International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2017)
T. Huang (Editor)

Poster

A Gaze-depth Estimation Technique with an Implicit and Continuous Data Acquisition for OST-HMDs

Youngho Lee[1,2] Thammathip Piumsomboon[2], Barrett Ens[2], Gun Lee[2], Arindam Dey[2], and Mark Billinghurst[2]

[1]Mokpo National University, South Korea
[2]University of South Australia, Australia

Abstract

The rapid development of machine learning algorithms can be leveraged for potential software solutions in many domains including techniques for depth estimation of human eye gaze. In this paper, we propose an implicit and continuous data acquisition method for 3D gaze depth estimation for an optical see-Through head mounted display (OST-HMD) equipped with an eye tracker. Our method constantly monitoring and generating user gaze data for training our machine learning algorithm. The gaze data acquired through the eye-tracker include the inter-pupillary distance (IPD) and the gaze distance to the real and virtual target for each eye.

CCS Concepts

•***Human-centered computing*** → ***Mixed / augmented reality;***

1. Introduction

The recent advancements and commercialization of the optical see-through head mounted displays (OST-HMD) and eye-trackers enable better Augmented Reality (AR) experience and higher precision gaze tracking. Both areas are crucial and active research topic that would influence how we interact with the machine in the near future [PLL*17, BGK*16]. Consumer graded eye trackers such as Pupil-labs, Tobii, SMI are more accessible and they recently offered the eye-trackers for HMDs. There are also HMDs that have the eye tracker integration [KPB14].

Past research have studied the acquisition of the 3D gaze plane. Geometrical calculations could be used to obtain 2D gaze position. However, due to the physiological characteristics of the human eye such as convergence and accommodation as well as the limitations of the current sensor, it has been challenging to obtain accurate and reliable continuous 3D gaze position. To alleviate this problem, the machine learning techniques such as Supporting Vector Machine (SVM) and Multi-layer Perceptron (MLP) were used to estimate 3D gaze depth [LSP*17]. To train this model, we require the gaze data for a continuous distance.

In this paper, we propose an implicit and continuous data acquisition method for 3D gaze-depth estimation for an OST-HMD with an eye-tracker. Our method generates training data for machine learning by acquiring the gaze-vector and the pupil distance of the two eyes and the distance to the object. Our procedure has three steps. Firstly, it measures the distance between eyes (IPD), while estimating the gaze vector of the eyes looking at the AR target at the same time. Secondly, it displays a virtual target on the OST-HMD and records the gaze vectors and the corresponding distance. This data between eyes and the real and virtual targets are used for training. Lastly, our technique aligns the virtual target to the real one.

The primary contribution of the proposed technique is the ability to implicitly and continuously collect the necessary data for adjusting the 3D gaze estimation in real-time. This method can increase the accuracy by acquiring continuous distance value from real and virtual. This offers an advantage over the conventional calibration method, which potentially reduce error over time causing by the HMD drifting.

2. Related Work

There have been few studies conducted for estimating gaze depth for OST-HMD. Kwon and Shul presented an experimental result for gaze-based 3D interaction to stereo image display [KS06]. Toyama et al. used eye-tracker to determine the focal plane for the multi-focal length OST-HMD [TOSK14]. Lee et al. determined the gaze plane of a multi-focal plane for OST-HMD to estimate the continuous gaze depths of human eyes [LSP*17].

Lee et al. found that machine learning algorithms could reduce the estimation error but required constant update of gaze data [LSP*17]. When the distance between eyes and a target object of the learning data is discrete, the error in the estimated value of the intermediate position grew larger. Therefore, the gaze vectors and the corresponding continuous distance values should be used for training. Moreover, HMD was found drifting after a period of

DOI: 10.2312/egve.20171364

time of usage and therefore, constant correction is required. They concluded that it is necessary to acquire and train data as often as possible.

3. Proposed Method

The machine learning techniques require data to learn and improve the prediction. To calculate the 3D gaze depth, gaze data from an eye-tracker, corresponding distance data, and pupil distance of the user must be collected. Figure 1 describes this process. Once those data is input into the neural network, a suitable model is generated. Then, we can predict an estimated distance, with a gaze data as an input.

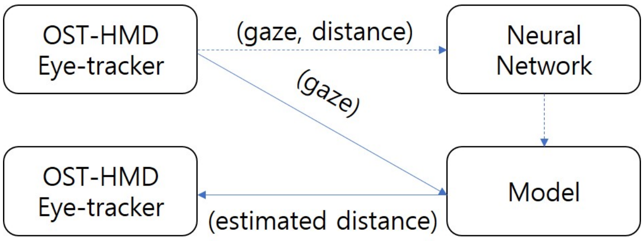

Figure 1: *Procedure for 3D gaze estimation. The dotted line is the learning process and the solid line is the estimation process.*

Our system's hardware comprises of the Microsoft Hololens and an eye-tracker [Mic16]. We used neural network as our machine learning algorithm. The system used the eye cameras and the world camera to track and measure the distance between eyes and an AR target. The OST-HMD has a 3D display so that it can show a virtual target at a specific distance. The distance between user's head and virtual target can be determined. The eye-tracker has to track both eyes and generate the gaze data in real time.

