Evaluation of Gaze-Depth Prediction Using Support Vector Machines

Choonsung Shin, Youmgmin Kim, Jisoo Hong,
Sung-Hee Hong, Hoonjong Kang
VR/AR Research Center
Korea Electronics Technologies Institute
Seoul, Republic of Korea
e-mail:{cshin, rainmaker, jhong, shhong, hoonjongkang}
@keti.re.kr

Youngho Lee
UVR Lab
Dept. of Computer Engineering
Mokpo National University
Jeonnam, Republic of Korea
e-mail: youngho@mokpo.ac.kr

***Abstract*—This paper presents the evaluation results of a gaze-depth prediction method for natural gaze interaction of wearable augmented reality. To calculate the gaze depth, we extracted the position of the center of the eyeball and the gaze vector of participants' eyes from a binocular eye tracker while the distance between participants' eyes and an object changed. We then applied support vector machines (SVM) to predict gaze depth. Based on our evaluation, prediction of gaze depth to actual focal distance was accurate to within +/- 20 cm.**

Keywords-gaze depth; eye tracking; augmented reality.

I. INTRODUCTION

Recently, augmented reality (AR) has received great attention from researchers and consumers due to the release of commercial devices from global companies. Smartphones are now commonly used in conjunction with AR SDKs (Software Development Kit), such as Vuforia and Kudan, thus developers easily make AR apps for them [1][2]. Furthermore, AR head-mounted displays (HMDs), such as Hololens have had a big impact on the possibility of AR for consumer business [3].

There have been many studies on improving interaction for AR glasses [4][5]. Researchers have integrated various interaction methods such as hand gesture interaction and gaze tracking for remote collaboration. However, it is still difficult to use hand gesture and control gaze direction. Moreover, Toyama et al. studied multi-focus estimation based on support vector regression for optical see-through HMDs [6]. However, depth estimation has still largely deviated from a user's focal length and thus has caused unstable user interaction.

In this paper, we present the evaluation results of gaze-depth estimation using support vector machines (SVM) based on a binocular tracker. We collected eye vectors and eye centers of both eyes. The collected data was analyzed and used for predicting eye-gaze depth using support vector regression (SVR) an SVM. The learned model was then used to analyze the accuracy of the prediction results.

This paper is organized as follows. In Section II, we first describe the prediction procedure and approach. We then introduce the evaluation result of the prediction approach in Section III. Finally we conclude with future work in Section IV.

II. PREDICTION OF EYE-GAZE DEPTH

Several studies have reported on estimations of gaze depth, and they have shown two categories [6]. One involves the use of 3D eye measurements geometrically based on vector intersection, and the other uses the SVR model. The vector intersection approach calculates eye depth by intersecting the eye vectors of the two eyes. However, the depth information from this method is inaccurate due to the unstable convergence of the human eye. The SVR approach uses a support vector machine that estimates the gaze depth by maximizing the margins of the support vectors. However, this regression still has a large estimation error due to the inherent problems of human eye convergence.

To predict gaze depth in a stable manner, the gaze-depth predicting method uses a binocular tracking device. The binocular eye tracker provides the eye position and eye line information of both eyes. Our research uses the position of the center of the eyeball and gaze information in order to predict eye gaze depth. We thus integrated smart eyeglasses and an eye tracker, as shown on the right side of Figure 1.

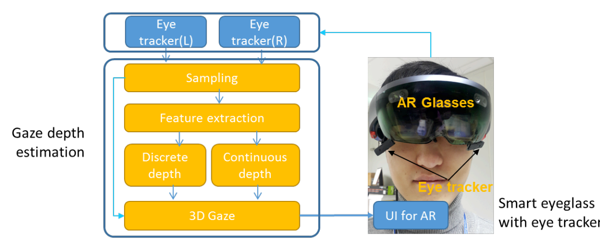

Figure 1. Gaze-depth prediction procedure for wearable 3D interaction

Using this equipment, we collected raw data from the eye tracker to predict eye depth as shown on the left of Figure 1. First, we extracted eye tracker information from the eyes of three test participants and selected features to obtain the gaze depth. Eye movement information includes the pupil position of each eye. Multiple pieces of eye information from the eyes were also collected. From this information, we analyzed and extracted the influential features for eye-depth prediction. Finally, the proposed method estimated the continuous line depth based on the SVR model and predicted the discrete line depth based on the SVM. The support vector model estimated the eye depth based on the maximum

 ISBN: 978-1-61208-570-8

margin of the given sample data. With this model, we trained and tested specimens with eye and depth information.

III. EVALUATION

To evaluate the accuracy of gaze depth in the distance, we installed a test bed consisting of a Pupil Labs binocular eye tracker [7]. This eye tracker has two eye cameras and one world camera. It tracks the eye movement and estimates gaze information of each eye and records world image frames at 120 Hz. We collected pupil and gaze data at focal distances of 1, 2, 3, 4, and 5 meters. There was a small panel located at each of the focal distances, and each user was asked to look at each panel. Three participants were involved in this experiment.

