

Enhancing Spatial Perception and User Experience in Video Games with Volumetric Shadows

Tuukka M. Takala¹ Perttu Hämäläinen¹ Mikael Matveinen¹ Taru Simonen¹ Jari Takatalo²

¹ Department of Media Technology, Aalto University Otaniementie 17, 02150 Espoo, Finland
² POEM Research Group, University of Helsinki Siltavuorenpenger 1 A, Helsinki, Finland
{tuukka.takala, perttu.hamalainen, mikael.matveinen, taru.simonen}@aalto.fi, jari.takatalo@helsinki.fi

ABSTRACT

In this paper, we investigate the use of volumetric shadows for enhancing three-dimensional perception and action in third-person motion games. They offer an alternative to previously studied cues and visual guides. Our preliminary survey revealed that from the games that require Kinect, 37% rely primarily on a third-person view and 9% on a first-person view. We conducted a user study where 30 participants performed object reaching, interception, and aiming tasks in six different graphical modes of a video game that was controlled using a Kinect sensor and PlayStation Move controllers. The study results indicate that different volumetric shadow cues can affect both the user experience and the gameplay performance positively or negatively, depending on the lighting setup. Qualitative user experience analysis shows that playing was found to be most easy and fluent in a typical virtual reality setting with stereo rendering and flat surface shadows.

Author Keywords

Depth perception; depth cues; stereoscopy; games; 3D user interface; game experience; volumetric shadows.

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces.

INTRODUCTION

Precise 3D input devices have become mainstream thanks to low-cost game hardware like Microsoft Kinect and PlayStation Move. However, the dream of immersive first-person virtual reality (VR) has not become true at homes, because display technology is lagging behind. Meanwhile, commercial motion games use a single television with the game world typically portrayed in a first-person view (Rise of Nightmares), a third-person over the shoulder view (Kinect Star Wars, Fighters Uncaged, The Fight: Lights Out) or a third-person mirror-like view (Your Shape Fitness Evolved, Dance Central).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

OzCHI'13, November 25 - 29 2013, Adelaide, Australia
Copyright 2013 ACM 978-1-4503-2525-7/13/11...\$15.00.

Our work is motivated by our experiences from prototyping a third-person Kinect action game where the object was to reach static and moving targets. In early tests using a television as the display, estimating distances and hitting targets in 3D seemed very difficult. It was clear that some form of additional cues or guidance was needed, preferably without the need for a stereo 3D display to reach a wider audience.

This paper presents a study that compares both proven and novel cues for enhancing target reaching and intercepting in a third-person perspective. Reaching and intercepting are tasks integral to most fighting, sports, and action games. Our main contribution is the study of volumetric shadows. We wanted to avoid possibly alienating and “brute force” visual guides such as rendering the predicted path of an object in the game world. We were particularly intrigued by the possibility of designing lighting so that it *enhances gameplay and creates positive user experience*.

According to our review of the literature, the use of volumetric shadows has not been studied in positioning or interception tasks, although different volumetric rendering techniques have been proposed for enhancing depth perception (Bruckner and Groller, 2007; Ropinski, Doring, and Rezk-Salama, 2010).

BACKGROUND

The human visual system combines a number of cues with information about the posture of the eyes to compute the distance of objects (Cutting and Vishton, 1995; Hubona et al., 1999; Wanger, Ferwerda, and Greenberg, 1992). Cutting and Vishton (1995) list a total of 15 sources of information about three-dimensional layout and distance: accommodation, aerial perspective, binocular disparity, convergence, height in visual field, motion perspective, occlusion, relative size, relative density, linear perspective, light and shading, texture gradients, kinetic depth, kinetic occlusion and disocclusion, and gravity. Increasing the amount of information generally increases the accuracy of depth judgments (Cutting and Vishton, 1995).

The effect of visual cues such as binocular disparity on motor behavior has been researched extensively. The importance of cues varies depending on the activity or sport. For example, Heinen and Vinken (2011) report that binocular vision is not necessary for experienced gymnasts.

On the other hand, Laby *et al.* (2011) report better stereoacuity in Olympic-level soccer and softball athletes compared to archers, and various ball catching and hitting studies have found binocular vision superior to monocular (Bennett *et al.*, 2000; van Hof, van der Kamp, and Savelsbergh, 2006; Mazyn *et al.*, 2007; Bulson, Ciuffreda, and Ludlam, 2009). However, monocular catching and hitting is still possible, and the importance of binocular vision varies as a function of, e.g., target size and speed (Bennett *et al.*, 2000; Mazyn *et al.*, 2007).

In the realm of computer graphics, Wanger *et al.* (1992) found shadows and perspective to help in a positioning task conducted using a single 2D display. Their study required the participants to match the position, rotation and size of an object with a reference object. In the context of this paper, moving the avatar's hand to touch a target can also be regarded as a position matching task. Wanger's follow-up study (Wanger, 1992) found shadow sharpness and size to have no significant effects on estimating object size and position. Hubona *et al.* (2004) reported that in monocular viewing, shadows enhance positioning accuracy to a level equivalent to stereo 3D (S3D) viewing without shadows. However, they also found that in S3D shadows increased both accuracy and response times.

There are many studies of using head tracking to aid depth perception: Teather and Stuerzlinger (2007), and Boritz and Booth (1997) have investigated both S3D and head-tracking in positioning tasks, and while their results support the importance of S3D, no significant effects were found for head tracking. However, it has also been shown that head tracking can help more than S3D in other tasks (Arthur, Booth, and Ware, 1993).

It should be noted that humans use focal and ambient vision differently for movement control. Focal vision is specialized for object identification and conscious movement control, and ambient vision is specialized for unconscious control based on, e.g., optical flow (Schmidt and Wrisberg, 2004). Conscious processing of focal vision is slow compared unconscious movement control based on ambient vision. In commercial motion games and our tests using a single television about 2-3 meters away from the user, the display is probably too small for ambient vision to be of much use.

