



# The Effect of Focal Distance, Age, and Brightness on Near-Field Augmented Reality Depth Matching

Gurjot Singh , Member, IEEE, Stephen R. Ellis, and J. Edward Swan II , Senior Member, IEEE

**Abstract**—Many augmented reality (AR) applications operate within near-field reaching distances, and require matching the depth of a virtual object with a real object. The accuracy of this matching was measured in three experiments, which examined the effect of focal distance, age, and brightness, within distances of 33.3 to 50 cm, using a custom-built AR haploscope. Experiment I examined the effect of focal demand, at the levels of collimated (infinite focal distance), consistent with other depth cues, and at the midpoint of reaching distance. Observers were too young to exhibit age-related reductions in accommodative ability. The depth matches of collimated targets were increasingly overestimated with increasing distance, consistent targets were slightly underestimated, and midpoint targets were accurately estimated. Experiment II replicated Experiment I, with older observers. Results were similar to Experiment I. Experiment III replicated Experiment I with dimmer targets, using young observers. Results were again consistent with Experiment I, except that both consistent and midpoint targets were accurately estimated. In all cases, collimated results were explained by a model, where the collimation biases the eyes' vergence angle outwards by a constant amount. Focal demand and brightness affect near-field AR depth matching, while age-related reductions in accommodative ability have no effect.

**Index Terms**—Perception and psychophysics, virtual and augmented reality, human performance, depth perception

## 1 INTRODUCTION

MANY compelling applications of augmented reality (AR) require interacting with real and virtual objects at reaching distances. Some examples include image-guided medical procedures (e.g., Kersten-Oertel et al. [1]), manufacturing (e.g., Caudell and Mizell [2]), and maintenance (e.g., Henderson and Feiner [3]). Among the factors that determine success is the accuracy with which observers can match the distance of a real object to an AR-presented virtual object. For example, a surgeon may need to cut to the depth indicated by an AR-presented tumor, or place a needle within the tumor. In order for AR to be useful for image-guided surgery of the brain, Edwards et al. [4] found that surgeons must be able to place a scalpel with a tolerance of 1 mm; and, in order for AR to be useful for a type of radiation therapy, Krempien et al. [5] found that a needle must also be placed with a tolerance of 1 mm.

In previous work motivated by this topic, Swan et al. [6] reported initial efforts to measure the accuracy of AR depth matching. An optical see-through AR display was used, and reaching distances of 24 to 56 cm were examined. The depth judgment was *perceptual matching*, where observers adjusted

a pointing object in depth, until they judged it to be the same distance from themselves as a target object. Fig. 1 summarizes these results, which were collected across three experiments. The pointer was always a real object, and therefore its distance from the observer could be objectively measured in the real world. In Fig. 1, the  $x$ -axis is the actual depth of the target object, and the  $y$ -axis is the depth error of the pointer. Here, error = 0 indicates that observers placed the pointer at the same depth as the target object; error > 0 indicates *overestimated* depth matches, where observers placed the pointer farther in depth than the target object; and error < 0 indicates *underestimated* depth matches, where observers placed the pointer closer than the target object. As a control condition, Swan et al. [6] examined the accuracy of matching a *real* target object, and found accuracies of 1.4 to 2.7 mm (Fig. 1a, the *real consistent* condition). However, when they examined matching a *virtual* target object, they found that observers systematically overestimated the matching distance, ranging from 0.5 cm at near distances to 4.0 cm at far distances (Fig. 1b, the *AR collimated* condition). Therefore, as illustrated in Fig. 1, there was a significant difference in depth matching real and virtual targets.

Swan et al. [6] determined that the likely reason for these results was that their AR display used collimating optics, which present virtual objects focused at optical infinity. They found the results to be very well described by a model, where this collimation causes the eyes' vergence angle to rotate outwards by a constant amount. Fig. 2 illustrates this model. Let the black points labelled  $\alpha$  and  $\alpha'$  be two real objects, with the first located close to the observer, and the

- G. Singh is with Fairleigh Dickinson University, Madison, NJ 07940. E-mail: gurjot@acm.org.
- S.R. Ellis was with NASA Ames Research Center, Moffett Field, CA 94035. He is currently retired. E-mail: ellisstephenr3@gmail.com.
- J.E. Swan II is with Mississippi State University, Mississippi State, MS 39759. E-mail: swan@acm.org.

Manuscript received 12 Nov. 2017; revised 28 Aug. 2018; accepted 30 Aug. 2018. Date of publication 13 Sept. 2018; date of current version 3 Jan. 2020. (Corresponding author: J. Edward Swan II.)  
Recommended for acceptance by B. H. Thomas.  
Digital Object Identifier no. 10.1109/TVCG.2018.2869729

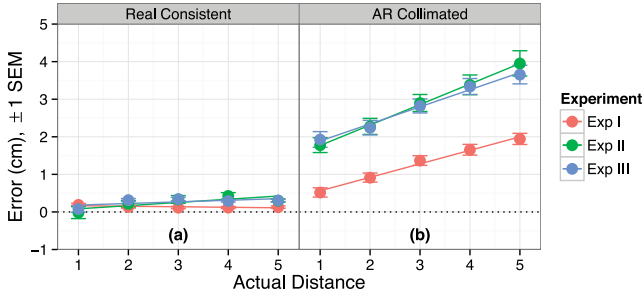


Fig. 1. The perceptual matching depth judgments from Swan et al. [6]. With increasing actual distance, real consistent targets were accurately matched, but AR collimated targets were increasingly overestimated. For Experiment I, the actual distances were 34, 38, 42, 46, and 50 cm, while for Experiments II and III the actual distances were 55, 63, 71, 79, and 87% of each observer's maximum reach.

second located farther away. And, let the red points labelled  $\beta$  and  $\beta'$  be two virtual objects, which are rendered to be the same distance as the real targets.  $\alpha$ ,  $\alpha'$ ,  $\beta$ , and  $\beta'$  also represent the angle of binocular parallax that the eyes make when the observer fixates on each object. Therefore, when fixating on the close real object, the angle of binocular parallax is  $\alpha$ , and if the fixation changes to the close virtual object, then the collimation causes the eyes to rotate outwards, reducing the angle to  $\beta$ . When placing the real pointer at the same depth as the virtual target  $\beta$ , observers' eyes rotate inwards and outwards as they fixate between the two objects, and therefore observers perceive them to be located at the same depth. The model predicts that this change in vergence angle,  $\Delta v = \alpha - \beta$ , is constant for reaching distances. Therefore, when fixating on the far target,  $\alpha > \alpha'$ , and this same change in vergence angle,  $\Delta v = \alpha' - \beta'$ , causes a larger depth distance between  $\alpha'$  and  $\beta'$  (Fig. 2). This model explains three properties of Swan et al.'s [6] results (Fig. 1): (1) because the collimating optics cause the eyes to rotate outwards, the depth judgments of the virtual targets are overestimated relative to the real targets, (2) the amount of overestimation increases with increasing distance, and (3) the results are very well fit with a linear model.

This analysis suggests that, for accurate depth placement, virtual objects need to be presented with a *focal depth*—also termed *accommodative demand*—that is consistent with their intended depth. Then, the eyes' vergence angle will not be biased, and depth matches will be more accurate. This paper reports three experiments that systematically examine this hypothesis.

However, it was not possible to conduct these experiments with the same AR display as Swan et al. [6]. That display, an NVIS Inc. nVisor ST60 model, contains unadjustable collimating optics, which always present virtual objects focused at optical infinity. This is consistent with the vast majority of commercially available AR displays, almost all of which have a focal distance that is set at the factory, and unadjustable by the end user.<sup>1</sup> Therefore, an *augmented reality haploscope*—an AR display mounted on an optical workbench, which allows

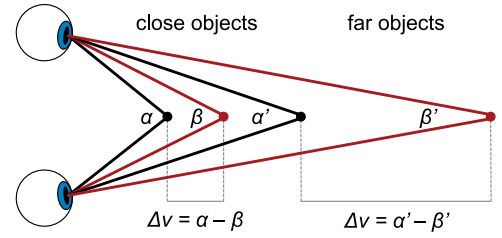


Fig. 2. The model that explains how a constant change in vergence angle,  $\Delta v$ , leads to matched distances of virtual objects (red:  $\beta$ ,  $\beta'$ ), relative to real objects (black:  $\alpha$ ,  $\alpha'$ ), that are increasingly overestimated with increasing distance.

accommodative demand and vergence angle to be independently and precisely adjusted—was developed and used for the experiments reported here.<sup>2</sup>

## 2 BACKGROUND AND RELATED WORK

### 2.1 Depth Perception and Depth Cues

The human visual system achieves a percept of perceived depth from *depth cues*—sources of perceptual information related to depth. At least nine depth cues have been identified (Cutting and Vishton [9]): *occlusion* (a closer object occludes farther objects), *binocular disparity* (an object projects to different locations on each retina), *motion perspective* (objects at different distances from a moving observer have different apparent velocities), *height in the visual field* (starting from the horizon, closer objects are lower in the visual field), *relative size* (among objects of the same size, the farther object projects to a smaller retinal angle), *accommodation* (the lens of the human eye changes shape to bring objects into focus), *vergence* (the two eyes rotate to fixate on an object (Fig. 2)), *relative density* (for a textured surface, at farther distances more objects are seen within a constant retinal angle), and *aerial perspective* (objects at great distances lose color saturation and contrast).

Depth cues differ in effectiveness based on various visual characteristics, such as scene content and distance from the observer. Nagata [10], and later Cutting and Vishton [9], organized the relative effectiveness of different depth cues according to distance. Within near-field reaching distances, they find that the operative depth cues, in approximate order of decreasing salience, are *occlusion*, *binocular disparity*, *motion perspective*, *relative size*, *accommodation and vergence*, and *relative density*.

Most of these depth cues can be categorized as *retinal*, because the cue information comes from the visual scene sensed on the retina. However, the cues of accommodation and vergence are *extra-retinal*, because the cue information comes from the muscles that control the lenses' shape and the eyes' vergence angle. In principle, the extra-retinal cues could provide absolute egocentric depth information (Gillam [11]). In contrast, retinal cues can only provide relative depth information between objects in the scene; these cues require an external reference to establish the scene's overall scale. However, when combined with extra-retinal cues, and an observer's constant interpupillary distance, retinal cues can

1. The authors of the current paper, who have been studying virtual and augmented reality for 9, 31, and 19 years, respectively, can only recall a single commercially-available AR display—the Microvision Nomad from the early 2000's—which came with a continuously variable focus adjustment (Gabbard et al. [7]).

