Localization of Virtual Objects in the Near Visual Field

Stephen R. Ellis, NASA Ames Research Center, Moffett Field, California, and **Brian M. Menges**, San Jose State University Foundation, San Jose, California

We examined errors in the localization of nearby virtual objects presented via see-through helmet-mounted displays as a function of viewing conditions and scene content in four experiments using a total of 38 participants. Monocular, biocular, and stereoscopic presentation of the virtual objects, accommodation (required focus), participants' age, and the position of physical surfaces were examined. Nearby physical surfaces were found to introduce localization errors that differ depending on the other experimental factors. These errors apparently arise from the occlusion of the physical background by the optically superimposed virtual objects, but they are modified by participants' accommodative competence and specific viewing conditions. The apparent physical size and transparency of the virtual objects and physical surfaces, respectively, are influenced by their relative position when superimposed. The design implications of the findings are discussed in a concluding section. Head-mounted displays of virtual objects are currently being evaluated as aids for mechanical assembly and equipment maintenance. Other applications include telesurgery, surgical planning, telerobotics, and visualization aids for robotic programming.

INTRODUCTION

Until recently, most studies concerning the design of see-through virtual image displays have considered systems that present virtual images focused only at large distances from the users' eyes (> 2 m). Research on the design of cockpit head-up displays (HUDs) is an example of such work (Weintraub & Ensing, 1992). The recent development of inexpensive head-mounted see-through displays, however, has extended their range of possible application to include many cases in which users manually interact with nearby virtual images. Possible applications include previewing mechanical assembly, surgical planning and training, and visualization of virtual objects produced by computer-aided design systems.

Computer-generated images, which behave like real objects in these applications, are often described as *virtual objects* and raise questions about optical design and information formatting that have not been previously confronted. The studies in this paper examine several such questions concerning some cues to distance that influence operator interaction with objects within arms' reach.

The perceptual cues to space have been classically separated in terms of the presence or absence of movement or binocular information. More recent analyses of depth perception focusing on the behavioral affordances of vision have more usefully reclassified the classical depth cues into three categories: those primarily important with respect to personal space, such as stereopsis (2 m, approximately 1–2 eye heights), those mainly relevant for action space, such as self-motion parallax (3–30 m, approximately 2–20 eye heights), and those especially relevant for vista space, such

Requests for reprints should be sent to Stephen R. Ellis, MS 262-2, Moffett Field, CA 94035-1000. **HUMAN FACTORS**, Vol. 40, No. 3, September 1998, pp. 415–431.

as aerial perspective (> 30 m, approximately > 20 eye heights; Cutting & Vishton, 1995).

This reclassification of sources of information concerning the spatial layout surrounding a viewer is particularly useful because it focuses attention on what vision is supposed to be used for in each of these distinct regions of space. Not surprisingly, observers' sensitivity to detecting changes in depth varies across the different cues, and the different cues have varying importance for different regions (Nagata, 1993). In particular, binocular convergence, accommodation (required focus), and associated reflexes play roles mainly relevant for the personal space associated with coordinated, manipulative activity. In this region, accommodation itself was classically not thought to be a direct or potent influence on perceived depth (Graham, 1965). However, through the accommodation-vergence reflex (Ciuffreda, 1992; Krishnan, Phillips, & Stark, 1973; Semlow & Hung, 1983), accommodation can indirectly but significantly influence perceived depth by causing a binocular-vergence-associated rescaling of perceptual space.

Other results suggest that some observers may be able to use accommodation itself as a depth cue (Fisher & Ciuffreda, 1988). Consequently, investigations of the localization of nearby virtual objects need to consider the role of accommodation and its associated reflexes.

Understanding the interaction of these physiological reflexes on depth perception will have growing importance as head-mounted displays of virtual objects are introduced into the industrial and medical workplace. Recent applications of such displays are designed to present nearby, spatially conformal, computer-generated virtual objects to their users for medical and manufacturing applications (Azuma & Bishop, 1994; Janin, Mizell, & Caudell, 1993; Rolland, Ariely, & Gibson, 1995). This work initially focused attention on precise calibration and alignment of the displays but now needs to be expanded to include the study of perceptual phenomena that might degrade operator performance even in well-calibrated systems.

The following four experiments explore such phenomena by examining participants' ability to adjust the distance of a physical pointer to match that of a nearby virtual object. The object is generated by a high-performance computer graphics system and presented by a head-mounted see-through display. This localization task was selected because it is close to the visual/manual manipulation expected of users of virtual objects and because its precision and accuracy can be easily and reliably measured. Preliminary testing, for example, showed that participants could set our mechanically displaced physical pointer to match the distance of physical targets within several millimeters and that this accuracy corresponded to their ability to match target distances with their fingers. Additionally, the use of our pointer allowed examination of targets just beyond arm's reach.

The initial experiment examines the effects of three different viewing conditions (monocular, biocular, and stereoscopic viewing) on participants' accuracy of placement of the physical pointer under a virtual object. These three conditions represent a range of cost and of image fidelity (i.e., the completeness with which an objects' physical characteristics are presented). The monocular condition presents a virtual image of an object as monocular helmet-mounted sights do and represents a minimal hardware/software rendering cost for such virtual-image displays. However, the monocular virtual images it presents are subject to visual suppression attributable to binocular rivalry.

The biocular condition, which presents two identical virtual images to the participants' left and right eyes, can avoid the rivalry problem. The images it presents are projected from a cyclopean position between the viewer's eyes but are shifted to allow unstressed fusion at a selected distance (see Figure 1). This condition halves the rendering cost of stereo displays but doubles the head-mounted display hardware requirements of monocular displays. It presents a pattern of disparity approximating that of a transparent, flat display surface and represents an intermediate-cost system with potentially more stable image brightness.

The third condition, conventional stereo display with parallel viewing vectors, presents the highest spatial fidelity but doubles both

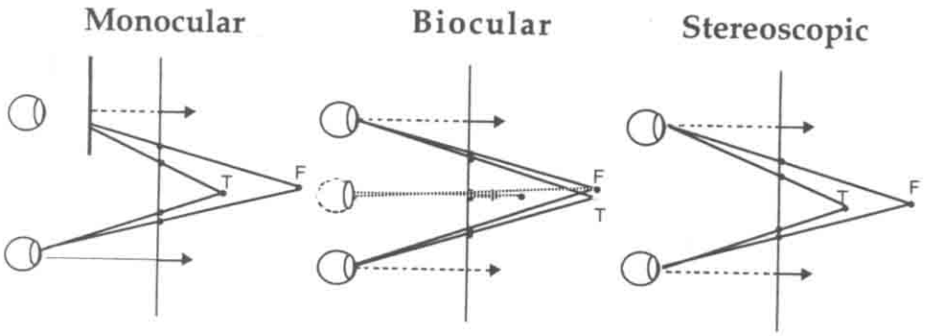

Figure 1. Viewing geometry for each viewing condition for a target at T and fixation at F.

the rendering and hardware display costs of the monocular condition.

