Binocular Eye-Tracking for the Control of a 3D Immersive Multimedia User Interface

Nikolaos Sidorakis[*] George Alex Koulieris[†] Katerina Mania[‡]

Technical University of Crete
Chania, Greece

ABSTRACT

In this paper, we present an innovative approach to design a gaze-controlled Multimedia User Interface for modern, immersive headsets. The wide-spread availability of consumer grade Virtual Reality Head Mounted Displays such as the Oculus Rift^{TM} transformed VR to a commodity available for everyday use. However, Virtual Environments require new paradigms of User Interfaces, since standard 2D interfaces are designed to be viewed from a static vantage point only, e.g. the computer screen. Additionally, traditional input methods such as the keyboard and mouse are hard to manipulate when the user wears a Head Mounted Display. We present a 3D Multimedia User Interface based on eye-tracking and develop six applications which cover commonly operated actions of everyday computing such as mail composing and multimedia viewing. We perform a user study to evaluate our system by acquiring both quantitative and qualitative data. The study indicated that users make less type errors while operating the eye-controlled interface compared to using the standard keyboard during immersive viewing. Subjects stated that they enjoyed the eye-tracking 3D interface more than the keyboard/mouse combination.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality;

1 INTRODUCTION

Virtual Reality (VR) is soon to become ubiquitous. The wide-spread availability of consumer grade VR Head Mounted Displays (HMDs) such as the Oculus Rift^{TM} transformed VR to a commodity available for everyday use. VR applications are now abundantly designed for recreation, work and communication. However, interacting with VR setups requires new paradigms of User Interfaces (UIs), since traditional 2D UIs are designed to be viewed from a static vantage point only, e.g. the computer screen [2]. Adding to this, traditional input methods such as the keyboard and mouse are hard to manipulate when the user wears a HMD. Using a keyboard and a mouse while immersed in a VR HMD is an erroneous extension of the desktop paradigm to VR, constituting a fundamental challenge that needs to be addressed [2].

Recently, various companies (e.g. SensoMotoric Instruments^{TM}, [10]) announced an eye-tracking add-on to the Oculus Rift Development Kit 2 (DK2) HMD. Novel, immersive 3D UI paradigms embedded in a VR setup that is controlled via eye-tracking can now be designed, implemented and evaluated. Gaze-based interaction is intuitive and natural, providing a completely immersive experience to the users. Tasks can be performed directly into the 3D spatial context without having to search for an out-of-view keyboard/mouse. Furthermore, people with physical disabilities, already depending on technology for recreation and basic communication, can now benefit even more from VR.

In this paper, we present an innovative approach to design a gaze-controlled Multimedia User Interface (MUI) [11] for a modern eye-tracking capable HMD. User fixations control the MUI. An on-screen cursor's orientation and position is directly manipulated by the gaze data. New types of immersive applications can be developed by employing this interaction paradigm. In our prototype implementation, we have developed six applications which cover commonly operated actions of everyday computing such as mail composing, photo viewing, music playing and gaming. Our approach is applicable to most 3D accelerated devices having a standard High-Definition Multimedia Interface (HDMI) port to drive a HMD (mobiles, tablets, laptops and desktops). We evaluate our system by conducting both a quantitative and a qualitative study.

[*]e-mail: nik.sidorakis@gmail.com
[†]e-mail: gkoulieris@isc.tuc.gr
[‡]e-mail: k.mania@ced.tuc.gr

2 RELATED WORK

2.1 3D User Interfaces

UI design for 2D applications is based on fundamental principles and best practices formed after many years of research. However, user interaction in a 3D spatial context introduces constraints due to the multiple degrees of motion freedom, requiring novel interaction metaphors such as "fly" and "zoom". These metaphors are not applicable in a standard 2D interface [1]. Bowman et al. [2] report an overview of metaphors, conventions and best practices for 3D UI design. 3D UIs are an integral part of Virtual Environments (VEs) in many interactive systems. In this work, we propose an eye-tracking paradigm for the control of modern 3D UIs deployed on commercial HMDs.

