Progress in off-plane computer-generated waveguide holography for near-to-eye 3D display

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

As Published
http://dx.doi.org/10.1117/12.2209649

Publisher
SPIE, the International Society of Optical Engineering

Version
Final published version

Accessed
Sun Feb 10 20:36:29 EST 2019

Citable Link
http://hdl.handle.net/1721.1/116373

Terms of Use
Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.

Detailed Terms
Progress in off-plane computer-generated waveguide holography for near-to-eye 3D display

Sundeep Jolly, Nickolaos Savidis, Bianca Datta, V. Michael Bove, Daniel Smalley


Event: SPIE OPTO, 2016, San Francisco, California, United States
Progress in off-plane computer-generated waveguide holography for near-to-eye 3D display

Sundeep Jolly\textsuperscript{a}, Nickolaos Savidis\textsuperscript{a}, Bianca Datta\textsuperscript{a}, V. Michael Bove, Jr.\textsuperscript{a}, and Daniel Smalley\textsuperscript{b}

\textsuperscript{a}MIT Media Lab, 77 Massachusetts Ave., Rm. E15-444, Cambridge, MA, United States
\textsuperscript{b}Department of Electrical and Computer Engineering, Brigham Young University, 459 Clyde Building, Provo, UT, United States

ABSTRACT

Waveguide holography refers to the use of holographic techniques for the control of guided-wave light in integrated optical devices (e.g., off-plane grating couplers and in-plane distributed Bragg gratings for guided-wave optical filtering). Off-plane computer-generated waveguide holography (CGWH) has also been employed in the generation of simple field distributions for image display. We have previously depicted the design and fabrication of a binary-phase CGWH operating in the Raman-Nath regime for the purposes of near-to-eye 3-D display and as a precursor to a dynamic, transparent flat-panel guided-wave holographic video display. In this paper, we describe design algorithms and fabrication techniques for multilevel phase CGWHs for near-to-eye 3-D display.

Keywords: holography, near-to-eye display, waveguide optics

1. INTRODUCTION

Near-to-eye displays act to directly project imagery into a viewer’s eye and can range in instantiation from extremely simple (such as an optical viewfinder) to more complex immersive displays for applications in virtual reality. Near-to-eye displays can be well-suited to recreate three-dimensional imagery for a viewer, typically through stereoscopic\textsuperscript{1} or other multiscopic means (e.g., multifocal\textsuperscript{1} or light-field\textsuperscript{2} displays). Like all 3-D display systems, near-to-eye 3-D display systems need to meet the challenge of displaying imagery such that a viewer correctly and comfortably perceives a 3-D scene at the output. In addition, near-to-eye 3-D displays additionally have to be designed for maximal viewer comfort and preferably in such a manner that limits bulkiness and weight. Historically, near-to-eye 3-D displays have been implemented using techniques such as stereoscopy (employing simple magnifiers),\textsuperscript{1} although more complex schemes for near-to-eye display based around lenticular imaging\textsuperscript{2} or factored light-field techniques\textsuperscript{3} have recently emerged.

1.1 Near-to-Eye Display Design Considerations

In typical glasses-bound stereoscopic, autostereoscopic, and automultiscopic displays, the accommodation-vergence conflict refers to the notion that a viewer’s eyes are focused at the distance of the screen while they are converged upon the distance to an apparent 3-D object (see Fig. 1 (a)). This mismatch often causes issues such as visual discomfort, fatigue, and distorted depth perception for the viewer.\textsuperscript{4} In many current see-through near-to-eye displays for augmented reality applications, there exists an accommodation-vergence mismatch between the real-world scenery and the displayed imagery (which is typically presented on a flat 2-D surface).\textsuperscript{5} The viewer is often made to accommodate on displayed imagery and also to accommodate and converge upon objects in reality (see Fig. 1 (b)). This is typically difficult because the accommodation distances of the displayed imagery and the object in reality are far apart such that simultaneous accommodation on both is impossible. For stereoscopic displays, a viewer focusing at the plane of the display will perceive all imagery – regardless of the apparent distance intended – as being in focus.\textsuperscript{6} In addition to an accommodation-vergence mismatch, the utility of accommodation for depth perception – a moderately strong cue for objects closer than 1 m to the eye – is greatly diminished in this situation.