The unnatural third-person perspective also changes the optical flow patterns. Although a first-person view could have less problems, being closer to our natural way of seeing the world, we agree with Oshita (2006) that a third-person view is optimal for many full body action games, because it allows the players to fully see how their movements are mapped to the avatar's movements and how the avatar reacts to the environment.

We went through a list of 95 games that required Kinect (Wikipedia, 2013) and viewed their gameplay videos online: 35 of the games relied primarily on a third-person

view, 9 relied primarily on a first-person view, 41 offered a mirror-like view (usually the mirror image was very small), and the other 10 games involved primarily 2D pointer interaction or did not fit into our categorization. Existing games often make it easier to hit or intercept objects using graphical guides. Sport Champions: Table Tennis draws the predicted path of the ball as a 3D curve in the game world. The Kinect Sports goalkeeping game transforms an interception task into a reaching task by displaying a static marker that shows where the ball will be a moment later.

In the onset of our study we conducted preliminary tests with a third-person Kinect action game. Using a 2D television as the display, estimating distances and hitting targets in 3D seemed much more difficult than in real life or in first-person, immersive VR system like the CAVE (Cruz-Neira, Sandin, and DeFanti, 1993). We found it possible to learn to hit static objects through trial and error, but moving objects were more often missed than hit. There was an apparent need for some form of additional cues or guidance, preferably without resorting to S3D displays so that wider audiences could be reached. This led us to experiment with volumetric shadows as depth cues. Our hypothesis is that volumetric lighting can be used to enhance spatial perception in real-time 3D applications. The purpose of our study was to test that hypothesis and to evaluate how different depth cues affect user experience.

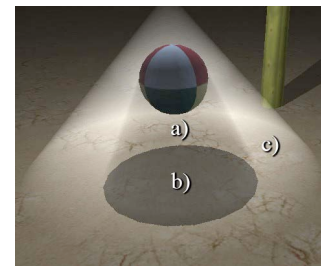


Figure 1. a) Volumetric shadow, b) surface shadow, c) volumetric light shaft.

Volumetric Shadows

In computer graphics realism of rendered images can be improved by taking into account light scattering effects of air and other participating media (Nishita, Miyawaki, and Nakamae, 1987). This results as images where visible beams of light emanate from the light source. If there are objects blocking the light beams, then no light scattering will occur in the umbral regions behind the objects, and so called *volumetric shadows* (Figure 1) will appear (Wyman and Ramsey, 2008).

For the sake of clarity, we use the term *surface shadow* to describe ordinary shadows that occur on illuminated surfaces when direct illumination from a light source is blocked by an object. It is worth noting that a light source can cast both surface shadows and volumetric shadows at the same time, as seen in Figure 1. To distinguish from standard light sources that emit only direct light, we use the

term *volumetric light source* for light sources that feature volumetric light shafts and shadows.

Surface shadows are prevalent in 3D video games, whereas few games use volumetric shadows. Those games that have volumetric shadows use them as mere eye-candy or for increased graphical immersion (e.g. *Uncharted 3*, *F.E.A.R. 2*). *Alan Wake* is a rare exception, as volumetric light sources play an important role in its gameplay, where it is necessary to illuminate the player character’s adversaries in order to defeat them.

Previous lighting and shadow studies about depth perception have been mostly limited to surface shadows. A surface shadow can resolve the depth/size ambiguity, but it is not ideal if the object of interest is far from the surface, because 1) the shadow cannot be seen accurately while keeping ones gaze fixed on the target and 2) the shadow may end up occluded or outside the camera view. In contrast, a volumetric shadow traces the linear path between the object and the surface shadow that it casts. By doing so, the volumetric shadow also points into the direction of the light source.

User Experience and Gameplay

Earlier video game lighting research has explored how lighting color affects gameplay performance and emotion (Knez and Niedenthal, 2008), and how lighting can be automatically adjusted to accommodate dramatic, aesthetic, and communicative goals of a game (El-Nasr and Horswill, 2004). Only a little is known how different depth cues affect user experience (UX) in games. Here, we refer the term UX as a subjective experience that stems from the use of technology that is in our case gameplay.

The concept of presence, namely the sense of being in a mediated environment, is broadly studied in different media (Lombard and Jones, 2007). Both S3D displays (IJsselstein et al., 2001; Takatalo et al., 2011) and head-tracking (Snow and Williges, 1998) have been found to enhance presence. Spatial awareness, attention, and realness/naturalness form the perceptual “Big three” sub-components of physical presence (International Society for Presence Research, 2000). In addition to users’ perceptions, their evaluations of the provided action are needed to study UX (Takatalo, Nyman, and Laaksonen, 2008; Takatalo et al., 2011).

Csikszentmihalyi’s concept of flow (1975) provides a good framework to analyze human activity in many contexts. The antecedents of flow – i.e. evaluated skills and challenges in the situation – indicate whether participants are experiencing mastery (skills above the challenges) or coping (skills below the challenges). In digital games, the mastery situation is found to have a positive effect on the enhanced motivation to continue on playing and to play again (Ryan, Rigby, and Przybylski, 2006). Thus, the flow framework integrates emotional and motivational layers on top of the perceptual-cognitive evaluation process.

USER STUDY

We organized a user study in which participants played a 3D video game in six different experimental conditions. We advertised the study via mailing lists and social media, and recruited a total of 35 participants from our university; students, researchers, and staff. Each participant received a movie ticket voucher after taking part in the study.

Game

The game used in the study was a simplified, video game version of wall tennis implemented with the Unity game engine. The player’s in-game avatar held a paddle racket in each hand, standing in front of an archery target with a diameter of 3 meters. The rackets were represented and tracked with PlayStation Move controllers. The avatar was controlled via OpenNI library using a Kinect sensor.

The objective of the game was to strike virtual balls and hit as close to the archery target’s center (bull’s-eye) as possible. Hitting directly at the bull’s-eye gave 100 points. The number of awarded points decreased linearly to zero on a disk with a radius of 4 meters, so that points were given even when hitting outside the archery target. Hit score was displayed on screen for two seconds after a hit was registered, so that players would strive to get better scores.

