INVITED PAPER

Full-color computational holographic near-eye display

Seyedmahdi Kazempourradi, Erdem Ulusoy and Hakan Urey

Optical Microsystems Laboratory, Koc University, Istanbul, Turkey

ABSTRACT

Near-eye displays (NEDs) are an excellent candidate for the future of augmented reality. Conventional micro-display based NED designs mostly provide stereoscopic 3D experience, which leads to visual discomfort due to vergence-accommodation conflict. Computational holographic near-eye displays (HNEDs) can simultaneously provide wide FOV, retinal resolution, attractive form-factor, and natural depth cues including accommodation. In HNEDs, computer-generated holograms (CGHs) are displayed on a spatial light modulators (SLMs). We propose a CGH computation algorithm that applies to arbitrary paraxial optical architectures; where the SLM illumination beam can be collimated, converging or diverging, and the SLM image as seen by the eye box plane may form at an arbitrary location, and can be virtual or real. Our CGH computation procedure eliminates speckle noise, which is observed in all other laser-based displays, and chromatic aberrations resulting from the light sources and the optics. Our proof-of-concept experiments demonstrate that HNEDs with simple optical architectures can deliver natural 3D images within a wide FOV (70 degrees) at retinal resolution (30 cycles-per-degree), exceeding 4000 resolvable pixels on a line using a printed binary mask. With the advances in SLM technology, HNEDs can realize the ultimate personalized display, meeting the demand of emerging augmented and virtual reality applications.

1. Introduction

Head-worn near-eye displays (NEDs) have promised an unprecedented experience in human–computer interaction with quite exciting augmented reality and virtual reality applications [1,2]. On the other hand, dynamic holographic displays based on the introduction of computer generated holograms (CGHs) in the late 1960s and the subsequent advances in the spatial light modulator (SLM) technologies, have promised the delivery of three-dimensional (3D) images with all natural depth cues [3,4].

Current NEDs are mostly fixed focus stereoscopic displays and present 2D or 3D virtual objects to two eyes, providing a light efficient, immersive and personalized experience. Almost all existing NEDs are designed as standard incoherent imaging systems [5]. For each eye, a flat panel microdisplay is deployed and the visual information is presented in the form of regular 2D images, mostly corresponding to perspective 2D views. This approach leads to two fundamental limitations: relay optics and visual fatigue problems [1,2]. The relay optics problem stems from the fact that human eyes cannot directly focus on a near-to-eye microdisplay since their near accommodation point is around 25 cm. Thus, optical conjugation of the microdisplay to the retina necessitates the placement of additional optical components between the microdisplay and the eye to provide supplementary refractive power. Lightweight designs (those that use single magnifier lenses or waveguides) are restricted to a small field-of-view (FOV), such as 15°–20° in order to limit the optical aberrations. Wide FOV designs require additional corrective components such as multi-part lens systems or free-form optics combiners, which significantly increase the size, complexity and weight of the NED [6–8]. Light field, or integral imaging based NEDs that use micro-lens or pinhole arrays as relay components, as well as NEDs that use directional backlights, have less bulk and wider FOV but also significantly reduced resolution [9–11]. In sum, microdisplay-based NEDs inevitably require a compromise in FOV, resolution and form factor. The visual fatigue limitation stems from the fact that microdisplay-based NEDs can only provide stereoscopic 3D. It is well-established that stereoscopic systems have a fixed focus distance and thus cause visual discomfort, which is triggered by the conflict between the accommodation (change of focal length...
of the eye lens) and convergence (rotation of the eyeballs towards a gazed object) responses [12]. Integral imaging or pinhole based NEDs have slightly reduced discomfort, but again at the cost of resolution [13]. Several solutions which form a multitude of focus distances by moving the microdisplays back and forth or by using variable focal length lenses in time-multiplexed schemes are limited only to small FOVs [14].

On the contrary to stereoscopic displays that provide perspective 2D images, a holographic display generates and presents the viewers directly a portion of the wave field that would emanate from the displayed 3D objects. As a result, viewers see the 3D objects with all natural depth cues. Since the viewers no longer focus on the display panel but rather on the gazed 3D object itself, the vergence-accommodation conflict (VAC) is eliminated. Majority of the proposed holographic display schemes have a table-top television-like configuration and aim to synthesize ghost-like 3D objects floating in mid-air around which several users can move and rotate [15–17]. Though this concept is definitely exciting, it is quite challenging as well. In particular, such schemes require large area SLMs with micron-level pixel pitches, corresponding to enormous SBP requirements that the current state of the SLM technology is far from meeting [4]. Not surprisingly, demonstrations are restricted to quite small objects and narrow viewing zones, with no sign of a significant improvement in the short run. A noteworthy solution is to use eye-tracking proposed by SeeReal Inc. to relieve the space-bandwidth-product (SBP) requirements [18]. SeeReal’s solution does not attempt to deliver the object waves within a large viewing area, but rather provides the object wave merely within two small windows conjugated to and steered with the eyes of the viewer. This way, the spatial bandwidth requirement of the holographic display is significantly reduced. However, the solution still requires a large panel SLM dedicated to a single user. Tracking the eyes of the viewer via cameras located on the distant display panel, which is about an arm’s length from the user, poses another complication. Further, the system allows user motion only within a limited region of space before the display needs to be rotated or translated.

The fundamental limitations associated with NEDs and holographic displays are significantly alleviated when the two domains are unified as holographic near-eye displays (HNEDs). This approach is largely ignored until recently except for a few attempts that do not seem to fully recognize the potential benefits. A holographic head-mounted display with RGB light emitting diode light source is investigated in [19], where the results do not provide a speckle-free full-color reconstruction. A bar-type HNED using two holographic optical elements with a limited FOV is reported in [20]. A full-color HNED is reported in [21], in which random phase assignment on virtual object points lead to computational speckle noise [21].

In this work, we first introduce the HNED basic principles. Then we review CGH computation algorithms and discuss details of our computation procedure to eliminate speckle noise and chromatic aberrations resulting from the light sources and the optics. Finally, we present experimental results for several optical architectures.

