

Michelson Holography: Dual-SLM Holography with Camera-in-the-loop Optimization

SUYEON CHOI^{1,2}, JONGHYUN KIM^{1,2*}, YIFAN PENG², AND GORDON WETZSTEIN²

¹NVIDIA, 2788 San Tomas Expressway, Santa Clara, CA 95051, USA

²Electrical Engineering Department, Stanford University, Stanford, California 94305, USA

* Corresponding author: jonghyunk@nvidia.com

Compiled September 23, 2020

We introduce Michelson Holography (MH), a holographic display technology that optimizes image quality for emerging holographic near-eye displays. Using two spatial light modulators, MH is capable of leveraging destructive interference to optically cancel out undiffracted light corrupting the observed image. We calibrate this system using emerging camera-in-the-loop holography techniques and demonstrate state-of-the-art holographic 2D image quality. © 2020

<http://dx.doi.org/10.1364/ao.XX.XXXXXX>

Near-eye displays in virtual and augmented reality systems should deliver high-quality imagery while supporting focus cues, a large field of view, and a reasonably sized eye box within a compact device form factor. Holographic near-eye displays promise to deliver these capabilities and have made remarkable progress over the last few years [1–11]. Unlike conventional displays, however, holographic displays using phase-only spatial light modulators (SLMs) can only indirectly show an image by shaping a wave field such that the target image is created through interference. This has been a challenging problem for decades and the image quality achieved by computer-generated holography (CGH) is often significantly lower than that of conventional technology, such as liquid crystal displays.

A fundamental challenge of phase-only SLMs used for holography is their low diffraction efficiency. Caused by their limited pixel fill factors, backplane architectures, and other factors, as much 20% of incident light may not be diffracted [12]. This creates the zeroth diffraction order, which typically interferes with the user-controlled diffraction orders and significantly degrades observed image quality.

In optics, on-axis and off-axis filtering schemes are the two most common techniques to minimize the zeroth diffraction order. On-axis filtering physically blocks the undiffracted beam at the Fourier plane, which inevitably also blocks some amount of the diffracted light that contributes to the low-frequency content. In addition, this blocking process is much more challenging when multiplexing three colors [13]. Off-axis methods suffer from reduced field of view (using half of the first diffraction order) or lowered efficiency (using higher diffraction orders)—both of these factors are crucial for near-eye displays [14]. In addition, approaches that compensate for the zeroth-order beam

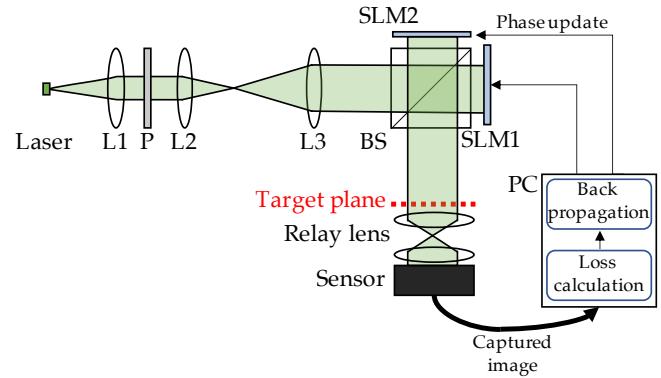


Fig. 1. Principle of Michelson Holography. L1: collimating lens, L2, L3: beam expanders, P: polarizer, BS: beam splitter. The sensor captures the intensity of the field at the target plane. The error between captured and target images is back-propagated, with a stochastic gradient descent optimization algorithm, to update the phase patterns of both SLMs.

have been described that model either a correction beam [15] or the pixelated structure of the SLM [16, 17]. Most recently, camera-in-the-loop (CITL) holography techniques have been described that can partially compensate the undiffracted light of an SLM using its diffracted component without having to explicitly model all of these terms [18, 19]. However, none of these software-only approaches accounts for the physical limitations of the actual SLM, limiting the degree to which destructive interference can be utilized to cancel out the zeroth order.

