Helen H. Hu

hhh@cs.utah.edu

School of Computing
University of Utah
50 Central Campus Drive
Room 3190
Salt Lake City, UT 84112

Amy A. Gooch

gooch@cs.utah.edu

School of Computing
University of Utah
50 Central Campus Drive
Room 3190
Salt Lake City, UT 84112

Sarah H. Creem-Regehr

sarah.creem@psych.utah.edu

Department of Psychology
University of Utah
380 South 1530 East, Room 502
Salt Lake City, UT 84112

William B. Thompson

thompson@cs.utah.edu

School of Computing
University of Utah
50 Central Campus Drive
Room 3190
Salt Lake City, UT 84112

Visual Cues for Perceiving Distances from Objects to Surfaces

Abstract

An accurate perception of the distance between an object and a nearby surface can increase a viewer's sense of presence in an immersive environment, particularly when a user is performing actions that affect or are affected by this distance. Two experiments were conducted examining the effectiveness of stereoscopic viewing, shadows, and interreflections at conveying this distance information. Subjects performed simple tasks based on the perception of the distance between a fixed virtual table and an approaching block in a virtual environment. In the first experiment, subjects lowered a virtual block to a virtual table. For this task, both stereoscopic viewing and shadows had statistically significant effects on subject performance. In the second experiment, subjects mechanically reported the perceived distance between a virtual block and virtual table. For this task, viewing condition, shadows, and interreflections were shown to be statistically significant distance cues.

1   Introduction

People are usually quite accurate at visually recognizing that a moving object is about to contact a surface in the environment. This perception of imminent contact is particularly important when a person is manipulating an object or otherwise performing tasks that are affected by contact between an object and a surface. Most current immersive environments do a poor job of presenting visual information sufficient to make these judgments.

In this paper we describe two experiments exploring whether three different visual cues are effective at conveying distance between an object and a surface. An object is situated within what Cutting and Vishton (1995) defined as the user's *personal space,* a range that is within a user's arm's reach and thus within the sphere of a user's direct control and contact. This range of space is especially relevant for an individual acting within an environment because it encompasses the area where the user will interact with and manipulate objects. In our two experiments, subjects performed tasks involving perception of distances between an object and a surface (*exocentric* distances), while stereoscopic viewing, shadows, and interreflections were either provided or removed as distance cues. (See figure 1.) In the first experiment, subjects controlled the movement of the object towards the stationary surface. In the second experiment, subjects were passive observers of the movement, reporting the final distance between the object and the surface through a mechanical task.

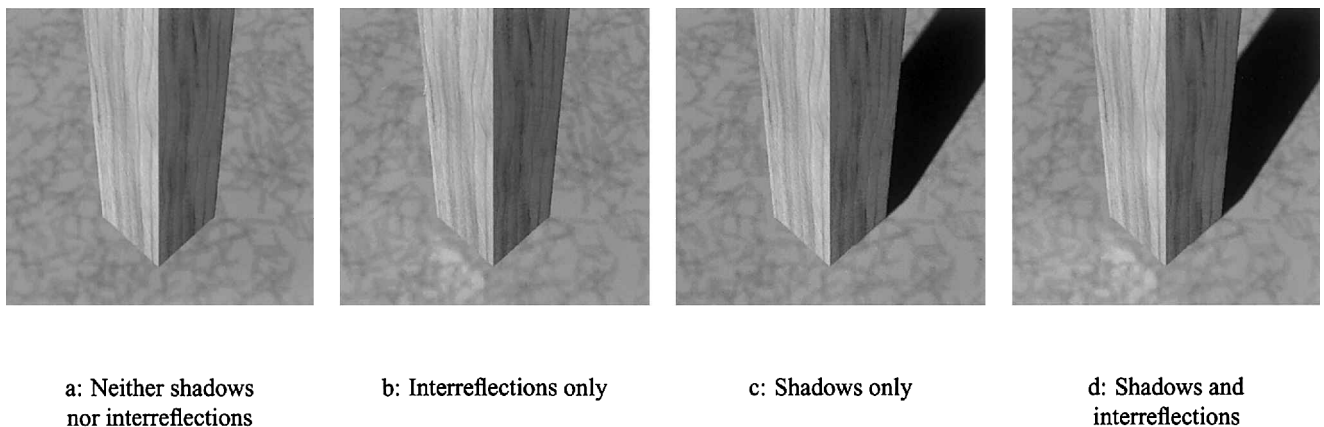It is often argued in the computer graphics and immersive environments

a: Neither shadows nor interreflections

b: Interreflections only

c: Shadows only

d: Shadows and interreflections

Figure 1. *Images similar to those used in both experiments. A virtual block is suspended 2 cm above a virtual table in all four images, but the presence and absence of shadows and interreflections are varied. Interreflections are visible here at the bottom left of images (b) and (d) and shadows are visible at the top left of images (c) and (d). In the actual experiments, the viewing condition (with or without stereoscopic viewing) was also varied, bringing the total number of conditions to eight.*

communities that stereo is important for effective distance judgments. In our experiments, we alternated between stereoscopic viewing and biocular viewing. For the biocular display, the same image was projected to both eyes. For the remainder of the paper, we use the term *stereo* to denote stereoscopic viewing.