2. Holographic near-eye display principles

The benefits of HNEDs are depicted in Figure 1 through the artistic illustration of the ultimate augmented reality HNED we envision. The HNED has a quite thin form factor due to its extremely simple optical architecture, which merely consists of a point light source that illuminates an SLM placed directly in front of the eye. In particular, no relay optics is used. Despite this simplicity, the HNED is able to provide a wide FOV. Further, due to its holographic nature, the HNED provides natural accommodation cues, as emphasized through the provided retinal images where a virtual text object appears in or out of focus along with the real object onto which it is augmented. The HNED is also quite flexible in providing additional functions without any hardware modification, such as the display of the always in-focus virtual object reading ‘Istanbul Tour’.

The possibility of achieving the HNED in Figure 1 is justified by the basic operational principles illustrated in Figure 2.

The point light source generates a narrowband, spatially coherent diverging wave that illuminates a reflective and semi-transparent SLM (they do not yet exist in sufficiently small pixel sizes). The SLM, by virtue of the properly computed CGH loaded on it, acts as an optical mask that transforms the illumination wave directly to the true light wave that would emanate from the virtual objects if they actually existed at their apparent locations. This latter wave, along with the wave emerging from the real objects, goes through the eye pupil and forms an image on the retina. Since the wave generated by the SLM is already shaped to possess the correct ray angles, the eye can readily form a focused image, without requiring the aid of relay lenses. The correct ray angles also lead to correct accommodation cues and remove the visual fatigue intrinsic to all stereoscopic displays. Virtual objects whose waves are delivered within the entire eye box (assumed to fully encapsulate the eye pupil) are viewed at the retinal resolution. Always-in-focus objects can be displayed by utilizing a smaller portion (around 1 mm) of the eye pupil. For full-color applications, a
Figure 1. Artistic illustration of the ultimate augmented reality HNED. The scene shows the Maiden’s Tower and Hagia Sophia as nearby and distant real objects, respectively. The HNED is envisioned to be in the form of an eyeglass and augments virtual objects on the real world view via the CGHs displayed on the reflective and semi-transparent SLMs, which can be fitted into the glass frame and illuminated by the point light sources placed on each side.Insets show the images formed on the retinas of the user. Two of the virtual objects (the large text boxes) appear in or out-of-focus along with the corresponding real objects, while the virtual object reading ‘Istanbul Tour’ is always in-focus. As expected, the eyeglass frame always appears out-of-focus. (Illustrated by Çağlar Genç, Design Lab, Koç University).

separate point light source should be turned on and off for each color channel in a time-sequential manner and the CGH on the SLM should be updated in synchronism. Similarly, functionalities such as compensation for the visual disorders of users can be provided via simple changes in the displayed CGHs, without requiring hardware modification.

From the perspective of holographic displays, Figure 2 illustrates a cost-effective and simple implementation in terms of hardware. The near-eye nature of the display minimizes the SLM size required for a given FOV. The space-bandwidth requirement of the SLM can be minimized by shrinking the eye-box size of the display down to the size of the user’s eye pupil size and employing pupil trackers to handle the eye movements. In this way, the amount of visual information the display delivers at a time, hence the SBP requirement of the SLM, is minimized. Compared to a distant-to-eye display, on a head-mounted system, the pupil trackers can be implemented in a much easier manner via inexpensive miniaturized cameras and infrared light sources. Finally, the system can be supplemented with already-existing head tracking units, also embedded within the display itself. In that case, the display can virtually provide all kinds of depth cues, including motion parallax, without imposing any restriction on the position of the user. The trackers and the associated content update procedures definitely increase the computational complexity of the system, but that is easier to handle and is definitely outweighed by the significantly decreased hardware cost.
3. Review of CGH computation methods

Holographic displays are coherent light processing systems, where visual information is fed via CGHs displayed on SLMs and synthesize wavefronts that are indistinguishable from those emanated by physical objects. A review of recent progress in computer-generated holography for three-dimensional scenes is provided in [22]. CGH computation methods can be categorized by the manner they express 3D objects. In wave-propagation based methods [23–34] objects are represented by point clouds or polygons, whereas in ray-tracing-based methods the 3D model is represented with RGB-Depth (RGBD) or multi-view images [35–39].

In wave-propagation methods, objects are represented by a set of self-emitting point sources (the point-based model) or polygons (mesh model). The optical wave field of a 3D scene is reconstructed by superposing the individual wave fields of primitive structures. Each surface of the 3-D object has a uniform diffusive or specular reflection. Hence the object has to be sampled by an enormous amount of points, usually in million scales. Wave-propagation methods generally suffer from two major problems: (i) the high computational cost of superimposing the wave fields of object points at each pixel of a CGH, (ii) the occlusion problem, i.e. for each viewpoint, occluded primitives of a scene should be identified and their contributions should be removed. If the latter computation is skipped, hidden surfaces become visible and the whole scene becomes transparent in an unnatural fashion.

Lucente proposed the first look-up table (LUT) approach [23], where a set of elementary fringe patterns of uniform sampled points in 3-D image space is pre-calculated and stored in a memory to improve the computational efficiency of point cloud method. Although the LUT successfully accelerates the CGH calculation, it requires large memory allocation. Recently, this large memory requirement has been reduced by several novel LUT (NLUT) methods [24,25]. Recurrence-relation-based CGH calculations have also been proposed in [26–28]. These methods are well-suited to hardware implementation [28].

