Gaze Direction in Virtual Reality Using Illumination Modulation and Sound

Eli Ben-Joseph and Eric Greenstein

EE 267 Poster Presentation

6/1/16

Introduction

Unlike the traditional 2D environments in which a user is limited to viewing the content on the screen in front of them, virtual reality (VR) brings the user into an immersive, 3D setting. Though virtual reality allows for a much more dynamic experience, content-creators now have less control over what the user is viewing.. Storytelling, especially in gaming, is essential to the experience. Unlike 2D displays, where content was pushed to the user, in VR the user can explore the scene as they please, making the job of a content creator more difficult.

Given the plethora of distractions a user may face in a VR scene, how can a content-creator ensure that the user is looking at the right place in the scene to continue the story, but not destroy the immersive experience with obvious cues? Seamless gaze direction is a problem that VR developers face.

To approach this problem, we attempted to guide a user's gaze with two strategies: a very subtle flicker in the periphery of the field of vision, or a 3D-localized sound.

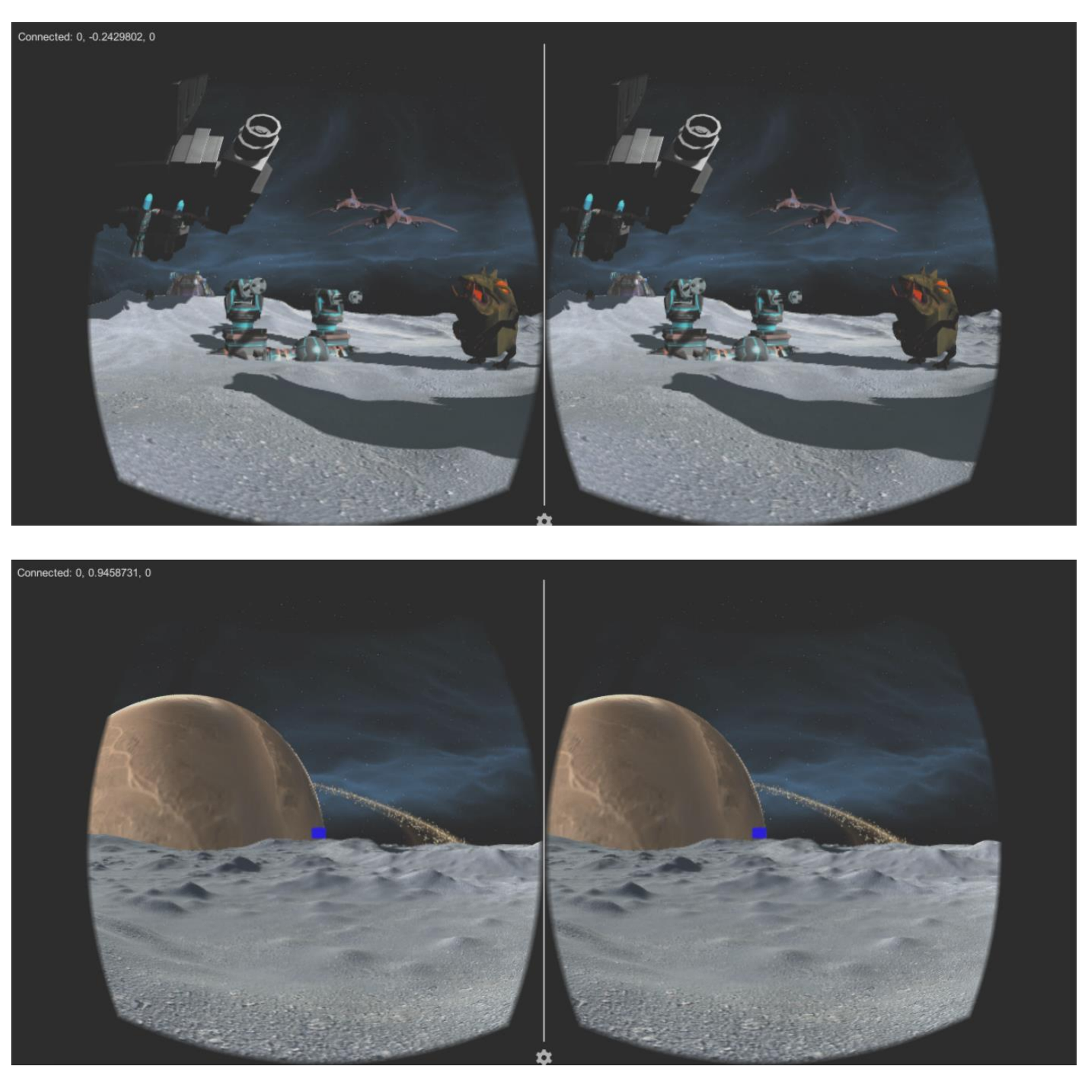

Figure 1: Experimental Scene

Methodology

Background Literature

Human gaze and the topic of gaze direction have been studied in various experiments. Factors such as image contrast, presence of faces and other informative regions, and high edge density have been known to control gaze [1, 2, 3]. Previous research has shown that illumination modulation can direct a user's gaze around a digital image [4]. There has also been substantial research on the connection between the visual and auditory systems and how gaze is mediated by sound.

However, much of this work has focused on static images. In today's digital world, there are new possibilities for gaze direction, particularly within virtual reality. In this experiment, we investigate how illumination modulation and 3D-localized sound can direct a user's gaze in virtual reality.

Experimental Setup and Procedure