2. Portions of these experiments are reported in a PhD dissertation (Singh [8]).

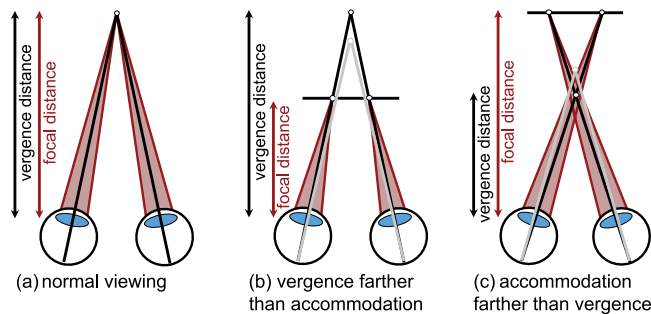


Fig. 3. The vergence-accommodation conflict, and its effect on perceived depth. (a) In normal viewing of real world objects, the vergence distance, required for zero binocular disparity, is the same as the focal distance, required for minimal focal blur. (b) When the vergence distance is farther than the focal distance, e.g., when viewing a virtual object beyond the surface of a stereo monitor, the vergence angle is biased inwards (grey lines), and the object is seen as closer than encoded by disparity. (c) When the vergence distance is closer than the focal distance, e.g., when viewing a virtual object in front of the surface of a stereo monitor, the vergence angle is biased outwards (grey lines), and the object is seen as farther than encoded by disparity.

also provide absolute depth information (Bingham and Pagano [12], Mon-Williams and Tresilian [13]). In general, the way the human visual system combines information from different depth cues to produce a stable percept of distance is subtle and not fully understood, although many theories have been advanced, and the collected evidence favors some theories over others (Landy et al. [14], Singh [8]).

## 2.2 Vergence and Accommodation

Visual perception requires a rapid series of precise eye movements (Leigh and Zee [15]), including *fixation* (hold an image steady on the fovea by minimizing eye movement), *saccadic* (quick movement that projects an object of interest to the fovea), *smooth pursuit* (retain fixation on an object during smooth movement of either the object or head), *vestibular* (hold vision steady during head movements), and *vergence* (the two eyes rotate to fixate on an object of interest (Fig. 2)). When changing fixation from a far to a near object, the eyes converge, the lenses become thicker, and the pupils constrict. These three actions—vergence, accommodation, and changing pupil size—are interlinked physiologically, and the mechanism of these three simultaneous reflexes is called the *near triad*. Because of the interlinkage, changes in either accommodation or vergence drive corresponding changes in the other two components of the triad (Semmlow and Hung [16]). Apart from the influence of accommodation and vergence, pupil diameter also changes according to scene illumination, becoming larger in dim settings and smaller in bright settings. Although these illumination-driven changes in pupil diameter affect the eye's optical depth of field, and therefore could potentially affect accommodation, little effect of changing pupil diameter on accommodation has been observed (Ripps et al. [17]). Therefore, in near field viewing, vergence and accommodation are the main depth reflexes, and the link between them is known as the *vergence-accommodation reflex*. Because of this reflex, accommodation and vergence operate in unison: changes in accommodation drive changes in vergence (*accommodative vergence*), and changes in vergence drive changes in accommodation (*vergence accommodation*) (Kersten and Legge [18]). Therefore,

the vergence reflex is driven both by binocular disparity (the eyes rotate to bring a fixated object to a level of zero binocular disparity), as well as accommodative vergence. Likewise, the accommodation reflex is driven both by focal blur (the lenses adjust to minimize blur), as well as vergence accommodation (Mon-Williams and Tresilian [19]).

The vergence-accommodation reflex is calibrated for viewing real world objects, which present consistent binocular disparity and focal blur cues (Fig. 3a). When viewing virtual objects, the binocular disparity and focal blur cues are often inconsistent, because the focal blur cue is fixed at the screen depth, while the depth of the binocular disparity cue varies, sometimes beyond the screen depth (Fig. 3b), and sometimes in front (Fig. 3c). This is called the *vergence-accommodation conflict*, and it is a ubiquitous aspect of all stereo displays with a single focal plane (Kruijff et al. [20]). The conflict causes visual fatigue (Gabbard et al. [7], Lambooij et al. [21]), hinders visual performance (Hoffman et al. [22]), and biases depth perception towards the screen depth (Fig. 3, Swenson [23], Mon-Williams and Tresilian [19]).

The contribution of vergence to perceived depth depends upon properties of the scene. At near-field distances, vergence has been conclusively found to provide egocentric depth information (Brenner and Van Damme [24], Owens and Liebowitz [25], Tresilian et al. [26], Viguier et al. [27], Foley [28]). Although vergence in isolation is not a very accurate depth cue, observers are very sensitive to changes in vergence, which generally allows accurately matching the depth of one object with another (Brenner and Van Damme [24]). Each individual has a different vergence resting point—their *dark vergence*—which is the vergence angle that their eyes assume when the controlling muscles are completely relaxed. In low light conditions, the egocentric depth specified by vergence is biased towards each individual's dark vergence distance (Owens and Liebowitz [25]). As a depth cue, vergence is most effective within near-field distances of 2 meters (Viguier et al. [27]), a distance range that encompasses ~90% of vergence eye movements (Tresilian et al. [26]). As other retinal depth cues become available, the contribution of vergence to perceived depth is reduced, but still present (Foley [28]).

As discussed above, accommodation influences perceived depth through the vergence-accommodation reflex. Although some studies have found evidence that accommodation alone can serve as a depth cue for some observers, these experiments require careful experimental setups to eliminate other depth cues, and the consensus remains that accommodation influences perceived depth through its effect on vergence (Mon-Williams and Tresilian [19]). Similar to dark vergence, each individual has a *dark focus*—the distance their eyes focus when the controlling muscles are in a relaxed state (Iavecchia et al. [29]). The dark focus biases the eye's focal response, resulting in a number of perceptual consequences, including perceived depth (Roscoe [30]). Generally, the dark focus and dark vergence distances vary independently, and for most individuals are not equal (Owens and Liebowitz [25]).

## 2.3 Accommodation and Age

Accommodative ability—the distance range where objects can be seen with minimal focal blur—decreases with



increasing age (Duane [31]), a condition known as *presbyopia*. It is primarily caused by hardening of the crystalline lens, although other physiological changes in the lens, connective tissue, and controlling muscles also play a role (Kasthurirangan and Glasser [32]). As measured by Duane [31], presbyopia begins by the age of 12, but through the early 30's the loss is minuscule—the closest distance of clear focus declines from  $\sim 8$  to  $\sim 13$  cm. However, the decline then accelerates, and by the age of 50 often surpasses 50 cm. At some point in the 40's, the closest distance of clear focus often surpasses standard reading distance, and reading glasses are required. By their mid-50's, most people have lost the ability to adjust the distance of clear focus.

It seems reasonable that this loss of accommodative ability would have perceptual consequences, and indeed, older people are worse than younger people at many perceptual tasks (Bian and Andersen [33]). However, accommodative vergence does not diminish with age; even as the visual system loses the ability to adjust accommodation, the eyes still verge properly in response to accommodative stimuli (Heron et al. [34]). Because vergence is the primary source of depth information from the vergence-accommodation reflex (Mon-Williams and Tresilian [19]), this suggests that depth perception could be unaffected by presbyopia. Indeed, Bian and Andersen [33] found that, when making judgments of medium-field egocentric distances, older people (average 73.4 years) were *more* accurate than younger people (average 22.5 years). This is one of a series of recent studies that have found that older observers preserve their abilities in tasks related to distance perception (Bian and Andersen [33]).

## 2.4 Accommodation and Scene Flatness

Another effect of the vergence-accommodation conflict in stereo displays is that the accommodative distance changes the perceived *flatness* of the scene (Andersen et al. [35], Nagata [10], Singh [8]). Specifically, when medium- to far-field scenery is shown on a display, but accommodative distance is in the near field, depth distances between scene objects are compressed, and the scene is perceived as being a flat window, positioned some depth distance from the observer. However, when the same scene is shown with collimation, these depth distances are no longer compressed, and the scene objects appear to extend in depth, with some closer to the observer and others farther. This is a reason why many augmented and virtual reality displays, especially those used for flight simulation and other far-field applications, present collimated light (Watt et al. [36]). Likewise, the NVIS nVisor ST60 used by Swan et al. [6], which also presents collimated light, was originally marketed for military training and forward observer tasks, which primarily involve medium- to far-field distances.

## 2.5 Depth Perception and Brightness

Among objects of the same size and distance, the brighter appear closer than the dimmer. This principal has long been known in art, and is discussed by Leonardo Da Vinci in his *Notebooks* (McCurdy [37]). The principal has been thoroughly studied, at both near field (Ashley [38], Farné [39]) and medium field (Coules [40]) distances, and in both monocular and binocular conditions (Coules [40]). In addition to brightness, the contrast between an object and the

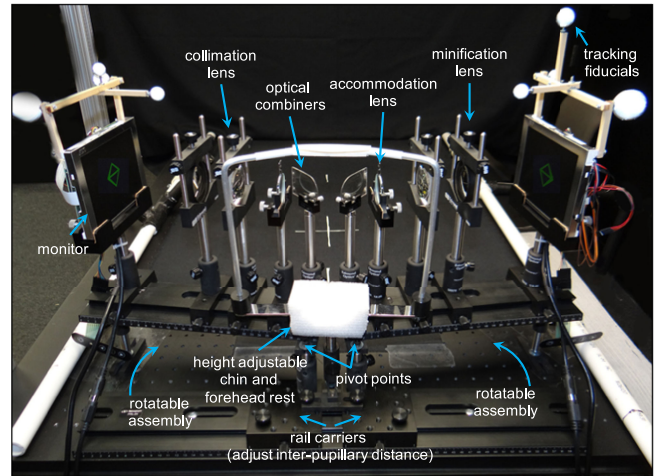


Fig. 4. The Augmented Reality (AR) haploscope. The physical design allows independent adjustment of vergence angle and focal distance.

background also effects perceived depth, so a dark object against a light background can appear closer than an object with less contrast (Farné [39]). Among the theories that explain this effect are that brighter objects stimulate a larger area on the retina, and that brighter objects effect pupil size, which then biases other near triad reflexes.

## 2.6 Related Work in Augmented Reality

To date, including Swan et al. [6], only a small number of papers have examined near-field AR depth matching. Ellis and Menges [41] measured the effects of vergence, accommodation, observer age, viewing condition (monocular, biocular stereo, binocular stereo), and the presence of an occluding surface. They found that accuracy is degraded by monocular viewing and an occluding surface. Using the same experimental setup, McCandless et al. [42] additionally studied motion parallax and latency in monocular viewing; they found reduced accuracy with increasing distance and latency. Singh et al. [43] found that an occluding surface has complex accuracy effects, and Rosa et al. [44] found increased accuracy with redundant tactile feedback.

