HoloBeam: Paper-Thin Near-Eye Displays

Kaan Akşit*
University College London

Yuta Itoh†
The University of Tokyo

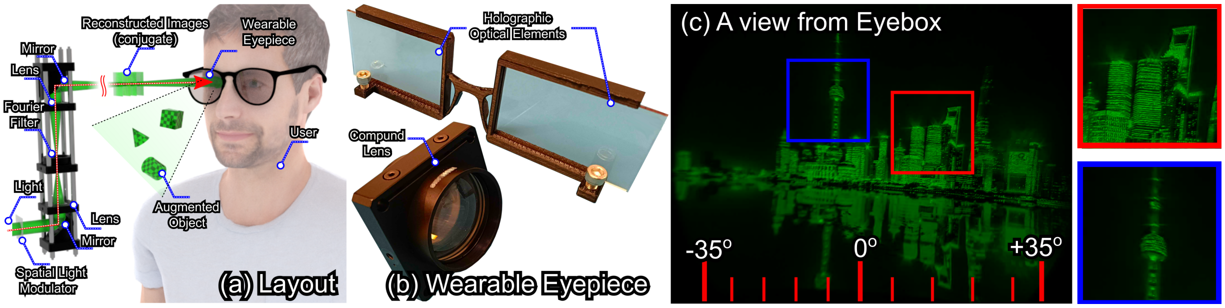

Figure 1: *HoloBeam* slim Virtual Reality headsets and Augmented Reality glasses with wide field of view and high resolutions. (a) Layout. A light beam from a light source gets modulated by a Spatial Light Modulator (SLM). Modulated beam reconstructs multiplane images in the vicinity of SLM (a few millimeters away). With the help of multiple lenses forming a projection system, conjugate images are formed at considerable throw distances (e.g., a few tens of cm to a meter or two). These conjugate images are relayed to a user's retina with the help of a slim and lightweight eyepiece. Thus, users can perceive high-resolution 3D virtual images overlayed on their real-world scene. (b) Wearable Eyepiece. *HoloBeam* glasses is made of wearable eyepieces that could come in the form of a submillimeter-thick holographic optical elements or conventional lenses with a reasonable thickness. (c) A view from eyebox. Provided photograph demonstrates the high image quality and wide field of view of a phase-only modulating *HoloBeam* prototype when viewed through eyepieces that use conventional lenses. Source image is from Flicker user gags9999.

ABSTRACT

An emerging alternative to conventional Augmented Reality (AR) glasses designs, Beaming displays promise slim AR glasses free from challenging design trade-offs, including battery-related limits or computational budget-related issues. These beaming displays remove active components such as batteries and electronics from AR glasses and move them to a projector that projects images to a user from a distance (1-2 meters), where users wear only passive optical eyepieces. However, earlier implementations of these displays delivered poor resolutions (7 cycles per degree) without any optical focus cues and were introduced with a bulky form-factor eyepiece (~ 50 *mm* thick). This paper introduces a new milestone for beaming displays, which we call *HoloBeam*. In this new design, a custom holographic projector populates a micro-volume located at some distance (1-2 meters) with multiple planes of images. Users view magnified copies of these images from this small volume with the help of an eyepiece that is either a Holographic Optical Element (HOE) or a set of lenses. Our *HoloBeam* prototypes demonstrate the thinnest AR glasses to date with submillimeter thickness (e.g., HOE film is only 120 μm thick). In addition, *HoloBeam* prototypes demonstrate near retinal resolutions (24 cycles per degree) with a 70 degrees-wide field of view.

Index Terms: Human-centered computing—Visualization—Visualization techniques—Treemaps; Human-centered computing—Visualization—Visualization design and evaluation methods

1 INTRODUCTION

Emerging as the future's interface [64], Virtual Reality (VR) and Augmented Reality (AR) glasses are expected to redefine how we interact with real and virtual environments by overlaying rendered images within our visual Field Of View (FoV). To achieve this vision, recent years have seen a strong push from the scientific community to build AR [59, 60] and VR [47, 68] glasses that are thin and lightweight. Video see-through AR [28] and VR headsets block the real-world view, whereas optical see-through AR has no such luxury to do so. Expectation from both AR glasses and VR headsets is that they should support high resolutions, wide FoV, and optical focus cues [1]. To date, VR headsets and AR glasses in the literature have struggled to balance these requirements, often yielding to issues related to one or more requirements (e.g., slim but low FoV, bulky but with optical focus cues).

An emerging design alternative to conventional designs for AR glasses, Beaming Displays [39], argues that designs should separate active and passive parts in AR glasses into two discrete components. Specifically, these Beaming Displays project images from a distance (1-2 m) to a user wearing a passive lightweight optical eyepiece. Thanks to removing active components from AR glasses, Beaming Displays fundamentally avoid heating, computation, and power budget-related issues. At the core, Beaming Displays aim to balance design requirements such as form-factor, resolution, FoV, and optical focus cues. However, the previous implementation of Beaming Displays [39] provided a limited resolution with 7 Cycles Per Degree (cpd) and 50 mm thick bulky AR glasses. Thus, the promises of Beaming Displays have yet to be fully realized.

This paper introduces a new Beaming Display named *HoloBeam* following recent Computer-Generated Holography (CGH) algoritmic methods and holographic eyepiece designs. In *HoloBeam*, a holographic projector reconstructs multiplanar images at a target location away from that projector (1-2 m). As the reconstructed images are small, a user wearing an eyepiece made from a Holographic Optical Element (HOE) perceives a magnified version of the reconstructed multiplanar images. Unlike previous implementations of Beaming Displays, *HoloBeam* delivers slim AR glasses with high resolutions (24 cpd) and wide FoV (70 degrees). However, this exploration in improving Beaming Displays with *HoloBeam* comes at the cost of users being fixated in front of the projector, where

*e-mail: k.aksit@ucl.ac.uk
†e-mail: yuta.itoh@iii.u-tokyo.ac.jp

Table 1: Comparison between AR and VR near-eye displays. Here, focus refers to the method used to support the optical focus cues. Optical see-through refers to the level of see-through in the real world. Wide eye-box refers to supporting varying gazes of users (above 10 mm). Moderate monocular fields of view are 20-50 degrees. Moderate resolution matches 10-20 cycles per degree. Although our work offers no mobility, it distinguishes itself as the slim and lightweight AR near-eye display, free from heating issues or limited power and computing resources.

	Focus	See-through	Eyebox	Field of View	Resolution	Form factor	Weight	Power and Compute	Heat	Mobility
This work	CGH	Clear	Small	Wide	High	Paper-Thin	Light	Expandable	No issue	Fixed
Beaming Displays [39]	Fixed	Moderate	Wide	Moderate	Low	Bulky	Regular	Expandable	No issue	Limited
Holographic VR [47]	CGH	Blocks	Small	Wide	High	Thin	Regular	Limited	Issue	Mobile
Video MR [28]	Multiplane	Video	Wide	Moderate	Moderate	Bulky	Regular	Limited	Issue	Mobile
Wide Étendue [54]	CGH	Moderate	Wide	Wide	Low	Bulky	Regular	Limited	Issue	Mobile
Foveated AR [48]	Varifocal	Moderate	Small	Wide	High	Bulky	Regular	Limited	Issue	Mobile
Scanning Eyebox [41]	CGH	Clear	Wide	Moderate	Moderate	Bulky	Regular	Limited	Issue	Mobile
Varifocal AR [6]	Varifocal	Moderate	Wide	Wide	Moderate	Bulky	Regular	Limited	Issue	Mobile
Holographic AR [59]	CGH	Clear	Small	Wide	High	Thin	Regular	Limited	Issue	Mobile
Lightfield VR [36]	Lightfield	Blocks	Wide	Moderate	Low	Bulky	Regular	Limited	Issue	Mobile
Pinlight Displays [60]	Lightfield	Moderate	Small	Moderate	Low	Thin	Regular	Limited	Issue	Mobile
Microlens VR [55]	Lightfield	Blocks	Small	Narrow	Low	Thin	Regular	Limited	Issue	Mobile

Beaming Displays provide room for mobility. Specifically, our work introduces the following contributions:

- **Thin Eyepiece Designs for *HoloBeam*.** We provide a theoretical analysis of our slim HOE-based eyepiece design (submillimeter). We verify our findings by building an in-house HOE recording hardware and demonstrate our HOEs in an actual *HoloBeam* prototype.
- **Learned CGH pipeline for *HoloBeam*.** We introduce a learned CGH pipeline that can calculate multiplanar holograms for our display prototypes at interactive rates (24 ms per frame). Uniquely, This learned CGH pipeline generates multiplanar holograms without using depth maps. Our method is the first to attempt a learned 3D hologram generation pipeline with RGB inputs but not RGBD.
- ***HoloBeam* Display Prototypes.** We instantiate two experimental display prototypes to demonstrate our outcomes. In the phase-only prototype, we use a 4k phase-only Spatial Light Modulator (SLM) and a conventional magnifier glass as an eyepiece. In the amplitude-only prototype, we show a cost-effective alternative to this prototype using amplitude-only SLM s and a HOE as an eyepiece.

Our current *HoloBeam* prototypes provide a limited eye-box. At this time, *HoloBeam* focuses on improving form factor, weight, FoV, and image quality, and our implementation do not offer users the freedom to move in front of a projector. Thus, user mobility and eye-box in our approach stand as two key research issues that we aim to deal with in the future. We believe our work could positively impact AR use cases in automotive and office work applications in their current form. Our code is also available at **GitHub:complight/multiholo**.

2 Related Work

Our work introduces a new holographic VR/AR glasses design paving the way for devices that ultimately appear to be ordinary glasses. Hence, we review relevant literature on AR Glasses, the use of HOEs in AR Glasses, and algorithmic approaches in CGH.

2.1 Augmented Reality Glasses

Existing AR glasses can be broadly categorized as conventional and holographic approaches. In both approaches, images generated by a light engine equipped with an SLM are relayed to a user's retina using an eyepiece. Curious readers can consult the work by Koulieris et al. [52] for a detailed survey on design approaches. We also suggest our readers review Tbl. 1 as they read through this section.

Conventional AR Glasses. In our definition, conventional AR glasses rely on static components to relay an incoherent image source to a user's retina. Most common types of these AR glasses rely on an optical relay called bird-bath optics [19, 27], a combination of a beam-splitter and a beam-combiner. Freeform optics [15] can help improve the optical qualities of these conventional AR glasses in terms of resolution, FoV, and form factor. On the other hand, folding optical paths using prism-like waveguides [18, 35] or diffractive waveguides [19] could also help improve the form factors of AR glasses. However, either freeform AR glasses that encapsulate unique diffractive components [63] or other waveguide-based AR glasses are typically fixed focus or often times do not converge to reliable benefits as they often arrive with tradeoffs in FoV, form-factor, and resolution. There are also variants of conventional AR glasses that can provide near-accurate optical focus cues [52]. Such VR headsets or AR glasses with optical focus cues are broadly categorized as varifocal [6, 31, 48], multiplane [25, 28, 57, 62, 78], focal surface [4, 33, 61], focus-invariant [51] and lightfield [5, 36, 55, 60, 74, 83]. Although all these types offer exciting solutions, to our knowledge, they largely suffer from problems related to weight, bulk, power, resolution, form factor, or heat.

Holographic AR Glasses. Unlike traditional focus supporting AR glasses, holographic AR glasses can manipulate both phase and amplitude of light, promising potential improvements in light efficiency, resolution, and dynamic range. Holographic AR glasses can offer sunglasses-like form factors HOEs [59] due to the use of thin diffractive optics. These holographic AR glasses can also offer high resolutions and support optical focus cues with continuous depth representations [59]. However, the eye-box of holographic AR glasses is typically limited due to the limited bandwidths of existing SLMs (étendue). To improve the eye-box of holographic AR displays, researchers explored opportunities in scanning a light source over an SLM or using multiple light sources [32, 41]. Unfortunately, these solutions arrive at the cost of hardware complexity and additional bulk.

