

High-resolution inset head-mounted display

Jannick P. Rolland, Akitoshi Yoshida, Larry D. Davis, and John H. Reif

A novel approach to inset superimposition in a high-resolution head-mounted display (HMD) is presented. The approach is innovative in its use of optoelectronic, nonmechanical devices in place of scanning mechanical devices commonly adopted previously. A paraxial layout of the overall HMD system is presented, and the benefit of employing hybrid refractive-diffractive optics for the optical component that generates the inset is discussed. A potential overall HMD design is finally presented to show the integrated system. The practical limitations of the designed system are discussed and an alternative approach is presented to compare the advantages and the limitations of these systems.

© 1998 Optical Society of America

OCIS codes: 090.2820, 220.0220, 050.1380, 230.0250.

1. Introduction

The field of virtual environments (VE's) and the use of head-mounted displays (HMD's) as three-dimensional (3D) visualization devices have recently received considerable attention due to the potential to create unique capabilities for advanced human-computer interaction.¹⁻³ Such advanced interaction can include interactive control and diagnostic, educational and training, teleoperation, and entertainment systems.⁴⁻⁹

Conventional HMD's typically do not utilize the full potential of VE technology. In particular, they are often limited in field of view and resolution, which is critical to the development of various VE applications. Using a small high-resolution inset and a large low-resolution background image can minimize the trade-off of field of view against an effective resolution.¹⁰ The gaze point of the user, which is determined from eye tracking, controls the position of the inset dynamically and thus high effective resolution over a large field is provided. The main advantages of such a dual-display approach are the

relatively low-cost displays as a substitution for expensive high-resolution displays, and the reduction in computation and bandwidth needed to update the scenes. Although the user can observe dynamic scenery over a large field at apparent high resolution, the image updates do not have to occur simultaneously at high resolution: the portion of the image near the gaze point may be updated quickly at high resolution; other portions of the image may be updated less frequently or at lower resolution to reduce both the computational load and the transmission bandwidth.

To match human visual acuity within the inset, the inset pixel size must subtend approximately 1 arc min at the eye.¹¹ From the knowledge that the human retina does not provide uniform visual acuity across its field of view, the resolution of the HMD needs to reach 1 arc min only inside the inset and can be lower over the entire background.^{12,13} De facto, human visual acuity degrades significantly as the distance from the fovea increases; at an angular distance of 5° from the center of the fovea, it is approximately a quarter of the highest acuity, and at an angular distance of 15°, it becomes only one seventh of the highest acuity.¹¹

Another common shortcoming of conventional HMD's is their lack of integrated effective interaction capabilities combining head and eye tracking. In fact, the interaction capability is ordinarily limited to the use of head tracking to measure the position and the orientation of the user's head and to generate scenery from the user's perspective.¹⁴ This capability allows the user to navigate through the virtual world and interact with the virtual objects with essentially 3D manual input devices. However, for sit-

J. P. Rolland and L. D. Davis are with the Center for Research and Education in Optics and Lasers, Department of Electrical and Computer Engineering, University of Central Florida, 4000 Central Florida Boulevard, Orlando, Florida 32816-2700. A. Yoshida is with the Computing Center at the University of Mannheim, D-68131 Mannheim, Germany. J. H. Reif is with the Department of Computer Science, Duke University, Durham, North Carolina 27708.

Received 13 January 1998; revised manuscript received 13 January 1998.

0003-6935/98/194183-11\$15.00/0

© 1998 Optical Society of America

uations that require response times less than 300 ms or difficult coordination skills, interaction capability supported by such manual input devices becomes inadequate. For those cases, eye movement could be used in conjunction with manual input devices to provide effective interaction methods. Various interaction methods can thus be realized through the use of hand, body, and eye movements.¹⁵⁻¹⁷ Since the eyes respond to stimulus 144 ms faster than the hand,¹⁸⁻²⁰ they can be used for fast and effective input, selection, and control methods.

An additional advantage of using eye tracking is that image rendering can take advantage of the physiological limitation of the eyes. It has been well known since Raymond Dodge²¹ in the 1900's that when the eyes move, information processing is suppressed. This is known, in the modern literature, as saccadic suppression.^{22,23} Therefore, although the gaze point is in rapid motion, the image update does not have to occur at full resolution and the fine detail of the scene can be rendered when the gaze point is basically fixed. With eye tracking, image rendering can be carried out according to both head and eye movement. Thus the use of eye tracking is not limited to finding the gaze point for positioning the inset, it can also play a fundamental role in providing another unique means of interaction in the VE. Combined with appropriate computer software, a HMD with eye-tracking capability will become an active vision HMD that gives the user the feeling of being immersed in the virtual environment and provides effective gaze-point-oriented interaction methods.

A common approach to producing a high-resolution inset is to employ large high-resolution displays, or light valves, and transport the high-resolution images to the eyes by imaging optics coupled to a bundle of optical fibers.^{24,25} In such an approach, scanning optics that employ, for example, mirrors as the scanning mechanism are used to position the inset at the user's point of gaze. Another more recent approach to a high-resolution inset uses only one light valve and renders the image at the gaze point more accurately than the surrounding image. Although systems employing either approach provide significant improvements over ordinary displays in terms of image quality, they are seldom put to use because they are heavy, expensive, and most important, nonportable.

We propose a radically different approach to the positioning of a high-resolution inset that is illustrated in Fig. 1.^{26,27} The approach allows the positioning of the inset without any mechanically moving parts. Instead, the inset is positioned optically with a lenslet array subsystem referred to as the duplicator and shown in Fig. 1. We predict that the use of fixed optoelectronic components allows the whole system to be fabricated with fewer alignment errors, to be immune to mechanical failure, to be more tolerant to vibrations, and to be portable, which is important. Such an approach leads to an optoelectronic high-resolution inset (OHRI) HMD.

In this paper, the optical insertion and superimpo-

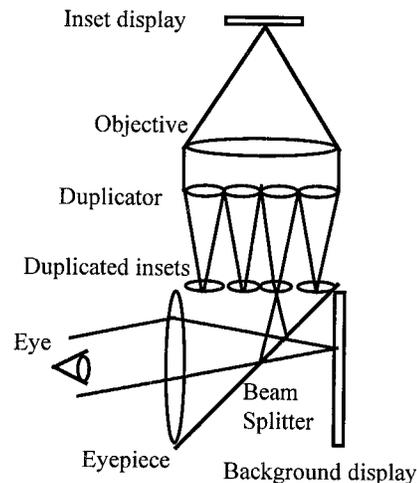


Fig. 1. Schematic diagram of the OHRI HMD.

sition schemes are first presented. The paraxial layout of the OHRI HMD is then given followed by a discussion of the potential use of binary optics for the duplicator. Next, a design configuration for the overall HMD is shown, its performance is demonstrated, and the integration of a low-cost eye-tracker system is presented. Finally, practical limitations of the system are discussed and alternative approaches to the design are presented to access the advantages and the limitations of the proposed system better.