Point methods suffer from the high-density sampling on the 3D object as well as on hologram plane. Polygon-based methods where the 3D object is represented by polygons, usually triangles, instead of points was proposed as a solution. It is also relatively easy to integrate texture and shading algorithms from computer graphics into polygon methods. Matsushima summarized the whole formulation of the traditional method, analyzed the influence of different interpolation approaches [29], and investigated the surface rendering with texture, shade [30], and specular reflection. For each arbitrary polygon,
an original triangle is depicted in its local coordinates. Then, a fast Fourier transform (FFT) is operated to get a local spectrum of the triangle. The global spectrum of the arbitrary polygon is obtained by remapping the spectrum to global coordinates which are implemented by a two-dimensional (2D) interpolation. A 2D shift and an angular spectrum propagation are performed to calculate the diffraction pattern on the hologram plane. In order to shorten the depicting process of the original polygon, the FFT computation, and the 2D interpolation, several methods were proposed. Kim and Ahrenberg proposed a full analytical polygon-based method where a primitive polygon and its spectrum are analytically obtained. The value of the spectrum at each frequency point of the diffracted wave is analytically derived by means of treating the 3D rotational transformation, 2D origin shift, and the angular spectrum propagation as an integrated step [31,32]. Liu reported a full analytical method by use of Fraunhofer diffraction to represent waves in the Fresnel region [33]. Polygon-based methods are advantageous in less amount of sampler compare to point methods. However, the FFT and 2D interpolation are heavy computational processes. Even though several improved polygon methods have been proposed, they either sacrificed texture capability or brought extra computation. There has been no real-time polygon-based method so far [34].

In ray-tracing-based methods, CGHs are computed by processing 3D scenes with computer graphics methods. In particular, for each viewpoint, computer graphics techniques are utilized to determine perspective views including occlusion effects in a fast and efficient manner. Unlike point cloud and polygon approaches, this approach requires neither hidden surface removal nor shading processing because both are automatically performed by the CG technique. Moreover, this approach readily manages natural 3D scenes and 3D computer graphics. However, in most implementations, the final CGH is formed merely as a collection of the 2D perspective views, losing the accommodation cue. Holographic stereograms and integral imaging based displays are well-known implementations of this type. Holographic stereograms are composites of elemental holograms generated from 2D images by FFT [35]. The computational cost of holographic stereograms is much lower than that of the point cloud and polygon approaches. On the other hand, although an observer can recognize a 3-D image reconstructed from elemental holograms, the quality of deep depth images is problematic. This is because of the fact that the reconstructions are essentially 2D images. To improve this method, researchers have proposed phase-added stereogram and accurate phase-added stereogram approaches [36,37]. Computational costs of original and modified holographic stereogram are almost identical. In such displays, even though binocular disparity and motion parallax cues may be delivered correctly along with appropriate occlusion effects, all objects appear sharp only when the viewer focuses at a single depth (usually the hologram plane), and apparently, out of plane objects appear blurred when focused on. As a result, such implementations are prone to the well-known vergence-accommodation conflict typical of stereoscopic displays. A method for computing realistic CGHs of 3D objects is presented, where the rendered view and the depth-map of the scene are used to handle occlusion, shading and parallax effects and avoid VAC [38].

All of the discussed methods compute the full-complex CGH that represents the virtual object wave within the eye box of a holographic display. However, additional encoding steps are required due to the fact that SLMs do not perform full complex but some restricted type of modulation, such as amplitude or phase only, or binary. Direct binary search, error diffusion, iterative Fourier transform algorithm are well-known procedures for encoding full complex CGHs into amplitude, phase or binary CGHs [39–41]. These algorithms perform successful encodings, yet, almost all of them have an iterative structure and involve repetitive FFT or pixel-by-pixel processing operations that pose significant computational complexity. A non-iterative algorithm, where simple Discrete Fourier Transform (DFT) relations are exploited to compute phase CGHs that exactly control half of the desired image samples via a single FFT is presented in [42].

4. CGH computation for HNEDs using scalar wave optics and paraxial approximation

In our work, CGHs are computed using scalar wave optics theory of light and paraxial approximation. Given a 3D scene, we utilize a graphical processing unit (GPU) and computer graphics methods to determine perspective images for a set of viewpoints. In addition, we extract and use the depth map information that is already computed during rendering calculations. Based on the depth map, we divide the scene into a number of slices, each corresponding to a specific depth. Then, for each slice, we compute an intermediate field using an FFT based Fresnel transform that represents the contribution of the depth slice to the field on the SLM plane. The total field on the SLM plane for a viewpoint is obtained by superposing the intermediate fields of depth slices. Afterward, linear and quadratic phase terms are superimposed to handle the light source and eye box positions. The proposed method computes a CGH with correct occlusion and accommodation effects in a fast manner. For phase-only SLM based experiments reported in this paper, the
resulting complex function is encoded into a phase-only pattern using the well-known iterative Fourier transform algorithm (IFTA) with around 10 iterations [39,40]. For the fabricated binary mask, the same procedure is used except that encoding is performed using the error diffusion algorithm [41].

Our computational procedure is also optimized for speckle-free image formation [43]. Our methodology starts with the observation that human eye images a point object not as a point but as an Airy disk on the retina, due to diffraction from the eye pupil. In case of an extended object illuminated with coherent light, the individual Airy disks of close-by object points interfere with each other. If the object is a typically physically existing real object, the granular surface structure leads to a random phase distribution on the object points. Hence, the individual Airy disks on the retina have random phase values as well, and the resulting interference pattern consists of rapidly varying destructive and constructive zones, which are seen to a viewer as speckle noise. However, in the case of a virtual object, there is total freedom in assigning a phase value to the object points. In particular, if the object points are assigned an initial phase value such that all the Airy disks arrive at the retina with the same phase, the retinal intensity distribution becomes smoothly interpolated and speckle noise is essentially eliminated. The resulting images appear closer to images that would be formed with incoherent light.

An HNED architecture consists of a point light source, an SLM, an eye box plane, and optical components in between. It is well known that within the realm of paraxial optics, any optical system can be modeled as a combination of a free-space propagation, a thin lens and another free-space propagation [44]. In this respect, a generic paraxial HNED architecture has the form depicted in Figure 3, which is the architecture we consider in this paper. As illustrated, light is emitted by a point source located at \( x_0 \) and propagates for a distance of \( d_{IL} \). The light passes through the illumination side lens with a focal length of \( f_{IL} \) and propagates for a distance of \( d_{IS} \) where it impinges on the SLM. The light gets modulated upon passage through the SLM, which is modeled as a thin multiplicative optical element. The rays propagate a distance of \( d_{SE} \) towards the eye-piece lens with a focal length of \( f_E \), followed by another propagation for a distance of \( d_{EP} \) after which they reach the pupil plane, where an eye box is formed around \( x_p \). The image replicas of the point source are formed on the pupil plane because in most SLMs, part of the incident light does not get modulated. Moreover, the pixelated structure of the SLM creates diffraction orders of both modulated and non-modulated beams. We assume in this paper that the eye box fits in the free area between the zeroth and first orders of the blue non-modulated beams, since the separation between two consecutive diffraction orders is minimum for blue wavelength.