Because the monocular and biocular viewing conditions degrade the fidelity with which distance is presented, we expect that the stereoscopic display will support the most accurate localization. The biocular display may, however, provide a competitive alternative for virtual objects with relatively little internal depth by avoiding the potential problem of binocular rivalry. The first experiment provides a descriptive study of relative localization accuracy of virtual objects presented via the alternative viewing conditions. Given that observer age and accommodative demand (required focus) could be expected to interact and influence localization accuracy, these display characteristics are also examined in the first study to develop designer guidelines for the adjustment of focus and selection of personnel to use virtual object displays. Subsequent studies consider the effect of the introduction of nearby physical surfaces on the localization of the virtual object, identify a phenomenon that introduces errors into such localization, and explore two alternative explanations for the phenomenon.

EXPERIMENT 1

Methods

Apparatus. All experiments reported in this paper used a custom head-mounted see-through display called an *electronic haploscope* by the authors, which is capable of presenting a 20° to 30° diameter circular monocular field to each eye with variable monocular overlap. In the following experiments, the system was used at 100% overlap and 20° field of view. The display system used two vertically mounted Citizen 1.5-inch (3.8 cm) 1000-line miniature cathode ray tubes (CRTs) in National Television Standards Committee (NTSC) mode that were driven by a Silicon Graphics (SGI) computer (4D/210GTXB) through custom video conditioning circuits. For the simple three-dimensional imagery used in the following experiments, the computer could maintain a 15-Hz graphics update rate. The CRT images were infinity collimated by standard glass telescope eyepieces (Erfl 32 mm and Ploessl 42 mm) mounted directly under the CRTs.

After the signal transformation from RGB to NTSC, individual pixels that corresponded in the current configuration to a 3 arcmin horizontal resolution measured from participants' eyes were easily discriminated. The collimated light could be modified by lenses and rotating prisms from a standard optometric trial lens set that allowed precise positioning with at least 5 arcmin resolution of the separate left and right images and allowed variation of the accommodative demand for each eye. The images were relayed to the participants' eyes by custom, partially silvered (15%) polycarbonate mirrors mounted at 45° directly in front of each eye. The left and right viewing channels could be mechanically adjusted between 55 and 71 mm separations for different participants' interpupillary distances. The video signal conditioning also allowed lateral adjustment of the video frame. Consequently, the display system could precisely position the center of each graphics viewport directly in front of the eyes of all participants for bore-sight alignment.

The entire display system, which was built around a snug fitting, rigid headband, is intended to be worn by a freely moving participant and weighs between 0.77 and 1.1 kg, depending on configuration. In the lightest configuration, the moments of inertia have been measured to be 0.0782 kg-m^2 vertical axis, 0.0644 kg-m^2 longitudinal axis, and 0.0391 kg-m^2 lateral axis when mounted on an erect head. In all of the experiments described in the following, the band was fitted to each participant's head and then supported by a special pivoted mount at the end of a 1.8-m table. This mount restricted horizontal movement but allowed some pitch movement. Participants sat at this end during the course of each experiment. The mount and chair were adjusted so that the virtual objects could be presented at eye level. Lateral head movement was restricted during all of the following experiments, but a residual pitch of ±10° was allowed for participant comfort. In practice, the participants were monitored by the experimenters so that they would keep their heads approximately level at an individually selected orientation during the course of the experiments.

Stimuli. A monocular, biocular, or stereoscopic virtual image of an upside-down, axially rotating (approximately 3 revolutions/min) pyramid was presented at a distance of 58 cm away from the participants' eyes by a head-mounted see-through display. It was seen against a wall covered by gray cloth that was 2.2 m from the participants. Preliminary experiments that examined varying the rotation rate of the pyramid for each trial showed that such variation had no effect on the localization of the virtual image. The reference distance of 58 cm was chosen for the experiment because it corresponded to a possible working distance for several industrial applications of interest to the authors. All displays were operated under moderate indoor artificial illumination (approximately 50 lux). The virtual image was presented with either 2 diopters of accommodative relief or at optical infinity. The stereo display was, however, calibrated (see the following) over a range of 30 to 110 cm.

The monocular display was simply the stereo channel that corresponded to the participants' dominant eye. The biocular display was produced by positioning the graphics eye point midway between the participant's eyes. The left and right images were identical copies of this view but were shifted laterally so that when the participants' eyes converged to the centers of each view port, they would have 0 disparity relative to the reference convergence point at 58 cm.

The plane of 0 disparity was thus set so that the participants could easily fuse the images when converged at 58 cm. This technique was used in general for all biocular stimuli at different depths that were experimentally interjected as described in the following. Though no keystone correction was applied for the several arc minutes of distortion caused by the image shift, the disparity pattern produced in this biocular image closely approximates that of a flat image of the target as if it were drawn on a transparent projection surface at the simulated convergence distance.

The wire-frame pyramid had a nominal 10-cm base and a 5-cm height. The width of the wire-frame lines was about 9 arcmin. The depicted size of the presented virtual object was randomly scaled from 70% to 130% of its nominal size for each trial to interfere with participants' possible use of angular size as a depth cue. The lines of the wire frame and all other computer-generated lines had a luminance of about 65 cd/m^2 and were seen against the gray cloth background of 2.9 cd/m^2, which had visible vertical seams. Although the presented luminance did not approach that used in aircraft HUDs, which must be visible against a 30 000 cd/m^2 background, it was just adequate for indoor work against most colored surfaces and is at least three times brighter than the luminance available in off-the-shelf, see-through head-mounted displays such as the IGlasses™ formerly made by VIO.

Out of every 30 judgments, 4 were based on unanalyzed random variations in the depicted depth of the pyramid. This variation was introduced to help ensure that the participants did not notice that the depicted depth was repeated. However, the major factors masking the repetition of the depicted depth were the perceptual effects that caused changes in the apparent depth with different viewing conditions. Because the viewing conditions could be unobtrusively intermixed and no feedback was given to the participants, there was no way for them to tell that the depicted distance was not changing. In fact, no naive participant in any of the following experiments reported noticing the repetition of the depicted target depth.

Task. The participants' task was to use a method of adjustment to position the binocularly visible physical cursor, a yellow-green LED (about 20 cd/m^2) pointer that was shaped like a pyramid (base = 0.5 cm, height = 1 cm), into vertical alignment with the apex of the inverted pyramidal virtual object. The physical cursor was moved on a rail by a chain and gear system and was positioned about 2° of visual angle below the virtual object. The distance to the pointer was automatically recorded through use of a shaft encoder interfaced to the display computer. The adjustment was self-paced, but participants were encouraged to take between 15 and 30 s for each adjustment and were allowed to take breaks at ½ to 1 h intervals as needed. As part of the standard procedure for use of human participants, all participants were informed that they could terminate the experiment at any time if they experienced any undue discomfort and were asked at the end if they experienced any viewing difficulties seeing the virtual objects.