2. Optical Insertion of the High-Resolution Inset

The basic concept of the OHRI HMD, which is illustrated in Figs. 1 and 2(a), is optical duplication of the inset image into a fixed array of nonoverlapping copies and selection of one copy by blocking the others. The selected copy of the inset image is then superimposed optically on the background image. When the original inset image is partitioned and permuted, as explained in Section 3, it can be positioned at arbitrary locations within the background image. The inset image traces the gaze point, thus the foveated part of the image is seen at high resolution.

The gaze point is determined with eye-tracking

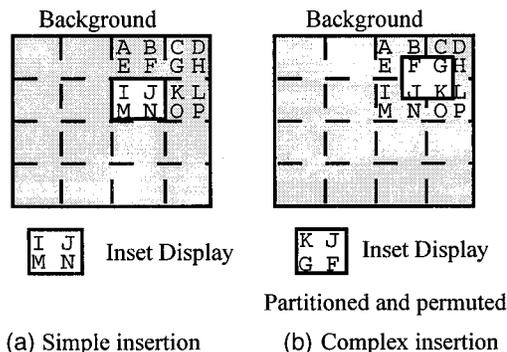


Fig. 2. Superimposition of the inset display: (a) simple insertion and (b) complex insertion with subinset resolution.

technology. Several methods that may differ in accuracy can be used to track eye movements and consequently the gaze points. High-quality head-mounted eye trackers typically produce an accuracy of 1° visual angle.^{28,29} For determining the gaze point for the sole purpose of superimposing the high-resolution inset on the background image, the required accuracy can be somewhat lower. For an inset subtending 10° , an accuracy of a few degrees will be sufficient to keep the fovea within the inset. To implement some of the complex human-computer interaction methods with the gaze point, the required accuracy of the eye tracker should essentially match that of the highest-quality commercial eye trackers. Once the gaze point is determined, the superposition of the high-resolution inset over the low-resolution background is carried out with fixed optical components as the imaging devices and liquid-crystal technology as the selection devices. The eye-tracker system proposed for the OHRI HMD is described in Section 8.

Figure 1 shows that two displays are required: one for the background and the other for the inset. The image of the inset display is duplicated optically to fill the entire background display, and a liquid-crystal device array located close to the duplicated images is used to select one element of the array. Only one copy of the inset display image passes through the liquid-crystal array, and all the other copies are blocked. The images of the inset display and the background display are then combined with a beam splitter. The inset may be simply superimposed on the background without blocking its low-resolution counterpart image behind the inset. In this case, the inset portion of the image becomes brighter than the peripheral image. The background image behind the inset may be dimmed electronically to control the effective brightness of the inset. To minimize the boundary effect at the periphery of the inset, one can soften the boundary of the inset either electronically at the inset display or optically with an optical filter inserted in the inset path. An alternative is to use another liquid-crystal device in front of the background display to block the low-resolution counterpart of the inset location. In this case, boundary effects can then be handled in a similar way.

The imaging sequence of the inset path can be described more precisely as follows: emitted light from the inset display is collimated by an objective lens. A first lenslet array then divides and focuses the collimated light to a set of identical images or duplicates of the inset display. This lenslet array constitutes the first layer of the duplicator. The chief rays from the duplicate image points are then set parallel (i.e., telecentric configuration) to the optical axis by another lenslet array that constitutes the second layer of the duplicator. This telecentric arrangement constrains the chief rays in eye space to converge to the location of the entrance pupil of the eye as described in Section 4 and illustrated in Fig. 3. An array of liquid-crystal shutters placed at the du-

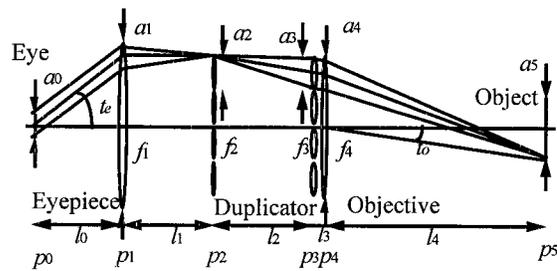


Fig. 3. Paraxial layout of the inset optical path, unfolded with respect to the beam splitter.

plicator passes one of the duplicated images and blocks the other images. Shutters with switching speeds of as much as 250 Hz are technically possible.³⁰ The duplicated images are located with respect to the optics so that they are the optical conjugates of the background display (i.e., they are located symmetrically to the background display with respect to the beam splitter). The eye can then see the combined inset-background image through the eyepiece. Although Fig. 1 is a schematic of the objective with a single lens, the duplicator with two lenslet arrays, and the eyepiece with another single lens, each subsystem (i.e., the objective, the duplicator, and the eyepiece) is implemented with multiple optical elements. A possible design configuration for the overall system is given thereafter.

3. Simple and Complex Insertion Schemes

The superimposition of the inset and the background is depicted in Figs. 2(a) and 2(b) by a simple and a more complex insertion scheme, respectively. In both figures the shaded areas correspond to the background and the bright areas correspond to the inset. The character symbols represent contents of the image, and the dashed lines represent the cell boundary of the duplicated images.

For a simple insertion scheme, the insertion may be made at discrete nonoverlapping cell locations. In this case, the liquid-crystal array, which may be placed anywhere inside the duplicator, blocks all the duplicated images except for one copy. When the inset image of I J M N is desired at a particular cell location, as shown in Figure 2(b), this image can be directly displayed from the inset display. The duplicated images of I J M N fill every cell, and the copy of this image at that cell location exits the duplicator to be superimposed on the background.

Interestingly, the system is not limited to discrete nonoverlapping inset locations. A complex insertion scheme can be used instead so that the insertion may be made at continuous locations, up to the pixel resolution of the liquid-crystal array. In this case, the size of the inset must be smaller or equal to the size of a single duplicated image. The liquid-crystal array, which must be now placed near the duplicated image plane, blocks all the duplicated images except for some portions of as many as four copies. When the inset image of F G J K is desired at a particular

location, as shown in Fig. 2(b), this image may be partitioned and permuted to K J G F, and the transformed image can be displayed from the inset display. The duplicated images of K J G F fill every cell, and the portions of the four adjacent copies at that location, which form the image of F G J K, exit the duplicator to be superimposed on the background.