The proposed method has two steps. Firstly, it matches the real and virtual targets displayed on to the user. This procedure is required to calibrate the coordinates of OST-HMD and eye-tracker. Secondly, it obtains the gaze data from the eye-tracker, the pupil distance, and the distance between eyes and the real AR target, the AR target moves in a specific interval. Figure 2 shows geometrical details.

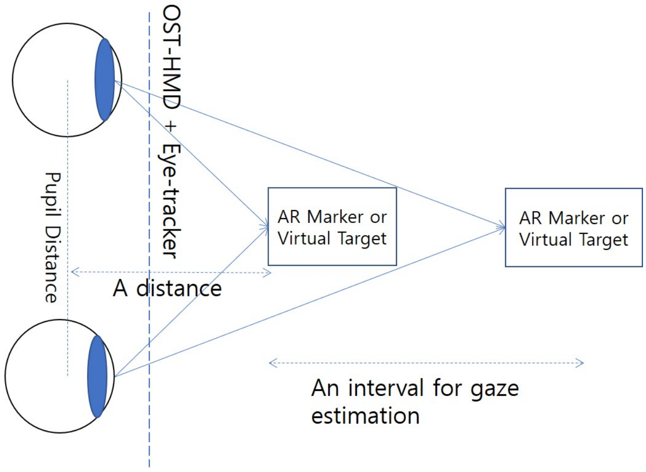

Figure 2: *Acquisition of gaze data according to continuous distance using an eye-tracker, a AR marker, and a virtual target image*

4. Conclusion

In this paper, we proposed a technique for 3D gaze depth estimation using an implicit and continuous gaze data acquisition. Our previous work found that we need gaze data corresponding to continuous distance and should be obtained as quickly as possible. Based on this finding, we proposed a method that allows us to capture gaze data at a continuous distance in the real and virtual worlds. In the future, we are going to implement a system to realize our proposed idea.

References

[BGK*16] BILLINGHURST M., GUPTA K., KATSUTOSHI M., LEE Y., LEE G., KUNZE K., SUGIMOTO M.: Is It in Your Eyes? Explorations in Using Gaze Cues for Remote Collaboration BT - Collaboration Meets Interactive Spaces. 177–199. URL: http://dx.doi.org/10.1007/978-3-319-45853-3_9, doi:10.1007/978-3-319-45853-3{_}9. 1

[KPB14] KASSNER M., PATERA W., BULLING A.: Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (2014), 1151–1160. URL: http://arxiv.org/abs/1405.0006, doi:10.1145/2638728.2641695. 1

[KS06] KWON Y. M., SHUL J. K.: Experimental researches on gaze-based 3D interaction to stereo image display. *Technologies for E-Learning and Digital Entertainment, Proceedings 3942* (2006), 1112–1120. doi:doi.org/10.1007/11736639{_}141. 1

[LSP*17] LEE Y., SHIN C., PLOPSKI A., ITOH Y., DEY A., LEE G., KIM S., BILLINGHURST M.: Estimating Gaze Depth Using Multi-Layer Perceptron. In *International Symposium on Ubiquitous Virtual Reality (ISUVR)* (2017), pp. 26–29. doi:10.1109/ISUVR.2017.13. 1

[Mic16] MICROSOFT: Microsoft Hololens Development Edition, 2016. URL: https://www.microsoft.com/microsoft-hololens/en-us. 2

[PLL*17] PIUMSOMBOON T., LEE Y., LEE G. A., DEY A., BILLINGHURST M.: Empathic Mixed Reality: Sharing What You Feel and Interacting with What You See. *2017 International Symposium on Ubiquitous Virtual Reality (ISUVR)*, section 4 (2017), 38–41. URL: http://ieeexplore.ieee.org/document/7988651/, doi:10.1109/ISUVR.2017.20. 1

[TOSK14] TOYAMA T., ORLOSKY J., SONNTAG D., KIYOKAWA K.: A natural interface for multi-focal plane head mounted displays using 3D gaze. *Proceedings of 12th International Working Conference on Advanced Visual Interfaces (AVI 2014) 2* (2014), 25–32. URL: http://dl.acm.org/citation.cfm?id=2598154, doi:10.1145/2598153.2598154. 1