Gaze-based interaction for VR environments

Patryk Piotrowski and Adam Nowosielski[0000-0001-7729-7867]

West Pomeranian University of Technology, Szczecin
Faculty of Computer Science and Information Technology
Żołnierska 52, 71-210, Szczecin, Poland
patryk.piotrowski19@gmail.com, anowosielski@wi.zut.edu.pl

Abstract. In this paper we propose a steering mechanism for VR headset utilizing eye tracking. Based on the fovea region traced by the eyetracker assembled into VR headset the visible 3D ray is generated towards the focal point of sight. The user can freely look around the virtual scene and is able to interact with objects indicated by the eyes. The paper gives an overview of the proposed interaction system and addresses the effectiveness and precision issues of such interaction modality.

Keywords: gaze-based interaction · virtual reality · eye tracking · gaze-operated games.

1 Introduction

Virtual reality systems are computer-generated environments where an user experiences sensations perceived by the human senses. These systems are based primarily on providing video and audio signals, and offer the opportunity to interact directly with the created scene with the help of touch or other form of manipulation using hands. Vision systems for virtual reality environments consist most frequently of head-mounted goggles, which are equipped with two liquid crystal displays placed opposite the eyes in a way that enables stereoscopic vision. The image displayed in the helmet is rendered independently for the left and right eye, and then combined into the stereopair. More and more solutions appear on the market, and the most popular include: HTC Vive, Oculus Rift CV1, Playstation VR (dedicated for Sony Playstation 4), Google Cardboard (dedicated for Android mobile devices).

The virtual reality solutions are delivered with controllers which aim is to increase the level of user's immersion with elements of the virtual environment. Interestingly, many novel interfaces offer hands-free control of electronic devices. The touchless interaction there is based on recognition of user actions performed by the whole body [1] or specific parts of the body (e.g. hands [2], head [3]). A completely new solution, not widely used and known, is the control through the sight, i.e. gaze-bazed interaction. The operation of such systems is based on eyetracking, a technique of gathering real-time data concerning gaze direction of human eyes [4]. The technology is based on tracking and analysing the movement of the eyes using cameras or sensors that register the eye area [4]. The latest solutions on the market introduce eye tracking capabilities to virtual reality environments [5].

In the paper we propose a novel steering mechanism based on ray-casting for human-computer interfaces. Based on the fovea region traced by the eye tracker, assembled into VR headset, the visible 3D ray is generated towards the focal point of sight. Thanks to head movements the user can freely look around the virtual scene and is able to interact with objects indicated by eyes.

The paper is structured as follows. In Sect. 2 the related works are addressed. Then in Sect. 3 the concept of gaze-based interaction in virtual reality environment is proposed. An example application is presented in Sect. 4. The proposed system is evaluated in Sect. 5. Final conclusions and a summary are provided in Sect. 6.

2 Related works

Most of the eye tracking solutions have been utilized for analysis of the eye movement for advertising industry, cognitive research and analysis of individual patterns of behaviours [6–8]. The eye tracking systems are recognized tools for analysis of the layout and operation efficiency of human-computer interaction systems [9]. They are now regarded also as input modality for people with disabilities [10]. For some users who suffer from certain motor impairments, the gaze interaction may be the only option available. Typical way of interaction in such systems assume fixations and dwell times. User is expected to look at specific element of the interface for a predefined time period called the dwell-time and after that the system assumes the selection (equivalent to mouse clicking). Such solution is used for navigating through graphical user interface or for eye-typing using the on-screen keyboard. Some innovations to this technique has been proposed. In [11] a cascading dwell gaze typing technique dynamically adjust the dwell time of keys in an on-screen keyboard. Depending on the probability during typing some keys are easier to select by decreasing their dwell times other are harder to choose (increased dwell times). A completely different approach was presented in [10]. Here, a dwell-free eye-typing technique has been proposed. Users are expected to look at or near the desired letters without stopping to dwell.

