



Finite-depth and vari-focal head-mounted displays based on geometrical lightguides

MIAOMIAO XU AND HONG HUA* 

3D Visualization and Imaging Systems Laboratory, James C. Wyant College of Optical Sciences, University of Arizona, 1630 East University Boulevard, Tucson, AZ 85721, USA

*hhua@optics.arizona.edu

Abstract: Existing waveguides and lightguides in optical see-through augmented reality (AR) displays usually guide collimated light, which results in a fixed image depth at optical infinity. In this paper, we explore the feasibility of integrating a lightguide with a varifocal optics engine to provide correct focus cues and solve the vergence-accommodation conflict in lightguide-based AR displays. The image performance and the cause of artifacts in a lightguide-based AR display with a varifocal optics engine are systematically analyzed. A non-sequential ray tracing method was developed to simulate the retinal image and quantify the effects of image focal depth on the image performance and artifacts for a vari-focal display engine of different depths. A prototype with varying image depths from 0 to 3 diopters was built and the experimental results validate the proposed system. A digital correction method is also proposed to correct the primary image artifact caused by the physical structure of the lightguide.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

An optical combiner is a key enabler to optical see-through head-mounted displays (OST-HMDs) for augmented or mixed reality (AR/MR) applications. Among the different technologies for constructing optical combiners, waveguides and lightguides demonstrate high promise and popularity due to their light weight, small volume, and relatively high efficiency. Both waveguides and lightguides consist of an in-coupler, a substrate, and an out-coupler. Waveguides and lightguides are generally classified by their coupler type. While waveguides use diffractive optical elements as their couplers, such as surface relief gratings (SRGs) [1–4], volume holograms (VHs) [5–9] and resonant waveguide gratings (RWGs) [10], lightguides use partially reflective mirror arrays (PRMAs) [11–13] or microstructure mirror arrays (MMAs) [14–16]. During the past few years, researchers have made several efforts on developing design methods for lightguides, optimizing their optical performance, and evaluating their image qualities. Cheng *et al.* analyzed the cause of stray light and proposed a method of minimizing stray light in PRMA lightguides [13]. Wang *et al.* analyzed angular tolerance and image artifacts caused by parallelism errors in PRMA lightguides [17]. Xu *et al.* proposed evaluation metrics for quantifying the key optical performances and artifacts of lightguides and a method of simulating and optimizing the image performance of MMA-based lightguides [18,19]. Several commercial products that adopt lightguides have also become available. Examples include the PRMA-based lightguide products by Lumus [10,11], the MMA-based lightguide product by Optinvent [14–16], and the fractured pinhole-like mirror array approach by LetinAR [20].

Despite the different types of the coupling optics being used, current waveguide- or lightguide-based systems share a similar characteristic: a collimator is required to collimate the ray bundles from a microdisplay before they are coupled into the substrate. Consequently, the out-coupled virtual image perceived at the eyepiece by a viewer appears to locate at the optical infinity in the visual space, since the total internal reflections (TIRs) in the substrate and the out-coupling process do not change the ray collimation, as shown in Fig. 1(a). The requirement for a collimated source before being coupled into a waveguide or lightguide substrate is mainly to minimize the various

artifacts caused by the propagation of non-collimated ray bundles through different imaging paths. As illustrated in Fig. 1(b), during the propagation in the lightguide, a diverged ray bundle is split into several segmented ray bundles (e.g. the red and the green bundles), each of which has a different optical path length (OPL) due to the different numbers of TIRs or reflections by different segmented micro-mirrors of a lightguide, and forms multiple images at different locations. The well-established conventional sequential ray tracing methods become inadequate for investigating the imaging artifacts and optimizing optical performance in a lightguide-based system. Requiring a collimated source naturally simplifies the design and ray tracing of waveguides and lightguides and avoids image artifacts caused by pupil expansion or split ray paths during propagation.

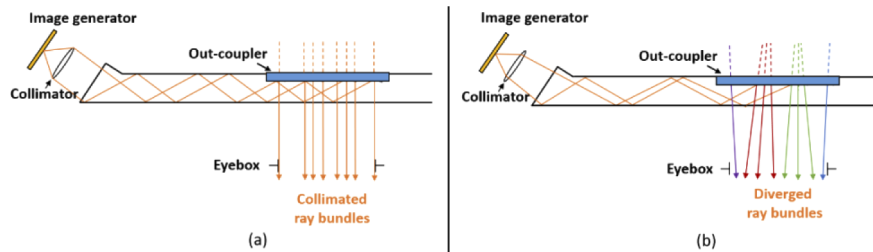


Fig. 1. The ray paths of geometrical lightguide when (a) the out-coupled image is at the optical infinity and the ray bundles are collimated and (b) the out-coupled image is at a finite depth and the ray bundles are diverging when coupled out.

However, there is a growing need for investigating waveguide- or lightguide-based systems with finite image depths rather than a fixed infinite image depth, especially in the context of addressing the vergence-accommodation conflict (VAC) problem in three-dimensional (3D) AR displays. Most of the current AR displays, including waveguide- or lightguide-based AR displays, use conventional stereoscopic techniques to display 3D digital images, where a pair of stereoscopic images with binocular disparity is rendered by a 2D image plane with a fixed focal depth. Such stereoscopic techniques, however, fail to render correct focus cues, including accommodation and retinal blur, and cause the decoupling of the naturally coupled actions of eye accommodation and convergence along with several other related visual cue conflicts, known as the VAC problem [21]. Studies have provided strong supportive evidence that the VAC problem may contribute to visual artifacts such as distorted depth perception, visual discomfort and fatigue [22–24].

Vari-focal display technology is one of the promising solutions to overcome the VAC problem and to render 3D images with correct focus cues in a wide depth range [21]. It dynamically varies the focal depth of a single-plane display, either by implementing an active optical element for focal power control [25–27] or by adjusting the distance between a microdisplay and its imaging optics [28–30]. However, it is not straightforward and not yet investigated to implement a vari-focal mechanism in a waveguide- or lightguide-based system offering a finite or variable focal depth due to the complexity involved in propagating a non-collimated beam through the substrate and the potential image quality degradation and artifacts associated with adopting a vari-focal display engine into a lightguide.

In this paper, we have proposed a systematic approach to analyze and evaluate the image quality change and artifacts induced by the focal depth changes in MMA-based lightguide displays and experimentally demonstrate methods for compensating the primary imaging artifacts in a vari-focal lightguide system. The significance of this study has multi-folds. First of all, as stated earlier, the study provides a comprehensive understanding of the effects of a vari-focal display engine in a waveguide- or lightguide-based AR display system and enables the development of vari-focal lightguide displays. Secondly, investigating the propagation of non-collimated ray paths through a flat substrate offers insightful guidance for specifying the system tolerances when

the ray bundles from a microdisplay are not perfectly collimated before being coupled into a lightguide or the substrate surfaces are not perfectly flat or parallel. Finally, developing methods for non-collimated ray propagation through a substrate may open up the possibilities of designing a curved lightguide with optical power. It is worth noting that the analysis methods may only be applicable to lightguide-based system, since some of the crucial effects in waveguide-based system (such as chromatic effects or angular degradation) are not included. On the other hand, similar image artifacts also arise in waveguide-based system when image focal depth changes.

The rest of the paper is organized as follows. Section 2 analyzed the optical path differences (OPDs) between different ray paths in a lightguide and illustrated the cause of image artifacts by OPDs; Section 3 demonstrated the non-sequential ray tracing method to simulate the imaging paths in a lightguide and demonstrated the simulations of the locations of multiple image points originating from the different ray paths; Section 4 quantified the image performance change by the modulation transfer function (MTF) and analyzed effects of focal depth on the displayed depth accuracy in both monocular and binocular configurations; Section 5 demonstrated a prototype and experimental results by capturing the out-coupled images at different image reconstruction depths to validate the image simulation model; and Section 6 demonstrated the results of our proposed digital correction method to compensate for the primary image artifacts induced by focal-depth change.