## 3 THE AUGMENTED REALITY HAPLOSCOPE

As motivated in Section 1, an *augmented reality haploscope* was designed and engineered.<sup>3</sup> The design was loosely based on the AR haploscopes described by Rolland et al. [45] and Ellis and Menges [41], but similar designs have a long history in the study of depth perception (e.g., Swenson [23]). Fig. 4 shows the AR haploscope. The primary structure is built on three optical rails: two 12-inch rails serve as mounting bases for left-eye and right-eye optical systems, and both 12-inch rails are mounted on a 24-inch rail using rail carriers, which can be adjusted to match the required interpupillary distance.

The goal of each optical system is to collimate the generated image, so the image is located at optical infinity, or 0 diopters (D).<sup>4</sup> Then, the collimated image can either be left

3. Additional technical details can be found in Singh [8].

4. Diopters, the standard unit of lens power  $p$ , are the inverse of focal distance  $f$  in meters:  $p = 1/f$ .

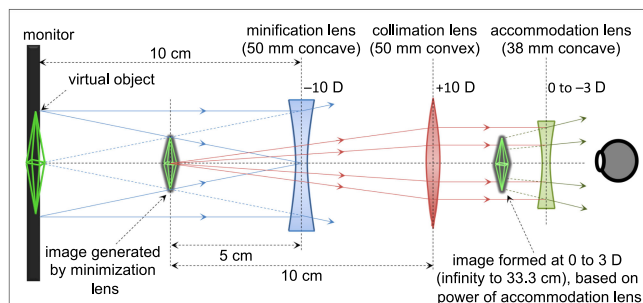


Fig. 5. The optical system of the AR haploscope. Changing the accommodation lens changes the focal distance.

at optical infinity, or a negative power lens can reduce the focal distance. Fig. 5 shows the optical design. The image is first generated by a monitor. Then, the image is minified by a  $-10$  D concave lens; without minification, only a small part of the monitor can be seen through the optical system. As shown in Fig. 5, when this  $-10$  D lens is placed 10 cm from the monitor, it creates a minified image at  $-5$  cm. This minified image is then collimated by a  $+10$  D convex lens, positioned 10 cm from the image. The collimated image is then passed through an accommodation lens. This comes from a standard optometric trial set; either a 0 D plain glass lens, which retains the collimation, or a negative power concave lens, which reduces the focal distance. In the experiments reported here, the strongest accommodation lens used was  $-3$  D, which resulted in a 33.3 cm focal distance. After generation, the images are reflected into the observers' eyes by 15 percent reflective optical combiners, mounted at  $45^\circ$  directly in front of each eye. Fig. 4 shows the monitors; the minification, collimation, and accommodation lenses; and the optical combiners.

Fig. 6 illustrates how the haploscope matches different vergence angles. The rail carriers are adjusted so that the distance between the pivot points matches the observer's interpupillary distance (Fig. 4). The chin and forehead rest is adjusted so that these pivot points are directly below the rotational centers of the observer's eyes. As illustrated in Fig. 6, when the left and right optical systems then rotate about the pivot points, for all convergence distances the view rays from the center of the two eyes stay in line with the principal axes of the optical systems. This allows presentation of a virtual object at any distance, near ( $n$ ), medium ( $m$ ), or far ( $f$ ), while the observer's view rays continue to pass through the middle of the optical systems, where optical distortion is minimized. In addition, as the virtual object distance changes, the distance between the observer's pupils also changes (Lee et al. [46]); this design models and tracks this changing interpupillary distance. To display a target object at a specific distance, the optical systems are rotated to the matching convergence angle  $1/2\alpha$  (Figs. 2, 6);  $1/2\alpha = \arctan(i/2d)$ , where  $i$  is the observer's interpupillary distance, and  $d$  is the target distance. The angle of each optical system is measured by a constellation of tracking fiducials attached to each monitor (Fig. 4).

## 4 EXPERIMENT I: ACCOMMODATION

As discussed in Section 1, Swan et al. [6] hypothesized that the linearly increasing overestimation they found with

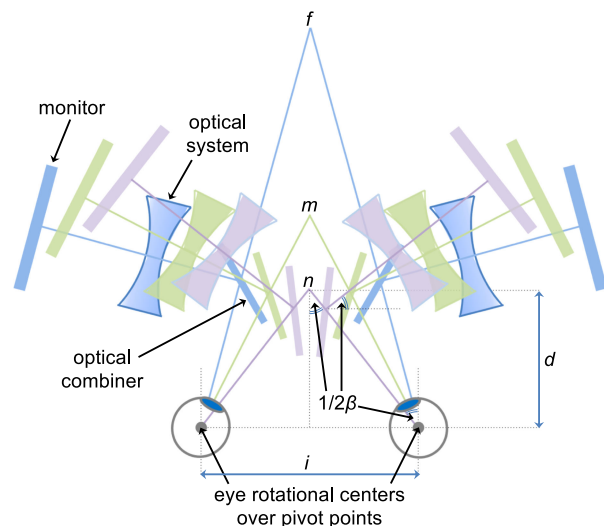


Fig. 6. Rotating the optical systems to match the correct vergence angle.

collimated AR graphics (Fig. 1), was caused by the collimation biasing the eyes' vergence angle to rotate outwards by a constant amount (Fig. 2). The purposes of Experiment I were to test aspects of this hypothesis, using the same matching task and within a similar range of near-field distances. Because Experiment I used a different display—the AR haploscope—the first purpose (1) was to replicate the *real consistent* and *AR collimated* conditions of Swan et al. [6]. If Experiment I found similar results, that would suggest that these results generalize to AR more broadly, and are not specific to the NVIS display used by Swan et al. [6]. The next purpose (2), the *AR consistent* condition, was to test whether presenting AR objects at a focal distance that was *consistent* with the distance specified by other depth cues, especially binocular disparity, would result in more accurate depth matches than what was seen in the AR collimated condition. If the depth matches are more accurate, that would further support the hypothesis that collimated graphics bias the eyes' vergence angle outwards. However, for many AR applications, always presenting virtual objects at consistent focal distances is unlikely to be practical. Therefore, the final purpose (3), the *AR midpoint* condition, was to test whether presenting AR objects at a focal distance equal to the *midpoint* of the tested range would result in performance similar to the consistent condition. If the performance is similar, this would suggest that, for accurate depth matching within reaching distances, the expense of making the focal demand consistent for every virtual object is not necessary.

## 4.1 Method

### 4.1.1 Apparatus and Task

Fig. 7 shows the experimental setup. The haploscope was mounted on the end of an optical breadboard, 244 cm long by 92 cm wide. The breadboard was supported by a custom-built aluminum table, with six legs. Mounted to the legs of the table were six hydraulic jacks, which could lift the entire table, so the surface could be adjusted to be between 104 and 134 cm above the ground. This adjustability allowed the table to be comfortably positioned for observers of many different heights. Aluminum arms extending above the table supported tracking cameras, as well as an overhead light (Fig. 7). Because the tracking cameras and light were attached to the table,



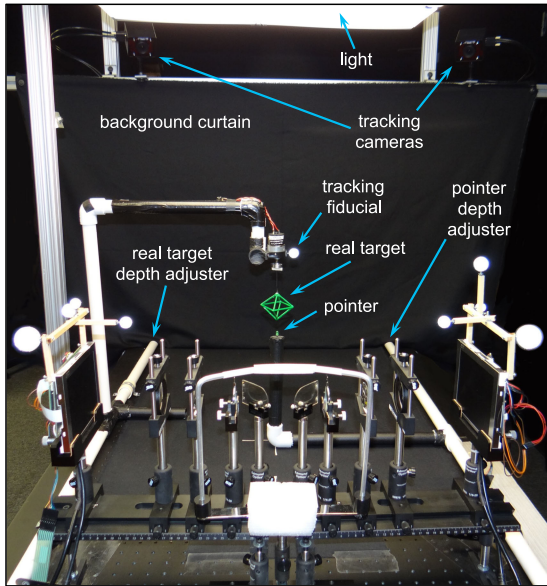


Fig. 7. The experimental setup. The AR haploscope was mounted on the end of an optical breadboard. Real and AR targets were positioned at different depths from the observer. The depth of the targets was matched by changing the position of the pointer.

when the table height was adjusted, their distance above the table top remained constant. Tracking was provided by a 2-camera TRACKPACK system, from A.R.T. GmbH.

On both sides of the table, *depth adjusters*—plastic pipes running through collars—could easily be slid back and forth in depth (Fig. 7). When the real target was presented, it hung from an arm attached to the left-hand depth adjuster. The real target was a wireframe octahedron, 5 cm wide by 6 cm high, constructed of balsa wood and painted green. An electric motor rotated the target at 4 rpm. Although slow, the rotation gave a definite sense of three-dimensional structure from motion, even when viewed monocularly. The depth position of the real target was precisely measured by a tracking fiducial mounted to the arm (Fig. 7). When an AR target was presented, the arm supporting the real target was removed. The AR target was identical to the real target: a green octahedron that rotated at 4 rpm, rendered and viewed through the haploscope optics. Only the green channel was used, which eliminated chromatic distortion. Careful calibration ensured that the AR target matched the real target in size and position at all tested distances. In addition, because accommodation lenses of different powers change the overall magnification of the optical system (Fig. 5), the calibration was repeated for every lens power. The targets were located 29 cm above the tabletop, and seen against a black curtain hung 1.2 meters from the observer (Fig. 7). The appearance of the real and AR targets was as similar as possible: the lighting and color of the real target made it appear to glow against an otherwise dark background, and it did not cast any visible shadows or reflections. The table was covered with black cloth, which created a smooth and featureless surface under the target.

The matching task from Swan et al. [6] was replicated. The pointer was made of green, translucent plastic, ~4 mm in diameter, with a rounded top, mounted on an arm attached to the right-hand depth adjuster (Fig. 7). Observers matched the target depth by sliding the depth adjuster until

the pointer was directly below the bottom point of the rotating target. The distance between the bottom of the target and the top of the pointer was ~1 cm. The depth position of the pointer was precisely measured by a tracking fiducial mounted on the arm (not visible in Fig. 7).

#### 4.1.2 Experimental Design

*Observers.* 40 observers were recruited from a population of university students and staff. The observers ranged in age from 18 to 38; the mean age was 20.9, and 18 were male and 22 female. 10 observers were paid \$12 an hour, and the rest received course credit.

*Independent Variables.* Observers saw 4 different conditions: real consistent, AR collimated, AR consistent, and AR midpoint. The target object appeared at 5 different distances from the observer: 33.3, 36.4, 40, 44.4, and 50 cm, which correspond to 3, 2.75, 2.5, 2.25, and 2 D. Observers saw 6 repetitions of each distance.