Unlike other displays, near-to-eye displays additionally face a particular challenge because of the accommodation range of the human eye. This means that, e.g., unaided microdisplays, cannot be resolved by the human eye at very close distances unless supporting optics are used. Simple lenses have often been used in conjunction with microdisplays to create virtual
Figure 1. Accommodation-vergence mismatch scenarios in near-to-eye augmented reality displays. (a) Accommodation-vergence mismatch resulting from the display of synthetic 3-D imagery by presenting disparate views to both of the left and right eyes, in which the eyes are forced to accommodate on the screen plane but converge upon the apparent depth of the synthetic 3-D object. (b) Accommodation-vergence mismatch between, e.g., a displayed 2-D plane of imagery at some apparent depth and a real-world object, in which the eyes are forced to accommodate on the displayed plane of imagery but converge upon the perceived depth of the real-world object.

images at focus planes falling at resolvable accommodation distances from the eye (see Fig. 2 (a), in which a magnified virtual image is created at a longer distance from the eye than the microdisplay); however, magnification of pixilated displays creates undesirable image artifacts due to decreased spatial resolution.

1.2 Waveguide Holography and Near-to-Eye Holographic Display

Waveguide holography typically refers to the use of holographic techniques for the control of guided-wave light in integrated optical devices (e.g., grating couplers and distributed Bragg gratings for guided-wave optical filtering) and for general control of guided-wave light via diffraction. Image display via optically-recorded waveguide Bragg holograms has been explored. Moreover, computer-generated waveguide holography (CGWH) has been demonstrated for the generation of free-space spot arrays and for the integrated implementation of several optical functions. Additionally, holographic optical elements have been used for near-to-eye wearable displays for the incoupling of microdisplay-modulated light into guided-wave modes and for the outcoupling of that light towards a viewer; however, such displays rely on waveguide and holographic techniques to relay the light from a microdisplay to a viewer and do not provide holographic imagery directly to a viewer.

Holographic display systems act to reconstruct light wavefronts emanating from 3-D scenes for a viewer and are naturally capable of handling all principal cues to depth – accommodation, vergence, motion parallax, and binocular disparity – in a visually and perceptually consistent manner. Of all the competing approaches for near-to-eye display, holography is the least explored for the display application. Goetz et al. explored holographic near-to-eye displays for vision correction but not for the general 3-D display application.

Recently, a waveguide-based platform for acousto-optic spatial light modulation for holographic video displays has emerged and shows several notable advantages over more conventional techniques for spatial light modulation for holographic video displays, such as the potential for no quantization noise or undesired diffracted orders upon reconstruction and full-color imagery via wavelength-division multiplexing (see Fig. 3). Although this modulation scheme has only to-date been demonstrated in a conventional scanning-type holographic display, the underlying surface acoustic wave technology can be extended to support a display requiring no additional supporting optics, effectively operating as a transparent, flat-panel holographic display.
Figure 2. Virtual image formation in microdisplay-based near-to-eye displays and holographic near-to-eye displays. (a) Virtual image generation from a microdisplay via a simple lens. Light rays reflected off a microdisplay located at distance $d_e + d_o$ from the eye are incident upon a plano- or bi-convex lens located at $d_e$ from the eye (i.e., eye relief distance), which acts to create a magnified virtual image of the microdisplay at a distance $d_i + d_o$ from the eye, where $d_i = f d_o / (f + d_o)$ and the magnification is $M = d_i / d_o = f / (f + f_o)$. While the original distance of the microdisplay $d_e + d_o$ is not within the accommodation range of the eye, the virtual image distance $d_i + d_o > d_o + d_e$ does fall within the accommodation range and the virtual image is therefore resolvable. (b) Virtual image generation from a near-to-eye hologram. The hologram itself recreates a wavefront of the desired image plane at an perceived image distance $d_i + d_e$ from the eye, without the need for any additional optics. Note that the distance of the virtual image from the eye falls within the accommodation range and the virtual image is therefore resolvable.
Figure 3. Structure of an anisotropic leaky-mode modulator, in which guided-mode light interacts with and is modulated by a nearly-collinear propagating surface acoustic wave containing holographic information. The interaction forces the guided-mode light to out-couple from the waveguide into a leaky-mode, which is then used as the spatially modulated light in a holographic display optical system.