Condition \ Depth cue	0	1	2	3	4	5
Stereo 3D		x				x
Motion parallax		x	x	x	x	x
Surface shadows		x	(x)	(x)	x	(x)
Ball volum. light				x		x
Racket volum. lights			x			
Above volum. light					x	

Table 1: Depth cues among the study conditions.

Our game had six different graphical modes that acted as test conditions in our study. Each condition had a unique set of depth cues, combining traditional cues and our volumetric shadows (Table 1). Screenshots of different conditions are presented in Figure 2. Condition 0 acted as a baseline, as it had the least amount of depth cues, providing only perspective cues. Condition 1 represented the “industry standard” of VR graphics with surface shadows and S3D. In Condition 2 both rackets had an omnidirectional volumetric light source with a unique color. In Condition 3 each ball acted as an omnidirectional volumetric light source. Condition 4 presented a volumetric light source placed high above the playing area. Condition 5 had S3D; otherwise it was identical to Condition 3. Surface shadows in conditions 2, 3, and 5 were fainter than in conditions 1 and 4, due to the additional lights.

We made two design decisions for easing the learning curve of aiming and striking: 1) There was no gravity in the physics simulation and striking a ball sent it on a linear path with constant velocity. 2) Balls interacted only with the

rackets and the in-game avatar passed through them without effect. Bouncing from ground and walls was left enabled for facilitating a basic sense of physical immersion.

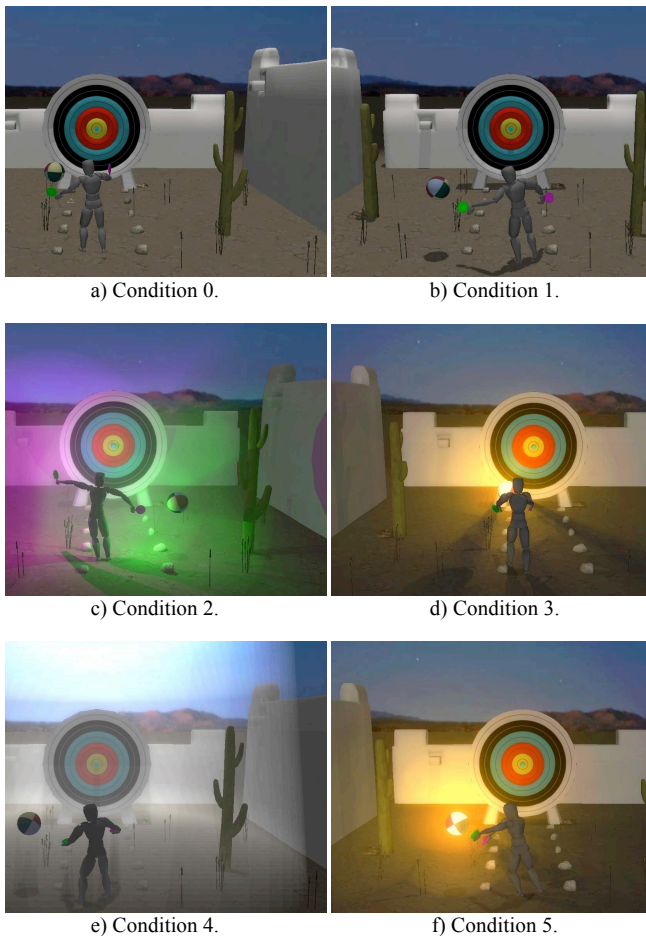


Figure 2. Six graphical modes that were the test conditions of our user study.

The game presented only one ball at any given moment. After striking a ball and registering a hit or miss, the ball would disappear. Before a new ball would appear, the player had to return to the center of the playing area, marked by a circular object. For our study this meant that the event of striking a ball became a repeatable trial with always the same outset, regardless of the participant, previously hit balls, or other conditions. The game had both static balls hovering in the air and balls that appeared from either side of the archery target, moving linearly through the playing area. In the latter case, the movement was constant with the speed of 2 m/s, which was determined to be challenging enough in our initial testing phase.

Natural Aiming and Depth Cues

As seen from Figure 2, volumetric lighting adds a sense of depth in conditions 2, 3, 4, and 5 of the game. Volumetric shadows in conditions 2, 3, and 5 form a cone frustum that is aligned along the line between the ball and the racket. In Condition 2, the light source in the racket casts a surface

shadow from the ball that can be used as a laser sight to aim the ball. A high-above volumetric light source in Condition 4 creates volumetric shadows that bind objects to their surface shadows on the ground.

Equipment

The game ran 50 frames per second on a computer with Windows 7, Intel Core2 6600, 3 GB of RAM, and Nvidia GeForce 8800GTX. The game was displayed on a 55” Panasonic TX-L55ET5Y television that could output S3D through passive, circularly polarized stereo-glasses. PlayStation Move controllers were connected to our computer via Move.me software running on a PlayStation 3 that was equipped with a PlayStation Eye camera. The coordinate system between Move controllers and Kinect was calibrated with RUIS library (Takala et al., 2011).

Environment

Playing area in front of the TV was marked with a 2 meter by 2 meter floor mat, from which the game could be played without stepping outside. Illuminance within that area varied between 235-320 lux, as measured with Konica Minolta Chroma Meter CL-200. An illuminance between 95-135 lux was measured by the back wall that was seen as background by Kinect and PlayStation Eye camera. The playing area center was at distance of 3.2 meters from the TV. There the luminous intensity was measured to be 157 cd/m², by pointing Konica Minolta Luminance Meter LS-110 at a white TV screen. The luminous intensity through the passive stereo glasses was 65 cd/m².

Design and Procedure

We used a randomized within-subject design with an incomplete counterbalancing, where the participants were exposed to each of the six conditions in a random order. Every other participant was exposed to the conditions in a reverse order compared to the previous participant, for counterbalancing learning effect over the conditions (Dey, 2010). We paid special attention that the participants would not discover our hypothesis or other study details before or during the study; we did not mention lights, shadows, depth, or aiming cues in our interview questions. Instead we let the participants report their findings in their own words, discussing only topics that they had brought up themselves.