2.2 Eye-tracking as an Input Device

Eye tracking has been used in the past as an input device to interact with a 2D UI [4, 5, 12], but not for an immersive 3D MUI. By moving their eyes, users can manipulate an on-screen cursor to point virtual items on the interface and then activate them via prolonged fixations or blinking. Previous research investigated eye-tracker based interfaces for disabled users [4, 5, 12]. Physically disabled users can benefit the most from VR technology since they usually greatly depend on computer aid for recreation or basic communication. However, most computer interfaces for the disabled are neither inexpensive nor easily accessible systems. Eye-tracking has also been used as a context-sensitive help system when a user had difficulty to comprehend a text during reading [9]. A head-mounted eye tracker has been used for interaction with real world planar digital displays [7]. A common issue when using an eye-tracking interface as an input device is known as the Midas' touch problem. Certain eye movements are involuntary and accidental interface activation is frequent [5]. Speech recognition has been used in the past to signify an event [6], however, it required accurate synchronization of gaze and speech data streams in order to be reliable. In our eye-tracking interface we deal with the Midas' touch issue

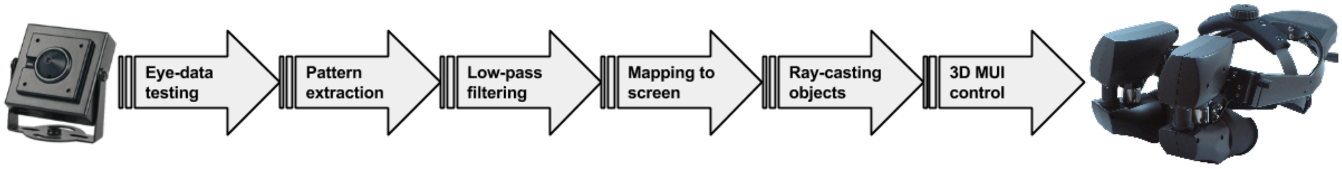

Figure 1: A schematic diagram of the proposed system, introducing principal software components.

by employing an additional mechanical input (switch) to signify a selection in the immersive environment.

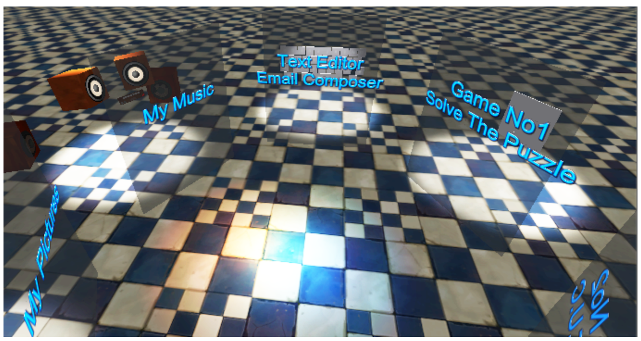

Figure 2: The application selector of the proposed 3D MUI. The user is placed at the center.

3 An Eye-Tracking 3D MUI

3.1 Design

In this section, the 3D user interface design and implementation will be presented. The 3D MUI elements are considered to be placed at static locations in the virtual space while a moving observer may interact with them from various locations. For our prototype implementation, the designed applications are centered around a main pivot point while the observer turns around and selects one of the applications (Figure 2). The project is implemented in Unity3D™ and deployed on a stereo NVisor™ SX111 HMD by setting up a virtual stereo camera rig. The rig simulates a virtual head model having two eyes and a neck, similar to the Oculus Rift head tracker implementation, however, without the positional 6DoF tracker [8]. The Inter-Pupillary Distance (IPD) can be individually adjusted for each user.

3.2 Implementation

Our system is comprised of six principal software components (Figure 1). The first, is a raw eye data calibration component performing clear pupil and glint tests, signal smoothing and filtering to eliminate noise. The second component is an eye scan pattern extraction system indicating the direction of the eye movement, fixations and blinks. The third, identifies that a clear glint signal was located and performs a low pass filtering to avoid flicker. The fourth component maps movements from eye space to screen coordinates. The fifth performs ray-casting over the 3D menus to identify fixated items. Finally, the 3D MUI receives data and executes control algorithms for cursor motion manipulation and menu item highlighting. Supplemental software components include a head tracker manager that provides 3DoF data of head movement and a SQLite database used for event logging and statistics.

