Ultra Low Cost Eye Gaze Tracking for Virtual Environments

Matthew Swarts[1] and Jin Noh[2]

[1]Georgia Institute of Technology, College of Architecture, Atlanta, GA, USA
matthew.swarts@coa.gatech.edu
[2]Gwinnet School of Mathematics, Science, and Technology, Lawrenceville, GA, USA
noh.jin2014@gmail.com

Abstract. In this paper we present an ultra-low cost eye gaze tracker specifically aimed at studying visual attention in 3D virtual environments. We capture camera view and user eye gaze for each frame and project vectors back into the environment to visualize where and what subjects view over time. Additionally we show one measure of calculating the accuracy in 3D space by creating vectors from the stored data and projecting them onto a fixed sphere. The ratio of hits to non-hits provides a measure of 3D sensitivity of the setup.

Keywords: Low Cost, Eye Tracking, Virtual Environments.

1 Introduction

In this paper we present a method for constructing an ultra-low cost eye gaze tracking system aimed at use in visual attention studies within 3D virtual environments. The motivation behind the development of low cost eye gaze tracking systems for virtual reality lies in the use of spatial analysis, human behavior research, and neuroscience to uncover how the structure and material of a space affects the understanding and cognition of space. Virtual reality has been used in the field of space syntax [1] to study the paths people take in a new environment. Eye tracking could be used to look deeper into the motivating factors of features of the space that entice people to take specific directional queues. Virtual reality has also been used in neuroscience [2] to replicate maze and visual puzzle experiments traditionally performed on mice. Eye tracking has been used in these instances as well. In museums, artifacts have been shown to be grouped both spatially as well as visually to create a rich set of connections [3]. These types of visual pairings and groupings could also be better understood through the use of eye gaze tracking. When the building no longer exists, the building was never constructed, the layout of the museum or building has changed through renovation, or for testing hypothetical spatial structures, eye tracking can be combined with virtual reality to gain understanding beyond what is available in the built environment.

Tracking eye movements allows us to see a fairly involuntary human response to an environment. Eye movements provide a more quantitative method of studying

R. Shumaker (Ed.): VAMR/HCII 2013, Part I, LNCS 8021, pp. 94–102, 2013.

human behavior and perception [4]. The eye moves in patterns of fixations and rapid saccades based on what is being view, when, and where. Capturing these movements allows us to see underlying structures in complex objects and patterns in the surrounding environment.

Eye tracking is used for a myriad of purposes including advertising and marketing, training, assistive technology, and psychological studies. Inspection and training both in the physical environment [5] and in 3D virtual environments [6, 7] can utilize eye tracking. Within the realm of virtual environments, eye tracking is used for object manipulation [8, 9] and user movement. It can also be used to show predicted eye movements based on visual attention cues [10]. Experiential effects, such as depth of field, can be improved using eye tracking [11]. A user's anticipation of a turn in active navigation of a virtual environment can be determined by observing eye movements [12]. Salient maps of features in an environment can be used to improve the accuracy of eye trackers [13-15] by applying attention theory [16]. Several devices also employ two eye trackers for binocular tracking to determine precise 3D location [6, 7, 17] and user movement [18].

Eye trackers are used for people with motor impairments or other disabilities, allowing them to interact with the physical and virtual environment [17, 19]. Along these lines, others have developed methods of reducing the size and weight of portable eye trackers [17, 20] and for making low-cost eye trackers [17, 21, 22] to increase the general accessibility of the technology.

2 System

Our system is composed of eye tracking hardware, eye tracking software, and a 3D virtual environment model with network messaging and analytics for processing and post-processing of the input data streams.

2.1 Eye Tracking Hardware

The eye tracking hardware is made from a camera, a lens, a filter, a clamp, a helmet mount, and infrared LEDs. The selection of each element was a balance among cost, availability, weight, and expected accuracy. The overall design was a helmet or head mounted eye tracker in which the image of the eye could be maximized for better accuracy. Other designs, such as a remote monitor mounted tracker, were considered, but the distance is an issue for maintaining spatial accuracy with the limited hardware.