Stereo calibration. The haploscope display system was adjusted by monocular superimposition of reference virtual images on an 18-cm diameter circumscribed circle presented at a distance of 2.2 m. In addition to position adjustment, this allowed adjustment of the field-of-view angle of the graphics system to match the total magnification of the system. This was done separately for each eye and for each participant before the experiment was started. Thus, we could account for any changes attributable to variation of accommodative demand and corrective lenses that might be worn by the participants. The participants' interpupillary distances (IPD) were measured with a binocular viewing device with digital readout (Varilux Model: Digital CRP). All displays and algorithms were adjusted to reflect the measured IPD values.

Preliminary test results for virtual targets placed between 33 and 108 cm showed that within the full range of adjustment used for the experiment, participants using a stereo display could align the cursor within ±0.3 cm of the depicted virtual object target depth (Ellis & Menges, 1997b). The distance responses were completely linear, unbiased, and unskewed and were conducted in the same full-room illumination as the experiment. Similar tests of the biocular viewing condition showed equally linear responses. Tests in the monocular condition showed inconsistent behavior, as expected. In further examination of the localization technique, pilot participants were asked to use the pointer to match the depth of physical targets. These tests showed linear, unbiased, unskewed estimates with statistical ranges of ±0.15 cm about depicted physical distances for targets used in the experiment.

Participants. The participants were five young (15–29 years) and five older (38–47 years) individuals. Participants in the older group could be presumed by population data to be at least early presbyopes (Moses, 1987). All but one young and one older participant (i.e., the authors) were naive with respect to the purpose of the experiment. The others were either paid participants recruited through the Ames Bionetics contractor or were laboratory personnel. All participants were screened on the Bausch & Lomb Orthorater stereo tests for stereoacuity better than 1 arcmin. Participants who normally wore prescription spectacles were allowed to wear them during the screening test and the experiment. During pilot testing for the experiment, inadvertent errors of 0.1 to 0.2 cm in modeling of participants' interpupillary distances in the graphics simulation produced easily noticeable artifacts. Precise stereo or biocular presentation of virtual objects evidently requires measurement and modeling of interpupillary distance with an accuracy on the order of ±0.1 cm.

Design. Viewing conditions (monocular, biocular, and stereoscopic) were crossed with accommodation (0 or 2 diopters) and nested within age groups. The experiment used a blocked design in which blocks of five replications of a given condition were presented for each of the three viewing conditions, producing uninterrupted sequences of 15 judgments. The sequence of viewing conditions was randomly assigned to each participant and thereafter systematically permuted after each set of three viewing conditions was presented. In general, it was possible to switch the viewing conditions solely through software. Thus, the participants were generally unaware of which viewing condition was presented, and the perceptual variation in apparent depth caused by variation in viewing condition was readily interpreted by them as variation in depicted depth. The viewing conditions were blocked for a given accommodative demand, which was switched by interrupting the experiment after every 15 trials to change the viewing lenses. The order of presentation of accommodative demand was permuted within subjects and balanced across participants.

Results

Analysis of variance (ANOVA) showed that the viewing conditions had a major effect on the bias of the participants' distance judgments, $F(2, 16) = 15.580$, $p < .001$. The mean stereoscopic and biocular localizations were almost completely correct, but a judgment bias appeared as an overestimate when the stereo depth cues associated with the virtual object were removed by monocular viewing. This effect interacted with accommodative demand and age as indicated in Figure 2, $F(2, 16) = 7.76$, $p < .004$. All other effects are related to this three-way interaction and will not be discussed individually. No participants reported any difficulties seeing the virtual objects during the experiment.

Discussion

The results of the ANOVA plotted in Figure 2 show that when depth cues to the virtual object are degraded to monocular conditions, judgments of its distance drop back toward the distance of the background wall at 2.2 m from the participants. A phenomenon that could explain the increased judged target distance in the monocular condition is the specific distance effect that causes visual targets of unknown physical size presented with weak or ambiguous depth cues to tend to appear in visually impoverished environments about 2 to 3 m away (Gogel & Tietz, 1973). This distance effect is also associated with tonic accommodation and vergence, which relax to approximately 1 to 2 m in the absence of

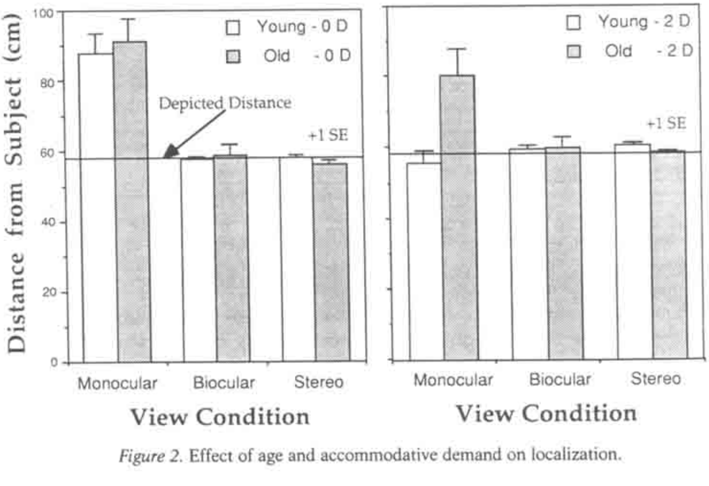

Figure 2. Effect of age and accommodative demand on localization.

distance cues (Owens & Liebowitz, 1976). Changes of convergence to a more-distant resting position could cause the localization of the virtual object to recede from the viewer. However, given that these effects are generally seen when targets appear against featureless backgrounds, they probably are not important for the present results.

Another phenomenon that is probably more relevant is the direct effect of change in ocular convergence on perceived depth (Owens & Liebowitz, 1983; Zuber, 1965). It is the kind of result to be expected if the participants' actual convergence was driven or attracted by the wall that provided a visually sharp, textured cloth with vertical seams at 2.2 m. In the monocular condition, the only source of information indicating that the virtual object is located any particular distance in front of the background is provided by accommodative demand. It is therefore not surprising that when 2 diopters of accommodative demand were provided, only the participants young enough to respond to this cue were able to localize the virtual object approximately correctly. In these participants, the accommodative-convergence reflex allowed them to maintain convergence on the virtual object. Older participants unable to respond to accommodative cues have only the disparity information provided by the background cloth to control their vergence. Thus, they would still diverge to fuse the background, and the monocularly presented virtual objects would still appear toward or on the background wall.

