Role of perceptually guided image quality metrics in computer-generated holography

Doğa Yılmaz[a], 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 relies on pixel-wise evaluation metrics for calculating holograms. Our work addresses a gap in the literature by evaluating various combinations of conventional pixel-wise metrics such as L2-norm and emerging perceptual metrics including FovVideoVDP[1] and ColorVideoVDP.[2] We report the effectiveness of each metric on the quality of the resulting holograms, identifying a set of recipes describing the optimal combination of visual metrics for high-quality hologram calculation.

Keywords: Computer-Generated Holography, Holographic displays

1. INTRODUCTION

Renewed interest in holographic displays in the recent years has triggered a boost in algorithmic research for CGH.[3] Primarly, these CGH algorithms aim to identify holograms that can reconstruct an accurate representation of a targeted 3D scene. These CGH algorithms could broadly be classified as optimization based,[4,5] learning based[6] and hybrid[7,8] methods. A typical hologram calculation process could be expressed as the following,

$$\hat{u}_t, \hat{l}_{(p,t)} \leftarrow \underset{u_t, l_{(p,t)}}{\text{argmin}}\ \mathcal{L}\left(\sum_{z=Z_0}^{Z}\sum_{p=1}^{P}\sum_{t=1}^{T}|l_{(p,t)}e^{i\frac{\lambda_p}{\lambda_{p_{\text{anchor}}}}u_t} \circledast h_p(\lambda_p, z, dx)|^2, sI_{(p,z)}\right), \tag{1}$$

where $Z \in \mathbb{R}_{\geq 0}$ and $Z_0 \in \mathbb{R}_{\geq 0}$ denote the largest and shortest light propagation distance, respectively, $P \in \mathbb{Z}$ denotes the total number of color primaries (i.e. $RGB \rightarrow 3$), $T \in \mathbb{Z}$ denotes the total number of subframes reproducing a full-color image (typically three), $l_{(t,p)} \in \mathbb{R}_{\geq 0}$ represents the light source power for the p-th primary at the t-th subframe, $\lambda_p \in \mathbb{R}_{700>\mathbb{R}>400}$ and $\lambda_{p_{\text{anchor}}} \in \mathbb{R}_{700>\mathbb{R}>400}$ denotes the wavelength of the active primary and the anchor primary, respectively, $dx \in \mathbb{R}_{\geq 0}$ represents the physical pixel pitch size of a hologram, $\hat{u}_t$, $I_{(p,z)} \in \mathbb{R}_{\geq 0}$ is a target image intensity, $s \in \mathbb{R}_{\geq 0}$ represents a parameter to scale the brightness of a target image intensity, $h_p \in \mathbb{C}$ denotes the wavelength, λ_p, and location, z, dependent kernel for propagating a field, and $\mathcal{L}$ represents the loss function that evaluates the intended target image for a given plane and a reconstructed image by a hologram. Note that $l \in \mathbb{R}_{t \times p}$ is always an identity matrix in the traditional holograms.

Our work explores a combination of pixel-wise metrics such as L2-norm and emerging perceptual metrics including FovVideoVDP[1] and ColorVideoVDP.[2] Specifically, in addition to conventional losses used in 3D hologram optimization[4] and learning,[6] we include these perceptual metrics to our $\mathcal{L}$ that is introduced in our hologram synthesis equation in Eq. 1. We weight the outcomes from traditional pixel-wise metrics used in multiplane CGH[4] and perceptual metrics [1,2] to regularize our hologram optimization process. In our experimentations, we identified a recipe to weight the loss components in $\mathcal{L}$ to improve our hologram optimization routines. Our practical findings suggest multipling the outcome from FovVideoVDP[1] and ColorVideoVDP[2] with -1×10^{-4} in addition to weighting existing pixel-wise metrics with one would lead to a slightly improved image quality. We provide a sample result in Fig. 1, where we also report the metric results. Note that in Fig. 1, P denotes Peak-to-Signal ratio in Decibels, S denotes Structured-Similarity Index Metric, L denotes LPIPS,[9] FV denotes FovVideoVDP[1] and CV denotes ColorVideoVDP.[2]

Further author information: (Send correspondence to Doğa Yılmaz and Kaan Akşit)
Doğa Yılmaz: doga.yilmaz.24@ucl.ac.uk
Kaan Akşit: kaanaksit@kaanaksit.com

Advances in Display Technologies XV, edited by Jiun-Haw Lee, Qiong-Hua Wang, Proc. of SPIE
Vol. 13388, 133880C · © 2025 SPIE · 0277-786X · doi: 10.1117/12.3040649