2. Characterizing focus-induced imaging artifacts

Figure 2 shows a typical structure of an MMA-lightguide AR display. It consists of an image generator, a collimator, and an MMA-lightguide. The image generator displays the digital image seen by a viewer, and the collimator magnifies the image from the image generator and forms a virtual image. The magnified virtual image is then coupled into the MMA-lightguide by a wedged-shaped in-coupler, propagates via TIR in the lightguide, and is coupled out to project rays towards the eyebox by the MMA out-coupler. The MMA out-coupler is composed of a one-dimensional micro-mirror array, which is an array of coated slanted groove tops spaced apart. The ray bundle from a pixel of the image generator usually propagates through the substrate via different TIR paths and hits several micro-mirrors when it is coupled out, which generates several sub-ray paths at the eyebox, as shown in different colors in Fig. 2. When the collimator is configured such that the focal depth of the virtual image is at optical infinity, the image seen by the viewer is not affected by the split of sub-ray paths, since all the sub-ray paths from the same field position have the same angle when coupled out. However, when the virtual image is located at a finite focal depth in the visual space, the split sub-ray paths have different OPDs and result in the formation of multiple out-coupled images at different locations. Consequently, the ray path splitting can cause image artifacts, which are similar to the ghost image artifacts in conventional imaging systems.

To analytically characterize the imaging artifacts induced by a non-collimated image source propagating through a light guide, let us assume that the ray bundle from each pixel on the virtual image appears to be diverging before being coupled into the lightguide. Figure 3 shows two different cases of ray path splitting for an in-coupled diverging ray bundle. In the first case, as illustrated in Fig. 3(a), the ray bundle from the source point P propagates through the lightguide by the same number of TIR reflections within the substrate but is coupled out by different micro-mirrors and thus split into different ray paths. For instance, the ray bundle is coupled out during the r^{th} reflection but reflected by two different micro-mirrors and thus split into two sub-paths illustrated with yellow and red colors, respectively. In the second case, as shown in Fig. 3(b), the ray bundle from the source propagates through the substrate via different number of TIR reflections but is coupled out by the same micro-mirrors. For instance, the red ray bundle (Path 2) is coupled out after the r^{th} reflection, while the yellow ray bundle (Path 1) is coupled out after two more TIRs than the red ray bundle. To differentiate the two types of ray

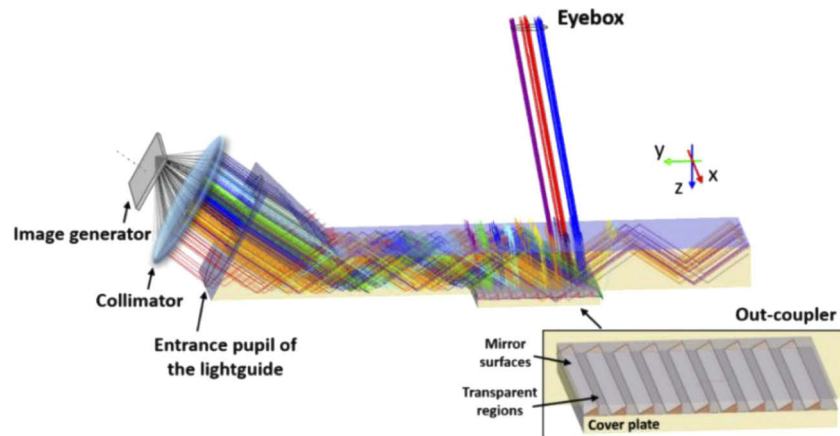


Fig. 2. Schematic layout of the MMA-based geometrical lightguide AR display.

path splitting, we defined the first case as the MMA-induced ray path splitting and defined the second case as the multiple reflections induced (MR-induced) ray path splitting. Practically both cases of ray splitting will likely occur simultaneously.

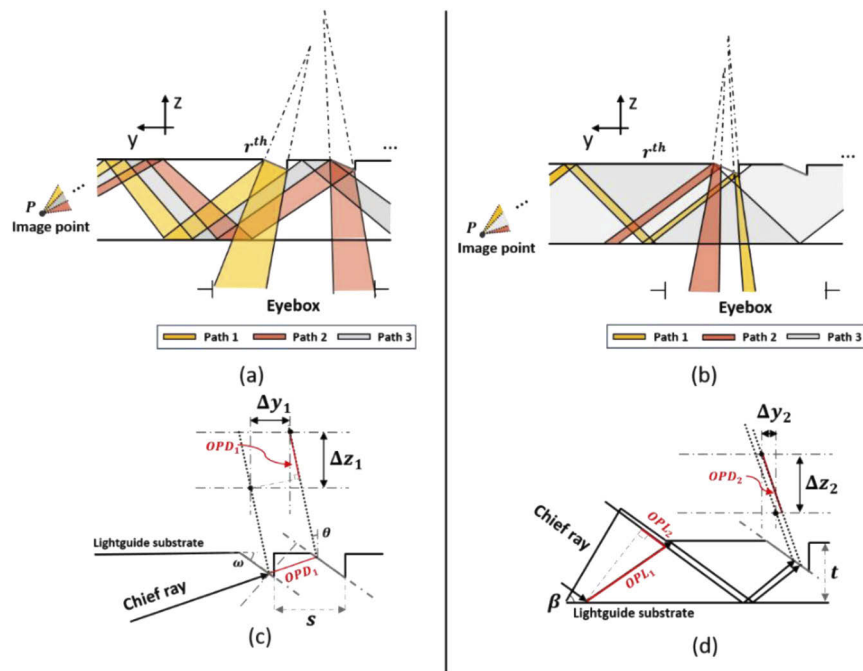


Fig. 3. Ray path diagrams of a ray bundle from the image generator. Different OPLs are introduced due to (a) micro-mirror displacement and (b) multiple TIRs. Zoomed-in chief ray path diagrams of each case are shown in (c) and (d), which illustrate the relationship between OPD_i and image displacements (Δy_i and Δz_i).

Figures 3(c) and 3(d) further illustrate the OPDs of the MMA-induced and MR-induced ray path splitting, respectively. In the MMA-induced case, the OPDs among the different split ray paths, denoted as OPD_1 , are primarily caused by the displacements among the micro-mirrors

along the chief ray direction of the ray bundle, and is related to the mirror-array spacing s and micro-mirror tapered angle ω as

$$OPD_1 = \frac{ns \sin \omega}{\cos(\omega + \theta')} \quad (1)$$

where n is the refractive index of the lightguide substrate, $\theta' = \sin^{-1} \left(\frac{\sin \theta}{n} \right)$ is the refractive angle in the lightguide substrate of a given field angle, θ , which is defined as the viewing direction of the eye located at the center of the eyebox as shown in Fig. 3(c). In the MR-induced case, The OPDs among the different paths, denoted as OPD_2 , are primarily due to the different number of TIR reflections, and is related to the lightguide structure parameters as

$$OPD_2 = OPL_1 - OPL_2 = \frac{nt}{\sin(90^\circ - \beta - \theta')} [1 - \sin(2\beta + 2\theta' - 90^\circ)] \quad (2)$$

where t is the lightguide thickness and β is the in-coupling wedge angle.

Each of the split ray paths is projected onto the eyebox at different locations and appears to form a separate virtual image in the visual space. In both cases, the OPDs among the split paths cause the separation of the images formed by the split ray paths in both the x and z directions. The image displacements in depth (Δz_j) and in plane (Δy_j) for both cases can be further calculated by projecting the OPD_j , which varies as a function of the field angle θ , in the two orthogonal directions as,

$$\begin{aligned} \Delta y_j &= OPD_j \cos(90^\circ - \beta - 2\theta') \cdot \cos \theta - OPD_j \cdot \sin \theta \\ \Delta z_j &= OPD_j \cos(90^\circ - \beta - 2\theta') \cdot \sin \theta + OPD_j \cdot \cos \theta \end{aligned} \quad (3)$$

where $j = 1, 2$ corresponds to one of the two cases in Fig. 3(c) and 3(d), respectively. From Eqs. (1)-(3), we can conclude that the MR-induced image artifact has much larger magnitude and thus more observable than the MMA-induced image artifact for a typical MMA-lightguide structure. Consider an MMA example with a 1mm mirror array spacing, 30° tapered angle, 4.4mm substrate thickness, and an in-coupled virtual image source with a focal depth of 1 diopter. For an object field angle of $\theta = +1^\circ$, the image displacement caused by the MMA-induced ray path splitting is 3.84 arc minutes in the OXY plane and 0.0013 diopters in its z -depth while the image displacement by the MR-induced case is 19.0 arc minutes in the OXY plane and an 0.0066 diopters in the z -depth. It can be also seen that the OPDs not only result in ghost-image-like artifacts, but also bring in depth displacements, which also affects the stereo acuity discussed in Section 4.