Inspired by the design of Michelson interferometers, we propose Michelson Holography (MH), a holographic display architecture that uses two phase-only SLMs and a newly developed CITL optimization procedure. The core idea of MH is to destructively interfere the diffracted light of one SLM with the undiffracted light of the other. As such, the undiffracted light can contribute to forming the target image rather than creating speckle and other artifacts. Our custom CITL optimization procedure captures the coherent superposition of all diffracted and undiffracted light of the display and backpropagates the error w.r.t. a target image into both SLM patterns simultaneously. This procedure does not require us to explicitly model the SLM pixel

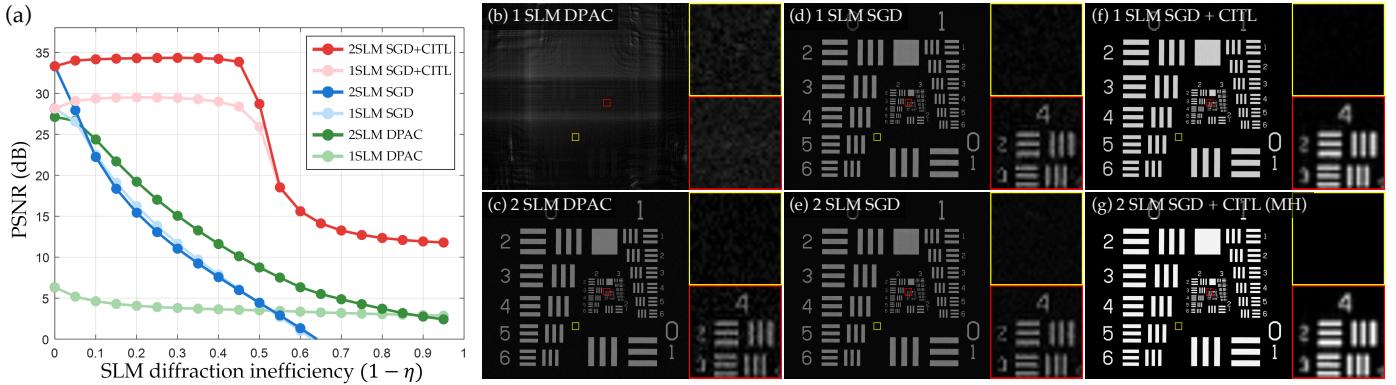


Fig. 2. Comparison of image quality achieved by several CGH techniques for varying amounts of SLM diffraction efficiency in simulation. (a) Quantitative evaluation of resolution chart. We report peak signal-to-noise ratio (PSNR). Double phase–amplitude coding (DPAC) with two SLMs works very well for high diffraction efficiencies, but the quality of DPAC quickly drops as the amount of undiffracted light increases. Similar trends are observed for the stochastic gradient descent (SGD) method. Using SGD with our CITL calibration method achieves high PSNR values for diffraction efficiencies as low as 50%. (b–g) The reconstructed images, here shown for $1 - \eta = 0.2$, show similar trends, with Michelson Holography achieving the highest image quality and contrast.

structure or the undiffracted light. We simply need a camera that captures intermediate images of this iterative CGH algorithm, which automatically optimizes the resulting phase patterns for both SLMs. Although holographic displays with multiple SLMs have been investigated for viewing angle enhancement [20], improving resolution [21], or complex modulation [5, 22–24], our approach is the first to leverage CITL optimization to optimize image quality of two phase-only SLMs by mitigating the effect of undiffracted light in a fully automatic manner.

Figure 1 shows the principle of MH. Here, a source field u_{src} generated by the laser is incident on both SLMs. Each SLM delays the phase of the incident light by $\phi_{1/2}$, respectively, with diffraction efficiency η . The fields reflected off of the SLMs are $u_{1/2} = [\eta \exp^{i\phi_{1/2}(x,y)} + (1 - \eta)]u_{src}$. The functions $g_{1/2}$ describe the unknown physical wave propagation from the SLMs to the target plane, including aberrations, whereas $\hat{g}_{1/2}$ describe their idealized models used in simulation. We use the angular spectrum method to model free-space wave propagation as

$$\begin{aligned} \hat{g}_{1/2}(\phi_{1/2}) &= \iint \mathcal{F}(u_{1/2}) \mathcal{H}(f_x, f_y) e^{i2\pi(f_x x + f_y y)} df_x df_y, \\ \mathcal{H}(f_x, f_y) &= \begin{cases} e^{i\frac{2\pi}{\lambda} \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2} z}, & \text{if } \sqrt{f_x^2 + f_y^2} < \frac{1}{\lambda}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (1)$$

where $\mathcal{F}(\cdot)$ denotes Fourier transform, f_x, f_y are spatial frequencies, λ is the wavelength, z is the propagation distance between SLM and target plane, and x, y are the 2D coordinates on the SLM plane. In practice, the pixelated SLMs modulate the discrete phase patterns $\phi_{1/2} \in \mathbb{R}^{N_x \times N_y}$, where $N_x \times N_y$ is the SLM resolution, and $g_{1/2}, \hat{g}_{1/2} : \mathbb{C}^{N_x \times N_y} \rightarrow \mathbb{C}^{N_x \times N_y}$.