Metrics

The following information was collected from each of the participants: vision, background, and subjective UX evaluations after each test condition. Because of the large number of test conditions, we attempted to keep the test procedure simple, but to still get as rich UX descriptions from the participants as possible. Thus, we used both qualitative interviews and quantitative scales to assess perceptions and actions related to UX.

After each condition, participants evaluated their “overall feeling” (0-10 Likert), “challenges of the task” (1-7 Likert),

and their own levels of skills (1-7 Likert). Our idea with *overall feeling* was to give the participants a chance to freely rate each condition. It provided a good one-dimensional comparison for the multidimensional interview data. In the past we have used challenges and skills measures to create a flow-space in digital games (Takatalo et al., 2010). Flow-space allows us to evaluate whether participants are experiencing mastery or coping.

The participants were interviewed and asked to describe the environment and the task. Different descriptions were compressed into 17 dichotomous variables. Each of the 17 variables had at least ten mentions among the participants. The 17 variables thus formed were analyzed in a correspondence analysis (CA), which is a multivariate descriptive data analytical technique for categorical data. CA shares similarities with the principal components analysis, which applies to continuous data. CA is helpful in depicting the relationship between two or more categorical variables in a 2-dimensional chart, summarizing and illustrating similarities and differences between categories and the associations between them (Greenacre, 1984).

Finally, the participants rated their feelings of Pleasure/Valence, Arousal, and Dominance/Control (PAD) using Self-Assessment Manikin's (SAM) (Lang, 1980). The PAD profile describes participants' degree of valence, level of activeness of an emotion, and experienced sense of control. Flow theory (Csikszentmihalyi, 1975) acknowledges PAD emotions as important flow factors.

Procedure

First the participants with corrected vision were asked to wear their eyewear if that helped them to see the TV better. Next we administered the standard TNO stereoscopic vision test (Lam ris Ootech, 1972), in order to determine how sensitive the participants' stereoscopic vision was on a scale of 15 - 480 arc seconds.

Before starting, participants practiced the game for 5 minutes in a graphical setting that had only surface shadows and monocular rendering. They were instructed that the goal of the game was to strike the ball so that it would hit as close to the bull's-eye as possible without bouncing it off from the ground or walls. The participants were also informed about tracking problems that would occur if the PlayStation Move controllers were turned away from PlayStation Eye camera's view. This was a problem for those participants who were using the controllers like a tennis racket, swinging their hand backwards and swiveling the controller away from the PlayStation Eye camera before striking. The problem was pronounced when striking balls in the farther edges of the playing area.

After practice the participants were exposed to six different conditions, each consisting of 30 trials of striking the game's virtual ball. Half of the balls were static and half were moving linearly, and these two sets were constant for each participant and each condition, but the order of balls

was randomized. Figure 3 illustrates the position of the static balls and trajectories of the moving balls. The static and moving balls had a varied distance from the ground and the moving balls' trajectories were parallel to ground plane.

Striking through 30 balls in each condition took on average 4 minutes. Participants wore the stereo-glasses even during the conditions that did not have S3D. After each condition, we interviewed the participant about the experience. These interviews lasted on average 8 minutes each. The mean duration for the whole experiment was 90 minutes.

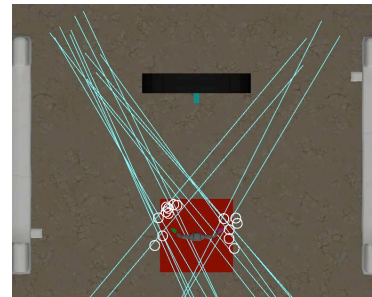


Figure 3. Static ball positions are displayed with white circles, and moving ball trajectories with blue lines. The red square represents the playing area.

RESULTS AND ANALYSIS

When calculating study results, we included 30 people with stereoscopic acuity better or equal to 120 arc seconds from the total of 35 participants. This way the possible differences between S3D and monocular rendering would not be obscured by participants with poor stereoscopic vision.

Participant Demographics

We included 24 males and 6 females in the analysis. Most of them (90%) were Finnish speaking students or research staff of the Aalto University. The mean age of the participants was 27.7 years (SD=4.04 years). Participants were rather familiar with the commercially available 3D user interfaces for digital games: 27% had used Sony Move, 43% Microsoft Kinect, and 87% Nintendo Wii.

Gameplay Performance Results

We explored statistically significant differences between conditions with the Kruskal-Wallis test, Friedman test, and a 3-way ANOVA whose factors were condition index, ball index, and subject index. Post-hoc tests were applied with a p-value of 0.05 using Tukey-Kramer correction. Results were obtained with Matlab's Statistics Toolbox.

Aiming Accuracy

We measured balls' hit accuracy with two variables: 1) Distance of the hit location on the archery target plane from the bull's-eye. If a ball had missed the target plane or hit the ground or side walls, its distance was set as the maximum distance that was recorded. By doing this, even the balls that missed the target plane were taken into account when

examining hit accuracy between conditions with a Kruskal-Wallis test. 2) We used a 3-way ANOVA over all the balls where binary outcome of the hit event was the dependent variable. Static and moving balls were examined separately.

Figure 4a is a boxplot presenting the number of static balls that hit the archery target without bouncing from ground or walls. No significant differences were found with the 3-way ANOVA, Kruskal-Wallis, or Friedman tests.

Number of moving balls that hit the archery target can be seen in the boxplot chart of Figure 4b. By comparing Figure 4a and Figure 4b it is clear that the participants were able to hit the target much more often when striking static balls. In fact, the participants missed many of the moving balls completely when trying to strike them (failed interception).

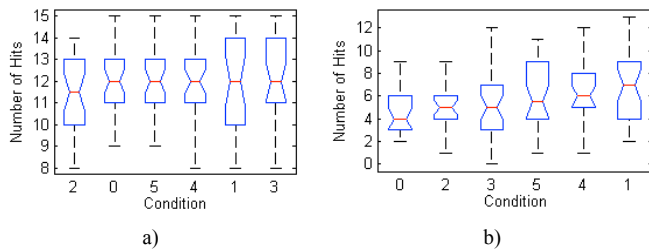


Figure 4. Number of (a) static and (b) moving balls that hit the archery target, struck by 30 participants. The conditions are sorted by medians.