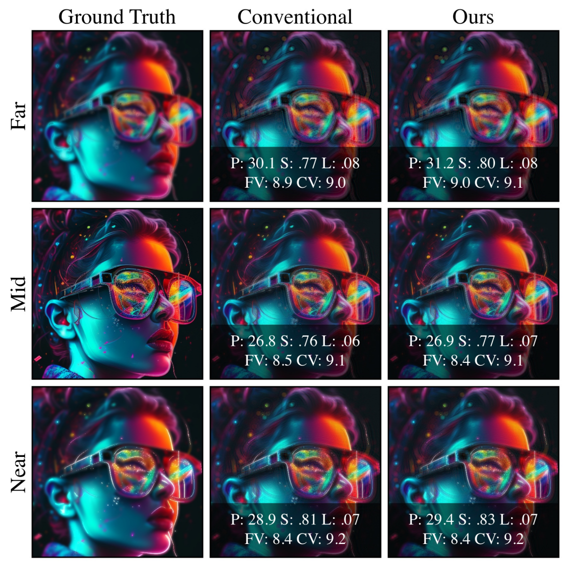

Figure 1. Including perceptual metrics in CGH optimization improves image quality slightly at the cost of a slight increase in computational complexity.

Our findings suggest that image quality of reconstructed 3D scenes could be improved when using selected perceptual metrics ,[1,2] see Fig. 1. However, we observed that the selected perceptual metrics[1,2] and emerging perceptual metrics[9] from the literature tend to require more computational resources. Thus, these perceptual metrics increase the computational complexity of hologram synthesis described in Eq. 1. To provide more insights, our Graphical Processing Unit (GPU) optimized CGH optimization code (located in *GitHub:complight/multi_color*, *perceptual_multiplane_loss* branch) basing on our differentiable toolkit[10] and NVIDIA RTX 3090 GPU, runs %15 slower (1:10 to 1:21) and requires %2 more GPU memory (6.02 GB to 6.11 GB).

We argue that CGH research community may find it beneficial to investigate on dedicated perceptually-guided visual quality metrics geared for 3D hologram synthesis problem described in Eq. 1 as there are novel artifacts in images displayed on a holographic display (e.g., ringing artifacts[4]). In addition, from our experiments, we speculate that there could be image quality related benefits in training models for hologram synthesis with these perceptually-guided visual quality metrics in the future.

REFERENCES

[1] Mantiuk, R. K., Denes, G., Chapiro, A., Kaplanyan, A., Rufo, G., Bachy, R., Lian, T., and Patney, A., "Fovvideovdp: A visible difference predictor for wide field-of-view video," ACM Transactions on Graphics (TOG) **40**(4), 1–19 (2021).

[2] Mantiuk, R. K., Hanji, P., Ashraf, M., Asano, Y., and Chapiro, A., "Colorvideovdp: A visual difference predictor for image, video and display distortions," ACM Trans. Graph. **43** (July 2024).

[3] Koulieris, G. A., Akşit, K., Stengel, M., Mantiuk, R. K., Mania, K., and Richardt, C., "Near-eye display and tracking technologies for virtual and augmented reality," in [Computer Graphics Forum], **38**(2), 493–519, Wiley Online Library (2019).

[4] Kavaklı, K., Itoh, Y., Urey, H., and Akşit, K., "Realistic defocus blur for multiplane computer-generated holography," IEEE VR 2023 (Mar. 2023).

[5] 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], SA '23 (Dec. 2023).

[6] Akşit, K. and Itoh, Y., "Holobeam: Paper-thin near-eye displays," in [2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)], 581–591 (Mar. 2023).

[7] Zhan, Y., Kavaklı, K., Urey, H., Sun, Q., and Akşit, K., "Autocolor: Learned light power control for multi-color holograms," SPIE (Jan. 2024).

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

[9] Zhang, R., Isola, P., Efros, A. A., Shechtman, E., and Wang, O., "The unreasonable effectiveness of deep features as a perceptual metric," in [Proceedings of the IEEE conference on computer vision and pattern recognition], 586–595 (2018).

[10] Akşit, K., Beyazian, J., Chakravarthula, P., Chen, Z., Doğan, M. D., Güzel, A. H., Itoh, Y., Kam, H., Karadeniz, A. S., Kavaklı, K., Shi, L., Spjut, J., Walton, D. R., Wu, J., Yılmaz, D., Yujie, W., Zhu, R., Xie, W., and Zhan, Y., "Odak, version 0.2.6," (Oct. 2024).