

Video See-through Design for Merging of Real and Virtual Environments

Emily K. Edwards, Jannick P. Rolland and Kurtis P. Keller
University of North Carolina, Computer Science Department
Chapel Hill, NC 27599-3175, USA

Abstract A head mounted display (HMD) that is optically opaque with respect to the outer world can be provided with see-through capability by mounting video cameras to the outside of the helmet. Stereoscopic views of the physical space surrounding the wearer are captured by the video cameras and are projected to the display screens inside the HMD. In this paper, the authors describe the conditions necessary for mounting the cameras to a pre-existing, non-see-through HMD before discussing the advantages and disadvantages of different mounting schemes. We finally describe the design of the video cameras lenses to be used with this specific application.

1. Introduction

A head-mounted display (HMD) consists of a stereoscopic viewer and a head motion tracking device mounted on a helmet. The viewer covers the eyes of the wearer and allows him or her to see magnified images of two screens inside the helmet where stereoscopic views of a scene of interest are displayed. The images are modified with input from the tracker which provides information to update the images in real time. One type of HMD is a totally immersive display in which the wearer of the HMD does not see the outside world. This paper deals with converting a totally immersive type of HMD into a video see-through HMD which involves the merging of computer generated images with real-world images (Bajura, et al., 1992). The real world, in this case, is acquired via two miniature video cameras mounted on the helmet. The acquired video signal from the miniature cameras is then transmitted to LCD screens mounted inside the helmet. The wearer sees both the view outside the helmet that he or she would see with naked eyes and computer-generated graphics that have been created, in our case, on Pixel-Planes 5, a massively

parallel graphic engine developed at the University of North Carolina at Chapel Hill under the direction of Henry Fuchs and John Poulton (Fuchs, et al., 1989). Each video-acquired image of the outside world and the computer generated graphics are combined into a single video signal before they serve as inputs to the LCD screens.

We shall first give an overview of the system that consists of the HMD combined with the two miniature cameras and discuss the necessary imaging conditions for merging the real world with a computer simulated world. A description of the video camera lenses is then given with its design considerations and performance.

2. System Considerations

The pre-existing totally immersive HMD that we shall refer to in this paper is the Flight Helmet from Vision Research which uses LEEP optics as the viewer and two color LCDs as displays for the video inputs (Vision Research), (LEEP, 1990). These video inputs are the combination of the computer generated output from Pixel-Planes 5 and the video output from the cameras. The video images of the outside scene are first acquired via charge coupled device (CCD) detectors mounted in the cameras and then fully painted on the LCD display screens after video transmission and passage through the LCD electronics. The term "fully painted" means that the overscan of the video signals transmitted from the cameras to the LCDs that would have been caused by the LCD electronics was measured to be negligible. This is especially important to the determination of the focal length of the video lenses because the focal length is a function of the field of view (FOV) and the imaging area on the CCD detector. If only part of the detector area were to be mapped to the LCD screens, the projected outside scene would appear magnified since the LEEP optics would project a smaller FOV to a larger one. Hence the camera lenses FOV must exactly match the LEEP optics FOV.

For the projected image of the outside scene to appear realistic to the wearers eyes and provide a view of the outside world identical to the view the wearer would see if he or she were not wearing the helmet, the two miniature cameras must act as the eyes of the person wearing the helmet. For

this to happen, the cameras have to be positioned with respect to the HMD as if they occupied the eyes physical locations. This can be theoretically accomplished using a 45° folding mirror per eye as shown in Figure 1. For a horizontal FOV of 75°, which is the LEEP optics' and video camera's FOV, a 530mm x 540mm mirror is then required (Robinett and Rolland, 1992); however, for stereo HMDs, the size of this mirror is impractical and also usable for only one eye. Vertically mounted camera arrangements yield similar, bulky packages.

Many different arrangements of cameras and mirrors were experimented with to find the best combinations of correct Inter-Pupillary-Distance (IPD), on-axisness and superimposition of the cameras nodal points (CNP) with the eyes nodal points (ENP). The latter condition will be referred to as CNP to ENP distance. Experiments showed that correct IPD is very important to the user for close to medium viewing distance in deciphering correct distance through stereoscopic viewing. Off-axisness (cameras viewing from above or below the eyes line of sight) can be hard to get used to when objects are close and can cause disorientation when walking. Placing the cameras in front of the eyes as in Figure 2a and 2b result in the image of the outside world seen through the HMD appear magnified and closer than they actually are. The difference in this case is that we can still perfectly register the real and virtual environments by choosing as the eye point for computing the stereo images of the computer generated scenes the actual location of the nodal points of the two cameras (instead of the physical locations of the wearer's eyes). The potential problem is that the two senses of vision and proprioception (i.e. hand position) will be mismatched.