The fact that the monocular virtual objects were not judged to be exactly on the wall reflects the response bias within the experiment originating from constraints on movement of the physical cursor, unavoidable guesses the observers may have made about the approximate size of the object, and the possibility that the observer converged to other closer objects, such as the pointer and its supports, which, though darkened to be less conspicuous, were still visible against the background. In fact, if the observers were to hold their eyes completely still, the monocular viewing situation would be similar to that of viewing a monocular afterimage in a demonstration of Emmert's law, the classic observation that an afterimage often appears at the distance of a physical surface against which it is projected (Brown, 1965). The readily available correct disparity information in the biocular and stereo conditions, however, provides the missing cue that allowed all observers to correctly judge the distance to the virtual object.

Finally, because no difficulties seeing the virtual objects were reported during the experiment, we found no evidence that binocular rivalry interfered with participants' ability to see the virtual objects in the monocular condition. This finding is consistent with all observations we have made of the monocular wire-frame virtual objects during preparation for the experiment. Apparently, the high contrast and motion of the objects we have examined easily overcome any binocular rivalry that might be present.

EXPERIMENT 2

Experiment 1 examined the effect of different viewing conditions on the localization of virtual objects superimposed on a physical surface 2.2 m away. However, given that new uses of virtual objects are likely to bring them closer to physical surfaces, Experiment 2 examines the effect of the introduction of a much closer physical surface. In view of the interacting roles of accommodation and convergence in the discussion of Experiment 1, one could reasonably expect the introduction of a nearby real surface to cause the observers to localize monocularly viewed virtual objects at the same distance as the introduced surface. If the accommodative demand for the virtual object is already matched to its displayed distance, one would expect that this improvement in the accuracy of localization would be larger for older observers than younger ones, who already would have accommodative cues to the virtual object distance (Ellis & Menges, 1997a). Accordingly, participants of different age groups were used when correct accommodative demand to the virtual object was provided.

In order to study the effects of introduction of a physical surface, participants must first judge the distance of the virtual image by itself. This first judgment is identical to those

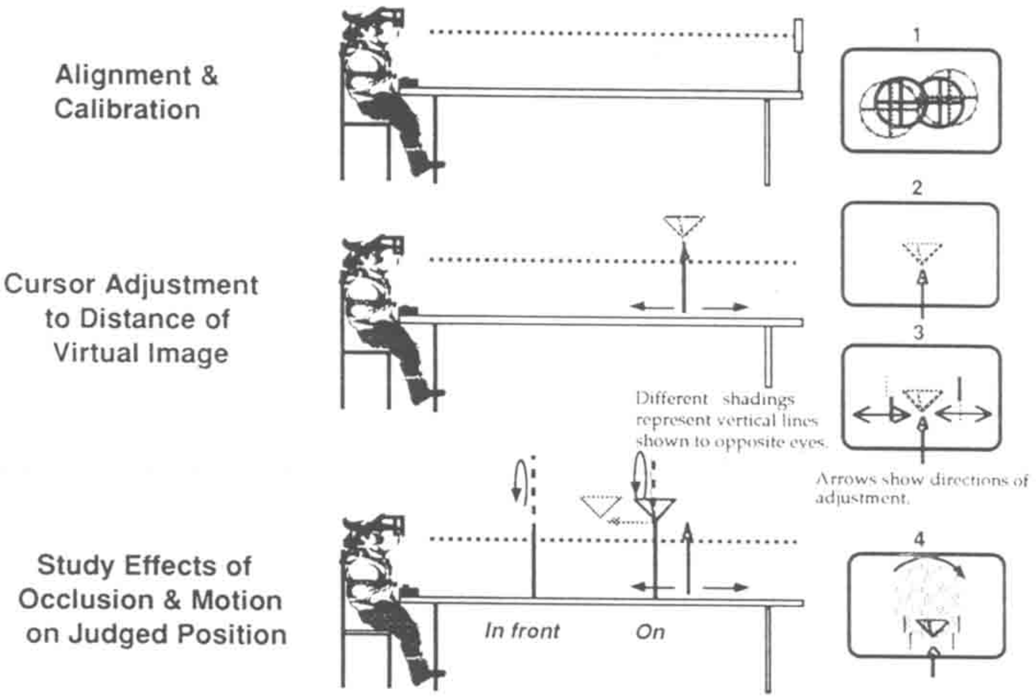

Figure 3. Experimental procedure illustration for Experiments 2, 3, and 4. Top: Alignment, magnification, and interpupillary adjustment. Middle: Initial localization of virtual object depth. Bottom: Testing conditions representing the on or in-front placement of the rotating checkerboard and the second localization of the virtual object depth. The rightmost panels 3 and 4 represent use of nonius lines to detect relative convergence in Experiment 3.

made with the viewing conditions in Experiment 1 with two-dimensional accommodative demand and provides a chance to replicate that part of the experiment.

Methods

Stimuli. The virtual image stimuli used in Experiment 2 were identical to those in Experiment 1 for the 2-diopter accommodation condition, but a new physical stimulus was introduced. This physical surface was a slowly, irregularly rotating checkerboard (approximately 2 revolutions/min) made of photocopied paper glued onto foam core and was mechanically introduced along the line of sight to the pyramid as illustrated in Figure 3. Motion of the checkerboard was introduced because preliminary testing showed that the changes in localization it produced were enhanced by motion.

The checkerboard was a disk 29 cm in diameter with 5 cm black and white checks having either 1.3 or 17.8 cd/m^2 luminance. It was positioned so that the virtual image of the pyramid could be seen against the lower rim of the disk in order to allow the participants to adjust the physical cursor to the apparent distance of the virtual image in the presence of the disk. Care was taken to ensure that the physical cursor was below the bottom of the virtual object and the edge of the disk, as in Experiment 1. As before, participants viewed the virtual objects with monocular, biocular, or stereoscopic view conditions.

Participants. Of the 13 participants, 7 were young (15–29 years) and 6 were older (38–47 years). All but one young and one older participant (the authors) were naive with respect to the purpose of the experiment. The others were either recruited through the Ames Bionetics contractor or were laboratory personnel. All participants were screened for stereo vision as in Experiment 1.

Task. The first part of the participants' task was to mechanically place the yellow-green LED pointer under the nadir of the slowly rotating, wire-frame virtual pyramid, which varied randomly in size for each trial as in Experiment 1. The second part of the task involved an adjustment of the pointer to match the pyramid's distance after the slowly, irregularly rotating, opaque checkerboard was introduced along the line of sight to the pyramid. The checkerboard was introduced at the previously judged distance of the apex of the virtual pyramid. This fact was unknown to all of the naive participants and remained unnoticed throughout the experiment. Although the virtual pyramid was also presented a second time at the same distance as the first localization, the experimental variations generally concealed this fact from the naive participants, who were led to believe that each trial, with or without the checkerboard, involved a potentially different depicted depth. As in Experiment 1, the occasional introduction of unanalyzed sham targets at different depths enforced the naive participants' belief that the virtual image could be displaced in various depths for every localization.

