Near-eye Light Field Holographic Rendering with Spherical Waves for Wide Field of View Interactive 3D Computer Graphics

LIANG SHI, NVIDIA and MIT CSAIL
FU-CHUNG HUANG, WARD LOPES, NVIDIA
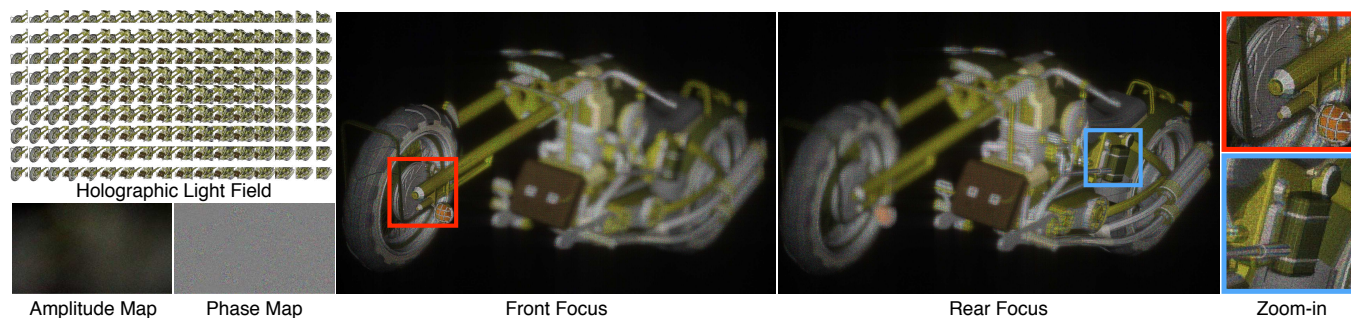WOJCIECH MATUSIK, MIT CSAIL
DAVID LUEBKE, NVIDIA

Fig. 1. Our holographic light field renders correct perspectives to generate a complex hologram with amplitude and phase maps at interactive rate (1080p). The simulated perceived images (from 4K hologram) demonstrate high spatial resolution and depths allowing eye to focus freely in space.

Holograms display a 3D image in high resolution and allow viewers to focus freely as if looking through a virtual window, yet computer generated holography (CGH) hasn't delivered the same visual quality under plane wave illumination and due to heavy computational cost. Light field displays have been popular due to their capability to provide continuous focus cues. However, light field displays must trade off between spatial and angular resolution, and do not model diffraction.

We present a light field-based CGH rendering pipeline allowing for reproduction of high-definition 3D scenes with continuous depth and support of intra-pupil view-dependent occlusion. Our rendering accurately accounts for diffraction and supports various types of reference illuminations for hologram. We avoid under- and over-sampling and geometric clipping effects seen in previous work. We also demonstrate an implementation of light field rendering plus the Fresnel diffraction integral based CGH calculation which is orders of magnitude faster than the state of the art [Zhang et al. 2015], achieving interactive volumetric 3D graphics.

To verify our computational results, we build a see-through, near-eye, color CGH display prototype which enables co-modulation of both amplitude and phase. We show that our rendering accurately models the spherical illumination introduced by the eye piece and produces the desired 3D imagery at the designated depth. We also analyze aliasing, theoretical resolution limits, depth of field, and other design trade-offs for near-eye CGH.

CCS Concepts: • **Computing methodologies** → **Mixed / augmented reality**;

Additional Key Words and Phrases: computer generated holography, light field, computational photography, vergence-accommodation conflict

1 INTRODUCTION

Making a comfortable visual experience is key to making practical daily use Virtual Reality (VR) and Augmented Reality (AR) successful. Providing a wide field of view, high resolution, interactivity, view-dependent occlusion, and being able to support continuous focus cues to avoid vergence-accommodation conflict are some of the most important issues for near eye displays.

Recently, near-eye light field displays have been explored as potential methods to fulfill the aforementioned requirements, in particular the continuous focus cues. Other near-eye display architectures have also been proposed to achieve a comparable viewing experiences, e.g., multi-focal displays and varifocal displays. However, combining continuous focus cues, resolution, space-time trade-offs, mechanical movement, and diffraction limits creates a challenging design space.

Holographic displays, especially using Computer Generated Holography (CGH), have been the hope of many to create the ultimate display since the late 60s and 70s. Holographic displays, though, are rarely used as mainstream displays because under plane wave illumination, the small diffraction angle directly leads to a narrow

field of view. For near-eye applications, using spherical wave illumination enables a trade-off between field of view, viewing distance to the display (eye relief), and size of exit pupil (eye box).

Holographic calculations are also notoriously computationally expensive. Various approximations, e.g. sacrificing depth fidelity or limiting the parallax to horizontal-only, are often required to speed up calculations enough to make computation time practical.

In this work, we show how to drive a CGH based near-eye display using light field rendering, We discuss the previously mentioned problems, the trade-offs between different illumination schemes, an implementation of holographic light field rendering optimized for high performance, and the benefits and the limitations of our choices made both in illumination and in the implementation.

Specifically, we make the following contributions:

- We propose a new rendering algorithm that avoids aliasing in both over-sampling, under-sampling and geometry clipping found in previous work. Our near-eye holographic paradigm supports continuous focus cues, intra-pupil occlusion, various reference wave illuminations and significantly improves the computational speed without sacrificing accuracy.
- We analyze theoretical limits on sampling, spatial resolution, and the connections between holographic displays and light field displays. We discuss the theoretical trade-off space for designing near-eye displays using CGH.
- We show a GPU implementation based on the Fresnel diffraction integral that is five orders of magnitude faster than the prior work [Zhang et al. 2015] and achieves interactive performance at high resolution and per-pixel depth for the first time.
- We build a complex-modulation see-through color display prototype supporting continuous focus cues and an expanded field of view to verify our rendering and calculation results.

2 PREVIOUS WORK

2.1 Fourier Holography

A Fourier hologram produces a flat image in the far field and is often viewed through a lens [Benton and Bove 2008]. In such holograms, an inverse Fourier-transformed target image is typically approximated by a binary amplitude mask [Brown and Lohmann 1966] or by an iteratively optimized phase modulation pattern [Fienup 1982; Gerchberg 1972; Lesem et al. 1969]. Multi-plane reconstructions are achieved by iterating between Fourier planes and applying phase retrieval techniques [Dorsch et al. 1994; Makowski et al. 2005] (without guarantee of convergence). Unfortunately, Fourier holograms provide little 3D information. A *Fresnel hologram*, as discussed below, can produce near-field 3D imagery.

2.2 Physically-based CGH

A physically-based CGH simulates optical propagation and interference to generate diffractive patterns that produce a desired light distribution [Benton and Bove 2008]. In such holograms, a digital 3D object is represented as the interference of waves from point-emitters (point-based method) [Waters 1966] or from polygonal tiles (polygon-based method) [Leseberg and Frère 1988].

The point-based method based on the Fresnel diffraction integral has been accelerated by elementary fringe mapping [Lucente 1993], texture-based GPU parallelization [Maimone et al. 2017; Masuda et al. 2006; Petz and Magnor 2003], and visibility-based non-uniform sampling [Chen and Wilkinson 2009]. In practice, such methods require a dense point cloud to reproduce a continuous surface and occlusion is often not well-handled, limiting virtual objects to simple geometries.

The polygon-based method leverages FFT-based angular spectrum (AS) decomposition [Goodman 2005] to efficiently calculate the optical field of polygons at the location of the hologram plane [Tommasi and Bianco 1993]. Remapping the surface property function to the angular spectrum allows texturing, shading, and reflection [Matsushima 2005]. Fully analytical frequency representations eliminate per-polygon AS transforms and reduce computational cost [Ahrenberg et al. 2008; Jia et al. 2014; Kim et al. 2008]. Silhouette masking during AS propagation accounts for self and mutual occlusion [Matsushima and Nakahara 2009; Matsushima et al. 2014]. Although polygon-based methods produce realistic 3D scenes with smooth surfaces, occlusion handling requires storing AS decompositions at intermediate planes and iterative optimization, which is difficult to parallelize.

2.3 Image-based CGH

The advance of computer graphics techniques has led to two types of image-based CGHs: the holographic stereogram and the layer-based method.

Conventional holographic stereogram (HS), either interferometry-based or computer-generated, partitions a hologram spatially into multiple horizontal bars with each recreating a distinct perspective captured from an imaging system or through rendering [DeBitetto 1969; Yatagai 1976]. Diffraction-specific calculations of HSs partition the hologram into elementary hologram patches, with each producing a local ray distribution that together reconstruct multiple views [Lucente and Galyean 1995]; a close analogy to integral imaging. The lack of focus cues and limited depth of field due to flat wavefronts is mitigated by introducing depth-dependent wavefront curvatures [Kang et al. 2016; Smithwick et al. 2010; Yamaguchi et al. 1993] or employing an intermediate ray sampling plane with angular spectrum propagation [Wakunami et al. 2013]. Resolution loss due to spatial partitioning and independent treatment of elementary hologram may be eliminated by combining the HS model with the point-based method [Zhang et al. 2015]. However, previous methods only generate discrete views from the center of each elementary hologram, causing aliasing, geometry clipping, and poor extensibility to spherical wave illumination, as shown in Section 5.1.

The layer-based method slices objects at multiple depths and superimposes the wavefronts from each slice on the hologram through inverse Fresnel transform or FFT-based AS decomposition [Bayraktar and Özcan 2010; Zhao et al. 2015]. View-dependent occlusion is achieved by angularly tiling groups of sliced layers [Chen and Chu 2015] or combined with the HS model to generate a layered-based elemental hologram for each spatial partition [Zhang et al. 2016].

Maimone et al. [2017] show a real-time single layer implementation on the GPU. Despite considerable computational saving for complex 3D objects, layer-based holograms cannot support continuous focus cues and accurate occlusion due to discrete plane sampling.

2.4 Holographic Display with Extended Field of View

The solid angle spanned by the maximum visible imagery to the eye defines the field of view (FoV). Chen and Chu [2015] demagnify the display pixel pitch via a 4f system to increase the Field of View at the cost of a smaller display panel. The magnification mismatch is compensated by pre-distorting the graphics model [Zhang et al. 2012]. Others have tiled multiple displays in a circular shape with each producing a specific part of scene towards the viewer [Hahn et al. 2008; Kozacki et al. 2012; Yaraş et al. 2011]. The Fourier optical system [Reichelt et al. 2012; Sato and Sakamoto 2012; Senoh et al. 2011] places a lens right against the display to produce a converging illumination and steers the viewing cones to the focal plane; the viewing zone is limited between finite depths in return for a shorter viewing distance and a wider field of view. A similar idea has also been demonstrated using holographic optical elements or compound microscopic optical systems to further expand the field of view [Li et al. 2016; Maimone et al. 2017].

2.5 Angular Displays and Light Field Displays

Surface materials with angular variations are achieved with programmable BRDFs using light field [Fuchs et al. 2008] or wave optics by directly optimizing diffraction [Glasner et al. 2014] or factorizing the Wigner Distribution Function [Ye et al. 2014]. Light field models the space-angle distribution in geometric optics, and have also been shown to support focus cues with dense angular sampling [Maimone et al. 2013] or additive multilayers [Lee et al. 2016; Narain et al. 2015]. Near eye displays have been shown using pinlights [Maimone et al. 2014], microlens arrays [Lanman and Luebke 2013], or attenuation based multilayer compression [Huang et al. 2015; Maimone and Fuchs 2013] to support intra-ocular occlusions which are important to depth perception [Zannoli et al. 2016]. However, diffraction limits further enhancement of spatial and angular resolution of light field displays. On the other hand, Ziegler et al. [2007] demonstrate a hologram transform based on off-line light field rendering, and we show a more efficient paradigm that generates high-resolution CGH interactively.

3 PRELIMINARY

In this section, we first review the mathematical foundations of computer generated holography and briefly discuss the limitations of using plane wave illumination. We then connect light fields to holograms via phase space modeling using the Wigner Distribution Function. We also discuss the benefits of using spherical wave illumination in Section 4 and introduce our proposed rendering and calculation in Section 5.

3.1 Computer Generated Holography (CGH)

A hologram converts an input "reference" light wave $E_R(x)$ to the desired output "object" light wave $E_O(x)$. This requires knowledge of both the reference wave and the object wave. Usually, the form of

the reference wave is a given, and the purpose of CGH is to compute the diffraction pattern which will do the conversion. This leaves the question of calculating the desired output waveform at the location of the hologram.

The desired output waveform at the location of the hologram may be calculated by starting at the position of the desired target 3D point and propagating light towards the hologram using the Fresnel diffraction integral. For object made up of points, this may be thought of as a summation of spherical waves originating from the points themselves:

$$E_O(x) = \sum_j \frac{A_j}{r_j(x)} e^{i\left(\frac{2\pi}{\lambda} r_j(x) + \phi_j\right)}, \tag{1}$$

where λ is the wavelength of the monochromatic light source, A_j is the amplitude of the point j at location (x_j, z_j), as shown in Figure 2, $r_j(x) = \sqrt{(x_j - x)^2 + z_j^2}$ is the Euclidean distance from the point to the pixel $(x, 0)$ on the hologram, and ϕ_j is the random initial phase associated with each diffuse point. Note that the resulting electric field is complex-valued.

A computer generated hologram modulates an input wavefront multiplicatively. For plane wave (collimated) illumination, of which the value of the wavefront is a constant across the hologram plane, the hologram to display is exactly the object electric field. For spherical wave illumination, the object electric field should be multiplied by a complex exponential with a quadratic phase to cancel out the quadratic phase of the spherical reference wave illumination.