In all of the camera arrangements, the IPD of the system is designed to retain that of the LEEP optics in the helmet. Minor changes in the CNP to ENP distance may prove to be an acceptable tradeoff to reduce the size of the system. To retain the full FOV in both X and Y dimensions, several methods were conceived but all had either size, CNP to ENP distance, or off-axisness problems. The final arrangement keeps correct IPD, maintains on-axisness and has a CNP to ENP distance of 0 (Figure 3a and 3b) at the expense of a reduction in vertical FOV. This optimized and folded mirror arrangement maintains the large horizontal FOV so difficult in wide

angle, side viewing systems, but reduces the vertical FOV to 40 degrees giving an image proportion overall of about 2 to 1 (H x V). This is comparable to proposed wide screen TV dimensions.

3. Design Specifications

In order for the image of the outside world to appear dimensionally correct and register accurately with the real world after projection through the LEEP optics, the FOV of the camera lenses and the LEEP optics must be the same as previously discussed. The FOVs of the two systems were matched in the following way. The vertical and horizontal FOVs of the viewer inside the helmet are specified by the vertical and horizontal dimensions of the images of the LCD screens optically formed through the optics and their distances with respect to the eyes of the wearer, respectively. Under the paraxial approximation, the FOV is given by 2θ with

$$\theta = \tan^{-1} [y / D] \quad (1)$$

where y is the height of the image of the LCD screens as measured from the optical axis, and D is their paraxial distance to the eyes of the observer (Longhurst, 1973). Given the height of the LCD screens used in the Flight Helmet, and their positions with respect to the LEEP optics, a paraxial raytrace through the LEEP optics yield values for y and D of 595.2mm and 1178.4mm, respectively. This is equivalent to a paraxial FOV of 26.8° according to Equation (1). Since this FOV needs to also be the FOV of the video lens, we use its value to calculate the correct focal length, f , for the video lens using Equation 1 where D equal f , y is now the height of the CCD detector as measured from its center and 2θ is the paraxial vertical FOV calculated for the Flight Helmet. We used the fact that the video lens is working at infinite conjugate in object space.. The camera detector size was given to be 6.4mm in width by 4.8mm in height, which yields a focal length value for the video lens of 4.75 mm (Panasonic Technical Support). The lenses must be optimized over the whole FOV at three well chosen FOV values along the largest dimension of the display, in this case the diagonal. Three FOV values are usually sufficient for rotationally symmetric systems.

A complication to the system, however, is that due to mechanical constraints, the LCD screens in the Flight Helmet are laterally offset from the optical axis of the LEEP optics by 6.5mm. One of the implications of such a horizontal offset is that the system needs to be optimized for a slightly larger FOV than otherwise necessary. The system was therefore optimized for a FOV of 43.2° instead of 39.8°, and we made sure that the system performed adequately at the corner of the screen farthest from the optical axis. A second even more critical implication of such an offset is that the center of distortion for the LEEP optics will not overlap with the center of distortion for the video cameras unless the CCD detectors are themselves offset by the appropriate amount with respect to the video lenses. We calculated that an offset of 0.75mm was necessary for the center of distortion to overlap while merging real and virtual scenes.

When considering the computer simulated images, a pre-distortion correction can be applied to them in order for the projected image through the LEEP optics to appear unwarped (Robinett and Rolland, 1992). The same cannot be done on the acquired video images using our current setup because of the direct output of the video signals to the LCDs. However, a fortunate occurrence was that the distortion of the LEEP optics as we proceed from the LCD screens to the outside world is of pincushion type and the distortion of the video optics as we proceed from the outside world to the CCD detector is of the barrel type. Thus while designing the video lens we constrained its distortion to equal to that of the LEEP optics. A plot of the distortion of the LEEP optics as we proceed from the outside world to the LCD screens and of the video optics as we proceed from the outside world to the CCD detectors are shown in Figures 4a and 4b respectively. The maximum distortion difference between the two plots was found to be 0.1% at the edge of the FOV.