Figures 4(a) and 4(b) share examples of both types of artifacts captured by a camera from a prototype setup where the virtual image plane is set at 1.8 diopters from the eyebox position. Details about the experimental setup will be discussed in Section 5. As shown in Fig. 4(a), the MMA-induced ghost images appear across the whole FOV and stay closer to each other, which can be observed only when fine image contents are displayed. On the contrary, the MR-induced ghost images, as shown in Fig. 4(b), separate further from each other and are more visible, but they only exist in a small field range. These results can be explained in ray tracing software as shown in Figs. 4(c) and 4(d), respectively. For most of the in-coupling field angles, the ray bundles propagate through the same number of TIRs but are split into three or four sub-ray paths at different micro-mirrors, as shown in Fig. 4(c). In this case, the image displacement is mainly induced by the adjacent mirror displacement, which causes the first-type image artifacts in Fig. 4(a). On the other hand, the MR-induced image artifacts arise only in some specific field angles where both the in-coupled ray bundles propagate through different numbers of reflections [e.g. reflection number of 3 and 5 for the blue and red rays in Fig. 4(d), respectively] when being coupled out. Since the OPDs induced by multiple reflections are much larger in magnitude (OPD shown in red arrows), the lateral image displacement of the MR-induced ghost image is more

visible since Δy_2 is much larger than Δy_1 based on Eq. (3). Besides image displacement, the ray path splitting also causes degradation in image quality and MTF, which is discussed in Section 4.

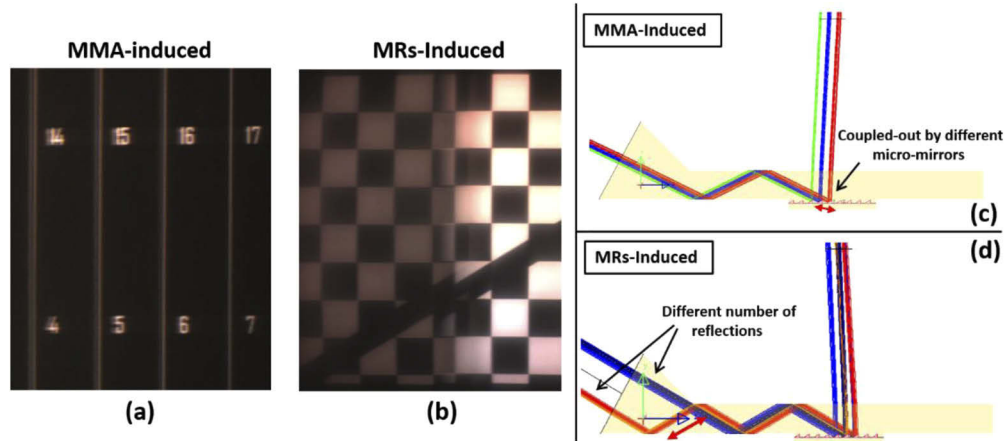


Fig. 4. Examples of both types of image artifacts when the image plane is at 1.8D from the viewing position. (a) Zoomed-in view of the MMA-induced ghost image when displaying fine contents at $+4^\circ$ to $+7^\circ$ field angles. (b) Zoomed-in view of the MRs-induced ghost image, where the ghost image appears at -3° with an angular width of 0.92° (the angular range and distribution depend on lightguide structure and image plane depth). The ray paths of MMA-induced ray splitting and MRs-induced ray splitting are shown in (c) and (d), respectively.

3. Ray tracing and image simulation method

To characterize and quantify the imaging artifacts induced by a non-collimated image source through a lightguide, it is necessary to establish ways of accurately simulating the ray splitting effects and perceived retinal image. The nature of non-sequential ray propagation through a geometric lightguide makes the process of performing ray tracing and image simulation much less straightforward and more challenging than the process of modeling and assessing conventional imaging optics which generally uses the sequential ray tracing method and calculates the diffraction effect based on a single aperture. First of all, the displayed image relayed by the lightguide generates multiple ray paths which may have non-negligible OPDs and form multiple spatially-separated virtual image points as analyzed in Section 2. The perceived image through the multiple segmented ray paths cannot be predicted by conventional sequential ray tracing methods. The second challenge is the simulation of diffraction effects. Each of the segmented ray paths projects a small footprint on the eye pupil due to the limited micro-mirror size induces significant diffraction effects. Typical non-sequential ray tracing software tools such as LightTools, however, do not simulate diffraction effects. Finally, the ray bundles from different field angles are reflected at different locations on the cascaded MMA, which generates field-dependent strip-shaped pupils that differ from the constant circular pupils of conventional imaging systems.

To overcome the aforementioned problems, we adopted the modeling method we previously developed for MMA-lightguide in [18,19] using LightTools. Instead of assuming the in-coupling source is perfectly collimated as in [18,19], a point source, representing a sampled field of a virtual image formed by the collimator, is placed at a distance d away from the entrance pupil of in-coupler, and is moved laterally to simulate different field positions on the virtual image plane. The Monte Carlo method was adopted to generate the rays from each of the point sources. The

virtual image distance, d , from the in-coupler wedge of the lightguide [shown in Fig. 5(a)] can be related to the focal depth, A , of the virtual image appearing in the visual space (in diopters) as

$$d = \frac{1000}{A} - Z_{EX} - nt - \frac{n}{\cos \theta'} \left(\frac{t_1}{2} + rt \right) \quad (4)$$

where Z_{EX} is the exit pupil distance measured from the inner surface of the lightguide to the eyebox, t_1 is the in-coupler wedge height, and r is the number of reflections from the in-coupler to out-coupler. The ray bundle from a pointing source was coupled into the lightguide through the entrance pupil of the lightguide, as shown in Fig. 5(a). A receiver at the eyebox of the lightguide collected the ray data, including ray path number, ray position, and direction cosines in the three-dimensional space. Figure 5(b) shows an example of a spot diagram recorded by the receiver, where each color dot represents a ray path. In this example, the input ray bundle was split into three different paths, represented by the blue, green, and red dots, respectively. We repeated this ray tracing and data collection process for each of total 1100 sampled fields.

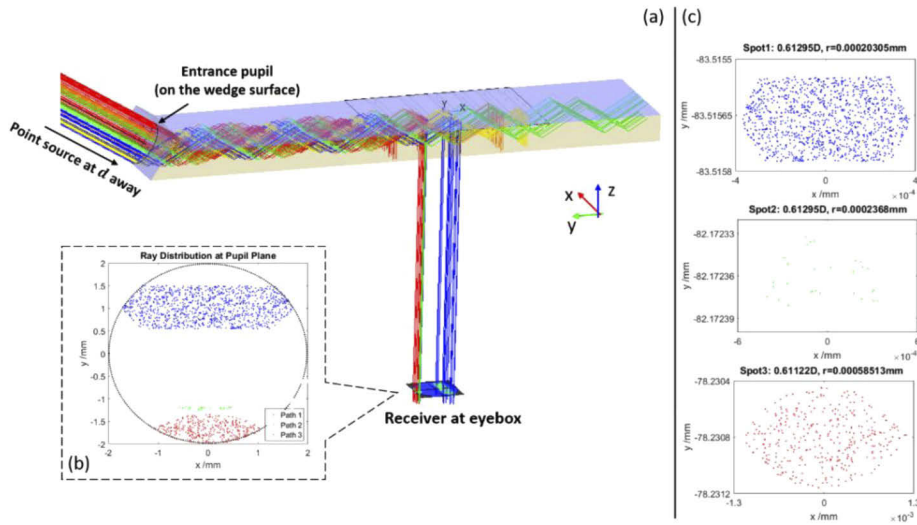


Fig. 5. (a) Schematic layout of the MMA-lightguide and the ray path simulation. (b) Spot diagram at the eyebox, where each ray path is designated with a unique color. (c) Spot diagrams of three segmented ray bundles at the depth with the minimum RMS spot size.

