

# Head-Turning Approach to Eye-Tracking in Immersive Virtual Environments

Andrei Sherstyuk\*  
University of Hawaii

Arindam Dey†  
University of South Australia

Christian Sandor‡  
University of South Australia

## ABSTRACT

Reliable and unobtrusive eye tracking remains a technical challenge for immersive virtual environment, especially when Head Mounted Displays (HMD) are used for visualization and users are allowed to move freely in the environment. In this work, we provide experimental evidence that gaze direction can be safely approximated by user head rotation, in HMD-based Virtual Reality (VR) applications, where users actively interact with the environment. We discuss the application range of our approach and consider possible extensions.

**Index Terms:** I.3.7 [Computer Graphics]: Three-dimensional Graphics and Realism—Virtual Reality; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

## 1 INTRODUCTION

One of the most attractive features of immersive VR is the level of control that the experimenter can impose on the environment and the users. VR systems provide an unmatched amount of information on users' activities, collected and processed in real-time by user interface modules. Practically unconstrained flexibility in creating content, along with detailed records on users' virtual experience made VR a tool of choice in many fields where human subjects are involved. These are: psychiatric research and treatment, physiological and neurophysiological rehabilitation, research on human cognition and perception, including studies on VR itself. During VR sessions, nearly all user actions may be detected, recorded and analyzed.

Because all input-output channels between the user and the environment are mediated via hardware, the amount and quality of information on user behavior depend on the choice of VR components. For immersive settings where users are allowed to move freely, HMDs and motion tracking equipment are commonly used. Depending on number of sensors or markers and the type of tracker, immersive VR systems provide accurate data on the location and orientation of the user's head and, optionally, other body parts. However, in order to track direction of the user's gaze, special eye-tracking hardware is required, which is less common. For a state of the art in eye-tracking technology, we refer readers to a recent survey by Hansen and Ji [3]. In general, adding an eye-tracking component into an HMD-based VR system often presents significant technological difficulties, due to spatial and operational constraints of the tracker.

In this paper, we present results of an experimental study in immersive VR, which suggest that under certain conditions, gaze direction may be sufficiently approximated by user head rotation, which allows to avoid using eye-tracking at all.

\*e-mail: andreis@hawaii.edu

†e-mail: Arindam.Dey@unisa.edu.au

‡e-mail: chris.sandor@gmail.com

## 1.1 Motivation and background

This work stemmed out from a number of research projects that aimed to improve the utility of HMD-based systems by making rendering gaze-sensitive. In one study [6], a view sliding algorithm was presented which shifted the viewports in screen space dynamically, towards the current point of interest. As a result, the subjects reported that their perceived field of view was larger than in control group. Also, subjects showed statistically significant improvements in tasks that involved near-field object access and manipulation. In the more recent work, an algorithm for dynamic eye convergence was presented for HMDs, which helped to reduce user efforts in fusing left and right images into one stereo view [7].

In order for both algorithms to work, the VR system needed to estimate the gaze direction, in real time. It turned out, that in both cases, that direction could be successfully replaced by head rotation, available from the regular motion tracker.

## 2 THE HYPOTHESIS: HEAD ROTATION $\approx$ GAZE DIRECTION

We base our hypothesis on two observations:

### 2.1 The restraining effect of "tunnel vision"

The width of human field of view (FOV) extends approximately 150° horizontally (60° overlapping, 90° to each side) and 135° vertically. With a few exception of high-end panoramic HMDs, most commercially available HMD models have a FOV ranging from 40 to 60 degrees diagonally. Examples include: *nVisorSX* and *V8* HMDs with 60°, *ProViewXL* with 50°, and a popular low-cost *eMagine* HMD with 40°. Low FOV values result in what is called "tunnel vision" effect, which is regarded as one of the most objectionable drawbacks of HMD-based VR systems. However, in our case, this deficiency turns into an advantage. When viewing the scene through a narrow HMD window, VR users are forced to rotate their head instead of moving their eyes.

### 2.2 Fixating on a current task

Studies in experimental neurophysiology show that human eyes always focus and converge on task-relevant objects, such as a hand, hand-held tools or locations of tool application [1]. Converging eyes in real-life brings the object of interest into the center of the field of view. When projected onto retina, the object's image falls onto a special area where most photoreceptors are located and spatial resolution is highest, which ensures the best viewing conditions. Similarly, in VR the best viewing conditions are achieved when both left and right images of the object of interest are located in the center of the HMD displays. That can only happen when the user rotates his or her head towards the object.

Basing on these arguments, we propose to approximate the user's gaze direction by orientation of his or her head. This approximation should remain valid as long as there is an active point of fixation on scene, such as a virtual hand, that is continuously being used for some task.

## 3 THE EXPERIMENT

The verification of this hypothesis came from a recent experimental study, conducted in September 2011 at University of Adelaide, Australia. The goal of the study was to measure the effects of dynamic eye convergence described in [7]. During the experiments,

the subjects were asked to spend 10 minutes on a virtual beach, catching butterflies that were flying around the scene, making occasional stops for rest at random locations. A successful capture was detected when user's virtual hand remained in continuous contact with the butterfly's wings for 2 seconds, while the butterfly was at resting position (see Figure 1.) All capture events were logged into a file for later processing, recording time and location of the user hand during capture in camera space. Equipment used: Flock of Birds motion tracking system from Ascension, for tracking user hand and head, running in standard range mode (4 feet radius); a stereo VH-2007 HMD from Canon, Ubuntu Linux PC, with Flatland open source VR engine [2].

