Wedge guides and pupil steering for mixed reality

Adrian R.L. Travis (SID Member)
Living Chen (SID Member)
Andreas Georgiou
Jiaqi Chu
Joel Kollin

Abstract — We suggest folding the optics of a classic-mixed reality display into a slim wedge light guide with holographic combiner so that the pupil of a projector near the ear is imaged onto the pupil of the eye. Such folding makes the pupil of the virtual image small, so when gaze wanders, the image pupil should be switched between appropriate positions by shearing the combiner. Modeling indicates that if a phase plate corrects distortion for the ray to the fovea center, the rest of the image can be scanned at adequate resolution provided that a cylindrical mirror varies focal power in the horizontal.

Keywords — virtual reality, augmented, wedges, holograms, light guides, foveal.

DOI #10.1002/jsid.629

1 Introduction

Virtual and mixed reality displays will only become commonplace if they look good, and for most people, that means they have to be slim, ideally like a pair of designer spectacles. Enthusiasts ask for unlimited field of view (e.g., 120°), and pragmatists ask for good resolution (e.g., 2000 pixels per radian), but the etendue of a slim system means that such image quality is only feasible with a very small eye box. The equivalent of a big eye box is needed if a display is to fit without fuss, and a successful solution has been to expand the pupil of the virtual image by reflecting it repeatedly in a slab waveguide. If but if the reflected versions are to mate, the guides must be highly parallel and flat thus limiting field of view.

We suggest that the image pupil instead be steered so as to always coincide with the pupil of the eye, as illustrated in Fig. 1. This relaxes guide tolerance so that the guide might be made of plastic and curved. Aberrations introduced by the curvature and flexure must be corrected by the projector, but such correction is in any case a requirement if the focus of the virtual image is to be variable in the manner needed to overlie mountains or tabletops.

2 Light guide

Many early head-up displays and augmented reality displays worked by pointing a projector at a partially reflective curved surface as shown in Fig. 2. The image from the projector is brought to a focus in the focal plane of the curved surface that itself is angled so as to allow the eye to see the virtual image. In cockpit head-up displays, the curved surface is a volume hologram because it reflects all light from the projector but is transparent to most light from the outside world.

Wedge light guides were first proposed as a way of folding the light between a video projector and screen for rear projection television. They can perform the same function when the image is virtual as illustrated schematically in Fig. 3. The principle of projection through a wedge is that each time a ray reflects off one surface, the ray’s angle with respect to the opposite surface reduces. Eventually, the critical angle is reached, and the ray emerges into air. Much as in free space, the angle at which the ray is injected will determine how far it travels before emerging and best is if all rays undergo the same number of reflections before exit because this prevents the projected image being broken into bands. For that to happen, thickness of the wedge versus length must be nonlinear, so Fig. 3 is strictly correct only if drawn on a curved surface.

In order to obtain a wide field of view, the guide must curve round the head of the viewer. A curved wedge guide was designed to be made from polycarbonate with an outside surface that has a constant radius of curvature of $r = 82$ mm. The distance between the outside and inside surfaces at $y = 0$ is $d = 5.211$ mm, and the inside surface is defined by the equation:

$$z = \frac{y^2/R}{1 + \sqrt{1 - y^2/R^2}} + \sum_{i=1}^{5} \beta_i y^i;$$

where radius, $R = 77$, $\beta_1 = -0.030$, $\beta_2 = -7.457 \times 10^{-4}$, $\beta_3 = 4.402 \times 10^{-6}$, $\beta_4 = 3.500 \times 10^{-8}$, and $\beta_5 = -2.654 \times 10^{-9}$ (all dimensions in millimeters).