Results

Analysis of the participants' first localization of the virtual object under the three viewing conditions with 2 diopters of accommodative demand closely replicates the findings of Experiment 1. The basic result is a significant two-way interaction of view condition and age, $F(2, 42) = 19.160$, $p < .0009$, in which age variation affects the judged distance only when the younger participants viewed monocularly (Figure 4).

Analysis of the offset of the mean judged distance to the virtual object associated with the introduction of the physical surface also showed a main effect of viewing condition, $F(2, 26) = 91.340$, $p < .0001$, and a significant interaction between viewing condition and age, $F(2, 26) = 21.921$, $p < .0001$ (Figure 4). These effects modulated the overall significant offset, $F(1, 13) = 90.623$, $p < .0001$, of the judged distance to the virtual object toward the viewer, which was caused by introduction of the physical surface. This effect appears for all viewing conditions as a closer localization of the target after interposition of the physical surface.

Discussion

The first virtual object localization shown in Figure 4 replicates the two-dimensional viewing conditions in Experiment 1, showing that the older participants were unable to use the accommodative information to estimate

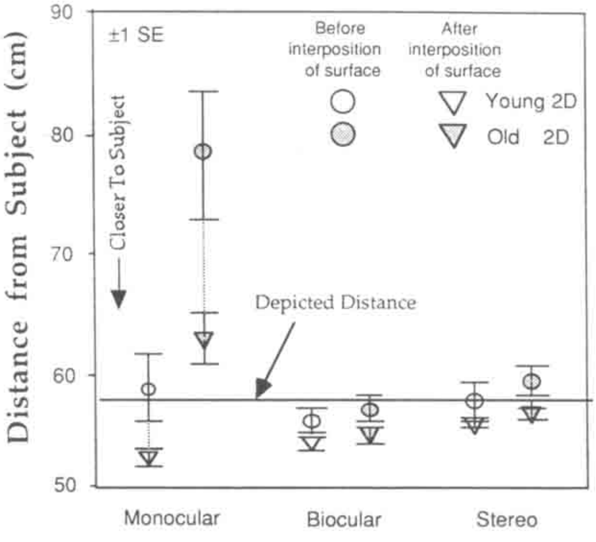

Figure 4. Interaction of age and viewing condition.

the virtual object distance and consequently localized the virtual object erroneously toward the background wall. Interestingly, interposition of the checkerboard at the judged distance of the monocularly viewed virtual object causes a substantial forward movement of the judged virtual object position (Figure 4). This change is what would be expected if insertion of the disparity and other cues to nearness of the checkerboard were to cause relative convergence with respect to the position of eyes before its introduction.

It is important to note that the forward movement of the localization is from the initial judged position for each appearance of the isolated virtual object. Given that the older participants tended to initially judge the virtual object to be too far away, the tendency for the checkerboard to bring the judged distance closer was corrective (see Figure 4). The younger participants viewing the virtual object monocularly did not significantly misjudge the virtual object distance, however, as evident from the error bars. For them, introduction of the checkerboard was detrimental, causing the virtual object to be judged too near.

In fact, the checkerboard insertion caused the virtual object to generally appear too near for the other two viewing conditions, which otherwise supported correct localization of the isolated virtual object. This was generally true for both age groups and suggests that introduction of the physical surface, the checkerboard, could be causing a small relative convergence under these conditions as well. Although disparity information for correct convergence to the virtual object is available under these two conditions, the virtual object providing this information is not of high visual quality. Its presentation corresponds roughly to 20/100 visual acuity.

Under such conditions, convergence based on stereoscopic cues might not be precise and could exhibit a fixation disparity. If this error were an exofixation disparity (i.e., it would tend to the distant side), introduction of the high-visual-fidelity checkerboard could correct it, causing a relative convergence and an associated decrease of the judged distance to the virtual object. If this small corrective convergence were incompletely compensated because of the breakdown of distance constancy, errors of localization could be expected. Thus the change in localization after introduction of the checkerboard could be attributable to a change in static convergence.

EXPERIMENT 3

Experiment 3 explicitly tested for such a change associated with the change in judged distance. Attention was focused on the stereoscopic viewing condition to see if the closer judged distance of the virtual object associated with introduction of the physical surface could be associated with an increase in static convergence. Because the amount of expected change was small (e.g., a change of 3 cm, from 58 to 55 cm), a sensitive measure of convergence was needed. Angular changes of monocular position of only about 3 arcmin would be expected if the change in localization were explainable by vergence change alone. Such a measurement is difficult to make without encumbering eye-tracking technology that preserves a clear visual field. However, it is conveniently just at the display resolution of the display configuration used. Therefore, a technique using nonius lines on the display itself was adopted (Ellis, Bucher, & Menges, 1995).

A nonius line is a line that is broken into two line segments, each of which is visible by only one eye. Such lines have typically been used to measure equivalent oculocentric directions to determine the position of the stereoscopic horopter. When the two segments are moved laterally during a period of constant convergence so that they appear to be collinear, their positions may be used to record a specific convergence position.

Methods

Stimuli. Given that pilot studies had suggested that longer distances enhance the effect of introduction of the checkerboard, the virtual pyramid was presented at 108 cm rather than 58 cm away from the participants' eyes. Such an enhancement was deemed helpful in increasing the detectability of any relative convergence

This display, otherwise similar to that used in Experiment 2, was operated under normal room illumination with 1 diopter of accommodative relief for the virtual image. Flanking nonius lines (lower right panels of Figure 3) of the same luminance and line width as the pyramid described in the following Task section were also occasionally presented to detect changes in static convergence.

Participants. Five men and one woman participating in the experiment had measured stereo resolution better than 1 arcmin. Some participants had vision corrected by contact lenses or glasses and were able to wear their corrections during the experiment. Participants' ages ranged from 17 to 47 years, and they included laboratory personnel as well as paid participants recruited by a contractor at Ames. Because of the computer control of the experiment, it was possible to conduct this experiment double blind.

Task. The participants' task had three basic parts: (a) localization of an isolated virtual pyramid and measurement of associated static ocular convergence, (b) relocalization of the virtual pyramid in the presence of a real surface either at or in front of the pyramid's apparent distance, and (c) measurement of changes in static convergence associated with the relocalization. The first part of the participant's task was to mechanically place the LED pointer under the nadir of the slowly rotating wire-frame pyramid. After aligning the pointer, the participants were presented with two sets of vertical nonius lines that were just flanking the pyramid (Figure 3, right panel 3). These lines were then adjusted to appear vertically collinear (i.e., to have equal visual directions on each side of the pyramid). This adjustment was made by moving the lower left and right segments with a joystick control (see Figure 3) and effectively recorded the participants' static convergence during this part of the experiment. Subsequent brief presentations of the nonius line will accordingly show how static convergence may have changed by revealing vertical misalignments.