We assume that full-color holograms are reconstructed by using the SLM in color field sequential mode. It is crucial to use the correct paraxial model parameters for each wavelength in order to eliminate chromatic aberrations. If not treated properly, chromatic aberrations result in lateral and axial shifts and size mismatches among color components. Actual lenses usually have slightly varying focal lengths and principal planes for different wavelengths, which should be taken into account.

Figure 4 summarizes all the steps of our CGH computation process. Our method starts with a 3D intensity image \( I(x, y, z) \) (a 2D image and a depth map) rendered for a single viewpoint \( (x_p, \text{the center of the eye box}) \). After taking the square root of the 3D intensity image \( I(x, y, z) \), the resulting 3D field amplitude \( E(x, y, z) \) is sliced into \( N \) planes (i.e. \( z_1, z_2, \ldots, z_N \)). The field amplitude at a distance of \( z_i \) (i.e. \( E(x, y, z_i) \)) is convolved with a lens term of focal length \( f_{HI,i} \) given as:

\[
f_{HI,i} = \frac{(f_E d_{EP} + d_{SE}) - d_{EP} d_{SE}}{z_i f_E^2}. \tag{1}
\]

In other words, for each field slice, we compute an intermediate field using an FFT based free-space propagation for a distance of \(-f_{HI}\). The intermediate fields of depth slices are then superimposed. The resulting full-complex hologram is encoded into a phase-only CGH. An option is to simply discard the magnitude information of the field. However, this will not result in a high-quality hologram. Hence, in this step, we use IFTA to compute phase-only CGHs. Finally, global tilt and lens terms are superimposed on the phase CGH to handle light source and eye box locations. The global lens term has a focal length of

\[
f_G = \frac{d_{SE}(f_E - d_{EP}) + d_{EP} f_E}{f_E - d_{EP}} \tag{2}
\]

and the global tilt term has an angle (in radians) of:

\[
\theta_G = \frac{x_p f_E}{f_E d_{SE} + f_E d_{EP} - d_{SE} d_{EP}}. \tag{3}
\]

The illumination lens term has a focal length of

\[
f_i = \frac{d_{IL}(f_{IL} - d_{IS}) + d_{IS} f_{IL}}{f_{IL} - d_{IS}} \tag{4}
\]

and the illumination tilt term has an angle (in radians) of

\[
\theta_i = \frac{x_p f_{IL}}{f_{IL} d_{IL} + f_{IL} d_{IS} - d_{IL} d_{IS}}. \tag{5}
\]

The physical interpretations of the lens and tilt terms defined above are as follows: the illumination lens and tilt
Figure 3. General schematic of the HNED. The light emits from the point source located at $x_s$ and propagates for a distance of $d_{L1}$. The light passes through the illumination side lens with a focal length of $f_{IL}$ and propagates for a distance of $d_{IL}$, as illustrated in Figure 3. The light gets modulated after passing the SLM, which is modeled as a thin multiplicative optical element. The rays propagate a distance of $d_{SE}$ toward the eye-piece lens with a focal length of $f_E$. The rays propagate for a distance of $d_{EP}$ and reach the pupil plane located at $x_p$ after passing the eye-piece lens. The image replicas of the point source are formed on the pupil plane when part of the incident light on SLM does not get modulated.

Figure 4. Block diagram of the phase-only CGH computation algorithm. For a binary amplitude CGH, the same procedure is used except that encoding is performed using the error diffusion algorithm instead of IFTA.

terms effectively convert the incoming illumination wave into a normally incident collimated plane wave. This normally incident plane wave is imaged to the center of the eye box by the global lens and tilt terms. The lens term $f_{H,i}$ is added to the CGH so that instead of the pupil plane, the point light source is imaged on a plane at a distance $z_i$ to the left of the pupil plane, forming a virtual object point there.

5. Experimental results and discussion

In order to verify the premises of the HNED concept depicted in Figures 1 and 2, we performed proof-of-concept experiments for the diverging beam and then converging beam illumination cases.

5.1. Experiment 1: diverging beam illumination using SLM

The experimental setup is depicted in Figure 5(a). The SLM is a reflective LCoS (liquid crystal on silicon) SLM (Holoeye Pluto-VIS, 8-bit phase-only modulation in the 0–2π range, 1920 × 1080 pixels with a pixel pitch of 8 μm). A Gaussian beam diverging from the tip of a single mode fiber (coupled to a He–Ne laser source with wavelength 632.9 nm) constitutes the point light source that illuminates the SLM. Since the SLM is not transparent, a beam splitter is used to combine the coherent light modulated and reflected by the SLM with the incoherent light from real objects (the two dice placed, respectively, 250 and 500 mm away from the eye). Other than this, the architecture is the same as in Figure 2. The combined light is captured by a camera in front of which an aperture with a diameter of 3 mm is placed to mimic the eye pupil. The distance between the aperture and the SLM is adjusted to 60 mm, and the distance between the point light source and SLM to 250 mm. The effective eye-box in this experiment is 4 mm × 4 mm square. The system has a diagonal FOV of 16.7 degrees. A single phase only CGH is computed for two virtual text objects designed to appear next to each dice. Figure 5(b,c) show the two images captured by the camera when the focus is adjusted.
Figure 5. (a) Setup for the first experiment (not-to-scale). The eye to SLM distance is 60 mm, while the point light source to SLM distance is 250 mm. The nearby and distant dices are at 250 and 500 mm away from the eye, respectively. Both dices have a side length of 14 mm. A 3 mm aperture is placed in front of the camera to mimic the eye pupil; (b,c) Image captured by a camera when the focus is adjusted to 250 and 500 mm, respectively. The experiment confirms that 3D images with correct focus cues are displayed without relay lenses at the retinal resolution. However, the images are degraded by the bright spots created by the un-modulated beam of the SLM and the intensity variation due to the non-uniform directivity of SLM pixels.

on each of the two dice. The results clearly confirm that the SLM placed much closer than the near point of the human eye can deliver focused retinal images despite the absence of relay lenses. Moreover, the system delivers natural accommodation cues and virtual objects appear in focus only when the eye focuses at the true depths. In order not to saturate the camera, we limited the laser power emanating from the single mode fiber to $< 20 \mu W$, and the laser power captured by the camera within the eye box is about 1 $\mu W$, which is already quite bright for the eye. This shows that the system is highly efficient in delivering light to the user’s eyes, verifying that low-power HNEDs using low-power laser light sources is possible.