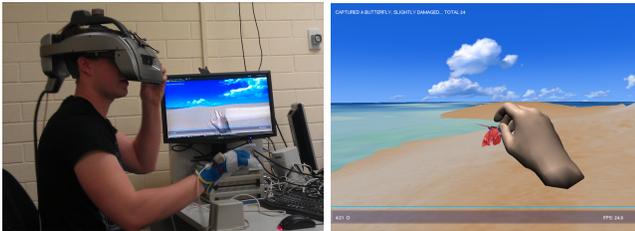


Figure 1: Left: Reenacted experimental session. The subject is wearing an HMD and a glove with motion sensors attached, looking for butterflies. The monitor displays the first person view of the scene. Right: a captured butterfly. Note that in both cases the hand is centrally positioned. Photo used with permission.

#### 4 RESULTS

We have processed session log files obtained from 12 subjects, with total of 425 recorded capture events. Figure 2 shows the distribution of azimuthal and elevation angles from the  $(0, 0, -1)$  direction in camera space, which corresponds to the head rotation in the word space. It turned out that most events happened near the center of the visual field. Notably, the relatively wide FOV of our HMD ( $60^\circ$  horizontal,  $47^\circ$  vertical), gave users a large "room" for intra-frame eye movements. Nevertheless, for high-precision hand manipulation, the users chose to rotate their heads instead, to bring the objects of interest close to the center of the viewable area, which agrees with our initial hypothesis.

#### 5 DISCUSSION

The proposed approach seems to be well suited for VR applications where users are expected or required to interact with objects using the virtual hand metaphor. The objects may be located at close range, accessible with traditional VR hand, or at large distances, if Go-Go extension of virtual hand is used [5]. In addition, for this approach to work, users must be able to move their head and hands freely, which is typical for many VR applications. However, in some cases user motion is restricted. One example is a VR-based pain reduction system [4], installed inside a Magnetic Resonance Imaging (MRI) scanner. During VR sessions, subjects are completely immobilized, because all motions over 1 mm in magnitude invalidate MRI scan data. Another class of VR applications that is unlikely to benefit from the proposed approach is free-style exploration, when users do not have objects of fixation and view the scene in a "sight-seeing" fashion.

Besides the virtual hand, there exist other metaphors that allow to pinpoint the current object of fixation. One of them is a cross-hair object, that can either slide freely across the screen space or remain fixed in the center. In the latter case, the user head rotation will strictly coincide with the gaze direction.

To conclude, we want to emphasize that every VR system is unique with respect to hardware components, scene content and user tasks. In order to optimize the hardware setup or make it

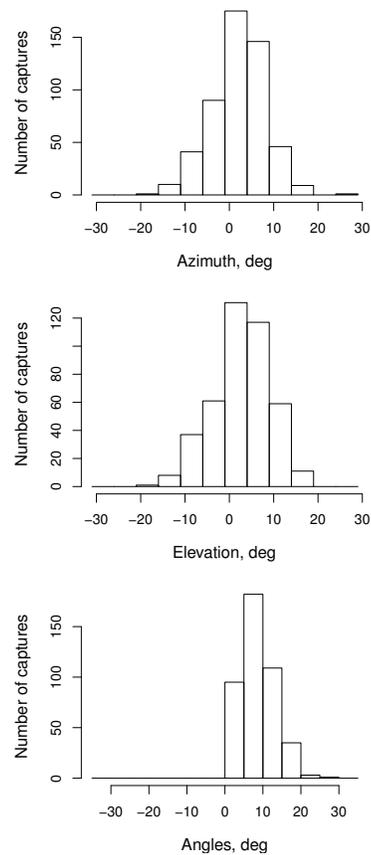


Figure 2: Distribution of viewing angles to hand locations, in camera space. Data collected from 12 subjects. Total number of samples 425. The slight shift towards +X and +Y directions is likely due to the fact the all subjects were right-handed and captured butterflies from the upper-right corner.

more cost-effective, it is worth considering if gaze direction (when required by research protocol) can be replaced by head rotation, while staying within allowed error bounds. The practice shows that in many cases, such approximation is quite sufficient.

#### REFERENCES

- [1] B. Biguer, M. Jeannerod, and P. Prablanc. The coordination of eye, head and arm movements during reaching at a single visual target. *Experimental Brain Research*, 46:301–304, 1982.
- [2] Flatland. *Albuquerque High Performance Computing Center (AHPCC), The Homunculus Project*. <http://www.hpc.unm.edu/homunculus>, 2002.
- [3] D. Hansen and Q. Ji. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on pattern analysis and machine intelligence*, 32(3):478–500, March 2010.
- [4] H. Hoffman, T. Richards, A. Bills, T. Oostrom, J. Magula, E. Seibel, and S. Sharar. Using fMRI to study the neural correlates of virtual reality analgesia. *CNS Spectrums*, 11:45–51, 2006.
- [5] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The Go-Go interaction technique: non-linear mapping for direct manipulation in vr. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*, UIST '96, pages 79–80, 1996.
- [6] A. Sherstyuk, C. Jay, and A. Treskunov. Impact of hand-assisted viewing on user performance and learning patterns in virtual environments. *The Visual Computer*, 27:173–185.
- [7] A. Sherstyuk and A. State. Dynamic eye convergence for head-mounted displays. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, VRST '10, pages 43–46, 2010.