The complexity of stereo display systems has prompted many researchers to question whether the advantages of stereo justify its implementation. For some tasks examined, stereoscopic viewing brought no improvements over monoscopic viewing in task performance (Kim, Ellis, Tyler, Hannaford, & Stark 1987; Reinhart, Beaton, & Snyder, 1990; Barfield & Rosenberg, 1995), whereas, for other tasks and situations, subjects performed better when stereo was present by learning a task more quickly (Drascic, 1991), by performing the task more quickly (Spain, 1990; Drascic, 1991; Yeh & Silverstein, 1992; Hsu, Babbs, Chelberg, Pizlo, & Delp, 1993; Ware & Franck, 1996), or by performing the task more accurately (Cole, Merritt, Fore, & Lester, 1990; Barfield & Rosenberg, 1995). The importance of stereo appears to be related to the difficulty of the task and the number of other visual depth cues available (Miller & Beaton, 1991; Brooks, Ince, & Lee, 1991), increasing in effectiveness with task difficulty and

impoverished scenes. Stereo also appears to be more effective for closer viewing distances (Surdick, Davis, King, & Hodges, 1997). Furthermore, the effectiveness of stereo sometimes differed between a static situation (more effective) and a similar dynamic situation (less effective) (Dosher, Sperling, & Wurst 1986).

Stereo is not the only visual cue that can be used in interactive systems to convey information about the distance between objects and surfaces. Shadows are also a source of such information (Yonas, Goldsmith, & Hallstrom, 1978; Kjelldahl & Prime, 1995; Kersten, Mamassian, & Knill 1997; Madison, Thompson, Kersten, Shirley, & Smits, 2001). The effectiveness of shadows varies widely between tasks (Wanger, Ferwerda, & Greenberg, 1992) and varies somewhat between subjects (Hu et al., 2000). Furthermore, the interaction of shadows with other depth cues often produces unexpected results. In some cases, shadows are strong enough to override other conflicting visual cues (Bloj, Kersten, & Hurlbert, 1999), whereas in other cases shadows actually degrade task performance in terms of both accuracy and speed when introduced to a stereoscopic display (Hubona, Wheeler, Shirah, & Brandt, 1999).

Fewer studies have examined interreflections as a

depth cue. Interreflections are the result of light reflecting from one surface to another, as seen on the table at the base of the block in figure 1(b). Although interreflections themselves are usually visually indistinct, there is evidence that people use interreflections perceptually (Kersten & Hurlbert, 1996), possibly as spatial cues (Madison et al., 2001). For still images, Madison et al. found interreflections to be as strong a perceptual cue for contact as shadows.

Modern graphics hardware has the power to interactively render reasonable quality shadows for some scenes. In limited cases, approximations to interreflections can also be rendered at interactive rates. On the other hand, stereo displays are awkward, complicate the rendering process, and have significant perceptual limitations (Wann, Rushton, & Mon-Williams, 1995). The remainder of this paper describes two experiments aimed at evaluating the utility of stereoscopic viewing, shadows, and interreflections as visual cues for the distance between an object and a surface in immersive environments.

2 General Method

We conducted two experiments. Our experimental design consisted of subjects performing an action task based on a visual stimulus rather than using verbal reports, because action tasks are more likely to provide information that is relevant to interactive immersive environments.

All subjects were first tested for stereo fusion using a random-dot stereogram before beginning the experiment. Subjects wore an nVision Datavisor HiRes head-mounted display, which has two 1280×1024 interlaced CRT displays with a $40.5°$ horizontal field of view. To prevent subjects from benefiting from motion parallax distance cues, translational head tracking was not utilized. By not activating rotational head tracking, we limited the subjects' field of view to exclude portions of the surface distant from the object of interest.

In the experiments reported here, we used a fixed interpupil distance of 6.5 cm for rendering the stereoscopic displays. Subjects' measured IPD ranged 5.5–7.0

cm, with an average of 6.35 cm. However, precise optical calibration of head-mounted displays (HMD) is quite difficult. The interpupil distance of the HMD can be only roughly controlled due to the lack of precise control over image position and accommodation in the HMDs as well as the lack of feedback from the user-adjusted HMD screen controls.

In both experiments, subjects viewed a virtual environment consisting of a block and a table. They were asked to perform a single task multiple times within each experiment, while three conditions were varied: viewing condition (stereoscopic or biocular viewing), shadows (present or absent), and interreflections (present or absent). The three independent variables allowed for eight (2^3) possible visual cue combinations, ranging from all three cues being present to all three cues being absent. To prevent subjects from using other cues to determine the distance between the virtual block and table, the table height and the direction of the light were varied randomly between trials. To maintain graphical realism at interactive rates, the tabletop images were precomputed using a standard Monte Carlo path tracer on an SGI Onyx2 R12000 and stored as a 3D texture, as described by Hu et al. (2000). Tabletop textures were adjusted to each table height so as to appear a constant texture size regardless of table height, thus preventing subjects from utilizing texture size as a distance cue.

Figure 2 demonstrates the virtual environment as viewed by a subject wearing the HMD. The same viewpoint and gaze direction were used for both experiments. The view was directed at $53.3°$ below the horizon, at a virtual table. Table height varied between trials, ranging from 46 cm to 60 cm below the subject's viewpoint. The 5×5 cm virtual block had infinite height (the top of the block was never visible) and was positioned 27.5 cm in front of the viewpoint. When stereoscopic viewing was provided, the block was placed in front of the midpoint between the subject's two viewpoints. Between trials, the table disappeared to prevent the subject from seeing the table "jump."