Besides verifying the premises of the HNED approach, the first experiment also demonstrates two technological limitations associated with currently existing SLMs [45]: (i) part of the light incident on the SLM remains un-modulated, forming mirror images of the point light source that are also intercepted by the camera aperture and degrade the final image (notice the bright spots in Figure 5(b,c)), (ii) the images have a noticeable intensity variation over the FOV, with dark bands near the edges. For the diverging beam case, Fourier transform of the pixel aperture results in $\text{Sinc}^2$ intensity profile across the reconstructed holographic image. This can be reduced by introducing an inverse brightness variation across the original image or by limiting the SLM pixel reflector areas. The first problem is due to the fact that inter-pixel gaps are not perfect absorbers/, while the second problem arises since SLM pixels have limited diffraction spread around specular reflections. Both of these problems will be relieved when SLMs with smaller pixels become available.

The most important shortfall of the demonstrations in Figure 5 is that the eye relief distance cannot be made shorter than about 60 mm (which is slightly impractical for near-to-eye configurations) due to the fact that the pixel density of the SLM used in the experiments is not sufficient to support closer distances. For closer distances, the supportable eye box size becomes smaller than the eye pupil, and undesired quantization noise, un-modulated and higher order replica beams also enter the eye, corrupting the retinal image. Fortunately, LCOS based SLMs with pixel pitches of 3.74 $\mu m$ have
Figure 6. Results of the second experiment: (a) 20 mm by 20 mm chrome on glass binary mask placed 20 mm away from the camera of a mobile phone; (b) The entire image is seen by the camera. The diagonal FOV is 70 degrees. More than 4000 by 4000 pixels are in the FOV; (c,d) Images at different zoom levels. The smallest resolvable text corresponds to 30 cycles per degree.

Figure 7. (a) Setup for the third and fourth experiment (not-to-scale). A plano-convex lens with a focal length of 60 mm is placed in front of the SLM. Point light source is located at 60 mm away from the SLM. Other distances are the same as the first experiment; (b,c) New images captured by the camera. The artifacts in the first experiment are successfully removed.
Figure 8. (a) Computer generated scene using Blender 3D modeling software; (b) The depth map of the scene; (c) The slice of the scene corresponding to 4 diopters (25 cm); (d) The slice of the scene corresponding to 3 diopters (33 cm); (e) The slice of the scene corresponding to 2 diopters (50 cm); (f) The slice of the scene corresponding to 1 diopter (100 cm).

already been demonstrated and are commercially available [46,47]. The eye relief distance can then be reduced to about 30 mm.

5.2. Experiment 2: diverging beam illumination using binary mask

To demonstrate the full potential of HNEDs, we present a second experiment with a binary CGH printed on chrome on glass mask with a minimum feature size of 2 μm and overall size of 20 mm by 20 mm. The mask is designed to operate at a practical eye relief distance of 20 mm, corresponding to a diagonal FOV of 70°. The effective eye-box in this experiment is 2.5 mm × 2.5 mm square. A mobile phone camera is placed at the position of the eye pupil, as in Figure 6(a). Figure 6(b) shows the wide FOV image captured by the camera, while the images in Figure 6(c,d) illustrate the details at different zoom levels. The smallest fonts correspond to 30 cycles per degree resolution, and the overall image contains more than 4000 by 4000 pixels. The image can also be viewed with the naked eye with high quality.

5.3. Experiment 3: converging beam illumination using SLM

Alternatively, the HNED architecture can be modified as in the third experiment depicted in Figure 7. The mere modification is that a plano-convex lens with a focal length of 60 mm is attached to the reflective SLM, and the point light source to SLM distance is adjusted to 60 mm so that the point light source is imaged on the eye box plane. The effective eye-box in this experiment is 4 mm × 4 mm. The system has a diagonal FOV of 16.7 degrees. A crucial point to recognize here is that the lens is not intended to function as a relay lens, but its mere purposes are (i) to concentrate the un-modulated beam to a region outside the eye box (so that it can be filtered out by the eye pupil), and (ii) to rotate the diffraction cones of SLM pixels so that the eye box receives uniform power from each pixel. Therefore, unlike the well-corrected relay lenses used in microdisplay-based approaches, the lens need not be free of aberrations and can be a simple single piece lens. Lens aberrations can be handled easily by the CGH computation procedure. As seen, the degradation due to the un-modulated beam and the intensity variation are removed.

5.4. Reconstruction of holograms with 4 depth plane

In this experiment, we used a He–Ne red laser of wavelength 632 to create a CGH of a 3D computer generated scene created from Blender 3D.
Figure 9. The experimental captured holograms: (a) The focus of camera is set to 4 diopters (25 cm); (b) The focus of camera is set to 3 diopters (33 cm); (c) The focus of camera is set to 2 diopters (50 cm); (d) The focus of camera is set to 1 diopter (100 cm).

Figure 10. (a) The speckle corrupted image of a virtual object with a random phase distribution; (b) The speckle-free image of a virtual object with an optimized phase distribution computed using our algorithm, verifying that the speckle artifact present in all laser displays can be eliminated in holographic HWDs using computational methods.