Significant main effect of conditions was found with Kruskal-Wallis in moving balls' hit distance from the bull's-eye ($\chi^2(5) = 17.7, p = 0.003$). The post-hoc test showed that accuracy was worse in Condition 0 when compared both to conditions 1 and 4. These results were repeated by the 3-way ANOVA followed by a post-hoc test.

Interception Rate of Moving Balls

We measured interception rate with the number of moving balls intercepted by each participant under different conditions (Figure 5). Interception rate was worst in Condition 0 as expected, followed by conditions 3, 5, and 2. Interestingly, interception rate was the highest in Condition 4, despite its monocular rendering. Condition 1 came second with its S3D.

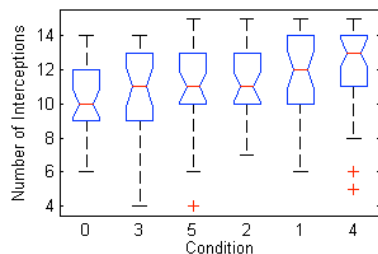


Figure 5. Number of moving balls intercepted by each of the 30 participants. The conditions are sorted by medians.

Significant main effect of conditions was found with Friedman test ($\chi^2(5) = 18.9, p = 0.002$). Post-hoc test revealed that interception rate was higher in Condition 4

when compared both to conditions 0 and 3. No other statistically significant differences were found. These results were repeated when treating binary outcome of a moving ball's interception event as a dependent variable and applying the 3-way ANOVA over all the moving balls, followed by a post-hoc test.

Acquisition Time for Static Balls

We also examined static balls' acquisition times; i.e. the elapsed time from each ball's appearance till it was struck by the participant. We found significant main effect of conditions ($F(5, 2651) = 19.9, p < 0.001$) using the 3-way ANOVA. According to a post-hoc test, acquisition time of Condition 2 was the longest (median of 3.1 seconds) of all the conditions (medians between 2.5 and 2.6 seconds). No other significant differences were found.

Kruskal-Wallis test revealed that those who reportedly used lighting cues for aiming (4 participants) had significantly better aiming accuracy ($\chi^2(1) = 4.9, p = 0.027$) in Condition 2 than those who did not use lighting cues (26 participants). The mean test order number of Condition 2 was 3.3 for the former group and 4.0 for the latter, and thus the group using lighting cues had had a little more practice with the task. This fact and the small group size means that we cannot say conclusively if using shadows for aiming improved accuracy in Condition 2.

UX Results

Based on our interviews, four participants out of 30 discovered and used the racket light sources' shadows for aiming in Condition 2. Two additional participants made the discovery, but they did not continue to aim with the shadows as they did not find them beneficial. Four other participants mentioned that the rackets' volumetric light sources improved spatial perception. In Condition 3 the balls acted as volumetric light sources, and 13 participants reported that this enhanced spatial perception. Condition 4's high-above volumetric light source was mentioned by 12 participants to improve spatial perception.

UX Interviews

The 17 dichotomous variables representing participants' perceptions about the environment were analyzed in a correspondence analysis (CA). CA provides orthogonal dimensions that are extracted in order to maximize the distance between row and column points (Greenacre, 1984). We used participants' perceptions as column variables, and the six lighting conditions as row variables.

Figure 6 presents the correspondence between the 17 perceptions and the six lighting conditions. We extracted two dimensions, with a proportion of 40% of inertia to the dimension one and 28% to the dimension two. The $\chi^2[96] = 286.4, p < .001$ supported the meaningful relationship between the row and column variables. Dimension one was named *Low Spatial Perception - Exciting but Unnatural*, and dimension two was called *Interactively Fluent -*

Visually Distracting, according to the corresponding variables. As the descriptions show, participants integrate their perceptions and actions when they are asked to describe their interactive environment.

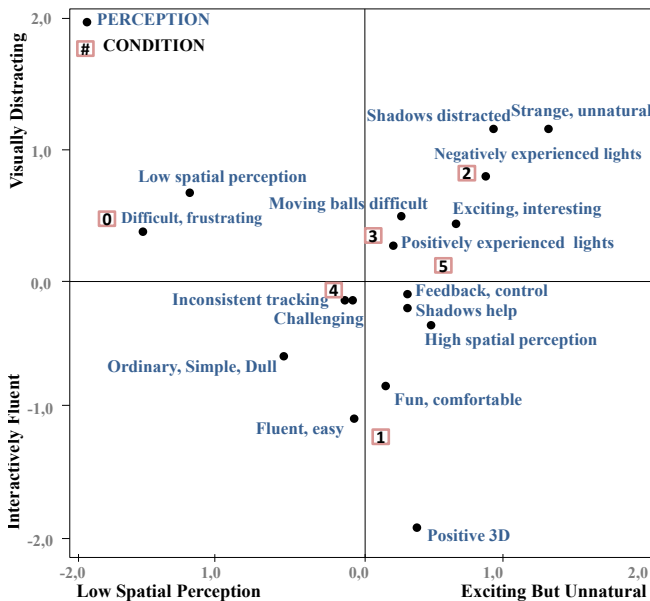


Figure 6. Correspondence analysis showing the relationships between the 17 perception variables and the six conditions.

The inspection of the Figure 6 shows that Condition 0 was perceived as spatially poor and difficult to play, mostly because of the lack of shadows. Condition 1 provided a positive 3D environment that was simple and sometimes even dull, in which playing was fluent and easy. Condition 2 was somewhat interesting and exciting, but for most of the participants the rackets' volumetric light sources and their colors were distracting and unnatural. This and the somewhat confusing shadows in Condition 2 made playing less comfortable and more difficult.