The second part of the task involved another adjustment of the pointer to the pyramid's depth after the slowly, irregularly rotating checkerboard was introduced along the line of sight to the pyramid. The pyramid was then presented a second time at the same depicted depth in this new configuration. As before, the experimental variations generally concealed this fact from the participants so that they believed each trial, with or without the checkerboard, involved a potentially different depicted depth. As before, unanalyzed trials with random variation in distance were introduced to maintain the participants' uncertainty.

After the second judgment of the pyramid's depth, the nonius lines were flashed briefly (for approximately 250 ms) next to the pyramid while the participants fixated it (Figure 3, right panel 4). The participants then made a forced choice indicating whether the upper or lower pair of the flashed nonius lines appeared laterally closer. The eye assignments of each segment of the nonius lines were randomly selected so that the meaning of the alternative possibilities in terms of convergence or divergence varied randomly across the trials. The assignment of the lower part of the left nonius line and the upper part of the right line to one eye and the other upper-lower pair to the other eye produced a differential effect, doubling the relative misalignment for any given vergence change and increasing the sensitivity of the technique for detecting changes in convergence. The participant reported which of the paired nonius lines were closer by a button press.

In fact, three different experimental conditions were used in the second part of the experiment because of the need for a control case. In the *on* condition, the checkerboard was mechanically introduced at the judged depth of the virtual pyramid object so that the pyramid appeared to be on the checkerboard. For the *in-front* condition, the checkerboard was introduced 30 cm in front of the judged depth. In the control condition, the second judgment was a replication of the first judgment in that the participant made a second judgment of the depth of the virtual object. This time, however, the participant made the forced-choice judgment of the nonius lines alignment without the addition of the checkerboard. Thus, the control was identical to the experimental conditions except that the checkerboard was not introduced into the line of sight. Therefore, this control provides an individual baseline for participants' judgment biases and changes of their convergence during the course of a trial.

Each condition was repeated 15 times for each participant in a randomized block design in which blocks of five replications of each condition were repeated. The six possible orders of the three conditions were distributed randomly across the six participants in the experiment.

The change in judged distance of the virtual object was analyzed in a single-factor repeated-measures ANOVA. Chi-square analyses were conducted on each individual participant's distribution of judgments of convergence/divergence for each of the three experimental conditions. Taking the control condition as a baseline, the relative strength of convergence could be measured by a ratio of the probability of convergence in each experimental condition to the probability of convergence in each participant's individual control. This ratio allows control for the possibility that participants might have an individual bias to converge or diverge simply because of repeated presentation of the virtual object.

Results

Single-factor repeated-measures analysis of the effect of superposition of the checkerboard and virtual images replicated the previous observations that the virtual object was moved closer to the viewer, $F(2, 10) = 7.549$, $p < .01$. Individual data are shown in Figure 5. This effect was somewhat stronger for the on condition than for the in-front condition and varied in strength across the six participants. Interestingly, one participant showed essentially no effect.

The cause of this individual participant's result is illuminated by considering all participants' tendency to relatively converge during judgment of the depth of the virtual object in

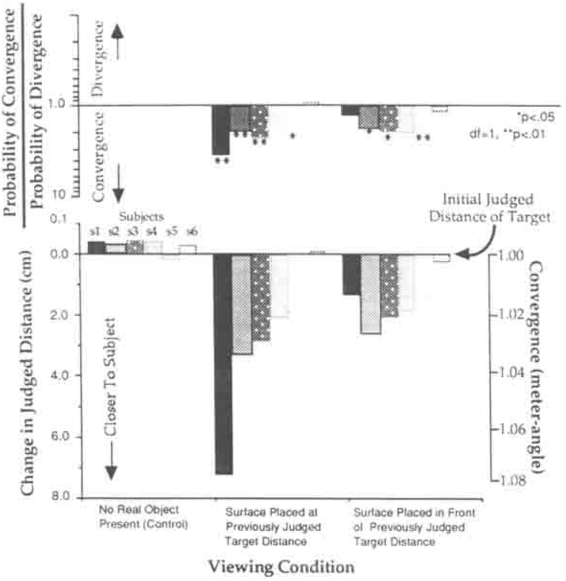

Figure 5. Relative convergence versus change in judged distance.

Figure 6. Slotted physical surface.

the presence of the checkerboard. This tendency is summarized for the experiment in Table 1, which displays the frequency of convergence or divergence indicated by the nonius judgments for all participants in the three experimental conditions, $\chi^2(2) = 20.37$, $p < .001$. The control case shows the expected 50:50 break collapsing across all participants, whereas the other two conditions show clear convergence – the on condition was somewhat stronger.

For further analysis, each participant's individual tendency to converge was computed separately as the ratio of their probability of convergence in an experimental condition to their probability of convergence in the control condition. These ratios are plotted in Figure 6 for each participant. Given that the control was used as the reference, all ratios for the control condition are 1. A 2 × 2 chi-square contingency was also computed to compare the distribution of convergence and divergence for each experimental condition with that of the control condition. This was done separately for each participant, for whom statistically significant differences in distributions are indicated by asterisks (Figure 5).

Discussion

The individual participant's localization errors in Figure 5 are sorted by the size of the change in the judged position of the virtual object for the on condition. These results can then be compared with the ratio of the convergence probabilities. As is clear from the figure, the two measurements are almost perfectly correlated across participants. The only participant not to show a displacement of the virtual object caused by the checkerboard is also the only one to show essentially no relative convergence. The participant showing the largest displacement attributable to introduction of the checkerboard is also the one with the strongest tendency to converge. The results for the in-front condition show a weaker apparent displacement of the virtual image, but they also show a correlation of convergence tendencies and changes in localization. The correlation of relative convergence with magnitude of displacement for the on and in-front conditions across participants and conditions is, in fact, $r(10) = .894$, $t = 6.31$, $p < .002$.

These results generally support the supposition that the change in judged depth could be related to a change in convergence, but the mechanism underlying this change remains to be clarified. Correction of a fixation disparity, for example, by introduction of a high-resolution physical stimulus, for example, is not the only possible mechanism.

One other possibility is that change in convergence is attributable to so-called perspective (Enright, 1991) or proximal vergence (Ciuffreda, 1992). These phenomena are changes in convergence attributable to changes in the apparent nearness of objects. They provide evidence that spatial interpretations of the distance of a visual image can simulate the

TABLE 1: Frequency of Convergence to Divergence during Depth Judgments of Virtual Objects

	Convergence	Divergence	Total Judgments	Ratio of Convergence to Divergence
On	84	21	105	1.58
In-front	70	35	105	1.32
Control	53	52	105	1.00

vergence system. Accordingly, the results from the present experiment, though showing that there is a clear oculomotor response associated with the error in judged depth, does not resolve its cause. A change in the apparent nearness of the virtual object attributable to its appearing to occlude the nearby checkerboard could be the cause of the measured convergence rather than its consequence. This question can be resolved only experimentally.