The experimental tasks studied in this research involve judging distances between points obliquely oriented with respect to the line of sight. Although more
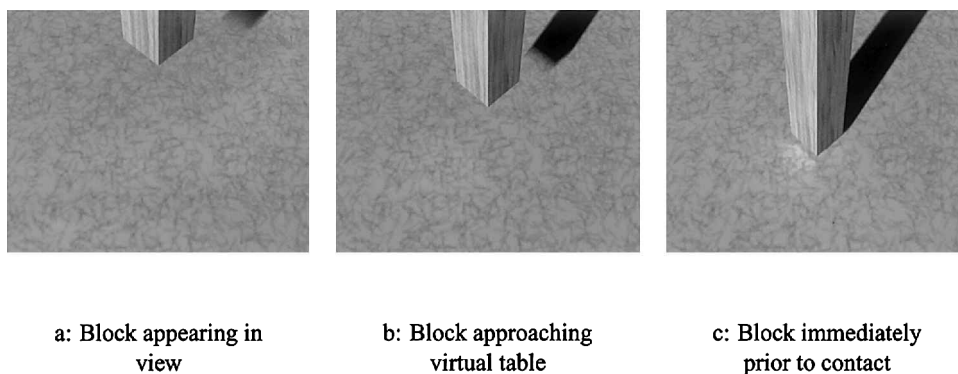
a: Block appearing in view | b: Block approaching virtual table | c: Block immediately prior to contact

Figure 2. *Several frames from the virtual environment seen by subjects wearing the head-mounted display. Each trial begins with only the virtual table in view. As the subject moves a physical block down, a virtual block appears and approaches the virtual table. Between trials, the table and block disappear from the scene.*

complex than judgments involving either distances along the line of sight or perpendicular to the line of sight, such viewing configurations are common in immersive environments involving object manipulation.

3   Experiment 1

3.1 Method

3.1.1 Participants. Six computer science graduate students (three men and three women) voluntarily participated in the experiment. All participants were tested individually. All had self-reported corrected vision without the use of prescription eyeglasses, and all were tested for stereo fusion before beginning the experiment. None knew of the hypothesis being tested, and none were authors of this paper.

3.1.2 Experiment 1: Procedure. A dynamic environment has visual cues, such as a discontinuous change in velocity, that provide information about contact but not about distance prior to contact. To focus on distance perception as opposed to event perception, subjects were instructed to bring a virtual block as close to a virtual table as possible without allowing the block to contact the table. They controlled the vertical movement of the virtual block by moving a physical block in

the real world, as seen in figure 3a. Once the experiment began, the physical block did not actually contact the physical table pictured at the base of the experimental apparatus in figure 3(a); the physical table was positioned low enough not to interfere with the experiment. (For the remainder of the paper, anytime we use the term, *the table*, we are referring to the virtual table.) The position of the physical block was measured by a SensAble Technologies PHANToM 1.5, which has a nominal spatial resolution of 0.03 mm. Measured using a timestamp method similar to that used by Liang, Shaw and Green (1991), system latency averaged 54 ms. Noise in position sensing of the object being manipulated averaged 0.50 mm with negligible drift.

After an initial training period, subjects performed the task 480 times spread through six sessions of trials. Subjects stood during each session but were allowed to rest up to five minutes between sessions. Pilot studies had demonstrated the difficulty that subjects experienced switching between stereo and biocular viewing; many subjects would lose fusion during biocular viewing. To ease subjects' eye strain and to alternate between stereo and biocular viewing without subjects' awareness, we turned the HMD off between each session of trials. Half the subjects began the experiment with stereoscopic viewing, whereas the other half began the experiment with a biocular display. Thus, within each session, two

a: Apparatus for Experiment 1   b: Apparatus for Experiment 2

Figure 3. *(a) In experiment 1, subjects controlled a virtual block by moving the physical block. They were instructed to bring the virtual block as close to a virtual table as possible without touching the virtual table. (b) In experiment 2, subjects reported the perceived distance between a virtual block and a virtual table by sliding their fingers along a scale.*

conditions—shadows (present or absent) and interreflections (present or absent)—were varied. A random order of the repeated cue combinations was presented to each subject.

For each trial, subjects had one second to bring the block down and start back up. The instructions emphasized that the motion should be natural and smooth. If the subject took too long or if the virtual block came in contact with the virtual table, the subject received negative feedback (the display became yellow or red, respectively) and the trial was discarded. This negative feedback improved the number of usable trials over the experimental design described by Hu et al., (2000). Overall, approximately 22% of the trials were discarded for exceeding the time limit or contacting the virtual table.

3.2 Experiment 1: Results

In this first experiment, virtual table height was used as an independent variable and lowest virtual cylin-

der height as a dependent variable. If a subject performed the task successfully, the block should have stopped just above the table and the bottom of the virtual block should have been slightly above the height of the virtual table surface. Figure 4 shows the data plot and regression analysis for subject 3′s data. Because we discarded every trial for which contact occurred between the block and the table, no data points appear below the $f(x) = x$ diagonal.

The slope of the regression line is an indication of the average scaling subjects applied between visually perceived positions in the virtual world and mechanically specified positions associated with manipulation of the physical block. A slope near 1.0 is an indication that subjects were able to recover the correct scale factor relating the visually and mechanically indicated positions.