Displaying the correct diffraction pattern requires devices which can modulate both the amplitude and the phase of the waveform with spatially varying patterns, e.g. liquid crystal (LC) displays or spatial light modulators (SLMs).

One major limitation of holographic displays using LC-SLMs is the narrow maximum deflection angle between the incident light angle θ_{in} and the outgoing angle θ_{out}, which is characterized by the grating equation:

$$\frac{\lambda}{2\Delta p} = |\sin \theta_{in} - \sin \theta_{out}|, \tag{2}$$

where Δp is the SLM pixel pitch. A typical $8\mu m$ pitch SLM can maximally deflect $532nm$ green light by $3.81°$, whereas a traditional analog hologram with nanometer-scale diffracting pattern can easily achieve a maximum angle greater than $100°$.

The small diffraction angle of SLMs severely impacts the practical application of SLMs to near eye displays under plane wave illumination, as shown in Figure 2. To observe the full extent of a SLM of width w_s, the minimal distance d_e^{min} from the eye to the SLM, called *eye relief*, is determined by the two converging rays deflected from the edges of the SLM, and is given by:

$$d_e^{min} = \frac{w_s}{2\tan\left(\sin^{-1}\left(\frac{\lambda}{2\Delta p}\right)\right)}. \tag{3}$$

The maximum angle α_{max} subtended by the SLM determines the maximum field of view (FoV):

$$\alpha_{max} = 2\tan^{-1}\left(\frac{w_s}{2d_e^{min}}\right) = 2\sin^{-1}\left(\frac{\lambda}{2\Delta p}\right). \tag{4}$$
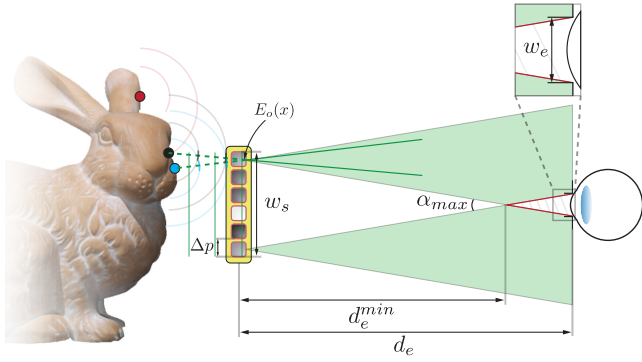
Fig. 2. Plane wave illuminated holographic display. Each display pixel diffracts wavefront into a spatially-uniform light cone. Rays are back-traced for each point emitter within the diffraction cone. The viewer can only observe the full extent of the displayed 3D image behind the intersection of two marginal cones diffracted from the edge of the display.
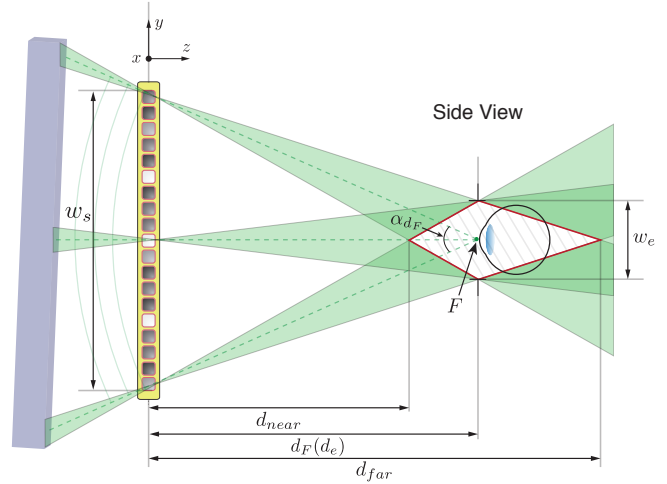

Fig. 3. Spherical wave illuminated holographic display. Each display pixel diffracts a light cone toward the focus. The entire display is observable by placing pupil in the diamond-shape viewing zone.

The red region allowing eye to move freely and still be able to see the entire hologram is called the *eye box*. When eye relief $d_e > d_e^{min}$, the eye box is obtained as

$$w_e = \frac{d_e - d_e^{min}}{d_e^{min}} w_s. \qquad (5)$$

As such, a 1080p SLM with $8\mu m$ pitch sets a minimum $265mm$ eye relief and a maximum field of view of only $3.81°$. Further, Equation (4) shows the FoV does not scale with SLM width, making CGH under plane wave illumination a less ideal candidate for near-eye displays. In section 4 we detail how spherical wave illumination surmounts these difficulties.

3.2 Holography, Wigner Distributions, and Light Fields

A light field $L(x, \theta)$ models the discrete ray space-angle distribution in geometric optics and is widely used as a content format for multiview displays. However, a high resolution light field display often suffers from diffractive blur [Huang et al. 2015]; since geometric optics models do not consider diffraction and interference, they cannot be used to display holograms. We will show how to transform light field content into a holographic wavefront in Section 5, and here we first discuss their similarity.

In light fields, the angle θ corresponds to the normal direction of the wavefront and can be related to the spatial-frequency $u = \theta/\lambda$ of the Fourier transformed object wave $\mathcal{F}(E_O(x))$, also known as the angular spectrum. Similar to light fields, the Wigner Distribution Function (WDF) [Alonso 2011; Bartelt et al. 1980; Bastiaans 1997] models both space and spatial-frequency of a scalar function, and is calculated by Fourier transforming the mutual intensity function of a scalar wave field $E(x)$:

$$W(x, u) = \int \left\langle E\left(x + \frac{f_u}{2}\right) E^*\left(x - \frac{f_u}{2}\right) \right\rangle e^{-i2\pi f_u u} df_u, \qquad (6)$$

where $\langle \cdot \rangle$ is the time-averaging operator for incoherent illumination. Note that the WDF is real but not necessarily positive due

to destructive interference; however, projecting the WDF along either axis results in a non-negative function. In particular, we obtain the intensity distribution by projecting along the spatial-frequency (u-axis):

$$\int_{-\infty}^{+\infty} W(x, u) du = |E(x)|^2. \qquad (7)$$

A non-negative WDF that resembles a light field can also be obtained by convolving the WDF of $E(x)$ with the inverted WDF of a finite aperture function [Zhang and Levoy 2009].

The definition of the WDF, unlike the light field, reveals an important property of holograms: the angular light distribution is inherently encoded in the spatial pattern of the scalar function $E(x)$, and no spatial resolution is sacrificed.

Despite the similarity and relationship between light fields and WDFs, reversing a light field to obtain the scalar function $E(x)$ is challenging: even though the inverse Fourier transform of the spatially-compressed WDF $\mathcal{F}^{-1}(W(x/2, u))$ gives the scalar wave field up to a constant scalar $E(x)E^*(0)$ [Claasen and Mecklenbräuker 1980], obtaining the WDF by deconvolving an arbitrary light field with the inverted WDF of a local aperture is impractical and does not always yield a valid WDF [Jagannathan et al. 1987]. Even with valid WDFs, inverting a high resolution (i.e. 1080p) WDF requires storing and processing trillions of samples which is intractable. Ye et al. [2014] approximate the nonnegative WDF via iterative global optimization. However, they sacrifice the spatial and angular resolution by partitioning the scalar field into independent segments.

With the aforementioned computational and physical limitations, we will not rely on the WDF to generate holograms. Nevertheless, the WDF's spatial-angular representation of phase space makes it ideal for analyzing holographic display characteristics both mathematically and graphically. We will revisit the WDF in Section 7.

4 SPHERICAL WAVE ILLUMINATION

CGHs under plane wave illumination give a narrow field of view and set a long eye relief, as discussed in Section 3.1. In holographic projection, using a diverging illumination expands the projected image field [Qu et al. 2015]. Similarly, we employ a converging beam to create a near eye CGH, achieving wider field of view and shorter eye relief, as shown in Figure 3.

We denote the focus of converging wavefront F, and the distance between the focus and the SLM d_F. Unlike under plane wave illumination, the eye box is confined between two depths determined by Equation (2):

$$d_{\{far,near\}} = \frac{w_s}{2\tan\left(\sin^{-1}\left(\sin\left(\tan^{-1}\left(\frac{w_s}{2d_F}\right)\right) \pm \frac{\lambda}{2\Delta p}\right)\right)}. \quad (8)$$

When the eye is located at the focus (eye relief $d_e = d_F$), the field of view α_{d_F} and the eye box $w_{e(d_F)}$ follow

$$\alpha_{d_F} = 2\tan^{-1}\left(\frac{w_s}{2d_e}\right), \quad w_{e(d_F)} = 2\tan\left(\sin^{-1}\left(\frac{\lambda}{2\Delta p}\right)\right)d_F. \quad (9)$$

Equation 9 shows the significance of using converging spherical wave illumination, that *(1) the eye relief only depends on the focal length of wavefront, (2) the field of view depends on both the SLM width w_s and the eye relief d_e, which is constrained by the demand for a larger eye box, and (3) for any given eye relief, a smaller pixel pitch results in a larger eye box.*

These relations offer more flexibility in designing near-eye displays, for example, using a shorter focal length wavefront allows for a wider field of view but a smaller eye box [Maimone et al. 2017]. In this paper, we require a minimal 10mm eye box coverage.

Using spherical wave illumination impacts the entire pipeline from the optical setup, the CGH calculation in Equation (1), and the light field rendering. In Section 5, we redesign light field rendering to adapt the diffraction geometry under both plane wave and spherical wave illumination, and show how to avoid the aliasing and clipping problem in the methods proposed by the previous work.

5 RENDERING AND CALCULATION FOR CGH

In the proposed CGH rendering pipeline, we start with polygon-based light field rendering such that occlusion is handled through the graphics pipeline using z-buffer. The output is a multiview holographic light field which contains color and position information intended for CGH calculation, as shown in Figure 4. We then transition to the point based method with local partitioning, which allows for per-pixel depth reconstruction and parallel computation on GPU at interactive speed in 1080p image quality for the first time.

In the following, we first describe the improved rendering setup under plane wave illumination, which avoids aliasing and geometry clipping inherited in the previously proposed rendering setup [Zhang et al. 2015]. We then extend the proposed algorithm to support spherical wave illumination, which enables near-eye CGHs. The Fresnel diffraction integral based calculation which converts a holographic light field into CGH is presented at the end.
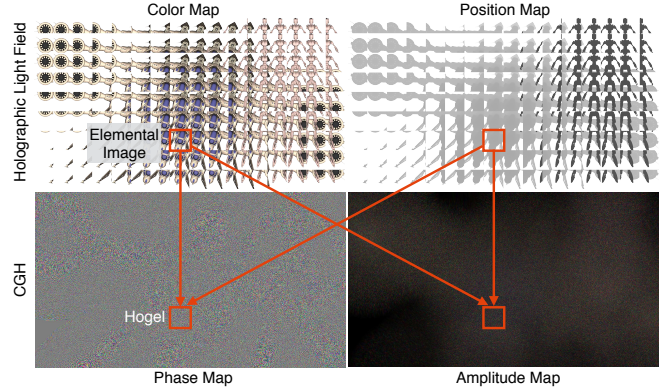

Fig. 4. We render the holographic light field with both color (top-left) and position (top-right) information . Each elemental image maps one-to-one to the associated hogel and this allows for parallelized acceleration. However, within each elemental-to-hogel pair, there is a dense mapping between every pair of pixels from both sides: every pixel in the hogel accumulates both phase and amplitude from every pixel in its elemental image, thus forming a complete bipartite compute graph per elemental-to-hogel pair.

5.1 Holographic Light Field Rendering

Calculating a full parallax hologram demands a densely sampled 4D light field [Huang et al. 2015; Maimone et al. 2013], where a local view (elemental image), which records depth-sorted visible scene points, is rendered for each SLM pixel from the its spatial location and based on the diffraction geometry. The same 4D light field can also be rendered through a dense angular view sampling at the SLM resolution, an approach that is often more preferred due to fewer rendering passes. Nevertheless, even under the small diffraction angle of existing SLMs, rendering a 1080p spatial resolution light field at an angular resolution equivalent to human visual acuity is computationally impractical for real-time graphics.

Assuming Lambertian surfaces, rendering a dense 4D light field generates highly similar views for adjacent SLM pixels and results in large redundancies: a single recording of each object point, regardless of the visibility to each SLM pixel, is sufficient to determine the point's wavefront over the entire SLM. Leveraging this observation, a hologram can be spatially partitioned into abutting grids, called holographic elements (hogels), as shown in Figure 4, such that only a single representative view, an elemental image, is rendered and used in calculation for each hogel, assuming all captured points are visible to all pixels in the hogel. The reduction in parallax introduces an error in occlusion bounded by the hogel size w_h (about 1mm in our prototype) within the eye box; this approximation substantially reduces the number of rendering pass (Section 7.6), and is widely adopted in holographic stereogram calculations (Section 2.3).

