Who do you look like? – Gaze-based authentication for workers in VR

Category: Research

ABSTRACT

Behavior-based authentication methods are actively being developed for XR. In particular, gaze-based methods promise continuous authentication of remote users. However, gaze behavior depends on the task being performed. Identification rate is typically highest when comparing data from the same task. In this study, we compared authentication performance using VR gaze data during random dot viewing, 360-degree image viewing, and a nuclear training simulation. We found that within-task authentication performed best for image viewing (72%). The implication for practitioners is to integrate image viewing into a VR workflow to collect gaze data that is viable for authentication.

Keywords: Eye Tracking, VR, Authentication, Future of Work

Index Terms: Human-centered computing—Systems and tools for interaction design; Computing methodologies—Virtual reality

1 INTRODUCTION

Behavior-based authentication methods are actively being developed in the XR community. In particular, user authentication based on gaze cues offers the promise of seamless and continuous authentication [6]. Current literature on gaze-based authentication in VR has focused on reading [4, 7] or image viewing [1] tasks. Maximum identification rates reported in these works range from 85% [1] to 97% [7]. Authentication algorithms performed best when the classifier was trained and tested on data from the same task [5].

In this study, we compared three scenarios for gaze based authentication: random dot viewing, image viewing, and completing a VR simulation of a nuclear reactor startup procedure. The first two scenarios were drawn from published literature [1, 3]. The third scenario was designed by our motivating context: remote VR-based training for nuclear reactor operators.

We hypothesized that we would observe identification rates comparable to published literature for the first two tasks. We hypothesized that identification rate would be lower for the simulation task, but above chance level. We found that identification rates are higher (72% within-task) for image viewing, which is consistent with prior literature. Identification for random dot viewing and the VR simulation task was marginally above chance level (15% and 12%). The implication for practitioners is to integrate image viewing into VR training, for example, by instructing users to familiarize themselves with sample environments before proceeding to the main task.

2 METHODOLOGY

An IRB approved experiment was used to collect eye-tracking data from various tasks in VR to explore the feasibility of gaze-based authentication within the context of nuclear engineering.

2.1 Equipment

The wireless Pico Neo 2 Eye head-mounted display (HMD) was used for this study due to its compatibility with Unity3D and native eye-tracking capabilities. The Pico HMD uses two handheld controllers for menu navigation and interaction within the training environment.

2.2 Participant Recruitment

Participants were recruited from the undergraduate student population at ANONYMOUS via email and flyers. Students were enrolled in departments related to nuclear engineering and materials science. Seventeen appointments were scheduled, and ten were fulfilled.

2.3 Study Flow

Upon arrival to the lab, participants heard an explanation of the steps of the protocol and affirmed informed consent. The HMD eye tracker was then calibrated using the default Tobii User Calibration. Participants viewed a plane with five circular targets that spanned three degrees of visual angle with a dynamic sphere visualizing the current gaze position. Participants were asked to view each target and indicate whether the sphere accurately followed their gaze. If the gaze sphere did not fall within the five targets then the calibration was repeated until gaze accuracy was validated.

The participant was provided instructions before each of the three authentication tasks (Sec. 2.4). Stimuli were always presented in the same order. Random dot viewing was the first task, followed by image viewing. A break of up to five minutes was provided at the midpoint of the image viewing set. The HMD eye tracker was then re-calibrated and validated. After image viewing, the participant took another break before moving on to the nuclear training simulation. Once all tasks were completed, an end of study survey gathered demographics and level of experience with VR technology.

2.4 Authentication Tasks

Figure 1 illustrates three VR authentication tasks. These tasks were motivated by past studies and the remote training workflow in VR.

2.4.1 Random Dot Viewing

A random dot viewing task was adapted to VR based on past studies on gaze-based authentication [3]. Participants followed the jumping dot with their eyes for 100 seconds. In the previous study, participants viewed a stationary screen using a chin rest to stabilize head position. In this study, we simulated the setup by presenting the random dot video on a virtual plane at a fixed distance in front of them without considering head movements or rotations.

2.4.2 Image Viewing

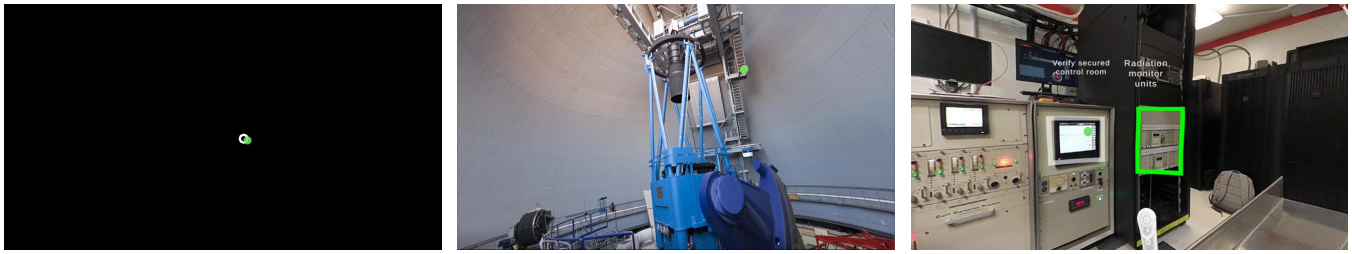
Participants viewed 50 randomly ordered 360-degree images. The number of stimuli was justified by previous work indicating that gaze-based authentication can be viable using 50 two-dimensional images [5]. Participants were instructed to view each omnidirectional image for 25 seconds with five second transitions after each image, for a total duration of 25 minutes. The image set includes 30 images of natural scenes previously used for authentication in VR [1] and 20 free-use images of laboratory and plant scenes.