Construction of an eye tracker generally requires a camera or two. For virtual reality that is presented on a single monitor, so only one camera is necessary for most setups. While binocular eye trackers exist for some virtual reality setups, they are not very useful for a single screen without true 3D capability. Tracking both eyes allows the capture of the user's focal plane. However, in using a single display screen, the focal plane can be assumed to be the screen itself. We selected the Sony PlayStation Eye camera, as it is possible to get speeds up to 187 frames per second (fps) at a resolution of 320x240 or 60fps at 640x480. The higher frame rate allows us to test higher temporal resolutions in capturing human eye movements than traditional web cameras running at 30 frames per second. The high frame rate capability of the PlayStation eye was developed by Sony by the maximizing the use of the USB 2.0 bandwidth limits. This limitation of data bandwidth is what distinguishes most low cost commercial off the shelf product (COTS) USB 2.0 color web-cameras from higher-end more expensive (>$500 USD) industrial application [23] single band cameras which use IEEE 1394 FireWire, Camera Link, or the new USB 3.0 specification with higher data transfer rates.

The camera was disassembled, and the plastic cover removed. The plastic lens holder was removed, and replaced by a lens mount with threading for m12 lenses. An infrared (IR) band-pass filter was inserted into the new camera lens mount to only allow IR light onto the imaging array. An 8mm optical lens was screwed into the new lens mount. This lens provided a larger, zoomed in view of the user's eye when mounted to the helmet, allowing more space for pupil view analysis, while keeping some distance from the user's eye.

Infrared light is typically used in eye tracking, because it is not in the visible spectrum and does not interfere with the user's vision. Additionally the human iris reflects infrared light, making the iris appear lighter regardless of the visible eye color.

Fig. 1. Camera assembly and helmet mount

This provides a way to easily discern the pupil from the iris. The pupil center can be more easily determined, which is a feature used for comparison in most eye tracking methods. An infrared light source was constructed from IR LEDs to illuminate the eye for use in indoor environments.

There are several options when it comes to mounting the camera assembly for monitoring computer display interaction. The typical setups include mounting to the bottom of the monitor as a remote camera, and mounting to the user's head using glasses or a helmet. We chose a head mounted approach using a helmet in order to increase the spatial resolution of the eye movements and to minimize calibration issues associated with large zoom lenses. Metal alligator clips on rods were taken from an electronics magnifier and used to hold the camera assembly onto the helmet while allowing for the slight adjustments needed for different users.

2.2 Eye Tracking Software

The core of the system relies of the work of the ITU GazeGroup [24]. The ITU Gaze Tracker is an open source eye tracker developed at IT University of Copenhagen [22] aimed at providing low cost alternatives to commercial systems and making it more accessible. The software is extremely flexible in terms of input hardware, hardware setup, and feature tracking. Their system also allows for control over the sensitivity of each feature. Additionally, there are several levels of calibration available. The element that is most important for our system is the ability to stream out the view location in screen coordinates over the network via Universal Datagram Protocol (UDP).

In addition to the ITU Gaze Tracker software, we also used drivers developed by Code Laboratories [25] specifically for the Sony PlayStation 3 Eye Camera. These drivers provide access to the higher frame rates available through the camera.

2.3 Virtual Environment

The Unity Game Engine [26] is a popular 3D video game engine used for developing and publishing video games on many platforms. It is free for non-commercial use, and extremely flexible with three powerful scripting languages, and the ability to bind to external libraries. We developed a set of scripts to capture the view points as well as the position, location, and field of view of the user along with the current system timestamp. The data is saved to a file, which can be loaded, analyzed, and visualized.