Depending on the required resolution of the inset imposed by the VE application, either of the two schemes described here can be adopted. In the case of the complex insertion, the required permutation can be implemented in hardware. A special purpose chip would be required, and the speed of superimposition update could be optimized. Such a chip can be designed by Evans and Sutherland.³¹

4. Paraxial Layout of the Overall Optical System

The overall optical system is folded into two optical paths: one path is dedicated to the generation of the inset and the other allows the generation of the low-resolution background. The two optical paths are combined through a beam splitter before reaching the eyes of the user. Figure 3 shows the basic configuration of the optoelectronic system employed for duplicating the inset image, unfolded with respect to the beam splitter. In this figure the rightmost element is the inset display and the leftmost element is the eye. A thin-lens paraxial model of this system is first derived. In this paraxial model, p_i , with $i = 1 \dots 5$, represents planes along the optical system (Fig. 3). More specifically, the eye pupil resides at p_0 and the display object resides at p_5 . Going from right to left on the figure, the first component is the objective lens, located at p_4 , which collimates the light from the display. The lens of the objective has a focal length f_4 .

The second component is a two-layer array of telecentric lenses located at p_3 and p_2 . We call this double-layer array of lenses the duplicator. The two lenslet arrays have focal lengths f_3 and f_2 . The first layer at p_3 duplicates the display image from the collimated light and forms an intermediary duplicated image of the object at p_2 , as also shown in Fig. 4(a). The second lenslet array, which aligns the principal rays parallel to the optical axis, as shown in both Figs. 3 and 4(a), is a catalyst for the chief rays to all pass through the center of the eye pupil. This array, however, introduces not only an additional alignment challenge but also a possible visual quality problem because it must be placed at or very close to the intermediate image plane and any scratches or dust on its surfaces significantly degrade the image quality. Furthermore, the lens boundaries may introduce an annoying blocking effect. An alternative to positioning the lens at the location of the intermediary image plane (i.e., plane p_2) is to defocus that lens slightly. Another alternative is to remove the lens, as shown in Figure 4(b). In this case, however, the bundle of rays suffers 50% vignetting at the edges.

The third component, which is located at p_1 , is the eyepiece. In the configuration shown in Fig. 3, the

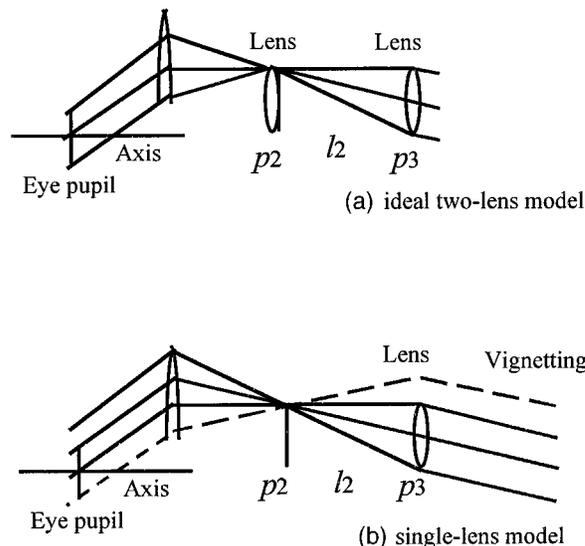


Fig. 4. (a) Ideal two-lens duplicator and (b) single-lens duplicator with vignetting.

intermediary images are located in the focal plane of the eyepiece lens, therefore, they are collimated in eye space. Collimation is not required, as long as the final images are within the range of accommodation of the user, which is typically from 250 mm to infinity. The eyepiece uses a lens with focal length f_1 , and the eye is placed at p_0 .

Let us denote the distance between planes p_i and p_{i+1} and the diameter of the aperture at p_i , l_i and a_i , respectively; the number of duplicated images along the vertical or the horizontal axes, k ; the largest chief-ray angle at the eye pupil, t_e ; and the largest angle of the inset object subtended at the apex of the objective lens, t_o . These parameters play essential roles in the paraxial model layout. The basic imaging condition just described and modeled in Fig. 3 yields the following paraxial relationships among the design parameters:

$$f_1 = l_0 = l_1, \quad (1)$$

$$l_0 \tan t_e = ka_2/2, \quad (2)$$

$$f_2 = f_3 = l_2, \quad (3)$$

$$l_2 \tan t_o = a_2/2, \quad (4)$$

$$f_4 = l_4, \quad (5)$$

$$l_4 \tan t_o = a_5/2, \quad (6)$$

$$a_0/a_3 = l_1/l_2, \quad (7)$$

$$a_3 = a_2, \quad (8)$$

An implementation of this paraxial layout is presented.

5. Paraxial Design Implementation

This specific implementation assumes two active-matrix liquid-crystal displays from Kopin³²: a 0.75-

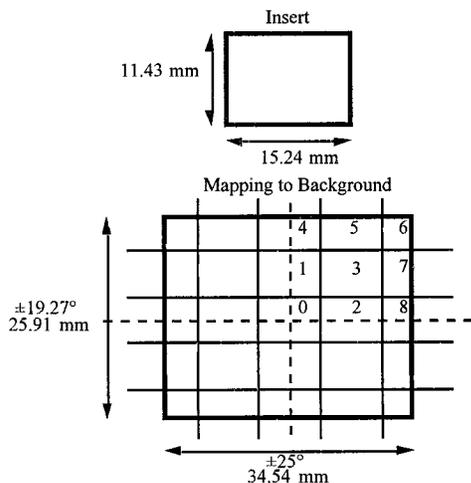


Fig. 5. Mapping of the inset on the background.

in. (1 in. = 2.54 cm) inset display and a 1.7-in. background display. Both displays have 640×480 gray-scale pixels. The contrast ratio of the display is 80:1 and each pixel can be addressed in 49 ns, yielding a frame rate of 72 frames/s. We discuss in Section 8 how this frame rate relates to requirements on eye-tracking technology. The system was optimized for 547 nm (± 5 nm). A color filter is used to produce monochromatic images from the displays. The typical spectral width is 10 nm for such a filter (e.g., standard bandpass filters from Coherent Ealing). The inset image is magnified (magnification is naturally < 1) into one sixteenth of the background image, as shown in Fig. 5. The size of the inset display is 15.24×11.43 mm², and the size of the background display is 34.54×25.91 mm². The area of the background display is partitioned into 25 nonoverlapping cells. The inset display is mapped optically into one of these cells for superimposition. There are 3×3 full-inset cells, 12 half-inset cells at both the horizontal and the vertical edges, and 4 quarter-inset cells at the corners. The cells at the upper right quadrant are numbered from 0 to 8. The other quadrants are symmetrical to this quadrant, and the performance at one quadrant is fully representative of the entire system. We investigated the upper right quadrant.