To quantify the variability between visually perceived position and mechanically specified positions, we report R^2, a measure of the amount of variance accounted for in the regression analysis, or the precision with which
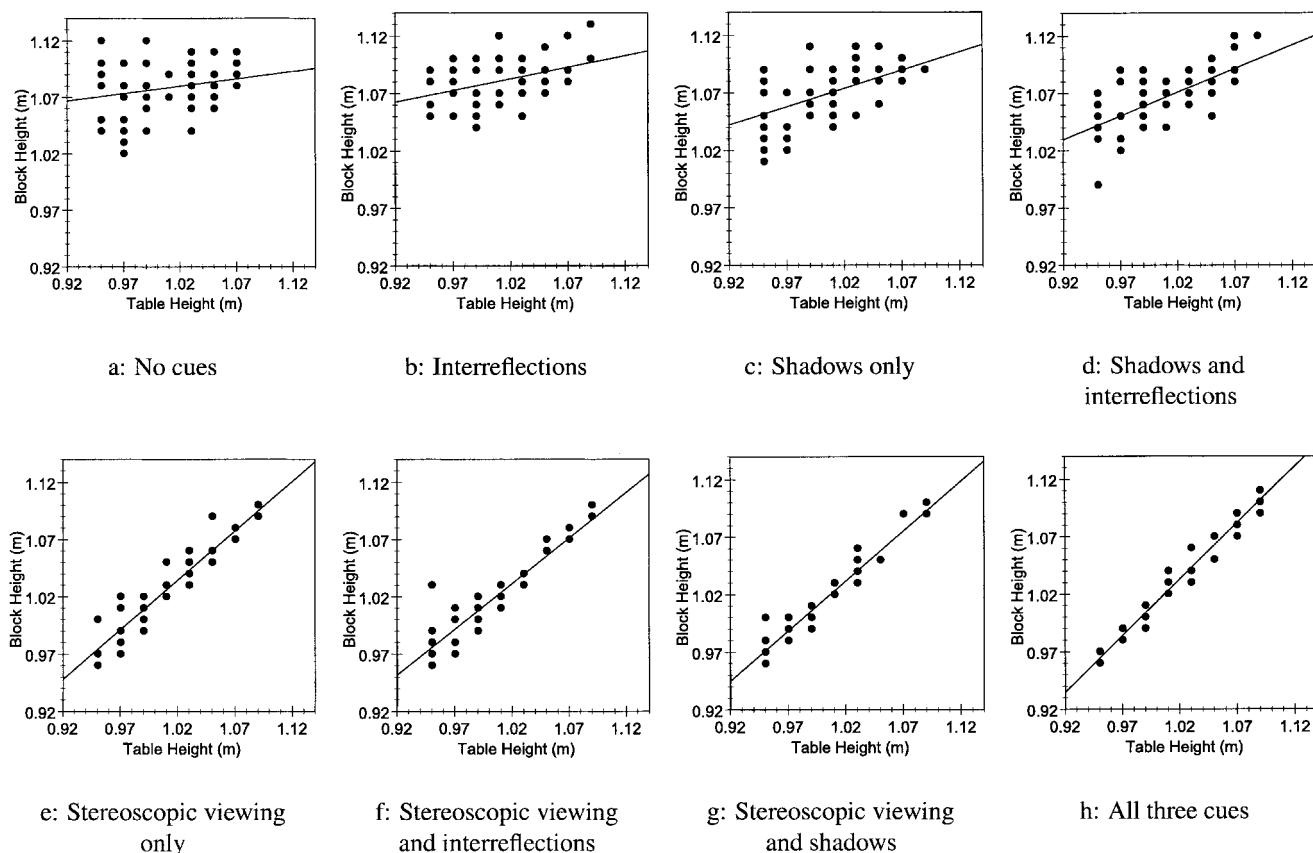
Figure 4. *Plot and regression of the experimental data for subject 3 in experiment 1. When stereoscopic viewing is available as a distance cue, the linear regression captures more of the variance in the data. Overlapping data points are indicated as one point.*

subjects were able to make judgments across trials. As the regression lines better approximate the data shown in figure 4 graphs, R^2 increases.

Shadows and interreflections are scale-invariant properties, providing information about relative distance, which is meaningful when distances and positions are compared. Shadows and interreflections do not provide information about absolute distance; that is, they cannot indicate distances and positions in some predefined and fixed standard. Stereo is capable of providing both relative distance (largely based on disparity) and absolute distance (largely based on vergence). When only relative distance cues are present, no information is available to enable the determination of the scaling of visual position and distance, although the human visual system will often behave as if absolute distance information is

available and assign an absolute distance scale. To be able to compare the absolute distance cue of stereo with the relative distance cues of shadows and interreflections, we focused on the R^2 values in the regression analysis, rather than the slope of the regression line.

Table 1 lists the amount of variability accounted for by the regression analysis (R^2) for each subject and each cue combination, and figure 5 graphs the mean R^2's and standard errors across all subjects. Although there are individual differences, all subjects performed better (higher R^2) with stereo viewing as compared to biocular viewing. A 2 (stereo) $\times$ 2 (shadows) $\times$ 2 (interreflections) repeated-measures analysis of variance (ANOVA) was performed on the mean R^2 values. The ANOVA revealed a statistically significant effect of stereoscopic viewing as a distance cue ($F(1, 5) = 12.58, p < .05$).

Table 1. *The amount of variance accounted for in the regression analysis (R^2) and significance level (p) for each cue condition for each subject in the first experiment. All subjects performed better with stereoscopic viewing. Some subjects performed the task with some success even without any distance cues at all, perhaps from a combination of the negative feedback provided and motor memory developed through the experiment*

Subject	Cues	R^2	p	Subject	Cues	R^2	p	Subject	Cues	R^2	p
1	SSI	0.919	0.000	2	SSI	0.530	0.000	3	SSI	0.974	0.000
	SS	0.867	0.000		SS	0.367	0.000		SS	0.955	0.000
	SI	0.060	0.111		SI	0.237	0.001		SI	0.887	0.000
	S	0.237	0.001		S	0.262	0.000		S	0.928	0.000
	BSI	0.888	0.000		BSI	0.020	0.330		BSI	0.443	0.000
	BS	0.766	0.000		BS	0.117	0.009		BS	0.295	0.000
	BI	0.257	0.000		BI	0.207	0.000		BI	0.170	0.004
	B	0.021	0.337		B	0.109	0.021		B	0.134	0.008
4	SSI	0.240	0.000	5	SSI	0.948	0.000	6	SSI	0.963	0.000
	SS	0.256	0.000		SS	0.744	0.000		SS	0.960	0.000
	SI	0.225	0.001		SI	0.723	0.000		SI	0.897	0.000
	S	0.226	0.001		S	0.684	0.000		S	0.927	0.000
	BSI	0.016	0.348		BSI	0.093	0.024		BSI	0.413	0.000
	BS	0.067	0.042		BS	0.211	0.000		BS	0.521	0.000
	BI	0.062	0.057		BI	0.186	0.002		BI	0.148	0.011
	B	0.033	0.176		B	0.030	0.258		B	0.271	0.000

Visual cue combinations:

SSI: stereo shadows and interreflections

SS: stereo shadows

SI: stereo interreflections

S: stereoscopic display without shadows nor interreflections

BSI: biocular shadows and interreflections

BS: biocular shadows

BI: biocular interreflections

B: biocular display without shadows nor interreflections

The ANOVA indicated no other significant effects or interactions. However, a nonparametric (Friedman) test indicated an effect of shadows ($p < .025$) as well as stereo viewing, ($p < .025$).[1]

---

1. Because not all of the cue combinations had normal distributions of R^2 values, we performed nonparametric analyses to confirm the results of the parametric ANOVA.