Therefore, optimizing the two SLM phase patterns in MH is done by solving the problem

$$\underset{\phi_1, \phi_2}{\text{minimize}} \mathcal{L}(s \cdot |\hat{g}_1(\phi_1) + \hat{g}_2(\phi_2)|, a_{target}), \quad (2)$$

where \mathcal{L} is a loss function, such as mean squared error, a_{target} is the amplitude of the target image, and s is a scale factor.

Most conventional CGH algorithms cannot easily model the proposed dual-SLM setup. Only the double phase–amplitude coding (DPAC) method [25] is specifically designed for this

case, because it represents a complex-valued field at the SLM plane $u = ae^{i\phi}$ as the sum of two phase-only fields with phases $\phi - \cos^{-1}a$ and $\phi + \cos^{-1}a$, respectively. Compared to the interleaved single-SLM DPAC variant [3], dual-SLM DPAC is expected to achieve more accurate results, which is why we use it as a baseline for our algorithm. Note that DPAC does not account for undiffracted light, optical aberrations, or SLM misalignment.

We build on the recently proposed CITL calibration approach [18] and adopt it to our dual-SLM setup. This approach builds on a stochastic gradient descent (SGD) solution for Eq. (2). Starting from some initialization or a previous estimate of the two phase patterns $\phi_{1/2}^{(k-1)}$, we display these patterns on the SLM, capture an image on the target plane with a camera, use that to evaluate the loss function \mathcal{L} , and then backpropagate the error using the gradients $\frac{\partial \mathcal{L}}{\partial \phi_{1/2}}$ to find the next estimates $\phi_{1/2}^{(k)}$. Because we use a recorded image, this includes both diffracted and undiffracted contributions of both SLMs. Specifically, the update rules of our SGD-based solver in iteration k are

$$\begin{aligned} \phi_{1/2}^{(k)} &\approx \phi_{1/2}^{(k-1)} - \alpha \left(\frac{\partial \mathcal{L}}{\partial g_{1/2}} \cdot \frac{\partial \hat{g}_{1/2}}{\partial \phi_{1/2}} \right)^T \times \\ &\quad \mathcal{L} \left(s \cdot |g_1(\phi_1^{(k-1)}) + g_2(\phi_2^{(k-1)})|, a_{target} \right). \end{aligned} \quad (3)$$

We evaluate several CGH algorithms, including single- and dual-SLM variants of DPAC and SGD with and without CITL. Figure 2 shows simulated results of a resolution chart at the target plane for varying SLM diffraction inefficiencies ($1 - \eta$). We simulate 1080p phase-only SLMs with $6.4 \mu\text{m}$ pixel pitch and $z = 10 \text{ cm}$ propagation distance between SLM and target plane. In most cases, the peak signal-to-noise ratio (PSNR) rapidly decreases as the undiffracted light increases; only our SGD method with CITL optimization achieves a constant and high image quality up to $\approx 50\%$ inefficiency for both single and dual-SLM setups. The proposed dual-SLM setup provides the best image quality of all methods, especially in the presence of undiffracted light. Figs. 2(b)–(g) show qualitative comparisons of a reconstructed resolution target for $(1 - \eta) = 0.2$. These simulations demonstrate that MH achieves the best image contrast and quality.

Our holographic display prototype uses two phase-only SLMs (HOLOEYE LETO and PLUTO) with $6.4\ \mu\text{m}$ and $8.0\ \mu\text{m}$ pixel pitch, respectively. The laser is a FISBA RGBbeam fiber-coupled module with three optically aligned laser diodes with wavelengths 638 nm, 520 nm, and 450 nm. All results and the images for the CITL algorithm were captured with a FLIR Grasshopper3 2.3 MP color USB3 vision sensor through two Nikon AF-S Nikkor 35mm f/1.8 relay lenses. The CITL optimization was run on each color channel separately and full-color results were combined in post-processing. Additional hardware and software details are discussed in Supplement 1.