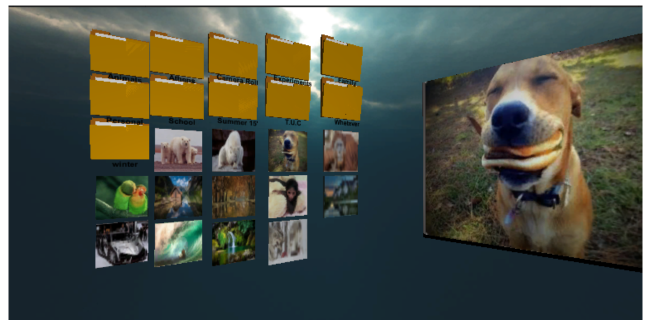

Figure 3: The eye-controlled Photo Gallery.

3.3 Developed Applications

Six applications based on the eye-tracking 3D UI paradigm have been implemented. User gaze substitutes the mouse pointer while the mechanical switch acts as a selector for the UI. A Photo Gallery allows the user to browse the pictures folder and manipulate photos (Figure 3). The application searches through an image folder and visualizes sub-folders and files on a virtual 3D slide-show supporting .jpg, .gif, .tiff, and .bmp formats. A 3D Music Player explores the user's music folder and exposes virtual 3D geometry for audio and playback control (Figure 4). The virtual speakers visualize the music and vibrate according to the music tempo. A 3D Email Composer consists of a 3D custom-made keyboard and a 3D email form (Figure 5). The keyboard is based on a standard mobile device keyboard layout supporting Latin characters and a set of symbols. In order to compose an email, the user fills his email and password, the receiver's email, the subject and the main body of the mail. A Word Processor allows the user to create, edit and save a .txt file. The scene consists of a 3D paper model representing a notepad and a Latin 3D keyboard. Finally, two immersive mini games were implemented. A puzzle game in which a user-selected picture is fragmented in tiles (Figure 6). The tile layout is then randomized and the user has to re-arrange the tiles to form the original picture and solve the puzzle. Finally, the provided action game is a 3D rendition of the classic flappy bird game including an airplane [3] (Figure 7).

4 User Study

4.1 Materials & Methods

We conducted a user study in order to evaluate the proposed eye-tracking interaction and 3D UI paradigm. We conducted a pilot study in order to identify which application required the greatest gaze-tracking accuracy. The study indicated that the Mail Composer required complex eye movements such as typing on a virtual keyboard and was selected to assess our method. A total of 7 people (2 female, mean age 24.3) participated in the experiment.

Apparatus The 3D UI was displayed on a NVisor™ SX111 HMD, having a stereo SXGA resolution and a Field-of-View (FoV)

Figure 4: The eye-controlled Music Player.

Figure 5: The eye-controlled Mail Composer.

of 102 degrees horizontal by 64 degrees vertical. Participants panned around the virtual environment using an InterSense™ InertiaCube3™ 3DoF head tracker attached to the HMD. Eye-tracking data was recorded using a twin-CCD binocular eye-tracker by Arrington Research™, also attached to the HMD updating at a frequency of 30Hz.

Procedure Participants sat on a swivel chair and were familiarized with the setup in a training session. Before acquiring data, every participant performed the standard eye tracker calibration procedure provided by Arrington Research™. Each individual calibration took approximately 20 seconds during which the subjects gazed at the center of 12 target squares displayed at different locations on the HMD. Following this, participants composed a dictated email by moving their eyes to select letters on a virtual keyboard and pressing a switch to select the letter they wished to write. Then the same email was written on a physical keyboard while still wearing the HMD. The keyboard was occluded by the HMD. Task accuracy and task completion times were recorded. All but one participants (that was excluded from the analysis due to insufficient eye tracking data) successfully executed both tasks by wearing the eye-tracking capable HMD.