3.3 Experiment 1: Discussion

When examined individually, almost all subjects showed a statistically significant effect for stereoscopic viewing, and some also showed a statistically significant effect for shadows. One subject did not have a significant effect for stereoscopic viewing but instead showed a significant effect for shadows, demonstrating that individual differences do exist and that this subject may have weak stereoacuity despite possessing stereo fusion. We
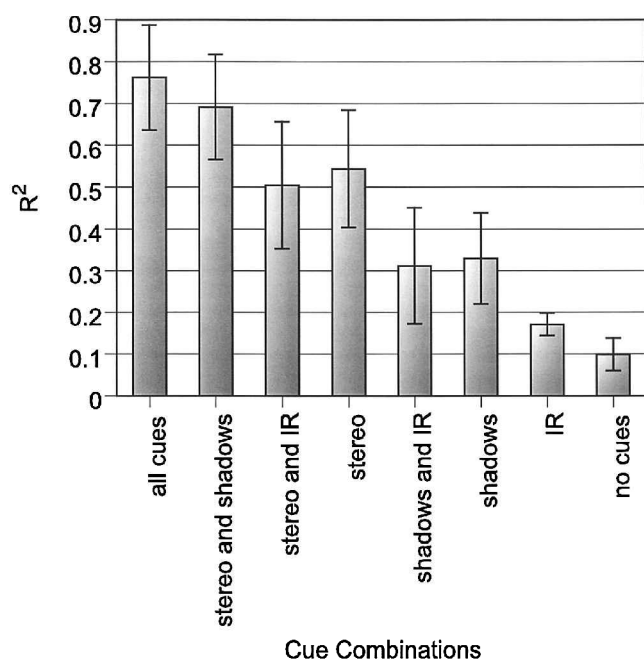
Figure 5. *Mean* R^2 *(±1 SE) for all subjects in the first experiment. When stereoscopic viewing is present, the amount of variance accounted for in the regression analysis is highest.*

designed the present experiment with multiple trials (480 total, 60 in each cue combination) to enable us to primarily analyze individual subjects because pilot studies indicated variability between subjects. However, a secondary analysis on group R^2 enabled us to make some generalizations across all subjects as well. In all, the group analysis indicated a strong effect of stereoscopic viewing as well as an effect of shadows. Interreflections did not have a statistically significant effect across all subjects, although the absence of a significant effect may be a result of the variability between subjects.

Providing negative feedback whenever the block touched the table could have potentially added bias to this experiment. Subjects could have learned the range of table heights through motor memory, and they may have then stopped the block an average distance high above the table, even if they did not know how far the table was from the block. This learning effect may explain subjects' better-than-random performances even when none of the three distance cues were available. To

assess whether there were statistically significant learning effects, R^2 was calculated for each cue condition after grouping the trials into three blocks (two sessions in each block). An analysis on each individual subject showed overall improvement in performance in one subject as a function of session. Grouped across all six participants, R^2 increased for biocular viewing with shadows and interreflections from block 1 to 2, and for biocular viewing without shadows or interreflections from block 1 to 3. Thus, there is some evidence for a possible motor learning bias when stereo was not present. Although there is a small effect of learning within the experiment, it is also possible that learning may have occurred during the training trials before the experiment began. The presence of other distance cues, such as linear perspective, would also improve subject performance for the null case. To counter these effects, extraneous distance cues were avoided or held constant whenever feasible for this experiment. The bias of stopping high above the table may have been stronger when only relative distance cues were available because any perception of absolute distance would have countered this bias. Because stereoscopic viewing was the only available absolute depth cue, this bias would have been stronger during biocular viewing, and may have weakened any effect of the illumination cues (shadows and interreflections).

A second experiment was designed to avoid potential bias due to explicit or implicit visual feedback to subjects.

4   Experiment 2

4.1 Method

4.1.1 Participants. Six computer science graduate students (three men and three women) voluntarily participated in the experiment. All participants were tested individually. All had self-reported corrected vision without the use of prescription eyeglasses, and all were tested for stereo fusion before beginning the experiment. None of the subjects knew of the hypothesis being tested, had participated in experiment 1, or were authors of this paper.

4.1.2 Experiment 2: Procedure. In this experiment, subjects watched a virtual block fall steadily towards a virtual table and stop a small distance above the table. They were instructed to slide their fingers along a scale until the distance between their index finger and their thumb matched the distance between the block and the table. (See figure 3(b).) The distance between the subjects' fingers was recorded by an experimenter. Subjects viewed the final still image of the block suspended over the table for one second before the scene disappeared.

After an initial training period, subjects performed the distance-measuring task 48 uninterrupted times with a biocular display (without stereoscopic viewing), during which the presence and absence of shadows and interreflections varied. As in experiment 1, the repeated cue combinations were presented in random order. Between one to three weeks later, subjects returned to perform the same distance-measuring task with stereoscopic viewing, and the same 48 variations of shadows and interreflections were presented to subjects in the same order. Practical considerations precluded mixing biocular and stereoscopic viewing trials in the same experimental session, as had been done for experiment 1.