Most recently, people have explored structured diffusers [54] and learned phase masks [13] to replace HOEs and expand the eye-box. But, these solutions typically arrive with image quality-related issues and require precision alignment in manufacturing. *HoloBeam* follows common holographic AR glasses that uses HOEs and shares similar shortcomings, namely eye-box-related issues. Unlike previous holographic AR glasses, *HoloBeam* offers significant improvements in fixing issues related to weight, form factor, power, and computation. Beyond our review, we refer our readers to a survey by Chang et al. [21] on holographic AR glasses. To our knowledge,

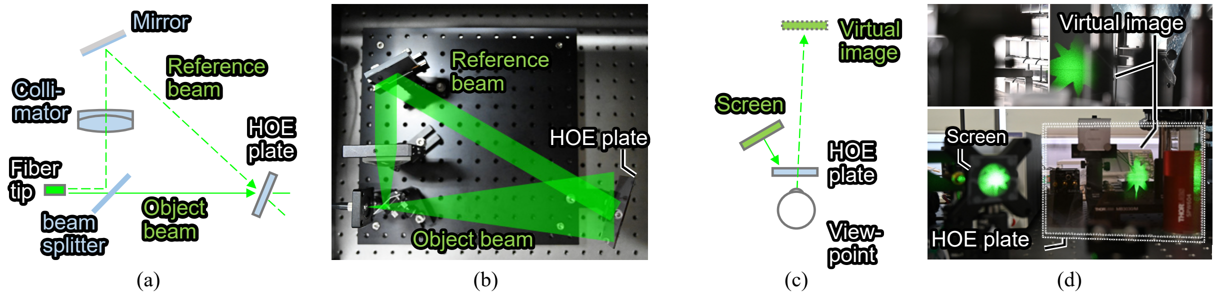

Figure 2: Our HOE recording setup and functional test of a recorded HOE lens. (a) A schematic diagram of our setup to record a transmissive HOE lens. (b) A birds-eye view of the actual recording optics. (c) A schematic diagram of the test, (d) showing the images from the viewpoint.

HoloBeam is the first holographic thin AR glasses that offers a new candidate to resolve issues related to form factor, resolution, FoV, weight, and power.

2.2 Thin Relay Optics for Augmented Reality Glasses

With thickness not exceeding a millimeter, HOEs are considered as thick diffraction gratings (thicker than a few multiples of visible light wavelengths) [24]. These wavelength-selective HOEs interfere efficiently with beams that match angles of incidence in their recording. Optical beams not meeting these conditions will pass through recorded HOEs leading to maintaining high transparency in seeing the real world through AR glasses. Therefore, HOEs can play a central role in AR glasses designs as one of the technologies to realize compact and lightweight optical beam combiners [12, 21]. The manufacturing methods of HOEs can be broadly classified into analog and digital methods. The former reproduce replicas of constitent actual lens systems [75]. In contrast, the latter reproduce a computationally designed optical system by controlling amplitude or phase modulation patterns programmatically in HOE recording [42]. *HoloBeam* follows HOE design and manufacturing methods in analog HOE recording.

2.3 Computer-Generated Holography

CGH comprises a family of methods that approach the hologram calculation problem with algorithmic strategies [16, 80]. Our readers could find an extensive survey of modern CGH methods in the work by Pi et al. [66]. The scene representation methods used at each CGH algorithm could help the categorization of CGH algorithms. Broadly, CGH can rely on scene representations based on the light field [30, 41, 71], multiplane [44, 72, 76], and point clouds [22, 59]. Alternative scene representations such as polygon based ones [38] are also available in the literature. There are also learned algorithmic methods that can help improve speed, accuracy, and image quality in CGH [20, 23, 45, 84]. In our work, we follow multiplane scene representations, and we use gradient-based optimization methods used in many earlier works to generate a dataset of holograms [84]. Using this dataset, we derive a new learned-CGH pipeline. Unlike previous learned-CGH methods, our CGH optimization pipeline does not require depth information of a scene in hologram calculation and decides on the targets by itself. Our CGH pipeline runs at interactive rates on an average compute resource.

3 *HoloBeam*: Holographic Beaming Displays

HoloBeam requires three primary components for a complete display system. These components are a holographic projector, an eyepiece, and a software pipeline to help calculate 3D multiplanar holograms. We also provide the layout of *HoloBeam* in Fig. 1.

3.1 Holographic Projector

Like other display systems, *HoloBeam* requires a light engine that contains light sources and an SLM. Given that *HoloBeam* is a holographic projector, it requires a coherent or partially coherent light source. In the meantime, the SLM of *HoloBeam*, could either modulate phase or amplitude of light or both - full complex. As depicted in the basic layout Fig. 1, the images generated using these SLMs could be imaged to the desired location using a set of lenses, forming 4f or 2f imaging systems or a more advanced multi-lens system. The conventional theoretical limits of Beaming Displays are covered in the original paper of Beaming Displays [39]. However, it should be noted that the limits of resolution in holographic approaches are an open scientific debate [53], which we also agree.

3.2 Eyepiece

Our holographic projector modulates light to reproduce the desired light field near the user's viewpoint. Since this light field is so close to the eye, the eye cannot focus on it. Besides standard optical components (e.g., lenses or mirrors), we can use HOE to convert this light field to images on the retina that users can accommodate and view.

Transmissive HOEs. We base our work on transmissive HOEs using photopolymer films as the material (see Xiong et al. [79] for more). When the photopolymer film is exposed to light, the monomers polymerize according to the interference fringes, creating an unevenly distributed polymer structure. Because of the difference in refractive index between the monomer and polymer, this structure behaves as a phase hologram. If the thickness of the film is sufficiently thicker than the wavelength of the light, we can treat it as a volume hologram. Each diffraction grating structure formed within a volume hologram reflects a portion of the incident light. At a given angle of incidence, if the optical path difference satisfies an integer multiple of the wavelength, all reflected light is intensified. In this case, the incident light is reflected almost entirely at one specific angle. This angle is called the Bragg angle, and the case where the incident light satisfies the Bragg angle is called the Bragg condition.

Based on this principle, diffractive optical elements using volume holograms are selective in wavelength and angle of incidence. Often the case, practically, it is sufficient to consider only the first-order reflected light in the numerical analysis of the reflection direction. If the incident light is sufficiently close to the Bragg angle and the wavelength at the hologram exposure, we can approximate its behavior well with Kogelnik's coupled wave theory [67]. In this theory, the k-vector closure method (KVCM) gives the ray behavior of the incident light (Fig. 3).

Bragg Diffraction Analysis From here, we briefly explain the theory of Bragg diffraction [29, 49, 50, 67]. Fig. 3(b) describes the behavior of incident light in a transmission volume hologram. In this explanation, we also assume the wave vectors are defined inside the photopolymer medium. In real use cases, we have to consider the refraction and reflection at the boundary between the air and the photopolymer plate that consists of several layers, including a protection layer.

Let $\mathbf{n}_\mathrm{r}$ be the wave vector of the reference beam incident on the photopolymer plate when creating a volume hologram, and $\mathbf{n}_\mathrm{s}$

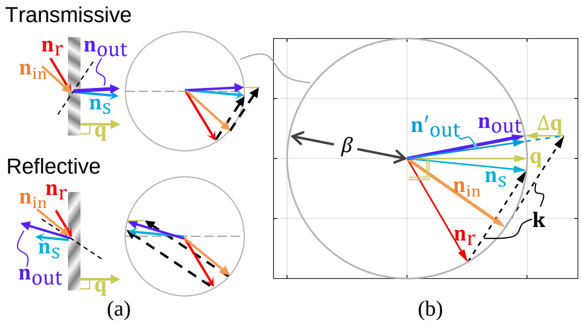

Figure 3: Describing the principle of volume holograms based on the Bragg condition. (a top) A transimissive case. (a bottom) A reflective case. (b) Relationship of wavenumber vectors as described in Sec. 3.2. Note that this setup is a special case when the wavelength of the light at the printing process of the volume holograms and that of the reconstruction are the same.

be the wave vector of the signal beam incident from the opposite side of $\mathbf{n}_\mathrm{r}$. Then, assuming that the two vectors have the same wavelength λ, their lengths or the wave numbers are given as a constant $\beta = 2\pi n_0/\lambda$, where n_0 is the average refractive index of the photopolymer. The two incident beams interfere with each other and create a volume hologram.

We define the k vector $\mathbf{k}$ as the wave vector extending in the direction of the interference fringes created by these two rays. These three vectors then satisfy:

$$\mathbf{n}_\mathrm{s} = \mathbf{n}_\mathrm{r} + \mathbf{k}. \tag{1}$$

Next, let us use this volume hologram to reproduce light. Let $\mathbf{n}_\mathrm{in}$ be the wave vector of the incident light. Then, from the Bragg condition above, we can calculate the reconstructed light $\mathbf{n}'_\mathrm{out}$ as follows:

$$\mathbf{n}'_\mathrm{out} = \mathbf{n}_\mathrm{in} + \mathbf{k}. \tag{2}$$

In general, however, $\mathbf{n}_\mathrm{in}$ does not align with $\mathbf{n}_\mathrm{r}$. Due to this discrepancy, the naive prediction result $\mathbf{n}'_\mathrm{out}$ is known to deviate from the observation and needs to be corrected:

$$\mathbf{n}_\mathrm{out} = \mathbf{n}_\mathrm{in} + \mathbf{k} + \Delta\mathbf{q}, \tag{3}$$

where $\Delta\mathbf{q}$ is a vector with the direction same as the surface normal. Its length is calculated so that $\mathbf{n}_\mathrm{out}$ is placed on the circle with radius β. We define a vector $\mathbf{q}$ perpendicular to the surface with length β.

The above equations only determine the output beam's direction. We are also interested in how much incoming light gets diffracted. The theory gives the diffraction efficiency η as follows [29, 50]:

$$\eta = \left[\nu\,\mathrm{sinc}\left(\sqrt{\nu^2+\xi^2}\right)\right]^2, \tag{4}$$

$$\nu = \frac{\pi n_1 d}{\lambda\sqrt{c_\mathrm{R} c_\mathrm{S}}}, \quad \xi = \frac{||\Delta\mathbf{q}||d}{2c_\mathrm{S}}, \tag{5}$$

$$c_\mathrm{R} = \mathbf{n}_\mathrm{in}^\mathrm{T}\mathbf{q}/\beta^2 (= \cos\theta), \quad c_\mathrm{S} = \mathbf{n}_\mathrm{out}^\mathrm{T}\mathbf{q}/\beta^2, \tag{6}$$

where c_R and c_S are called the obliquity factors [50]. Note that θ is the incident angle of $\mathbf{n}_\mathrm{in}$, n_1 is the refractive index modulation of the photopolymer, and d is the photopolymer thickness (See also eq. (42, 43) in [50]).

As a special case, if there is no slant, i.e. the gratings are aligned with the photopolymer normal, and the input beam is ideal, i.e., $\mathbf{n}_\mathrm{in} = \mathbf{n}_\mathrm{r}$, then $\xi = 0$ and $c_\mathrm{R} = c_\mathrm{S}$. This simplifies the diffraction efficiency to the following as given as eq. (45) in [50],

$$\eta = [\nu\,\mathrm{sinc}(\nu)]^2 = \sin^2\nu = \sin^2\left(\frac{\pi d n_1}{\lambda\cos\theta}\right). \tag{7}$$

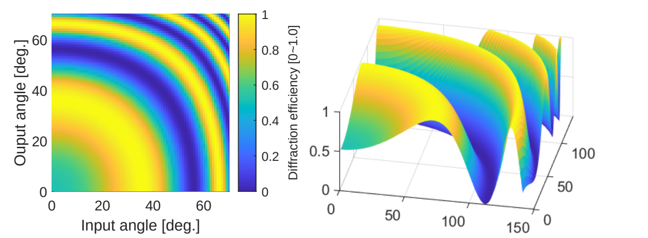

Figure 4: Diffraction efficiency simulation of a transmissive volume hologram. (left) 2D visualization. (right) 3D visualization.