Figure 4. Space-bandwidth product (base-10 logarithm) as a function of eye-display distance (i.e., eye relief) and viewer rotation angle for a fixed field of view $\theta_{\text{max}} = 3^\circ$ and operating wavelength $\lambda = 600$ nm. Note that for a typical free-standing holographic video display located at close to 1 m from a viewer with an allowable viewer rotation of up to $45^\circ$, the SBP approaches values in excess of $10^{13}$. However, for a typical near-to-eye instantiation of a holographic video display with an eye relief distance of approximately 15 mm and eye rotation angle of $15^\circ$, the SBP takes values between $10^7$ and $10^8$ – several orders of magnitude lower than that for a typical free-standing HVD.
Recent work has furthermore shown the potential of holographic display technologies in improving perceptual performance over competing techniques for 3-D display.\textsuperscript{14} Fig. 2 (b) depicts display holography in a near-to-eye configuration, in which no optics beyond the hologram (i.e., spatial light modulator) itself are required, as the wavefront information carried by the hologram recreates wavefronts at proper apparent distances (i.e., within the eye’s accommodation range) – and importantly, with consistent visual cues to depth (i.e., no accommodation-vergence mismatch) – making holographic displays compelling for the near-to-eye application due to their potential to be implemented with a limited footprint and complexity in optics. Furthermore, because of naturally-matched accommodation and vergence cues and high reconstruction fidelity, holography is well-poised as a contender for future near-to-eye 3-D displays.

1.3 Space-Bandwidth Product Considerations

The space-bandwidth product of a spatial light modulator dictates the achievable angular field-of-view and spatial extent for the output of a holographic video display built upon it. Briefly, this unitless number is typically a function of the number of pixels and the pixel pitch (which dictates the maximal diffracted angle, as mapped from maximal spatial frequency as \( \theta_{max} = \sin^{-1}(f_{max} \lambda) \), where \( \lambda \) is the operating wavelength) for a pixellated spatial light modulator (such as an LCoS or DMD device) or a function of the temporal operating bandwidth for an analog acousto-optic modulator. Alternatively, one can consider directly the space-bandwidth required of a holographic display to support a viewing configuration for a given viewer-display distance, viewer rotation angle, and desired field of view.\textsuperscript{15}

In order to illustrate possible space-bandwidth product requirements of a near-to-eye holographic video display, we consider a field of view of 3\(^\circ\), which is consistent with the field of view in which retinal resolution is highest (i.e., observations are made with foveal acuity).\textsuperscript{16} We also consider only the case in which the viewer can rotate but not translate laterally or axially relative to the display plane (e.g., the eye can rotate but will be fixed both laterally and axially relative to a head-mounted display). Following the analysis in Onural et. al.,\textsuperscript{15} the total 1-D spatial extent of a spatial light modulator located a distance \( z_H \) away from a viewer (in the near-to-eye case, the eye) with an ocular field of view of \( \theta_{max} = 3\,^\circ \) and allowable eye rotation angle \( \theta_R = 15\,^\circ \) can be found as

\[
l_{total} = 2z_H \tan(\theta_{max} + \theta_R)
\]

and the maximal spatial frequency required for that ocular field of view and allowable eye rotation angle can be found as

\[
f_{max} = \frac{\sin(\theta_{max} + \theta_R)}{\lambda}
\]

where \( \lambda \) is the operating wavelength. For the two-dimensional case, the spatial extent becomes a circle of diameter \( l_{total} \) in the spatial domain and the bandwidth becomes a circle of diameter \( 2f_{max} \) in the Fourier domain. Then, the 2-D space-bandwidth product is found as

\[
SBP = \pi \left( \frac{l_{total}}{2} \right)^2 \times \pi \left( \frac{2f_{max}}{2} \right)^2 = \left[ \frac{\pi z_H \sin(\theta_{max} + \theta_R) \tan(\theta_{max} + \theta_R)}{\lambda} \right]^2
\]