When the in-coupled beam has a finite focal depth, each of the split ray paths will form a spatially displaced virtual image. To reconstruct the image position in visual space, the ray position (x_i, y_i) in the plane perpendicular to the viewer's eyesight at a distance z in front of the eyebox is calculated by:

$$\begin{aligned} x_i &= x_{pi} + \frac{z}{N_i} \cdot L_i \\ y_i &= y_{pi} + \frac{z}{N_i} \cdot M_i \end{aligned} \quad (5)$$

where (x_{pi}, y_{pi}) and (L_i, M_i, N_i) are the recorded ray position and the direction cosines at the eyebox. The perceived image positions and depths were computed by finding a depth z that gives the minimum root-mean-square (RMS) footprint sizes for each ray path. Figure 5(c) shows the minimum RMS spot diagrams of each ray path in Fig. 5(b). Note that the image points from different ray paths are noticeably separated in Y-direction indicated by the y coordinates in Fig. 5(c) and may be located at different z depths due to the OPDs where the z -depth of each image point is indicated by the number above each spot diagram.

The perceived retinal image is the integral sum of the piecewise rays propagated through the different ray paths integrated by the eye lens. We adopted a computational method similar to the one described in our previous work in [19] to simulate the retinal point spread function (PSF) and perceived retinal image. Figure 6(a) shows a schematic layout of lightguide coupled with an eye model placed at the eyebox plane for retinal image simulation. In this illustration, a ray bundled from a diverging point source is coupled into the lightguide and is split into two different ray paths which consequently form two spatially separated virtual image points, P_1' and P_2' , on the virtual image planes (ξ, η) and (ξ', η') . Since the ray bundles are piece-wisely reflected by the MMA, the footprints of the split ray bundles on the pupil plane of the eye model may only occupy a small portion of the eye pupil, as shown in Fig. 5(b), and thus the retinal image of each split ray paths is subject to significant diffraction effects. Moreover, the ray bundles may hit different locations on the MMA which may have a spatially varying reflectance, and thus the projected field amplitudes of different ray paths on the eye pupil are depending upon its incident locations on the MMA. To model these effects, we treat each of the split ray paths and its associated virtual image as an incoherent source and model its retinal PSF separately. The incoherent retinal PSF of the segmented ray bundle (m, n) in the n^{th} sub-ray path and the m^{th} field with the eye accommodated at the depth A is expressed as

$$PSF_{m,n}(u, v; \xi, \eta; A) = \left| \frac{1}{j\lambda^2 z_{\text{eye}}} \exp[-j\frac{k_m}{2z}(u^2 + v^2)] \exp[-j\frac{k_m}{2z_{\text{eye}}}(\xi^2 + \eta^2)] \int \int E_{r,mn}(x, y; \xi, \eta) \cdot P(x, y) \cdot \exp[jk_m W_{ab}(x, y; \xi, \eta)] \cdot \exp[j\frac{k_m}{2}(\frac{1}{z} - \frac{1}{A})(x^2 + y^2)] \exp\{-jk_m[(\frac{\xi}{z} + \frac{u}{z_{\text{eye}}})x + (\frac{\eta}{z} + \frac{v}{z_{\text{eye}}})y]\} dx dy \right|^2 \quad (6)$$

where (ξ, η) , (x, y) and (u, v) are the coordinates on the image plane, pupil plane and retinal plane, respectively, z_{eye} is the distance between the pupil plane and the retinal plane, k_m is the wave vector of the m^{th} field, $P(x, y)$ is the pupil function, W_{ab} is the compound aberration phase term in the eye model for which we chose to use the Arizona Eye Model, which is able to model the clinical levels of aberrations and population-average ocular parameters [31]. $E_{r,mn}(x, y; \xi, \eta)$ is a function introduced to model a field dependent amplitude distribution function of each ray path of a given field on the pupil plane. It can be calculated as $E_{i,mn}(x, y; \xi, \eta) \cdot r_A(x, y; \xi, \eta)$, where $E_{i,mn}(x, y; \xi, \eta)$ is the incident ray amplitude distribution incident on the MMA and $r_A(x, y; \xi, \eta)$ is the projected amplitude reflectance on the pupil plane.

As illustrated in Fig. 6(a), the sub-ray paths, such as the Path n_1 and Path n_2 , from the ray bundle m_1 of the same object field, are projected to two spatially separated image points in the visual space. These image points correspond two slightly different field angles with respect to the eye pupil center due to their lateral shift and form two independent images on the retina. Figure 6(b) shows the retinal PSFs of the three segmented ray bundles shown in Fig. 5(b) when the eye is accommodated at $A = 0.6D$. The overall retinal PSF of the m^{th} field, $PSF_m(u, v; \xi, \eta; A)$, may be calculated as the sum of all three PSFs calculated for the sub-ray paths as

$$PSF_m(u, v; \xi, \eta; A) = \sum_n PSF_{m,n}(u, v; \xi, \eta; A) \quad (7)$$

Figure 6(c) plots the integrated retinal PSF for the example shown in Fig. 5(b). As expected, due to the spatial shifts of the sub-image points, the overall PSF consists of multiple peaks. The method of calculating the overall retinal PSF with Eq. (7) is reasonable for an individual object field, but it overlooks the potential crosstalk effects of adjacent object fields. With more than one adjacent point sources, the image points of the sub-ray paths from the m^{th} field in the object space may overlap with the sub-image points of several different fields. An example is shown in Fig. 6(d), where two ray bundles (denoted in orange and green color), from two different object points m_1 and m_2 , form an overlapping image point on the retina, although their corresponding

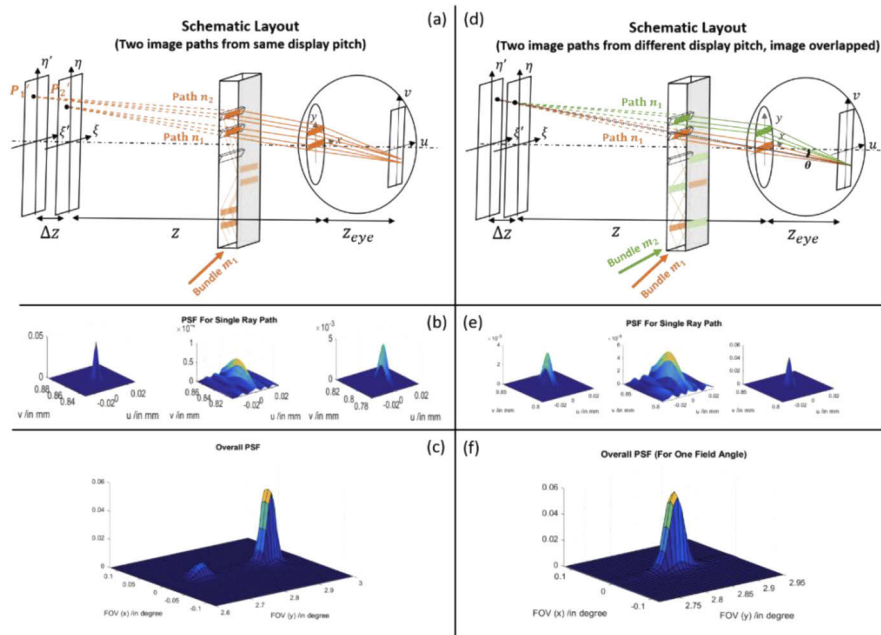


Fig. 6. (a) Schematic layout of the retinal image simulation model for a single object field. (b) The simulated retinal PSFs for each segmented ray bundle as shown in Fig. 5(b). (c) The overall PSF that is weighted over all ray paths in one ray bundle in (b). (d) Schematic layout of the retinal image simulation model for multiple object fields where some of the sub-ray bundles form an overlapping image point on the retina. (e) Retinal PSFs from the ray bundles of different object fields with overlapping image position on the retina. (f) The overall PSF calculated by the weighted sum of the overlapping PSF in (e).

image points in the visual space are separated along the view direction by Δz . To account for these effects, we modified the Eq. (7) and defined the overall retinal PSF by summing up the sub-image PSFs of all the overlapping images points on the retina. As shown in Fig. 6(d), (a) field angle, θ , is used to characterize the chief ray angle of a field position on the retina space with respect to the center of the eye pupil, and is defined as $\theta = \frac{v}{z_{eye}}$. The overall retinal PSF corresponding to a given field angle θ , denoted as $PSF_{\theta}(u, v; \xi, \eta; A)$, is calculated as the weighted sum of all the retinal PSFs of the sub-ray paths overlapping on the retina but from different object fields, expressed as

$$PSF_{\theta}(u, v; \xi, \eta; A) = \sum_{i=1}^{m_{\theta}} w_i \cdot PSF_{\theta_i}(u, v; \xi, \eta; A) \quad (8)$$

where the weighting factor $w_i = \frac{L_i}{\sum_{i=1}^{m_{\theta}} L_i}$, m_{θ} is the total number of sub-image points overlapping on the θ direction, L_i is the luminance of the i^{th} image point in the θ direction, A is the dioptric accommodation depth of the eye.