EXPERIMENT 4

One approach to analyzing whether the oculomotor effect (i.e., the convergence) observed in Experiment 3 is caused by the superposition of the virtual object on the background is to devise a stimulus condition that both strengthens the oculomotor cues to convergence and weakens the visual evidence for occlusion, thus reducing the likelihood of convergence caused by proximal vergence.

We attempted to create such a stimulus by cutting an annular slot 8-cm wide out of our rotating checkerboard so that the virtual pyramid would be just able to fall through the resulting hole (see Figure 6). The outer rim of the checkerboard was supported by thin radial wire that matched the color of the background wall and therefore was invisible to the observers. This stimulus triples the number of moving edges that provide the strong disparity discontinuity, which could be the stimulus to convergence that could cause the change in static convergence observed in Experiment 3. If the better stimulus to convergence provided by the checkerboard caused the change in static convergence, this stimulus should strengthen the effect. However, the slotted hole virtually eliminates the visual evidence for occlusion. If proximal vergence triggered by occlusion caused the change in judged distance, one would expect not to find a change in the judged distance of the virtual object when it is presented in the slot.

Methods

Participants. We used nine participants (ages 23–47) who were either laboratory personnel or paid participants provided by Bionetics.

This experiment was conducted using a methodology equivalent to that of Experiment 2 for stereoscopic virtual objects. Two different depicted distances of the virtual objects, 83 cm and 108 cm, were randomly ordered into blocks of 20 runs. In each block either a solid checkerboard or a slotted checkerboard was introduced along the line of sight of the virtual object. After introduction of the checkerboard, the change in the judged distance to the virtual object was measured. Block types were alternated for all participants. All but one participant, from whom half of the data was lost, experienced four blocks, making a total of 80 judgments per participant. The order of presentation of the two checkerboard types was counterbalanced across participants.

Results

Analysis of variance showed that whereas the solid disk caused a previously observed offset of the judged virtual object distance toward the observer ($M = 2.80$ cm; $SE = \pm 0.75$ cm), the slotted disk caused a mean change of only 0.72 cm (SE ±0.49 cm). This difference was statistically significant, $F(1, 8) = 19.605$, $p < .002$. The offset for the slotted disk condition, though close to 0, was statistically significantly less than 0, $t(8) = -2.59$, $p < .04$.

The size of the offset was larger for the greater depicted distance. The offset for the 83-cm virtual object was 1.49 cm (SE ±0.43), and the offset was 2.03 cm (SE ±0.57) for the 108-cm object. This difference was just significant, $F(1, 8) = 5.349$, $p < .05$. There was no statistical interaction, $F(1, 8) = 3.068$, $p > .05$, so the effects of placing the virtual object in the slot were statistically indistinguishable for the two presentation distances. Accordingly, the data from the two presentation distances may be collapsed as in Figure 7. This figure shows the full distribution of all of the participants' responses illustrating the effect of the presence of the slot on the judged distance to the virtual object.

Discussion

As is clear from Figure 7, introduction of the slot in the checkerboard that removed the occlusion between the virtual object and the

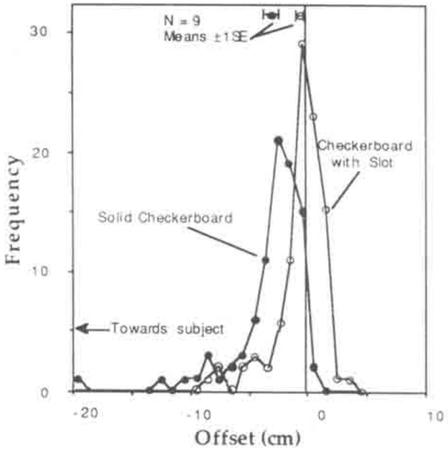

Figure 7. Effect of hole on judged distance to virtual objects.

checkerboard greatly reduced the offset in judged virtual object position produced by the checkerboard introduction. This reduction makes the oculomotor explanation of the change less likely because the binocular depth cues of the checkerboard would be expected, if anything, to strengthen any vergence response. Given that the overlap of the virtual object contours and those of the checkerboard seem to be the key features causing the shift in the virtual object's judged position, the occlusion cue placing the virtual object in front of the checkerboard seems to be the best explanation for the nearer localization of the virtual objects. Thus, the change in static convergence appears to be a consequence of proximal vergence and resembles effects in recent reports of convergence being driven by kinetic depth effects, which produce a perception of nearness (Ringach, Hawken, & Shapley, 1997).

Another possible interpretation of the results, however, could be based on the proposal that the alternative explanations of the offset of the judged distance of the virtual object are not mutually exclusive (W. Shebilske, personal communication, December 1997). In this view the introduction of the slotted checkerboard stimulus could introduce a proximal vergence tendency to fixate farther away, because any objects seen through the slot would necessarily be more distant than the slot. If this tendency for a more distant fixation were to occur, it would oppose any tendency of the visual information to reduce oculomotor bias. The net effect of the opposition could account for the almost negligible forward displacement of the virtual object observed when participants were exposed to the slotted checkerboard.

One way to dissociate the two influences could be to study the individual oculomotor biases of each participant, measuring, for example, their phoria with a Maddox Rod test or an equivalent test. To the extent that participants exhibit exophoria, one could expect a exofixation disparity while viewing the virtual object. Its removal by presentation of the checkerboards could be associated with the change in judged distance. An explanation of the observed effects of the checkerboards solely in terms of proximal vergence would not predict a correlation between the phoria tests and the participants' individual phoria variations. Accordingly, future investigations could examine individual participant's oculomotor biases as a technique to more precisely determine the cause of the offset of the change in judged distance associated with superimposition of virtual objects against physical surfaces.

DESIGN CONSIDERATIONS

1. The present results were observed with a static eye point and can be expected to change when significant lateral head movement producing motion parallax is introduced. Nevertheless, it is important to realize that because many of the new applications of head-mounted see-through displays will, in fact, involve relatively static viewing, the conditions used remain practically relevant.
2. Given that weight and cost considerations may encourage the use of monocular displays, these displays are likely to be initial candidates for many applications. Accordingly, such displays should have a variable focus control to appropriately direct the convergence of prepresbyopic users. Designers and supervisors should be aware that operators over 40 years of age will generally not benefit from the variable focus adjustment.
3. Biocular and stereo displays should be used with a bore-sighting procedure in which focus is adjusted to a reference target so as to correct for any errors in depth attributable to inappropriate vergence.
4. Computer-generated or other targets presented binocularly should have individually tailored stereo disparity to correct their spatial localization so as to compensate for the tendency of virtual objects to appear to float in front of the surfaces that they are seen against.