The rest of the conditions were closer to each other experientially. Although some of the participants in Condition 3 perceived the ball's volumetric light source and its colors positively, in most cases it diminished the spatial perception of the ball. This made hitting more difficult, especially with moving balls. On the contrary, volumetric light source in Condition 4 made hitting a bit fluent and easier, but still challenging. Some participants complained Condition 4 to be too simple and even dull. Because of the S3D, high spatial perception was mentioned more often in Condition 5 compared to Condition 3. However, the volumetric light in the ball divided participants' opinions: some of them liked it but some thought that it distracted them. Problems with racket tracking were reported quite equally across all the conditions, while Condition 1 received the largest number of mentions. Tracking problems may have been ignored more in other conditions because of the other, possibly stronger perceptions that caught the participants' attention.

UX Scales

First, we studied UX in the playing order of the conditions with General Linear Model Repeated Measures Anova (PASW Statistics 18, pairwise comparisons between conditions were applied with a p-value of 0.05 using Bonferroni correction). *Overall feeling* increased significantly as the experiment proceeded (Wilk's Lambda = .33, $F(5,25) = 10.10$, $p < .001$, $\eta^2 = .67$). Similarly, the evaluated *skills* increased (Wilk's Lambda = .49, $F(5,25) = 5.17$, $p < .01$, $\eta^2 = .51$). Since the participants evaluated *challenges* the same way throughout the test, they experienced a clear learning curve from coping towards mastery (Wilk's Lambda = .45, $F(5,25) = 6.01$, $p < .01$, $\eta^2 = .55$). The sense of control was the only PAD scale that increased significantly when the experiment proceeded (Wilk's Lambda = .59, $F(5,25) = 3.52$, $p < .05$, $\eta^2 = .41$). Notably, good *overall feeling*, *skills*, and sense of *control* were easier to obtain in some of the conditions than others, regardless of the order of the condition.

The six conditions were significantly different in *overall feeling* (Wilk's Lambda = .31, $F(5,25) = 11.10$, $p < .001$, $\eta^2 = .69$). Pairwise comparison shows that the *overall feeling* was significantly the lowest in Condition 0. *Overall feeling* in Condition 1 was higher compared to conditions 0, 2, and 4. Conditions 2, 3, 4, and 5 scored equally.

The flow-space (Figure 7) shows that the participants were coping in Condition 0, which was rated as significantly more challenging than all the other conditions but Condition 2. Conditions 1, 2, 3, 4, and 5 did not differ from each other in this regard (Wilk's Lambda = .54, $F(5,25) = 4.30$, $p < .01$, $\eta^2 = .46$). Moreover, in Condition 1 the participants experienced mastery. The evaluated *skills* in Condition 1 were significantly higher than in conditions 0 or 2 (Wilk's Lambda = .61, $F(5,25) = 3.19$, $p < .05$, $\eta^2 = .39$).

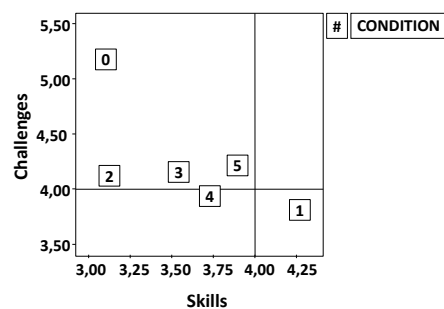


Figure 7. The means of the skills and challenges of each condition plotted in the flow-space.

PAD profiles show that participants were equally aroused across conditions (Figure 8). Condition 1 was the highest in *valence* and conditions 0 and 2 were the lowest (Wilk's Lambda = .24, $F(5,25) = 15.73$, $p < .001$, $\eta^2 = .76$). There were no difference between conditions 3, 4, and 5 in *valence*. Condition 1 was also significantly higher in the sense of *control* compared to Condition 0 (Wilk's Lambda = .60, $F(5,25) = 3.29$, $p < .05$, $\eta^2 = .40$). The three

dimensions of the PAD profile reveal how equal degree of *arousal* across the conditions affects UX differently if it is accompanied either with low *control* and *valence* (Condition 0) or high *control* and *valence* (Condition 1).

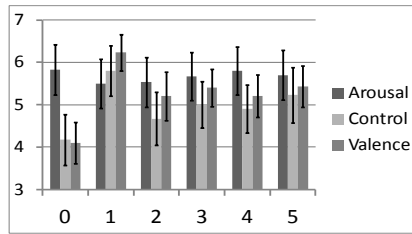


Figure 8. The means and 95% confidence intervals of the PAD profiles in the six conditions.

Taking all the qualitative and quantitative UX measures into account reveals how each condition was perceived, evaluated, and finally experienced. The UX and its causes are rather univocal in conditions 0, 1, and 2. Condition 4 has some clear characteristics of its own, but conditions 3 and 5 are difficult to distinguish. These subjective findings are in line with our gameplay performance results.

DISCUSSION

Compiling the UX and gameplay performance data that we gathered gives a rich description of the UX and gameplay in our six different conditions. Analysis of the interviews revealed both perceptual and action dimensions. Moreover, the analysis revealed the “big three” physical presence dimensions, that is spatial awareness (low spatial perception), attention (visually distracting), and realism/naturalness (exciting but unnatural) (International Society for Presence Research, 2000).

Condition 0 lacked shadows and was the worst condition in terms of UX and gameplay performance results. In Condition 2 some participants used shadows to aim the balls towards the bull’s-eye, which contributed to the static ball acquisition time that was significantly the longest. Our results were inconclusive whether this aiming improved hit accuracy. Similar to Condition 0, the *challenges* and *skills* balance in Condition 2 was towards coping. Although the ratio between arousal and control was in a better balance, playing was evaluated as uncomfortable and uneasy.

Participants achieved the best UX in Condition 1 with its typical VR setting of S3D and surface shadows; they experienced mastery (*skills* above the *challenges*) in the perceptually positive 3D environment. Condition 4 had volumetric shadows instead of S3D and it was the closest to Condition 1 in mastery and similar in other aspects: both received mentions about being fluent and easy to play, while being simple or even dull. On the other hand, conditions 1 and 4 had very different PAD profiles: equally high level of arousal was associated with lower valence and control in Condition 4, whereas in Condition 1 it was associated with higher control and valence. This difference might be related to the lower *overall feeling* in Condition 4.