As shown in Figure 8(a), the scene in Blender consists of a meshed bowling ball, a meshed bowling bottle together with a ‘Bowling’ text on top of it, a ‘Koc University, OML’ text and a soccer ball with a ‘Soccer’ text. These objects are respectively placed at 25 cm, 33 cm, 50 cm and 1 m (1 diopter separation in between) distances from the eye plane. The depth map of the scene is shown in Figure 8(b). The depth map (z-buffer) data of the scene helps us to extract perspective view and the 3D position of the object. The depth range is divided into 4 planes and the slices corresponding to the object planes are provided in Figure 8(c–f), respectively.

Figure 9 illustrates the experimental images taken from the reconstructed hologram. Figures 9(a–d) demonstrate the reconstructed images where the focus of camera is changed from the first object to the last object,
ranging from 25 to 100 cm, respectively. The results verify the method and show that the proposed technique is capable to handle both accommodation and occlusion cues. The effective eye-box in this experiment is \(4 \text{ mm} \times 4 \text{ mm}\). The system has a diagonal FOV of 16.7 degrees.

5.5. Converging beam illumination using SLM and speckle reduction

Use of coherent laser light source might be seen as a fundamental limitation of the HNED approach since laser-based displays suffer from speckle noise [48]. Firstly, we note that the system does not require a perfectly monochromatic illumination. It can work under LED illumination, just as many other holographic displays that have been demonstrated in the literature, provided that LED source is partially coherent, i.e. sufficiently narrowband and has small emission area. Indeed, the relatively low coherency of LED sources helps wash out the speckle patterns, though at the cost of slight loss in resolution. Secondly, using our computational methodology discussed in the previous section, it is possible to obtain speckle-free and natural looking retinal images even under laser illumination, as can already be verified from the images presented so far. An experimental result comparatively illustrating this point is provided in Figure 10(a,b). The image in Figure 10(a) suffers from speckle noise as a result of having a random phase distribution assigned to the virtual object points, which is the main practice in commonly used iterative CGH algorithms. On the other hand, the image in Figure 10(b) is almost free of speckle due to the assignment of an optimized phase distribution calculated using our methodology. However, we point that a minor discrepancy still exists between the predicted and the achieved images, which can be attributed to the crosstalk of SLM pixels and other small artifacts from the optical components.
5.6. Experiment 4: full color with aberration correction using SLM

As the last set of experiments, we captured a full-color hologram using the same setup illustrated in Figure 7(a). The effective eye-box in this experiment is 3 mm x 3 mm (reduced due to wavelength scaling) and the diagonal FOV is 16.7 degrees. A time-division multiplexing method is used to reconstruct the full-color holograms. Red, green and blue lasers are coupled into a single mode fiber. We used an Arduino to sequentially switch the lasers on and off in synchrony with the corresponding trigger signal from the SLM. The individual images of each red, green and blue holograms are captured using a 5 Mega-pixels color JAI camera. We combined the captured images into a single RGB image using MATLAB, as shown in Figure 11. The optical reconstruction of CGHs is based on the diffraction theory and the use of different wavelengths results in chromatic aberration. Without considering the correct compensation factors the reconstructed holographic images do not overlap to form a full-color hologram. One of the main sources of the chromatic aberration is neglecting the change of refractive indices of lenses with respect to the wavelength, which can be compensated using the thick lens approximation in the CGH computation algorithm. Figure 11(a,b) demonstrates the results when we do not use the thick lens approximations in our paraxial CGH computation. As evident in the magnified views, the colors are separated. In Figure 11(a) the camera is focused at 25 cm, wherein Figure 11(b) the focus of the camera is set to 100 cm.

Figure 11(c,d) illustrates the correct full-color CGH reconstruction experimental results after applying proper corrections. In Figure 11(c) the camera is focused at 25 cm and the magnified views show the correct overlap. Similarly, color-overlap is evident in Figure 11(d), where the camera is focused at 100 cm.

6. Conclusions

Our work shows and demonstrates that breakthrough achievements can be obtained through the unification of the near-eye and holographic display domains: On one side, NEDs, which have suffered from bulkiness and created visual discomfort, can become impressively compact to achieve eye-glass form factors while providing quite wide FOVs, retinal resolution and natural 3D. Here, holography helps the NED domain by removing the necessity of using relay optics and by bringing natural accommodation cues. On the other side, holographic displays, which have been believed to be a dream that is many years away due to the enormous space-bandwidth product (SBP) requirements, become realizable even with today’s technology. Here, the NED domain helps holography by minimizing the eye box size and parallax that needs to be delivered to a user at one time. The HNED configuration also provides the most convenient setting for pupil and head trackers and provides significantly enhanced light efficiency and motion freedom.

The current state of the SLM technology impedes the enjoyment of the full potential of holographic displays. The non-transparent nature of LCoS SLMs prohibits the placement of the SLM directly in front of the eye in augmented reality applications. Unfortunately, currently available transmissive SLMs have pixel pitches around 20–40 μm, which is too large for near-to-eye configurations. Even if the pixel pitches of transmissive SLMs shrink down to 5–10 microns, the real world light will also get modulated by the CGH while passing through it, which may create visual artefacts. However, if the SLM pixel pitch decreases to about 1 μm, perhaps via several newly emerging technologies, the above problem can be avoided. Large number of pixels and larger eye box size of 10–15 mm can be supported and pupil trackers can be avoided as well. Though such developments will definitely take some time, much faster progress can be possible if the display industry focus on small-pixel SLMs with low pixel crosstalk. The high computational cost of CGHs pose yet another challenge, especially when integration with pupil and head trackers is taken into account. However, we believe that once the technological issues are overcome, the unification will realize the ultimate interactive personalized display of all times.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work is supported by European Research Council (ERC) under the Seventh Framework Programme [grant number FP7/2007-2013] ERC advanced grant agreement [grant number 340200], with acronym Wear3D.

Notes on contributors

Seyedmahdi Kazempouradi received the B.Sc. degree from Sharif University of Technology, Tehran, Iran, in 2011 and M.Sc. from Bilkent University, Ankara, Turkey, in 2014, all in Electrical Engineering. His master thesis includes developing algebraic acceleration techniques and algorithms for the solutions of electromagnetics problems, integral equations, iterative methods, and high-performance computing. He joined Optical Microsystems Laboratory (OML) at Koç University in 2014 as a Ph.D. student. He worked in the prestigious Wear3D project funded by the
European Research Council Advanced Grant (ERC-AdG) to develop next generation wearable and 3D display technologies, during his Ph.D. studies.