Figure 4 showcases a simulation of volume hologram for Bragg conditions. In this simulation, we set the wavelength $\lambda = 532$ nm, the grating thickness $d = 30[\mu\mathrm{m}]$, and the maximum refractive index modulation $n_1 = 0.04$. The parameters are intended to replicate a commercial photopolymer Beyfol HX120 from Covestro AG. The figure shows when the reference and the signal beams take incident angle pairs between 0 to 70 degrees. One can see that there are angle pairs that give $\eta = 1$, meaning such a condition can direct the input beam to the output direction with almost 100% efficiency.

3.3 Learned Computer-Generated Holography Pipeline

Our learned CGH pipeline builds upon state-of-the-art optimization methodology of hologram generation for dataset creation [23, 26, 44, 47, 56, 73, 84]. Firstly, we introduce this optimization method. Then, we will explain how our learned method differs from these methods and their learned derivatives. To our knowledge, our learned method is the first attempt towards generating 3D holograms only using RGB inputs but not RGBD inputs.

3.3.1 Hologram Dataset Generation

A *HoloBeam* projector generates multiplane images at the desired projection distance by optically relaying images generated close to an SLM. The optical relay here refers to the set of lenses used in front of an SLM as depicted in Fig. 1 layout. As ideally, the optics take care of relaying operation in *HoloBeam*, the remaining computational challenge is calculating the ideal hologram pattern, O_h, for a phase-only SLM to generate 3D images at various image planes $Z_0, Z_1, ..., Z_n$ in close vicinity of a SLM. This calculation could be achieved by propagating a hologram pattern, O_h, using a light transport model. Such a model could simply be written as a convolution operation [77],

$$u(x,y,z) = O_h(x,y) * h(x,y,z), \tag{8}$$

where u represents the complex amplitude at a target image plane, z represents the distance between a hologram and a target image plane, and h represents a convolutional kernel that simulates light transport. Note that h could also be learned from actual hardware using camera captures [45], or could be replaced with a Convolutional Neural Network (CNN) [23] using these captures. The resultant intensity, $|u|^2$, could then be compared against a target image, T, using a loss function, $\mathscr{L}(|u|^2, T)$. Over a few successions of iterations, a hologram could be generated from target images dedicated to each plane $(T_0, T_1, ..., T_n)$. These target images are typically generated using an image or a photograph and their corresponding depth maps (e.g., [23, 44, 47, 84]).

3.3.2 Our Learned Method

Our learned CGH pipeline aims to produce multiplane holograms *without requiring the scene's depth information*. Typical and conventional scenes or images often do not arrive with depth information. Although their depth information could be estimated reliably using a CNN [11], given the wide availability of 2D images and photographs, we see value in easing a user's workflow in generating holograms

```
def estimate(x, n):
    """
    Parameters
    ----------
    x : torch.tensor
        Single color images [kx1xmxn].
    n : torch.nn.modulelist
        Network model used in training.

    Returns
    -------
    O_h : torch.tensor
         Estimated holograms [kx1xmxn].
    """
    y = n.forward(x)
    a = y[:, 0]
    b = y[:, 1]
    ϕ[:, :, 0::2, 0::2] = a[:, :, 0::2, 0::2]
    ϕ[:, :, 1::2, 1::2] = a[:, :, 1::2, 1::2]
    ϕ[:, :, 0::2, 1::2] = b[:, :, 0::2, 1::2]
    ϕ[:, :, 1::2, 0::2] = b[:, :, 1::2, 0::2]
    ϕ → O_h
    return O_h
```

Listing 1: The learned differentiable model used in estimating multiplane phase-only holograms (Pythonic abstraction). This routine runs for every color primary separately.

without access to the depth information. This way, in the future, existing 2D digital content (e.g., games, movies) could potentially be converted to 3D holograms without having two go through multiple steps of estimating depth and generating holograms.

The optimizations discussed previously [23, 44, 47, 84] could be used to generate a hologram dataset from RGBD data. A hologram generation model could then be trained using RGB images as input while discarding their depth channel and their corresponding optimized holograms. Such a model could be trained and tested using the forward model found in Listings 1. In this forward model, a CNN could estimate an output with two channels from a single color of an input image.

In the meantime, we should note that holograms generating images in proximity to an SLM typically generated using the Double-Phase coding method [34, 59, 71]. Inspired by Double-Phase coding, these estimated output channels could then be compiled into a phase-only hologram pattern following a checkerboard-like pattern, maintaining the look of a Double-Phase coded hologram.

4 Implementation

This section will detail the making of HOE in our prototypes, our two prototypes used in our evaluations, and our learned CGH pipeline.

4.1 *HoloBeam* Thin Eyepiece

We used photopolymer sheets as the recording material for our HOE. Specifically, we used holographic film from Litiholo, a 2x3 inch plate consisting of a photopolymer applied to a 2.0 mm glass plate. The photopolymer consists of a 60-micron tri-acetyl cellulose (TAC) film substrate, a 60-micron recording layer, and a protective film layer. The total thickness, including the glass substrate, was 2.28 mm. In HOE recording, the reference and object beams hit the plate simultaneously. If these beams reach the same side of the plate, we obtain a transmissive HOE; if they reach from both sides, we obtain a reflective HOE.

In general, according to the principle of Bragg diffraction, a transmissive HOE requires a shallower diffraction structure than a reflective HOE. In other words, the design tolerance of the diffraction structure is larger for transmission HOEs, and they are more resistant to disruptions such as physical vibrations during fabrication.

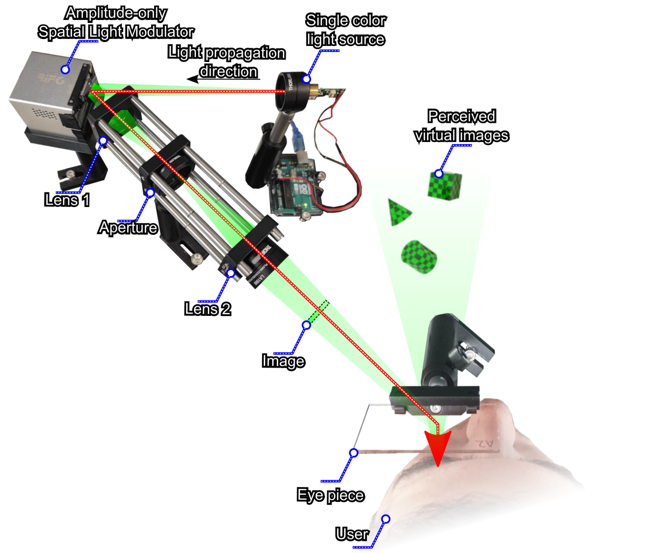

Figure 5: Amplitude-only *HoloBeam* prototype. Here, a single-color laser illuminates an amplitude-modulating SLM. The generated image on the SLM is relayed to an image plane. HOE eyepiece relays the image on that image plane to a user's retina, and the user perceives fixed-focus virtual images. The distance between the projector assembly and the eyepiece is 30 cm and not drawn at true scale to avoid a wider figure.

Therefore, we adopted the transmissive design in this study. For ordinary AR display applications, reflective HOEs are more often used due to the advantages of stray light prevention, high transmission efficiency, and placement of the built-in display.

4.2 *HoloBeam* Prototypes

Amplitude-only *HoloBeam* Our amplitude-only prototype is built to demonstrate how compact, slim, and cost-effective our solution could be. This prototype uses a green laser diode with a 532 nm wavelength and 10 mW optical power. We harvested a laser diode from a generic laser pointer and drove it with an IRF540N MOSFET and an Arduino microcontroller. We collimate this coherent light with a 100 mm focal length lens, Thorlabs LA1509. The collimated beam then arrives at an amplitude-only SLM, specifically a Digital Micromirror Device (DMD) with 854 by 480 pixels and 5.4μm pixel pitch. This specific DMD is harvested from a pico projector, RIF 6 Cube, and we can only push 8-bit frames without any access to timing or individual binary frames.

The modulated beam from DMD is relayed to an image plane with a throw distance of 30 cm using 4f optics. Following the path from DMD towards our image plane, we use a 50 mm plano-convex lens, Thorlabs LA1131, an adjustable aperture at the Fourier plane, Thorlabs SM1D2D, and a 150 mm bi-convex lens, LB1437. This prototype has no beam steering capability and uses a fixated HOE eyepiece at a location 30 cm away from the projection assembly. Fig. 5 shows the entire assembly of our amplitude-only *HoloBeam* prototype, and Fig. 7 shows an example see-through capture. As demonstrated in our supplementary materials, we could also build a vertical-layout assembly.

Phase-only *HoloBeam* We constructed our phase-only *HoloBeam* prototype to emonstrate a system with high image quality and a wide FoV (see Fig. 6). The prototype uses fiber-coupled light sources (420, 520, and 638 nm - Fisba ReadyBeam) and is also equipped with an incoherent RGB LED light source, which can be activated as needed. The phase-only SLM in our prototype is a Jasper Display SLM Research kit (2400 by 4094 pixels and 3.74μm

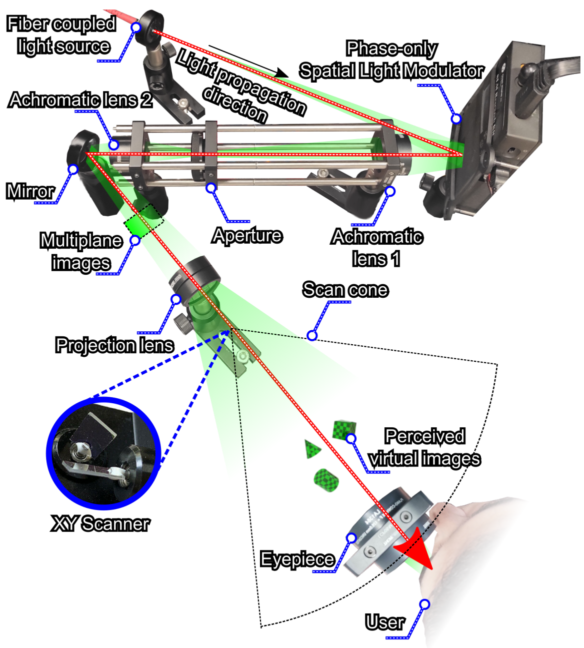

Figure 6: Phase-only *HoloBeam* prototype. A fiber-coupled multicolor light source illuminates a phase-only SLM. Multiplane images generated from this SLM are filtered with a pinhole using an aperture and a 4f imaging system. These images are then projected at a meter distance using a 2f imaging system as a projection lens. Note that an XY scanner follows the projection lens to steer the beam toward a user. A user wearing an eyepiece composed of a cascade of lenses perceives images with multiple focuses. The distance between the projector assembly and the eyepiece is a meter in reality and not drawn at the actual scale to avoid a wider figure.

pixel pitch). The modulated beam from the SLM generates multiplanar images after passing through a 4f system and reflecting off a mirror used for path folding (Thorlabs ME1-P01). From the SLM to the path folding mirror, we used the following optical components: an achromatic lens (Thorlabs AC254-100A-ML), an aperture located at the Fourier plane (Thorlabs SM1D12D), and another achromatic lens (Thorlabs AC254-050-A-ML).