The following parameter values were selected according to available display devices and design trade-offs, and they are summarized in Table 1. We selected a_0 , the diameter of the eye pupil, to be 10 mm; a_5 , the diameter of the high-resolution display, to be 15.24 mm; ka_2 , the diameter of the array of duplicated images, to be 34.54 mm; t_e , the maximum

Table 2. Design Value Parameters (mm)

i	a_i	f_i	l_i
0	10.000	—	37.036
1	—	37.036	37.036
2	8.635	31.980	31.980
3	8.635	31.980	—
4	—	56.442	56.442
5	15.240	—	—

^aGiven in millimeters.

chief ray angle at the eyepoint, to be 25° ; and k , the number of duplicated images along the duplicated array diameter, to be 4. From these values and Eqs. 1–8, the maximum half-field angle of the objective lens, t_o , is 7.69° . All parameter values a_i , f_i , and l_i are determined and are listed in Table 2. These parameters are used to design each component of the high-resolution inset system (i.e., the objective lens, the duplicator, and the eyepiece). As smaller displays become available on the market, smaller and more compact systems may be built. In fact, Kopin has recently manufactured a 0.24-in. display with 320×240 pixels that may be used with their 0.75-in. display to build such a compact system.

6. Design of the Duplicator with Binary Optics

The duplicator arrays can employ either conventional or binary optics technology. The disadvantage of using conventional optics for the duplicator lenslet arrays is the high sensitivity to misalignment. An alternative is to use binary optics where the lenslet array is fabricated on one substrate. This approach solves the problem of aligning the individual lenses in the arrays. A fill factor of 100% is assumed for the binary optics array because spacing of less than 10 μm can be achieved in binary optics arrays fabrication. The minimum feature size determined by the fabrication facility limits the smallest f -number of the lens to be fabricated. For analytic binarization, the minimum feature size occurs at the edges of the lens.³³ Starting with the grating equation, the definition of the f -number, and the assumption that the feature size is small compared with the lens diameter and the focal length, gives as the minimum feature size of p

$$p = \frac{\lambda}{m} \sqrt{1 + 4(f_\#)^2}, \quad (9)$$

where λ is the wavelength, m the number of phase levels, and $f_\#$ the f -number of the lens.

Table 1. Basic Design Parameters

Parameters	Background	Inset
Display size	34.54 mm \times 25.91 mm	15.24 mm \times 11.43 mm
Pixels resolution	640 \times 480	640 \times 480
Field of view	$50^\circ \times 38.55^\circ$	$13.30^\circ \times 10.05^\circ$ (at center)
Angular resolution	4.69 arc min \sim 12.5 pixels/deg	1.25 arc min \sim 48 pixels/deg

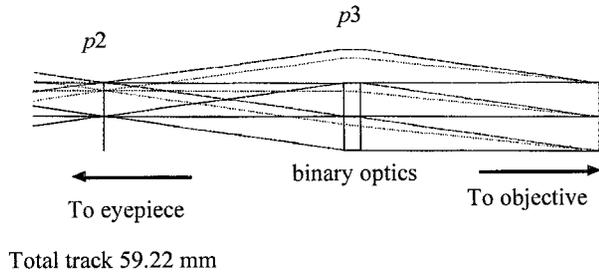
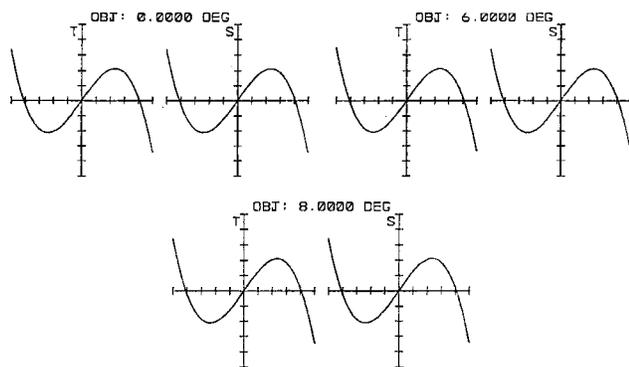


Fig. 6. Duplicator layout with binary optics.

The number of phase levels must be sufficiently large for the fabricated lens to yield high efficiency. The f -number of the duplicator is calculated to be 3.6. When four levels ($m = 4$) are assumed, the minimum feature size p is $1 \mu\text{m}$ for a wavelength of 550 nm . An element with this feature size can be easily fabricated. In this case, an element with one binary optics surface can be designed. However, the theoretical efficiency of the lens at its edges is limited to 81%.³² Note that potential edge effects for either binary optical arrays or elements are automatically accounted for in the computation of the modulation transfer function (MTF) by the optical design software, given that apertures are set correctly.

To have higher efficiency at the expense of an increase in the cost of fabrication, one can adopt more phase levels. For eight levels, p becomes half a micrometer. The efficiency then increases to 95%. Note that a feature size of half a micrometer may be difficult to achieve in some facility, but it is commonly achieved with state-of-the-art technology. If it is the case, however, the power of this element can be distributed in more than one surface. When two surfaces with 95% efficiency are used, the efficiency of the element is kept at 90%. The design given in Fig. 6 shows two silica binary optic surfaces located in the plane p_3 , and the minimum feature size is set to $1 \mu\text{m}$ in this case. Figures 7 and 8 show the rayfan and the MTF plots of the binary optics duplicator, respectively. The required spatial frequency at the duplicator is 37 line pairs/mm. ($320/8.635 = 37.1$ line pairs/mm, where 320 is the largest number of



Maximum scale: $20 \mu\text{m}$

Fig. 7. Duplicator rayfan.

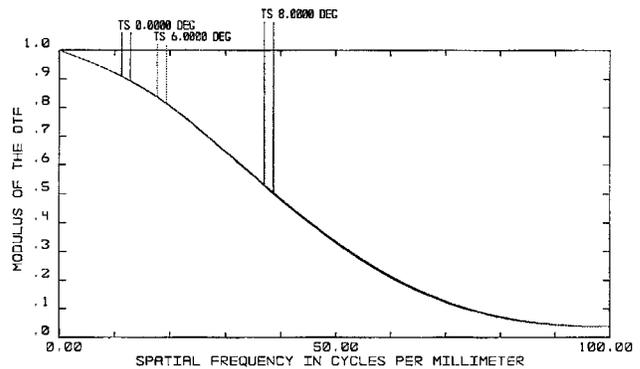


Fig. 8. Duplicator MTF.

line pairs in the display (640×380), and a_2 equals 8.635 mm as shown in Table 2.)

An efficiency of less than 100% results in scattered light in the optical system that causes reduction in the MTF from its designed nominal value. Frequently, a 90% efficiency yields 10% of scattered light. The effect on the MTF is an overall decrease of the MTF curve by a constant amount. In our example, the value of the decrease in MTF can be estimated when the value of the MTF at a low frequency of 1 or 2 cycles/mm is multiplied by 10%. Let us consider the MTF at position 0 on the axis shown in Fig. 11. This is the lowest plain curve for position zero. At 2 cycles/mm, the MTF value is 0.975. A 10% decrease corresponds to a reduction of approximately 0.1 in the MTF that must be subtracted from the values of the MTF across the spectrum. A plot of the estimated effective value of the MTF to account for 10% scatter is shown in Fig. 11 as the lowest dash curve. The effect of scatter for a HMD is minimal, and our experience with designing and building such systems has revealed that it is negligible. This would not be the case of night-vision HMD's in which light is scarce and scatter can significantly degrade the performance of such systems.