Figure 6(e) shows examples of independent $PSF_{\theta_i}(u, v; \xi, \eta; A)$ formed by the sub-ray paths of 3 different object fields but overlapping spatially on the retina, while Fig. 6(f) plots the corresponding overall retinal PSF_{θ} calculated with Eq. (8). The retinal image of each object field is then calculated as the convolution of the displayed image and the retinal PSF given by Eq. (8). The retinal modulation transfer function (MTF) for a given field angle θ , denoted as

$MTF_{\theta}(\xi_u, \eta_v)$, can be calculated by applying Fourier transform to Eq. (8) as

$$MTF_{\theta}(\xi_u, \eta_v; A) = \left| \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} PSF_{\theta}(u, v; \xi, \eta; A) \exp[-j2\pi(\xi_u u + \eta_v v)] dudv}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} PSF_{\theta}(u, v; \xi, \eta; A) dudv} \right| \quad (9)$$

Given that the overlapping sub-image points on the retina may be distributed at slightly different depths in the visual space, the depth of the fused image from the sub-images is determined by evaluating the value of $PSF_{\theta}(u, v; \xi, \eta; A)$ in a range of accommodation depth A and finding out the depth with maximum corresponding $MTF_{\theta}(\xi_u, \eta_v; A)$ at a selected mid-frequency between 10-15 cycles per degree.

4. Simulation results

Based on the method described in Section 3, we simulated the image performance of a commercial MMA-lightguide in LightTools. The lightguide has a dimension of 51mm (L) x 4.39mm (W) x 16mm (H), with a 61.67° in-coupler wedge angle and a 13.3mm (L) x 1.53mm (W) x 16mm (H) out-coupler area. The out-coupler of the lightguide is composed of an array of equally spaced slanted mirrors sandwiched between the lightguide substrate and a cover plate. The MMA has a width of 0.77mm and the spacing between adjacent mirrors is 1.5mm. A microdisplay with a resolution of about 1100 by 660 pixels and a maximum luminance of 200 cd/m² serves as the image generator. The collimator has a focal length of 20.82mm, which gives a FOV of 19.11° x 11.49°. By iteratively placing a point source at the center of each pixel on the image formed by the collimator, we simulated the ray bundles from all different field angles. A circular receiver with a diameter of 4mm and a bin count of 31 × 31 was placed 23mm away from the inner surface of the lightguide and acted as the eyebox of the system. The retinal plane has a sampling density of 0.47 arcmin/pixel.

Figure 7(a) shows a simulated retinal image and the image profile (a cross section of the image) when a uniform white image is displayed at a focal depth of 0.6D. Figure 7(b) plots the number of overlapping sub-image points on the retina across the field (in red dots) and their depth distributions over the FOV (in green dots). Using the method described in Section 3, Fig. 7(c) plots the depth of the fused image plane from the sub-images of different depths across the FOV. The average depth of the image plane is 0.602D with a depth variation (the dioptric depth difference across the FOV, Δd) of 0.003D. We further computed the MTF for each field position across the FOV using Eq. (9). For the convenience of evaluating the optics performance with a simple metric, the displayable threshold frequency of a given field angle is defined as the frequency when the MTF value remains above 0.2. Figure 7(d) plots the display threshold frequency as a function of field angle. Across the FOV of ±9.55°, the threshold frequency varies between 15 and 20 cycles per degree, with a noticeable dip around the field angle of around -3.5 degrees. This indicates that an MMA-based display with a finite focal depth exhibits non-uniform image resolution and performance across its FOV and some region is expected to suffer more quality degradation.

To study the image performance of a varifocal MMA-lightguide when the focal depth of the display collimation engine varies in a wide range, we further simulated the image profiles, the MTF cut-off frequencies, and the fused depths of the image planes when the focal depth of the displayed image is located at 0D, 0.6D, 1.2D, 1.8D, 2.4D, and 3D, respectively. These depths have equal dioptric spacing and can provide acceptable image quality in a wide depth range [32,33]. Figure 8(a) shows the image profiles across the FOV for the focal depth range varying from 3D to infinity. The periodic shape of the image profile results from the periodic arrangement of the MMA. Figure 8(b) shows the displayable threshold frequency same as defined earlier over

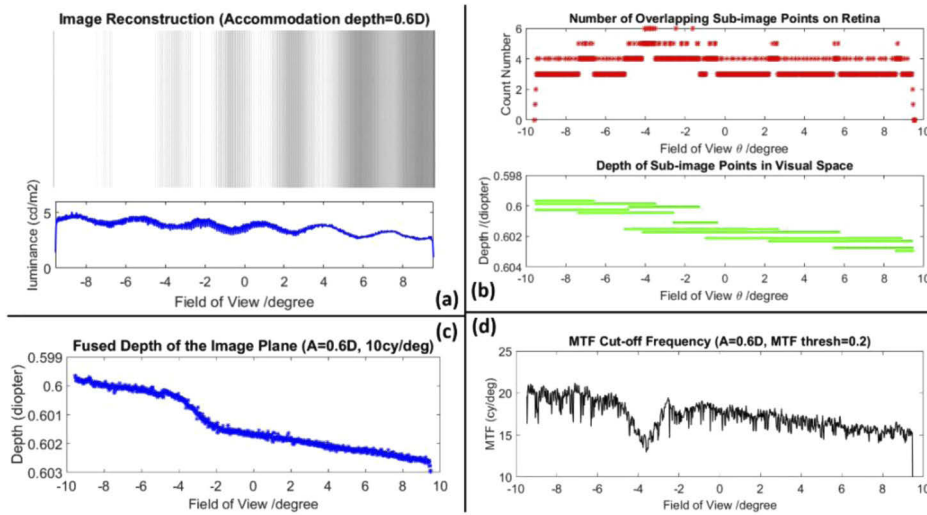


Fig. 7. Image simulation results of the MMA-lightguide when the image plane is located at 0.6D away from the eye. (a) The simulated retinal image profile of the MMA-lightguide. (b) Depth of the out-coupled image points (green dots) and the number of image points generated by different ray paths over the FOV (red dots). (c) Depth of the image plane found by calculating the depth fused $PSF_{\theta}(u, v; \xi, \eta; A)$. (d) Threshold angular frequencies on the image plane when MTF is above 0.2.

the FOV for different focal depths. The MTF threshold frequency does not change significantly with the depth, but there is a valley between -5° and -2° for all depths. This is because both the MMA-induced and the MRs-induced ghost image artifacts arise in the area and more ray path splitting narrows down the aperture size of each ray path, which enhances the diffraction effects. Figure 8(c) shows the rendered depths of the fused image plane after being imaged through MMA lightguide for the focal depth of 1.2D, 1.8D, 2.4D and 3D, respectively.