We project the reconstructed multiplane image to an eyepiece using a 2f system consisting of Thorlabs AC254-150-A-ML. Although there is an optional galvanometer scanner (AT20-2278) available, we did not utilize it in this work. At a meter throw distance, our user perceives multiplane images by looking through our eyepiece composed of two lens cascades. From the projection toward the user, these lenses in our eyepiece are Thorlabs LA1384 and Thorlabs LA1050. For more details, please consult our supplementary.

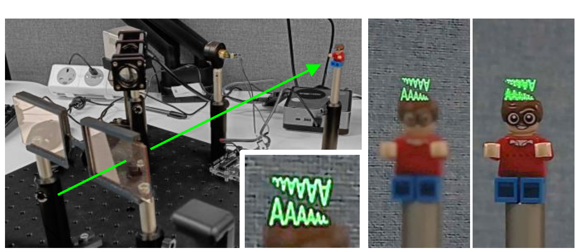

Figure 7: An actual photograph of a see-through image from our amplitude-only *HoloBeam* prototype. The insets demonstrate how the generated virtual image is focused at a far plane (1.5 meters away).

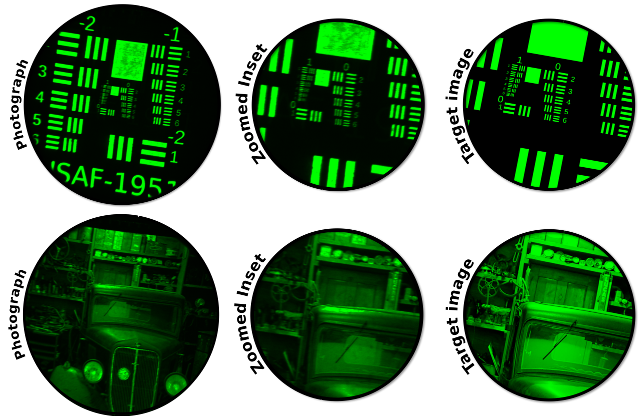

Figure 8: Captured photographs from the eyebox of our phase-only *HoloBeam* prototype. These photographs are captured using a XIMEA MC245MG-SY-UB image sensor equipped with an adjustable $5-50$ mm lens while using $20ms$ exposure times. During these captures, only a green LED light source is used in the display prototype. The figure also shows zoomed-insets and their target images. The provided images are circular as the aperture of our eyepiece in this prototype is circular. The source image on the right side is from DIV2K [2].

4.3 *HoloBeam* Software

We choose to dedicate our amplitude-only prototype to showcasing thin eyepieces. Thus, this particular prototype relies on projecting conventional 2D images to deliver its message related to form factor (not amplitude-only holograms). In the meantime, we dedicate our phase-only prototype to delivering 3D images while providing a wide FoV and large throw distances. We choose not to use HOE in the phase-only prototype to avoid the heavy engineering work that poses an engineering resource challenge (e.g., recording HOE in multicolor, replicating optics of the display for the reference beam of a recording setup).

Our learned CGH pipeline is simply dedicated to our phase-only prototype, and it bases on PyTorch [65] and a CGH toolkit [9, 43, 46]. Training of our learned CGH pipeline is conducted using a learning rate of 0.0001 for ten epochs (Source code: **GitHub:complight/multiholo** [7]). Our model relied on a U-Net [69] with 28 hidden channels as our neural network. To generate our dataset for training our learned CGH pipeline, we first create depth maps for the DIV2K image dataset [3] following the work by Alhashim and Wonka et al. [10,11]. We resize these images and their depth to the resolution of our SLM. Using their estimated depth, we use these images to generate multiplane holograms with six planes following the work by Akşit et al. [8,44] (plane separation is 1 mm). We discard the estimated depth, and we use 900 input images and their corresponding multiplane holograms in our training while 100 of them are in validation. As the training complete, our estimation routine takes about 20-28 ms to generate a single phase-only hologram at the resolution of our SLM. We use a computer with NVIDIA GeForce RTX 2080 GPU with 12 GB memory and an Intel i7, 3.9 GHz CPU to drive our holographic display prototype. When we display our holograms in our phase-only prototype, we update the calculated O_h with a linear phase grating term to avoid undiffracted light,

$$O'_h(x,y) = \begin{cases} e^{-j(\phi(x,y)+\pi)} & \text{if } y = \text{odd} \\ e^{-j\phi(x,y)} & \text{if } y = \text{even}, \end{cases} \tag{9}$$

where ϕ, x, y represents the original phase of O_h. In Listing 1, we provide a simplified learned estimation model that used in our training and estimation routines. Note that the estimation routine relies on a CNN. Specifically, we use a U-Net as our CNN [69]. This CNN takes a single color of an image as input. The output of the CNN is a tensor with two channels. The resultant constrained tensor

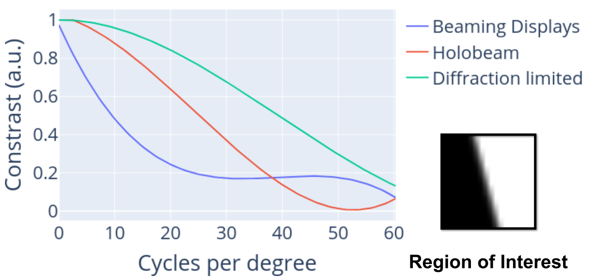

Figure 9: Modulation Transfer Function Analysis for phase-only *HoloBeam* prototype with LED illumination. Beaming Displays [39] provided 8 cpd at their central FoV. Phase-only *HoloBeam* prototype supports 24 cpd at half contrast in its central FoV. Thus, *HoloBeam* improves the resolution three folds while expanding the FoV with respect to Beaming Displays [39].

represents the phase component of a phase-only hologram. During a training session, the output of this estimation model is compared against a ground-truth hologram using an L2 loss function. Unlike the recent literature [23], no reconstruction losses are involved.

5 Evaluation

Using our two *HoloBeam* prototypes, we assess the practical limits of our approach. For analyzing the optical qualities of our *HoloBeam* approach, we use our phase-only prototype, whereas for the demonstration of a slim eyepiece built using an in-house recorded HOE, we will use our amplitude-only prototype.

5.1 Quality Analysis

In this section, we will rely on optimized phase-only holograms for assessing absolute resolution and FoV characteristics of our proposed method. For demonstrating, 3D images from our prototype, we will rely on both our learned method and optimization method.

Resolution and Field of View. The eye relief of our phase-only prototype is 35 mm. The aperture of our eyepiece is 50.8 mm, leading to FoV of 70 degrees as depicted in Fig. 1. We also provide additional results from various scenes in Fig. 8. To assess the resolution quality of this prototype, we rely on a standard Modulation Transfer Function (MTF) analysis [17]. Across our evaluations, we capture photographs with a XIMEA MC245MG-SY-UB image sensor and an adjustable $5-50$ mm lens while using 20 ms exposure times to approximate a human observer's experience. Our MTF analysis suggests that the phase-only *HoloBeam* prototype can support up to 24 cpd in resolution when used with LED illuminations (at central FoV). A healthy Human Visual System (HVS) demands 30 cpd or more for realistic-looking resolutions. Such resolutions could be met as we use the existing lasers in our phase-only *HoloBeam* prototype. However, we observe that perturbating fringe patterns shadow the image quality as depicted with an extra capture at our supplementary. We believe these perturbations mostly originated from optics starting from the end of 4f lenses towards the eyepiece. *HoloBeam* could potentially provide higher resolutions while using LEDs with the increasing aperture size of SLM and lenses in the future.

Multiplane Images. Light source coherency used in a holographic display dictates the depth of field of the reconstructed images [58]. Thus, incoherent broadband sources increase the depth of field and degrade optical focus cues in reconstructed images. Given the situation with the depth of field of images, the conventional way to demonstrate multiple images in a holographic display involves using a laser light source. We also use laser in our phase-only *HoloBeam* prototype to demonstrate capabilities related to multiplane images.

However, as the lenses between our 4f system and eyepiece generate unintended aberrations and distortions, our final image contains fringes when lasers are used (see our supplementary document for evidence). These distortions could be fixed in the future using learned approaches that could account for imperfections in holographic display hardware [20, 23, 45] and by designing and manufacturing dedicated projection optics like in many projector products. Thus, to avoid any visual artifacts caused by the eyepiece and additional lenses, we capture reconstructions of our phase-only holograms right after the 4f imaging system using a bare XIMEA MC245MG-SY-UB image sensor with $50ms$ exposure. We provide a sample capture demonstrating a focus change as in Fig. 10. Note that the optimized version shown in Fig. 10 resembles the image quality of most recent standard literature [23, 44, 73]. For more results, please consult our supplementary materials.

In our observation, the learned model shrinks the depth range of a scene (e.g., six multiplane images mostly focused at two planes) and tends to distribute depth levels that are not faithful to an original depth map (e.g., windmills door becoming sharp at near focus rather than far focus in Fig. 10). We believe this learned model promises encouraging first results towards a hologram generation routine where a 2D image is transformed into a 3D multiplanar hologram without requiring scene-depth information or multiple perspective images.

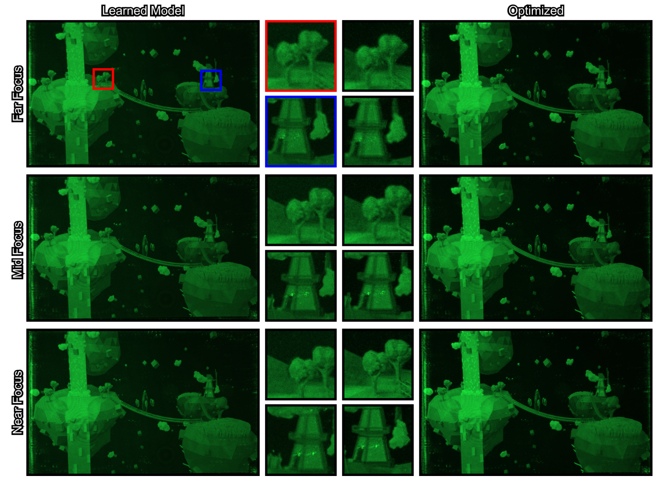

Figure 10: Learned Model for Multiplane Holograms. Our learned hologram generation method can help reconstruct high-quality images with depth support from conventional images. However, the generated hologram's depth is not necessarily faithful to the original depth as compared with the optimized hologram. We observe that the learned model tends to shrink the depth range (e.g., six multiplane images in this example are mostly focused on one to three planes). Note that the optimized case resembles a typical outcome from recent standard literature [23, 44, 73]. For more results, consult our supplementary.

5.2 HOE Lens Analysis

In *HoloBeam*, the spatial relationships among the projection system, the eyepiece, and the eye affect the final image quality and visibility. This section analyzes our HOE design with simulations that mimics our amplitude-only prototype.

Simulation Setup In our simulation on MATLAB 2022a, we use an open-source ray optics library. Since this library does not implement HOE simulations, we implemented a volume hologram class based on Sec. 3.2. For all simulations, we set the wavelength to $\lambda = 532$ nm, the average refraction index to $n_0 = 1.5$, the maximum refractive index modulation to $n_1 = 0.04$, and the photopolymer thickness $= 30\mu$m, respectively.