4. Design Performance

A layout of the lens is shown in Figure 5. The ray fan plots in Figure 6 for four representative field angles show the performance of the lens over the full width of the detector. The design shows good performance over the full entrance pupil width of 2.8mm for a lens open at F/1.7. We

allowed for some vignetting at the edge of the field to produce a good MTF performance which is shown in Figure 7. The maximum frequency of 50 cycles corresponds to a 6.4mm x 4.8mm detector with 320 x 240 pixels. The field plot in Figure 8 shows that the residual astigmatism is well balanced over the full FOV.

Conclusion

The design and system considerations of a video see-through optical system using the Flight Helmet HMD from Vision Research has been discussed. The most important considerations are that the two miniature video cameras must ideally simulate the eyes of the person wearing the helmet. Thus they should be positioned with respect to the HMD as if they occupy the eyes physical locations. Due to the current design of the HMD, compromises in camera location and/or FOV were necessary. It is believed that a reduction in vertical FOV while maintaining perfect position of the cameras will present an easier to adapt display for the wearer. Further work with the system after assembly will include testing of this hypothesis.

Acknowledgments

The authors would like to thank especially Mike Bajura and Vern Chi for their stimulating discussions and suggestions concerning the specifications of the system. We would also like to acknowledge David Harrison for his assistance and expertise in the area of video electronics. Finally, we are very grateful to the principal investigator for the project, Henry Fuchs, who made this research possible. This work was funded under DARPA grant number DAEA18-90-C-0044.

References

Bajura, M., H. Fuchs, and R. Ohbuchi (1992). *Merging Virtual Objects with the Real World*, Computer Graphics (Proceedings of SIGGRAPH '92), **26** (2), pp 203 -10.

Fuchs, H., J. Poulton, J. Eyles, T. Greer, J. Goldfeather, D. Ellsworth, S. Molnar, and L. Israel (1989). *Pixel-Planes 5: A Heterogeneous Multiprocessor Graphics System Using Processor-Enhanced Memories*, Computer Graphics (Proceedings of SIGGRAPH '89), **23** (3), pp 79-88.

LEEP (1990). *Cyberface II applications note*. LEEP Systems/ Pop-Optix Labs, 241 Crescent St., Waltham, Massachusetts 02154.

Longhurst, R. S. (1973). *Geometrical and Physical Optics*. New York: Longman.

Robinett, W. and J. P. Rolland (1992). *A Computational Model for the Stereoscopic Optics of a Head Mounted Display*, Presence, Vol 1.

Vision Research. Flight Helmet / 3193 Belick Street #2, Santa Clara, CA 95054.

Panasonic Technical Support. Matsushita Electric Corp. of America, One Panasonic Way, Secaucus, NJ.

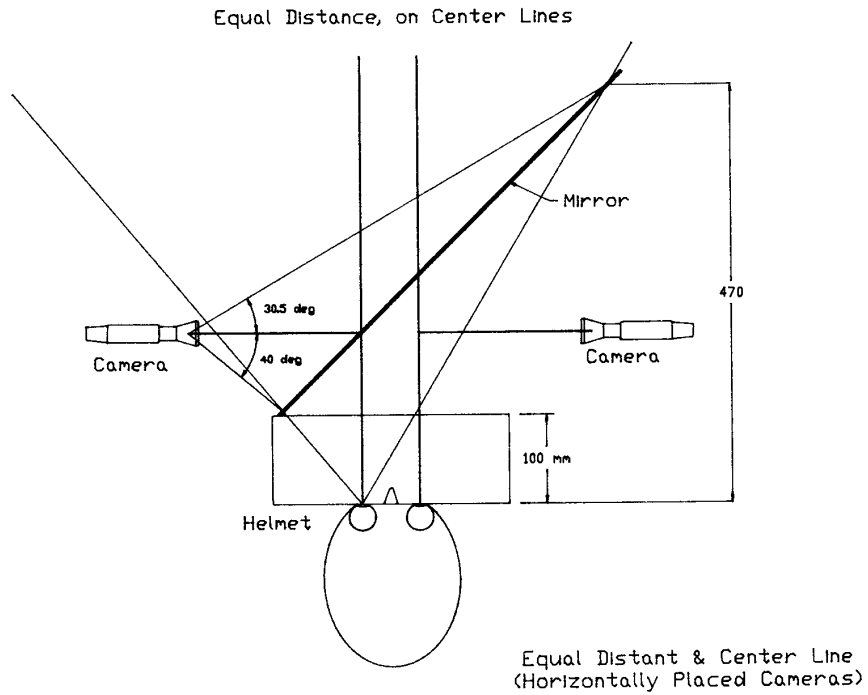


Figure 1: HMD with the miniature cameras mounted such that they simulate the eyes of the observer. This configuration is not feasible since the mirrors must be very large.