Based on the results of Figs. 7(c) and 8(c), a small depth shift can be observed cross the FOV for the different focal depths. The results further show that the positive fields appear to be closer to the eye than the negative fields at all of the simulated image, which can be explained by the tendency of increasing OPLs from the left to the right side of the fields in the lightguide. Figure 9(a) further plots the depth variation (Δd) across the FOV as a function of the focal depth of the collimator image plane varying from 0 to 3 Diopters. It can be seen that the depth variation increases as the image depth becomes closer to the eye. The magnitude of the depth variation is rather small (less than 0.1 diopters for the range of 3 diopters) and is expected to have negligible impact on the depth perception in the case of a monocular varifocal engine. However, it can cause significant errors of vergence angle in a stereoscopic system and affect the perceived depth of stereoscopic rendering. Figure 9(b) plots the error of the vergence angle as a function of the reconstructed depth rendered via binocular disparities when the focal depth A [shown in Fig. 9(c)] of the 2D image plane for both arms of a binocular MMA-based display system is varied from 0.6 to 3 diopters, assuming that the inter-pupillary distance (IPD) of human eyes is 64mm [34]. The result shows that the binocular vergence angle error can increase dramatically from less than 2 arcminutes for the binocular rendering depth of 0.6 diopters to nearly 50 arcminutes for the rendering depth of 2 diopters. Such large error is beyond the stereo acuity of normal adult human eyes by several orders of magnitude [35], which is significantly perceivable. To display an image with the corrected binocular vergence angle, depth calibration and digital image correction should be employed to account for the depth shift effects.

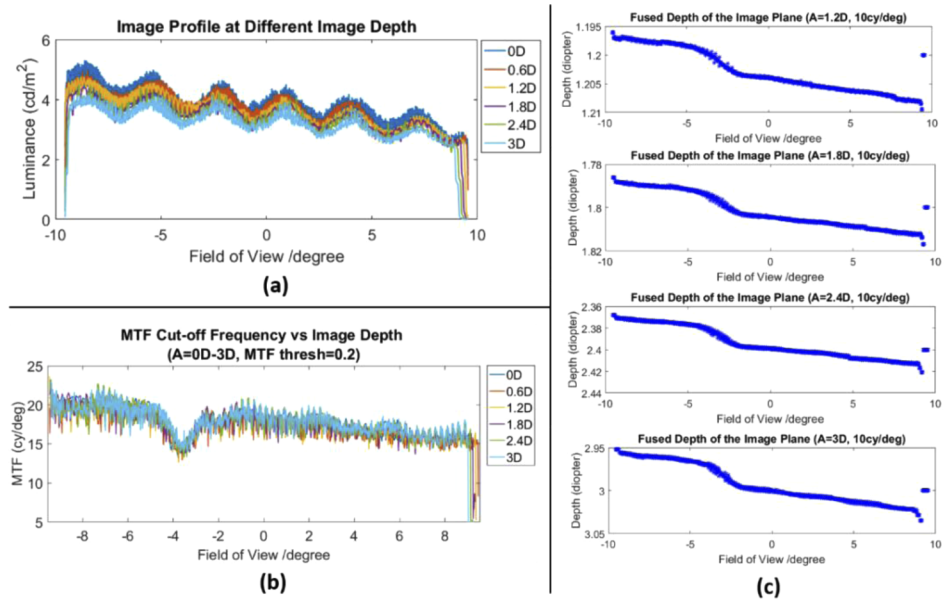


Fig. 8. Simulated image performance of the MMA-lightguide when the focal depth of the image plane varies from 0D to 3D. (a) The image profiles over the FOV for different focal depths. (b) The maximum displayable frequencies over the FOV when the MTF is above 0.2. (c) The fused depths of the image plane after being imaged through the lightguide for the focal depth of 1.2D, 1.8D, 2.4D and 3D, respectively, while the result for the focal depth of 0.6D is plotted in Fig. 7(c).

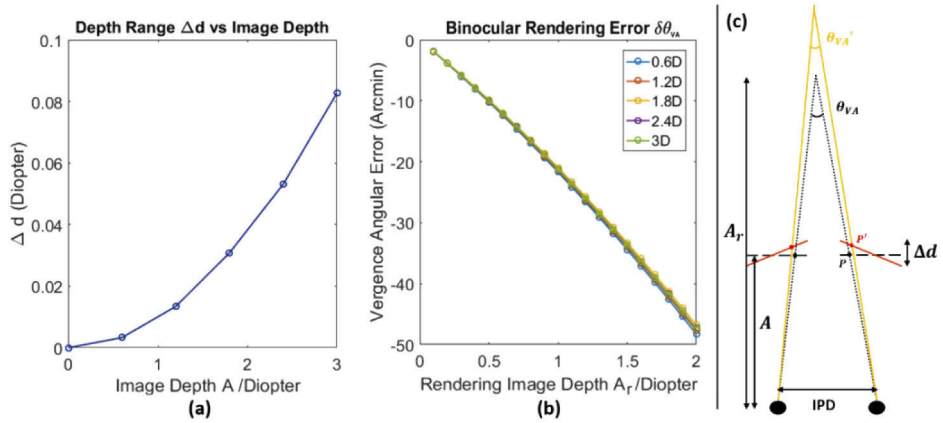


Fig. 9. (a) Depth variation Δd of the image plane as a function of central image plane depth. (b) Reconstruction error of vergence angle as a function of binocular image reconstruction depth in stereoscopic vision. (c) Schematic layout of the vergence angle error, which relates to image plane depth A and the rendering image depth A_r .

5. Experimental results

To validate the ray tracing and simulation results described in Section 4, we built a prototype, as shown in Fig. 10. A ferroelectric liquid crystal on silicon (F-LCoS) microdisplay with built-in RGB LED light source serves as the image generator. A commercial MMA-lightguide [Fig. 10(a)] and its collimator were used as the imaging optics. To match the original design performance of

the imaging optics, only the central area of the LCoS was used, which was 0.32" in diagonal and had a resolution of 1100 by 660. The LCoS was mounted on a one-dimensional translation stage so that the focal depth of the virtual image plane could be controlled by varying the axial distance between the LCoS and the collimator. In the experiment, six image depths were tested, including 0D, 0.6D, 1.2D, 1.8D, 2.4D and 3D, which correspond to the simulated image depths in Section 4. A periodic black/white pattern with a period of 2.78 degrees per cycle was displayed on the LCoS as the test image. A 4K camera with a 12mm, f/4 camera lens was set at the exit pupil of the lightguide to capture the out-coupled image. To measure the image depth, three physical resolution targets were placed in the see-through path of the lightguide system. As shown in Fig. 10(b), Target1 was translated on an optical rail between 1.2D-3D, while Target2 and Target3 were fixed at 0.6D and 0.25D where Target3 was used as a reference for 0D.

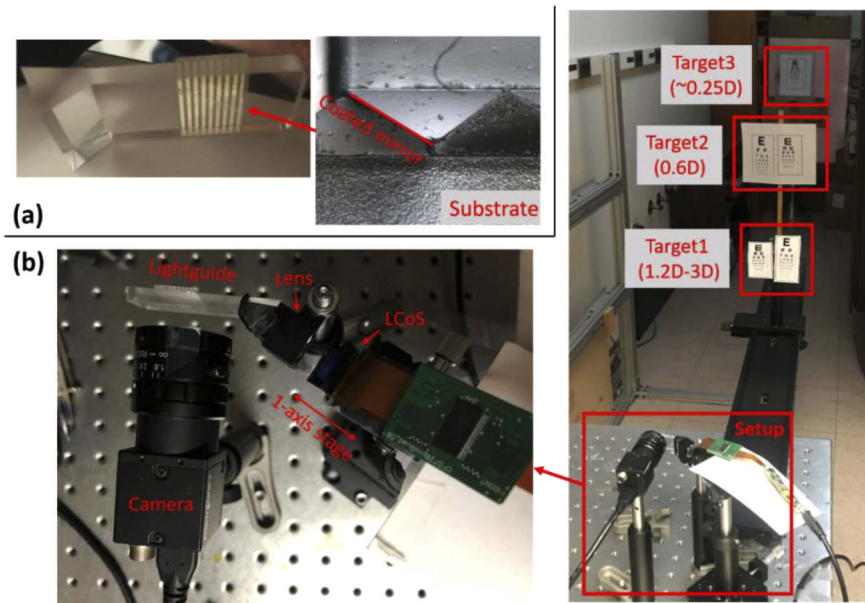


Fig. 10. (a) MMA-lightguide (left) and the MMA structure captured under microscope (right). (b) Experimental setup of the finite depth MMA-lightguide AR system. Three targets at different depths are used as references.