Under plane wave illumination, the previous work renders a hogel's view by setting a virtual camera directly on the SLM surface and at the hogel's center, as shown in Figure 5 (bottom). The view frustum, as shown by the pink cone, has a field of view equaling the maximum diffraction angle defined by the SLM pixel. However, this scheme cannot provide an accurate per-pixel wavefront accumulation as it only renders the object points visible to the hogel's center
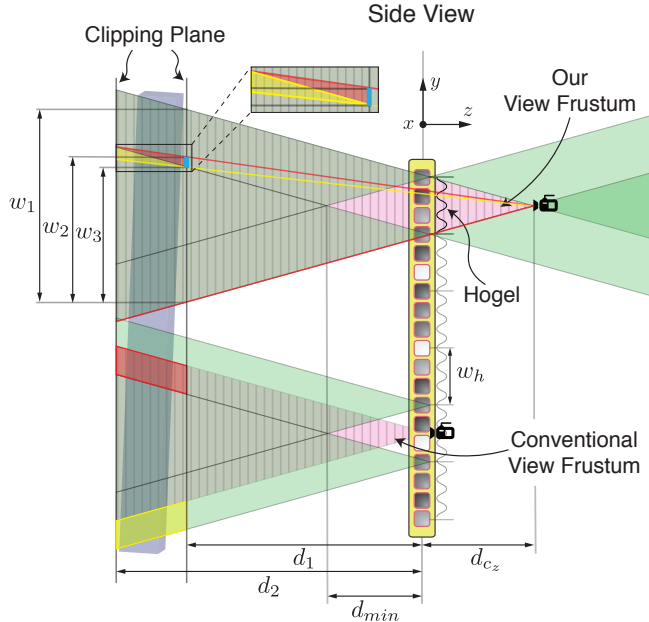
Fig. 5. Holographic light field rendering for plane wave illumination. The proposed view frustum (top, pink cone) covers the entire space in front of the SLM, as opposed to the convention setup (bottom, pink cone), where geometry clipping happens when objects are located between two adjacent hogels and within a distance d_{min} to the SLM. In our setup, the rendering contains a larger area w_1 than what each SLM pixel sees, thus we use a sliding window w_2 based on the maximum diffraction angle to disambiguate the blue segment and select the valid fraction from the rendered elemental image for each SLM pixel under consideration.

pixel and also extends poorly to spherical wave illumination. Consider the hogel's bottom SLM pixel, the rendering of objects from the top red parallelogram is mistakenly incorporated, and that from the bottom yellow parallelogram is incorrectly excluded. Meanwhile, to prevent geometry clipping, the separation w_h between adjacent view frustums sets an implicit depth limit

$$z \leq -d_{min} \equiv -\frac{w_h}{2 \tan\left(\sin^{-1}\left(\frac{\lambda}{2\Delta p}\right)\right)}, \tag{10}$$

to the objects and the near clipping plane; artifacts are shown in Figure 11. We propose a unified framework, holographic light field rendering, in the following, to support accurate view estimation under various illuminations.

To gather the part of 3D scene visible to the entire hogel, we laterally shift the virtual camera in the z direction, as shown in Figure 5 (top), to the intersection of the two marginal rays, defined by the diffraction cone associated with the hogel's top and bottom pixel. The lateral offset to the SLM is given by $d_{c_z} = d_{min}$. The offset camera projects all the scene points, visible to the pixels in the hogel, to the area $w_1 = 2 \sin^{-1}(\lambda/2\Delta p)(d_1 + d_{c_z})$ for further processing. Again, for the hogel's bottom pixel, we incorrectly incorporate the top portion of the scene points when using the entire view w_1.

To address this problem, a conservative sliding window considers only the valid fraction in the rendered hogel view during the final

wave calculation. This sliding window, $w_3 = \left(1 - \frac{d_{c_z}}{d_1 + d_{c_z}}\right) w_1$, on the front clipping plane is defined by projecting a cone of maximum diffraction angle from the SLM pixel under consideration (the bottom hogel pixel in fig. 5 (top)) to the scene object. Since we approximate the projection geometry by replacing the per-pixel rendering with the per-hogel rendering, the 3D projection to the offset virtual camera also introduces ambiguity segments, as shown in the inset of Figure 5. For the bottom hogel pixel, an example is shown in the blue segment, where both the valid yellow triangle (should be considered) and the invalid red triangle (should be excluded) are projected to the same location. During the final wave calculation, we disambiguate this region by testing whether or not the angle to the SLM pixel under consideration is smaller than the maximum diffraction angle. To test the visibility, we include the ambiguity segment by extending the sliding window to the following:

$$w_2 = \left(1 - \frac{d_{c_z}}{d_2 + d_{c_z}}\right) w_1. \tag{11}$$

Arranging a camera array under the proposed configuration allows for unrestricted disposition of scene objects. The lateral separation between camera and SLM ensures that adjacent camera views overlap immediately in front of the SLM, and the tiled frustum array fully covers the field of view of the entire hologram, allowing near clipping plane to be set at arbitrary depth in front of SLM. The configuration works well for rendering under plane wave diffraction geometry. For rendering under spherical wave illumination, the prior discussion still applies, but the camera geometry has to consider the limited eye box and the expanded field of view as discussed in Section 4.

5.2 Spherical-Holographic Light Field Rendering

Using spherical wave illumination enables a wider field of view. However, the view-frustum must be updated to accommodate the new diffraction geometry. Specifically, the frustums undergo spatially-varying transforms: the wavefront curvature introduces an off-axis rotation to the diffraction cones and centers each frustum on the local incident ray direction, as shown in Figure 6 (top). Meanwhile, at a larger incident angle, the diffraction cone and the associated frustum shrinks according to the grating equation (eq. (2)), but stills covers the eye box defined by the diffraction cone of the SLM's center pixel. Overall, the updated viewing frustums of all SLM pixels collectively widen the field of view. In the conventional setup (bottom), the cameras' viewing frustums are largely disjoint to the pixels' frustums and fail to provide a properly generalized view.

Extending the proposed camera arrangement to spherical illumination sets the virtual camera at the intersection of the marginal rays restricted by the eye box (eq. (9)) and skews the field of view. The separation between camera and SLM is given by

$$d_{c_z} = \frac{d_F w_h}{w_{e(d_F)} + w_h}. \tag{12}$$

The offset between camera and hogel center along x- and y-axis depends on each hogel's position relative to the eye box. Assuming $2M + 1$ by $2N + 1$ partitioning of the SLM along x and y, the displacement from the (m, n)-th hogel center to its virtual camera is
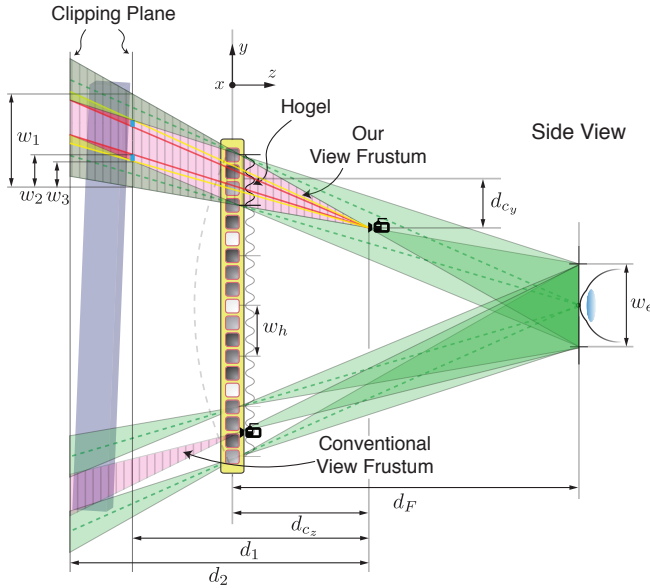
Fig. 6. Holographic light field rendering for spherical wave illumination. The camera view frustums are spatially non-uniform and steered off-axis to capture hogel's entire field of view. Calculating the per-SLM-pixel view is similar to that in the plane wave illumination setup. Points in the red triangles (projected to the blue segment) are disambiguated by examining the angles with respect to the central ray (dashed green line) of the diffraction cone for the pixel under consideration.

given by

$$d_{c_x} = \frac{m w_h d_{c_z}}{d_F}, \qquad d_{c_y} = \frac{n w_h d_{c_z}}{d_F}. \tag{13}$$

In camera space, we can define the off-axis projection matrix as

$$P_{\{m,n\}} = \begin{bmatrix} \frac{2d_{c_z}}{w_h} & 0 & \frac{2d_{c_x}}{w_h} & 0 \\ 0 & \frac{2d_{c_z}}{w_h} & \frac{2d_{c_y}}{w_h} & 0 \\ 0 & 0 & -\frac{d_2+d_1}{d_2-d_1} & -\frac{2d_2 d_1}{d_2-d_1} \\ 0 & 0 & -1 & 0 \end{bmatrix}. \tag{14}$$

The sliding window w_2 inside each elemental image to disambiguate the projected pixels is

$$w_2 = \left(1 - \frac{d_{c_z}(d_2 + d_F - d_{c_z})}{d_2 d_F}\right) w_1. \tag{15}$$

We also analyze the error and performance gain due to choices of large hogel size by examining the error-free fraction in the sliding window

$$\frac{w_3}{w_2} = 1 - \frac{(d_2 - d_1)d_{c_z}}{d_1(d_2 - d_{c_z})}; \tag{16}$$

more details are discussed in Section 7.6. We emphasize two extreme choices of hogel size. When each pixel is treated as an individual hogel, the proposed holographic light field rendering renders the exact view perceived by each SLM pixel and produces a dense 4D light field as that in Huang et al. [2015]. On the other hand, when the entire SLM is treated as a single hogel, depth cue from intra-occular occlusion [Zannoli et al. 2016] is ignored and the approach is the same as that adopted by Maimone et al. [2017].

With the rendered holographic light field, as shown in Figure 4 (top row), we can compute the interference patterns (bottom row) to be shown on the SLM pixel by accumulating the amplitudes and phases in the cropped elemental image using the sliding window.

5.3 Calculating Object Wave for Displaying on SLM

In holographic light fields, the one-to-one mapping between an SLM hogel and its visible elemental image, as shown in Figure 4, facilitates parallel computation of the Fresnel diffraction integral.

Let p denote a SLM pixel in the (m, n)-th hogel at a displacement $(\Delta x, \Delta y)$ to the hogel center. In its associated camera space under spherical wave illumination, p's spatial coordinate is given by $(\Delta x + d_{c_x}, \Delta y + d_{c_y}, -d_{c_z})$. Let p's estimated view be a sliding window of $P_{w_2} \times P_{w_2}$ pixels and define q_j be the j-th elemental pixel whose rendered point is located at $(x_{q_j}, y_{q_j}, z_{q_j})$ with amplitude A_{q_j} and initial phase ϕ_{q_j}. The Euclidean distance between p and q_j is given by

$$r(p, q_j) = \sqrt{(x_{q_j} - \Delta x - d_{c_y})^2 + (y_{q_j} - \Delta y - d_{c_y})^2 + (z_{q_j} + d_{c_z})^2}. \tag{17}$$

Since the CGH converts a reference wave $E_R(x)$ to the object wave $E_O(x)$, we need to pre-multiply the object wave with the conjugate reference wave—a reference wave that propagates away from the SLM. The conjugate spherical wave also has to be evaluated in a per-pixel basis that requires the distance from p to the focus of spherical wavefront F:

$$r(p, F) = \sqrt{d_F^2 + (m w_h + \Delta x)^2 + (n w_h + \Delta y)^2}. \tag{18}$$

Substituting Equation (1) with variables defined for an on-axis spherical wave illumination establishes the complex amplitude on the SLM at location p:

$$E(p) = \underbrace{\left(\sum_{j=1}^{P_{w_2}^2} \frac{A_{q_j}}{r(p, q_j)} e^{i2\pi \frac{r(p, q_j)}{\lambda} + \phi_{q_j}}\right)}_{\text{Object Wave } E_O(x)} \underbrace{\left(\frac{A_F}{r(p, F)} e^{i2\pi \frac{r(p, F)}{\lambda}}\right)}_{\substack{\text{Conjugate Reference} \\ \text{Wave } E_R^*(x)}}, \tag{19}$$

where A_F denotes the amplitude of conjugate spherical wave. Note that the CGH under plane wave illumination is simply the object wave $E_O(x)$ in the equation, where $r(p, q_j)$ is reformulated according to plane wave rendering geometry. For paraxial region, Equation (19) can be simplified by the Fresnel approximation to remove the square root operation. In this paper we aim for an accurate quadruple sum between ps and qs on the SLM and the rendered scene plane respectively.

6 IMPLEMENTATION AND RESULT

Our implementation depends on the desired trade-off between field of view and eye box described in Equation (9) and the available off-the-shelf hardware. In this section, our design requires a minimum eye box of $10mm$ size, and we will discuss more design trade-off spaces for software in Section 7.6 and hardware in Section 7.7.