Received 12/19/17; accepted 02/09/18.
Adrian R.L. Travis is with the Clare College, Cambridge University, Cambridge CB2 1TL, UK; e-mail arlt1@cam.ac.uk.
Living Chen is with the Futurewei Technologies Inc, Huawei R&D USA, 10900 NE 8th St, Suite 500, Bellevue, WA 98004, USA.
Andreas Georgiou is with the Microsoft Research Cambridge, 21 Station Road, Cambridge CB1 2FB, UK.
Jiaqi Chu is with the Engineering Department, Cambridge University, Cambridge CB2 1PZ, UK.
Joel Kollin is with the Microsoft Corporation, One Microsoft Way, Redmond, WA 98052, USA.
© 2018 Society for Information Display 1071-0922/18/2609-0629$1.00.
The guide was modeled as a nonsequential component in Zemax by a back-to-back pair of Biconic Zernike Lenses whose shared face was the Front Radius. The radial height was 56.096 mm, and X Half-Width was 15 mm. A pair of prisms was added at the thick end by way of an input bevel with a mirror of cylindrical curvature 82 mm extending the outer radius through the prisms. The hologram was modeled by an optically fabricated hologram that Zemax allowed only in sequential optics. The nonsequential components were therefore placed within a sequential environment with the exit port just outside the prism pair and the input port configured as an elliptical grating adjacent to the hologram. The hologram was specified to diffract all rays launched at the center of the thick end of the wedge into the center of an eye looking at the wedge from a distance of 14 mm. Sequential optics are needed to define an optically fabricated hologram, and in order to do this, the guide was expressed as a repeated pair of Biconic Zernike surfaces. The deformable mirror of Fig. 10 was expressed by a flat mirror plus a Zernike Standard Phase surface. By way of verification, the guide and hologram were also modeled in Optalix.

The model predicted that rays injected at the thick end will undergo 14 internal reflections before reaching the critical angle provided that they are launched via a pupil spanning 3.05 mm and within an angular range of 0.695 radians (40°), that is, a cylindrical etendue of 2.12 mm radians. This is sufficient for our virtual image to have the intended resolution and field of view (120°) where etendue must be 1 mm × 2.09 radians.

Figure 4 shows the model, and a perfect virtual image is assumed where each pixel comprises a ray bundle with diameter sufficient for a resolution of 2000 pixels per radian, that is, the diameter of the ray bundle is 2000 wavelengths which is approximately 1 mm. Rays are traced backwards through the hologram and guide to the entry and for a field of view of 115°, and the ray trace predicts that all rays will reach the input.

The guide surfaces can have spherical or cylindrical symmetry – cylindrical was chosen because it proves easier to make. Either way, the model predicts considerable astigmatism as shown by Figs 5 and 6 where rays from the thick end are passed through an ideal lens whose focal length was 7.143 mm in the sagittal plane and 10 mm in the tangential plane.

A curved wedge with similar profile was machined from acrylic ($r = 82$, $d = 4.127$, $R = 77$, $\beta_1 = -0.036$, $\beta_2 = -7.150 \times 10^{-4}$, $\beta_3 = 4.233 \times 10^{-6}$).

The guide surfaces can have spherical or cylindrical symmetry – cylindrical was chosen because it proves easier to make. Either way, the model predicts considerable astigmatism as shown by Figs 5 and 6 where rays from the thick end are passed through an ideal lens whose focal length was 7.143 mm in the sagittal plane and 10 mm in the tangential plane.

A curved wedge with similar profile was machined from acrylic ($r = 82$, $d = 4.127$, $R = 77$, $\beta_1 = -0.036$, $\beta_2 = -7.150 \times 10^{-4}$, $\beta_3 = 4.233 \times 10^{-6}$).

The guide was modeled as a nonsequential component in Zemax by a back-to-back pair of Biconic Zernike Lenses whose shared face was the Front Radius. The radial height was 56.096 mm, and X Half-Width was 15 mm. A pair of prisms was added at the thick end by way of an input bevel with a mirror of cylindrical curvature 82 mm extending the outer radius through the prisms. The hologram was modeled by an optically fabricated hologram that Zemax allowed only in sequential optics. The nonsequential components were therefore placed within a sequential environment with the exit port just outside the prism pair and the input port configured as an elliptical grating adjacent to the hologram. The hologram was specified to diffract all rays launched at the center of the thick end of the wedge into the center of an eye looking at the wedge from a distance of 14 mm. Sequential optics are needed to define an optically fabricated hologram, and in order to do this, the guide was expressed as a repeated pair of Biconic Zernike surfaces. The deformable mirror of Fig. 10 was expressed by a flat mirror plus a Zernike Standard Phase surface. By way of verification, the guide and hologram were also modeled in Optalix.