We first analyzed influential features for predicting eye-gaze depth. For this purpose, we calculated the importance information gain of each feature with respect to gaze depth. As seen in Figure 2, among the 17 features analyzed, the position of the center of the eyeball is the feature most highly related to eye-gaze depth. The direction of each eye's gaze is also related to eye-gaze depth.

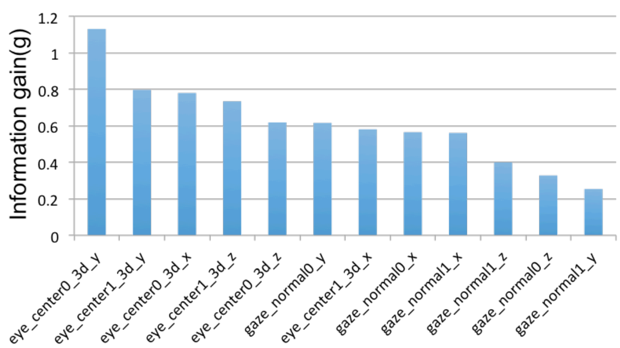

Figure 2. Information gained from features on gaze depth

We then analyzed the performance of continuous gaze-depth prediction obtained from the SVR model. Figure 3 illustrates the relationship between the actual and the predicted gaze depth. The model's prediction of gaze depth to actual focal distance was accurate to within +/- 20 cm (absolute mean error) except for the 2-meter distance.

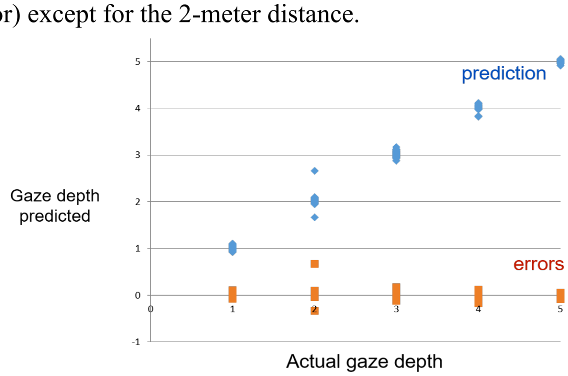

Figure 3. Gaze-depth estimation using the SVR model

Lastly, we evaluated the performance of prediction of discrete gaze depth obtained from the SVM. The overall prediction accuracy was 99% in classifying focal depth. As seen in Table 1, the gaze depths were well classified based on the focal distance. There was only an error in predicting the gaze depth at 4 meters. This indicates that the depth can be predicted by machine learning and discrete prediction might be practically useful for user interaction. However, the results should be tested with a greater number of participants, since only a small number of participants were involved.

TABLE I. CONFUSION MATRIX OF DEPTH CLASSIFICATION FROM SVM

Gaze depth	1	2	3	4	5
1	5971	0	0	0	0
2	0	5874	0	0	0
3	0	0	6415	0	0
4	0	0	0	6109	1
5	0	0	0	0	6449

IV. CONCLUSION

In this paper, we presented the results of the evaluation of gaze depth using a supporting vector machine for natural interaction of wearable AR systems. We set up smart eyeglasses with a binocular eye tracker and then collected pupil and gaze data from the eye tracker. We then evaluated the performance of the gaze-depth prediction based on SVMs.

This work is the first step towards supporting natural gaze interaction based on eye-gaze information for wearable computers. There are still technical problems that need to be improved. We first would like to find features that are more influential in estimating gaze-depth information. We would also like to find more stable and accurate estimation methods for predicting gaze-depth information.

ACKNOWLEDGMENT

This work was supported by the Ministry of Science, ICT and Future Planning (Cross-Ministry Giga Korea Project).

REFERENCES

[1] Vuforia. https://vuforia.com (access date: 2017 Feb. 24)

[2] Kudan. https://www.kudan.eu (access date: 2017 Feb. 24)

[3] Hololens. https://www.microsoft.com/microsoft-hololens/en-us (access date: 2017 Feb. 24)

[4] Y. Lee, K. Masai, K. Kunze, M. Sugimoto and M. Billinghurst, "A Remote Collaboration System with Empathy Glasses," *(ISMAR-Adjunct)*, 2016, pp. 342-343.

[5] M. Billinghurst et al., "Is It in Your Eyes? Explorations in Using Gaze Cues for Remote Collaboration," Collaboration Meets Interactive Spaces, pp.177-199, 2016.

[6] T. Toyama, J. Orlosky, D. Sonntag, and K. Kiyokawa, "A natural interface for multi-focal plane head mounted displays using 3D gaze," In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces* (AVI '14). USA, 25-32.

[7] M. Kassner, W. Patera, and A. Bulling, "Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction," UbiComp Adjunct, pp. 1151-1160, 2014.