6.1 Computing Hardware and Software

We render holographic light fields on a PC with an Intel core i7 3.3GHz CPU and 16GB of RAM, and a NVIDIA TitanX GPU. The multiview holographic light field is rendered with Simultaneous

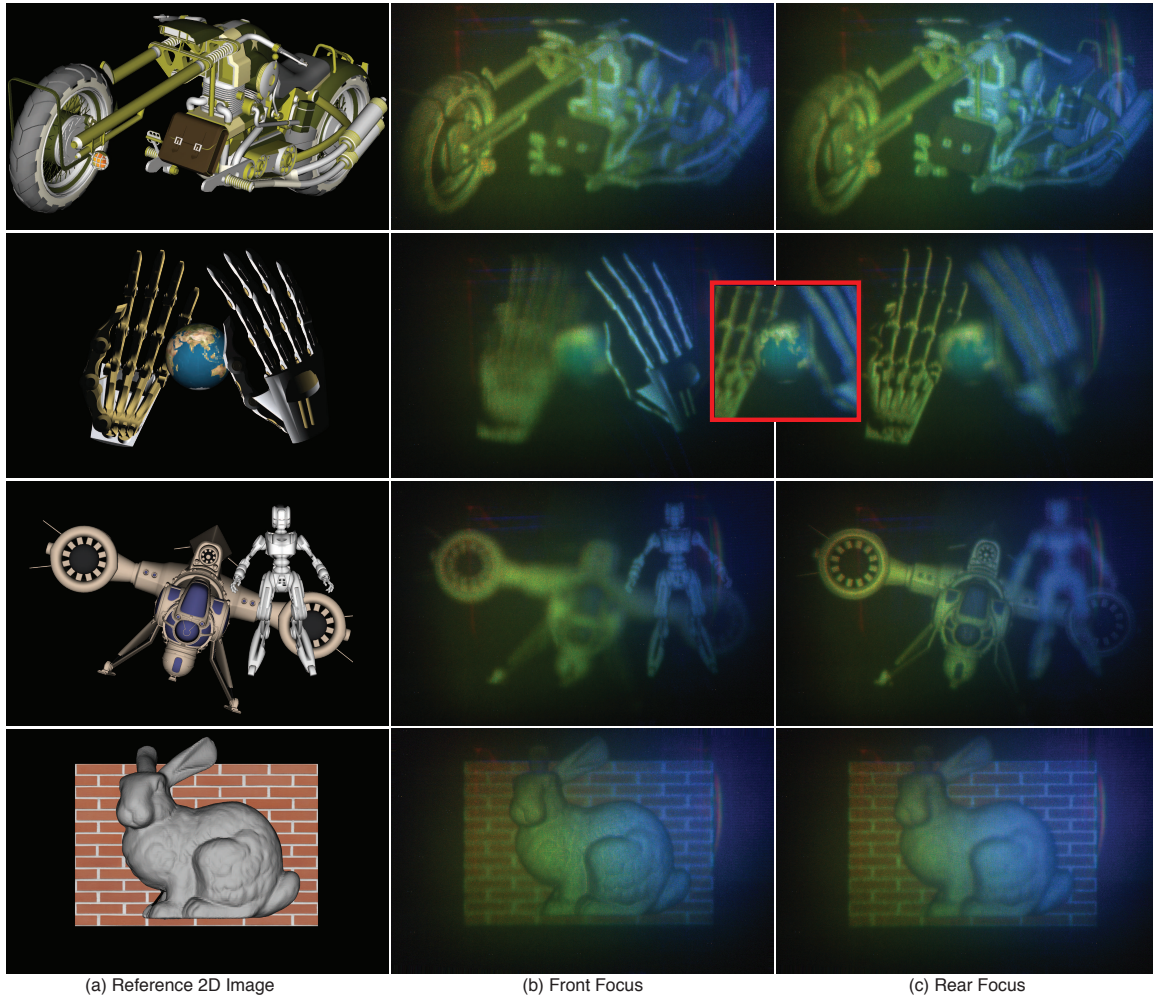| (a) Reference 2D Image | (b) Front Focus | (c) Rear Focus |

Fig. 7. We verify our rendering and CGH calculation on an optical color hologram setup. Given our core contribution is the novel rendering and fast calculation, the color artifacts due to laser calibration do not affect the accuracy of our results. We show the reference 2D image in (a). When the object is displayed on a hologram, we can support accommodation (b and c). Our computation uses per-pixel depth allowing continuous focusing, as shown in the inset for the second example.

Multi-projection at the geometry shader stage. The elemental image resolution bounds the rendering cost. Unlike in the rendering of full resolution light fields [Huang et al. 2015], the view size is bounded by the relatively small diffraction cone. For 4K SLMs, the overall rendering to a resolution of 5376×3024 light field (under 16×9 hogels) takes less than $20ms$.

We use CUDA-OpenGL Inter-Op to calculate the CGH, and the pseudo kernel code that runs on each SLM pixel is outlined in Algorithm 1, where P_{w_s}, P_{w_h}, P_{w_1}, and P_{w_2} denote the number of pixels in the SLM, the hogel, the elemental image, and the sliding window respectively. The heavy computation in the loop (line 5-12) hides the latency of repeated texture fetching (line 6 and line 9). The major performance bottleneck comes from three parts: calculating the Euclidean distance and phase shift between 3D image points and SLM pixels (line 8), accumulating complex value with trigonometric operations (line 11), and conditional branching (line 10). We

re-formulate the maximum diffraction angle comparison to avoid division and minimize the branching code. We further parallelize the compute-looping (line-5) using the prefix-sum algorithm and using on-chip shared memory to avoid memory bank conflicts. Together, we obtain an additional 40% performance gain compared to an unoptimized GPU CUDA code, 100× to a GLSL shader implementation of the same algorithm, and more than 13000× to a C++ implementation on CPU [Zhang et al. 2015]. Maimone et al. [2017] show similar performance to our GLSL code for offline computation of the point based method. We profiled the code to be heavily compute-bounded (nearly 11 trillion floating point operation).

6.2 Optical Setup

The optical setup is shown in Figure 8 with three incoming narrow band red, green, and blue lasers operating at wavelengths $642nm$,

Algorithm 1 Fringe Calculation Kernel with Spherical Wave

1: $\mathbf{i} \leftarrow pixel\ index\ in\ SLM$
2: $\phi_{E_R^*(x)} \leftarrow length\left(\left\{\left(\left(\mathbf{i} - \frac{P_{w_s}}{2}\right) \cdot \Delta p, D_F\right)\right\}\right)/\lambda \cdot 2\pi$
3: $\mathbf{i}_{Hogel} \leftarrow mod(\mathbf{i}, P_{w_h})$
4: $\mathbf{i}_{Elem} \leftarrow \lfloor \mathbf{i}/P_{w_h} \rfloor \cdot P_{w_1} + \mathbf{i}_{Hogel}/P_{w_h} \cdot (P_{w_1} - P_{w_2})$
5: **for** $Pixel\ index\ in\ sliding\ window\ \mathbf{j} = 0\ to\ P_{w_2} \times P_{w_2}$ **do**
6: $\quad \mathbf{q} \leftarrow LightFieldPositionTex[\mathbf{i}_{Elem} + \mathbf{j}]$
7: $\quad \Delta\mathbf{q} \leftarrow \left\{\left(\mathbf{i} - \frac{P_{w_h}}{2}\right) * \Delta p + \mathbf{d}_{c_{\{x,y\}}}, -d_{c_z}\right\}$
8: $\quad \phi \leftarrow Length(\mathbf{q} - \Delta\mathbf{q})/\lambda \cdot 2\pi + \phi_{E_R^*(x)} + \phi_{Init}$
9: $\quad A \leftarrow LightFieldColorTex[\mathbf{i}_{Elem} + \mathbf{j}] \cdot (1/r)$
10: $\quad$ **if** Angle < MaxDiffractionAngle **then**
11: $\quad\quad \mathbb{E}(\mathbf{i}) \leftarrow \mathbb{E}(\mathbf{i}) + A \cdot e^{i\phi}$
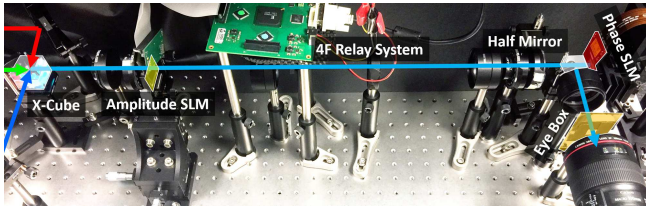12: $\quad$ **end if**
13: **end for**

Fig. 8. The optical setup relays amplitude modulation to phase modulation via a 4F system. Color reconstruction is achieved by field sequential modulation. The aperture stop at the eye box filters the undiffracted zero order beam and higher order diffractions.

515nm, and 442nm respectively. We obtain field-sequential color reconstruction by displaying the color specific holograms at the same time as illuminating with that specific color and cycling through the colors red, green, and blue. We achieve complex-modulation by relaying an amplitude SLM to a phase SLM via a 4F system. The amplitude SLM, made by bbs bild und lichtsysteme GmbH (part# LCD L3C07U-85G13) , has 1920 × 1080 resolution, 8.5μm pixel pitch, and operates in transmissive mode. We calibrate the non-linear amplitude modulation response and the accompaning phase shift. The phase SLM is made by HoloEye (part# GAEA) and factory pre-calibrated. It has 4094 × 2464 resolution, 3.74μm pixel pitch and operates in reflective mode. The hologram is computed at 4K resolution, and the amplitude is down-sampled to 1080p using nearest neighbor filtering for display. We capture the photographed results with a Cannon EOS 5D MIII DSLR camera with a 100mm f/2.8 lens.

A Thorlabs 75mm achromatic lens is placed 50mm behind the phase SLM and before the eye box. The lens, effectively, converts the plane wave illumination to a converging illumination [Reichelt et al. 2012] focused at the back focal point centered at the eye box. Our rendering and calculation are adjusted with respect to the magnified virtual SLM and the spherical wave illumination.

Though we only demonstrate a bench-top prototype, our algorithm is fully compatible with the prototypes demonstrated by Maimone et al. [2017], including the compact head-mounted design. To our knowledge, companys like Silicon Light Machine are also developing complex-modulation SLMs which will remove the necessity
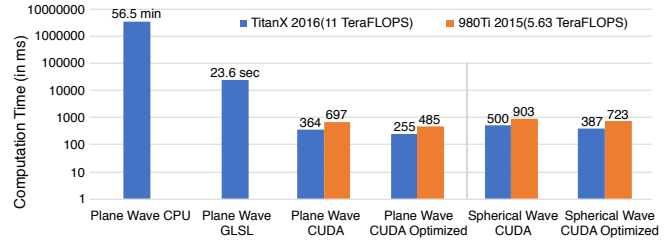

Fig. 9. Comparing performance under CPU and different GPU implementations, and different illuminations. We implement Zhang el al.'s algorithm in C++, GLSL (left two blue bars) under plane wave illumination and our algorithms in CUDA (right 4 pair of bars) under both illumination schemes. The elemental image resolution is 128 × 128 (2048 × 1152 overall points cloud) in the plane wave method, and 160 × 160 (2560 × 1440 overall) in the spherical wave method; both methods achieve 30 pixels per degree angular sampling on a 1920 × 1080 resolution SLM.

of relaying different SLMs with 4F system and largely reduce the form factor of our display prototype.

We verify our rendering and CGH calculation using the above-mentioned optical setup; photographed results are shown in Figure 7. Color artifacts are due to laser filtering, alignment, and imprecise beam stop, and do not affect the accuracy of our rendering results. Similar to the teaser, results focusing on the front tire (b) and on the rear exhaust pipe (c) of the motorcycle matches the refocusing video in the supplementary material. The front, middle, and rear focusing in the second row demonstrate the per-pixel depth scene reconstruction which supports continuous focus. High frequencies in geometry and texture are well preserved in the third and fourth row examples.

7 EVALUATION AND ANALYSIS

We evaluate the performance, model the aliasing, theoretical resolution and sampling limits through the WDF analysis, and discuss the physical limitations of the SLM panel in the design trade-off space.

7.1 Compute Performance

We evaluate the performance of the implementations discussed in Section 6, under different illumination schemes. The performance results, as shown in Figure 9, are obtained at an angular sampling rate of 30 pixels per visual degree, similar to the angular resolution found in commercial VR/AR devices. Specifically, the CPU implementation of plane wave method by Zhang et al. [2015], which we also reimplemented on CPU, takes nearly an hour to compute a 3-color-channel 1080p hologram, where an optimized CUDA implementation of our algorithm takes only 250ms, achieving more than 5 orders of magnitude speed up. The performance is compute-bounded, thus a Pascal Titan X with 11 TeraFLOPS nearly doubles the frame rate of using a GeForce 980Ti with 5.63 TeraFLOPS[1].
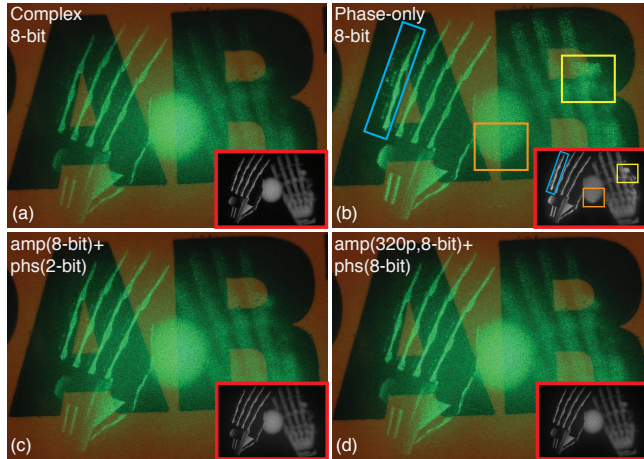
236:10 • L.Shi et al.

Fig. 10. Comparisons of (a): the ideal amplitude and phase hologram with (b): phase-only hologram, (c): lower bit depth phase, and (d): lower amplitude resolution. We compare the simulated results (insets) using a 1080p SLM.

7.2 Phase-only Holograms and Quantization

With diffuse objects (i.i.d random phases), the resulting hologram exhibits a near-uniform amplitude and most information is transfered to the phase [Goodman 2005]. Similar to most prior work, we construct a phase-only hologram by discarding the amplitude, as shown in Figure 10 (b). The reconstruction closely resembles the reference result (a) with minor artifacts marked in blue, orange and yellow. Adding a low-resolution amplitude mitigates these artifacts significantly (d).

The bit-level of phase precision can also be reduced [Zhao et al. 2015] for other purposes, e.g. higher refresh rate or resolution, without noticeably sacrificing the image quality. The image reconstructs well with a slight contrast loss even at 2-bit quantization (c). We simulate different quantization levels in the Appendix, Figure 18.

7.3 Geometry Clipping and Aliasing Comparison

Unlike previous work [Zhang et al. 2015], our rendering, as discussed in Section 5.1, avoids geometry clipping (fig. 5, bottom) for sufficiently close objects ($z \leq d_{min}$ in Equation (10)) and aliasing due to over- and under-sampling outside the view frustum in the center of hogel. We visualize these benefits in Figure 11 by reconstructing a spring that spans between $0.9d_{min}$ (13mm) and $2.1d_{min}$ (30mm) to the SLM. The front focus on the CGH produced by Zhang et al.[2015] (c) exhibits visible geometry loss and strong aliasing due to geometry clipping and over- and under-sampling, as analyzed in Section 7.4. However, their method also performs well, as shown in (d), outside the distance d_{min} for the majority of the viewing volume.