When considering Fig. 6, note that the duplicator solution presented in Fig. 4(b) was selected over that presented in Fig. 4(a). The vignetting occurring at the duplicator was discussed in Section 4. We adopted this configuration, as it is the simplest configuration of the two. Moreover, it contributes to minimizing potential alignment problems that might occur when a lenslet array in the plane p_2 and a binary lenslet array (or a pair of them as discussed above) in the plane p_3 are simultaneously used. In the implementation presented here, a single binary optics element was adopted for the duplicator to demonstrate the entire design in its simplest form at the expense of vignetting. With increase in cost, an additional lenslet array may be added in the plane p_2 to reduce vignetting if maximizing light output is an important design criterion.

One of the potential downfalls of using binary optics in any optical system is the degradation of image quality, which is a result of incorrectly handling substantial chromatic aberrations. Because binary optics has the opposite chromatic aberrations as

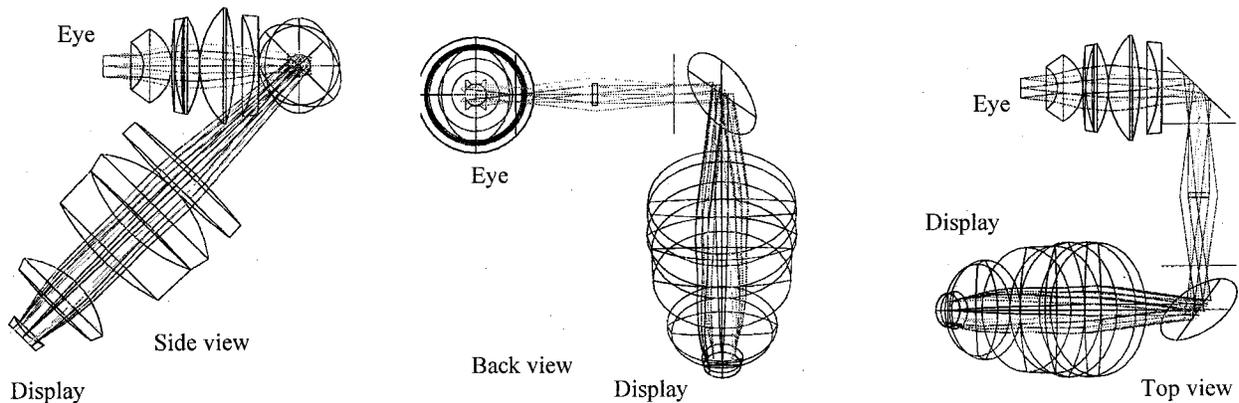


Fig. 9. Possible design and geometric configuration of the OHRI HMD.

refractive optics, it is most commonly used in combination with conventional optics for imaging systems. Moreover, imaging power may be distributed among diffractive and refractive elements to balance aberrations; although this is possible, it is not usually done. In a first effort to demonstrate the conceptual design of this new approach to inset displays, we assume monochromatic light and employ only diffractive elements at the duplicator for simplicity. For multispectrum light, some refractive elements might be required in the duplicator design, and variations in diffraction efficiencies over the spectrum will need to be accounted for in the design and the performance assessment. A complete assessment of the use of diffractive elements for a color display is beyond the scope of this paper.

7. Design and Overall Performance of the Optoelectronic High-Resolution Inset Head-Mounted Display

A design configuration of the entire system is shown in Fig. 9. The system is folded at several places to keep the center of gravity low and close to the head. Figures 10 and 11 show the performance of the entire system at several duplicator positions, indexed from 0 to 8 according to Fig. 5. These figures show that the system is capable of resolving spatial frequencies above the required level of 21 line pair/mm ($320/15.240 = 21.0$ line pair/mm, where 320 is the largest number of line pairs in the display (640×380), and a_5 equals 15.240 mm as shown in Table 2) ($MTF > 20\%$) except at the largest field angle ($MTF > 10\%$). Although the system demonstrated here was obtained after optimization of each optical component (i.e., objective, duplicator, and eyepiece) independently, the performance of the system may be improved by optimizing all components simultaneously.

8. Proposed Eye-Tracker System for the Optoelectronic High-Resolution Inset Head-Mounted Display

The OHRI HMD system requires that the relative orientation of the eyes be known. Therefore an understanding of eye movement might appear critical.

Although the primary eye movements are saccadic, smooth-pursuit, and vergence and the fastest motions are that of saccades (initiated approximately 200 ms after the target object leaves the fovea³⁴ at speeds of as much as 900 deg/s), the parameter of interest for an inset HMD is the duration of a fixation for which the inset will be positioned. The speed of smooth-pursuit movements is typically slower, reaching 100 deg/s.³⁴ Other types of eye movements, including the vestibular-ocular reflex, optokinetic response, and nystagmus combine aspects of the saccadic and the smooth-pursuit movements.

It is widely accepted in the vision literature that it takes typically 100 ms to process new visual information (numbers from 80 to 150 ms are argued among visual scientists). As a result, a fixation is typically defined as a 100-ms pause in eye movement.³⁵ It is known that microsaccades of average amplitude 30 arc min do occur during fixation, but unless the deviation is greater than 2.5° , we shall consider that it is one single fixation for our application. For high-resolution inset HMD's, the time required for a fixation to occur (i.e., 100 ms) or, more conservatively, the time required for the average number of fixations per second (i.e., 4 fixations/s) determines the required bandwidth of the eye tracker. Therefore a 10-Hz bandwidth is required based on the time of a fixation, or a 4-Hz bandwidth is required based on the average number of fixations per second. With a display device updating information at 72 frames/s, we shall be limited in most tracking schemes by the tracking device and not by the display frame rate. Note, however, that current frame rates of tracking devices are largely sufficient to monitor changes in fixations. Furthermore, the range of saccadic and smooth-pursuit eye movements without head motion is $1\text{--}30^\circ$ relative to the current eye position,²⁹ requiring our eye tracker to measure as high as 30° displacements accurately.