The model predicted that rays injected at the thick end will undergo 14 internal reflections before reaching the critical angle provided that they are launched via a pupil spanning 3.05 mm and within an angular range of 0.695 radians (40°), that is, a cylindrical etendue of 2.12 mm radians. This is sufficient for our virtual image to have the intended resolution and field of view (120°) where etendue must be 1 mm × 2.09 radians.

Figure 4 shows the model, and a perfect virtual image is assumed where each pixel comprises a ray bundle with diameter sufficient for a resolution of 2000 pixels per radian, that is, the diameter of the ray bundle is 2000 wavelengths which is approximately 1 mm. Rays are traced backwards through the hologram and guide to the entry and for a field of view of 115°, and the ray trace predicts that all rays will reach the input.

The guide surfaces can have spherical or cylindrical symmetry – cylindrical was chosen because it proves easier to make. Either way, the model predicts considerable astigmatism as shown by Figs 5 and 6 where rays from the thick end are passed through an ideal lens whose focal length was 7.143 mm in the sagittal plane and 10 mm in the tangential plane.

A curved wedge with similar profile was machined from acrylic ($r = 82$, $d = 4.127$, $R = 77$, $\beta_1 = -0.036$, $\beta_2 = -7.150 \times 10^{-4}$, $\beta_3 = 4.233 \times 10^{-6}$).

The guide surfaces can have spherical or cylindrical symmetry – cylindrical was chosen because it proves easier to make. Either way, the model predicts considerable astigmatism as shown by Figs 5 and 6 where rays from the thick end are passed through an ideal lens whose focal length was 7.143 mm in the sagittal plane and 10 mm in the tangential plane.

A curved wedge with similar profile was machined from acrylic ($r = 82$, $d = 4.127$, $R = 77$, $\beta_1 = -0.036$, $\beta_2 = -7.150 \times 10^{-4}$, $\beta_3 = 4.233 \times 10^{-6}$).
\[ \beta_4 = -1.342 \times 10^{-8} \text{, and } \beta_5 = -1.926 \times 10^{-9} \]. A thin film of ultraviolet-curing adhesive was squeezed between each surface, and a sheet of thin (100–200 \( \mu \)m thick) float glass that itself was supported by a mold of the surface machined from acrylic transparent to ultraviolet. The film was left 10 min for thickness to equalize then cured with ultraviolet, and the result was surfaces with the accuracy of numerical machining (\( \pm 25 \, \mu \)m) but the smoothness of float glass (from 1 to 3 nm Ra).

A pair of achromats relayed the image from a beam-scanning projector into the wedge entrance, and the distance between these achromats was adjusted for tangential focus while a cylindrical lens was added for sagittal focus: the result is shown in Fig. 7. At the point of optimal focus, resolution was only 140 \( \mu \)m versus a target of 7 \( \mu \)m (14 mm back vertex distance divided by 2000 pixels per radian). A video projector makes a poor image source, and both a better source and more carefully adjusted lenses will allow us to measure the properties of the guide more precisely.

Volume holograms of the kind needed to concentrate light emerging from the guide into the viewer’s eye have been recorded, but the quality is still not high enough for useful measurements to be made.

### 3 Pupil steering

In order for the pupil of the virtual image to be steered, it is proposed that holograms for several pupil positions be recorded in the combiner and that this is then subjected to shear so that only one of these holograms is active at any instant. By shear, we mean that the top surface of the combiner is translated relative to the bottom surface such that, were the combiner to be homogeneous, the material in between would be translated by a distance proportional to its distance from the bottom surface.

The hologram that deflects light from the guide to the viewer is a volume grating, and such gratings diffract only if the Bragg condition is met as shown in Fig. 8. If the grating is sheared, that is, if the top surface of the film in which the grating is recorded is moved in a parallel but opposite direction to the bottom surface, then the planes that constitute the grating must rotate. The component of spatial frequency of the grating in the plane of the film cannot alter, so it must be the component of spatial frequency perpendicular to this plane that alters, that is, the...
grating vector shears in a direction orthogonal to that of the shear of the film. With sufficient shear, the Bragg condition is no longer met, and the grating ceases to diffract. Many volume holograms can be recorded in a single holographic film, and the intention here is that each hologram becomes active at a different degree of shear.