The user's view is captured by the ITU Gaze Tracker, which sends the coordinates to Unity via UDP. During each frame of the virtual environment, the position, rotation, and field of view of the user's camera is captured. If there has been a new view coordinate received since the last frame, then the 2D screen coordinates from the ITU Gaze Tracker are concatenated with the 3D camera data, and the entire data set is stored to file. This ensures that there is always 3D camera data associated with a 2D view coordinate. For this association we assume that the virtual environment operates at a higher frequency than the eye tracker. In our tests, this was the case, as the environments were highly optimized, and the limiting factor was the frequency of the eye tracking camera. However, we do record both the camera data and the eye tracking data separately as well for cases where a custom post-synchronization step is necessary.

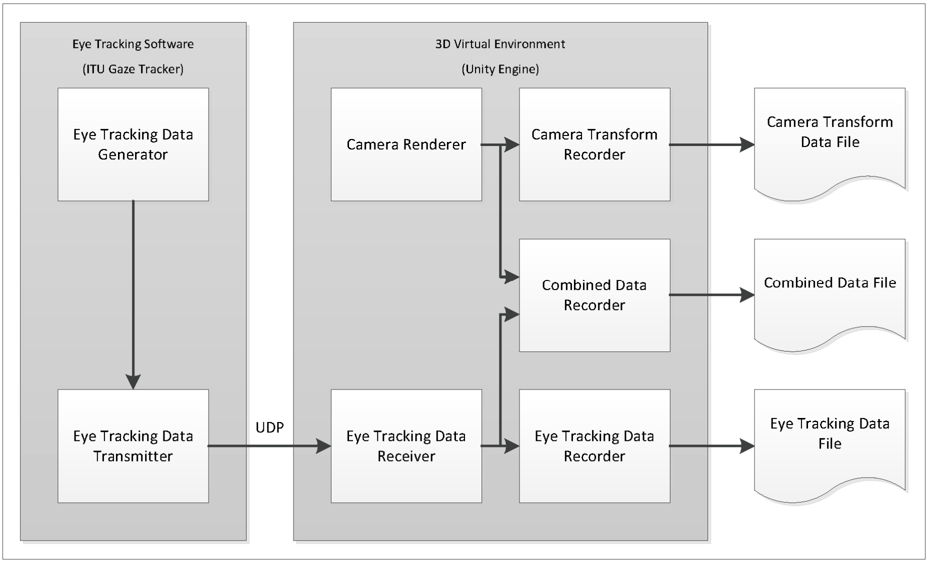

Fig. 2. Overall software architecture. Eye tracking software passes eye tracking data through UDP. The 3D Virtual Environment receives the eye tracking data and records it to a file. It also captures the current camera matrix and saves the 3D position and view data to a file.

2.4 Visualization

One typical method for analyzing eye tracking data in two dimensions is through the use of a heat map. A heat map of eye movements or mouse movements is an aggregation of time spent in each location of the 2D space. This is accomplished by applying a circle with a radial gradient of transparency to each recorded location with an intensity of the duration spent in that location. This 2D visualization method is not particularly well suited to 3D space without much more data points to aggregate.

The visualization in Figure 3 is a 3D virtual environment constructed using Autodesk 3D Studio Max and Adobe Illustrator. Our initial motivations prompted the use of virtual museums for initial testing. As an example we selected Tadao Ando's Pulitzer Foundation of the Arts, for its ability to produce alternative visual and spatial interpretations using space, light, and color [27].

To visualize the eye tracking data in 3D space we use the camera projection matrix, which includes the position, rotation, and field of view. We then take the 2D view coordinate data and using the recorded dimensions and aspect of the display, we calculate a projection vector. Using a screen-to-world operation we project the vector into the 3D environment from the camera location until it hits a surface. At that hit location in 3D space, we create a colored sphere. The center of the sphere corresponds to where the user was looking at that point in time. The radius of the sphere corresponds to the distance between the camera and the hit location. Smaller spheres represent smaller viewing distances, while larger spheres represent larger viewing distances. Lastly, the color ranges between a spectrum of two hue values in the Hue-Saturation-Value (HSV) color space to represent the time within the trial.