4.2 Experiment 2: Results

A linear regression was performed on each subject for each cue combination, assessing how well the actual distance between the virtual block and virtual table predicted the reported distance between the two. Figure 6 shows the data plot and regression analysis for subject 2. All subjects had similar R^2 values as subject 2. Unlike the first experiment in which some data were discarded because of contact between the object and the table, all data were retained in this experiment. Table 2 lists the R^2 values for each subject and each cue combination. Not only did the regression lines account for more variance in the data as more cues were provided, but the regression lines also showed steeper slopes. Subject 2's data in figure 6 changes from an almost horizontal line when none of the three cues are available (figure 6(a)) to a diagonal line if shadow cues are present (figures 6(c) and (d)) or

stereoscopic viewing is present (figures 6(e) (f), (g), and (h)).

Figure 7 graphs the mean R^2's and standard errors for the second experiment. In general, subjects performed best either with stereo viewing or when both shadows and interreflections were present. A 2 (stereo) × 2 (shadows) × 2 (interreflections) ANOVA on the mean R^2's revealed a significant effect for stereo ($F(1, 5) = 71.39, p < .001$) and shadows ($F(1, 5) = 52.79, p < .001$), and a marginal effect of interreflections[2] ($F(1, 5) = 5.4, p < .07$). There was a significant stereo×shadows interaction ($F(1, 5) = 95.15, p < .001$) and a significant stereo×interreflections interaction ($F(1, 5) = 9.79, p < .05$). The shadows×interreflections interaction was not significant ($F(1, 5) = 0.14, p = .729$).

To examine the interactions of shadows and interreflections with viewing condition, separate 2 (shadows) × 2 (interreflections) ANOVAs for each viewing condition (stereoscopic or biocular viewing) were performed. We found a significant effect for shadows during biocular viewing ($F(1, 5) = 110.01, p < .001$) and for interreflections during biocular viewing ($F(1, 5) = 7.28, p < .05$). There were no significant effects for shadows or interreflections during stereoscopic viewing.

A post hoc comparison between the stereo cue combinations and both illumination cues (shadow and interreflections) without stereo indicated no difference in performance ($p = .34$). For the biocular conditions, although average performance with shadows and interreflections combined was better than with shadows alone, the difference was not statistically significant ($p = .16$). These results indicate that, although stereoscopic viewing is a powerful cue for distance, shadows and interreflections combined or shadows alone provide information that leads to equivalent performance. Interreflections alone improved performance compared to no cues at all, but to a much smaller degree than shadows.

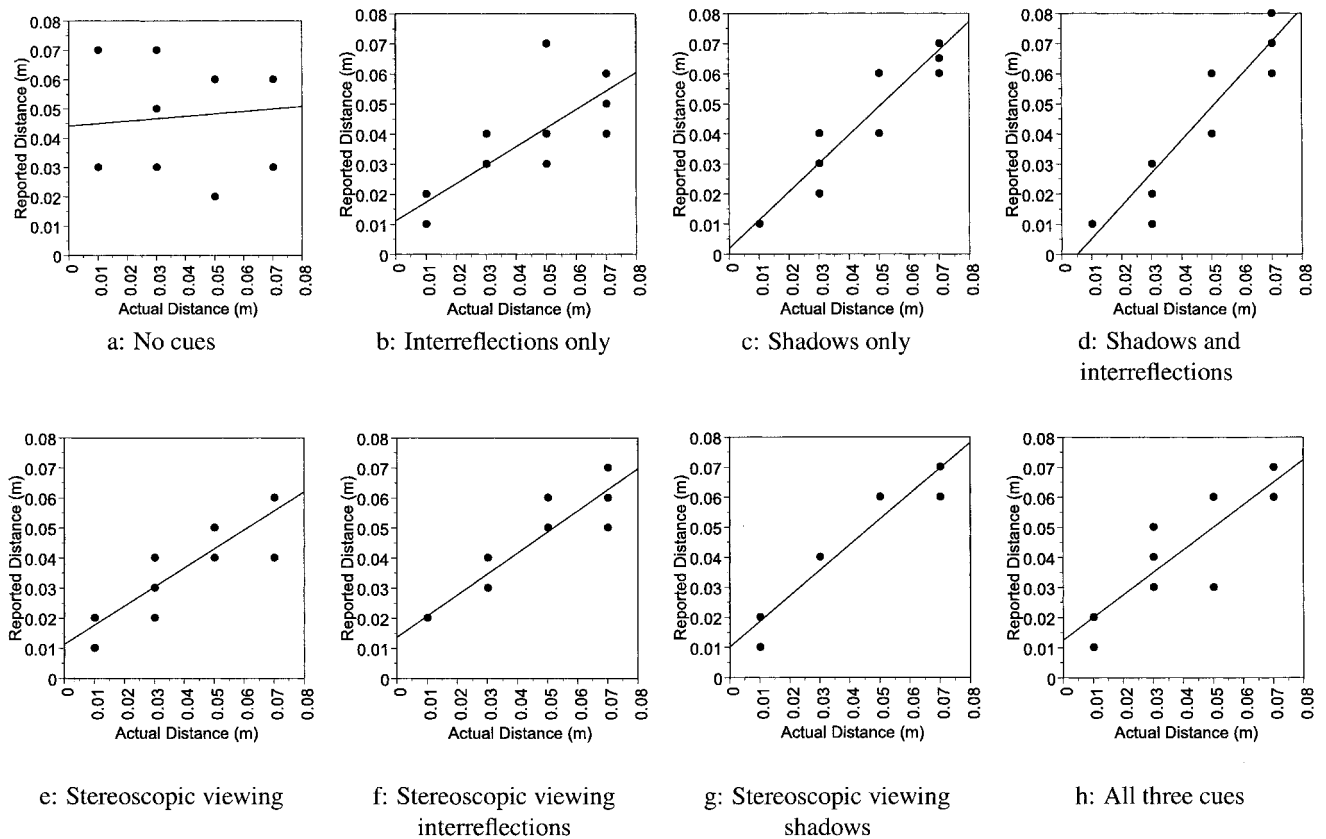2. A nonparametric Friedman test confirmed the main effects of stereo and shadows ($p < .025$).