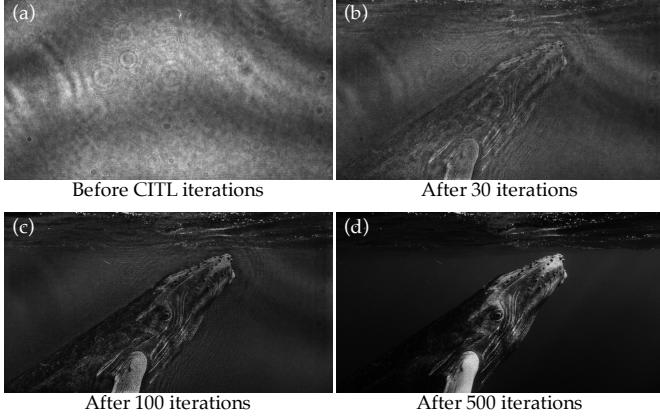


Fig. 3. Convergence of experimental results. (a) The undiffracted light from two SLMs naturally creates a fringe pattern. Our CITL optimization iteratively improves the observed image, as shown for (b) 30, (c) 100, and (d) 500 iterations (Visualization 1).

Similar to a Michelson interferometer, the two SLMs of MH create fringe patterns on the target plane when no pattern is displayed (Fig. 3(a)). This is what makes calibrating a dual-SLM setup so challenging, but it is also this interference that gives MH the ability to produce higher-contrast images than single-SLM setups. Using the proposed CITL optimization, the target image is found after about 500 iterations (Figs. 3(b)-(d), Visualization 1).

Figure 4(a) and Table 1 show qualitative and quantitative experimental evaluations of contrast and image quality for several different optical and algorithmic holographic display variants. Here, Weber contrast is defined as $(I_{\max} - I_{\min}) / I_{\min}$ and Michelson contrast is $(I_{\max} - I_{\min}) / (I_{\max} + I_{\min})$, where I is the intensity of the display. Similar to our simulations, the proposed dual-SLM setup combined with CITL optimization achieves the best results in all cases. Especially for the blue channel, where

Table 1. Experimental results evaluating contrast listing Weber / Michelson contrast obtained from captured sinusoidal patterns (cf. Fig. 4(a)). MH shows the highest contrast for all wavelengths.

Wavelength	1 SLM		2 SLM
	SGD	SGD + CITL	SGD + CITL
R (638 nm)	27.3 / 0.93	33.4 / 0.94	38.2 / 0.95
G (520 nm)	15.5 / 0.89	28.4 / 0.93	37.0 / 0.95
B (450 nm)	8.75 / 0.81	9.51 / 0.83	20.9 / 0.91

the SLM diffraction efficiency is lower than at other wavelengths, MH shows a big improvement in contrast. The proposed method shows the uniform efficiency in the spectral domain which can correct the color shift as well.

Figures 4(c)-(h) show full-color experimental holographic images. Whereas non-CITL approaches, including DPAC and SGD, suffer from speckle and other artifacts, CITL optimization can significantly reduce these. However, using a camera in the loop allows us to automatically account for optical aberrations, phase nonlinearities, and other imperfections not directly modeled by the image formation of Eq. (1). Comparing the CITL-calibrated single- and dual-SLM setups demonstrates again that the proposed MH approach optimizes image quality and contrast, while reducing remaining artifacts (Visualizations 2 and 3).

The optical paths of the two SLMs in our setup are only roughly aligned. We primarily rely on a homography calibration of their respective contributions to the target plane. Here, the camera captures test patterns displayed on each of the SLMs and automatically derives a homography mapping between them (see Supplement for details). To study how robust MH is for slight misalignments of the SLMs, we ran several experiments in simulation for $\eta = 0.8$. As seen in Fig. 5, misalignments as small as one pixel in the lateral or axial dimension result in strong image degradation for DPAC. SGD is more robust along the axial but not along the lateral dimension. SGD with CITL shows a similar trend for axial and lateral robustness, although with a significantly higher absolute quality.

In summary, we demonstrate that Michelson Holography, i.e., a dual-SLM holographic display with CITL calibration, provides superior image quality than existing CGH approaches both in simulation and experiment. Yet, our method is not without limitations. Currently, the CITL algorithm requires an iterative optimization process with a camera for each target image. This procedure takes around 500 seconds for a 1080p target image with our PyTorch implementation executed on an NVIDIA RTX 2080Ti. Although we could likely achieve real-time framerates using a CITL-calibrated neural network [18], we have not attempted to do that in this letter. Our hardware prototype is currently limited by using SLMs with slightly different pixel pitches. This mismatch might cause unwanted errors and adversely affect the image quality of our results. However, it also implies that MH works robustly with different SLMs. Finally, we only demonstrate 2D holographic images in this letter and leave a detailed exploration of per-pixel 3D MH for future work.