All experiments were carried out on the head-mounted display (HMD) that we built during class. We tested 15 users overall. Each user was shown our scene, and selected to be in the control, flicker, or sound groups at random. An object-of-interest (a cube) was shown at a random position outside the field-of-view of the user, and the user was instructed to simply observe and explore the scene. The time it took for the user to notice the cube, as measured by the head orientation of the user, was recorded.

Flicker

Previously, we have conducted experiments demonstrating that intensity modulation, which we call flicker, can be used to direct a user around simple scenes. In this environment, flicker was implemented in the right periphery of the user's right eye, with the goal of inducing a head turn in that direction. A rectangular band of pixels, stretching the whole height of the screen, was modulated. The flicker parameters were chosen to be slightly noticeable in the periphery and were based off our previous research.

Methodology (Cont.)

Sound

To guide people to the cube, we attached a sound file to the cube in Unity, and used the built-in 3D sound model to compute stereo sound that adjusted for the user's position and orientation. A laser sound was chosen because it was fitting for our space battle scene. It played in a loop that repeated approximately every 2.5 seconds.

Results

Flicker had the fewest seconds per degree (0.0632 sec/deg), followed by control (0.1164) and sound (0.1418). Though the average time to find the cube was lowest for subjects in the flicker group, it only held a p-value of 0.18 when compared to the control group. However, given the limited amount of data gathered, this number should be taken with a grain of salt. With more data, we may find flicker to have no effect, or to have a statistically significant effect. The initial results are encouraging, but not yet conclusive. The average time to find the cube was higher for subjects in the control and sound groups. Comparing the sound vs. control groups with a t-test yielded a p-value of 0.67, and comparing the sound vs. flicker groups yielded a p-value of 0.17.