Figure 11 shows the layout of our simulation. We used the library's default eye optics parameters for the eye model, and the pupil diameter was set to 3 mm. The eyeball center is set to 30 mm from the HOE lens. We assume that the projection optics relays a

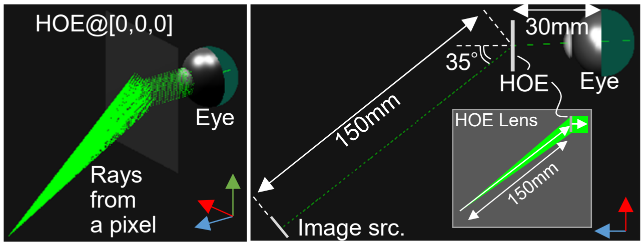

Figure 11: Visualization of our simulation in Sec. 5.2. The configuration shows the default layout of each component. We used left-handed coordinates where the z-axis faces away from the viewpoint.

virtual image from the amplitude-only SLM and the virtual image forms at 150 mm from the HOE with a pan angle of 35 degrees. The HOE is designed to form a lens with f=150 mm and a 35-degree angle tilt. Note that we scaled the virtual image size by 3, assuming that the projection optics relay the SLM image plane, leading to the pixel pitch of the relayed image as $p = 5.4 * 3 = 16.2\mu$m. This results in the rays' diffraction angle from each pixel becoming $2\,\mathrm{asin}(1.22\lambda/p)$, about 4.6 degrees. We randomly generated 100 rays for each pixel along the central ray direction within the diffraction angle.

We also assume an ideal projection system that can guide images to the default pupil center position. We thus let the central rays target the center of the HOE, i.e. (0,0,0) in the world coordinates. We refer to this positional relationship as the default layout. As a displayed image, we set 17x17-point grids uniformly spanning in an 801x801 pixel image (Fig. 12 left column top).

We evaluate two major factors related to misalignments: the eyebox and head alignment. The eyebox analysis evaluates when the eye locations change while the virtual image source and the eyepiece are fixed. The head alignment analysis evaluates when the head system, i.e., the HOE eyepiece and the eye, changes its orientations and positions against the virtual image source.

We set the disturbance ranges of each setup in the following. For the eyebox position: ± 4 mm with 0.25 mm step in the x-y directions; for the head orientation: ± 10 degrees with 0.5-degree step for the pan-tilt angles; and for the head translation: ± 4 mm with 0.25 mm step in the x-y directions.

For the eyebox and head translation analyses, one may also explore the z direction. We, however, fixed the z-axis parameter in our analyses to keep the explanations concise since our pilot analysis results did not give much difference in the z-axis analysis compared to the x-y space analyses.

Simulation Results Figure 12 shows an overview of our simulations, where the second to fourth rows correspond to eyebox analysis (x-y plane), head orientation analysis (pan-tilt angles), and head translation analysis (x-y plane), respectively. The second column shows the total brightness of the observed images for the given misalignment parameters. The total brightness is relative to that of the image taken with the default layout, i.e., the center pixel of each figure. The third column is a color visualization of whether or not the light rays hit the retina at each viewpoint. The images accumulated hit from every 10 viewpoints over the misalignment parameter space. And the rays from the same viewpoint will be the same color regardless of their intensity. The fourth columns instead accumulate ray brightness and show colormaps of the total brightness of all given view conditions. Please also refer to our supplementary viewpoint videos of the three analyses.

The eyebox analysis shows that the brightness level is consistent over eye position changes. This is understandable given the diffraction angle of the virtual image is about 4.6 degrees, which sufficiently covers the pupil, as we can also observe in Fig. 11.

The head orientation analysis shows that our setup is more robust in the pan (horizontal) rotation than in the tilt (vertical) orientation. A possible reason for this tendency is that the HOE lens is designed for input rays' virtual point source to be placed on the x-z plane; thus, the diffraction efficiency could radically decrease (off-Bragg) along vertical angle errors.

Finally, the head translation analysis shows that our setup is more robust in head translation errors than head orientation errors. However, the image intensities change periodically along vertical head misalignment. The image thus may appear to flicker in dynamic tracking environments, depending on the tracking accuracy. This tendency may stem from the periodical structure of the Bragg diffraction efficiency over input angles, as in Fig. 4. Our informal, practical observation with the amplitude-only prototype aligns with the simulation results.

6 Discussion and Limitations

Beaming Displays [39] design methodology aims to resolve a vital issue in AR glasses designs and enable practical AR applications that could converge to comfortable and realistic AR experiences. Although *HoloBeam* provided evidence that Beaming Display designs could provide resolutions and FoV meeting more of HVS demands, there are still outstanding major issues related to Beaming Displays design methodology. We list these issues and their potential solutions in the following paragraphs.

Eyebox and cost Holographic approaches imply small eyebox traditionally, and also true for our implementation. Most recently, the work by Jang et al. [40] addresses the eyebox issue while maintaining a true 3D holographic AR glasses, which could help inspire a fix of the issue. Holographic displays lead to a complex hardware design that is highly costly due to niche equipment such as a phase-only SLM [85]. Our amplitude-only prototype aims to improve on cost of holographic displays by relying on cheaper but less capable hardware. However, this is not a true fix to the cost problem as cheaper alternatives for a phase-only SLM have to be invented.

Improving mobility. Beaming Displays [39] provided evidence that delivering images to a moving user could be possible to some extent, however, *HoloBeam* does not provide any tracking of users and the same case is not replicated as *HoloBeam* focused on improving image quality and form factor. The original Beaming Displays work tends to relay an image on a screen integrated into an eyepiece, whereas *HoloBeam* is a standalone optical relay that magnifies images projected in mid-air. Thus, *HoloBeam* is much less forgiving of any misalignment in the beaming path. We believe this requires special attention in future works and the functionality of a diffuser in the original Beaming Displays has to be embedded into future variants *HoloBeam* to benefit from both mobility and image quality improvements.

If the functionality of a diffuser could be transferred to a new eyepiece design, we trust that mobility-related issues (e.g., head position and orientation of a user) could be resolved without sacrificing the form factor. This item is our next step in this line of research.

Improving the eyepiece. The image brightness provided by a HOE eyepiece is sensitive to misalignments due to users' head orientations. We believe there are potential research directions that could help mitigate these alignment-related issues. A potential solution could be to investigate embedding the trends in curved or free-form HOE designs [14, 42] to our HOE designs in the future. Another potential research direction is to bring the properties of bird-bath optics used in the original Beaming Displays [39] into the HOE design space. In such a case, we could explore the embedding properties of a diffuser with a lens in HOE designs [37, 81, 82].

***HoloBeam*'s target applications.** Given the limitations in tracking and alignment for *HoloBeam*, we identify applications that could make the most sense for what is possible with *HoloBeam*

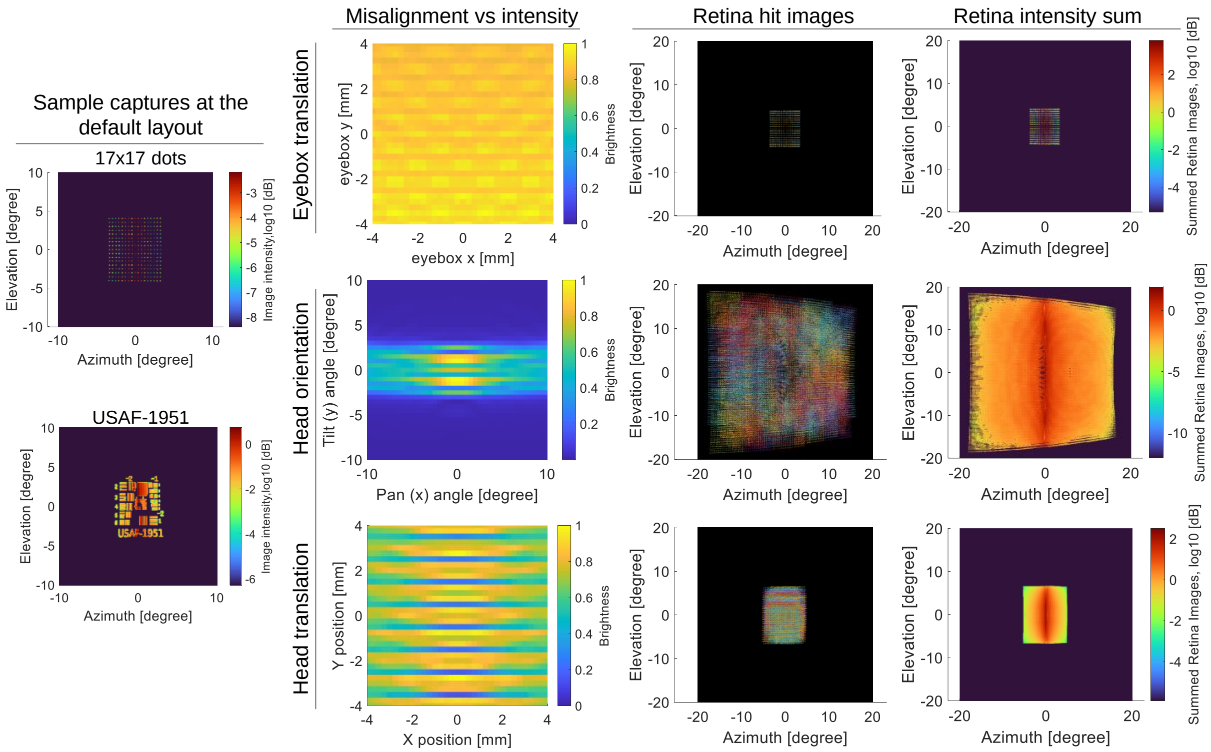

Figure 12: A summary of the simulation analysis in Sec. 5.2. The first column is simulated retinal views on the default layout when displaying a 17x17 point grid used in the analysis and a USAF1951 chart as another example. For the column between the second and fourth, the three rows for the eyebox, the head orientation, and the head translation analysis, respectively. The three columns visualize the relative image intensity over misalignment parameter spaces, visualization of the retinal hit count of the rays at sampled viewpoints, and of accumulated intensities.

today. Applications that do not require a user to be frequently moving could be more forgiving with the lack of tracking. Thus, we believe *HoloBeam* could be a good fit for work-related applications (e.g., desktop displays). There are existing 3D desktop displays for similar purposes (e.g., Brelyon[1]), which we believe *HoloBeam* could provide a new slim version. A potential application area in the future could also be automotive displays, namely heads-up displays. Note that a driver in a car is more or less stationary and must constantly gaze at the road. A future variant of *HoloBeam* could provide more freedom in movement and could be a potential new way to build a heads-up display that does not require fiddling with the windshield or dashboard of a car.

Improving Depth Generation of Learned CGH method. The depth generation of our learned holograms is limited as the depth information of a scene is not provided as an input to our learned algorithm. We designed our learned algorithm as a single constrained U-Net. However, a depth estimation network could also be used to condition our method or a more detailed analysis could be conducted in the future regarding the contribution of each layer in that single U-Net. This way, an educated way to improve our attempts could be achieved in the future.

Complementing contact lens AR. In recent years, there has been a push in research related to AR contact lenses [70], which could potentially transform AR to a seamless setting without any glasses. In the future, our work could also be transformed into an optical component embedded in a contact lens that helps register beams from a projector to a user's retina.

7 Conclusion

Future AR glasses must be thin and lightweight while supporting high resolutions, wide FoV, and optical focus cues. Today's AR glasses struggle to balance these requirements, often yielding to design challenges. However, the experts largely agree that these requirements are compulsory to maintain comfortable visual experiences with AR glasses. So, we ask whether the way out of this struggle could lie in changing the perspectives in AR glasses design.

This paper introduces a new stand leading to a bold change in AR glasses design which we named *HoloBeam*. For the first time in the literature, we demonstrate very slim AR glasses based on our *HoloBeam* method. These AR glasses support 3D images with near-correct optical focus cues and provide resolutions matching commonly accepted retinal resolution criteria (30 cpd). In this design, there are still open research issues that require proper attention from the relevant research communities. These issues include tracking the accuracy of users beaming accuracy of projectors and designs, leading to more freedom in user movement. Although these outstanding issues exist, this work is a significant milestone that could pave the road to a new body of work toward the ultimate AR glasses with comfortable and realistic visual experiences.

Supplementary Material

Our code is available at **GitHub:complight/multiholo**. Dependencies are available at **GitHub:kaanaksit/odak** [43, 46]. Additional media, results and materials are distributed separately as a supplementary documents.