In the *real consistent* condition, observers saw the real target object (Fig. 7), which, by definition, was always presented at a focal distance that was consistent with its actual distance. In the remaining conditions, the AR target was seen. In the *AR collimated* condition, a 0 D plain glass accommodation lens was used, presenting the target at optical infinity. In the *AR consistent* condition, the accommodation lens power—3, 2.75, 2.5, 2.25, or 2 D—was always consistent with the target's presented distance. Finally, in the *AR midpoint* condition, the 2.5 D accommodation lens was used, presenting the target at a focal distance of 40 cm.

*Dependent Variables.* The primary dependent variable was *judged distance*—the measured position of the pointer (Fig. 7). In addition,  $error = judged\ distance - actual\ distance$  was also calculated (Fig. 1).

*Design.* A mixed design was used, with condition varying between observers, and distance and repetition varying within each observer. There were 10 observers in each condition, and the presentation order of condition varied in a round-robin fashion, so each group of 4 observers covered all conditions. For each observer, distance  $\times$  repetition was randomly permuted, with the restriction that the distance changed every trial. Therefore, each observer completed 5 (distance)  $\times$  6 (repetition) = 30 trials, and the experiment collected a total of 40 (observers)  $\times$  30 (trials) = 1200 datapoints.

#### 4.1.3 Procedure

After receiving an explanation of the experimental procedures, an observer gave informed consent. Then, they took a stereo vision test, which measured their sensitivity to depth changes encoded by binocular disparity. Next, the observer's interpupillary distance was measured, using a pupilometer set to optical infinity, and the haploscope was adjusted to match this distance. The task was then explained, using the real target and the pointer. If the observer indicated that, when working at the demonstrated distances, they would normally wear corrective optics (glasses or contacts), they were instructed to wear the optics. Observers then donned safety goggles, which easily fit over glasses. The goggles had 3.5 cm circular openings for each eye, and were otherwise covered with black gaffer tape. The size of these openings was calibrated so that, when looking

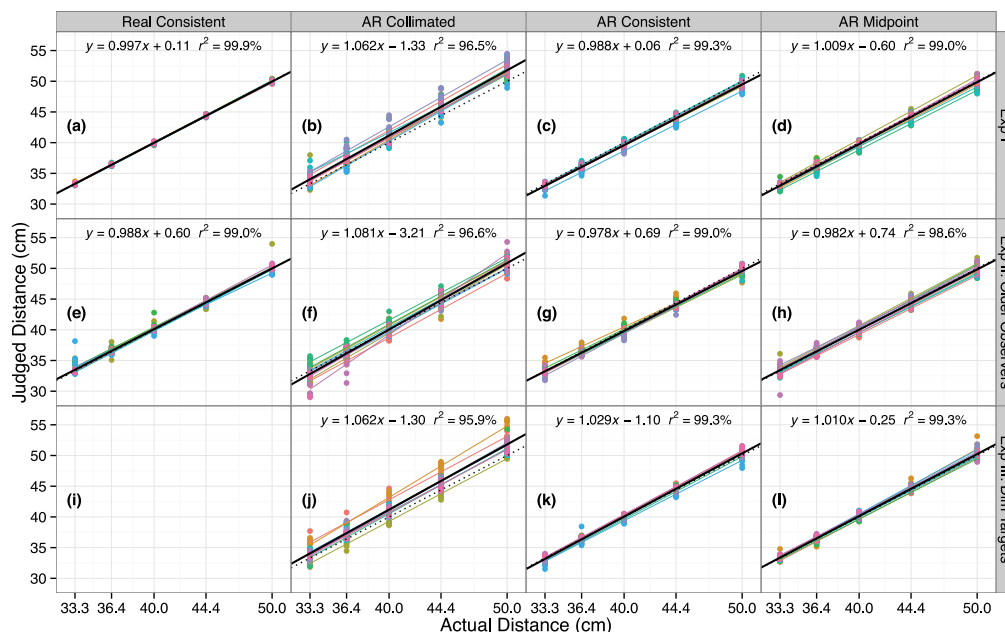


Fig. 8. The results for all three experiments, plotted as judged against actual distance. Each panel shows the individual data points, color coded according to observer, and fit with a color-coded regression line per observer. For each panel, the dotted line represents veridical performance, and the solid black line is the overall regression line. The corresponding regression equation is at the top of each panel. The  $r^2$  values indicate that the data in each panel is very well described by the regression equation. A separate group of 10 observers provided the data for each panel. The graph summarizes  $N = 1200$  (Experiments I and II) and  $N = 900$  (Experiment III) data points.

through the haploscope optics, observers could see the complete field of view provided by the optical combiners, but their peripheral view of the rest of the haploscope was blocked. The chinrest and forehead rest were adjusted so that the observer's eyes were approximately centered within the optical combiners, and the haploscope pivot points were approximately centered under the eyes' rotational centers (Figs. 4, 6). The table and chair heights were adjusted so the observer was sitting comfortably.

The observer then completed one of the four conditions. The pointer was placed at a random position within the trackable distance of 23 to 67 cm from the observer, and the experimenter then displayed the first target distance. Using their right hand to manipulate the pointer depth adjuster (Fig. 7), the observer moved the pointer from this starting position to match the target's depth. The observer then closed their eyes, and the experimenter displayed the next target distance. The observer then opened their eyes, and moved the pointer from the previously matched distance to the new distance. This pattern continued until all trials were completed. To display distances with the real target, the experimenter used the real target depth adjuster to slide the real target to the correct position. For the AR target, the experimenter adjusted the angle of each haploscope arm, and swapped out the accommodation lenses as needed. Regardless of condition, the procedures were as similar as possible, and the time required for each trial was approximately equal. During real consistent trials, observers looked through the haploscope optics, even though the monitors were switched off.

After the trials, the observer was debriefed. The overall experiment took approximately one hour.

## 4.2 Analysis

Similar to Swan et al. [6], the data was analyzed by examining the slopes and intercepts of linear equations that predict

judged distance from actual distance. Multiple regression methods determine if the slopes and intercepts significantly differ (Pedhazur [47], Cohen et al. [48]). For data with this structure, multiple regression methods are preferable to ANOVA analysis, because multiple regression allows the prediction of a continuous dependent variable (judged distance) from a continuous independent variable (actual target distance), as well as a categorical independent variable (condition). In contrast, ANOVA analysis only examines categorical independent variables, which results in a significant loss of power when an independent variable is inherently continuous (Pedhazur [47]). In addition, multiple regression yields slopes and intercepts, which as descriptive statistics are more useful than means, because they directly describe functions that predict judged distances from actual target distances. Finally, multiple regression methods focus on effect size, as opposed to significance; an analytic approach advocated by many in the applied statistics community (Cohen et al. [48]).<sup>5</sup>

Figs. 8a, 8b, 8c, 8d and 9 show the results from Experiment I, plotted as a scatterplot of judged against actual distance (Fig. 8), as well as mean error against distance (Fig. 9). Both figures indicate that the data is very well fit by linear regressions; note the  $r^2$  values in Fig. 8. Fig. 10 shows multiple regression analysis, which compares pairs of panels from Fig. 8 against each other; each panel in Fig. 10 examines two independent variables: a continuous variable (actual distance), and a categorical variable (a pair of panels from Fig. 8). To properly account for repeated measurements, for each observer at each distance, the

5. Custom analysis software, developed by the third author, was used, which implements methods described by Pedhazur [47]. A more detailed discussion of the application of multiple regression methods to depth perception data is available in Swan et al. [6].

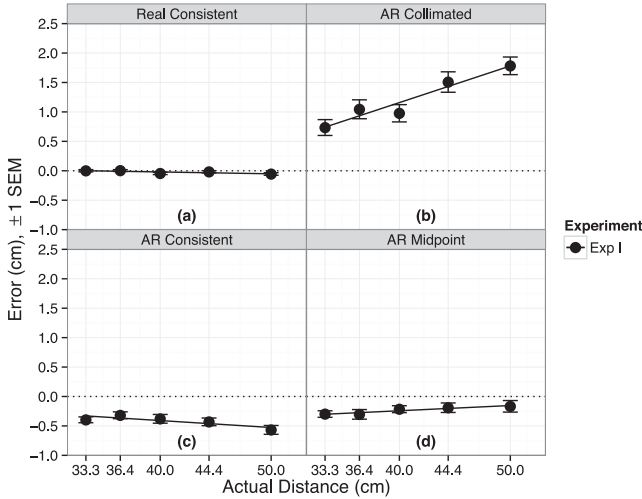


Fig. 9. Experiment I, plotted as mean error against distance ( $N = 1200$ ).

responses were averaged over the 6 repetitions, reducing the size of the analyzed dataset from 1200 to 200 points—note the reduced density of points in Fig. 10 relative to Figs. 8a, 8b, 8c, 8d.

Each panel in Fig. 10 compares two regression equations from Fig. 8. The multiple regression analysis operates in the following manner: First, the *slopes* of the equations are tested to see if they significantly differ. If they do, as in Fig. 10a, both equations from Fig. 8 are reported as the best overall description of the data in the panel. If the slopes of the equations do not significantly differ, then the *intercepts* of the equations are tested to see if they significantly differ. This test first sets the slopes of the equations—which do not differ—to a common value. If the intercepts significantly differ, as in Fig. 10b, two regression equations, with slopes adjusted to a common value, are reported as the best overall description of the data in the panel. If neither the slopes nor the intercepts significantly differ, as in Fig. 10c, then the data from both panels is combined, and a regression over the combined data is reported as the best overall description of the data in the panel. Therefore, this multiple regression analysis yields three possible outcomes, which by chance are illustrated in the first three panels of Fig. 10: (1) the slopes significantly differ (Fig. 10a), (2) the slopes do not differ but the intercepts significantly differ (Fig. 10b), or (3) neither the slopes nor the intercepts significantly differ (Fig. 10c).

In each case, the panel also indicates two measures of effect size: (1) the overall  $R^2$  value, the percentage of variation in the panel explained by the linear regressions, and (2)  $dR^2$ , the percentage of variation explained by the change in the categorical variable. If  $dR^2$  is too small, hypothesis testing is not performed, because any statistical differences would be too small to be meaningful (Pedhazur [47]). Based on the results reported in this paper, hypothesis testing is only conducted when  $dR^2 \geq 0.08\%$ . Finally, for each panel, if there is a statistical difference in either slope or intercept, then the distance,  $d$ , in cm, between the fitted regression lines is also reported. When there is a difference in slope, as in Fig. 10a,  $d$  is reported at the minimum and maximum  $x$  values (33.3 and 50 cm). When there is a difference in intercept, as in Fig. 10b, then the regression lines are the same distance apart for every  $x$ , and only one  $d$  value is reported.