Holographic displays are maturing into a practical technology. Artificial intelligence has enabled real-time framerates for 2D CGH [18, 26–28] and fully automatic display calibration for optimized image quality [18, 19]. Here, we demonstrate how these neural holography techniques also enable novel holographic display hardware architectures to further improve holographic image quality, which slowly approaches that of conventional displays based on liquid crystal or organic light-emitting diode technology.

ACKNOWLEDGMENTS

We thank Ward Lopes, Morgan McGuire, and David Luebke for helpful discussions and advice. This project was partially supported by Ford, NSF (awards 1553333 and 1839974), a Sloan Fellowship, an Okawa Research Grant, and a PECASE by the ARO.

See Supplement 1 for supporting content.

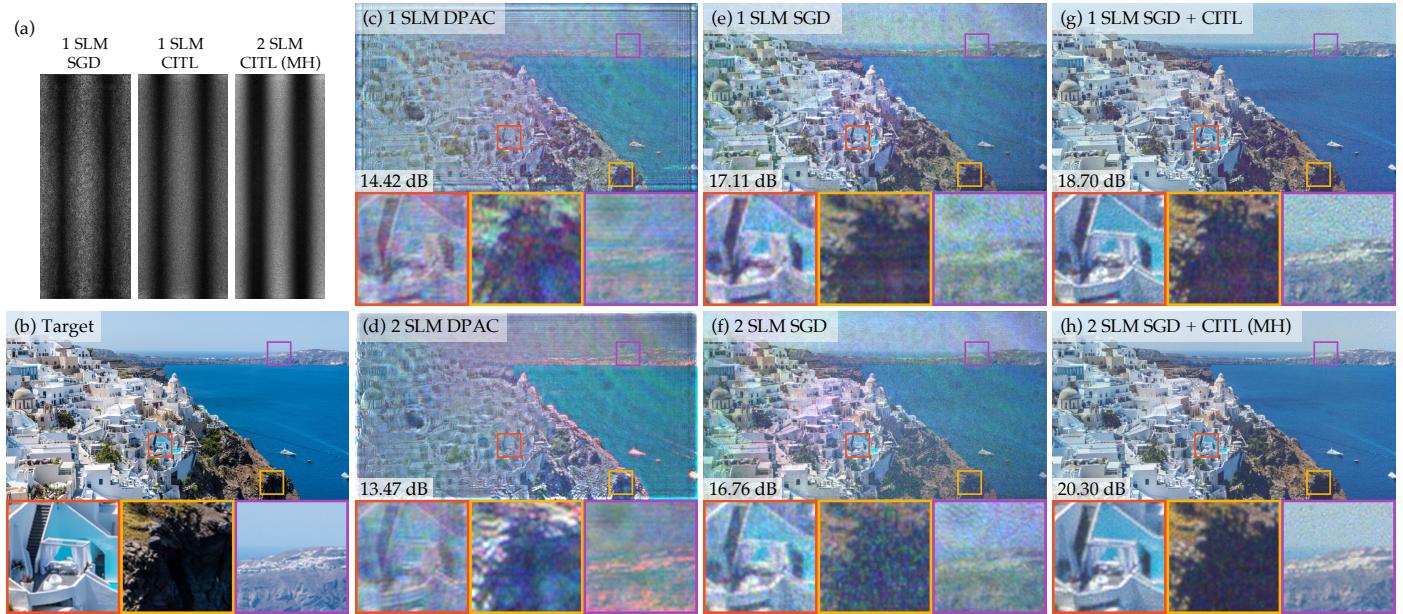


Fig. 4. Experimentally captured results. (a) Qualitative results of the contrast experiments shown in Table 1. SGD with CITL calibration achieves a higher contrast than other CGH approaches. (c)-(g) Experimental holographic images as created by several different CGH methods. Numbers indicate the PSNR w.r.t. the target image (b). Michelson Holography shows about 1.5 dB quantitative improvement over the next-best method, i.e., a single-SLM CITL variant of MH [18], and significant improvements in image quality, contrast, and speckle reduction compared with all other CGH methods (Visualizations 2 and 3).