After the experiment ended, participants were asked to rate their experience when using the 3D MUI on a 1-7 Likert Scale, by answering four questions, commonly used in qualitative assessments of 3D User Interfaces [13]: (i) The 3D UI is as comfortable as the traditional mouse-keyboard paradigm. (ii) I felt more tired when using the 3D UI than with the traditional mouse-keyboard paradigm. (iii) The 3D UI was more interesting when using it compared to traditional mouse-keyboard paradigm. (iv) I prefer the 3D UI more than the traditional mouse-keyboard paradigm.

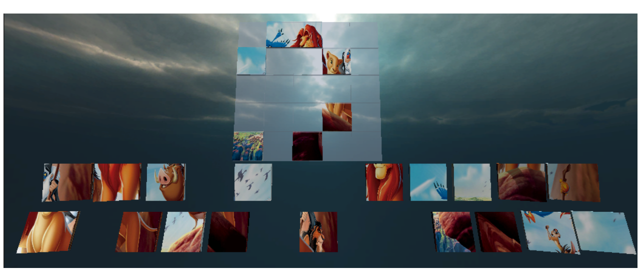

Figure 6: The eye-controlled Puzzle game.

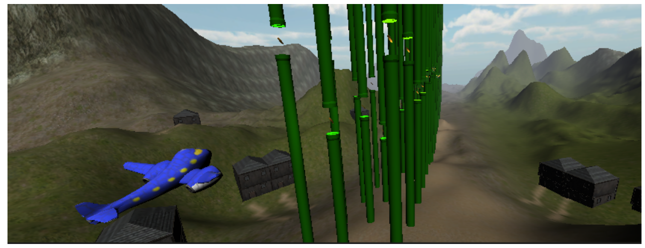

Figure 7: The eye-controlled airplane in the Flappy airplane game.

4.2 Data Analysis & Discussion

An independent-samples t-test was conducted to compare type rate with the eye-tracker versus occluded keyboard input. There was a significant difference in the scores for the eye-tracker (M=257.6s, SD=5570.3) and keyboard (M=123.3s, SD=7169.0) conditions; $t(6) = 2.91$, $p < 0.05$. These results suggest that participants type about 2 times slower on average (Figure 8) with the eye-tracker interface. However, analysing the type error rate for both interfaces (9.3% for the eye-tracker vs 54.19% for the occluded keyboard) indicated that despite the fact that users type faster when utilizing the out-of-view keyboard, they make many more errors (Figure 8).

The results of the qualitative analysis (Figure 9) and thorough discussion with the test subjects indicated that the proposed interaction method is much more enjoyable than a standard keyboard. Fatigue levels where found to be the same for both interfaces, with the exception of one subject (Figure 9). Usage of a 3D UI keyboard does not require good typing skills. After interaction with the rest of the applications, participants stated that they would certainly opt for the 3D UI since it feels more futuristic and therefore more exciting for the users.

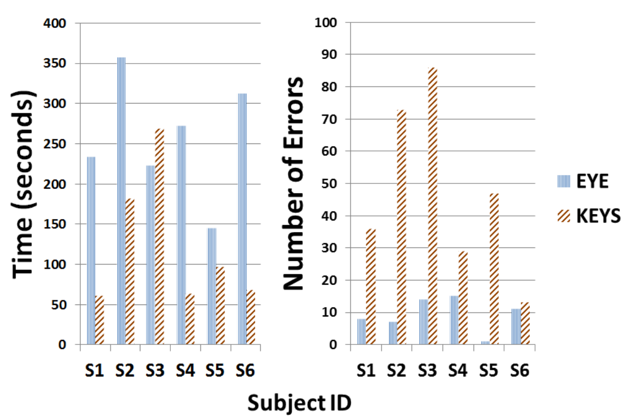

Figure 8: Left: Comparison of task completion times for the eye-tracker input versus blind keyboard input for all subjects. Right: Comparison of type errors for the eye-tracker input versus blind keyboard input for all subjects.