Acknowledgments

The authors wish to thank Koray Kavaklı for fruitful discussions. Kaan Akşit is supported by the Royal Society's RGS\R2\212229 - Research Grants 2021 Round 2 in building the hardware prototype and Meta Reality Labs inclusive rendering initiative for building the rendering pipeline. Yuta Itoh is supported by JST FOREST Grant Number JPMJFR206E and JSPS KAKENHI Grant Number JP20J14971, 20H05958, and 21K19788, Japan.

[1]https://brelyon.com/

REFERENCES

[1] A. Aghasi, B. Heshmat, L. Wei, and M. Tian. Optimal allocation of quantized human eye depth perception for multi-focal 3d display design. *Optics Express*, 29(7):9878–9896, 2021.

[2] E. Agustsson and R. Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 126–135, 2017.

[3] E. Agustsson and R. Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.

[4] K. Akşit, P. Chakravarthula, K. Rathinavel, Y. Jeong, R. Albert, H. Fuchs, and D. Luebke. Manufacturing application-driven foveated near-eye displays. *IEEE transactions on visualization and computer graphics*, 25(5):1928–1939, 2019.

[5] K. Akşit, J. Kautz, and D. Luebke. Slim near-eye display using pinhole aperture arrays. *Applied optics*, 54(11):3422–3427, 2015.

[6] K. Akşit, W. Lopes, J. Kim, P. Shirley, and D. Luebke. Near-eye varifocal augmented reality display using see-through screens. *ACM Transactions on Graphics (TOG)*, 36(6):1–13, 2017.

[7] K. Akşit. complight/multiholo. `https://github.com/complight/multiholo`, 2022.

[8] K. Akşit. complight/realistic_defocus. `https://github.com/complight/realistic_defocus`, 2022.

[9] K. Akşit, A. S. Karadeniz, P. Chakravarthula, W. Yujie, K. Kavaklı, Y. Itoh, D. R. Walton, and R. Zhu. Odak, May 2022. If you use this software, please cite it as below. doi: 10.5281/zenodo.6528486

[10] I. Alhashim. ialhashim/densedepth. `https://github.com/ialhashim/DenseDepth/`, 2022.

[11] I. Alhashim and P. Wonka. High quality monocular depth estimation via transfer learning. *arXiv preprint arXiv:1812.11941*, 2018.

[12] T. Ando, T. Matsumoto, H. Takahasihi, and E. Shimizu. Head mounted display for mixed reality using holographic optical elements. *MEMOIRS-FACULTY OF ENGINEERING OSAKA CITY UNIVERSITY*, 40:1–6, 1999.

[13] S.-H. Baek, E. Tseng, A. Maimone, N. Matsuda, G. Kuo, Q. Fu, W. Heidrich, D. Lanman, and F. Heide. Neural etendue expander for ultra-wide-angle high-fidelity holographic display. *arXiv preprint arXiv:2109.08123*, 2021.

[14] K. Bang, C. Jang, and B. Lee. Curved holographic optical elements and applications for curved see-through displays. *Journal of Information Display*, 20(1):9–23, 2019.

[15] A. Bauer and J. P. Rolland. Two all reflective, freeform, optical see-through head-worn displays. In *International Optical Design Conference*, pp. ITh3A–6. Optica Publishing Group, 2014.

[16] B. Brown and A. Lohmann. Computer-generated binary holograms. *IBM Journal of research and Development*, 13(2):160–168, 1969.

[17] P. D. Burns et al. Slanted-edge mtf for digital camera and scanner analysis. In *Is and Ts Pics Conference*, pp. 135–138. Society for Imaging Science & Technology, 2000.

[18] O. Cakmakci, Y. Ha, and J. P. Rolland. A compact optical see-through head-worn display with occlusion support. In *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 16–25. IEEE, 2004.

[19] A. Cameron. The application of holographic optical waveguide technology to the q-sight family of helmet-mounted displays. In *Head-and Helmet-Mounted Displays XIV: Design and Applications*, vol. 7326, pp. 150–160. SPIE, 2009.

[20] P. Chakravarthula, E. Tseng, T. Srivastava, H. Fuchs, and F. Heide. Learned hardware-in-the-loop phase retrieval for holographic near-eye displays. *ACM Transactions on Graphics (TOG)*, 39(6):1–18, 2020.

[21] C. Chang, K. Bang, G. Wetzstein, B. Lee, and L. Gao. Toward the next-generation vr/ar optics: a review of holographic near-eye displays from a human-centric perspective. *Optica*, 7(11):1563–1578, 2020.

[22] R. H.-Y. Chen and T. D. Wilkinson. Computer generated hologram from point cloud using graphics processor. *Applied optics*, 48(36):6841–6850, 2009.

[23] S. Choi, M. Gopakumar, Y. Peng, J. Kim, and G. Wetzstein. Neural 3d holography: Learning accurate wave propagation models for 3d holographic virtual and augmented reality displays. *ACM Transactions on Graphics (TOG)*, 40(6):1–12, 2021.

[24] D. Close. Holographic optical elements. *Optical Engineering*, 14(5):408–419, 1975.

[25] W. Cui and L. Gao. Optical mapping near-eye three-dimensional display with correct focus cues. *Optics Letters*, 42(13):2475–2478, 2017.

[26] V. R. Curtis, N. W. Caira, J. Xu, A. G. Sata, and N. C. Pégard. Dcgh: Dynamic computer generated holography for speckle-free, high fidelity 3d displays. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 1–9. IEEE, 2021.

[27] D. Dunn, C. Tippets, K. Torell, P. Kellnhofer, K. Akşit, P. Didyk, K. Myszkowski, D. Luebke, and H. Fuchs. Wide field of view varifocal near-eye display using see-through deformable membrane mirrors. *IEEE transactions on visualization and computer graphics*, 23(4):1322–1331, 2017.

[28] C. Ebner, S. Mori, P. Mohr, Y. Peng, D. Schmalstieg, G. Wetzstein, and D. Kalkofen. Video see-through mixed reality with focus cues. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2256–2266, 2022.

[29] M. Fally, J. Klepp, and Y. Tomita. An experimental study on the validity of diffraction theories for off-bragg replay of volume holographic gratings. *Applied Physics B*, 108(1):89–96, 2012.

[30] S. Hamann, L. Shi, O. Solgaard, and G. Wetzstein. Time-multiplexed light field synthesis via factored wigner distribution function. *Optics letters*, 43(3):599–602, 2018.

[31] T. Hamasaki and Y. Itoh. Varifocal occlusion for optical see-through head-mounted displays using a slide occlusion mask. *IEEE transactions on visualization and computer graphics*, 25(5):1961–1969, 2019.

[32] M. K. Hedili, B. Soner, E. Ulusoy, and H. Urey. Light-efficient augmented reality display with steerable eyebox. *Optics Express*, 27(9):12572–12581, 2019.

[33] Y. Hiroi, T. Kaminokado, S. Ono, and Y. Itoh. Focal surface occlusion. *Optics Express*, 29(22):36581–36597, 2021.

[34] C.-K. Hsueh and A. A. Sawchuk. Computer-generated double-phase holograms. *Applied optics*, 17(24):3874–3883, 1978.

[35] X. Hu and H. Hua. High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics. *Optics express*, 22(11):13896–13903, 2014.

[36] F.-C. Huang, D. P. Luebke, and G. Wetzstein. The light field stereoscope. In *SIGGRAPH emerging technologies*, pp. 24–1, 2015.

[37] L. Hwang, G. Hur, J. Kim, P. Gentet, S. Kwon, and S. Lee. Uniformity improvement of a reconstructed-holographic image in a near-eye display system using off-axis hoe. *Optics Express*, 30(12):21439–21454, 2022.

[38] D. Im, E. Moon, Y. Park, D. Lee, J. Hahn, and H. Kim. Phase-regularized polygon computer-generated holograms. *Optics Letters*, 39(12):3642–3645, 2014.

[39] Y. Itoh, T. Kaminokado, and K. Akşit. Beaming displays. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2659–2668, 2021.

[40] C. Jang, K. Bang, M. Chae, B. Lee, and D. Lanman. Waveguide holography: Towards true 3d holographic glasses. *arXiv preprint arXiv:2211.02784*, 2022.

[41] C. Jang, K. Bang, G. Li, and B. Lee. Holographic near-eye display with expanded eye-box. *ACM Transactions on Graphics (TOG)*, 37(6):1–14, 2018.

[42] C. Jang, O. Mercier, K. Bang, G. Li, Y. Zhao, and D. Lanman. Design and fabrication of freeform holographic optical elements. *ACM Transactions on Graphics (TOG)*, 39(6):1–15, 2020.

[43] K. Kavaklı and K. Akşit. Introduction to odak: a differentiable toolkit for optical sciences, vision sciences and computer graphics. In *Frontiers in Optics*, pp. FTu1A–1. Optica Publishing Group, 2022.

[44] K. Kavaklı, Y. Itoh, H. Urey, and K. Akşit. Realistic defocus blur for multiplane computer-generated holography. *arXiv preprint arXiv:2205.07030*, 2022. doi: 10.48550/arXiv.2205.07030

[45] K. Kavaklı, H. Urey, and K. Akşit. Learned holographic light transport. *Applied Optics*, 61(5):B50–B55, 2022.

[46] K. Kavakli, D. R. Walton, N. Antipa, R. Mantiuk, D. Lanman, and

K. Akşit. Optimizing vision and visuals: lectures on cameras, displays and perception. In *ACM SIGGRAPH 2022 Courses*, pp. 1–66. 2022.

[47] J. Kim, M. Gopakumar, S. Choi, Y. Peng, W. Lopes, and G. Wetzstein. Holographic glasses for virtual reality. In *Proceedings of the ACM SIGGRAPH*, pp. 1—-9, 2022.

[48] J. Kim, Y. Jeong, M. Stengel, K. Aksit, R. A. Albert, B. Boudaoud, T. Greer, J. Kim, W. Lopes, Z. Majercik, et al. Foveated ar: dynamically-foveated augmented reality display. *ACM Trans. Graph.*, 38(4):99–1, 2019.

[49] N. Kim, Y.-L. Piao, and H.-Y. Wu. Holographic optical elements and application. *Holographic Materials and Optical Systems*, 5:99–131, 2017.

[50] H. Kogelnik. Coupled wave theory for thick hologram gratings. *Bell System Technical Journal*, 48(9):2909–2947, 1969.

[51] R. Konrad, N. Padmanaban, K. Molner, E. A. Cooper, and G. Wetzstein. Accommodation-invariant computational near-eye displays. *ACM Transactions on Graphics (TOG)*, 36(4):1–12, 2017.

[52] G. A. Koulieris, K. Akşit, M. Stengel, R. K. Mantiuk, K. Mania, and C. Richardt. Near-eye display and tracking technologies for virtual and augmented reality. *Computer Graphics Forum*, 38(2):493–519, 2019. doi: 10.1111/cgf.13654

[53] T. Kozacki. On resolution and viewing of holographic image generated by 3d holographic display. *Optics Express*, 18(26):27118–27129, 2010.

[54] G. Kuo, L. Waller, R. Ng, and A. Maimone. High resolution étendue expansion for holographic displays. *ACM Transactions on Graphics (TOG)*, 39(4):66–1, 2020.

[55] D. Lanman and D. Luebke. Near-eye light field displays. *ACM Transactions on Graphics (TOG)*, 32(6):1–10, 2013.

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

[57] S. Lee, Y. Jo, D. Yoo, J. Cho, D. Lee, and B. Lee. Tomographic near-eye displays. *Nature communications*, 10(1):1–10, 2019.

[58] S. Lee, D. Kim, S.-W. Nam, B. Lee, J. Cho, and B. Lee. Light source optimization for partially coherent holographic displays with consideration of speckle contrast, resolution, and depth of field. *Scientific reports*, 10(1):1–12, 2020.