New solution replacing the traditional technique of fixations and dwell time has been proposed in [12]. Authors has proposed gaze gestures. In contrast to the classical way of interaction using the eye sight, a gaze gesture uses eye motion and are insensitive to accuracy problems and immune against calibration shift [12]. This is possible because gaze is not used directly for pointing and only the information of relative eye movements are required. The gaze gestures have also been reported to be successful in interacting with games [13].

Apart from using gaze to control computer systems other interesting applications can be found in the scientific literature. To accelerate the raytracing in [14] fewer primary rays are generated in the peripheral regions of vision. For the fovea region traced by the eye tracker the sampling frequency is maximised [14].

The above examples show the multitude of applications of eye tracking systems. The novelty is now installing these systems in virtual reality headsets which offers new possibilities of application.

3 Gaze-driven ray-cast interface concept

The concept of using an eye tracker for steering in the virtual reality environment assumes the usage of the eye focus point to interact with objects of the virtual scene. An overview of the system built upon this concept is presented in Fig. 1 and the process of interaction consists of the following steps:

- mapping the direction of the user's eye focus on the screen coordinates,
- generating a primary ray (raycasting using the sphere) from the coordinates of the user's eye focus direction,
- intersection analysis with scene objects,
- indication of the object pointed by the sight,
- handling the event associated with the object,
- rendering a stereopair for virtual reality goggles.

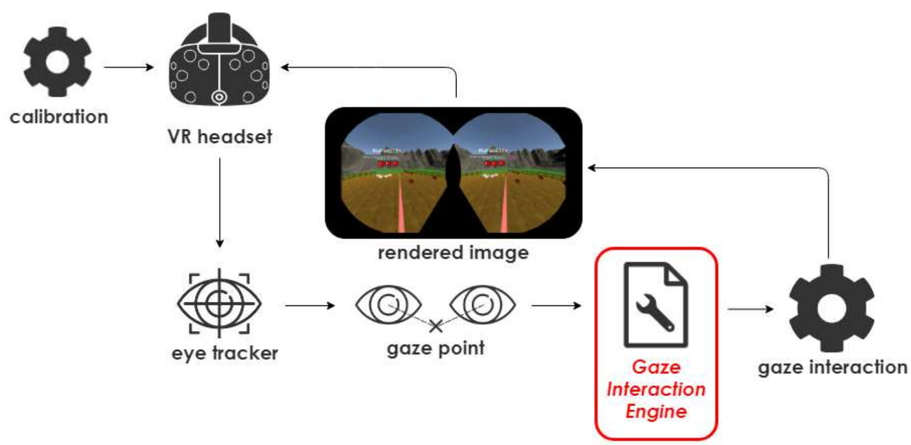

Fig. 1. Scheme of the VR system with gaze-based interaction.

The main idea, then, is to generate a ray which takes into account the position, rotation and direction of the eye's focus on the virtual scene. During the initial mapping, coordinates are taken for the left and right eyes independently, and the final value of the focal point is the result of their averaging. In case of intersection detection with the scene object the appropriate procedures are executed.

Figure 2 presents diagram of the interaction process with a scene object using eye focus direction. Four states can be distinguished for the object: no interaction, beginning, continuing (during), and ending interaction. Start of the interaction is crucial since it might be triggered with the eyesight solely (after a predefined dwell time) or with the use of hand operated controller.

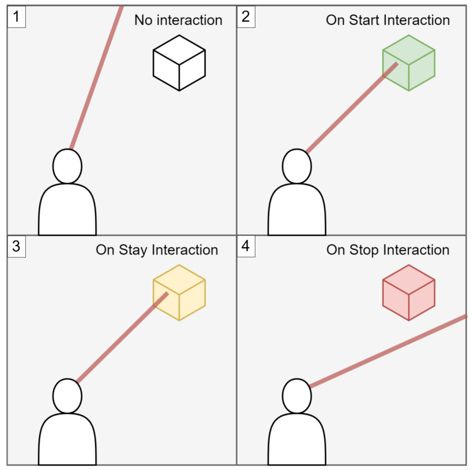

Fig. 2. Diagram showing the interaction process with a scene object using eye focus direction.