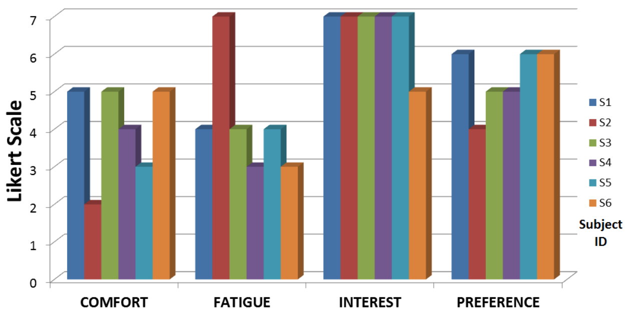

Figure 9: Responses on the four questions of the qualitative questionnaire for the 3D UI keyboard.

5 Conclusion

We presented an innovative gaze-controlled MUI for an eye-tracking capable headset, suitable for modern consumer-grade HMDs. We developed six applications which cover commonly operated actions of everyday computing such as mail composing, music playing and photo viewing. We performed a user study by acquiring both quantitative and qualitative data. The type error rate was lower when utilizing the proposed 3D UI in comparison to an occluded keyboard. The qualitative study indicated that users enjoy considerably more the proposed 3D UI over the traditional input methods such as a keyboard and mouse. A middle-ware API will soon be provided for eye-tracking and blink handling that can be extended to additional MUIs designated for the Oculus Rift^{TM} and Samsung Gear VR^{TM}.

Acknowledgements

This research has been co-financed by the European Union (European Social Fund - ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Heracleitus II: Investing in knowledge society through the European Social Fund.

European Union
European Social Fund
OPERATIONAL PROGRAMME
EDUCATION AND LIFELONG LEARNING
investing in knowledge society
MINISTRY OF EDUCATION & RELIGIOUS AFFAIRS, CULTURE & SPORTS
MANAGING AUTHORITY
NSRF 2007-2013
programme for development
EUROPEAN SOCIAL FUND
Co- financed by Greece and the European Union

References

[1] D. A. Bowman, J. Chen, C. A. Wingrave, J. F. Lucas, A. Ray, N. F. Polys, Q. Li, Y. Haciahmetoglu, J.-S. Kim, S. Kim, et al. New directions in 3d user interfaces. *IJVR*, 5(2):3–14, 2006.

[2] D. A. Bowman, E. Kruijff, J. J. LaViola Jr, and I. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley, 2004.

[3] T. F. B. game. Flappy bird, Feb. 2015.

[4] T. E. Hutchinson, K. P. White Jr, W. N. Martin, K. C. Reichert, and L. A. Frey. Human-computer interaction using eye-gaze input. *Systems, Man and Cybernetics, IEEE Transactions on*, 19(6):1527–1534, 1989.

[5] R. J. Jacob and K. S. Karn. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind, Oxford Journal*, 2(3):4, 2003.

[6] M. Kaur, M. Tremaine, N. Huang, J. Wilder, Z. Gacovski, F. Flippo, and C. S. Mantravadi. Where is it? event synchronization in gaze-speech input systems. In *Proceedings of the 5th international conference on Multimodal interfaces*, pages 151–158. ACM, 2003.

[7] D. Mardanbegi and D. W. Hansen. Mobile gaze-based screen interaction in 3d environments. In *Proceedings of the 1st conference on novel gaze-controlled applications*, page 2. ACM, 2011.

[8] Oculus. Oculus developer guide, Feb. 2015.

[9] J. L. Sibert, M. Gokturk, and R. A. Lavine. The reading assistant: eye gaze triggered auditory prompting for reading remediation. In *Proceedings of the 13th annual ACM Symposium on User Interface Software and Technology (SUIT)*, pages 101–107. ACM, 2000.

[10] SMI. Sensomotoric instruments announces eye-tracking upgrade for oculus rift, Feb. 2015.

[11] A. Sutcliffe. Multimedia user interface design. *Human-Computer Interaction: Design Issues, Solutions, and Applications*, page 65, 2009.

[12] M. Yamada and T. Fukuda. Eye word processor (ewp) and peripheral controller for the als patient. *IEEE Proceedings on Physical Science, Measurement and Instrumentation, Management and Education-Reviews*, 134(4):328–330, 1987.

[13] B. Yoo, J.-J. Han, C. Choi, K. Yi, S. Suh, D. Park, and C. Kim. 3d user interface combining gaze and hand gestures for large-scale display. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pages 3709–3714. ACM, 2010.