This expression is plotted against eye-display distance and viewer rotation angle in Fig. 4 for a fixed field of view \( \theta_{max} = 3\,^\circ \) and operating wavelength \( \lambda = 600 \text{ nm} \). Highlighted are regions of interest for typical near-to-eye holographic video displays and typical free-standing or tabletop holographic video displays. Note that the SBP requirement is mitigated from upwards of 10\textsuperscript{13} for a free-standing HVD to 10\textsuperscript{7} for a near-to-eye instantiation – and thereby reveals that the problem of near-to-eye holographic video display – while perhaps not trivially easy – is considerably more manageable, from a light-modulation and computational perspective, relative to that of free-standing or tabletop HVDs.

Fig. 5 depicts the space-bandwidth product given in Eq. 3 as a function of field of view angle \( \theta_{max} \) for a fixed eye relief distance \( z_H = 15 \text{ mm} \) and maximal eye rotation angle \( \theta_R = 15\,^\circ \). While still remaining relatively manageable (in an order-of-magnitude sense), the SBP requirement increases monotonically for increasing field of view. One can therefore conclude that the most immersive near-to-eye holographic experiences will therefore require increased capability – both from a light-modulation and computational perspective – and that FOV should be an important consideration in the design of near-to-eye HVDs.
Having established basic feasibility of near-to-eye holographic video display from the perspective of the space-bandwidth product, we now explore techniques for static, waveguide-based computer-generated holography for near-to-eye display. As a precursor to dynamic, waveguide-based holography implemented as a transparent flat-panel display, prototyping such computer-generated waveguide holograms in a static form will allow for investigation into illumination techniques, computational considerations, and perceptual considerations important for a future near-to-eye, transparent, flat-panel dynamic holographic display based around guided-wave acousto-optics. In order to inform the development of such a future display, the current work intends to address questions related to the optimization of light delivery over the active area of the hologram, tailoring of the diffraction efficiency over the hologram active area in order to ensure a uniformly available dynamic range, and computation of holograms for the near-to-eye instantiation provided limitations on the human visual system and already mentioned considerations on space-bandwidth product resulting from near-to-eye placement of the display.

2. METHODS

For the purposes of the current experiment, the computer-generated holograms we employ are computed such that the displayed output consists of a 2-D plane of imagery. This geometry is depicted in Fig. 6.

Our computation scheme employs the Fresnel-transforming Gerchberg-Saxton method for phase-only hologram computation. This algorithm produces a phase-only CGH pattern that generates a desired plane of imagery at some distance from the hologram plane. A single loop of this iterative process is depicted in Fig. 7. Briefly, we begin with a complex field consisting of a randomly-initialized phase pattern having unit amplitude. This phase pattern is Fresnel transformed to yield the propagated field at the image plane. The Fresnel field arriving at the image plane is Fresnel transformed via the 2-D Fresnel diffraction integral:

\begin{equation}
\mathbf{E}(\mathbf{r}) = \mathcal{F}[\mathcal{F}^{-1}[\mathbf{P}] \mathcal{F}^{-1}[\mathbf{E}] - \mathbf{P}]\end{equation}

where \( \mathbf{E} \) is the electric field, \( \mathbf{P} \) is the phase pattern, and \( \mathcal{F} \) is the Fresnel transform.
Figure 6. Diffraction geometry used in generating the CGH pattern. Assuming planar input illumination, the Fresnel field at the hologram plane is enforced to be phase-only via the Gerchberg-Saxton algorithm. The desired imagery is generated at a distance $z_I$ from the hologram plane.