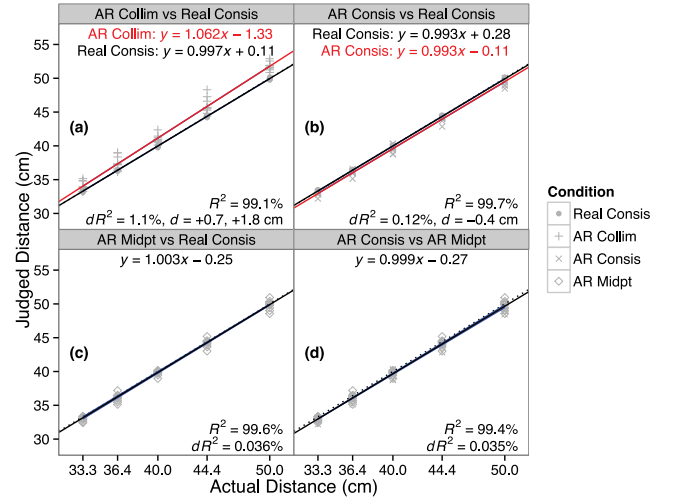


Fig. 10. Experiment I, multiple regression analysis, plotted as a scatter-plot of judged against actual distance, with  $N = 200$  ghosted data points. The thin dashed lines represent veridical performance. Blue lines represent fitted regression equations from Fig. 8. Black and red lines represent the linear regressions shown in each panel. Blue lines are not visible when overlaid by black or red lines; the degree of blue line visibility is a graphical indication of how closely the regressions in each panel agree with the regressions from Fig. 8.

$d$  is a signed value;  $d > 0$  indicates a distance farther from the observer, and  $d < 0$  closer to the observer.

### 4.3 Results

*Real Consistent Very Accurate.* Fig. 9a indicates that observers were extremely accurate in the real consistent condition. The mean error is  $-0.02$  cm, and the slope of the regression for Fig. 9a,  $y = -0.003x + 0.11$ , does not significantly differ from 0 ( $F_{1,48} = 1.63, p = 0.21$ ). Note that this is statistically equivalent to testing whether the slope in Fig. 8a differs from 1.

*AR Collimated Increasingly Overestimated.* When AR collimated is compared to real consistent (Fig. 10a), the slopes significantly differ ( $F_{1,96} = 10.7, p = 0.001$ ), indicating that the AR collimated targets were overestimated, from  $+0.7$  to  $+1.8$  cm (Fig. 9b).

*AR Consistent Underestimated.* When AR consistent is compared to real consistent (Fig. 10b), the slopes do not significantly differ ( $F_{1,96} = 0.68, p = 0.41$ ), but the intercepts do ( $F_{1,97} = 41.3, p < 0.001$ ), indicating that the AR consistent targets were underestimated by a constant  $-0.4$  cm (Fig. 9c).

*AR Midpoint Equivalent to Real Consistent.* When AR midpoint is compared to real consistent (Fig. 10c), the effect size of the difference is 0.036 percent of the variation, which is too small for any statistical differences to be meaningful. Therefore, the joint data is best fit by a single equation, indicating that AR midpoint targets were accurately matched (Fig. 9d).

*AR consistent and AR midpoint equivalent:* When AR consistent is compared to AR midpoint (Fig. 10d), the effect size is 0.035 percent, also too small for any statistical differences to be meaningful. Therefore, matches of AR consistent and AR midpoint targets were equivalent (Figs. 9c, 9d).

### 4.4 Discussion

The first purpose (1) of Experiment I was to replicate the real consistent and AR collimated conditions of Swan et al. [6] (Fig. 1). The pattern in Figs. 9a, b indeed matches Fig. 1.



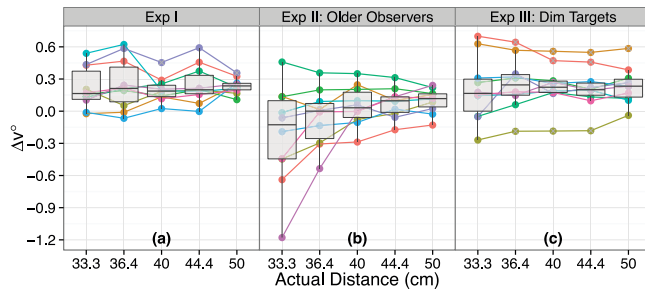


Fig. 11. For the *AR collimated* condition, the change in vergence angle  $\Delta v = \alpha - \beta$  (Fig. 2), when an observer has matched the depth of the virtual target  $\beta$  with the real pointer (Fig. 7). Each line in each panel is a different observer. For all  $N = 30$  observers,  $\Delta v$  is approximately constant across all tested distances. The boxplot gives the value for the median observer.

Given the many differences between the AR haploscope and the NVIS display used by Swan et al. [6], this replication is consistent with the idea that this pattern of results generalizes to any collimated AR or stereo display. In addition, Swan et al. [6] hypothesized that collimation biases the eyes' vergence angle to rotate outwards by a constant amount (Fig. 2). For each distance, Fig. 11a shows  $\Delta v$ , the change in vergence angle,<sup>6</sup> for the 10 AR collimated observers. For all observers  $\Delta v$  changes less than  $0.5^\circ$ , and the median observer, seen in the boxplot, changes less than  $0.072^\circ$ . These small angular changes are consistent with the hypothesis that, within these reaching distances, the vergence angle bias is constant.

The next purpose (2) was to test whether presenting AR objects at a focal distance that was *consistent* with the distance specified by other depth cues, especially binocular disparity, would result in more accurate depth matches than what was seen in the AR collimated condition. Figs. 9a, b, and c, as well as the analysis in Figs. 10a and b, confirm this hypothesis: AR consistent is much more accurate than AR collimated, and for a consistent focal distance, real and AR targets do not differ in slope (Fig. 10b).

The final purpose (3) was to test whether presenting AR objects at a focal distance equal to the midpoint of the tested range would result in similar performance as the consistent condition. Figs. 9c and d, as well as the analysis in Figs. 10c and d, indicate that, when the focus was set to the midpoint, matching was indeed just as accurate.

## 5 EXPERIMENT II: AGE

As discussed in Section 2.3, increasing age leads to presbyopia. While Experiment I found significant negative effects of collimation, all of the observers were young, with a mean age of 20.9, and therefore likely not presbyopic. In addition, also discussed in Section 2.3, although older people are worse than younger people at many perceptual tasks, recent studies have found that older people preserve their abilities in many tasks related to distance perception. Therefore, it was unclear if older observers would replicate the effects

6.  $\Delta v = \alpha - \beta$ ,  $\alpha = 2 \arctan(i/2x)$ , and  $\beta = 2 \arctan(i/2y)$ , where  $i$  is the observer's interpupillary distance,  $x$  is the actual target distance, and  $y$  is the judged distance (Fig. 8). Note that using  $x$  assumes that observers would match a real object with perfect accuracy, but the very accurate and precise results for the real consistent condition suggest this assumption is reasonable.

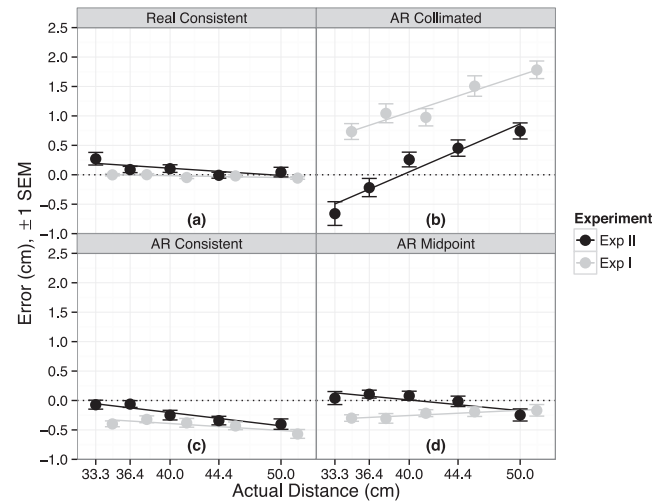


Fig. 12. Experiment II: older observers, plotted as mean error against actual distance ( $N = 1200$ ). For comparison, Experiment I's results are also shown in light grey, offset along the  $x$ -axis for clarity.

observed in Experiment I (Fig. 9). Furthermore, this work was primarily inspired by medical AR applications, and the majority of medical professionals are old enough to suffer some degree of presbyopia. Therefore, the purpose of Experiment II was to replicate Experiment I, using presbyopic observers, aged 40 and older.

### 5.1 Method

Other than the age of the observers, the methods of Experiment II were identical to Experiment I. 40 observers were recruited from a population of university and community members. The observers ranged in age from 41 to 80; the mean age was 55.6, and 19 were male and 21 female. 6 observers were paid \$10 an hour, 33 were paid \$12 hour, and one was not paid. Each observer completed 5 (distance)  $\times$  6 (repetition) = 30 trials, and the experiment collected a total of 40 (observers)  $\times$  30 (trials) = 1200 datapoints.

### 5.2 Results

Fig. 8e–h shows the results from Experiment II as scatterplots; the  $r^2$  values indicate that the data continues to be very well fit by regression equations. Fig. 12 shows the results as error, with Experiment I's results also shown for comparison. Figs. 13 and 14 show the results of multiple regression analysis.

*Older and Younger Only Differ in AR Collimated.* Fig. 13 compares Experiment I to Experiment II condition by condition. For the AR collimated condition (Fig. 13b), the slopes do not significantly differ ( $F_{1,96} = 0.38, p = 0.54$ ), but the intercepts do ( $F_{1,97} = 35.7, p < 0.001$ ); the older observers matched collimated AR targets a constant  $-1.1$  cm closer to themselves than the younger observers (Fig. 12b). For the remaining conditions, the effect size of the difference between Experiments I and II, 0.013 percent (Fig. 13a), 0.031 percent (Fig. 13c), and 0.056 percent (Fig. 13d), is too small for any statistical differences to be meaningful. Therefore, for the real consistent, AR consistent, and AR midpoint conditions, the results for the older observers and the younger observers are equivalent.

*Real Consistent Very Accurate.* Fig. 12a indicates that older observers were very accurate when matching the distance

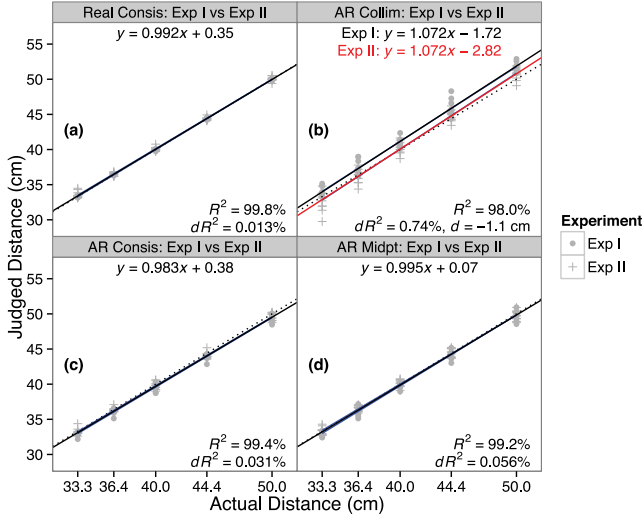


Fig. 13. Experiment I versus II, the effect of age, multiple regression analysis, with  $N = 400$  ghosted data points. See the caption for Fig. 10.

of real targets. The mean error is  $+0.04$  cm, and the slope of the linear model for Fig. 12a,  $y = -0.012x + 0.60$ , does not significantly differ from 0 ( $F_{1,48} = 2.5, p = 0.12$ ). Note that this is statistically equivalent to testing whether the slope in Fig. 8e differs from 1.