Figure 6. *Plot and regression of the experimental data for subject 2 in experiment 2. The linear regression captures more of the variance in the biocular data if shadows or interreflections are available as a distance cue. However, with stereoscopic viewing, shadows and interreflections do not have a significant effect. Overlapping data points are indicated as one point. (There are three points per distance in each figure.)*

4.3 Experiment 2: Discussion

In the original experiment, subject performance was better than random when no depth cues were present. In this revised experiment, subject performance without any depth cues is closer to the expected poor performance, suggesting the second experimental design eliminated some of the biases of the first. This improvement exists despite the seemingly less natural probe of sliding fingers along a scale.

Every subject demonstrated a statistically significant effect for stereoscopic viewing when examined individually. All subjects' data also showed a statistically significant effect for shadows during biocular viewing, and some subjects' data showed a statistically significant effect for interreflections during biocular viewing. The pooled data reported similar results: stereoscopic viewing, shadows with biocular viewing, and interreflections with biocular viewing had statistically significant effects as distance cues.

The R^2 values in table 2 and the mean R^2 values in figure 7 demonstrate that subjects performed similarly when stereoscopic viewing was the only available cue and when both illumination cues were present without stereoscopic viewing, suggesting that relative distance cues like shadows and interreflections may compensate for the absence of absolute distance cues in some situations.

Table 2. *The amount of variance accounted for (R^2) and significance level (p) for each cue condition for each subject in the second experiment. With biocular viewing, all subjects performed better with illumination cues. With stereoscopic viewing, most subjects performed equally well regardless of the presence or absence of illumination cues*

Subject	Cues	R^2	p	Subject	Cues	R^2	p	Subject	Cues	R^2	p
1	SSI	0.794	0.000	2	SSI	0.802	0.000	3	SSI	0.873	0.000
	SS	0.549	0.006		SS	0.924	0.000		SS	0.892	0.000
	SI	0.618	0.002		SI	0.888	0.000		SI	0.894	0.000
	S	0.735	0.000		S	0.817	0.000		S	0.846	0.000
	BSI	0.735	0.000		BSI	0.905	0.000		BSI	0.779	0.000
	BS	0.380	0.033		BS	0.891	0.000		BS	0.700	0.001
	BI	0.484	0.012		BI	0.632	0.002		BI	0.166	0.188
	B	0.007	0.798		B	0.011	0.746		B	0.010	0.760
4	SSI	0.876	0.000	5	SSI	0.850	0.000	6	SSI	0.823	0.000
	SS	0.874	0.000		SS	0.838	0.000		SS	0.824	0.000
	SI	0.717	0.001		SI	0.512	0.009		SI	0.794	0.000
	S	0.790	0.000		S	0.580	0.004		S	0.759	0.000
	BSI	0.743	0.000		BSI	0.722	0.000		BSI	0.871	0.000
	BS	0.598	0.003		BS	0.869	0.000		BS	0.487	0.012
	BI	0.062	0.436		BI	0.177	0.174		BI	0.184	0.164
	B	0.159	0.200		B	0.051	0.0478		B	0.040	0.534

Visual cue combinations:

SSI: stereo shadows and interreflections

SS: stereo shadows

SI: stereo interreflections

S: stereoscopic display without shadows or interreflections

BSI: biocular shadows and interreflections

BS: biocular shadows

BI: biocular interreflections

B: biocular display without shadows nor interreflections

5   Conclusions

We conducted two experiments examining distance perception through two different experimental designs. The presence and absence of stereoscopic viewing, shadows, and interreflections were varied for a total of eight different distance cue combinations. In both experiments, stereo was found to be a strong and statistically significant distance cue for the task of perceiving distance between an object and a surface. Nonparametric statistics suggested that shadows were a significant effect in the first experiment. In the second experiment, the combination of shadow and interreflection cues proved to be approximately as effective as the use of binocular stereo. Shadows alone were also effective without stereo, although there were greater individual differences.

Other individual differences were found between subjects. Many subjects perceived relative distances equally well, but their perceived scaling ratios varied widely. This variation may have been caused by some subjects having weaker stereoacuity, despite being capable of stereo fusion. In addition, some subjects appeared to
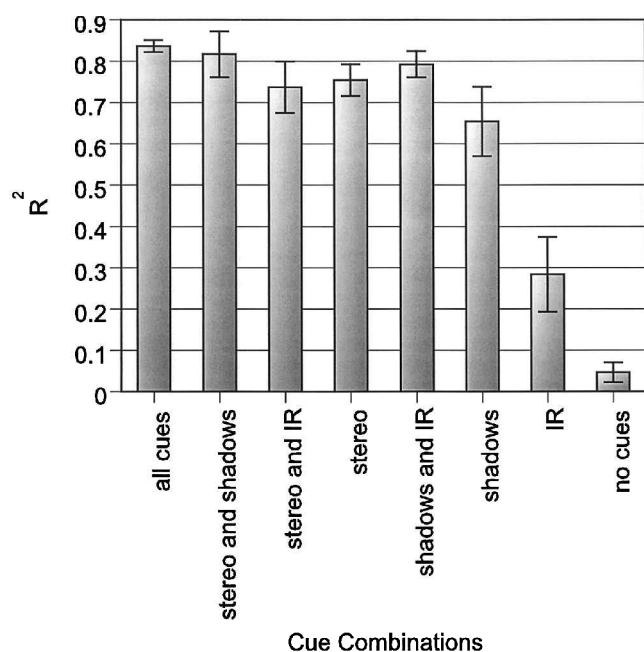
Figure 7. *Mean* R^2 *($\pm$1 SE) for all subjects in the second-experiment. The correlation in the data between the actual stopping distance and the reported stopping distance is highest with stereoscopic viewing.*

use shadows more effectively, whereas other subjects appeared to more effectively use interreflections.