Figure 2a: HMD with cameras mounted in feasible locations. This causes a mismatch between the senses of vision and proprioception.

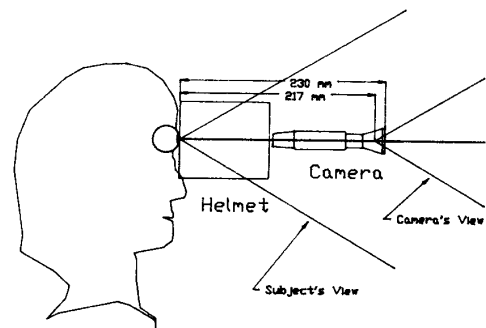
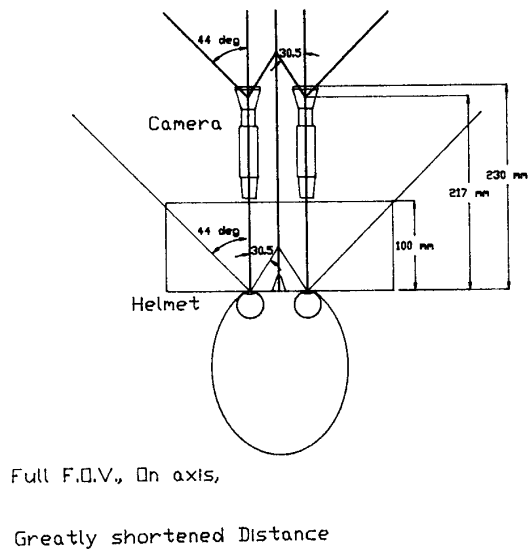


Figure 2b: Same as 2a but side view

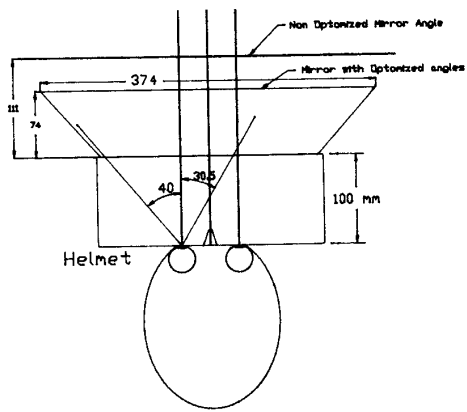


Figure 3a: Final folded mirror arrangement with matched IPD, on-axisness, and a null camera nodal points to eye nodal points distance

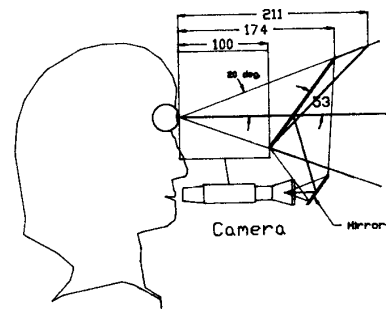


Figure 3b: Same as 3a but side view showing optimized angle of folds.

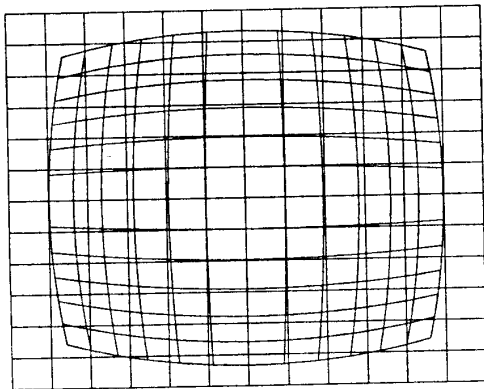


Figure 4a: Distortion of the video lens within the paraxial FOV.

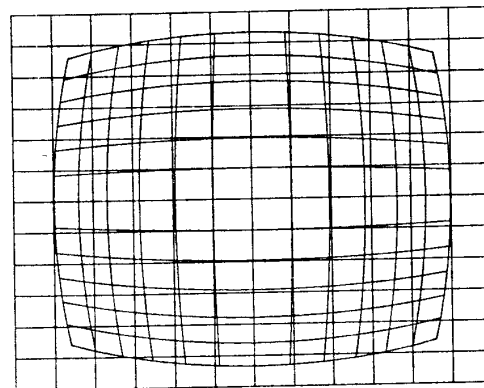


Figure 4b: Distortion of the LEEP optics within the paraxial FOV.