*AR Collimated Increasingly Overestimated.* For the older observers, when AR collimated is compared to real consistent (Fig. 14a), the slopes significantly differ ( $F_{1,96} = 13.8, p < 0.001$ ); the AR collimated errors ranged from  $-0.7$  to  $+0.9$  cm (Fig. 12b).

*AR Consistent Underestimated.* For the older observers, when AR consistent is compared to real consistent (Fig. 14b), the slopes do not significantly differ ( $F_{1,96} = 0.56, p = 0.46$ ), but the intercepts do ( $F_{1,97} = 16.6, p < 0.001$ ); the AR consistent targets were underestimated by a constant  $-0.3$  cm (Fig. 12c).

*AR Midpoint Equivalent to Real Consistent.* For the older observers, when AR midpoint is compared to real consistent (Fig. 14c), the effect size is 0.0093 percent, which is too small for any statistical differences to be meaningful. Therefore, the AR midpoint targets were accurately matched (Fig. 12d).

*AR Consistent and AR Midpoint Equivalent.* For the older observers, when AR consistent is compared to AR midpoint (Fig. 14d), the effect size is 0.036 percent, also too small for any statistical differences to be meaningful. Therefore, the matches of the AR consistent and AR midpoint targets were equivalent (Fig. 12c, d).

### 5.3 Discussion

The purpose of Experiment II was to replicate Experiment I, using older, presbyopic observers. Although near focus was not measured in either experiment, according to Duane [31], the younger observers in Experiment I had an expected near focus of  $\sim 8.3$  cm ( $\sim 11.5$  D), while for these older observers the expected near focus was  $\sim 68$  cm ( $\sim 1.5$  D).

Experiment II's results only differ for the AR collimated condition. For collimated targets, older observers showed less overestimation than younger observers, with matches shifted towards the observer by a constant  $-1.1$  cm (Fig. 13b). Older observers had a mean error of  $+0.12$  cm, while younger observers had a mean error of  $+1.2$  cm (Fig. 12b), and

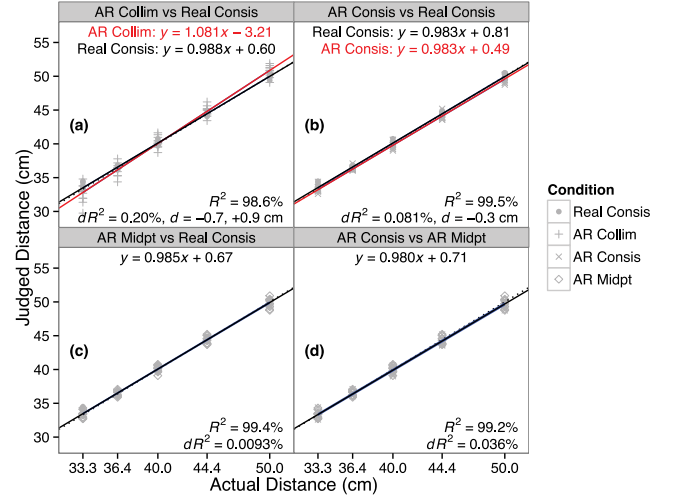


Fig. 14. Experiment II, multiple regression analysis, older observers, with  $N = 200$  ghosted data points. See the caption for Fig. 10.

therefore older observers were on average *more* accurate than younger observers. However, the slope,  $b = 1.072$ , is the same for both sets of observers (Fig. 13b), and differs significantly from the slope for the real consistent condition (Figs. 10a and 14a). Therefore, for both younger and older observers, matches of collimated targets were inaccurate, and increasingly overestimated with increasing distance. In addition, for each distance, Fig. 11b shows  $\Delta v$ , the change in vergence angle for the 10 older AR collimated observers. For 9 of the 10 observers  $\Delta v$  changes less than  $0.6^\circ$ , for the outlying observer it changes  $1.4^\circ$ , and the median observer changes less than  $0.25^\circ$ . These small angular changes are consistent with the hypothesis that, for both younger and older observers, the vergence angle bias is constant.

For the other conditions, the observer's age—and therefore the observers' ability to accommodate to different focal demands—made no difference. Older observers were just as accurate as younger observers in matching the distance to real targets, as well as to AR targets with both consistent and midpoint focal cues. These results are consistent with previous work that has found that older observers preserve their abilities in many tasks related to distance perception (Bian and Andersen [33]).

## 6 EXPERIMENT III: BRIGHTNESS

However, a conflicting finding from both experiments is that AR consistent was underestimated, while AR midpoint was accurate and AR consistent was equivalent to AR midpoint. This is true for both Experiment I (Fig. 10) and Experiment II (Fig. 14). Although it is not known why AR midpoint is more accurate, it may be related to the complexity of calibrating for the AR consistent condition, where each accommodation lens must be calibrated separately (Section 4.1.1). Nevertheless, the AR consistent underestimation is statistically significant, and was replicated by both younger and older observers.

As discussed in Section 2.5, brighter objects appear closer than similar-sized dimmer objects. Figs. 15a and b show photographs, taken through the haploscope optics, of the real and AR targets used in Experiments I and II. The AR target

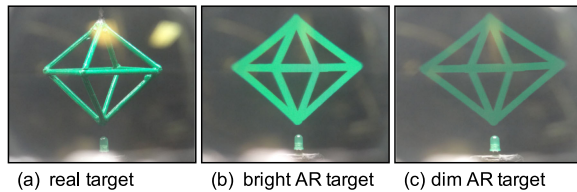


Fig. 15. Experiment III examined target brightness. (a) The real target. (b) The bright AR target used in Experiments I and II. (c) The dim AR target used in Experiment III.

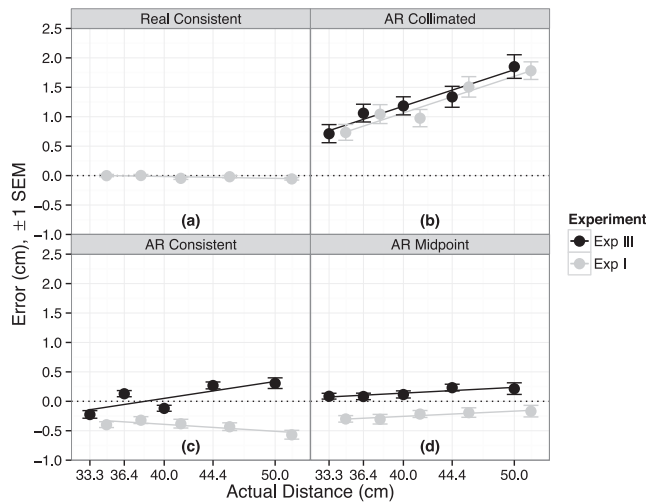


Fig. 16. Experiment III: dim targets, plotted as mean error against actual distance ( $N = 900$ ). For comparison, Experiment I's results are also shown in light grey, offset along the  $x$ -axis for clarity.

appeared brighter than the real target.<sup>7</sup> For Experiment III, the brightness was reduced, until the AR and real targets appeared to have the same brightness (Fig. 15c). The purpose of Experiment III was to determine if the dim AR target would increase the accuracy of the AR consistent condition.

## 6.1 Method

Other than the brightness of the AR target, the methods of Experiment III were identical to Experiment I. Because the real target object did not change, that condition was not replicated. To facilitate comparison with Experiment I, younger observers were recruited, from a population of university students and staff. The 30 observers ranged in age from 17 to 24; the mean age was 19.8, and 21 were male and 9 female. 6 observers were paid \$12 an hour, and the rest received course credit. Each observer completed 5 (distance)  $\times$  6 (repetition) = 30 trials, and the experiment collected a total of 30 (observers)  $\times$  30 (trials) = 900 datapoints.

## 6.2 Results

Fig. 8j-l shows the results from Experiment III as scatter-plots; the  $r^2$  values indicate that the data continues to be very well fit by regression equations. Fig. 16 shows the same results as error, with Experiment I's results also shown for comparison. Figs. 17 and 18 show the results of multiple regression analysis.

7. Note that brightness is the perceptual experience of luminance, and cannot be directly measured or captured with a camera. The luminance of the targets was measured (Singh [8]).

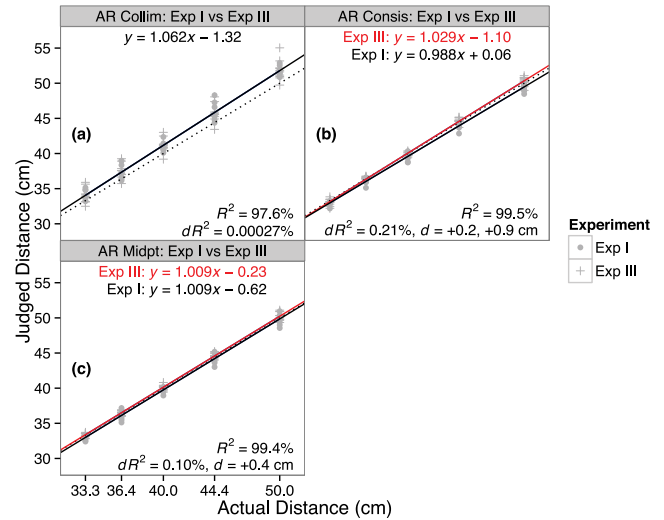


Fig. 17. Experiment I versus III, the effect of brightness, multiple regression analysis, with  $N = 300$  ghosted data points. See the caption for Fig. 10.

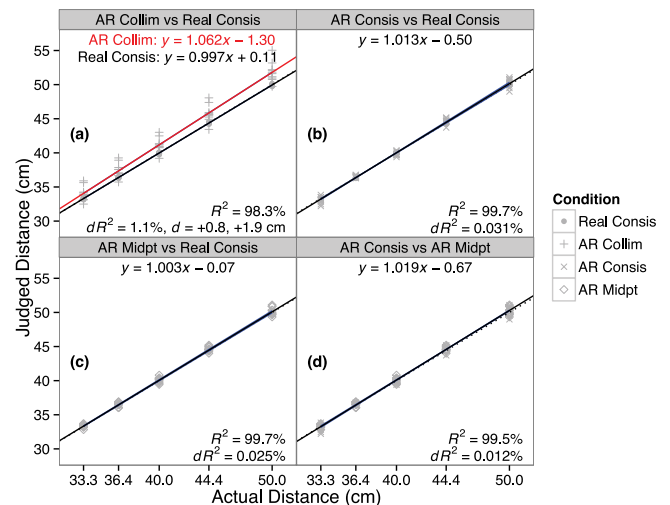


Fig. 18. Experiment III, multiple regression analysis, dim targets, with  $N = 200$  ghosted data points. See the caption for Fig. 10. The real consistent data is repeated from Experiment I.