As a result of these observations, an eye tracker based on the principle of electro-oculography (EOG) satisfies our requirements. EOG is a method of recording voltage changes due to eye rotation.¹⁹ In the eye, a potential difference of as much as 1.0 mV exists

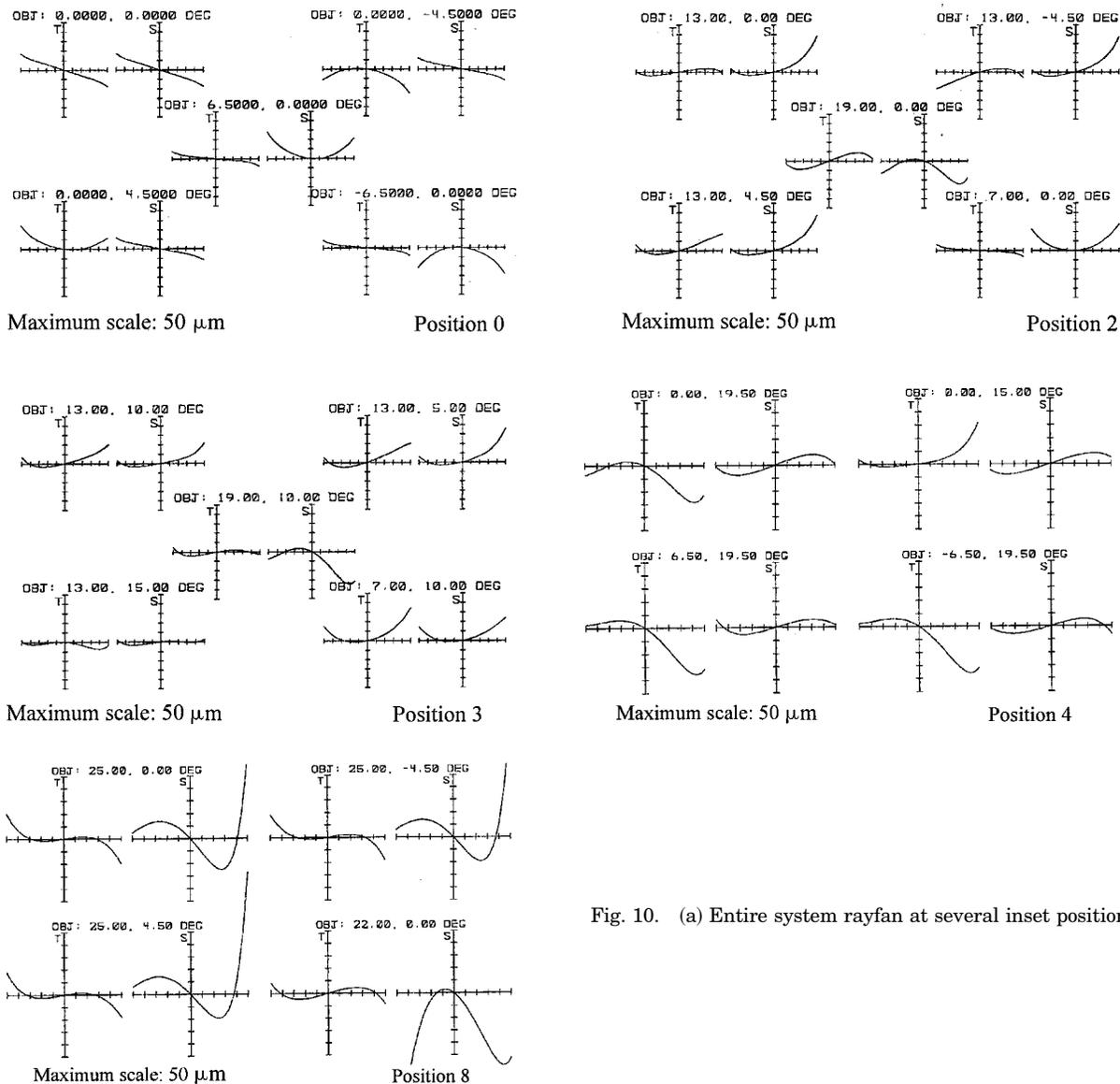


Fig. 10. (a) Entire system rayfan at several inset positions.

between the cornea and the retina, with the cornea being positive with respect to the retina.²⁹ When the eye moves, this corneo-retinal potential (CRP) causes a change in potential in the area immediately surrounding the eye. By measuring this change in potential at certain locations around the eye, one can determine eye position indirectly. With proper electrode placement, which is detailed hereafter, a tracking accuracy of ± 1 degrees is achievable,¹⁹ with a range of $\pm 70^\circ$.²⁹ In addition, a frequency response of as much as 15 Hz is achievable with EOG. (Note that infrared eye trackers, with rates from 60 to 240 Hz, can also be implemented.³⁵) Finally, the recording electronics used for EOG can be miniaturized to decrease overall weight and system size.

The design of the eye tracker involves minimizing electrical interference while providing adequate output response. Electromuscular signals, galvanic skin response (GSR), and electroencephalography (EEG) signals can cause interference when measur-

ing the CRP.³⁶ To lessen the effects of electromuscular signals (caused by facial movements), Ag-AgCl electrodes are placed as close to the orbital bone boundary as possible. Placing the electrodes close to the eye also increases the amplitude of the CRP signal. However, the CRP is still small and amplification of the signal is necessary. Additionally, the presence of EEG signals at the electrodes prevents linear amplification. To minimize EEG interference, a differential amplifier configuration with a high common-mode rejection ratio is used.³⁶ Removing oils from the perspective electrode location can minimize GSR. Still, measuring the CRP from the same locations each time better reduces the effects of GSR by providing a repeatable resistance. Because the CRP may change as a factor of lighting and metabolic rate, a quick calibration must be performed for each use of the system.

EOG eye-tracking technology is attractive for the OHRI HMD because it is positioned around the eyes