General Appearance of Nearby Virtual Objects Close to Physical Surfaces

In addition to affecting the localization of a virtual object, a proximate physical surface can markedly affect its appearance. For example, when a fixed-size, monocularly viewed virtual object is projected against physical surfaces at different distances, its apparent size appears to grow and shrink as the distance of the surface increases or decreases. The head restraint, the discrete nature of specific testing conditions, and the size randomization used in the present experiments prevented our participants from noticing this effect during testing. It can be very prominent, however, and it was readily noticed during pilot testing and when participants were allowed to continuously view virtual objects while their heads were unrestrained.

Another prominent phenomenal effect occurred during presentation of the opaque physical surface in front of the location of the stereoscopic virtual object in Experiment 3. In this case, the surface appears to change its physical properties and become transparent in the regions where the virtual object overlaps. The placement of the surface 30 cm in front of the virtual object was selected to ensure that this transparency would be induced for all participants. The effect is very striking because its onset can be very sudden and can depend on the distance at which the surface is placed in front of the virtual object. It is most strikingly demonstrated with a moving surface that is initially behind a virtual object. As the surface is moved forward, the object is "pushed" along in front. However, after a critical distance varying from participant to participant, the virtual object will abruptly "fall back" behind the surface and, in doing so, will impart a sense of transparency to the surface.

This transformation is quite striking, especially when the surface is an observer's hand. When the demonstration is conducted close to the observer so that changes in their ocular convergence are easily visible, a divergence associated with the falling back of the virtual object is easily seen. As is shown in Figure 5 from Experiment 3, the localization of the virtual object is nevertheless brought closer to the observer by the surface, even when the surface becomes transparent. This observation thus underscores the particular difficulty faced if a designer of an optically overlaid virtual object wishes to place the object precisely inside or behind a physical surface or object. The resultant distortions in the virtual object's localization may be correctable by a reduction in its disparity, but this technique will have to be verified by future research.

ACKNOWLEDGMENTS

Part of the head-mounted display used in this experiment was based on a mechanical design by Ramon Alarcon. Experiment 4 arose from discussions with Charles Neveu. Preliminary reports of the experiments described in this paper were made at the 1995 and 1996

annual meetings of the Human Factors and Ergonomics Society, the 1995 meeting of the Psychonomics Society, and the 1995 meeting of the International Federation of Automatic Control.

REFERENCES

Azuma, R., & Bishop, G. (1994). Improving static and dynamic registration in an optical see-through HMD. In *Proceedings of SIGGRAPH '94* (pp. 197–204). New York: ACM.

Brown, J. L. (1965). Afterimages. In C. H. Graham (Ed.), *Vision and visual perception* (p. 485). New York: Wiley.

Ciuffreda, K. J. (1992). Components of clinical near vergence testing. *Journal of Behavioral Optometry, 3*, 3–13.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Handbook of perception and cognition* (Vol. 5, pp. 71–117). New York: Academic.

Ellis, S. R., Bucher, U. J., & Menges, B. M. (1995). The relationship of binocular convergence to error in the judged distance of virtual objects. In *Proceedings of the International Federation of Automatic Control* (pp. 297–301). Laxenburg, Austria: International Federation of Automatic Control.

Ellis, S. R., & Menges, B. M. (1997a). Effects of age on the judged distance to virtual objects in the near visual field. In W. A. Rogers (Ed.), *Designing for an aging population: Ten years of human factors/ergonomics research* (pp. 15–19). Santa Monica, CA: Human Factors and Ergonomics Society.

Ellis, S. R., & Menges, B. M. (1997b). Judged distance to virtual objects in the near visual field. *Presence 6*, 452–460.

Enright, J. T. (1991). Paradoxical monocular stereopsis and perspective vergence. In S. R. Ellis, M. K. Kaiser, & A. J. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (1st ed., pp. 567–576). London: Taylor & Francis.

Fisher, S. K., & Ciuffreda, K. J. (1988). Accommodation and apparent distance. *Perception, 17*, 609–621.

Gogel, W. C., & Tietz, J. D. (1973). Absolute motion parallax and the specific distance tendency. *Perception and Psychophysics, 13*, 284–292.

Graham, C. H. (1965). Space perception. In C. H. Graham (Ed.), *Vision and visual perception* (pp. 519–520). New York: Wiley.

Janin, A. L., Mizell, D. W., & Caudell, T. P. (1993). Calibration of head-mounted displays for augmented reality applications. In *Proceedings of IEEE Virtual Reality Annual International Symposium '93* (pp. 246–255). New York: IEEE.

Krishnan, V. V., Phillips, S., & Stark, L. (1973). Frequency analysis of accommodation, accommodative vergence, and disparity vergence. *Vision Research, 13*, 1545–1554.

Moses, R. A. (1987). Accommodation. In R. A. Moses & W. M. Haret, Jr. (Eds.), *Adler's physiology of the eye* (pp. 291–310). Washington, DC: Mosby.

Nagata, S. (1993). How to reinforce perception of depth in single two-dimensional pictures. In S. R. Ellis, M. K. Kaiser, & A. J. Grunwald (Eds.), *Pictorial communication in virtual and real environments* (2nd ed., pp. 527–545). London: Taylor & Francis.

Owens, D. A., & Liebowitz, H. W. (1976). The specific distance tendency. *Perception and Psychophysics, 20*, 2–9.

Owens, D. A., & Liebowitz, H. W. (1983). Perceptual and motor consequences of tonic vergence. In C. M. Schor & K. J. Ciuffreda (Eds.), *Vergence eye movements: Basic and clinical aspects* (p. 50). Boston: Butterworth.

Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). Binocular eye movements caused by the perception of three dimensional structure from motion. *Vision Research, 36*, 1479–1492.

Rolland, J. P., Ariely, D., & Gibson, W. (1995). Towards quantifying depth and size perception in 3D virtual environments. *Presence 4*, 24–49.

Semlow, J. L., & Hung, G. K. (1983). The near response: Theories of control. In C. M. Schor & K. J. Ciuffreda (Eds.), *Vergence eye movements: Basic and clinic aspects* (pp. 175–196). Boston: Butterworth.

Weintraub, D. J., & Ensing, M. (1992). *Human factors issues in head-up display design: The book of HUD*. Wright-Patterson AFB, OH: CSERIAC.

Zuber, B. (1965). *Physiological control of eye movements in humans*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Stephen R. Ellis is head of the Advanced Displays and Spatial Perception Laboratory in the Human and Systems Technologies Branch of the Flight Systems and Human Factors Division of NASA Ames Research Center. He received a Ph.D. in psychology in 1974 from McGill University.

Brian M. Menges is an aerospace engineer at Lockheed-Martin Corporation in Sunnyvale, California. He received his B.S. in 1997 in aeronautical engineering from San Jose State University.

Date received: August 6, 1997
Date accepted: April 2, 1998