*Dim targets Differ in AR Consistent and AR Midpoint.* Fig. 17 compares Experiment I to Experiment III condition by condition. For the AR collimated condition (Fig. 17a), the effect size of the difference is 0.00027 percent, much too small for any statistical differences to be meaningful. Therefore, the results for the dim targets and the bright targets are equivalent (Fig. 16b). For the AR consistent condition (Fig. 17b), the slopes significantly differ ( $F_{1,96} = 7.7, p = 0.007$ ), and therefore the dim targets were matched +0.2 to +0.9 cm farther than the bright targets (Fig. 16c). And finally, for the AR midpoint condition (Fig. 17c), the slopes do not significantly differ ( $F_{1,96} < 0.01, p = 0.96$ ), but the intercepts do ( $F_{1,97} = 17.7, p < 0.001$ ), and therefore the dim targets were matched +0.4 cm farther than the bright targets (Fig. 16d).

*AR Collimated Increasingly Overestimated.* When dim AR collimated is compared to real consistent from Experiment I (Fig. 18a), the slopes significantly differ ( $F_{1,96} = 5.5, p = 0.021$ ), indicating that the dim AR collimated targets were overestimated from +0.8 to +1.9 cm (Fig. 16b).



*AR Consistent Equivalent to Real Consistent.* When dim AR consistent is compared to real consistent from Experiment I (Fig. 18b), the effect size of the difference is 0.031 percent, which is too small for any statistical differences to be meaningful. Therefore, the dim AR consistent targets were accurately matched (Fig. 16c).

*AR Midpoint Equivalent to Real Consistent.* When dim AR midpoint is compared to real consistent from Experiment I (Fig. 18c), the effect size of the difference is 0.025 percent, which is too small for any statistical differences to be meaningful. Therefore, the dim AR midpoint targets were accurately matched (Fig. 16d).

*AR Consistent and AR Midpoint Equivalent.* When dim AR consistent is compared to dim AR midpoint (Fig. 18d), the effect size is 0.012 percent, also too small for any statistical differences to be meaningful. Therefore, the matches of dim AR consistent and dim AR midpoint targets were equivalent (Figs. 16c, 16d).

### 6.3 Discussion

The purpose of Experiment III was to determine if a dim AR target, which has the same apparent brightness as the real target (Fig. 15), would increase the accuracy of the AR consistent condition. When comparing Experiment III to Experiment I (Fig. 17), there was no difference in matching AR collimated targets, but increased accuracy for both AR consistent and AR midpoint targets. In addition, when combined with the real consistent data from Experiment I (Figs. 18), matches for the dim AR target were overestimated in the AR collimated condition, but accurate in both the AR consistent and AR midpoint conditions. Therefore (Fig. 16), for the AR consistent and AR midpoint conditions, the dim AR targets were as accurately matched as the real targets.

In addition, the pattern of AR consistent being underestimated, while AR midpoint was accurate and AR consistent was equivalent to AR midpoint, occurred for both the younger observers of Experiment I (Fig. 10) and the older observers of Experiment II (Fig. 14). While Experiment III only tested younger observers, the results are consistent with the hypothesis that older observers would also accurately match the depth of dim AR consistent and dim AR midpoint targets.

Finally, for the dim AR collimated targets, for each distance Fig. 11c shows  $\Delta v$ , the change in vergence angle, for the 10 AR collimated observers. For all observers  $\Delta v$  changes less than  $0.4^\circ$ , and the median observer changes less than  $0.12^\circ$ . These small angular changes are consistent with the hypothesis that, for the dim AR targets, the vergence angle bias is constant.

## 7 GENERAL DISCUSSION

*Constant Vergence Angle Bias.* As discussed in Section 1, Swan et al. [6] proposed that collimated, near-field AR targets bias the eyes' vergence angle to rotate outwards by a constant amount. All of the experiments reported here replicated this result, and strongly support this hypothesis. In addition, these findings are consistent with the prediction, by Mon-Williams and Tresilian [19], that an inconsistent accommodative cue biases perceived depth in the same direction as the accommodative cue (Fig. 3). However, Swan et al. [6] did not measure this change in vergence

angle, and it was not measured here. In a future experiment, the change in vergence angle should be directly measured.

*Brightness.* The experiments found the most accurate matches for dim AR targets, which more closely matched the brightness of real targets. These results are consistent with previous work that finds brighter objects appear closer than dimmer objects (Ashley [38], Farnè [39], Coules [40]). However, it is interesting that the experiments replicated this effect in a *between-subjects* design: different groups of 10 observers saw the real targets, the bright AR targets, and the dim AR targets. It would have been less surprising to have replicated the effect in a *within-subjects* design, where observers made a judgment about two targets with different brightnesses, viewed simultaneously (e.g., Ashley [38], Farnè [39], Coules [40]). Like collimation, brightness could also bias vergence angle, and therefore a future experiment should examine whether the brightness of an AR object directly influences vergence angle. These experiments suggest that brightness is another component of accurate depth presentation that must be considered by AR practitioners.

*Midpoint Focus.* While the AR consistent condition was accurate for the dim AR targets, the AR midpoint condition was accurate across all of the experiments. Although the changing accuracy of the AR consistent condition could be related to calibration errors, these experiments do not settle this question. Nevertheless, the practical implication is that, because the AR midpoint condition was at least as accurate as the AR consistent condition, for AR applications requiring accurate near-field depth matching, it is sufficient to set the focus to the middle of the working volume.

The positive results for the AR midpoint condition also suggest comparison with light-field displays, which can simultaneously present multiple virtual objects at different focal distances. In addition to solving the vergence-accommodation conflict, light-field displays are predicted to eventually become the dominant technology for all kinds of 3D experiences (Balram [49]). However, although the technology is rapidly developing, AR light-field displays face many fundamental challenges and design tradeoffs, in areas such as depth range, color resolution, spatial resolution, computational demands, and data throughput requirements (Wu et al. [50]). Therefore, the AR midpoint results suggest that the engineering complexity of light-field displays may not always be necessary, especially for AR applications where the most important perceptual task is matching near-field distances (e.g., Edwards [4], Krempien et al. [5]).

*Future Work.* The errors detected in these experiments are likely explained by vergence angle biases. Therefore, useful future work would replicate the experiments while measuring vergence angle. Possible measurement methods include binocular eye tracking (Wang et al. [51]) and nonius lines (Ellis and Menges [41]).

In addition, because the AR haploscope was mounted to a tabletop, these experiments could not examine the depth cue of motion perspective. Although some AR applications, such as the operating microscope described by Edwards et al. [4], are also mounted and therefore lack motion perspective, it is a very salient depth cue (Nagata [10], Cutting and Vishton [9]), and should be examined in future experiments. A head-mounted AR haploscope, such

as the one used by McCandless et al. [42], would allow a replication of these experiments to include motion perspective.

**Practical Implications.** For accurate near-field depth matching, the experiments reported here have the following implications:

- (1) Collimated graphics should not be used.
- (2) A focal distance set to the middle of the depth range is as good as a focal distance optimized for every virtual object.
- (3) The brightness of virtual objects needs to match the brightness of real objects.
- (4) Observers old enough to suffer age-related reductions in accommodative ability are just as accurate as younger observers.

## ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation, under awards IIS-0713609, IIS-1018413, and IIS-1320909, to J. E. Swan II. This work was conducted at the Institute for Neurocognitive Science and Technology, and the Center for Advanced Vehicular Systems, at Mississippi State University.