Figure 9 shows the far-field pattern as shear was applied to a stack of three films of commercial photopolymer glued together. There was one grating in each film, each orientated to diffract the input ray in a unique direction, and each chosen so as to become active at a unique amount of shear.

There are electronic ways of controllably deflecting a ray of light, for example, liquid crystal prisms and electro-wetting prisms,7 that might seem less prone to failure than a mechanical approach like shearing. But the pupil steerer must be transparent, not only so that the viewer’s eyes can see outside but also so that they can be seen, important for key applications like teleconferencing. Volume holograms excel for mixed reality because they are transparent to most transient light and, importantly, have a spatial frequency too high for the sun or car headlights to form rainbow images due to second-order diffraction. Each of the recorded holograms can deflect light to any desired angle, whereas alternatives tend to struggle to achieve more than a few degrees of deflection, whether variable or fixed.8 Even so, mechanical actuation tends to consume much energy - how much energy does shearing use?

Commercial off-the-shelf photopolymer9 is 20 microns thick and has a shear modulus G of 200 kPa, that is, three times more flexible than rubber. The equations of Kogelnik10 indicate that a shear of 3 microns should be sufficient for an 85:1 extinction ratio, that is, a shear angle of \( \theta = 3/20 \) radians, and if the area A of the film is 30 mm by 50 mm, then the force needed to shear it should be \( F = GA\theta = 45 \) N. The energy needed to switch between two holograms is therefore predicted to be 135 \( \mu \)J. A 10 mm piezoelectric transducer will extend 10 microns at standard voltages (0–150 V) with a blocking force of 330 N.

In reality, a shear of several hundred microns is needed because off-the-shelf photopolymer comes sandwiched between thick protective layers. Furthermore, switching between the 10th and 11th recorded hologram will require 10 times the force and 10 times the extension as to switch between the first and second. The extension of a piezoelectric can be magnified by, for example, a factor of 10 by tilting it at a small angle in a constrained package, but the force behind the actuation is reduced by at least the same factor. The simplest way to reduce the force requirement would be to thicken the photopolymer layer because this reduces the shear angle between holograms: photopolymer as thick as 1.5 mm has been reported.11 More radical would be to make photopolymer that is less stiff: volume holograms can be recorded by ultraviolet light in off-the-shelf polydimethylsiloxane doped with benzophenone12, and the stiffness of polydimethylsiloxane can be reduced to a few kPa by diluting the mixing ratio.13

4 Adaptive projector

The classic way to project a pre-distorted image is with a liquid crystal display and an appropriately designed lens, but both would likely be big for so much distortion over such a wide field of view. In any case, there are further requirements: firstly, the linewidth of the red, green, and blue constituents of the image must be narrow if the
combiner is to redirect rays without blur. Secondly, variable focus is desirable so that the virtual image can appear at the same distance from the viewer as the real object on which the image is superposed. Lastly, left and right eye views must be aligned very precisely, a requirement so stringent that the inherent alignment in pupil expansion systems has been a key factor in promoting their development.

Narrow linewidths are the preserve of lasers, so while holographic projection and beam scanning are options, it will require imagination for this approach to work well with light emitting diode arrays. Holographic projection can in principle provide any pre-distortion and focus that is desired. But the calculation of holograms takes time and power, and advances will be needed to reduce these to acceptable levels. Laser beam scanners are small, while the image can be updated as it is written so scanners would be convenient if they could be configured to have adequate pre-distortion and variable focus.

A laser beam could be pre-distorted by reflection off an appropriately shaped surface before launch into the curved guide so that the beam reaches the fovea undistorted. The beam has to be scanned, so the surface shape must change quickly, but deformable mirrors developed for adaptive optics can change shape in less than a millisecond. Figure 10 shows a model with rays traced backwards through the polycarbonate wedge, reflected off such a mirror, and then focused via an ideal toroidal lens (X-power = 0.11 mm⁻¹, Y-power = 0.059 mm⁻¹, distance to focal plane = 10.4 mm, angle of focal plane = 5.51°). The surface of the mirror was optimized to eliminate all distortion, and then the angle of the ray leaving the pupil of the eye was varied both in the vertical and the horizontal.

Spot diagrams at the focal plane are shown in Fig. 11, and within 6° above and below the origin, the spot remains within the diffraction limit (delineated in red). At >6°, the spot becomes blurred but also shown on the diagram is how the resolution of the human eye degrades with angle, and it is clear that the laser beam can be scanned in the vertical without dynamic correction.