7.4 Sampling Outside the Maximum Diffraction Cone

In Section 5.1, a sliding window w_2 limits wavefront sampling; we argue that sampling the wavefront of any point outside the maximum diffraction cone dictated by the highest spatial frequency

---

[1]The upcoming NVIDIA Volta GV100 2017 has 120 TeraFLOPS in tensor operations.

ACM Transactions on Graphics, Vol. 36, No. 6, Article 236. Publication date: November 2017.
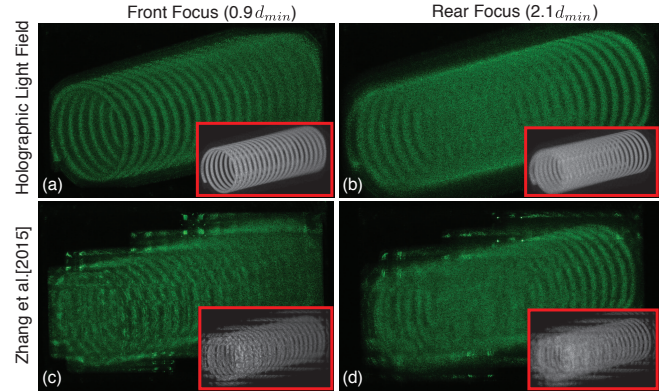

Fig. 11. The proposed method (top row, photographed) provides an accurate rendering setup (Section 5.1) such that no geometry is clipped near the SLM. The algorithm used by Zhang et al. [2015] results in geometry clipping when the object is closer to the SLM than d_{min} (c), and aliasing due to over and under-sampling (c and d). The captured results from a 1080p SLM closely resemble the simulations (red insets).
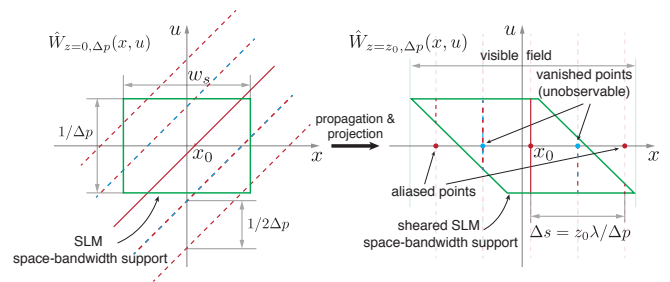

Fig. 12. WDF analysis of aliasing caused by sampling outside the maximum diffraction cone. (Left): The dashed lines denote replicas from discrete sampling. The display's space-bandwidth product in green rectangle is bounded by the spatial extent w_s and the maximum spatial frequency $\pm 1/2\Delta p$. (Right): The visible field is the lateral extent of the sheared space-bandwidth product. Each constructively interfered replica in red dash line projects to an aliased point at a displacement depicted by the Nyquist sampling rate.

$|u^{max}| = \frac{1}{2\Delta p}$ results in aliasing. Assuming paraxial region where the WDF description is accurate, we analyze the reconstruction of a single spherical point emitter on a 1D hologram. In the analysis, we also discard the spatially-varying amplitude for simplicity, by noticing that the derived resolution matches the simulation using complex amplitude wavefront well. Consider the emitter at (x_0, z_0), its wavefront at $z = 0$ and the WDF (eq. (6)) are:

$$E(x) = e^{i\pi \frac{(x-x_0)^2}{z_0\lambda}}, \quad W_{z=0}(x,u) = \delta\left(\frac{x-x_0}{z_0\lambda} - u\right). \quad (20)$$

Sampling the wavefront at a discrete interval Δp on the SLM generates infinite replicas (fig. 12 (left)) of doubly sampled WDF $W_{z=0}(x,u)$
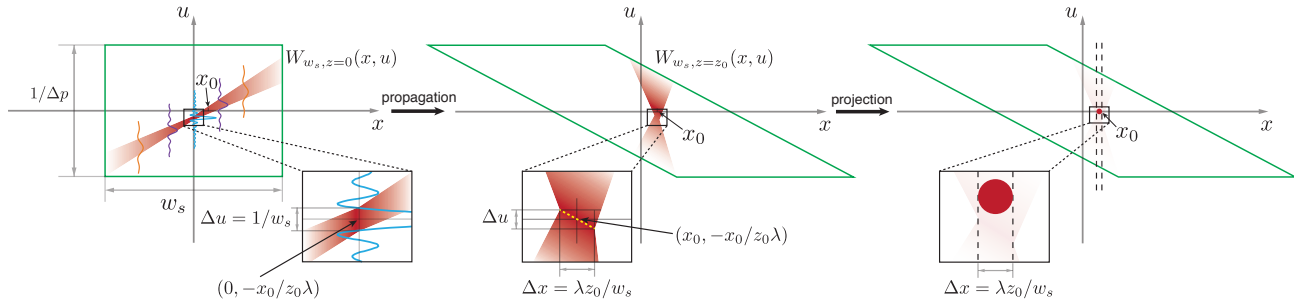
Fig. 13. WDF analysis of resolution limit outside the DoF. (Left): Frequency is minimally dilated to the main lobe of the sinc function near the center, while the peripheral dilation (orange sinc) has wider spread and weaker peak amplitude. (Middle): The yellow dash line marks the sheared minimally dilated frequency at the observed distance. (Right): The dash lines bound the region where sinc response is always positive, and the projected circle bounds the resolution.

along u-axis at a displacement of $\frac{1}{2\Delta p}$ [Stern and Javidi 2004]:

$$\hat{W}_{z=0,\Delta p}(x,u) = \frac{1}{2\Delta p}\Big(\sum_k \sum_l \delta(x - k\Delta p)W_{z=0}(x, u - l/2\Delta p)$$

(21)

$$+ (-1)^k \delta(x - (k+1/2)\Delta p)W_{z=0}(x, u - l/2\Delta p)\Big).$$

The space-bandwidth product of the SLM, shown as the green rectangle in the WDF chart, bounds the reproducible wavefront. To observe the point emitter, we propagate the wavefront back to z_0; the WDF, similar to a light field, is sheared along x-axis [Goodman 2014]. The intensity distribution at z_0 is obtained by projecting along the u-axis from the visible WDF enclosed by the sheared space-bandwidth support (eq. (7)), as shown in Figure 12 (right). For odd k-th replicas, no spatial point is observed because alternating signs (dash line interleaved with red and blue) destructively interfere; for even k-th replicas, aliased points are generated within the field of view. The displacement between adjacent aliased emitters obeys the Nyquist theorem and is jointly characterized by the maximal spatial frequency and the depth of the target emitter as $\Delta s = z_0\lambda/\Delta p$.

To avoid aliasing, the even k-th replicas should be located outside the sheared space-bandwidth support of the display: if the point emitter is located within a minimal distance to the SLM

$$z_0 \leq z_{min} \equiv (w_s + |2x_0|)\Delta p/\lambda,$$

(22)

we avoid sampling outside the maximum diffraction cone using a sliding window w_2; artifacts due to oversampling are shown in the Appendix, Figure 17. **Equation (22) also defines the *Depth of Field* (DoF)**. In Section 7.5, we discuss the limits of reconstructed resolution inside and outside the DoF (z_{min}).

7.5 Resolution Limits on the Depth of Field (DoF)

We show that, like light field displays, holographic displays also have limited spatial resolution outside the depth of field, beyond which spatial resolution exhibits a depth-dependent degradation. The effect is due to sampling of a continuous wavefront over SLM's finite spatial extent. We distinguish two cases: *spatially-limited* sampling when a point emitter is outside z_{min} (eq. (22)) and *frequency-limited* when it is inside.
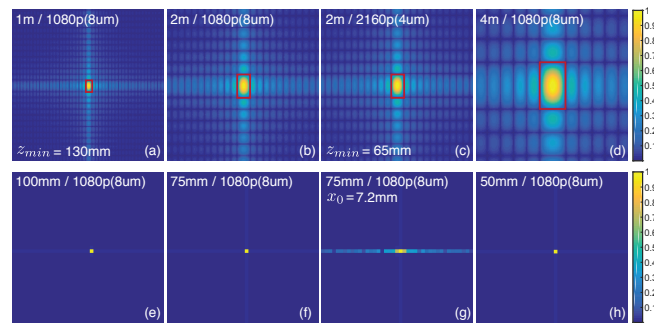

Fig. 14. Simulation of spatial resolution limit. (First Row): *spatially-limited sampling outside the DoF*. The red bounding box marks the width of the main lobe as resolution limit, and the bounding box scales proportional to the distance between the point emitter and the display. The observed resolution limit (eq. (25)) is the same for display with the same spatial extent but smaller pixel pitch (b,c); however, their DoF (eq. (22)) are different. (Second Row): *frequency-limited sampling within the DoF*. Spatial resolution in (e,f,h) remains at Δp when the projected maximum diffraction cone is not cropped by SLM boundary. When the display boundary crops the projected maximum diffraction cone, reconstruction degrades as shown in (g) due to under-sampling.

Spatially-limited sampling. When the object point is outside the depth of field ($z_0 \geq z_{min}$), the SLM is smaller than the maximum diffraction cone projected from the object point, and all pixels sample the wavefront:

$$E_{w_s}(x,0) = e^{i\pi\frac{(x-x_0)^2}{z_0\lambda}} \sqcap \left(\frac{x}{w_s}\right),$$

(23)

where $\sqcap(\cdot)$ is a rectangular function modeling SLM's finite spatial extent. The corresponding WDF can be derived as

$$W_{w_s,z=0}(x,u) = 2w_s\left(1 - \frac{|2x|}{w_s}\right)\text{sinc}\left(w_s\left(1 - \frac{|2x|}{w_s}\right)\left(\frac{x-x_0}{z_0\lambda} - u\right)\right)$$

$$= \delta\left(\frac{x-x_0}{z_0\lambda} - u\right) \otimes_u 2w_s\left(1 - \frac{|2x|}{w_s}\right)\text{sinc}\left(2w_s\left(1 - \frac{|2x|}{w_s}\right)u\right),$$

which is a convolution along u-axis between the WDF of the continuous wavefront and the WDF of the rectangular function. The convolution dilates the frequency response from a Dirac delta function to a spatially varying sinc function, as shown in Fig. 13 (left).

Because sinc degrades rapidly after its first zero-crossing, we characterize the frequency dilation by the width of the main lobe:

$$\Delta u = \frac{1}{w_s \left(1 - \frac{|2x|}{w_s}\right)}, \quad x \in (-w_s/2, w_s/2). \quad (24)$$

The minimal frequency dilation $\Delta u_{min} = 1/w_s$ is obtained at $x = 0$ and centers around $u = \frac{-x_0}{z_0 \lambda}$. To observe the point emitter, we propagate the WDF to $z = z_0$ and the propagation shears the frequency dilation to create a spatial dilation. In Figure 13 (middle), the yellow dash line marks the sheared $W_{w_s, z=0}(x = 0, u)$ and bounds the minimal spatial dilation

$$\Delta x_{min} = z_0 \lambda \Delta u_{min} = \frac{z_0 \lambda}{w_s} \equiv \frac{z_0 \lambda}{(P_{w_s} \Delta p)}. \quad (25)$$

Calculating the exact intensity distribution at $z = z_0$ requires integrating the sheared spatially varying sinc along u-axis (eq. (7)). However, we notice that for $x \in (x_0 - \frac{1}{2}\Delta x_{min}, x_0 + \frac{1}{2}\Delta x_{min})$, as shown by the region defined between the two dash lines in Figure 13 (right), the WDF is always positive as the main lobes sheared from different spatial locations are at least Δx_{min} wide. Thus, the spatial resolution of the point is upper-bounded by Δx_{min}; this width agrees with the closed-form solution $\text{sinc}^2 \left(\frac{(x-x_0)}{\Delta x_{min}}\right)$ obtained by Kizacki [2010] and shows a depth-dependent degradation.

To validate, we reconstruct a point emitter at distances $1m$, $2m$, $4m$ away from a 1920×1080 SLM of $8um$ pixel pitch, all three distances are outside the SLM's DoF as shown in Figure 14 (a, b, and d), in simulation [Matsushima and Nakahara 2009], and we find that the resolution of reconstructed points scales proportionally to the distance, as predicted by Equation (25). For these point emitters, doubling the SLM resolution while halving the pitch results in the same degraded spatial dilation as the DoF reduces (eq. (25)). As an example, a point reconstructed at $2m$ to a 3840×2160 SLM of $4um$ pixel pitch SLM (c) has the same dilation as that from a 1920×1080 SLM of $8um$ pixel pitch SLM (b). **Increasing display size w_s, e.g. by tiling, expands the DoF and avoids resolution degradation.**

Frequency-limited sampling. When the object point is within the depth of field ($z_0 < z_{min}$), the active sampling region is bounded by the intersection between the projected maximum diffraction cone and the SLM boundary. We leave the detailed derivation to the Appendix Section A and summarize the result here.