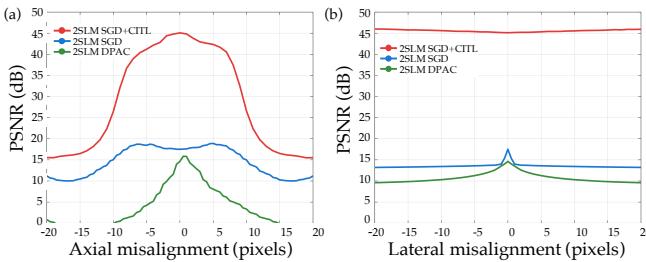


Fig. 5. Evaluating SLM alignment robustness in simulation. The axial and lateral misalignment between two SLMs can cause a significant image degradation, even for a misalignment on the order of the size of a single pixel. Whereas DPAC is very sensitive to small amounts of SLM alignments, SGD approaches are more robust to these with the CITL-calibrated SGD variant achieving by far the best and most robust results.

REFERENCES

1. J.-S. Chen and D. Chu, Opt. Express **23**, 18143 (2015).
2. G. Li, D. Lee, Y. Jeong, J. Cho, and B. Lee, Opt. Lett. **41**, 2486 (2016).
3. A. Maimone, A. Georgiou, and J. S. Kollin, ACM Trans. Graph. (SIGGRAPH) **36**, 85 (2017).
4. H. Yu, K. Lee, J. Park, and Y. Park, Nat. Photonics **11**, 186 (2017).
5. L. Shi, F.-C. Huang, W. Lopes, W. Matusik, and D. Luebke, ACM Trans. Graph. (SIGGRAPH Asia) **36**, 236:1 (2017).
6. P. Chakravarthula, Y. Peng, J. Kollin, H. Fuchs, and F. Heide, ACM Trans. Graph. (SIGGRAPH Asia) **38** (2019).
7. J.-H. Park and M. Askari, Opt. express **27**, 2562 (2019).
8. N. Padmanaban, Y. Peng, and G. Wetzstein, ACM Trans. Graph. (SIGGRAPH Asia) **38** (2019).
9. B. Lee, D. Yoo, J. Jeong, S. Lee, D. Lee, and B. Lee, Opt. Lett. **45**, 2148 (2020).
10. G. Kuo, L. Waller, R. Ng, and A. Maimone, ACM Trans. Graph. **39** (2020).
11. D. E. Smalley, Q. Smithwick, V. Bove, J. Barabas, and S. Jolly, Nature **498**, 313 (2013).
12. E. Ronzitti, M. Guillon, V. de Sars, and V. Emiliani, Opt. express **20**, 17843 (2012).
13. J. Cho, S. Kim, S. Park, B. Lee, and H. Kim, Opt. Lett. **43**, 3397 (2018).
14. H. Kim, C.-Y. Hwang, K.-S. Kim, J. Roh, W. Moon, S. Kim, B.-R. Lee, S. Oh, and J. Hahn, Appl. optics **53**, G139 (2014).
15. D. Palima and V. R. Daria, Appl. optics **46**, 4197 (2007).
16. J. Liang, S.-Y. Wu, F. K. Fatemi, and M. F. Becker, Appl. optics **51**, 3294 (2012).
17. W. Improso, G. A. Tapang, and C. A. Saloma, in *Photonics, Optics, and Laser Technology*, (2017), pp. 208–214.
18. Y. Peng, S. Choi, N. Padmanaban, and G. Wetzstein, ACM Trans. Graph. (SIGGRAPH Asia) (2020).
19. P. Chakravarthula, E. Tseng, T. Srivastava, H. Fuchs, and F. Heide, ACM Trans. Graph. (SIGGRAPH Asia) (2020).
20. J. Hahn, H. Kim, Y. Lim, G. Park, and B. Lee, Opt. Express **16**, 12372 (2008).
21. F. Yaraş, H. Kang, and L. Onural, in *3DTV-Conference*, (IEEE, 2010), pp. 1–4.
22. R. D. Juday and J. M. Florence, in *Wave Propagation and Scattering in Varied Media II*, , vol. 1558 (1991), pp. 499–504.
23. S. Park, J. Roh, S. Kim, J. Park, H. Kang, J. Hahn, Y. Jeon, S. Park, and H. Kim, Opt. express **25**, 3469 (2017).
24. A. Macfaden and T. Wilkinson, JOSA A **34**, 161 (2017).
25. C.-K. Hsueh and A. A. Sawchuk, Appl. optics **17**, 3874 (1978).
26. R. Horisaki, R. Takagi, and J. Tanida, Appl. optics **57**, 3859 (2018).
27. M. H. Eybposh, N. W. Caira, M. Atisa, P. Chakravarthula, and N. C. Pégard, Opt. Express **28**, 26636 (2020).
28. J. Lee, J. Jeong, J. Cho, D. Yoo, B. Lee, and B. Lee, Opt. Express **28**, 27137 (2020).