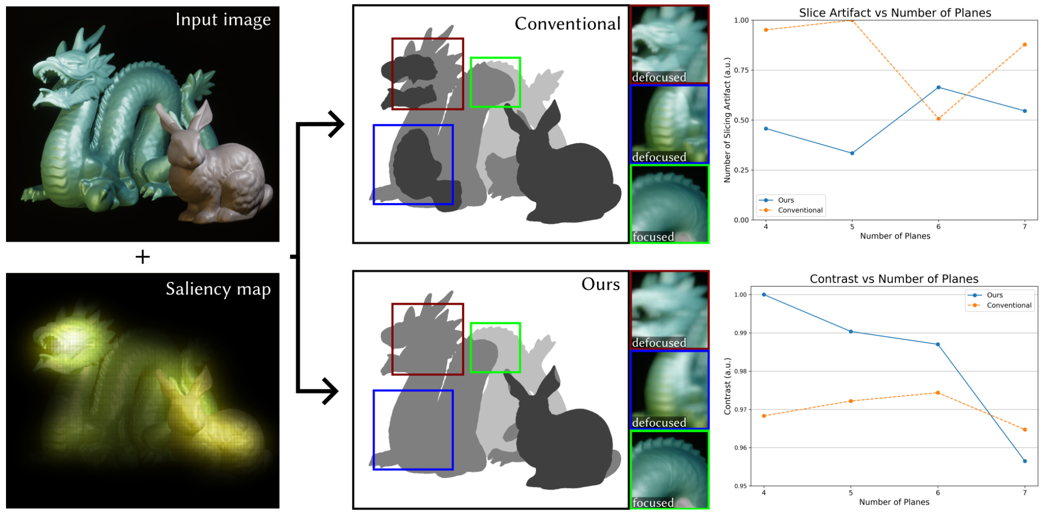

Figure 1. (left)The input image and the saliency map generate the focal stack images. (middle) The quantized depthmaps of the the input image using the conventional method (middle-top), and our method (middle-bottom). (right) The plots show the contrast and the number of slicing artifacts relative to the number of planes. Note that these arbitrary values are normalized by diving the maximum value. Dragon, Bunny: credit to Stanford Computer Graphics Laboratory.

sensitivity is multiplied per spatial frequency band to map the physical contrast to the perceived contrast domain. Thus, the overall perceived contrast of a focal stack can be calculated by the function $C:\mathbb{R}^{N\times H\times W}\rightarrow\mathbb{R}^{H\times W}$:

$$C(\mathsf{F})=\frac{1}{L}\sum_{I_f\in\mathsf{F}}^{N}\sum_{l=1}^{L}\frac{\mathfrak{L}_l(I_f)}{\mathcal{U}(\mathcal{G}_{l+1}(I_f))}*\mathbf{s}_l, \tag{3}$$

where, the L is the number of pyramid levels, (8 in our experiment), and $\mathcal{U}$ is the upsampling operator. $\mathfrak{L}_l$ is the Laplacian pyramid, $\mathcal{G}_l$ is the Gaussian pyramid, and $\mathbf{s}_l$ is the contrast sensitivity corresponding to the l-th spatial frequency band. We employed Barten's contrast sensitivity model[5] for the contrast sensitivity. We evaluate the generated holograms by propagating them to a fixed range of distances and compare the reconstruction contrast and image quality metrics with the densely rendered focal stack from a modern graphics pipeline. For each chromatic channel, we simulated speckle-free, high-quality reconstructions using the time-multiplexed CGH pipeline[6] with 12 random-phase frames, then combined the channels for the final simulation. The hyperparameters are set as $\lambda_1=1e^{-2}$ and $\lambda_2=1e^{-7}$ for the contrast loss and slicing artifact loss. As shown in the middle section of Fig.1, the optimizer avoids directly slicing through the salient areas (e.g., the heads of the dragon and bunny) and reduces the number of artifacts by slightly adjusting the slicing position (highlighted by the green box). The normalized plots in the right section of Fig.1 illustrate that the proposed optimization avoids slicing artifacts and improves contrast, given a different number of planes being used.

Our method reduces the slicing artifact up to 48%, see Fig.1. Additionally, our method shows up to 2% improvements in contrast for different numbers of focal planes. The comparison results show that our reconstruction scheme achieves similar performance in image quality metrics, including PSNR, SSIM, LPIPS, and DISTS. In the future, this optimization can be accelerated using a learned approach and integrated into 3D CGH pipelines. Additionally, new targeting scheme on focal surfaces can be investigated by taking advantage of the emerging focal surface beam propagation model.[7]

References

[1] Kavaklı, K., Itoh, Y., Urey, H., and Akşit, K., "Realistic defocus blur for multiplane computer-generated holography," in [2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)], 418–426, IEEE (2023).

[2] Kavaklı, K., Shi, L., Urey, H., Matusik, W., and Akşit, K., "Multi-color holograms improve brightness in holographic displays," in [SIGGRAPH Asia 2023 Conference Papers], 1–11 (2023).

[3] Lee, B., Kim, D., Lee, S., Chen, C., and Lee, B., "High-contrast, speckle-free, true 3d holography via binary cgh optimization," Scientific reports **12**(1), 2811 (2022).

[4] Peli, E., "Contrast in complex images," JOSA A **7**(10), 2032–2040 (1990).

[5] Barten, P. G., [Contrast sensitivity of the human eye and its effects on image quality], SPIE press (1999).

[6] Kim, D., Nam, S.-W., Choi, S., Seo, J.-M., Wetzstein, G., and Jeong, Y., "Holographic parallax improves 3d perceptual realism," ACM Transactions on Graphics (TOG) **43**(4), 1–13 (2024).

[7] Zheng, C., Zhan, Y., Shi, L., Cakmakci, O., and Akşit, K., "Focal surface holographic light transport using learned spatially adaptive convolutions," in [Proceedings of SIGGRAPH Asia 2024 Technical Communications (SA Technical Communications '24)], SA '24 (December 2024).