Together conditions 1 and 4 were better or on par with the other conditions in gameplay performance results. No statistically significant differences were found between the two. Interestingly, Condition 4 had a significantly better interception rate compared to conditions 0 and 3, whereas the S3D-equipped Condition 1 did not. Since we did not have a test condition with monocular rendering and mere surface shadows from a high-above light source, it is not clear how much Condition 4’s volumetric shadows might have improved the result against mere surface shadows.

S3D was the only setup distinction between conditions 3 and 5. Usually S3D increases the experience of physical presence (IJsselsteijn et al., 2001; Takatalo et al., 2011), but we did not find any clear differences in either the UX or gameplay performance results. This implies that the S3D’s positive effect on depth perception in Condition 5 could have been diminished due to its volumetric lighting setup. Although the volumetric light sources used in conditions 3 and 5 were experienced mostly positively, it seems that they decreased the object-background contrast and thus weakened spatial perception (Schor and Wood, 1983). We suspect that this is why the interception rate in Condition 3 was significantly lower than in Condition 4.

Conditions 1 and 5 were the only ones with S3D. There were no significant differences between these two conditions in quantitative UX or gameplay performance results. Mastery was experienced only in Condition 1 which received the most mentions about being fluent and easy. Condition 1 also had a unique PAD profile while Condition 5’s profile resembled that of conditions 2, 3, and 4.

These pairwise comparisons between conditions 1 and 5 and conditions 3 and 4 suggest the following: volumetric light source inside the target object could 1) negatively affect the UX and 2) possibly impair gameplay performance when compared to a high-above volumetric light source. Our results imply that the high-above volumetric light source is the best choice from the three different types of volumetric lighting setups of our game in terms of UX and gameplay performance results.

We found no significant differences between conditions when examining hit accuracy of static balls. It appears that the dominant depth cue with static balls was occlusion; participants often moved their hand in xy-plane until it was in front or behind of the ball and then adjusted the z-position until the ball was hit.

Our task of striking balls towards a bull’s-eye might not be optimal for eliciting depth perception related performance differences, although many motion games have a (often simplified) version of this task. Precise aiming in our game required the racket to be swung in 3D so that its collision with the ball would result in a trajectory towards the bull’s-eye. For this the participants had to sense the 3D location of the ball and the 6D pose of the racket simultaneously. We suspect that the positioning and scaling tasks from prior

studies (Hubona et al., 1999; Wanger, Ferwerda, and Greenberg, 1992) could have led to more clear results.

Color of lighting was notably different in conditions 2, 3, and 5. Due to the high number of conditions and already long experiment duration, we decided to focus on depth cues and ignored colors as possibly contributing factors.

Lighting Guidelines for Improving Spatial Perception

Based on our results and observations during the study, we composed a short list of guidelines to aid lighting design in 3D applications where spatial perception is important:

- 1) Objects should be well contrasted against their background. Volumetric lighting and other lighting techniques can reduce this contrast and make it difficult to clearly distinguish visual border of an object (conditions 3 and 5).
- 2) Surface shadows that are meant to improve depth perception should be clearly visible. Additional illumination such as that of volumetric lighting can weaken these shadows (conditions 2, 3, and 5).
- 3) Two moving light sources can distract the user and negatively affect the UX (Condition 2). This is in line with a study by Hubona et al. (1999), who reported that two light sources can impair task performance.

CONCLUSION

In this paper we introduced novel lighting cues that can be used to assist reaching, interception, and aiming tasks, as well as enhance spatial perception. The cues are natural and blend into the rendered images because they are based on realistic rendering of volumetric lighting. This offers an alternative to traditional visual guides that are augmented over images and may appear out of place.

We presented a user study with 30 participants where the lighting cues were tested. Our results indicate that volumetric shadows can affect gameplay performance and UX positively or negatively, depending on the lighting setup. Statistically significant differences in our gameplay performance results imply that volumetric shadows can affect depth perception. A high-above volumetric light source with monocular rendering (Condition 4) did not differ from our best S3D setup with mere surface shadows (Condition 1) in terms of gameplay performance. Conversely, Condition 4 had a significantly better interception rate when compared to two other conditions whereas Condition 1 did not. Further studies are needed to quantify how much volumetric shadows can increase depth perception in monocular and S3D conditions when compared to surface shadows.

We analyzed UX with CA, Likert-scales, and flow-space metrics. Nearly half of the participants reported enhanced spatial perception in conditions with volumetric lighting. The use of volumetric light sources in our game divided the study participants' experiences however: some were

pleased with exciting and interesting lighting conditions, others were distracted by them. We found indications that a poor choice of volumetric lighting could diminish S3D's positive effects on UX and depth perception. Overall, the most pleasing game experience was achieved with S3D and surface shadows (Condition 1).

Our study sets a starting point for further research on volumetric shadows as visual guides. Future studies need to confirm our findings for applications with first-person viewpoints.

ACKNOWLEDGMENTS

This work was supported by Finnish Doctoral Program in User-Centered Information Technology (UCIT) and Helsinki Institute of Science and Technology Studies (HIST).