**Erdem Ulusoy** received his BS and PhD degrees both from Bilkent University, Department of Electrical and Electronics Engineering, in 2004 and 2012, respectively. During his PhD, he was a member of the EC funded 3DTV Network of Excellence project. His research interests include Fourier Optics, Computer Generated Holography, Light Field Synthesis with Spatial Light Modulators and Signal and Image Processing Problems in Diffraction and Holography.

**Hakan Urey** received the B.Sc. degree from Middle East Technical University, Ankara in 1992, and M.Sc. and Ph.D. degrees from Georgia Institute of Technology in 1996 and in 1997, all in Electrical Engineering. After completing his Ph.D., he joined Microvision Inc.-Seattle as a Research Engineer and he played a key role in the development of the Retinal Scanning Display technology. He was the Principal System Engineer when he left Microvision to join the faculty of engineering at Koç University in 2001. He was promoted to Associate Professor in 2007 and Professor in 2010. He published about 70 journals and 100 international conference papers, 7 edited books, 4 book chapters, and has more than 50 issued and several pending patents. His research interests are in the area of optical MEMS, micro-optics and optical system design, 2D/3D display and imaging systems and biosensors. He is a member of SPIE, IEEE, and OSA. He received an advanced grant from the European Research Council (ERC-AdG) in 2013, Outstanding Faculty Award from Koç University in 2013, TÜBİTAK-Encouragement Award in 2009, Outstanding Young Scientist Award from the Turkish Academy of Sciences (TOBA) in 2007 and Werner Von Siemens Excellence Award in 2006.

**ORCID**

Seyyedmahdi Kazempourradi  [http://orcid.org/0000-0002-4583-1452](http://orcid.org/0000-0002-4583-1452)

**References**

[1] J. Rolland, K.P. Thompson, H. Urey and M. Thomas, See-through Head Worn Display (HWD) Architectures, in *Handbook of Visual Display Technology* (Springer Berlin Heidelberg, 2012), pp. 2145–2170.

[2] J. Rolland and O. Cakmakci, Head-worn Displays: The Future through New Eyes, Opt. Photonics News 20 (4), 20–27 (2009).

[3] C. Slinger, C. Cameron and M. Stanley, Computer-generated Holography as a Generic Display Technology, Computer 38, 46–53 (2005).

[4] F. Yaras, H. Kang and L. Onural, State of the Art in Holographic Displays: A Survey, J. Disp. Technol. 6 (10), 443–454 (2010).

[5] B. Kress and T. Starner, A Review of Head-mounted Displays (HMD) Technologies and Applications for Consumer Electronics, in *SPIE Defense, Security, and Sensing* (International Society for Optics and Photonics, Baltimore, Maryland, United States, 2013), pp. 87200A–87200A.

[6] D. Cheng, Y. Wang, H. Hua and M.M. Talha, Design of an Optical See-through Head-mounted Display with a Low F-number and Large Field of View Using a Freeform Prism, Appl. Opt. 48 (14), 2655–2668 (2009).

[7] D. Cheng, Y. Wang, H. Hua and J. Sasian, Design of a Wide-angle, Lightweight Head-mounted Display Using Free-form Optics Tiling, Opt. Lett. 36 (11), 2098–2100 (2011).

[8] K.P. Thompson and J. Rolland, Freeform Optical Surfaces: A Revolution in Imaging Optical Design, Opt. Photonics News 23 (6), 30–35 (2012).

[9] H. Hua and B. Javid, A 3D Integral Imaging Optical See-through Head-mounted Display, Opt. Express 22 (11), 13484–13491 (2014).

[10] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke and H. Fuchs, Pinlight Displays: Wide Field of View Augmented Reality Eyeglasses Using Defocused Point Light Sources, in *ACM SIGGRAPH 2014 Emerging Technologies* (ACM, Vancouver Canada, 2014), p. 20.

[11] A. Maimone and H. Fuchs, Computational Augmented Reality Eyeglasses, in *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on* (IEEE, Adelaide, 2013), pp. 29–38.

[12] H. Urey, K.V. Chellappan, E. Erden and P. Surman, State of the Art in Stereoscopic and Autostereoscopic Displays, Proc. IEEE 99 (4), 540–555 (2011).

[13] H. Hua and B. Javid, Augmented Reality: Easy on the Eyes, Opt. Photonics News 26 (2), 26–33 (2015).

[14] S. Liu and H. Hua, Time-multiplexed Dual-focal Plane Head-mounted Display with a Liquid Lens, Opt. Lett. 34 (11), 1642–1644 (2009).

[15] F. Yaraş, H. Kang and L. Onural, Real-time Phase-only Color Holographic Video Display System Using LED Illumination, Appl. Opt. 48 (34), H48–H53 (2009).

[16] F. Yaraş, H. Kang and L. Onural, Circular Holographic Video Display System, Opt. Express 19 (10), 9147–9156 (2011).

[17] H. Sasaki, K. Yamamoto, K. Wakunami, Y. Ichihashi, R. Oi and T. Senoh, Large Size Three-dimensional Video by Electronic Holography Using Multiple Spatial Light Modulators, Sci. Rep. 4, 61771–61778 (2014).

[18] R. Häussler, A. Schwirddner and N. Leister, Large Holographic Displays as Alternative to Stereoscopic Displays, in *Electronic Imaging 2008* (International Society for Optics and Photonics, San Jose, California, United States, 2008), pp. 68030M–68030M.

[19] E. Moon, K. Myeongiae, R. Jinyoung, H. Kim and J. Hahn, Holographic Head-mounted Display with RGB Light Emitting Diode Light Source, Opt. Express 22 (6), 6526–6534 (2014).