[59] A. Maimone, A. Georgiou, and J. S. Kollin. Holographic near-eye displays for virtual and augmented reality. *ACM Transactions on Graphics (Tog)*, 36(4):1–16, 2017.

[60] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs. Pinlight displays: Wide field of view augmented reality eyeglasses using defocused point light sources. In *ACM SIGGRAPH 2014 Emerging Technologies*, SIGGRAPH '14. Association for Computing Machinery, New York, NY, USA, 2014. doi: 10.1145/2614066.2614080

[61] N. Matsuda, A. Fix, and D. Lanman. Focal surface displays. *ACM Transactions on Graphics (TOG)*, 36(4):1–14, 2017.

[62] R. Narain, R. A. Albert, A. Bulbul, G. J. Ward, M. S. Banks, and J. F. O'Brien. Optimal presentation of imagery with focus cues on multi-plane displays. *ACM Transactions on Graphics (TOG)*, 34(4):1–12, 2015.

[63] D. K. Nikolov, A. Bauer, F. Cheng, H. Kato, A. N. Vamivakas, and J. P. Rolland. Metaform optics: Bridging nanophotonics and freeform optics. *Science Advances*, 7(18):eabe5112, 2021.

[64] J. Orlosky, M. Sra, K. Bektaş, H. Peng, J. Kim, N. Kos'myna, T. Höllerer, A. Steed, K. Kiyokawa, and K. Akşit. Telelife: The future of remote living. frontiers in virtual reality 2 (2021), 147. *URL: https://www. frontiersin. org/article/10.3389/frvir*, 2021.

[65] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in pytorch. In *NIPS 2017 Workshop on Autodiff*, 2017.

[66] D. Pi, J. Liu, and Y. Wang. Review of computer-generated hologram algorithms for color dynamic holographic three-dimensional display. *Light: Science & Applications*, 11(1):1–17, 2022.

[67] M. Prijatelj, J. Klepp, Y. Tomita, and M. Fally. Far-off-bragg reconstruction of volume holographic gratings: A comparison of experiment and theories. *Physical Review A*, 87(6):063810, 2013.

[68] J. Ratcliff, A. Supikov, S. Alfaro, and R. Azuma. Thinvr: Heterogeneous microlens arrays for compact, 180 degree fov vr near-eye displays. *IEEE transactions on visualization and computer graphics*, 26(5):1981–1990, 2020.

[69] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer, 2015.

[70] J. Sano and Y. Takaki. Holographic contact lens display that provides focusable images for eyes. *Optics Express*, 29(7):10568–10579, 2021.

[71] L. Shi, F.-C. Huang, W. Lopes, W. Matusik, and D. Luebke. Near-eye light field holographic rendering with spherical waves for wide field of view interactive 3d computer graphics. *ACM Transactions on Graphics (TOG)*, 36(6):1–17, 2017.

[72] L. Shi, B. Li, C. Kim, P. Kellnhofer, and W. Matusik. Towards real-time photorealistic 3d holography with deep neural networks. *Nature*, 591(7849):234–239, 2021.

[73] L. Shi, B. Li, and W. Matusik. End-to-end learning of 3d phase-only holograms for holographic display. *Light: Science & Applications*, 11(1):1–18, 2022.

[74] W. Song, Q. Cheng, P. Surman, Y. Liu, Y. Zheng, Z. Lin, and Y. Wang. Design of a light-field near-eye display using random pinholes. *Optics Express*, 27(17):23763–23774, 2019.

[75] W. C. Sweatt. Describing holographic optical elements as lenses. *JOSA*, 67(6):803–808, 1977.

[76] A. Symeonidou, D. Blinder, and P. Schelkens. Colour computer-generated holography for point clouds utilizing the phong illumination model. *Optics express*, 26(8):10282–10298, 2018.

[77] M. Sypek. Light propagation in the fresnel region. new numerical approach. *Optics communications*, 116(1-3):43–48, 1995.

[78] G. Tan, T. Zhan, Y.-H. Lee, J. Xiong, and S.-T. Wu. Polarization-multiplexed multiplane display. *Optics letters*, 43(22):5651–5654, 2018.

[79] J. Xiong, K. Yin, K. Li, and S.-T. Wu. Holographic optical elements for augmented reality: principles, present status, and future perspectives. *Advanced Photonics Research*, 2(1):2000049, 2021.

[80] T. Yatagai. Three-dimensional displays using computer-generated holograms. *Optics Communications*, 12(1):43–45, 1974.

[81] J. Yeom, J. Hong, and J. Jeong. Projection-type see-through near-to-eye display with a passively enlarged eye-box by combining a holographic lens and diffuser. *Optics Express*, 29(22):36005–36020, 2021.

[82] J. Yeom, J. Jeong, J. Hong, and K.-s. Choi. Analysis on image quality of a holographic lens with a non-converging signal wave for compact near-eye displays. *Optics Express*, 30(20):36632–36643, 2022.

[83] T. Zhan, Y.-H. Lee, and S.-T. Wu. High-resolution additive light field near-eye display by switchable pancharatnam–berry phase lenses. *Optics express*, 26(4):4863–4872, 2018.

[84] J. Zhang, N. Pégard, J. Zhong, H. Adesnik, and L. Waller. 3d computer-generated holography by non-convex optimization. *Optica*, 4(10):1306–1313, 2017.

[85] Z. Zhang, Z. You, and D. Chu. Fundamentals of phase-only liquid crystal on silicon (lcos) devices. *Light: Science & Applications*, 3(10):e213–e213, 2014.

Supplementary Materials for "HoloBeam: Paper-Thin Near-Eye Displays"

Kaan Akşit* — University College London

Yuta Itoh† — The University of Tokyo

1 Contributions from each author

Here we list contributions from each author in this specific research work:

Yuta Itoh:

- Designing and building a HOE recording setup in hardware.
- Compiling extensive study on HOE design space.
- Compiling figures for HOE recording setup and functional test of recorded HOEs.
- Review and edits for the entire documentation.

Kaan Akşit:

- Initiation of project idea of *HoloBeam*, this includes blueprints for projector designs, conventional eyepiece designs and algoritmic designs.
- Design and manufacturing of amplitude-only and phase-only *HoloBeam* prototypes,
- Coding and training of the machine learning model.
- Writing abstract, introduction, related work, conclusion sections entirely.
- Compiling figures for a simplified layout diagram of *HoloBeam*, teaser, phase-only and amplitude-only hardware in the manuscript and supplementary documentation.
- Review and edits for the entire documentation.

2 Additional related works

2.1 Thin Relay Optics for Augmented Reality Glasses

Holographic Optical Elements There are various photosensitive materials for holograms [14]. For example, Silver halide emulsions are used for amplitude modulating Holographic Optical Element (HOE)s. Yet another example, Photorefractive crystals (dichromate gelatin or lithium niobate) and photopolymers are organic materials used for phase modulating HOEs. Among these materials, photopolymers are relatively easy to handle because they do not require the development process required for other materials and can be bleached by incoherent light. Attracting attention as a replacement for HOEs, metasurfaces are artificial sub-wavelength subwavelength structures promising superior optical performance, for example, large aperture, wide bandwidth, and polarization dependence [9, 10]. However, these systems are in their early stages in proving their practicality in Augmented Reality (AR) glasses.

*e-mail: k.aksit@ucl.ac.uk

†e-mail: yuta.itoh@iii.u-tokyo.ac.jp

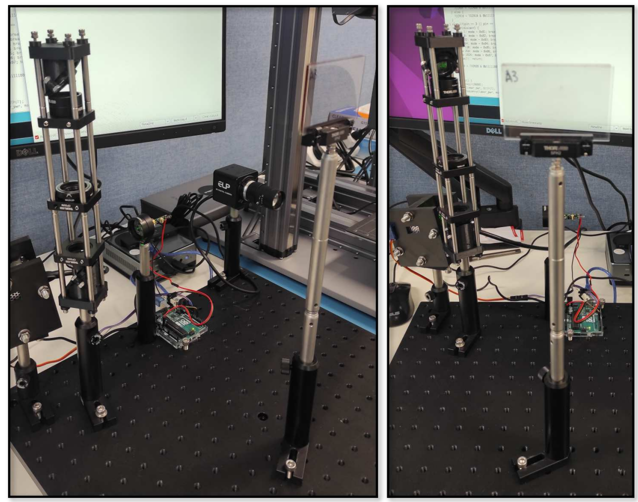

Figure 1: Vertically built amplitude-only *HoloBeam* prototype. Our projection assembly's optical path is folded using two mirrors in this configuration.

2.2 Augmented Reality Glasses

Beaming AR Glasses. Most recently, a new design method called Beaming Displays [6] is proposed, where active components are separated from AR glasses designs to address issues in conventional AR glasses providing a fixed focus image. In their work, a user wears bird-bath optics similar to the ones in the literature, and the image is projected into this bird-bath optics from a distance. However, their results [6] demonstrate a moderate Field Of View (FoV), limited resolutions, and bulky form factor with no support in optical focus cues. Our work follows the design approach in Beaming Displays in AR glasses. However, *HoloBeam* merges the benefits of holographic techniques with Beaming Displays. This way, *HoloBeam* improves upon Beaming Displays by supporting optical focus cues, improved resolutions and unmatched form-factors.

Other AR apporaches. We do not cover all other AR displays approaches in the literature that may have a loose relevance to our work [3]. Both previous literature [6] and *HoloBeam* share similarities with projection based spatial AR displays such as head-up displays in automative [4]. The primary difference in our approach with projection based spatial AR displays is that the eyepiece is close to a user's eye. Although there are AR displays having eyepiece close to a user's eye [8], unlike theirs our solution do not require a large projection screen. Thus, *HoloBeam* brings the possibility to widen FoV and making displayed images visible only to a user wearing our eyepiece. In recent years, there has been a push in research related to AR contact lenses [12], which could potentially transform AR to a seamless setting without any glasses. We understand that such contact lenses based approaches pose primary challenges in practical aspects and approvals. In the future, our work could also be transformed into an optical component embedded in a contact lens

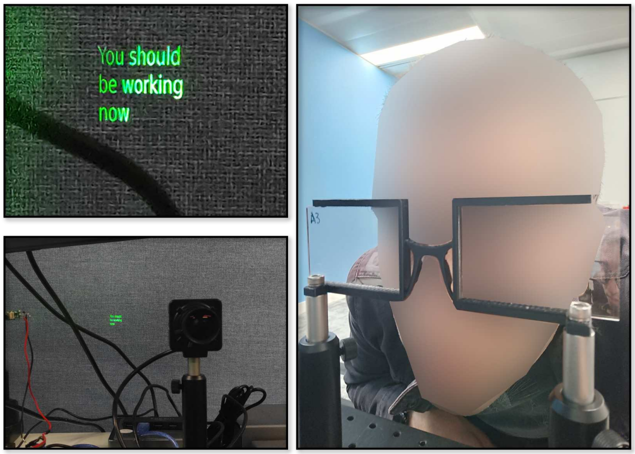

Figure 2: Various photographs showing usage and view from the HOE eyepiece. Note that this eyepiece is dedicated to amplitude-only *HoloBeam* prototype.

that helps register beams from a projector to a user's retina. Thus, *HoloBeam* has the potential to be complementary to the contact lens based AR designs of the future.

3 Amplitude-only *HoloBeam* prototype

Our amplitude-only *HoloBeam* prototype could also be built in a vertical layout as demonstrated in Fig. 1. This way, we demonstrate that our projection assembly could be assembled in various forms using optical path folding (e.g., with mirrors or prisms). Such path folding could be helpful when integrating *HoloBeam* to an automative application in a potential future.

It should also be noted that we provide sample photographs from our benchtop prototype in Fig. 2.