2.4.3 Nuclear Training Simulation

The final task was a nuclear training simulation for powering up a reactor. This was justified as a prototype of a VR-based nuclear training scenario. Participants viewed a 360-degree image of a nuclear training reactor. Areas of Interest (AOIs) were outlined to indicate which equipment can be selected. Duration ranged from four to eight minutes depending on the participants' progress. To advance, participants performed point and click button presses on the AOIs corresponding to the current step in powering up the reactor.

2.5 Authentication Model

Authentication was performed using a Radial Basis Function Network and a set of features from fixation and saccade events previously applied to VR data [1]. Features were generated for each participant and task and then segmented into blocks. Random dot viewing, image viewing, and simulation were segmented by time,



Random Dots Image Viewing Simulation

Figure 1: VR environments for each task. Gaze position is visualized as a green sphere in this figure, but was not visible to participants.

Table 1: Identification rate for authentication across tasks.

	Test: Dots	Test: Images	Test: Sim.
Train: Dots	15%	8%	9%
Train: Images	10%	72%	25%
Train: Simulation	11%	15%	12%

image, and simulation step, respectively. These blocks were then randomly selected to compose the training dataset which was used to fit the model, and the testing dataset which was used to evaluate the accuracy of predictions on participant identity. For within-task evaluation, 50% of the data was used for training and 50% for testing. Likewise, for between-task evaluation 50% of the data from each task was used for training and testing. Identification rate was determined by classifying features from each individual in the testing set to make a single prediction of their identity. If the model matched the individual's features to the correct identity, the classification succeeded. This classification was done for each individual, and the percentage of correct classifications was computed as identification rate. The evaluation was repeated ten times for each combination of tasks with a random selection of training and testing blocks.

3 Results

Table 1 presents the identification rates within and between each task. Most attempts achieved an authentication rate better than the chance rate (10%). Within-task authentication generally had greater performance than between-task. The within-task identification rate for image viewing was highest at 72%. The best between-task authentication (25%) was achieved from training on image viewing and testing on simulation.

4 Discussion

We evaluated within-task and between-task authentication using eye movements in a VR environment. Within-task authentication was generally more accurate than between-task authentication. Between-task authentication was highest for similar tasks. Within-task authentication for image viewing had the highest identification rate at 72%. The image viewing task was longer than the other tasks, generating about 21 minutes of data. We hypothesize that image viewing performance was a result of the volume of data and elicitation of repetitive exploratory behavior across many images as seen in past studies [1, 5].

Random dot viewing and simulation under-performed image viewing for within-task authentication. Higher rates for random dot viewing were expected based on past work that achieved 96% accuracy. However, our experiment varied in that our eye tracker sampling rate of 90Hz was much lower than 1000Hz. Also, we only showed the random dot sequence to the user once, resulting in a data volume of 100 seconds compared to 200 seconds, generating training and testing datasets that did not contain repeated dot movements [3].

The simulation and image-viewing tasks both involved exploring 360-degree images, but simulation had lower within-task accuracy. This could be due to the volume of data, which varied based on how long participants took to complete the procedure and was at most eight minutes. Varying experience levels among nuclear engineering students may also impact whether eye movements were more exploratory or direct, influencing the observed behavior patterns. Between-task identification rates were lower than within-task results. The results indicate that comparing tasks which elicit exploratory behavior with tasks based on prescribed behavior may negatively impact identification rates. Further, the best between-task performance (25%) was achieved from training on image viewing and testing on simulation, suggesting that the tasks' similarity has a positive impact and potential for identification if more data were available.

Future Work Our observations suggest that both the type and duration of task impact authentication performance. To further explore duration's effect, we could analyze subsets of the image-viewing dataset. Additionally, the random dot viewing task may be executed twice for each participant to elicit repetitive eye movements, as was done in previous work [3]. Exploring the eye movement behaviors and subsequent feature distributions elicited by each task would be a valuable step in understanding individual differences in between-task authentication. The use of expanded feature sets including pupil biometrics, different classification models, and methods that map feature distributions between task would support this understanding [2]. In this exploration, image viewing was found to be the most viable task for authentication. In a potential job training authentication pipeline, integrating image-viewing scenes early in the training program could permit within-task authentication. For example, an exploration period at the beginning of the simulation task could produce data for authentication in a natural manner.

References

[1] B. David-John, D. Hosfelt, K. Butler, and E. Jain. A privacy-preserving approach to streaming eye-tracking data. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2555–2565, 2021.

[2] S. Eberz, G. Lovisotto, K. B. Rasmussen, V. Lenders, and I. Martinovic. 28 blinks later: Tackling practical challenges of eye movement biometrics. In *Proceedings of ACM CCS*, pp. 1187–1199, 2019.

[3] O. V. Komogortsev and I. Rigas. Bioeye 2015: Competition on biometrics via eye movements. In *IEEE BTAS*, pp. 1–8. IEEE, 2015.

[4] D. J. Lohr, S. Aziz, and O. Komogortsev. Eye movement biometrics using a new dataset collected in virtual reality. In *Proceedings of ACM ETRA*, pp. 1–3, 2020.

[5] C. Schröder, S. M. K. Al Zaidawi, M. H. Prinzler, S. Maneth, and G. Zachmann. Robustness of eye movement biometrics against varying stimuli and varying trajectory length. In *Proceedings of ACM CHI*, pp. 1–7, 2020.

[6] I. Sluganovic, M. Roeschlin, K. B. Rasmussen, and I. Martinovic. Analysis of reflexive eye movements for fast replay-resistant biometric authentication. *ACM TOPS*, 22(1):1–30, 2018.

[7] J. Steil, I. Hagestedt, M. X. Huang, and A. Bulling. Privacy-aware eye tracking using differential privacy. In *Proceedings of ACM ETRA*, 2019.