$$U_o(x, y, \Delta z) = \frac{e^{jk\Delta z}}{j\Delta z} e^{\frac{jk}{\Delta z} (x'^2+y'^2)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U_i(x', y') t(x', y') e^{\frac{jk}{\Delta z} (x'^2+y'^2)} e^{-\frac{2\pi}{\lambda z}(xx'+yy')} dx' dy'$$  \hspace{1cm} (4)

where $U_o(x, y, \Delta z)$ is the Fresnel field arriving at the hologram plane, $U_i(x, y)$ is the plane wave illumination field with form $U_p(x, y) = A_0 \exp(jk\phi_0)$, $t(x, y)$ is the amplitude transmittance function of the diffracting aperture (i.e., the planar image), $k = 2\pi/\lambda$ is the wavenumber, $\Delta z$ is the propagation distance, $(x, y)$ are the spatial coordinates at the input plane, and $(x', y')$ are the spatial coordinates at the input plane. This integral is evaluated as a discrete Fourier transform using the Fast Fourier Transform algorithm.

At this point, the first estimate of the approximated intensity profile at the image plane is found via the squared-modulus of the complex Fresnel field at the image plane. On subsequent iteration loops, this estimate is compared to the desired image reconstruction and the results of comparison are used as a basis for ending the optimization (i.e., $\|I_{desired} - I_{estimate}\|^2 < \epsilon$, where $I_{desired}$ is the intensity profile of the desired 2-D imagery, $I_{estimate}$ is the approximated intensity profile arriving at the image plane via Fresnel transform of the optimized phase CGH pattern, $\|\cdot\|^2$ indicates the square of the $L^2$ norm, and $\epsilon$ is a small error threshold). If the criterion for ending the optimization is not met, the target intensity is used to amplitude-enforce the Fresnel field at the image plane (i.e., the square root of the target intensity is set to be the new amplitude) and that field is inverse Fresnel-transformed to find the complex Fresnel field arriving at the hologram plane. In order to meet fabrication constraints, the number of allowable phase levels is set to 8 and the constrained phase profile is the resulting phase CGH after an iteration. On a new iteration, the constrained phase profile is equalized with a unit-amplitude plane wave and that complex field is the input for a Fresnel transform again to the image plane. Here, another image comparison can take place.

Using target imagery of the MIT Media Lab logo, the Fresnel-transforming Gerchberg-Saxton algorithm was employed over 50 iterations. The resulting phase CGH profile and approximated intensity arriving at the image plane are depicted in Figs. 8 and 9.

In order to prototype computer-generated waveguide holograms for near-to-eye applications, our approach employs electron-beam lithographic techniques to fabricate phase-retarding diffractive patterns in PMMA on quartz substrates. Fig. 10 depicts a typical fabrication process for such a prototype. After beginning with a quartz (SiO$_2$) substrate, 600 nm poly(methyl methacrylate) (PMMA) is spin-coated on and post-baked. In order to avoid charging of the substrate via the electron-beam write process, 10 nm of chrome is deposited on top of the PMMA via evaporation. Using an Elionix ELS-F125 electron-beam lithography system operating at 125 KeV and employing a beam current of 100 nA, the grayscale phase CGH is...
Figure 7. Iteration loop for phase-only CGH computation via the Fresnel-transforming Gerchberg-Saxton method. Beginning with a random phase profile with unit amplitude, this complex hologram is Fresnel transformed to find the complex field arriving at the image plane. The intensity profile is compared with the desired imagery, and if the error criterion is not satisfied, the iteration loop continues as the target imagery is used to amplitude-enforce the Fresnel field at the image plane. The field is then inverse Fresnel-transformed to find the field at the hologram plane. After constraining the phase-profile to 8 phase levels, the resulting phase-only CGH is equalized with a unit-amplitude plane wave and a new iteration loop begins.

Figure 8. Optimized phase profile at hologram plane after 50 iteration of the Fresnel-transforming Gerchberg-Saxton algorithm with desired imagery of the MIT Media Lab logo.