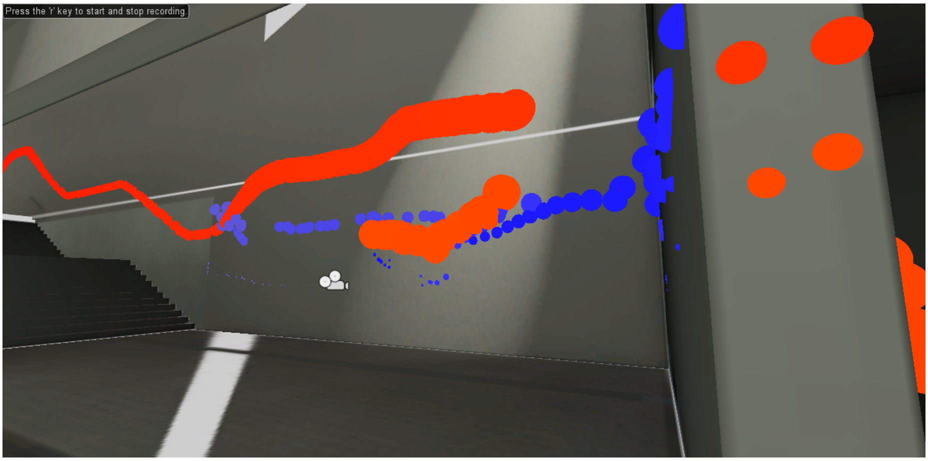

Fig. 3. Screen shot of a visualization of the eye tracked data over one user session in a 3D virtual environment

3 3D Accuracy Testing

As a preliminary test for accuracy, we designed a test environment in which the subject is asked to focus on a specific object. The environment, see Figure 4, consists of a rectangular room 24m wide by 24m in depth and 4m in height. A red sphere of 0.5m diameter is placed at the center of the room. The camera is set at an eye height of 1.8m. As the subject moves around the space, keeping the red sphere in view, the user is asked to keep their eyes on the sphere.

After a user has moved through the space, viewing the center sphere, we are able to project rays from each recorded view to a location in the space. By varying the size of the center sphere, we are able to see and count what ratio of hits contacted with the sphere, and which did not. Since the user continuously views the center sphere, the trajectory of the ray should tend to intersect the sphere, but with variance due to the accuracy of the eye tracking, and of the ability of the subject to remain in position. Figure 5 shows a screen capture of projections from 1 subject trial, using 5 different size diameter center spheres. The ideal case is that all of the hit spheres would be attached to the center sphere when its diameter is equal to 0.5m, the same as in the experiment. As we increase the sphere size, we get a sense of the error falloff in three dimensions.

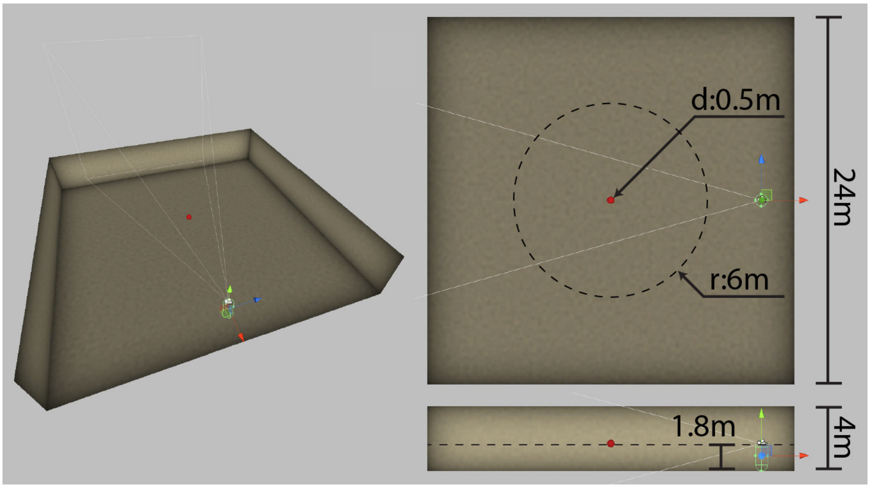

Fig. 4. The layout and dimensions of the test environments with a sphere located at the center of a large rectangular room