4 Implementation and Application

Based on the concept presented in Sect. 3 a sight-operated interaction system, named *Gaze Interaction Engine* (marked with a red border in Fig. 1), for the virtual reality environment has been developed. This solution has the form of a *UnityAsset* module for the Unity environment. It is hereby made available to the public and can be accessed through the web page [15]. The developed gaze-based interaction system is designed for the virtual reality HTC Vive hardware and eye tracker from Pupil Labs [5]. In our research the eye tracker has been set to receive 640 x 480 pixel infrared eye image with 120 frames per second.

Our interface can be employed to create computer games and multimedia applications. A good example of using the tool was presented during the event devoted to games creation *GryfJam* in Szczecin (Poland) on 17th and 18th of May 2019. One of the authors of the paper, Patryk Piotrowski, with the help of Michał Chwesiuk developed a simple game for virtual reality glasses with manipulation only with the gaze. The game, named *KurzoCITY*, belongs to the genre of arcade games. The player's goal is to collect as many grains as possible on the farm under the pressure of competition from virtual poultry (see Fig.3 for a game preview).

The eye tracker used in the helmet analyzes the movement of the eyeballs. For each frame a ray is generated from the player's eyes to the focal point. A

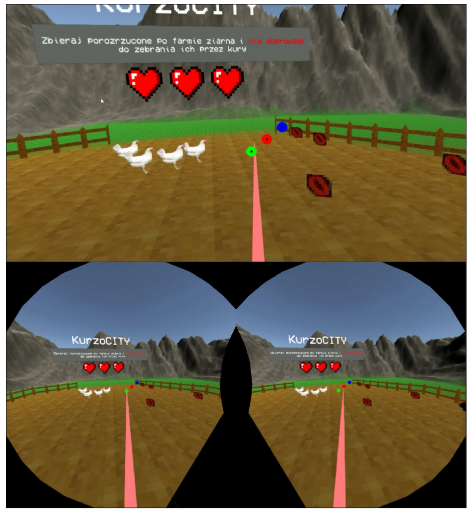

Fig. 3. The use of Gaze Interaction Engine in the *KurzoCITY* game: screen view (top) and stereo pair for virtual reality googles (bottom).

look at the grain allows it to be collected. Over time, the level of difficulty of the game increases by adding new opponents and raising the number of grains to collected by the player. The game ends when the poultry collect a total of three seeds.

5 Evaluations

The game described in the preceding section, as already mentioned, had been developed during the *GryfJam* event. Using the developed game and event's participants, tests of the effectiveness of the *Gaze Interaction Engine* were conducted. We observed high playability which indicates that proposed gaze-steering mechanism is successful. Nevertheless, some problems and imperfections of the eye tracking system have been noticed. Among over 30 participants of all our experiments, we found 2 who were not able to pass the calibration process entirely. The greatest setup difficulty was fitting the helmet and adjusting the distance between the lenses which ensure correct detection of the pupil. With an unmatched arrangement, the position of the pupil can not be determined correctly and the examples are presented in Fig. 4. The top left sample, for comparison purposes, contains a correct case. The eye is in the center, corneal reflections are visible, the center of pupil is annotated with the red dot and the pupil border is surrounded by a red border.

The second problem encountered was decalibration of the eye tracker at the time of use. Expressions which appear on the face may cause slight shifts of the entire headset and in the effect render the eye tracker erroneous.