Figure 9. Approximated intensity of the MIT Media Lab logo arriving at the image plane after Fresnel-transforming the optimized phase profile in Fig. 8 found after 50 iterations of GS.

Written in the PMMA with a base dose of 1400 µC/cm² (note the actual dosage varies from that base dose according to the gray level of the phase CGH). The chrome layer is removed via a chrome etchant and the exposed PMMA is developed in a solution of 4:1 isopropanol:methylisobutylketone for 30 s. An SEM image of a portion of an example written CGH in PMMA on top of silicon is shown in Fig. 11.

The replay process for our written computer-generated waveguide holograms is depicted in Fig. 12. Briefly, laser light is edge-coupled into the quartz substrate onto which the CGH is written at an angle higher than the critical angle for
Figure 10. Fabrication process for a static off-plane computer-generated waveguide hologram. (a) Single-crystal quartz or fused silica is used as the substrate. (b) 600 nm of PMMA is deposited via spin-coating and post-baking. (c) The PMMA is exposed via grayscale electron-beam lithography. (d) The exposed PMMA is developed in an IPA:MIBK solution.

Figure 11. Scanning electron micrograph of a sample lithographed phase-only CGH in PMMA resist on silicon substrate. Square pixels are 500 nm in dimension.
the air-quartz interface (for single crystal quartz with \( n = 1.54 \), \( \theta_c = \sin^{-1}(1.00/1.54) \approx 40.49^\circ \)). By definition, the quartz substrate acts as a multi-modal waveguide and confines the illumination light via total internal reflection. When the totally internally reflected light encounters the written CGH, it is outcoupled in a similar manner to waveguided light that encounters a grating outcoupler. The CGH acts as a Fourier sum of many single-frequency grating outcouplers and is therefore well-suited to outcouple (and spatially phase) the guided light. The diffracted light from the CGWH then propagates away from the device for viewing.

Fig. 13 depicts the edge-coupled laser light in the quartz substrate, the outcoupling via the CGH towards a screen, and the formation of strongly-scattered diffracted orders. Note that this diffraction results from a non-image forming CGH and is intended to depict the outcoupling performance of a CGWH. Fig. 14 depicts the optical reconstruction of the phase CGH depicted in Fig. 8. Aside from some strong zero-order diffracted scatter, the reconstruction presents good discriminability and verifies the computational and fabrication processes.

Fig. 12. Replay process for a static off-plane computer-generated waveguide hologram. (a) Laser light is edge-coupled into the glass substrate at an angle higher than the critical angle. (b) The coupled light is confined into a guided-mode via total internal reflection. (c) The coupled light encounters the CGH and is outcoupled from the substrate and modulated with the holographic information, enabling reconstruction.

3. CONCLUSIONS AND FUTURE WORK

We have presented a space-bandwidth analysis of near-to-eye holographic displays and depicted computational and fabrication techniques for static near-to-eye displays based around computer-generated waveguide holograms. We anticipate that future work resulting from such techniques will continue to inform the design and development of waveguide-based, dynamic near-to-eye holographic displays.

ACKNOWLEDGEMENTS

This research has been supported by consortium funding at the MIT Media Laboratory, by the Center for Terrestrial Sensing at the MIT Media Laboratory, and by Air Force Research Laboratory contract FA8650-14-C-6571. The authors gratefully acknowledge facility use and technical assistance by the MIT Nanostructures Laboratory, the MIT Scanning Electron-Beam Lithography facility, and the MIT Research Laboratory of Electronics. The authors would also like to thank NVIDIA for their generous donation of the GPU hardware used in this work.
Figure 13. Diffractive outcoupling from edge-lit computer-generated waveguide hologram, depicting total internal reflection of edge-coupled illumination light and the formation of strongly-scattered diffracted orders on a screen.

Figure 14. Diffractive reconstruction of phase-only CGH computed using the Fresnel-transforming Gerchberg-Saxton algorithm and fabricated in PMMA on quartz substrate via e-beam lithography.
REFERENCES