An expanded version of Fig. 11 is shown in the upper part of Fig. 12, and some form of dynamic refocusing is needed in the horizontal. In the lower part of Fig. 12, the y-power of the toroidal lens is changed versus horizontal angle showing that a change in cylindrical focus alone is sufficient to keep spot size finer than the resolution of the human eye. This might, for example, be effected by a cylindrical mirror resonating at the same frequency as horizontal scan; if fast scan were in the vertical, horizontal mirror curvature might be altered by little more than a speaker coil.

---

**FIGURE 10** — A phase modulating deformable mirror and ideal toroidal lens was added to the model.

**FIGURE 11** — Modeling predicts that if a ray is pre-distorted for the fovea looking straight ahead, then the ray can scan in the vertical without exceeding peripheral resolution of the eye.

**FIGURE 12** — Modeling predicts that when scanning in the horizontal, a mirror with variable cylindrical focus will keep projected resolution finer than peripheral resolution.
As the orbit rotates away from looking straight ahead, the phase-modulating mirror must change so that the beam continues to reach the fovea undistorted. Figure 13 shows a plot of the phase profiles less tilt that are needed at the extremes of view. In each diagram, the difference between blue (minimum) and red (maximum) phase is given by the number of waves specified underneath.

The variation is everywhere less than 54 waves, small enough to be corrected by, for example, the deformable mirror DM69-08 from Alpao except that its diameter is 6.1 mm versus the 4.4 mm diameter assumed in our model. We calculated the power that would be needed by the DM69-08 for each deformation, and it ranged from 80 mW for the central pixel (0°, 0°) to 2.2 W for the nose-side downward pixel (−30°, −20°). These powers are very high, but the orbit rotates slowly, so there is no need for the switching speed of a deformable mirror. A hologram displayed on a liquid crystal panel would suffice to produce the patterns of Fig. 13, and these holograms could be recorded in memory, avoiding the power consumed when holograms are calculated in real time. It may be, therefore, that a holographic/beam scanning hybrid gives the lowest power solution to the pre-distorting projector needed for the curved guide.

The scanning mirror in a laser projector typically has dimensions of 1 mm by 1 mm, whereas the entry pupil of the curved guide is 3.4 mm by 4.2 mm. Even if this were not so, there is the further problem that the resolution of a 120° by 90° image at 2000 pixels per radian is approximately 4 k by 1.5 k, whereas beam scanners typically manage little more than 500 lines at 60 Hz. Once again, the low resolution of peripheral vision may be of help. The figure of merit governing laser beam scanners is the product of mirror dimension, scan angle, and scanning frequency, so a large mirror is in principle permissible provided that the scanning angle is small.

A large mirror and small scanning angle are exactly what is needed for the fovea. As for peripheral vision with its lower resolution, this might be addressed by a smaller mirror scanning over a large angle albeit at much the same scanning frequency. Figure 14 shows a graph of angle per pixel versus scanning angle for a suggested combination of three projectors operating at frame rates of 120 Hz (in brown) underneath a curve showing the resolution typical of the human eye. In order to scan ±60° about the fovea, the number of lines that each projector would need to scan is 419 (about fovea with 1 mm pupil at eye), 333 (intermediate with 0.24 mm pupil at eye), and 359 (periphery with 0.09 mm pupil at eye).

![Figure 13](image1.png)

**FIGURE 13** — Number of waves of predistortion required for the display of various virtual pixels.

![Figure 14](image2.png)

**FIGURE 14** — Separate projectors might be used for the fovea, periphery and extreme periphery of vision.
If the hope is to display high resolution only to the fovea, then a steering mirror is needed to deflect light from the high resolution projector to the region under observation by the fovea. The problem with this is that the angle of this steering mirror must be known to great accuracy, otherwise the low and high resolution parts of the image will not mate. This is a similar problem to that of aligning left-eye and right-eye views, and no doubt, there will be other distortions as the optics flex: how do we measure them? A similar problem is faced in astronomy where fluctuations in air temperature distort the wave fronts of light arriving from a star. The solution is to project an artificial “guide star” on the sodium layer 90 km up in the atmosphere and to measure the distortion of wave fronts arriving at the telescope from this point. That makes it possible to deduce the aberrations that are introduced by the atmosphere then to deform a mirror so that it corrects these aberrations. The same mirror can then correct wave fronts arriving from the star under observation as they reflect off the mirror. This technique, known as adaptive optics, can in principle be applied to projection displays over a small field of view.