The highest spatial resolution is achieved inside a triangular region in front of the SLM, defined by the Appendix Equation (37), within which the projected maximum diffraction cone is not cropped by the SLM boundary. For object points in this region, the active sampling region on the SLM is given by $w'_s = z_0 \frac{\lambda}{\Delta p}$ and substitutes it to Equation (25) gives

$$\Delta x_{min} = \Delta p. \quad (26)$$

This reveals a key observation: **inside the depth of field, the highest spatial resolution remains constant and equals the pixel pitch.** Outside this triangular region, the resolution degrades due to spatial cropping by the SLM boundary. A generalized resolution is given by the Appendix Equation (36) and visualized by the Appendix Figure 21.
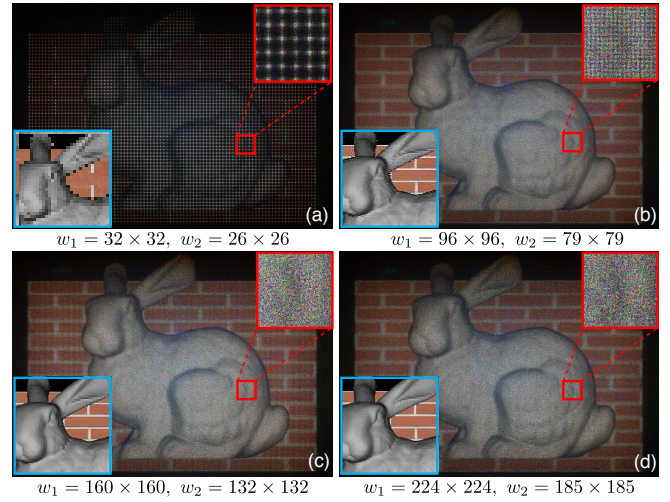

$w_1 = 32 \times 32, \ w_2 = 26 \times 26$ (a)

$w_1 = 96 \times 96, \ w_2 = 79 \times 79$ (b)

$w_1 = 160 \times 160, \ w_2 = 132 \times 132$ (c)

$w_1 = 224 \times 224, \ w_2 = 185 \times 185$ (d)

Fig. 15. Comparisons of elemental image resolution w_1 (blue inset). We simulate focusing on the bunny on a 1080p $8\mu m$ CGH. The angular sampling rate is (a): 6, (b): 18, (c): 30, and (d): 45 pixels per degree respectively. The reconstructions exhibit obvious aliasing on the top row (red inset) but are smoothed out greatly above 30 pixels per degree (c).

In Figure 14, we reconstruct a point emitter at distances $100mm$, $75mm$, $50mm$ away from a 1920×1080 SLM of $8um$ pixel pitch, all three distances are within the SLM's DoF and in the triangular region of highest spatial resolution. As shown, all reconstructions are of single pixel wide (e, f, and h). When the point is laterally offset, the under-sampling due to spatial cropping by the SLM boundary results an under-resolved point as a line in (g).

Resolution Limits under Spherical Wave Illumination. The above-mentioned analysis assumes plane wave illumination. Placing a lens of focal length d_F right behind the SLM magnifies the optical setup and creates spherical wave illumination focused at the back focal plane. With the magnification, the depth of field under spherical wave illumination is also magnified and given by:

$$\overline{z_{min}} = \frac{d_F}{d_F - z_{min}} z_{min}, \quad (27)$$

where the overline operator ‾ denotes magnified geometry under spherical wave illumination. We still distinguish two cases: when the depth of field under plane wave illumination is beyond the front focal plane ($z_{min} \geq d_F$), the CGH under spherical wave illumination has unlimited depth of field. On the other hand, the reconstructed point between the depth of field and the front focal plane ($z_{min} \leq z_0 \leq d_F$) has a magnified point spread at the virtual image plane. We summarize the reconstructed resolution as following:

$$\overline{\Delta x} = \begin{cases} \frac{d_F}{d_F - z_0} \Delta p, & \text{if } z_0 \leq \min(d_F, z_{min}) \\ \frac{d_F}{d_F - z_0} \Delta x_{min}, & \text{otherwise,} \end{cases} \quad (28)$$
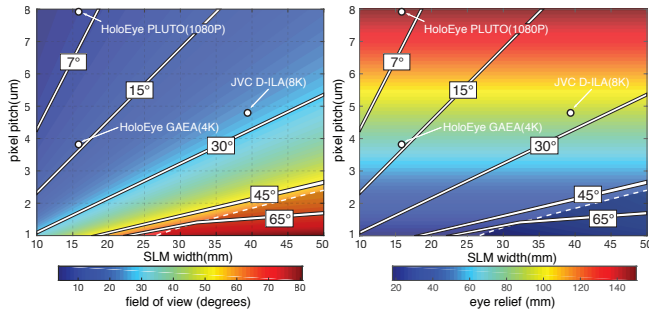
where Δx_{min} is defined in Equation (36).

Fig. 16. Design trade space between the pixel pitch Δp and the SLM width w_s assuming eye box $w_e = 10mm$ and $f/\# = 1$. (Left) The maximum FoV increases with a smaller pitch and a wider SLM. The white dash line separates the above pitch bounded case from the lower lens bounded case (shaded region); details are discussed in Section B. (Right): The minimum eye relief (focal length) decreases with a smaller pitch (eq. (9)) for a given eye box requirement, and the current off-the-shelf SLM is far from the lens-bounded region.

7.6 Elemental Image Resolution & Hogel Partitioning

The elemental image resolution gives the angular sampling rate under spherical wave illumination. We derive the maximum angular resolution $\Delta\alpha_{max} = \Delta p/d_F$ in Section C; human visual acuity or the angular resolution determine the required upper bound on the angular sampling rate $\beta = \max((1/60)^\circ, \Delta\alpha_{max})$. For a hogel partition of width w_h, the maximum hogel viewing cone φ is given by:

$$\varphi = 2\tan^{-1}\left(\frac{w_h}{2d_{c_z}}\right) = 2\tan^{-1}\left(\frac{w_{e(d_F)} + w_h}{2d_F}\right), \quad (29)$$

and the maximum elemental image resolution is obtained as:

$$k = \varphi/\beta. \quad (30)$$

For a 10mm eye box, a 3840×2160 SLM of $4um$ pixel pitch achieves an angular sampling $4\times$ the visual acuity — too expensive and impractical to render for interactive applications. In Figure 15, we find the threshold above 30 pixels per degree gives a good approximation without strong perceivable aliasing.

Hogel Partition. Although a small hogel size and dense partitioning increases the number of rendered views, the reduced ambiguity region (eq. (16)) produces accurate perspectives for intra-ocular occlusion. To balance this trade-off, we evaluate the hogel size w_h based on the ratio between the error-free segment w_3 and the approximated view sliding window w_2. Choosing $w_h \approx 1mm$ in our prototypes reduces the ambiguity region to less than 0.16% for a two-dimensional view and gives a 16×9 hogel partitioning. A larger pixel pitch requires denser hogel partitioning; however, rendering is not the bottleneck.

7.7 Practical Design Trade-offs

Using spherical wave illumination opens a design trade space not possible under plane wave illumination: for a given eye relief, increasing the FoV does not require simultaneously increasing the size of the display panel and decreasing the pixel pitch (eq. (9)).

In Figure 16, we show the design space of spherical wave illumination based holographic displays according to Equation (9). We

assume a minimum eye box $w_e = 10mm$, wavelength $\lambda = 532nm$, and a f-number limited to one. The f-number is the ratio of the focal length of the lens to the diameter of the lens. The design trade-off is often described as the "etendue". Small f-number and small eye box creates even wider FoV as shown in the Appendix [Maimone et al. 2017]. We plot off-the-shelf SLMs in the figure; SLMs with pitch $\leq 3um$ or 8K resolution exists [Bleha and Lei 2013].

For a given pixel pitch Δp and eye box w_e, we also derive the bound below which FoV and eye relief are solely limited by maximum diffraction angle. The bound for the SLM width is give by

$$\overline{w_s} = \left(\frac{\tan\left(\sin^{-1}\left(\sin\left(\tan^{-1}\left((f/\#)/2\right)\right) + \frac{\lambda}{2\Delta p}\right)\right)}{\tan\left(\sin^{-1}\left(\frac{\lambda}{2\Delta p}\right)\right)} - 1\right)w_e. \quad (31)$$

As shown in Figure 16, the SLM width in white dash line separates the pitch-bounded case from the the lens-bounded case (shaded region below dash line); details are in Section B. This limitation is not strict, since holograms are capable of compensating aberrations introduced by optical elements with smaller f-number. A compound microscopic pupil-forming design or a holographic optical element (HOE) based design allows for a higher angular magnification in trade for a smaller eye box and more significant optical aberrations. With such technique, tiling three GAEA SLMs allows for a 66.7° FoV and 35mm eye relief with a 5mm eye box. In summary, a smaller pitch or larger SLM improves the visual experience under spherical illumination.

Comparison with Maimone [2017]. We make many different design and implementation choices on both hardware and software. We require the minimum eye box to be 10mm. This results in a smaller FoV, compared to Maimone et al.'s prototype which maximizes the FoV to 80° at a cost of $\leq 1mm$ eye box. Their software implementation targets real-time performance and assumes constant depth for applying a separable Fast Fourier Transform. We present a Fresnel diffraction integral based implementation that supports per-pixel depth reconstruction and runs interactively (vs. their off-line implementation). The 10mm eye box requires us to render a dense near-eye light field to support intra-ocular occlusion [Zannoli et al. 2016], and our algorithm avoids geometry clipping, under- and over-sampling in the previous work.

8 CONCLUSION AND FUTURE WORK

CGHs using plane wave illumination encounter limitations when applied to near-eye displays; we advocate using spherical wave illumination to break many of these constraints. We show an accurate rendering algorithm by offsetting the camera and using a sliding window to provide a pixel-accurate rendering perspective, and we extend the al gorithm to support spherical wave illumination. Computing the Fresnel diffraction integral on a GPU provides over five orders of magnitude performance improvement compared to the CPU implementation in the previous work, making interactive CGH possible. We derive theoretical resolution limits and trade-off analysis for designing a near-eye display in-depth. In this work, we show the first practical near-eye CGH for high resolution at interactive

rates, modeling intra-pupil occlusion and providing continuous focus cues with per-pixel depth. We believe CGH has the potential to be the ultimate display and we hope to stimulate graphics research incorporating holography. Below we discuss limitations and future work.

Dense Angular Sampling. Holograms are capable of showing highly view-dependent effects like occlusion, specular, or mirror surfaces; these are critical to the depth perception in VR or AR [Toth et al. 2015; Zannoli et al. 2016]. Our rendering considers sparse views at hogel width intervals, and is imprecise for highly specular materials. Spectral analysis on materials could help determine the necessary sampling rate and possibly support non-uniform sampling density and hogel partitioning.

Hardware Accelerating with Tensor Operations. Our software currently computes at 4 FPS. New GPUs have accelerated tensor operations for machine learning. As with light fields [Wetzstein et al. 2012], holographic computation could be substantially accelerated.

Simultaneous Amplitude-Phase Modeling. We use both amplitude and phase SLMs. Alternatively, the double phase methods or super pixel methods enable simultaneous modulation by sacrificing spatial resolution. Incorporating the phase-change within the amplitude SLM [van Putten et al. 2008] may allow for joint optimization.

Wavelength Mixing. Most CGH use field sequential colors, but multiplexing multiple wavelengths has been shown [Glasner et al. 2014; Levy et al. 2001]. Content-adaptive optimization for the entire hologram could increase the frame rate two to three times.

Practical Limitation on FoV. In Section 4 and Section 7.7, we show that the FoV is primarily limited by the spatial extent of the SLM. Similar to tiling SLMs circularly to create a spherical display [Hahn et al. 2008], tiling on a flat plane under spherical wave illumination also increases the field of view. For a Giga-pixel SLMs, bandwidth and physical implementation are engineering challenges. Although smaller pixel pitch also increase FoV, cross-talk between liquid crystals becomes non-negligible and requires further optimization.

Correcting Individual Aberrations. Wearing a corrective lens inside the near-eye display is cumbersome. Holograms and light fields can incorporate personalized inverse aberrations [Huang et al. 2014; Lanman and Luebke 2013; Maimone et al. 2017; Pamplona et al. 2012, 2011]. We envision such system [Padmanaban et al. 2017] will soon be required for a better-than-reality near-eye displays.

ACKNOWLEDGEMENTS

REFERENCES

Lukas Ahrenberg, Philip Benzie, Marcus Magnor, and John Watson. 2008. Computer generated holograms from three dimensional meshes using an analytic light transport model. *Applied optics* 47, 10 (2008), 1567–1574.

Miguel A. Alonso. 2011. Wigner functions in optics: describing beams as ray bundles and pulses as particle ensembles. *Adv. Opt. Photon.* 3, 4 (Dec 2011), 272–365.

HO Bartelt, K-H Brenner, and AW Lohmann. 1980. The Wigner distribution function and its optical production. *Optics Communications* 32, 1 (1980), 32–38.

Martin J Bastiaans. 1997. Application of the Wigner distribution function in optics. *The Wigner Distribution—Theory and Applications in Signal Processing* (1997), 375–426.

Muharrem Bayraktar and Meriç Özcan. 2010. Method to calculate the far field of three-dimensional objects for computer-generated holography. *Appl. Opt.* 49, 24 (Aug 2010), 4647–4654. https://doi.org/10.1364/AO.49.004647

Stephen A. Benton and Jr. Bove, V. Michael. 2008. *Holographic Imaging.* Wiley-Interscience, New York, NY, USA.

William P Bleha and LiJuan Alice Lei. 2013. Advances in liquid crystal on silicon (LCOS) spatial light modulator technology. (2013), 87360A–87360A.

Bryon R Brown and Adolf W Lohmann. 1966. Complex spatial filtering with binary masks. *Applied Optics* 5, 6 (1966), 967–969.

J-S Chen and DP Chu. 2015. Improved layer-based method for rapid hologram generation and real-time interactive holographic display applications. *Optics express* 23, 14 (2015), 18143–18155.

Rick H-Y Chen and Timothy D Wilkinson. 2009. Computer generated hologram from point cloud using graphics processor. *Applied optics* 48, 36 (2009), 6841–6850.