REFERENCES

- Arthur, K.W., Booth, K.S., and Ware, C. Evaluating 3D task performance for fish tank virtual worlds. *ACM Trans. Inf. Syst.* 11, 3 (1993), 239–265.
- Bennett, S., van der Kamp, J., Savelsbergh, G.J.P., and Davids, K. Discriminating the role of binocular information in the timing of a one-handed catch. *Experimental Brain Research* 135, 3 (2000), 341–347.
- Boritz, J. and Booth, K.S. A study of interactive 3D point location in a computer simulated virtual environment. *Proc. ACM symposium on Virtual reality software and technology*, ACM (1997), 181–187.
- Bruckner, S. and Groller, M.E. Enhancing Depth-Perception with Flexible Volumetric Halos. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2007), 1344–1351.
- Bulson, R., Ciuffreda, K.J., and Ludlam, D.P. Effect of Binocular vs. Monocular Viewing on Golf Putting Accuracy. *Journal of Behavioral Optometry* 20, 2 (2009), 31–34.
- Cruz-Neira, C., Sandin, D.J., and DeFanti, T.A. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, (1993), 135–142.
- Csikszentmihalyi, M. *Beyond boredom and anxiety*. Jossey-Bass Publishers, 1975.
- Cutting, J.E. and Vishton, P.M. Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. *Perception of space and motion* 5, (1995), 69–117.
- Dey, A. *Incomplete Block Designs*. WORLD SCIENTIFIC PUB, 2010.
- El-Nasr, M.S. and Horswill, I. Automating Lighting Design for Interactive Entertainment. *Computers in Entertainment* 2, 2 (2004), 15.
- Greenacre, M.J. *Theory and applications of correspondence analysis*. Academic press, 1984.

- Heinen, T. and Vinken, P.M. Monocular and binocular vision in the performance of a complex skill. *Journal of Sports Science and Medicine* 10, (2011), 520–527.
- Van Hof, P., van der Kamp, J., and Savelsbergh, G.J.P. Three- to eight-month-old infants' catching under monocular and binocular vision. *Human Movement Science* 25, 1 (2006), 18–36.
- Hubona, G.S., Shirah, G.W., and Jennings, D.K. The Effects of Cast Shadows and Stereopsis on Performing Computer-Generated Spatial Tasks. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 34, 4 (2004), 483–493.
- Hubona, G.S., Wheeler, P.N., Shirah, G.W., and Brandt, M. The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. *ACM Trans. Comput.-Hum. Interact.* 6, 3 (1999), 214–242.
- Jsselsteijn, W., Ridder, H. de, Freeman, J., Avons, S.E., and Bouwhuis, D. Effects of Stereoscopic Presentation, Image Motion, and Screen Size on Subjective and Objective Corroborative Measures of Presence. *Presence: Teleoperators and Virtual Environments* 10, 3 (2001), 298–311.
- International Society for Presence Research. The Concept of Presence: Explication Statement. (2000). Retrieved 30 May 2013 from <http://ispr.info/>.
- Knez, I. and Niedenthal, S. Lighting in Digital Game Worlds: Effects on Affect and Play Performance. *CyberPsychology & Behavior* 11, 2 (2008), 129–137.
- Laby, D.M., Kirschen, D.G., and Pantall, P. The Visual Function of Olympic-Level Athletes—An Initial Report. *Eye & Contact Lens: Science & Clinical Practice* 37, 3 (2011), 116–122.
- Laméris Ootech. *TNO Test for Stereoscopic Vision*. Netherlands Organization for Applied Scientific Research, 1972.
- Lang, P.J. Behavioral treatment and bio-behavioral assessment: Computer applications. In *Technology in Mental Health Care Delivery Systems*. Ablex Pub. Corp., 1980, 119–137.
- Lombard, M. and Jones, M.T. Identifying the (Tele)Presence Literature. *PsychNology Journal* 5, 2 (2007), 197–206.
- Mazyn, L., Lenoir, M., Montagne, G., Delaey, C., and Savelsbergh, G. Stereo vision enhances the learning of a catching skill. *Experimental Brain Research* 179, 4 (2007), 723–726.
- Nishita, T., Miyawaki, Y., and Nakamae, E. A shading model for atmospheric scattering considering luminous intensity distribution of light sources. *ACM SIGGRAPH Computer Graphics*, (1987), 303–310.
- Oshita, M. Motion-capture-based avatar control framework in third-person view virtual environments. *Proc. ACM SIGCHI international conference on Advances in computer entertainment technology*, ACM (2006), 2.
- Ropinski, T., Doring, C., and Rezk-Salama, C. Interactive volumetric lighting simulating scattering and shadowing. *IEEE Pacific Visualization Symposium*, (2010), 169 – 176.
- Ryan, R., Rigby, C., and Przybylski, A. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30, 4 (2006), 344–360.
- Schmidt, R.A. and Wrisberg, C.N. *Motor Learning and Performance*. Human Kinetics Publishers, 2004.
- Schor, C.M. and Wood, I. Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research* 23, 12 (1983), 1649–1654.
- Snow, M.P. and Williges, R.C. Empirical Models Based on Free-Modulus Magnitude Estimation of Perceived Presence in Virtual Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 40, 3 (1998), 386–402.
- Takala, T.M., Pugliese, R., Rauhamaa, P., and Takala, T. Reality-based User Interface System (RUIS). *3D User Interfaces (3DUI), 2011 IEEE Symposium on*, (2011), 141–142.
- Takatalo, J., Häkkinen, J., Kaistinen, J., and Nyman, G. Presence, Involvement, and Flow in Digital Games. In R. Bernhaupt, ed., *Evaluating User Experience in Games*. Springer London, 2010, 23–46.
- Takatalo, J., Kawai, T., Kaistinen, J., Nyman, G., and Häkkinen, J. User Experience in 3D Stereoscopic Games. *Media Psychology* 14, 4 (2011), 387–414.
- Takatalo, J., Nyman, G., and Laaksonen, L. Components of human experience in virtual environments. *Computers in Human Behavior* 24, 1 (2008), 1–15.
- Teather, R.J. and Stuerzlinger, W. Guidelines for 3D positioning techniques. *Proc. conference on Future Play*, ACM (2007), 61–68.
- Wanger, L.R., Ferwerda, J.A., and Greenberg, D.P. Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics and Applications* 12, 3 (1992), 44 –58.
- Wanger, L.R. The effect of shadow quality on the perception of spatial relationships in computer generated imagery. *Proc. 1992 symposium on Interactive 3D graphics*, (1992), 39–42.
- Wikipedia. List of Kinect games. (2013). Retrieved 27 June 2013 from http://en.wikipedia.org/wiki/Kinect_games.
- Wyman, C. and Ramsey, S. Interactive volumetric shadows in participating media with single-scattering. *IEEE Symposium on Interactive Ray Tracing*, (2008), 87–92.