Figure 11(a) shows a set of the out-coupled test images captured by the camera corresponding to the different focal depths from 0.6 to 3 diopters, respectively. For the purpose of comparison, we simulated the retinal images using the simulation method described in Section 3. In the simulation, the input image was the same periodic black/white target image as the one displayed through the prototype and the focal depth of the input image was set to match with the corresponding sampled depths for experiment. In each of sub-figures for different focal depths in Fig. 11(a), the corresponding simulated image and its intensity profile are also shown. It can be observed that the simulated image profiles match well with the captured images except for some image non-uniformity, which may be due to the propagation losses and pupil mismatch. It can be further seen that the MRs-induced ghost image is visible between -5° and -2° fields, and the horizontal displacement between the MRs-induced ghost images becomes larger as the focal depth of the image becomes closer to the eye. The horizontal displacement changes the image period as well as the image contrast. Figure 11(b) plots the angular subtense of the MRs-induced ghost image separation as a function of the focal depth of the image plane. The angular subtense can also be calculated by the OPDs between different ray paths and the lateral image displacement

[Eqs. (1)-(3)]. The theoretical calculation results are also plotted in Fig. 11(b), showing a good accordance with the experimental results. Figure 11(c) shows the experimental results of the MTF cut-off frequency (MTF is above 0.2) across the whole FOV when the image focal depth varies from 0D to 3D. Slanted edge is used for the testing at different field angles. Compared with the simulation results shown in Fig. 8(b), the cut-off frequency of the experimental results are lower due to the aberrations and collimation errors of the system. However, the experimental results show similar tendency, and the image performance does not degrade much as the image focal depth changes.

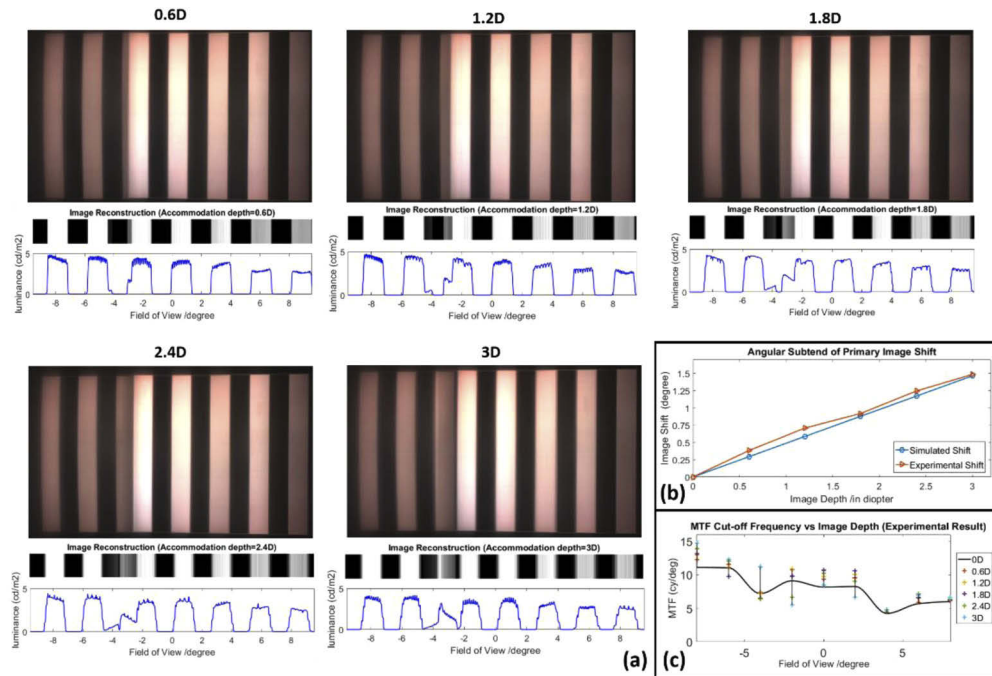


Fig. 11. (a) Experimental results (top subplot) and simulated results (bottom subplot) of the out-coupled image from the MMA-lightguide. The displayed image is a periodic rectangular pattern with a period of 2.78 degrees per cycle (80 pixels per strip). (b) The experimental results and theoretical predictions of the lateral image displacement (in degrees) between the MRs-induced ghost images [based on Eqs. (1)-(3)]. (c) Experimental results of the MTF cut-off frequency (threshold is 0.2) across the field of view.

6. Digital correction

From the experimental results shown in Section 5, we can see that the major image artifact of the MMA-lightguide with a finite image focal depth is the irregularity and artifacts of the image period induced by different numbers of TIR propagation by the substrate. Between -5° and -2° fields in the visual space, the MR-induced ghost images have comparable intensities to the actual image content and they have lateral displacement between each other, which breaks the original image period. Figures 12(c) and 12(e) further illustrate the problem. While a periodic bar pattern with a constant spatial frequency [Fig. 12(c)] is displayed and a similar out-coupled image is supposed to be seen, the actual out-coupled image [Fig. 12(e)] shows an altered spatial frequency at the region from -5° to -2° due to the lateral displacement caused by the MR-induced artifact [third bar in Fig. 12(e)]. To minimize the image artifact and correct the displayed image period, the digital input of the microdisplay can be pre-processed to compensate for the image artifacts.

Since the artifacts only arise in a small angular subtense, we can segment the input image and optimize the input values in the region where the MR-induced ghost image is the most observable [ROI in Fig. 12(e)]. If the ghost images have comparable intensities and they have k pixels of lateral displacement, we can optimize the digital input by the linear least-square method (LLS) given by Eq. (10)

$$Q = \min_B \|T \cdot B - I\|_2^2 \quad (10)$$

where $B = [b_1, b_2, \dots, b_n]^T$ is the image profile in ROI, b_n is the digital input value of n^{th} pixel in ROI, $I = [I_1, I_2, \dots, I_m]^T$ is the desired display intensity in the optimized region ($m = n + k$ due to the image displacement), T is an m by n sparse, binary-valued weighting matrix that denotes the pixels in B that contribute to I . Figures 12(a) and 12(b) show an example of a desired image profile in the ROI and the optimized results of the ROI by LLS. The image segment on the left side of the ROI should be cropped and translated to compensate for the lateral image displacement. After the regional optimizations, all segments are stitched together. Figures 12(c) and 12(d) show the digital input before and after the digital correction, when a black/white pattern with a period of 4.17 degrees/cycle is displayed at 2.4D away. The camera-captured out-coupled images before and after digital correction are shown in Figs. 12(e) and 12(f), respectively. It can be clearly seen that the image period is corrected and the image artifacts are compensated, though some residual artifacts of image non-uniformity can still be observed. These errors are due to the algorithm, since the captured image profile matches the predicted image profile in Fig. 12(b). Note that the performance of the digital correction highly depends on the image contents as well as the image displacement. Additionally, note that the optimized image performance may contain large rendering errors under some specific image periods or image depths because of the residual in optimization.

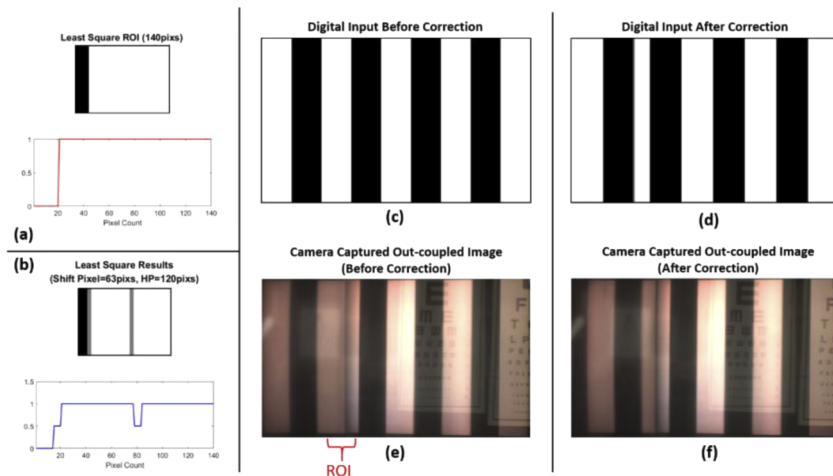


Fig. 12. Digital correction of the MRs-induced ghost image. (a) The desired image profile of the local area where LLS optimization is adopted. (b) The optimization results of the least square optimization region. (c) and (d): Digital inputs before and after digital correction. (e) and (f): Out-coupled images captured by camera before and after digital correction. The red bracket denotes the area of optimization as shown in (a) and (b).