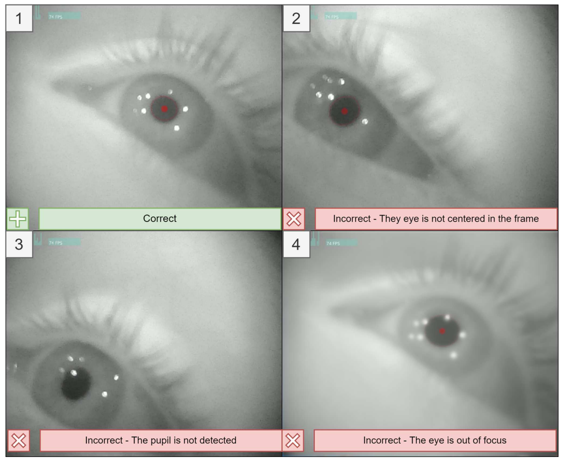

Fig. 4. Calibration problems: the appearance of the eye seen by the eye tracking system mounted in the virtual reality helmet.

Encountered problems, described above, can be classified as hardware related. To evaluate the accuracy of the eye-based interaction an additional experiment has been conducted. We prepared a grid with 26 separate interactable buttons (divided into three rows, occupying approximately half field of view vertically and 100% field of view horizontally). The goal of each participant was to press the highlighted button by focusing the eyes on it with the dwell-time equal 600 ms and visual progress indicator provided. We measured the time of pressing randomly highlighted buttons and the accuracy of the process itself. There were 17 participants (volunteers from students and employees of our university) who performed between 2 and 6 sessions. There have been 70 sessions in total and each sessions consisted of pressing 23.6 buttons on average. Results are presented in the graphical form in Fig. 5.

The averaged time of pressing a random button equal 1.79 second. It includes the 600 ms dwell-time, required for the interaction to take place. The precision seems to be more problematic here. We registered the averaged (over all participants and sessions) error rate of 5.51%. The error have been calculated as the ratio of pressing the improper button (most often the adjacent one) for the total number of presses. These results indicate that interfaces composed of many components arranged close to each other may be problematic to operate using current eye tracking solutions for the virtual reality helmets. However, when the number of interactive elements in the scene decreases and the size of these elements increase the interaction is quite convenient. The proposed game is a good proof here. With the relatively small dwell time (set to 200 ms compared to 600 ms in the pressing buttons experiment) very high level of interaction among participants have been observed.

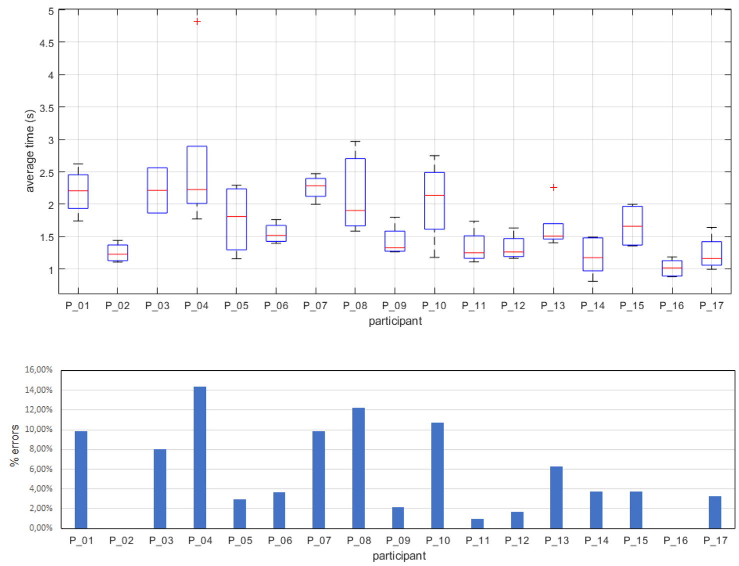

Fig. 5. Evaluation results: performance (top) and accuracy (bottom) of 17 participants.

6 Conclusion

The proposed interaction system for virtual reality environments enables the effective implementation of multimedia applications and games, operated using the eyesight. The visible 3D ray is generated towards the focal point of sight to facilitate user with the interaction process where free head movements are present. The eye control is faster compared to, for example, additional hand-operated controllers. In order for the motor reaction to take place, a stimulus and a nerve impulse are required after visual observation. These stages are eliminated. Eye trackers mounted in the VR headsets can significantly help people with disabilities offering unusual possibilities and for a wide range of recipients can offer new opportunities for interaction in human-computer interfaces and games.