TACM Claasen and WFG Mecklenbräuker. 1980. TIME-FREQUENCY SIGNAL ANALYSIS. *Philips Journal of Research* 35, 6 (1980), 372–389.

Dominic J DeBitetto. 1969. Holographic panoramic stereograms synthesized from white light recordings. *Applied optics* 8, 8 (1969), 1740–1741.

Rainer G Dorsch, Adolf W Lohmann, and Stefan Sinzinger. 1994. Fresnel ping-pong algorithm for two-plane computer-generated hologram display. *Applied optics* 33, 5 (1994), 869–875.

James R Fienup. 1982. Phase retrieval algorithms: a comparison. *Applied optics* 21, 15 (1982), 2758–2769.

Martin Fuchs, Ramesh Raskar, Hans-Peter Seidel, and Hendrik P. A. Lensch. 2008. Towards Passive 6D Reflectance Field Displays. In *ACM SIGGRAPH 2008 Papers (SIGGRAPH '08).* ACM, New York, NY, USA, Article 58, 8 pages.

Ralph W Gerchberg. 1972. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik* 35 (1972), 237.

Daniel Glasner, Todd Zickler, and Anat Levin. 2014. A Reflectance Display. *ACM Trans. Graph.* 33, 4, Article 61 (July 2014), 12 pages. https://doi.org/10.1145/2601097.2601140

Joseph W Goodman. 2005. *Introduction to Fourier optics.* Roberts and Company Publishers.

Joseph W. Goodman. 2014. Holography Viewed from the Perspective of the Light Field Camera. (2014), 3–15. https://doi.org/10.1007/978-3-642-36359-7_1

Joonku Hahn, Hwi Kim, Yongjun Lim, Gilbae Park, and Byoungho Lee. 2008. Wide viewing angle dynamic holographic stereogram with a curved array of spatial light modulators. *Optics express* 16, 16 (2008), 12372–12386.

Fu-Chung Huang, Kevin Chen, and Gordon Wetzstein. 2015. The Light Field Stereoscope: Immersive Computer Graphics via Factored Near-eye Light Field Displays with Focus Cues. *ACM Trans. Graph.* 34, 4, Article 60 (July 2015), 12 pages.

Fu-Chung Huang, Gordon Wetzstein, Brian A. Barsky, and Ramesh Raskar. 2014. Eyeglasses-free Display: Towards Correcting Visual Aberrations with Computational Light Field Displays. *ACM Trans. Graph.* 33, 4, Article 59 (July 2014), 12 pages.

R Jagannathan, R Simon, ECG Sudarshan, and R Vasudevan. 1987. Dynamical maps and nonnegative phase-space distribution functions in quantum mechanics. *Physics Letters A* 120, 4 (1987), 161–164.

Jia Jia, Juan Liu, Guofan Jin, and Yongtian Wang. 2014. Fast and effective occlusion culling for 3D holographic displays by inverse orthographic projection with low angular sampling. *Applied optics* 53, 27 (2014), 6287–6293.

Hoonjong Kang, Elena Stoykova, and Hiroshi Yoshikawa. 2016. Fast phase-added stereogram algorithm for generation of photorealistic 3D content. *Appl. Opt.* 55, 3 (Jan 2016), A135–A143. https://doi.org/10.1364/AO.55.00A135

Hwi Kim, Joonku Hahn, and Byoungho Lee. 2008. Mathematical modeling of triangle-mesh-modeled three-dimensional surface objects for digital holography. *Applied optics* 47, 19 (2008), D117–D127.

Tomasz Kozacki. 2010. On resolution and viewing of holographic image generated by 3D holographic display. *Opt. Express* 18, 26 (Dec 2010), 27118–27129.

Tomasz Kozacki, Małgorzata Kujawińska, Grzegorz Finke, Bryan Hennelly, and Nitesh Pandey. 2012. Extended viewing angle holographic display system with tilted SLMs in a circular configuration. *Applied optics* 51, 11 (2012), 1771–1780.

Douglas Lanman and David Luebke. 2013. Near-eye Light Field Displays. *ACM Trans. Graph.* 32, 6, Article 220 (Nov. 2013), 10 pages.

Seungjae Lee, Changwon Jang, Seokil Moon, Jaebum Cho, and Byoungho Lee. 2016. Additive Light Field Displays: Realization of Augmented Reality with Holographic Optical Elements. *ACM Trans. Graph.* 35, 4, Article 60 (July 2016), 13 pages.

Detlef Leseberg and Christian Frère. 1988. Computer-generated holograms of 3-D objects composed of tilted planar segments. *Appl. Opt.* 27, 14 (Jul 1988), 3020–3024.

LB Lesem, PM Hirsch, and JA Jordan. 1969. The kinoform: a new wavefront reconstruction device. *IBM Journal of Research and Development* 13, 2 (1969), 150–155.

U Levy, E Marom, and D Mendlovic. 2001. Simultaneous multicolor image formation with a single diffractive optical element. *Optics letters* 26, 15 (2001), 1149–1151.

Gang Li, Dukho Lee, Youngmo Jeong, Jaebum Cho, and Byoungho Lee. 2016. Holographic display for see-through augmented reality using mirror-lens holographic optical element. *Opt. Lett.* 41, 11 (Jun 2016), 2486–2489. https://doi.org/10.1364/OL.41.002486

Mark Lucente and Tinsley A. Galyean. 1995. Rendering Interactive Holographic Images. (1995), 387–394. https://doi.org/10.1145/218380.218490

Mark E Lucente. 1993. Interactive computation of holograms using a look-up table. *J. Electronic Imaging* 2, 1 (1993), 28–34.

A. Maimone and H. Fuchs. 2013. Computational augmented reality eyeglasses. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 29–38.

Andrew Maimone, Andreas Georgiou, and Joel Kollin. 2017. Holographic Near-Eye Displays for Virtual and Augmented Reality. *ACM Transactions on Graphics* 36 (July 2017), 85:1–85:16.

Andrew Maimone, Douglas Lanman, Kishore Rathinavel, Kurtis Keller, David Luebke, and Henry Fuchs. 2014. Pinlight Displays: Wide Field of View Augmented Reality Eyeglasses Using Defocused Point Light Sources. *ACM Trans. Graph.* 33, 4, Article 89 (July 2014), 11 pages.

Andrew Maimone, Gordon Wetzstein, Matthew Hirsch, Douglas Lanman, Ramesh Raskar, and Henry Fuchs. 2013. Focus 3D: Compressive Accommodation Display. *ACM Trans. Graph.* 32, 5, Article 153 (Oct. 2013), 13 pages.

MichaĂĔĂĆ Makowski, Maciej Sypek, Andrzej Kolodziejczyk, and Grzegorz MikuĂĔĂĆa. 2005. Three-plane phase-only computer hologram generated with iterative Fresnel algorithm. *Optical Engineering* 44, 12 (2005), 125805–125805–7.

Nobuyuki Masuda, Tomoyoshi Ito, Takashi Tanaka, Atsushi Shiraki, and Takashige Sugie. 2006. Computer generated holography using a graphics processing unit. *Optics Express* 14, 2 (2006), 603–608.

Kyoji Matsushima. 2005. Computer-generated holograms for three-dimensional surface objects with shade and texture. *Applied optics* 44, 22 (2005), 4607–4614.

Kyoji Matsushima and Sumio Nakahara. 2009. Extremely high-definition full-parallax computer-generated hologram created by the polygon-based method. *Applied optics* 48, 34 (2009), H54–H63.

Kyoji Matsushima, Masaki Nakamura, and Sumio Nakahara. 2014. Silhouette method for hidden surface removal in computer holography and its acceleration using the switch-back technique. *Optics express* 22, 20 (2014), 24450–24465.

Rahul Narain, Rachel A. Albert, Abdullah Bulbul, Gregory J. Ward, Martin S. Banks, and James F. O'Brien. 2015. Optimal Presentation of Imagery with Focus Cues on Multi-plane Displays. *ACM Trans. Graph.* 34, 4, Article 59 (July 2015), 12 pages.

N. Padmanaban, R. Konrad, T. Stramer, E.A. Cooper, and G. Wetzstein. 2017. Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. *Proceedings of the National Academy of Sciences* (2017).

Vitor F. Pamplona, Manuel M. Oliveira, Daniel G. Aliaga, and Ramesh Raskar. 2012. Tailored Displays to Compensate for Visual Aberrations. *ACM Trans. Graph.* 31, 4, Article 81 (July 2012), 12 pages.

Vitor F. Pamplona, Erick B. Passos, Jan Zizka, Manuel M. Oliveira, Everett Lawson, Esteban Clua, and Ramesh Raskar. 2011. CATRA: Interactive Measuring and Modeling of Cataracts. In *ACM SIGGRAPH 2011 Papers (SIGGRAPH '11)*. ACM, New York, NY, USA, Article 47, 8 pages.

Christoph Petz and Marcus Magnor. 2003. Fast hologram synthesis for 3D geometry models using graphics hardware. In *Electronic Imaging 2003*. International Society for Optics and Photonics, 266–275.

Weidong Qu, Huarong Gu, Hao Zhang, and Qiaofeng Tan. 2015. Image magnification in lensless holographic projection using double-sampling Fresnel diffraction. *Appl. Opt.* 54, 34 (Dec 2015), 10018–10021.

Stephan Reichelt, Ralf Häussler, Gerald Fütterer, Norbert Leister, Hiromi Kato, Naru Usukura, and Yuuichi Kanbayashi. 2012. Full-range, complex spatial light modulator for real-time holography. *Optics letters* 37, 11 (2012), 1955–1957.

Yusuke Sato and Yuji Sakamoto. 2012. Calculation method for reconstruction at arbitrary depth in CGH with Fourier transform optical system. In *SPIE OPTO*. International Society for Optics and Photonics, 82810W–82810W.

Takanori Senoh, Tomoyuki Mishina, Kenji Yamamoto, Ryutaro Oi, and Taiichiro Kurita. 2011. Viewing-zone-angle-expanded color electronic holography system using ultra-high-definition liquid crystal displays with undesirable light elimination. *Journal of display technology* 7, 7 (2011), 382–390.

Quinn YJ Smithwick, James Barabas, Daniel E Smalley, and V Michael Bove Jr. 2010. Interactive holographic stereograms with accommodation cues. (2010), 761903–761903.

Adrian Stern and Bahram Javidi. 2004. Sampling in the light of Wigner distribution. *J. Opt. Soc. Am. A* 21, 3 (Mar 2004), 360–366. https://doi.org/10.1364/JOSAA.21.000360

Tullio Tommasi and Bruno Bianco. 1993. Computer-generated holograms of tilted planes by a spatial frequency approach. *JOSA A* 10, 2 (1993), 299–305.

Robert Toth, Jon Hasselgren, and Tomas Akenine-Möller. 2015. Perception of Highlight Disparity at a Distance in Consumer Head-mounted Displays. In *Proceedings of the 7th Conference on High-Performance Graphics (HPG '15)*. ACM, New York, NY, USA, 61–66. https://doi.org/10.1145/2790060.2790062

E. G. van Putten, I. M. Vellekoop, and A. P. Mosk. 2008. Spatial amplitude and phase modulation using commercial twisted nematic LCDs. *Appl. Opt.* 47, 12 (Apr 2008), 61–66.

Koki Wakunami, Hiroaki Yamashita, and Masahiro Yamaguchi. 2013. Occlusion culling for computer generated hologram based on ray-wavefront conversion. *Opt. Express* 21, 19 (Sep 2013), 21811–21822. https://doi.org/10.1364/OE.21.021811
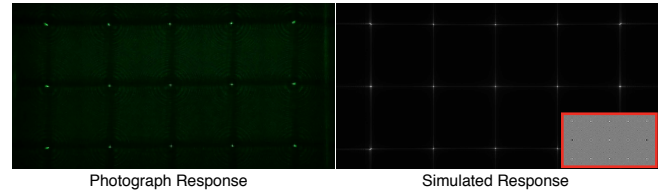
Fig. 17. Photographed aliasing response (left) and simulated aliasing response (right) due to sampling outside the maximum diffraction cone as discussed in Section 7.4. Multiple Fresnel zone plates are presented in the over-sampled diffraction pattern (red inset). In the non-paraxial region, the aliased points no longer reconstructed in a circular shape.

James P Waters. 1966. Holographic image synthesis utilizing theoretical methods. *Applied physics letters* 9, 11 (1966), 405–407.

Gordon Wetzstein, Douglas Lanman, Matthew Hirsch, and Ramesh Raskar. 2012. Tensor Displays: Compressive Light Field Synthesis Using Multilayer Displays with Directional Backlighting. *ACM Trans. Graph.* 31, 4, Article 80 (July 2012), 11 pages.

Masahiro Yamaguchi, Hideshi Hoshino, Toshio Honda, and Nagaaki Ohyama. 1993. Phase-added stereogram: calculation of hologram using computer graphics technique. (1993), 25–31.

Fahri Yaraş, Hoonjong Kang, and Levent Onural. 2011. Circular holographic video display system. *Optics Express* 19, 10 (2011), 9147–9156.

Toyohiko Yatagai. 1976. Stereoscopic approach to 3-D display using computer-generated holograms. *Appl. Opt.* 15, 11 (Nov 1976), 2722–2729.

Genzhi Ye, Sundeep Jolly, V. Michael Bove, Jr., Qionghai Dai, Ramesh Raskar, and Gordon Wetzstein. 2014. Toward BxDF Display Using Multilayer Diffraction. *ACM Trans. Graph.* 33, 6, Article 191 (Nov. 2014), 14 pages.