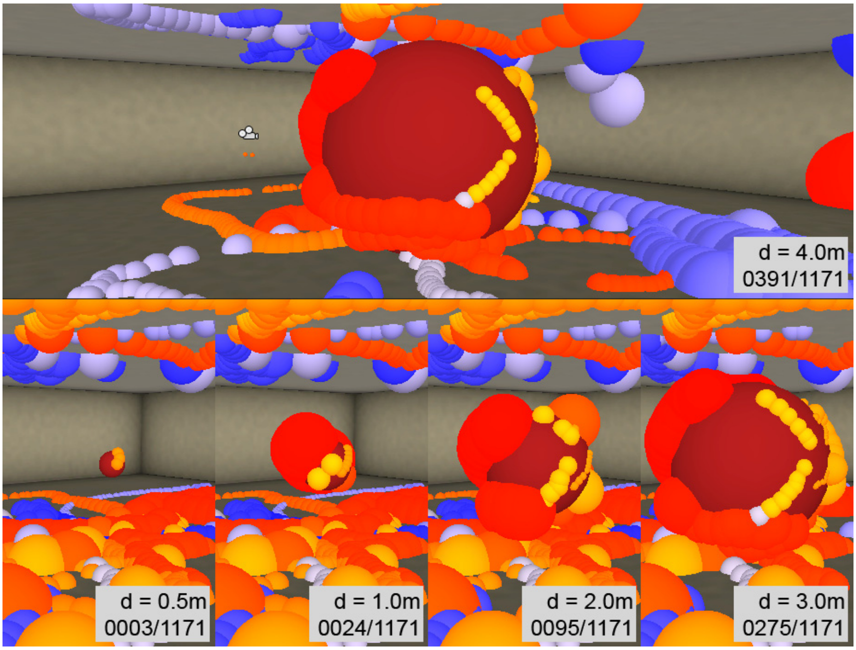

Fig. 5. A sequence of screen captures of projected spheres of increasing size for one subject trial. Larger spheres identify a larger error capture zone. 391 out of 1171 samples were captures by a sphere of diameter 4.0m.

4 Conclusions and Future Work

We have demonstrated that an eye tracker for virtual reality can be developed at low cost, which can be used for some areas of research within 3D virtual environments. Additionally we provided a new measure for 3D sensitivity to indicate the accuracy of the system for a particular eye tracker setup.

Future work could incorporate an open-source model of 3D printed glasses specifically for the PS3 Eye Camera. Additionally the integration of salient maps to lock onto the most likely candidates for visual attention, may be useful for better accuracy in low cost systems. Lastly more measures of 3D movement, including angular velocities and angular accelerations, could be incorporated into a more robust regression model to determine which aspects of interaction with the 3D virtual environment are likely to cause the most errors in hit detection for low cost systems.