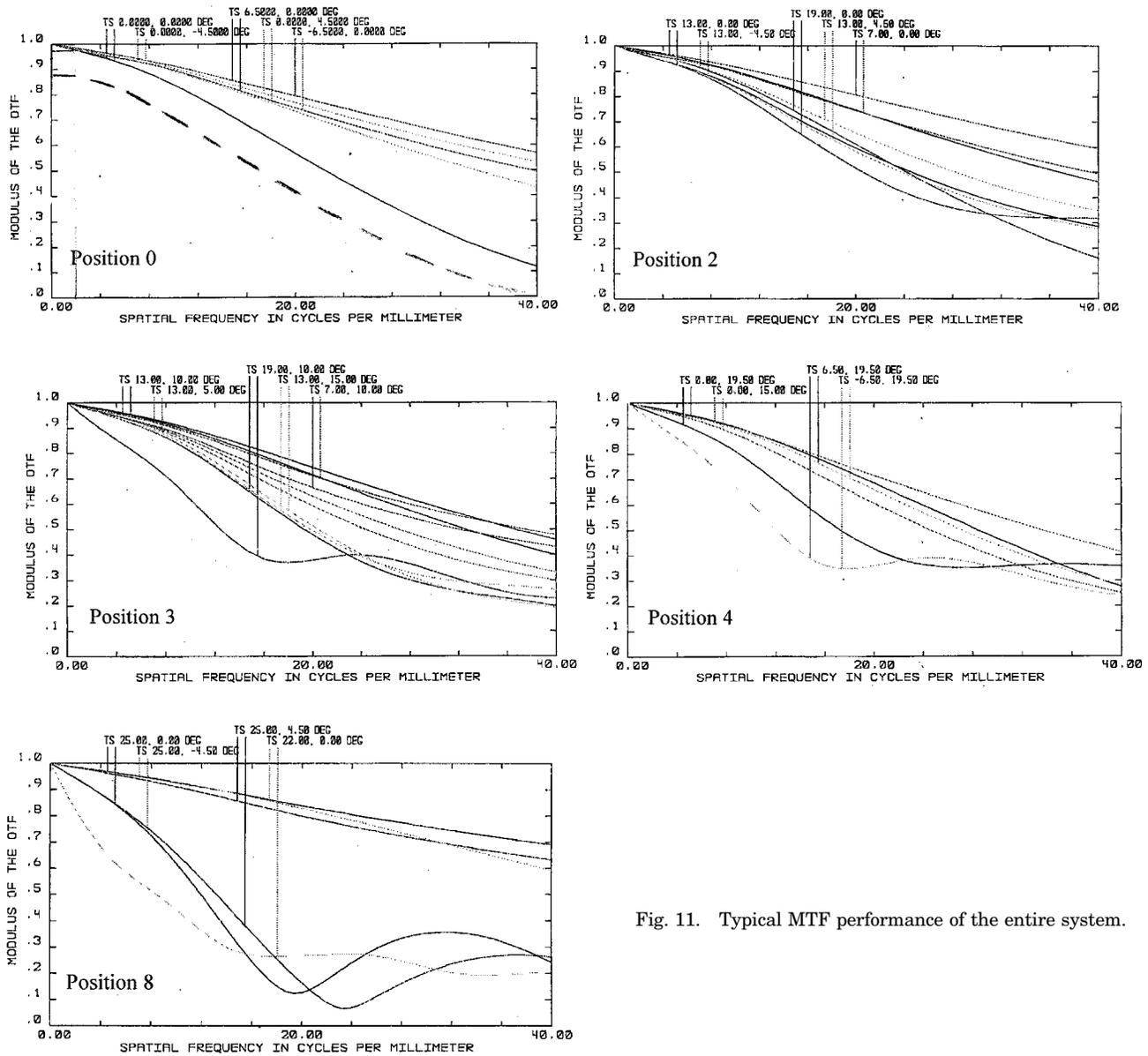


Fig. 11. Typical MTF performance of the entire system.

and therefore is nonobtrusive. In addition, it requires no exhaustive user preparation and can be integrated easily with our prototype. Finally, the wide availability of the electronics required, the capability of miniaturization, and the low cost of such devices suggest that the use of EOG in HMD's will become commonplace.

9. Practical Limitations

Although the proposed approach may achieve our goal as a prototype, it has several practical limitations. These issues are first discussed and alternative approaches as well as their respective advantages and limitations are described.

One of the difficulties is to keep the duplicated images free of both distortion and field curvature. When there is distortion, the gaps between the duplicated images become highly visible and it becomes impossi-

ble to form an inset image with four neighboring duplicate images as described earlier. The visual quality is unevenly degraded by field curvature as well. Furthermore, the alignment issue also becomes critical for the objective and the duplicator systems that must generate duplicated images from a single inset image. Finally, once constructed, the entire system must be packaged in a light and rigid helmet system to guarantee the portability and the durability of the system. Diffraction efficiency at the duplicator, as discussed in Section 6, and reflection at multiple surfaces decrease the MTF ratios. We estimate that decreases of less than 15% overall can be achieved with current technologies. Antireflection coatings can be effectively used to reduce reflection at lens surfaces, for example, and more levels in diffractive optical elements can be utilized to yield higher efficiencies at the expense of increase in fabrication cost.

Recent advances in producing miniature active matrix displays and their prospected wide use in various fields seems to indicate that using multiple displays, instead of optically duplicated ones, may be advantageous. This approach eliminates the alignment problem and decreases the volume and the weight of the system. This approach differs from other multiple displays system because the entire view is rendered by a single background display and only the inset is rendered by one of the inset displays at a time. The benefit of separating two sources, namely, the background and the inset, is maintained. Furthermore, the inset size can be enlarged by software within the bandwidth limit. However, the problem of completely eliminating the gaps between the inset displays remains a challenge. For the miniature displays in consideration, the gaps are relatively large (stretching approximately 20% of the displayable area in each direction), leading to a non-viable or at least a highly suboptimal approach. However, since the eye position and the visual angle are limited, if a display with a very small peripheral area was made available, one can minimize these peripheral gaps by using an array of lenses with each lens directly attached to each display.

Another problem with the proposed design is the merging of two sources. The current eyepiece that must allocate a large space between the eyepiece lens and its focal plane introduces a scalability problem for field of view. This space was necessary for merging the two optical paths. This construction is not necessary if a transparent display is used for the background as well as the insets. In this case, both the background and the multiple inset displays can be illuminated from behind. However, the problem of positioning two surfaces, the background and the insets, on a single plane is not possible and a lack in exact superimposition yields parallax errors with various eye locations behind the HMD. Given the physical dimensions of such displays, an analysis of the magnitude of such errors needs to be conducted if such a scheme is to be considered.

10. Conclusion

The approach to the design of a novel high-resolution inset HMD referred to as the OHRI HMD was presented. The principle of insertion for the high-resolution inset, which involved no mechanical components but rather solely optoelectronic devices, was shown. The apparent benefit of the OHRI HMD is its potential for providing a relatively large field of view (i.e., 630) as well as a high-resolution image (i.e., 1.25 arc min for inset and 4.7 arc min for background) in addition to being portable. Moreover, such a system also supports various gaze-point-oriented interaction methods unlike systems with no eye-tracking capability.

The authors thank Reinhardt Männer and Steffen Noehte at the University of Mannheim for their interesting discussions and equipment support, Hudson Welch at Digital Optics Corporation for his

assistance with binary optics specification and fabrication, Applied Science Laboratories for their assistance with eye-tracker requirements, and Robert Kennedy from RSK Assessments, Inc. for pointing us to the literature in saccadic suppressions and the literature on electro-oculography. This research was supported in part by the Office of Naval Research, grant N000149710654 and the National Library of Medicine under grant 1-R29-LM06322-01A1.