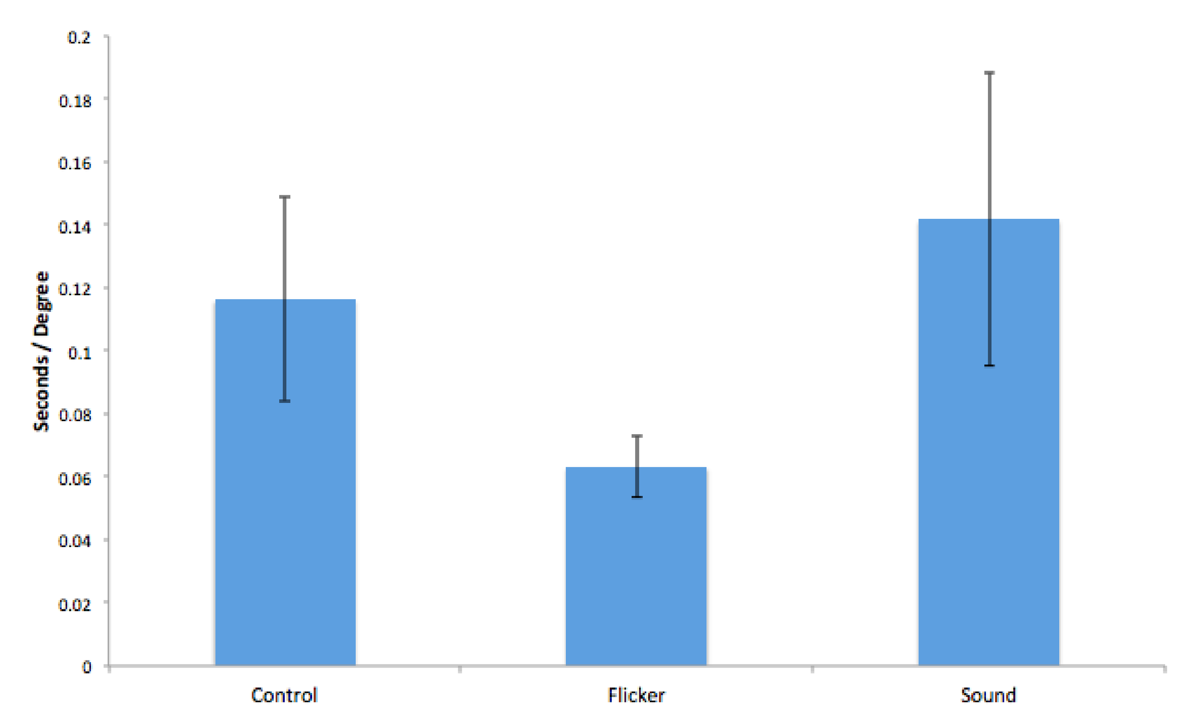

Figure 2: Average Response Time to Gaze-Directing Stimulus

Results (Cont.)

Comparison	P-value
Flicker vs. control	0.1809
Sound vs. control	0.6664
Flicker vs. sound	0.1674

Table 1: Comparison of Different Gaze Directing Techniques

Future Work

There are numerous subjects to explore going forward. First and foremost, more data should be gathered across flicker, sound, and control to determine if the improvement brought on by the flicker holds a statistically significant advantage over the other methods, and if the sound strategy is indeed no different than control. Within the flicker case specifically, more fine tuning of the parameters (size, color, and frequency) can be done to ensure that the flicker is subtle (no subject notices it) yet effective. Using a more state-of-the-art VR system would likely help, as their framerates are >60fps (from previous experiments, a minimum 60fps is required for ideal flicker modulation effects), and their orientation tracking systems will be faster and more accurate (running our IMU and Unity scene resulted in significant drop in frame rate). Within the sound case, it is likely that certain sounds will work better than others in drawing attention (e.g., human voice may be more effective at attention-grabbing according to literature [5]). Also, it is possible that certain frequencies and loop intervals are more effective than others.

References

[1] Mannan, Sabira K., Keith H. Ruddock, and David S. Wooding. "The relationship between the locations of spatial features and those of fixations made during visual examination of briefly presented images." Spatial vision 10.3 (1996): 165-188.
[2] Parkhurst, Derrick J., and Ernst Niebur. "Scene content selected by active vision." Spatial vision 16.2 (2003): 125-154.
[3] Mackworth, Norman H., and Anthony J. Morandi. "The gaze selects informative details within pictures." Perception & Psy- chophysics 2.11 (1967): 547-552.
[4] Bailey,Reynold,etal."Subtlegazedirection."ACMTransactions on Graphics (TOG) 28.4 (2009): 100.
[5] Populin, Luis C. "Human sound localization: measurements in untrained, head-unrestrained subjects using gaze as a pointer." *Experimental brain research* 190.1 (2008): 11-30.