References

1. Dalton, R.C.: The secret is to follow your nose: route path selection and angularity. Environment & Behavior 35, 107–131 (2003)
2. Moffat, S.D., Resnick, S.M.: Effects of age on virtual environment place navigation and allocentric cognitive mapping. Behavioral Neuroscience 116, 851–859 (2002)
3. Zamani, P.: Views across boundaries and groupings across categories (electronic resource): the morphology of display in the galleries of the High Museum of Art 1983-2003. Georgia Institute of Technology, Atlanta (2008)
4. Yarbus, A.L.: Eye movement and vision. In: Haigh, B. (trans.) Plenum Press, New York (1967)
5. Megaw, E.D., Richardson, J.: Eye movements and industrial inspection. Applied Ergonomics 10, 145–154 (1979)
6. Gramopadhye, A.K., Melloy, B.J., Nair, S.N., Vora, J., Orhan, C., Duchowski, A.T., Shivashankaraiah, V., Rawls, T., Kanki, B.: The use of binocular eye tracking in virtual reality for aircraft inspection training. Int. J. Ind. Eng.-Theory Appl. Pract. 9, 123–132 (2002)
7. Duchowski, A.: Eye tracking methodology: Theory and practice. Springer (2007)
8. Bowman, D.A., Hodges, L.F.: An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In: Proceedings of the 1997 Symposium: Interactive 3D Graphics, vol. 35 (1997)
9. Tanriverdi, V., Jacob, R.: Interacting with Eye Movements in Virtual Environments, pp. 265–272. Association for Computing Machinery (2000)
10. Peters, R.J., Itti, L.: Computational Mechanisms for Gaze Direction in Interactive Visual Environments, pp. 27–32. ACM (2006)
11. Hillaire, S., Lecuyer, A., Cozot, R., Casiez, G.: Using an Eye-Tracking System to Improve Camera Motions and Depth-of-Field Blur Effects in Virtual Environments. In: Virtual Reality Conference, VR 2008, pp. 47–50. IEEE (2008)
12. Hillaire, S., Lécuyer, A., Breton, G., Corte, T.R.: Gaze behavior and visual attention model when turning in virtual environments. In: Proceedings of the 16th ACM Symposium: Virtual Reality Software & Technology, vol. 43 (2009)
13. Hillaire, S., Breton, G., Ouarti, N., Cozot, R., Lecuyer, A.: Using a Visual Attention Model to Improve Gaze Tracking Systems in Interactive 3D Applications. Comput. Graph. Forum 29, 1830–1841 (2010)

14. Hillaire, S., Lecuyer, A., Regia-Corte, T., Cozot, R., Royan, J., Breton, G.: Design and Application of Real-Time Visual Attention Model for the Exploration of 3D Virtual Environments. IEEE Trans. Vis. Comput. Graph 18, 356–368 (2012)
15. Lee, S., Kim, G.J., Choi, S.: Real-Time Tracking of Visually Attended Objects in Virtual Environments and Its Application to LOD. IEEE Trans. Vis. Comput. Graph. 15, 6–19 (2009)
16. Treisman, A., Gelade, G.: A feature-integration theory of attention. In: Wolfe, J., Robertson, L. (eds.) From perception to consciousness: Searching with Anne Treisman, pp. 77–96. Oxford University Press, New York (2012)
17. Abbott, W.W., Faisal, A.A.: Ultra-low-cost 3D gaze estimation: an intuitive high information throughput compliment to direct brain-machine interfaces. J. Neural Eng. 9 (2012)
18. Munn, S.M., Pelz, J.B.: 3D point-of-regard, position and head orientation from a portable monocular video-based eye tracker. Eye Tracking Research & Application 181 (2008)
19. Adjouadi, M., Sesin, A., Ayala, M., Cabrerizo, M.: Remote eye gaze tracking system as a computer interface for persons with severe motor disability. In: Miesenberger, K., Klaus, J., Zagler, W.L., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 761–769. Springer, Heidelberg (2004)
20. Babcock, J.S., Pelz, J.B.: Building a Lightweight Eyetracking Headgear, pp. 109–114. Association of Computing Machinery, New York (2004)
21. Kassner, M., Patera, W.: PUPIL: constructing the space of visual attention. Dept. of Architecture, vol. Masters, pp. 181. Massachusetts Institute of Technology (2012)
22. San Agustin, J., Skovsgaard, H., Mollenbach, E., Barret, M., Tall, M., Hansen, D.W., Hansen, J.P.: Evaluation of a low-cost open-source gaze tracker. Eye Tracking Research & Application 77 (2010)
23. Point Grey Research, `http://ww2.ptgrey.com/`
24. ITU GazeGroup, `http://www.gazegroup.org/`
25. Code Laboratories > CL Studio Live, `http://codelaboratories.com/`
26. Unity - Game Engine, `http://unity3d.com/`
27. Bafna, S., Losonczi, A., Peponis, J.: Perceptual Tuning of a Simple Box. In: Eighth International Space Syntax Symposium (2012)