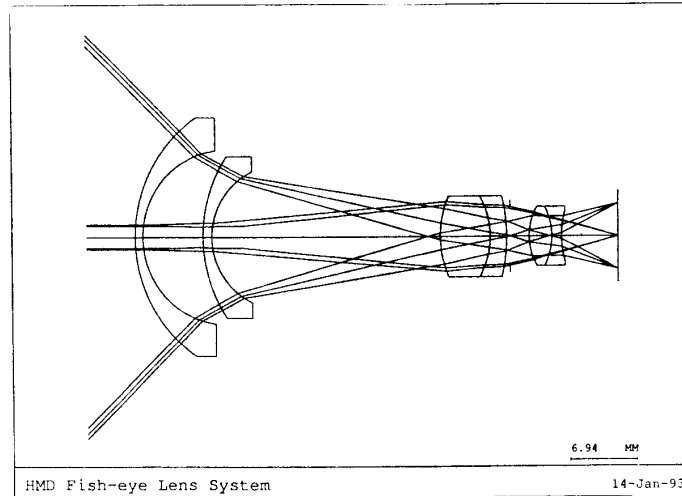


Figure 5: Layout of the video lens.

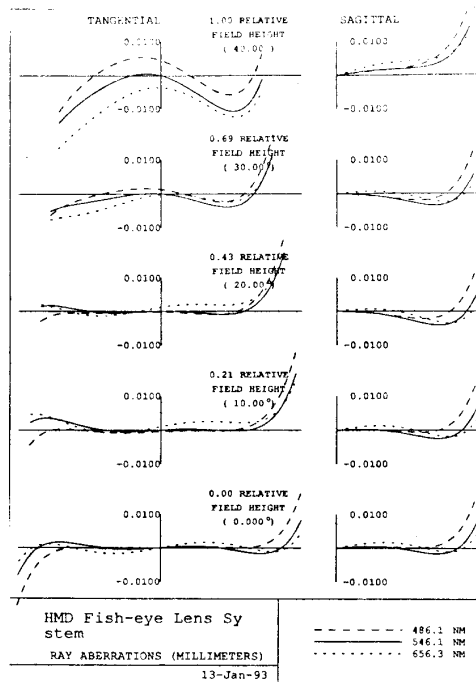


Figure 6: Ray fan plots for the following field angles: 10°, 20°, 30°, and 40°. The ray fan plot is evaluated at all three wavelengths: 486.1nm, 546.1nm, and 656.3nm.

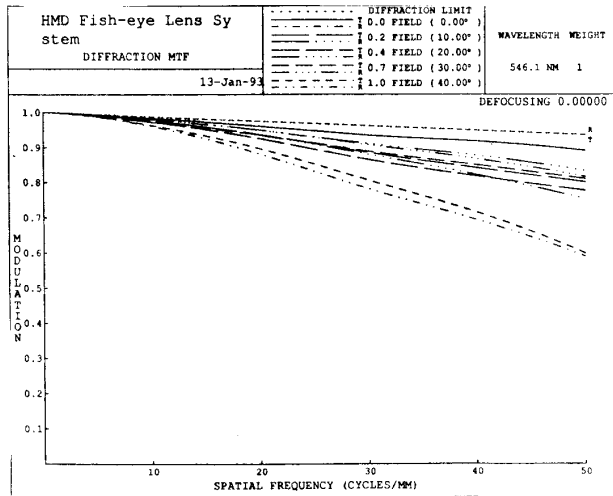


Figure 7: Geometrical MTF plot for the following field angles: 10°, 20°, 30°, and 40°. The MTF is evaluated at the median wavelength of 546.1nm.

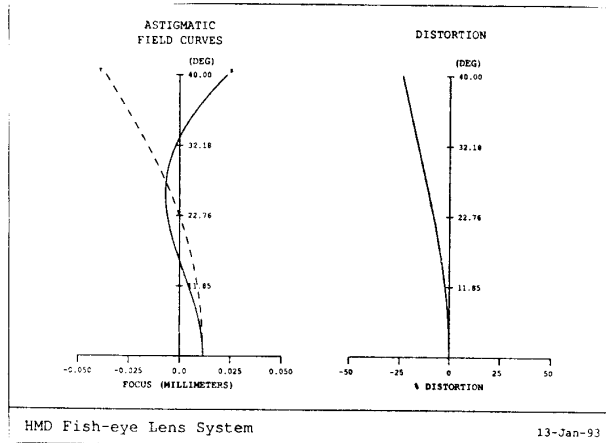


Figure 8: Field plot of video lens for a 40° field angle.