Many headsets already project an infrared grid of dots—the equivalent of a guide star—onto the viewer’s surroundings in order to help registrate the headset to its environment. The direction of rays in a wedge light guide and hologram—the equivalent of an aberrating atmosphere—is easily reversed, and the guides can act equally well as cameras as they can as projection displays. The suggestion here therefore is to look at the infrared dots via the steering mirror, guide, and hologram in order to find the exact position of the steering mirror and register the left eye view to right eye view. This is shown conceptually in Fig. 15.

5 Conclusions

Modeling of a slim curved wedge guide predicts a virtual image with a wide field of view but a small eye box and the need for much pre-distortion. It is suggested that the eye box be enlarged by sheering the holographic combiner so as to steer the pupil of the virtual image into the pupil of the eye. Modeling indicates that a beam-scanning projector could adapt to the position of the fovea if the beam is pre-distorted by a liquid crystal hologram and varifocal cylindrical mirror.

Acknowledgments

The authors would like to thank Neil Emerton and Tim Large who developed techniques for flat wedges that have made much of this work possible.

References


FIGURE 15 — Project an infrared grid in front of the curved guide and return rays might be coupled in via a hologram so that waves emerging from the thick end define the distortions imposed by the guide.
Adrian Travis was a lecturer at Cambridge University from 1988 to 2007 then a scientist at Microsoft Corporation until 2017. He is a fellow emeritus of Clare College, Cambridge University.

Liying Chen is a senior architect in optics at Futurewei Technologies, Inc. She provides innovative solutions and technical expertise to customers in the areas of imaging, display, and sensing. Dr. Chen was an optics researcher in the Applied Sciences Laboratory at Microsoft from 2009 to 2016, where she studied wedge optics and diffractive optical elements and their applications in imagers, biometric sensors, and head-mounted displays. In 2004, she earned her PhD in optical sciences from the University of Arizona, where she remained as a research scientist in the Center for Gamma-Ray Imaging until 2009. She invented a small animal imager to capture particle emissions directly in vivo while working there. Dr. Chen is a member of the OSA, SPIE, and SID.

Andreas Georgiou received his first degree and PhD from the Engineering Department of University of Cambridge in 2001 and 2006, respectively. He started working in Holography during his Master’s in 2000 and continued to do so during his PhD where he investigated ways to optimize holographic optical switches for telecoms. After his PhD, he continued working in the Cambridge of University in collaboration with Alps Electric where within a team built one of the first color holographic projectors. In 2008, he moved to the Open University working in the Planetary and Space Sciences team. There, he designed the optics of the ultraviolet-visible spectrometer of the then planned ExoMars mission to the Red Planet. Between 2010 and 2012, he worked at Imperial College, London, where he developed optical instruments for minimally invasive surgery. From 2012 onwards, he has been with Microsoft Research, Cambridge, looking at novel display and sensing technologies. He is particularly interested in how digital and analogue holography can be used in future consumer display devices.

Jiaqi Chu received her BSc in Electrical Engineering from Shanghai Jiao Tong University and then did a PhD in Photonics at University of Cambridge. In her doctoral studies, she focused on orbital angular momentum of light to carry more information and implement three-dimensional displays. She has worked on carrying two-dimensional amplitude information with orbital angular momentum modes in the past, and she has recently started to study integrating holograms with orbital angular momentum multiplexing.

Joel Kollin has been pioneering the next generation display and human-computer interface technology for over 30 years, including the original holovideo system at the MIT Media Lab, the Virtual Retinal Display system at the University of Washington’s Human Interface Technology Lab, and dynamic parallax barrier and foveal display technologies at the NYU Media Research Lab. He has also designed large-scale immersive installations for artistic and educational purposes. He is currently a Principle Research Engineer in the Hardware, Devices, and Experiences group at Microsoft Research NExT in Redmond WA, working on the future of immersive display with some incredibly talented colleagues.