4 Phase-only *HoloBeam* prototype

Additional photographs from our phase-only *HoloBeam* prototype are provided in Fig. 3.

5 Optical beam propagation

A phase-only hologram is described as a two-dimensional array filled with phase values and typically described with a complex notation as $O_h = e^{j\phi(x,y)}$, where ϕ represents the phase delay introduced by each pixel at a phase-only hologram. Holographic displays typically represent holograms, O_h with programmable SLMs. Meanwhile, a coherent beam illuminating a phase-only hologram, U_i, is also described as a two-dimensional array. Note that U_i is an oscillating electric field described as $U_i = A_0 e^{j(k\vec{r}+\phi_0(x,y))}$, where A_0 represents the amplitude of the optical beam, k represents the wavenumber that can be calculated as $\frac{2\pi}{\lambda}$, λ represents the wavelength of light, and ϕ_0 represents the initial phase of the optical beam. In calculation, A_0 is often considered $A_0 = 1$ for an ideal collimated beam, while ϕ_0 is assumed to be a two-dimensional array filled with random values between zero to 2π. Finally, leading to simplification of U_i as $e^{j\phi_0}$. In simple terms, as U_i illuminates O_h, U_i by modulated with O_h, forming a new modulated beam U_m that is calculated as

$$U_m = U_i O_h = e^{j(\phi(x,y)+\phi_0(x,y))}. \quad (1)$$

A modulated beam, U_m, has to propagate in free-space from the hologram plane (SLM plane) towards a target depth plane to reconstruct images at various depth planes. Propagation of optical beams from one plane to another follows the theory and method introduced by Rayleigh-Sommerfeld diffraction integrals [5]. This diffraction integral's first solution, the Huygens-Fresnel principle, is expressed as follows:

$$u(x,y) = \frac{1}{j\lambda} \int\int u_0(x,y) \frac{e^{jkr}}{r} cos(\theta) dxdy, \quad (2)$$

where resultant field, $U(x,y)$ is calculated by integrating over every point across hologram plane, $U_0(x,y)$ represents the optical field in the hologram plane for every point across XY axes, r represents the optical path between a selected point in hologram plane and a selected point in target plane, θ represents the angle between these points. The angular spectrum method, an approximation of the Huygens-Fresnel principle, is often simplified into a single convolution with a fixed spatially invariant complex kernel, $h(x,y)$ [13],

$$u(x,y) = u_0(x,y) * h(x,y) = \mathscr{F}^{-1}(\mathscr{F}(u_0(x,y))\mathscr{F}(h(x,y))). \quad (3)$$

In our implementations, we rely on a fundamental library for optical sciences [2] , which provides a differentiable version of various optical beam propagation methods. Therefore, our methods can work with other beam propagation approximations.

6 Additional results from Phase-Only *HoloBeam* Prototype

In this section, we provide additional results from the eyebox of our prototype. Fig. 4 depicts the effect of using coherent and incoherent illumination sources in the phase-only *HoloBeam* prototype.

Fig. 5 demonstrates a full color image reconstruction using our phase-only *HoloBeam* prototype. Please note that this demonstrator does not have color distortion correction implemented, yet. Therefore, in Fig. 5, readers will observe color distortions and misalignments in the final image.

Fig. 6 provides additional captures from our phase-only *HoloBeam* prototype, demonstrating the image quality.

For more results from our prototypes, please consult to provided compressed files that contains videos and raw photographs from our experiments.

References

[1] E. Agustsson and R. Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 126–135, 2017.

[2] K. Akşit, A. S. Karadeniz, P. Chakravarthula, W. Yujie, K. Kavaklı, Y. Itoh, D. R. Walton, and R. Zhu. Odak, May 2022. If you use this software, please cite it as below. doi: 10.5281/zenodo.6528486

[3] O. Bimber and R. Raskar. Modern approaches to augmented reality. In *ACM SIGGRAPH 2006 Courses*, SIGGRAPH '06, p. 1–es. Association for Computing Machinery, New York, NY, USA, 2006. doi: 10.1145/1185657.1185796

[4] M. K. Hedili, M. O. Freeman, and H. Urey. Microlens array-based high-gain screen design for direct projection head-up displays. *Applied optics*, 52(6):1351–1357, 2013.

[5] J. C. Heurtley. Scalar rayleigh–sommerfeld and kirchhoff diffraction integrals: a comparison of exact evaluations for axial points. *JOSA*, 63(8):1003–1008, 1973.

[6] Y. Itoh, T. Kaminokado, and K. Akşit. Beaming displays. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2659–2668, 2021.

[7] C. Kim, H. Zimmer, Y. Pritch, A. Sorkine-Hornung, and M. H. Gross. Scene reconstruction from high spatio-angular resolution light fields. *ACM Trans. Graph.*, 32(4):73–1, 2013.

[8] S. Kimura, D. Iwai, P. Punpongsanon, and K. Sato. Multifocal stereoscopic projection mapping. *IEEE Transactions on Visualization and Computer Graphics*, 27(11):4256–4266, 2021.

[9] G.-Y. Lee, J.-Y. Hong, S. Hwang, S. Moon, H. Kang, S. Jeon, H. Kim, J.-H. Jeong, and B. Lee. Metasurface eyepiece for augmented reality. *Nature communications*, 9(1):1–10, 2018.

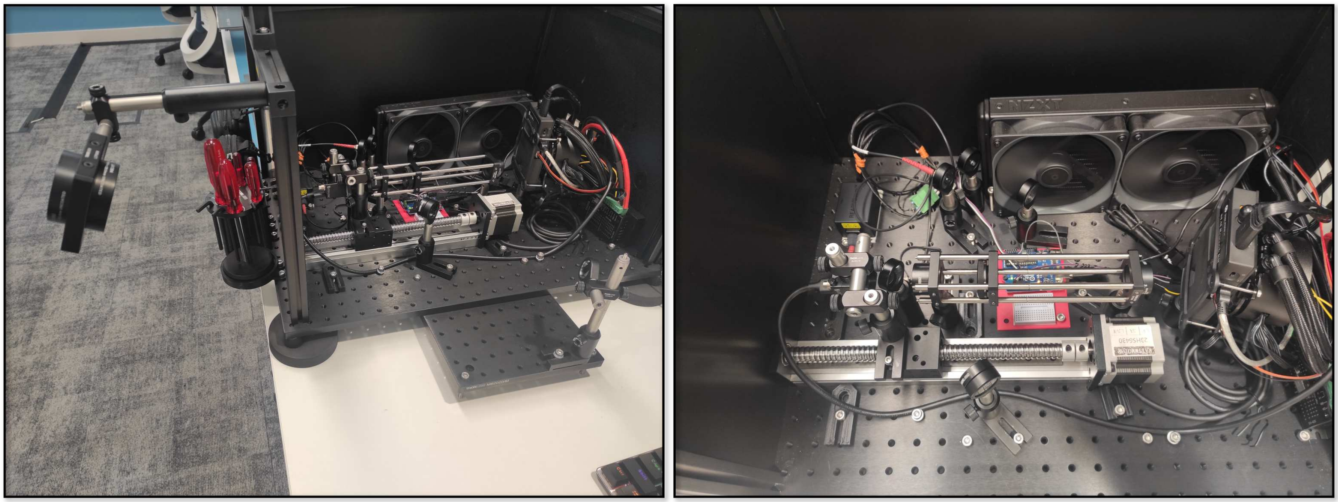

Figure 3: Additional photographs from our phase-only *HoloBeam* prototype. Images shows a true scale assembly and close look up at the projector assembly. Cooler fan in this hardware setup helps keeping the Phase SLM stable at a fix temperature. In a final product, these components are not necessarily needed, and there are much custom compact holographic projectors in some existing products (e.g., Holoeye).

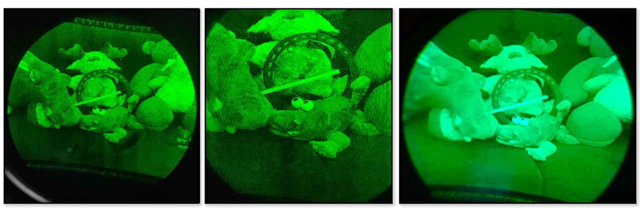

Figure 4: Photographs from the eyebox of the phase-only *HoloBeam* prototype. (Left) A photograph showing the resultant reconstructed image when a single color of the laser is used. (Middle) A close-up view from the very same scene shown in the left image. (Right) A photigraph showing the resultant reconstructed image when a single color LED is used. Lasers cause unintended perturbations in the image as demonstrated in this figure, while LEDs lead to a smoother image. Note that these photographs were taken using a smartphone camera.

[10] S. C. Malek, A. C. Overvig, A. Alù, and N. Yu. Multifunctional resonant wavefront-shaping meta-optics based on multilayer and multi-perturbation nonlocal metasurfaces. *Light: Science & Applications*, 11(1):1–13, 2022.

[11] T. Roosendaal. Big buck bunny. In *ACM SIGGRAPH ASIA 2008 computer animation festival*, pp. 62–62. 2008.

[12] J. Sano and Y. Takaki. Holographic contact lens display that provides focusable images for eyes. *Optics Express*, 29(7):10568–10579, 2021.

[13] M. Sypek. Light propagation in the fresnel region. new numerical approach. *Optics communications*, 116(1-3):43–48, 1995.

[14] J. Xiong, K. Yin, K. Li, and S.-T. Wu. Holographic optical elements for augmented reality: principles, present status, and future perspectives. *Advanced Photonics Research*, 2(1):2000049, 2021.

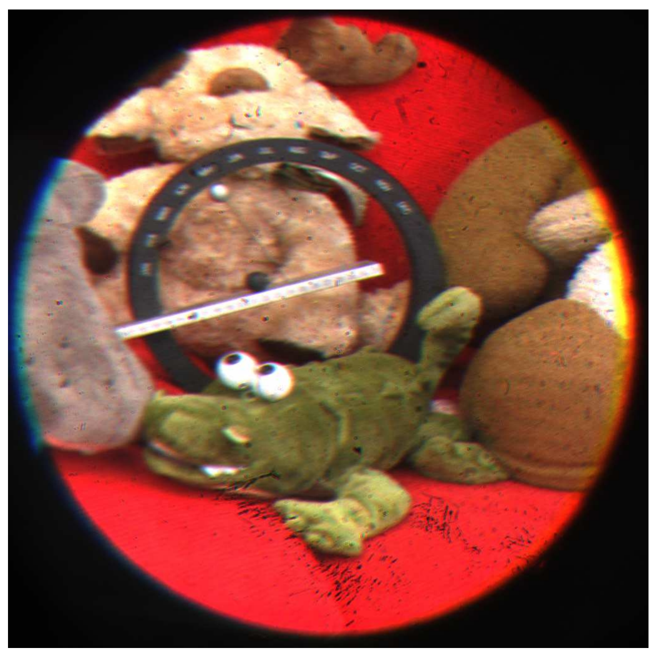

Figure 5: Photographs from the phase-only *HoloBeam* prototype. Full color photographs from our *HoloBeam* prototype. Note that full color images do not have the correct distortion correction in-place, thus leading to abberated images.

Figure 6: Resolution Quality of phase-only *HoloBeam* prototype. The first row shows captured photographs from the eyebox of our phase-only *HoloBeam* prototype. These photographs are captured using a XIMEA MC245MG-SY-UB image sensor equipped with an adjustable $5-50$ mm lens while using $20ms$ exposure times to approximate a human observer's view. During these captures, only a green LED light source is used in the display prototype. The second row shows zoomed-insets from the first row, while the third row shows the target images. The provided images are circular as the aperture of our eyepiece in this prototype is circular. The image quality of our phase-only *HoloBeam* approximates the ground truth images with a slight loss in contrast. The source image in the second, third and fourth column are from Big Buck Bunny [11], DIV2K [1] and Changil et al. [7], respectively.