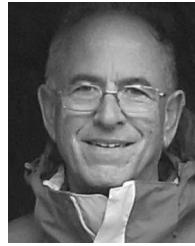
## REFERENCES

- [1] M. Kersten-Oertel, P. Jannin, and D. L. Collins, "The state of the art of visualization in mixed reality image guided surgery," *Comp. Med. Imag. Graph.*, vol. 37, no. 2, pp. 98–112, 2013.
- [2] T. P. Caudell and D. W. Mizell, "Augmented reality: An application of heads-up display technology to manual manufacturing processes," in *Proc. Hawaii Int. Conf. Syst. Sci.*, 1992, pp. 659–669.
- [3] S. J. Henderson and S. Feiner, "Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret," in *Proc. IEEE Int. Symp. Mixed Augmented Reality*, 2009, pp. 135–144.
- [4] P. J. Edwards, A. P. King, C. R. Maurer Jr, D. A. de Cunha, D. P. Caudell, D. L. G. Hill, R. P. Gaston, M. R. Fenlon, A. Jusczyck, A. J. Strong, C. L. Chandler, and M. J. Gleeson, "Design and evaluation of a system for microscope-assisted guided interventions (MAGI)," *IEEE Trans. Med. Imag.*, vol. 19, no. 11, pp. 1082–1093, Nov. 2000.
- [5] R. Krempien, H. Hoppe, L. Kahrs, S. Daeuber, O. Schorr, G. Eggers, M. Bischof, M. W. Munter, J. Debus, and W. Harms, "Projector-based augmented reality for intuitive intraoperative guidance in image-guided 3D interstitial brachytherapy," *Int. J. Radiation Oncology Biol. Phys.*, vol. 70, no. 3, pp. 944–952, 2008.
- [6] J. E. Swan II, G. Singh, and S. R. Ellis, "Matching and reaching depth judgments with real and augmented reality targets," *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 11, pp. 1289–1298, Nov. 2015.
- [7] J. L. Gabbard, D. G. Mehra, and J. E. Swan II, "Effects of AR display context switching and focal distance switching on human performance," *IEEE Trans. Vis. Comput. Graph.*, 2018, doi: 10.1109/TVCG.2017.2754257.
- [8] G. Singh, "Near-field depth perception in optical see-through augmented reality," Ph.D. dissertation, Dept. Comput. Sci. Eng., Mississippi State Univ., Mississippi State, MS, USA, Aug. 2013.
- [9] J. E. Cutting and P. M. Vishton, "Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth," in *Handbook of Perception and Cognition: Perception of Space and Motion*, vol. 5, W. Epstein and S. Rogers, Eds. New York, NY, USA: Academic Press, 1995, pp. 69–117.
- [10] S. Nagata, "How to reinforce perception of depth in single two-dimensional pictures," in *Pictorial Communication in Virtual and Real Environments*, S. R. Ellis, M. K. Kaiser, and A. J. Grunwald, Eds. New York, NY, USA: Taylor & Francis, 1991, pp. 527–545.
- [11] B. Gillam, "The perception of space layout," in *Perception of Space and Motion*, 2nd ed., W. Epstein and S. Rogers, Eds. New York, NY, USA: Academic Press, 1995, pp. 23–67.
- [12] G. P. Bingham and C. C. Pagano, "The necessity of a perception-action approach to definite distance perception: Monocular distance perception to guide reaching," *J. Exp. Psychology: Hum. Perception Perform.*, vol. 24, no. 1, pp. 145–168, 1998.
- [13] M. Mon-Williams and J. R. Tresilian, "Some recent studies on the extraretinal contribution to distance perception," *Perception*, vol. 28, no. 2, pp. 167–181, 1999.
- [14] M. S. Landy, L. T. Maloney, E. B. Johnston, and M. Young, "Measurement and modeling of depth cue combination: In defense of weak fusion," *Vis. Res.*, vol. 35, no. 3, pp. 389–412, 1995.
- [15] R. J. Leigh and D. S. Zee, *The Neurology of Eye Movements*, 5th ed. Oxford, United Kingdom: Oxford Univ. Press, 2015.
- [16] J. L. Semmlow and G. K. Hung, "The near response: Theories of control," in *Vergence Eye Movements: Basic & Clinical Aspects*, C. M. Schor and K. J. Ciuffreda, Eds. London, United Kingdom: Butterworth, 1983, pp. 175–195.
- [17] H. Ripps, N. B. Chin, I. M. Siegel, and G. M. Breinin, "The effect of pupil size on accommodation, convergence, and the AC/A ratio," *Invest. Ophthalmology Visual Sci.*, vol. 1, no. 1, pp. 127–135, 1962.
- [18] D. Kersten and G. E. Legge, "Convergence accommodation," *J. Opt. Soc. America*, vol. 73, no. 3, pp. 332–338, 1983.
- [19] M. Mon-Williams and J. R. Tresilian, "Ordinal depth information from accommodation?," *Ergonom.*, vol. 43, no. 3, pp. 391–404, 2000.
- [20] E. Kruijff, J. E. Swan II, and S. Feiner, "Perceptual issues in augmented reality revisited," in *Proc. IEEE Int. Symp. Mixed Augmented Reality*, 2010, pp. 3–12.
- [21] M. Lambooi, M. Fortuin, I. Heynderickx, and W. IJsselstein, "Visual discomfort and visual fatigue of stereoscopic displays: A review," *J. Imag. Sci. Technol.*, vol. 53, no. 3, pp. 030201–030201–14, 2009.
- [22] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *J. Vis.*, vol. 8(3), no. 33, pp. 1–30, 2008.
- [23] H. A. Swenson, "The relative influence of accommodation and convergence in the judgment of distance," *J. Gen. Psychology*, vol. 7, no. 2, pp. 360–380, 1932.
- [24] E. Brenner and J. M. Van Damme, "Judging distance from ocular convergence," *Vis. Res.*, vol. 38, no. 4, pp. 493–498, 1998.
- [25] D. A. Owens and H. W. Liebowitz, "Accommodation, convergence, and distance perception in low illumination," *Optometry Vis. Sci.*, vol. 57, no. 9, pp. 540–550, 1980.
- [26] J. R. Tresilian, M. Mon-Williams, and B. M. Kelly, "Increasing confidence in vergence as a cue to distance," *Proc. Roy. Soc. London B*, vol. 266, pp. 39–44, 1999.
- [27] A. Viguier, G. Clement, and Y. Trotter, "Distance perception within near visual space," *Perception*, vol. 30, pp. 115–124, 2001.
- [28] J. M. Foley, "Binocular distance perception," *Psychological Rev.*, vol. 87, no. 5, pp. 411–434, 1980.
- [29] J. H. Iavecchia, H. P. Iavecchia, and S. N. Roscoe, "Eye accommodation to head-up virtual images," *Hum. Factors*, vol. 30, no. 6, pp. 703–712, 1988.
- [30] S. N. Roscoe, "Bigness is in the eye of the beholder," *Hum. Factors*, vol. 27, no. 6, pp. 615–636, 1985.
- [31] A. Duane, "Normal values of the accommodation at all ages," *J. Amer. Med. Assoc.*, vol. LIX, pp. 1010–1013, 1912.
- [32] S. Kasthurirangan and A. Glasser, "Age related changes in accommodative dynamics in humans," *Vis. Res.*, vol. 46, pp. 1507–1519, 2006.
- [33] Z. Bian and G. J. Andersen, "Aging and the perception of egocentric distance," *Psychology Aging*, vol. 28, no. 3, pp. 813–825, 2013.
- [34] G. Heron, W. N. Charman, and C. M. Schor, "Age changes in the interactions between the accommodation and vergence systems," *Optometry Vis. Sci.*, vol. 78, no. 10, pp. 754–762, 2001.
- [35] G. J. Andersen, A. Saidpour, and M. L. Braundstein, "Effects of collimation on perceived layout in 3-D scenes," *Perception*, vol. 27, no. 11, pp. 1305–1315, 1998.
- [36] S. J. Watt, K. Akeley, M. O. Ernst, and M. S. Banks, "Focus cues affect perceived depth," *J. Vis.*, vol. 5, no. 10, pp. 834–862, 2005.
- [37] E. McCurdy, *The Notebooks of Leonardo da Vinci, Volume II*. New York, NY, USA: Reynal & Hitchcock, 1938.
- [38] M. L. Ashley, "Concerning the significance of intensity of light in visual estimates of depth," *Psychological Rev.*, vol. 5, no. 6, pp. 595–615, 1898.
- [39] M. Farné, "Brightness as an indicator to distance: Relative brightness per se or contrast with the background?" *Perception*, vol. 6, pp. 287–293, 1977.
- [40] J. Coules, "Effect of photometric brightness on judgments of distance," *Exp. Psychology*, vol. 50, no. 1, pp. 19–25, 1955.

- [41] S. R. Ellis and B. M. Menges, "Localization of virtual objects in the near visual field," *Hum. Factors*, vol. 40, no. 3, pp. 415–431, 1998.
- [42] J. W. McCandless, S. R. Ellis, and B. D. Adelstein, "Localization of a time-delayed, monocular virtual object superimposed on a real environment," *Presence: Teleoperators Virtual Environments*, vol. 9, no. 1, pp. 15–24, 2000.
- [43] G. Singh, J. E. Swan II, J. A. Jones, and S. R. Ellis, "Depth judgment measures and occluding surfaces in near-field augmented reality," in *Proc. ACM Symp. Appl. Perception Graph. Vis.*, 2010, pp. 149–156.
- [44] N. Rosa, W. Hürst, P. J. Werkhoven, and R. C. Velthkamp, "Visuotactile integration for depth perception in augmented reality," in *Proc. ACM Int. Conf. Multimodal Interaction*, 2016, pp. 45–52.
- [45] J. P. Rolland, W. Gibson, and D. Ariely, "Towards quantifying depth and size perception in virtual environments," *Presence: Teleoperators Virtual Environments*, vol. 4, no. 3, pp. 24–49, 1995.
- [46] S. Lee, X. Hu, and H. Hua, "Effects of optical combiner and IPD change for convergence on near-field depth perception in an optical see-through HMD," *IEEE Trans. Vis. Comput. Graph.*, vol. 22, no. 5, pp. 1540–1554, May 2016.
- [47] E. J. Pedhazur, *Multiple Regression in Behavioral Research*, 2nd ed. New York, NY, USA: Holt, Rinehart and Winston, 1982.
- [48] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, *Applied Multiple Regression / Correlation Analysis for the Behavioral Sciences*, 3rd ed., Mahwah NJ, USA: Lawrence Erlbaum Associates, 2003.
- [49] N. Balram, "The next wave of 3-D-light-field displays," *Inf. Display*, vol. 30, no. 6, 2014, Art. no. 4.
- [50] W. Wu, K. Berkner, I. Tošić, and N. Balram, "Personal near-to-eye light-field displays," *Inf. Display*, vol. 30, no. 6, pp. 16–22, 2014.
- [51] R. I. Wang, B. Pelfrey, A. T. Duchowski, and D. H. House, "Online 3D gaze localization on stereoscopic displays," *ACM Trans. Appl. Perception*, vol. 11, no. 1, pp. 1–21, 2014.



**Gurjot Singh** received the MS and PhD degrees in computer science from Mississippi State University, in 2010 and 2013, respectively. He is an assistant professor of computer science with the Department of Mathematics, Computer Science, and Physics, Fairleigh Dickinson University. Before joining Fairleigh Dickinson University in 2015, he spent two years as a postdoctoral associate with the 3DI Research Group, Virginia Tech. His research interests include augmented and virtual reality, vision and perception, human-computer interaction, user studies and empirical methods, serious games, and technology-enhanced learning. He has served on conference committees for IEEE VR, ACM SAP, and IARIA VISUAL. He is an active member of ACM and the Upsilon Pi Epsilon honor society. He is a member of the IEEE.



**Stephen R. Ellis** received the AB degree in behavioral science from U.C. Berkeley, in 1969, and the PhD degree in psychology from McGill University, in 1974, and postdoctoral fellowships from Brown University and U.C. Berkeley. He has headed the Advanced Displays and Spatial Perception Laboratory at the NASA Ames Research Center. Before joining the civil service at NASA Ames, he was an NRC fellow. He has published more than 200 journal publications and reports on user interaction with spatial information, and was in the forefront of the introduction and study of perspective and 3D formats into user interfaces for aerospace applications. More recently, he was among the first researchers to use virtual environments to study human performance, most recently to study the interaction of communication delay and display-control misalignment, as can occur during teleoperation. He has served on the editorial boards of *Presence*, *Virtual Reality*, *Frontiers-Virtual Environments*, and *Human Factors*, reviewed for SIGGRAPH, IEEE VR, and HFES, and has edited a book, *Pictorial communication in virtual and real environments*, Taylor and Francis, London, 2nd Ed., 1993.



**J. Edward Swan II** received the BS degree in computer science from Auburn University, in 1989, and the MS and PhD degrees in computer science from Ohio State University, in 1992 and 1997, respectively, where he studied computer graphics and human-computer interaction. He is a professor of computer science and engineering, and faculty with the Center for Advanced Vehicular Systems, Mississippi State University. Before joining Mississippi State University in 2004, he spent seven years as a scientist with the Naval Research Laboratory in Washington, D.C. His research and scholarship has encompassed augmented and virtual reality, perception, human-computer interaction, human factors, empirical methods, volume graphics, and visualization. His research has been funded by the National Science Foundation, the Department of Defense, the National Aeronautics and Space Administration, the Naval Research Laboratory, and the Office of Naval Research. He is a senior member of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at [www.computer.org/csdl](http://www.computer.org/csdl).**