References

1. S. S. Fisher, "Virtual interface environments," in *The Art of Human-Computer Interface Design*, B. Laurel ed. (Addison-Wesley, Menlo Park, Calif., 1990), 423-438.
2. D. Foley, "Interfaces for Advanced Computing," *Sci. Am.* **257**(4), 126-135 (1987).
3. J. C. Chung, M. R. Harris, F. P. Brooks, H. Fuchs, M. T. Kelley, J. Hughes, M. Ouh-Young, C. Cheung, R. L. Holloway, and M. Pique, "Exploring virtual worlds with head-mounted displays," in *Three-Dimensional Visualization and Display Technologies*, S. S. Fisher and W. E. Robbins, eds., Proc. SPIE **1083**, 42-52 (1989).
4. G. Burdea and P. Coiffet, *Virtual Reality Technology* (Wiley, New York, 1994).
5. R. E. Cole, C. Ikehara, and J. O. Merritt, "A low cost helmet-mounted camera/display system for field testing teleoperator tasks," in *Stereoscopic Displays and Applications III*, J. O. Merritt and S. S. Fisher, eds., Proc. SPIE **1669**, 228-235 (1992).
6. S. S. Fisher, M. McGreevy, J. Humphries, and W. Robinett, "Virtual environment display system," presented at the Association for Computing Machinery Workshop on Interactive 3D Graphics, Chapel Hill, North Carolina, 23-24 Oct. 1986.
7. S. S. Fisher, M. W. McGreevy, J. Humphries, and W. Robinett, "Virtual interface environment for telepresence applications," in *Proceedings of the American Nuclear Society International Topical Meeting on Remote Systems and Robotics in Hostile Environments*, J. D. Berger, ed. (American Nuclear Society, Lagrange Park, Ill., 1987).
8. C. Herot, "Spatial management of data," *ACM Trans. Database Systems* **5**(4), 493-514 (1980).
9. D. Thalmann, "Using virtual reality techniques in the animation process," in *Virtual Reality Systems*, R. A. Earnshaw, M. A. Gigante, and H. Jones, eds. (Academic, Reading, Mass., 1993).
10. E. M. Howlett, "High-resolution inserts in wide-angle head-mounted stereoscopic displays," in *Stereoscopic Displays and Applications III*, J. O. Merritt and S. S. Fisher, eds., Proc. SPIE **1669**, 193-203 (1992).
11. H. Davson, *Physiology of the Eye*, 5th ed. (Pergamon, New York, 1990).
12. R. A. Moses, *Adlers Physiology of the Eye* (Mosby, St. Louis, Mo., 1970).
13. G. Westheimer, "The eye as an optical instrument," in *Handbook of Perception and Human Performance* 1(4) (Wiley-Interscience, New York, 1986), Vol. 1, Chap. 4.
14. F. J. Ferrin, "Survey of helmet tracking technologies," in *Large Screen Projection, Avionic, and Helmet-Mounted Displays*, H. M. Assenheim, R. A. Flasck, T. M. Lippert, J. Bentz, and W. Groves, eds., Proc. SPIE **1456**, 86-94 (1991).
15. R. A. Bolt, "Gaze-orchestrated dynamic windows," *Comput. Graphics*, **15**(3), 109-119 (1981).
16. S. Bryson, "Interaction of objects in a virtual environment: a two-point paradigm," in *Stereoscopic Displays and Application II*, J. O. Merritt and S. S. Fisher, eds., Proc. SPIE **1457**, 180-187 (1991).

17. H. Jacoby and S. R. Ellis, "Using virtual menus in a virtual environment," *Proc. SPIE* **1668**, 39–47 (1992).
18. T. P. Colgate, "Reaction and response time of individuals reacting to auditory, visual, and tactile stimuli," *Research Quarterly* **39**, 783–784 (1968).
19. P. J. Oster and J. A. Stern, "Measurement of Eye Movement," in *Techniques of Psychophysiology*, I. Martin, and P. H. Venables, eds. (Wiley, New York, 1980).
20. H. Girolamo, "Notional helmet concepts: a survey of near-term and future technologies," U.S. Army NATICK Technical Report NATIK/TR-91/017 (U.S. Army, 1991).
21. R. Dodge, "Five types of eye movement in the horizontal meridian plane of the field of regard," *Am. J. Phys.* **8**, 307–329 (1903).
22. F. Volkman, L. A. Riggs, K. D. White, and R. K. Moore, "Contrast sensitivity during saccadic eye movements," *Vis. Res.* **18**, 1193–1199 (1978).
23. F. Volkman "Human visual suppression," *Vis. Res.* **26**, 1401–1416 (1986).
24. R. Burbidge and P. M. Murray, "Hardware improvement to the helmet-mounted projector on the visual display research tool at the Naval Training Systems Center," in *Helmet-Mounted Displays*, J. T. Carollo, ed., *Proc. SPIE* **1116**, 52–60 (1989).
25. M. L. Thomas, W. P. Siegmund, S. E. Antos, and R. M. Robinson, "Fiber optic development for use on the fiber optic helmet-mounted display," in *Helmet-Mounted Displays*, J. T. Carollo, ed., *Proc. SPIE* **1116**, 90–101 (1989).
26. A. Yoshida, J. P. Rolland, and J. H. Reif, "Design and applications of a high-resolution insert head-mounted-display," *Proceedings of the IEEE Virtual Reality Annual International Symposium* (Institute of Electrical and Electronics Engineers, New York, 1995), 84–93.
27. A. Yoshida, J. P. Rolland, and J. H. Reif, "Optical design and analysis of a head-mounted display with a high-resolution insert," in *Novel Optical Systems Design and Optimization*, J. M. Sasian, ed., *Proc. SPIE* **2537**, 71–82 (1995).
28. Applied Science Laboratories, *Eye Tracking Systems Handbook* (Applied Science Laboratories, Waltham, Mass., 1992).
29. L. Young and D. Sheena, "Survey of eye movement recording methods," *Behav. Res. Methods Instrum. Comput.* **7**, 397–429 (1975).
30. G. Sharp, Boulder Nonlinear Systems, Inc., 1898 South Flatiron Court, Boulder, Colo. 80301 (personal communication, 1997).
31. P. K. Doenges, Evans and Sutherland, 600 Komas Drive, P.O. Box 58700, Salt Lake City, Utah 84158 (personal communication, 1997).
32. J. P. Salerno, "Single crystal silicon AMLCDs," presented at the International Workshop on Active Matrix Liquid Crystal Displays, 14th International Display Research Conference, Monterey, Calif., 1–7 September 1994.
33. J. Jahns, "Diffractive optical elements for optical computers," in *Optical Computing Hardware*, J. Jahns and S. H. Lee, ed. (Academic, Boston, Mass., 1994).
34. M. E. Goldberg, H. M. Eggers, and P. Gouras, "The Ocular Motor System," in *Principles of Neural Science*, 3rd ed, E. R. Kandel, J. H. Schwartz, and T. M. Jessell, ed. (Appleton & Lange, Norwalk, Conn., 1991).
35. J. Borah, Applied Science Laboratories, 175 Middlesex Turnpike, Bedford, Mass. 01730-1428 (personal communication, 1997).
36. M. A. Wiedl, "Some Practical Consideration in Eye Movement Recording Methodologies" (Pacific Missile Test Center, Point Mugu, Calif., 1977).