7. Conclusion

In this paper, we have investigated the imaging artifacts and performance for an MMA-based lightguide system with a finite or variable focal depth. We have analyzed and classified the

ray path splitting and imaging artifacts induced by the change of the image focal depth, and presented a systematic approach to simulate the perceived retina image for quantifying the effects of focal depth on the image quality and artifacts by using a non-sequential ray tracing method and diffraction theory. Through the simulation, we have established estimation of the depth and position of the perceived image as well as the retinal image performance. Through a prototype set up with a commercially available MMA lightguide, we experimentally demonstrated these effects and the experimental results show good agreement with the simulation results. Finally, we present a digital correction method to correct the major artifacts induced by the non-collimated ray bundles focused at finite depths. Future work may be done for further developing a novel image generator or designing a new lightguide structure to minimize image artifacts in hardware. Moreover, user studies regarding to the actual depth perception and visual comfort should be done to prove the validity of the proposed vari-focal HMD system.

Disclosures

Dr. Hong Hua has a disclosed financial interest in Magic Leap Inc. The terms of this arrangement have been properly disclosed to The University of Arizona and reviewed by the Institutional Review Committee in accordance with its conflict of interest policies.

References

1. B. C. Kress and W. J. Cummings, "11-1: Invited Paper: Towards the Ultimate Mixed Reality Experience: HoloLens Display Architecture Choices," *SID Symp. Dig. Tech. Pap.* **48**(1), 127–131 (2017).
2. T. Levola and P. Laakkonen, "Replicated slanted gratings with a high refractive index material for in and outcoupling of light," *Opt. Express* **15**(5), 2067–2074 (2007).
3. P. Äyräs, P. Saarikko, and T. Levola, "Exit pupil expander with a large field of view based on diffractive optics," *J. Soc. Inf. Disp.* **17**(8), 659–664 (2009).
4. P. Saarikko, "Diffractive exit-pupil expander for spherical light guide virtual displays designed for near-distance viewing," *J. Opt. A: Pure Appl. Opt.* **11**(6), 065504 (2009).
5. Z. Lv, J. Liu, J. Xiao, and Y. Kuang, "Integrated holographic waveguide display system with a common optical path for visible and infrared light," *Opt. Express* **26**(25), 32802–32811 (2018).
6. J. D. Waldern, A. J. Grant, and M. M. Popovich, "17-4: DigiLens AR HUD Waveguide Technology," *SID Symp. Dig. Tech. Pap.* **49**(1), 204–207 (2018).
7. J. Xiao, J. Liu, Z. Lv, X. Shi, and J. Han, "On-axis near-eye display system based on directional scattering holographic waveguide and curved goggle," *Opt. Express* **27**(2), 1683–1692 (2019).
8. C. Yoo, K. Bang, C. Jang, D. Kim, C. K. Lee, G. Sung, and B. Lee, "Dual-focal waveguide see-through near-eye display with polarization-dependent lenses," *Opt. Lett.* **44**(8), 1920–1923 (2019).
9. J. D. Waldern, R. Morad, and M. M. Popovich, "Waveguide Manufacturing for AR Displays, Past, Present and Future," *Frontiers in Optics, FW5A-1* (2018).
10. G. Quaranta, G. Basset, O. J. Martin, and B. Gallinet, "Recent advances in resonant waveguide gratings," *Laser Photonics Rev.* **12**(9), 1800017 (2018).
11. Y. Amitai, "P-21: Extremely Compact High-Performance HMDs Based on Substrate-Guided Optical Element," *SID Symp. Dig. Tech. Pap.* **35**(1), 310–313 (2004).
12. Y. Amitai, "P-27: A Two-Dimensional Aperture Expander for Ultra-Compact, High-Performance Head-Worn Displays," *SID Symp. Dig. Tech. Pap.* **36**(1), 360–363 (2005).
13. D. Cheng, Y. Wang, C. Xu, W. Song, and G. Jin, "Design of an ultra-thin near-eye display with geometrical waveguide and freeform optics," *Opt. Express* **22**(17), 20705–20719 (2014).
14. B. Pascal, D. Guilhem, and S. Khaled, "Optical guide and ocular vision optical system," U.S. Patent, No. 8,433,172 (2013).
15. K. Sarayeddine, K. Mirza, P. Benoit, and X. Hugel, "Monolithic light guide optics enabling new user experience for see-through AR glasses," *Photonics Applications for Aviation, Aerospace, Commercial, and Harsh Environments V* 9202-92020 (2014).
16. K. Sarayeddine and K. Mirza, "Key challenges to affordable see-through wearable displays: the missing link for mobile AR mass deployment," *Photonics Applications for Aerospace, Commercial, and Harsh Environments IV* 8720-87200 (2013).
17. Q. Wang, D. Cheng, Q. Hou, Y. Hu, and Y. Wang, "Stray light and tolerance analysis of an ultrathin waveguide display," *Appl. Opt.* **54**(28), 8354–8362 (2015).
18. M. Xu and H. Hua, "Ultrathin optical combiner with microstructure mirrors in augmented reality," *Proc. SPIE* **10676**, 1067614 (2018).

19. M. Xu and H. Hua, "Methods of optimizing and evaluating geometrical lightguides with microstructure mirrors for augmented reality displays," *Opt. Express* **27**(4), 5523–5543 (2019).
20. <https://letinar.com/>
21. H. Hua, "Enabling focus cues in head-mounted displays," *Proc. IEEE* **105**(5), 805–824 (2017).
22. M. Lambooi, M. Fortuin, I. Heynderickx, and W. IJsselstein, "Visual discomfort and visual fatigue of stereoscopic displays: A review," *J. Imaging Sci. Technol.* **53**(3), 030201 (2009).
23. D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-Accommodation Conflicts Hinder Visual Performance and Cause Visual Fatigue," *J. Vis.* **8**(3), 33 (2008).
24. S. J. Watt, K. Akeley, M. O. Ernst, and M. S. Banks, "Focus cues affect perceived depth," *J. Vis.* **5**(10), 7 (2005).
25. S. Liu, H. Hua, and D. Cheng, "A novel prototype for an optical see-through head-mounted display with addressable focus cues," *IEEE Trans. Visual. Comput. Graphics* **16**(3), 381–393 (2010).
26. D. Dunn, C. Tippets, K. Torell, P. Kellnhofer, K. Akşit, P. Didyk, and H. Fuchs, "Wide field of view varifocal near-eye display using see-through deformable membrane mirrors," *IEEE Trans. Visual. Comput. Graphics* **23**(4), 1322–1331 (2017).
27. Y. Jo, S. Lee, D. Yoo, S. Choi, D. Kim, and B. Lee, "Tomographic projector: large scale volumetric display with uniform viewing experiences," *ACM Trans. Graph.* **38**(6), 1–13 (2019).
28. K. Akşit, W. Lopes, J. Kim, P. Shirley, and D. Luebke, "Near-eye varifocal augmented reality display using see-through screens," *ACM Trans. Graph.* **36**(6), 1–13 (2017).
29. T. Shibata, T. Kawai, K. Ohta, M. Otsuki, N. Miyake, Y. Yoshihara, and T. Iwasaki, "Stereoscopic 3-D display with optical correction for the reduction of the discrepancy between accommodation and convergence," *J. Soc. Inf. Disp.* **13**(8), 665–671 (2005).
30. S. Shiwa, K. Omura, and F. Kishino, "Proposal for a 3-D display with accommodative compensation: 3DDAC," *J. Soc. Inf. Disp.* **4**(4), 255–261 (1996).
31. J. Schwiegerling, *Field guide to visual and ophthalmic optics* (SPIE, 2004).
32. X. Hu and H. Hua, "High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics," *Opt. Express* **22**(11), 13896–13903 (2014).
33. S. Liu and H. Hua, "Extended depth-of-field microscopic imaging with a variable focus microscope objective," *Opt. Express* **19**(1), 353–362 (2011).
34. J. P. Rolland and H. Hua, "Head-mounted display systems," *Encyclopedia of optical engineering* 2 (2005).
35. P. E. Romano, J. A. Romano, and J. E. Puklin, "Stereoacuity development in children with normal binocular single vision," *Am. J. Ophthalmol.* **79**(6), 966–971 (1975).