Much additional work will be required before the generality of these results is known with confidence. Experiment 1 involved a dynamic, visuomotor task with closed-loop control, whereas experiment 2 used an open-loop matching task involving both dynamic and stationary views of the scene. Different tasks and either static viewing or different sorts of visual motion might well produce different results. In addition, experiment 1 was potentially subject to biases due to interactive latencies. Distance judgments between points along the line of sight and perpendicular to the line of sight are both likely to depend on visual cues different from those involved in the oblique viewing reported here. Finally, the effectiveness of illumination cues such as shadows and interreflections almost certainly depends on both the geometry, surface markings, and materials in the scene, as well as how faithfully these global illumination effects are rendered in the computer graphics.

References

Barfield, W., & Rosenberg, C. (1995). Judgements of azimuth and elevation as a function of monoscopic and binocular depth cues using perspective display. *Human Factors, 37,* 173–181.

Bloj, M. G., Kersten, D., & Hurlbert, A. C. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature, 402,* 877–879.

Brooks, T., Ince, I., & Lee, G. (1991). Vision issues for space teleoperations assembly and servicing (vistas). (Tech. Rep. STX/ROB/91-01). STX Corporation.

Cole, R. E., Merritt, J. O., Fore, S., & Lester, P. (1990). Remote manipulator tasks impossible without stereo TV. *Proc. SPIE Stereoscopic Displays and Applications, 1256,* 255–265.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distance: The integration, relative potency and contextual use of different information about depth. In W. Epstein & S. Rogers (Eds.), *Perception of Space and Motion* (pp. 69–117). New York: Academic Press.

Dosher, B. A., Sperling, G., & Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research, 26*(6), 973–990.

Drascic, D. (1991). Skill acquisition and task performance in teleoperation using monoscopic and sterescopic video remote viewing. *Proc. Human Factors Society's 35th Annual Meeting,* 1367–1371.

Hsu, J., Babbs, C. F., Chelberg, D. M., Pizlo, Z., & Delp, E. J. (1993). A study of the effectiveness of stereo imaging with applications in mammography. *Proc. SPIE Human Vision, Visual Processing, and Digital Display IV,* 154–165.

Hu, H. H., Gooch, A. A., Thompson, W. B., Smits, B. E., Rieser, J. J., & Shirley, P. (2000). Visual cues for imminent

object contact in realistic virtual environments. *Proceedings IEEE Visualization 2000,* 179–185.

Hubona, G. S., Wheeler, P. N., Shirah, G. W., & Brandt, M. (1999). The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. *ACM Transaction on Computer-Human Interaction CHI 99, 6,* 214–242.

Kersten, D. & Hulbert, A. (1996). Discounting the color of mutual illumination: A 3D shape-induced color phenomenon. *Investigative Opthamology and Visual Science,* 37(3).

Kersten, D., Mamassian, P., & Knill, D. C. (1997). Moving cast shadows induce apparent motion in depth. *Perception, 26*(2), 171–192.

Kim, W. S., Ellis, S. R., Tyler, M. E., Hannaford, B., & Stark, L. (1987). Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks. *IEEE Trans. on Systems, Man, and Cybernetics, SMC-17,* 61–71.

Kjelldahl, L., & Prime, M. (1995). A study on how depth perception is affected by different presentation methods of 3D objects on a 2D display. *Computers & Graphics, 19*(2), 199–202.

Liang, J., Shaw, C., & Green, M. (1991). On temporal-spatial realism in the virtual reality environment. *Proc. UIST, ACM Symp. on User Interface Software and Technology,* 19–25.

Madison, C., Thompson, W., Kersten, D., Shirley, P., & Smits, B. (2001). Use of interreflection and shadow for surface contact. *Perception & Psychophysics, 63,* 187–193.

Miller, R. H., & Beaton, R. J. (1991). Some effects on depth-position and course-prediction judgments in 2-D and 3-D displays. *Proc. SPIE Stereoscopic Displays and Applications II, 1457,* 248–258.

Reinhart, W. F., Beaton, R. J., & Snyder, H. L. (1990). Depth cueing for visual search and cursor positioning. *Proc. SPIE Stereoscopic Displays and Applications, 1256,* 12–21.

Spain, E. H. (1990). Stereo advantage for a peg-in-hole task using a force-feedback manipulator. *Proc. SPIE Stereoscopic Displays and Applications, 1256,* 244–254.

Surdick, R. T., Davis, E. T., King, R. A., & Hodges, L. F. (1997). The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances. *Presence: Teleoperators and Virtual Environments, 6*(5), 513–531.

Wanger, L. R., Ferwerda, J. A., & Greenberg, D. P. (1992). Perceiving spatial relationship in computer-generated images. *IEEE Computer Graphics & Applications, 12*(3), 44–58.

Wann, J. P., Rushton, S., & Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. *Vision Research, 35*(19), 2731–2736.

Ware, C., & Franck, G. (1996). Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Computer Graphics, 15*(2), 121–140.

Yeh, Y.-Y., & Silverstein, L. D. (1992). Spatial judgements with monoscopic and stereoscopic presentation of perspective displays. *Human Factors, 34,* 583–601.

Yonas, A., Goldsmith, L. T., & Hallstrom, J. L. (1978). Development of sensitivity to information provided by cast shadows in pictures. *Perception, 7,* 333–341.