References

1. Giorio, C., Fascinari, M.: Kinect in Motion – Audio and Visual Tracking by Example. Packt Publishing, Birmingham (2013)
2. Nowosielski, A.: Evaluation of Touchless Typing Techniques with Hand Movement. In: Burduk, R., et al. (eds) Proceedings of the 9th International Conference on Computer Recognition Systems CORES 2015. AISC, vol. 403, pp. 441–449. Springer, Cham (2016)
3. Nowosielski, A.: 3-Steps Keyboard: Reduced Interaction Interface for Touchless Typing with Head Movements. In: Kurzynski M, Wozniak M, Burduk R (eds.) Pro-

ceedings of the 10th International Conference on Computer Recognition Systems CORES 2017. AISC, vol. 578, pp. 229–237. Springer, Cham (2018)
4. Mantiuk, R., Kowalik, M., Nowosielski, A., Bazyluk, B.: Do-It-Yourself Eye Tracker: Low-Cost Pupil-Based Eye Tracker for Computer Graphics Applications. LNCS, vol. 7131, pp. 115–125 (2012)
5. Pupil Labs GmbH. Eye tracking for Virtual and Augmented Reality, https://pupil-labs.com/vr-ar/. Last accessed 15 Jun 2019
6. Wedel, M., Pieters, R.: A Review of Eye-Tracking Research in Marketing. In: Malhotra N.K. (ed.) Review of Marketing Research (Review of Marketing Research, Volume 4), Emerald Group Publishing Limited, pp. 123–147 (2008)
7. Berkovsky, S., Taib, R., Koprinska, I., Wang, E., Zeng, Y., Li, J., Kleitman, S.: Detecting Personality Traits Using Eye-Tracking Data. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19. pp. 221:1–221:12. ACM New York, NY, USA (2019)
8. Jankowski, J., Ziemba, P., Watróbski, J., Kazienko, P.: Towards the Tradeoff Between Online Marketing Resources Exploitation and the User Experience with the Use of Eye Tracking. In: Nguyen NT, Trawiński B, Fujita H, Hong TP (eds) Intelligent Information and Database Systems. ACIIDS 2016. LNCS, vol. 9621, pp. 330–343. Springer, Berlin, Heidelberg (2016)
9. Jacob, R.J.K., Karn, K.S.: Commentary on Section 4 - Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises. In: Hyönä J, Radach R, Deubel H (eds) The Mind's Eye, pp. 573–605. North-Holland (2003)
10. Kristensson, P.O., Vertanen, K.: The potential of dwell-free eye-typing for fast assistive gaze communication. In: Spencer SN (Ed.) Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12), pp. 241–244. ACM, New York, NY, USA (2012)
11. Mott, M.E., Williams, S., Wobbrock, J.O., Morris, M.R.: Improving Dwell-Based Gaze Typing with Dynamic, Cascading Dwell Times. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17), pp. 2558–2570. ACM, New York, NY, USA (2017)
12. Drewes, H., Schmidt, A.: Interacting with the Computer Using Gaze Gestures. In: Baranauskas C., Palanque P., Abascal J., Barbosa S.D.J. (eds) Human-Computer Interaction – INTERACT 2007. INTERACT 2007. LNCS, vol. 4663, pp. 475–488. Springer, Berlin, Heidelberg (2007)
13. Istance, H., Hyrskykari, A., Immonen, L., Mansikkamaa, S., Vickers, S. Designing gaze gestures for gaming: an investigation of performance. Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications (ETRA '10), pp. 323–330. ACM New York, NY, USA (2010)
14. Siekawa, A., Chwesiuk, M., Mantiuk, R., Piórkowski, R.: Foveated Ray Tracing for VR Headsets. MultiMedia Modeling. LNCS, vol. 11295, pp. 106–117 (2019)
15. Piotrowski, P., Nowosielski, A. (2019) *Gaze Interaction Engine* (project page), https://github.com/patryk191129/GazeInteractionEngine