Marina Zannoli, Gordon D. Love, Rahul Narain, and Martin S. Banks. 2016. Blur and the perception of depth at occlusions. *Journal of Vision* 16, 6 (2016), 17.

Hao Zhang, Qiaofeng Tan, and Guofan Jin. 2012. Holographic display system of a three-dimensional image with distortion-free magnification and zero-order elimination. *Optical Engineering* 51, 7 (2012), 075801–1–075801–5.

Hao Zhang, Yan Zhao, Liangcai Cao, and Guofan Jin. 2015. Fully computed holographic stereogram based algorithm for computer-generated holograms with accurate depth cues. *Opt. Express* 23, 4 (Feb 2015), 3901–3913.

Hao Zhang, Yan Zhao, Liangcai Cao, and Guofan Jin. 2016. Layered holographic stereogram based on inverse Fresnel diffraction. *Applied optics* 55, 3 (2016).

Zhengyun Zhang and Marc Levoy. 2009. Wigner distributions and how they relate to the light field. (April 2009), 1–10.

Yan Zhao, Liangcai Cao, Hao Zhang, Dezhao Kong, and Guofan Jin. 2015. Accurate calculation of computer-generated holograms using angular-spectrum layer-oriented method. *Optics express* 23, 20 (2015), 25440–25449.

Remo Ziegler, Simon Bucheli, Lukas Ahrenberg, Marcus Magnor, and Markus Gross. 2007. A Bidirectional Light Field-Hologram Transform. In *Computer Graphics Forum*, Vol. 26. Wiley Online Library, 435–446.

A FREQUENCY-LIMITED RESOLUTION LIMITS ON DEPTH OF FIELD

For object points within the DoF of the display ($z_0 < z_{min}$), the SLM is wider than the maximum diffraction cone projected from the object point to the SLM. The anti-aliased sampling limits the active sampling area inside the intersection of the projected maximum diffraction cone and the SLM's finite spatial extent

$$x_l = \max\left(x_0 - \frac{z_0\lambda}{2\Delta p}, -\frac{w_s}{2}\right), \quad x_r = \min\left(x_0 + \frac{z_0\lambda}{2\Delta p}, \frac{w_s}{2}\right), \quad (32)$$

as shown in fig. 19 (left). For simplicity, we denote $x_c = (x_l + x_r)/2$ and $w_s' = (x_r - x_l)$. The continuous wavefront is given by

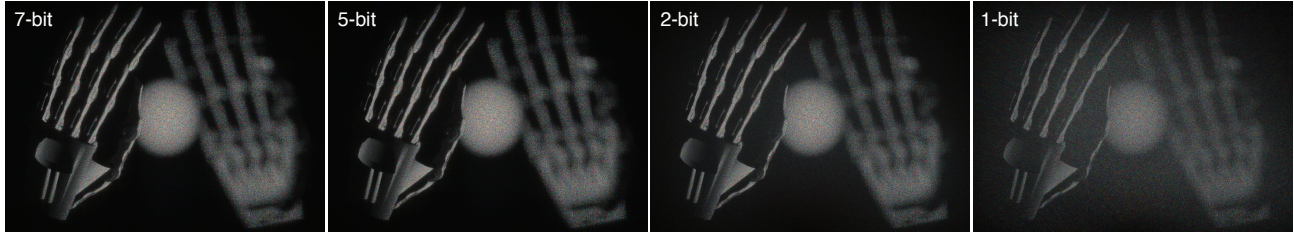$$E_{w_s'}(x,0) = e^{i\pi\frac{(x-x_0)^2}{z_0\lambda}} \sqcap\left(\frac{x - x_c}{w_s'}\right). \quad (33)$$

Fig. 18. Comparison of phase bit levels and simulated scene focused on the front hand from a 1080p, $8um$ pixel pitch, CGH. The image quality presents noticeable degradation below 3-bit; we show the 2-bit result again as a transition stage. In our prototype, the effective phase control is between 6-bit and 7-bit and the degradation with respect to 8-bit is negligible.
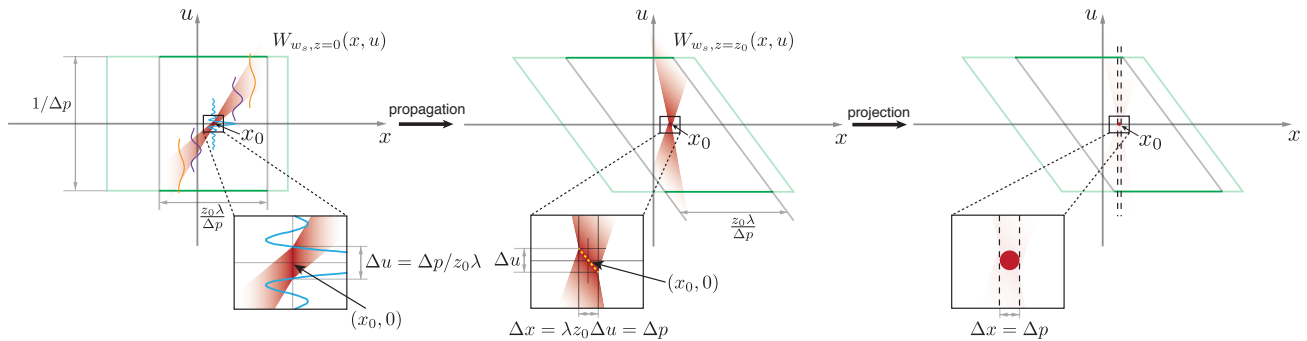

Fig. 19. WDF analysis of resolution limits under frequency-limited sampling inside the DoF. (Left): The active sampling region is limited to the intersection of the projected maximum diffraction cone and the SLM's finite spatial extent. In this example, the projected maximum diffraction cone falls fully inside the SLM. The frequency is minimally dilated to the main lobe of the sinc function at x_0. (Middle): The yellow dash line marks the sheared minimally dilated frequency at the observed distance. The minimum spatial dilation is Δp. (Right): The dash lines bound the region where the sinc response is always positive, and the projected WDF inside this region interferes constructively.
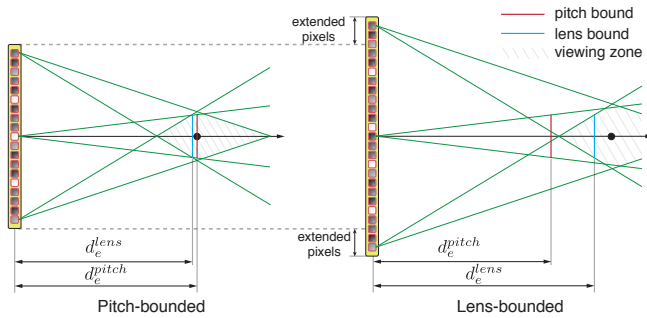

Fig. 20. Illustration of (left): pitch-bounded and (right): lens-bounded display modes. When the spherical illumination is constrained by the $f/\#$ (right), the diffraction cones emitted from the edges of the SLM shift away, and eye relief increases on a wider SLM.

and the corresponding WDF can be derived as

$$
\begin{aligned}
W_{w'_s, z=0}&(x, u) \\
&= 2w'_s \left(1 - \frac{2|x - x_c|}{w'_s}\right) \mathrm{sinc}\left(2w'_s \left(1 - \frac{2|x - x_c|}{w'_s}\right)\left(\frac{x - x_0}{z_0\lambda} - u\right)\right) \\
&= \delta\left(\frac{x - x_0}{z_0\lambda} - u\right) \otimes_u 2w'_s \left(1 - \frac{2|x - x_c|}{w'_s}\right) \mathrm{sinc}\left(2w'_s \left(1 - \frac{2|x - x_c|}{w'_s}\right)u\right),
\end{aligned}
$$

(34)

from which the frequency dilation is characterized as the width of the main lobe

$$
\Delta u = \frac{1}{w'_s \left(1 - \frac{2|x - x_c|}{w'_s}\right)}, \quad x \in (x_l, x_r). \tag{35}
$$

The minimal frequency dilation $\Delta u_{min} = 1/w'_s$ is achieved at $x = x_c$, which depends on both (x_0, z_0) and w_s(fig. 19 left). Propagating the WDF to $z = z_0$ shears the minimal frequency dilation to the minimal spatial dilation

$$
\Delta x_{min} = z_0\lambda\Delta u_{min} = z_0\lambda/w'_s, \tag{36}
$$

which gives the maximum spatial resolution(fig. 19 middle, right). We note that $\Delta x_{min} = \Delta p$ when the active sampling region $w'_s = z_0\frac{\lambda}{\Delta p}$, which is solely limited by the projected maximum diffraction cone (by (x_0, z_0)). This condition is satisfied when

$$
z_0 \leq (w_s - |2x_0|)\Delta p/\lambda. \tag{37}
$$

Beyond this distance, cropping due to the SLM's finite spatial extent reduces spatial resolution (fig. 14 (g)). To conclude, we show an overall resolution distribution as a function of SLM width and pixel pitch in Figure 21.
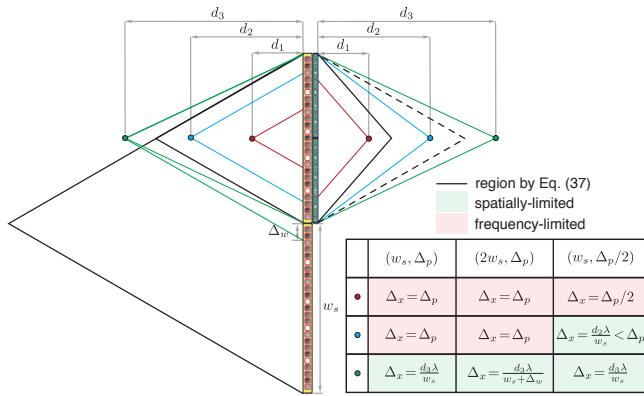
Fig. 21. Illustration of spatial resolution as a function of SLM width and pixel pitch. For $(w_s, \Delta p)$ SLM (yellow top), red and blue point inside the black triangle (region defined by eq. (37)) are reproduced at Δp resolution. For $(w_s, \Delta p/2)$ SLM (dark blue top), resolution inside the region defined by eq. (37) is doubled(red point). Sampling of blue point becomes spatially-limited, however, spatial resolution is still higher than $(w_s, \Delta p)$ case because the active sampling area is the entire SLM. For both cases, the green point subtends to the entire SLM and are reproduced at same spatial resolution. For $(2w_s, \Delta p)$ SLM (yellow top+bottom), the expanded region defined by eq. (37) encapsulates larger space with Δp resolution and produces higher resolution, i.e. green point.

B   LENS-PITCH BOUND DERIVATION

In Section 7.7, we show the FoV and eye relief bounds under spherical wave illumination, and, we contrast two different cases when the maximum FoV and shortest eye relief are bounded by either the pitch pixel (maximum diffraction angle, fig. 20 (left)) or the lens ($f/\#$, fig. 20 (right)). We detail the two cases in the following.

Given a required eye box w_e, the minimum eye relief d_e^{pitch} enforced by the central diffraction cone is solely determined by the pixel pitch:

$$d_e^{pitch} = \frac{w_e}{2\tan\left(\sin^{-1}\left(\frac{\lambda}{2\Delta p}\right)\right)}. \qquad (38)$$

On the other hand, the minimum lens-bounded eye relief d_e^{lens} enforced by the peripheral cones is jointly determined by the SLM width w_s, focal length d_F, and pixel pitch Δp according to Equation (2) and similar triangle relation:

$$d_e^{lens} = \frac{w_s + w_e}{2\tan\left(\sin^{-1}\left(\sin\left(\tan^{-1}\left(\frac{w_s}{2d_F}\right)\right) + \frac{\lambda}{2\Delta p}\right)\right)}. \qquad (39)$$

Equation (39) makes the assumption that the f-number is not bounded by 1. It is easy to show that, under the f-number limitation, if we let $d_F = d_e^{pitch}$ and this focal length d_F is valid ($f/\# \geq 1$), we always have $d_e^{lens} \leq d_e^{pitch}$ under this given d_F (fig. 20 (left)). Thus it is pitch-bounded.

A wider SLM (fig. 20 (right)) with the same pixel pitch has the same pitch-bounded eye relief d_e^{pitch}, but the lens-bounded eye relief extends due to the $f/\# \geq 1$ limitation (fig. 20). Substituting

the $f/\#$ in Equation (39) and equating Equations (38) and (39) gives the maximum pitch-bounded SLM width $\overline{w_s}$ (eq. (31)).

C   MAXIMUM ANGULAR RESOLUTION

The spatial resolution analysis derived in Section 7.5 is based on plane wave illumination, but it can be transformed to spherical wave illumination according to Section 6.2. We consider a magnifying glass placed right after the SLM with focal length d_F and the pupil located at the back focal plane. The scene reconstructed under spherical wave illumination can be considered as a miniature scene lying in between the SLM and the front focal plane reconstructed under plane wave illumination. Let Δx_{min} be the spatial resolution at z_0 in front of SLM under plane wave illumination. When $z_0 \leq d_F$, a virtual image of spatial resolution $\frac{d_F}{d_F - z_0}\Delta x_{min}$ is created at $\frac{d_F}{d_F - z_0}z_0$, and the angular resolution is obtained from the angle subtended by the spatial resolution of the virtual image to the pupil:

$$\Delta\alpha = \left(\frac{d_F}{d_F - z_0}\Delta x_{min}\right) / \left(\frac{d_F z_0}{d_F - z_0} + d_F\right) = \Delta x_{min}/d_F. \qquad (40)$$

When z_0 satisfies eq. (37), we have $\Delta x_{min} = \Delta p$ and the maximum angular resolution is obtained as $\Delta\alpha_{max} = \Delta p/d_F$.