[20] H.-J. Yeom, H.-J. Kim, S.-B. Kim, H. Zhang, B. Li, Y.-M. Ji, S.-H. Kim and J.-H. Park, 3D Holographic Head Mounted Display Using Holographic Optical Elements with Astigmatism Aberration Compensation, Opt. Express 23 (25), 32025–32034 (2015).

[21] A. Mainmone, A. Georgiou and J. Kollin, Holographic Near-Eye Displays for Virtual and Augmented Reality, ACM Trans. Graph 36 (4), 85 (2017).
[22] J.-H. Park, Recent Progress in Computer-generated Holography for Three-dimensional Scenes, J. Inf. Disp. 18 (1), 1–12 (2017).

[23] M.E. Lucente, Interactive Computation of Holograms Using a Look-up Table, J. Electron. Imag. 2 (1), 28–34 (1993).

[24] S.-C. Kim and E.-S. Kim, Effective Generation of Digital Holograms of Three-dimensional Objects Using a Novel Look-up Table Method, Appl. Opt. 47 (19), D55–D62 (2008).

[25] S.-C. Kim, J.-M. Kim and E.-S. Kim, Effective Memory Reduction of the Novel Look-up Table with One-dimensional Sub-principle Fringe Patterns in Computer-generated Holograms, Opt. Express 20 (11), 12021–12034 (2012).

[26] K. Matsushima and M. Takai, Recurrence Formulas for Fast Creation of Synthetic Three-dimensional Holograms, Appl. Opt. 39 (35), 6587–6594 (2000).

[27] H. Yoshikawa, Fast Computation of Fresnel Holograms Employing Difference, Opt. Rev. 8 (5), 331–335 (2001).

[28] T. Shimobaba, S. Hishinuma and T. Ito, Special-purpose Computer for Holography Horn-4 with Recurrence Algorithm, Comput. Phys. Commun. 148 (2), 160–170 (2002).

[29] K. Matsushima, H. Schimmel and F. Wyrowski, Fast Calculation Method for Optical Diffraction on Tilted Planes by Use of the Angular Spectrum of Plane Waves, J. Opt. Soc. Amer. A 20 (9), 1755–1762 (2003).

[30] K. Matsushima, Computer-generated Holograms for Three-dimensional Surface Objects with Shade and Texture, Appl. Opt. 44 (22), 4607–4614 (2005).

[31] L. Ahrenberg, P. Benzie, M. Magnor and J. Watson, Computer Generated Holograms from Three Dimensional Meshes Using an Analytic Light Transport Model, Appl. Opt. 47 (10), 1567–1574 (2008).

[32] H. Kim, J. Hahn and B. Lee, Mathematical Modeling of Triangle-meshmodeled Three-dimensional Surface Objects for Digital Holography, Appl. Opt. 47 (19), D117–D127 (2008).

[33] Y.-Z. Liu, J.-W. Dong, Y.-Y. Pu, H.-X. He, B.-C. Chen, H.-Z. Wang, H. Zheng and Y. Yu, Fraunhofer Computer-generated Hologram for Diffused 3D Scene in Fresnel Region, Opt. Lett. 36 (11), 2128–2130 (2011).

[34] Y. Pan, J. Liu, X. Li and Y. Wang, A Review of Dynamic Holographic Three-dimensional Display: Algorithms, Devices, and Systems, IEEE Trans. Ind. Inform. 12 (4), 599–1610 (2016).

[35] P.St-Hilaire, M.E. Lucente, J.D. Sutter, R. Pappu, C.J. Sarpell and S.A. Benton, Scaling up the MIT Holographic Video System, Proc. SPIE 2333, 374–380 (1995).

[36] L. Onural, G. Bozdagi and A. Atalar, New High-resolution Display Device for Holographic Three-Dimensional Video: Principles and Simulations, Opt. Eng. 33 (3), 835–844 (1994).

[37] D.E. Smallley, Integrated Optics for Holographic Video, M. S. thesis, Elect. Eng. Comput. Sci., Mass. Inst. Tech., Cambridge, MA, 2006.

[38] S. Kazempourradi, E. Ulusoy and H. Urey, Fast Computer-generated Hologram Computation Using Rendered Depth Map Image, in Proceedings of SPIE Photonics West Conference, Practical Holography XXXI: Materials and Applications, 10127 (1), 10127Q (2017).

[39] F. Wyrowski and O. Bryngdahl, Iterative Fourier-transform Algorithm Applied to Computer Holography, JOSA A 5 (7), 1058–1065 (1988).

[40] J.R. Fienup, Iterative Method Applied to Image Reconstruction and to Computer-generated Holograms, Opt. Eng. 19 (3), 193297–193297 (1980).

[41] R. Eschbach, Comparison of Error Diffusion Methods for Computer-generated Holograms, Appl. Opt. 30 (26), 3702–3710 (1991).

[42] D. Mengu, E. Ulusoy and H. Urey, Non-iterative Phase Hologram Computation for Low Speckle Holographic Image Projection, Opt. Express 24, 4462–4476 (2016).

[43] F. Wyrowski and O. Bryngdahl, Speckle-free Reconstruction in Digital Holography, JOSA A 6 (8), 1171–1174 (1989).

[44] E.L. Pedrotti, L.M. Pedrotti and L.S. Pedrotti, Introduction to Optics (Cambridge University Press, Cambridge, United Kingdom, 2017).

[45] Z. Zhang, Z. You and D. Chu, Fundamentals of Phase-only Liquid Crystal on Silicon (LCOS) Devices, Light Sci. Appl. 3 (10), e213 (2014).

[46] N. Collings, T. Davey, J. Christmas, D. Chu and B. Crossland, The Applications and Technology of Phase-only Liquid Crystal on Silicon Devices, J. Disp. Technol. 7 (3), 112–119 (2011).

[47] W.P. Bleha and L.A. Lei, Advances in Liquid Crystal on Silicon (LCOS) Spatial Light Modulator Technology, in SPIE Defense, Security, and Sensing (International Society for Optics and Photonics, Baltimore, Maryland, United States, 2013), pp. 87360A–87360A.

[48] K.V. Chellappan, E. Erden and H. Urey, Laser-based Displays: A Review, Appl. Opt. 49 (25), F79–F98 (2010).