Physics-Inspired Input Method for Near-Field Mixed Reality Applications Using Latent Active Correction

Zhenliang Zhang[1], Yue Li[1], Dongdong Weng[1,2,*], Yue Liu[1,2], Yongtian Wang[1,2]

[1] Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing, China, 100081

[2] Advanced Innovation Center for Future Visual Entertainment, 4 Xitucheng Rd, Haidian, Beijing, China, 100088

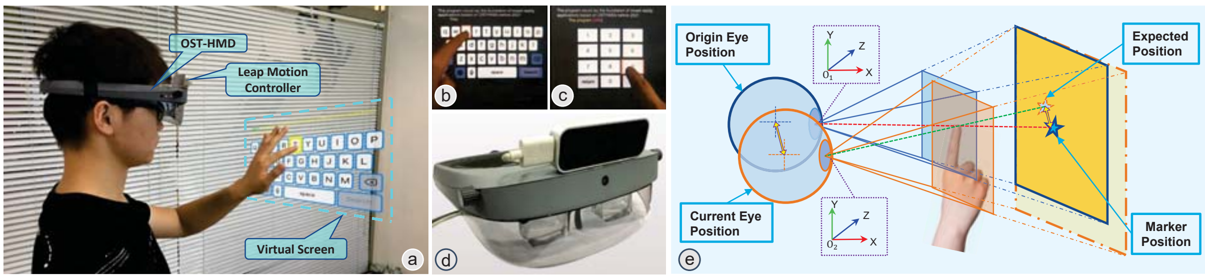

Figure 1: System overview. (a) Direct physics-inspired interaction. (b) First-person view of English keyboard. (c) First-person view of numerical keyboard. (d) The binocular OST-HMD. (e) Dynamic eye-HMD model considering eye movement during use.

Abstract

Calibration accuracy is one of the most important factors to affect the user experience in mixed reality applications. For a typical mixed reality system built with the optical see-through head-mounted display (OST-HMD), a key problem is how to guarantee the accuracy of hand-eye coordination by decreasing the instability of the eye and the HMD in long-term use. In this paper, we propose a real-time latent active correction (LAC) algorithm to decrease hand-eye calibration errors accumulated over time. Experimental results show that we can successfully use the LAC algorithm to physics-inspired virtual input methods.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Graphical user interfaces

1 Introduction

In a typical mixed reality system involving information exchange, the input method is a crucial factor to determine the user experience. For example, the direct physics-inspired interaction can provide the user a similar experience with that in physical world. However, since there is a certain relative movement between the eye and the head-mounted display, a real-time correction method should be proposed to ensure the long-term accurate fusion of virtual and real elements. Most existing calibration methods for a typical hand-eye coordination interaction system only have an off-line calibration module, but it is obvious that the calibrated parameters may change due to eye movement. For instance, Zhang et al. [5] revised the typical manual calibration with the consideration of eye movement. To perform recalibration in real time, Itoh et al. [2] used an eye tracker to estimate the 3D positions of user's eyes, which could modify the parameters according to eye positions. Different from this, we instead plan to do correction without any additional devices.

In this paper, we firstly calibrate the system based on a monocular SPAAM [4]. If needed, it can be transferred to binocular OST-HMDs using the stereo calibration [1]. The calibration is realized based on the alignment of the image point with a single 3D point (such as a

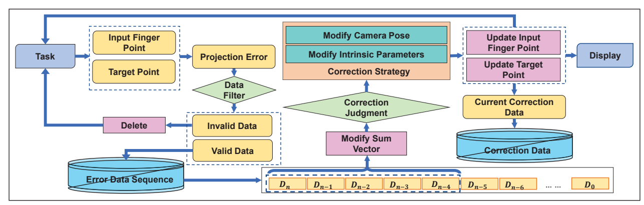

Figure 2: The workflow of real-time latent active correction.

fingertip) in the world coordinate system from various viewpoints. To construct the system, the Leap Motion controller has been used to obtain the hand skeletons, which is similar to Moser's works [3]. Based on this calibration, a real-time latent active correction (LAC) algorithm is proposed to decrease calibration errors accumulated over time. To conclude our contributions based on the system as shown in Fig. 1, we highlight our works as follows. (1) Construct an accurate model to describe the hand-eye coordination process, which simulates the eye-HMD model considering the eye movement. (2) Propose a real-time LAC algorithm to decrease the registration error which comes from the relative movement between the eye and the HMD. (3) Implement a task about text entry and evaluate three input methods regarding their reaction time.

2 Dynamic Hand-Eye Coordination Model

2.1 Details of Dynamic Hand-Eye Model

In near-field mixed reality applications, an eye movement would have a noticeable effect on the registration of virtual and real objects. The hand interaction mentioned here is the direct physics-inspired interaction, which means that the real hand can touch the virtual objects at the level of visual perception. In this way, the hand is required to be accurately registered into the space of mixed reality. This is different from gesture-based interaction methods due to the strong requirements of accurate hand-eye registration.

To provide a comfortable viewing experience, current OST-HMDs generally put the virtual image plane to a far distance, such as 3 meters. However, the hand interaction often occurs in a near distance, which may create a registration error between the hand and the eye once the eye movement occurs. As shown in Fig. 1 (e), if an eye movement occurs, the position of the eyeball is changed and the

*crgj@bit.edu.cn

2018 IEEE Conference on Virtual Reality and 3D User Interfaces
18-22 March, Reutlingen, Germany
978-1-5386-3365-6/18/$31.00 ©2018 IEEE

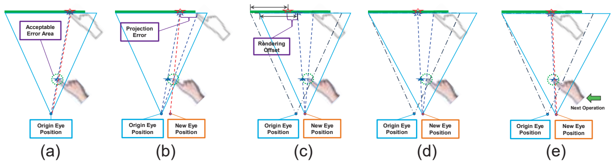

Figure 3: Detailed description of latent active correction. The hollow blue star denotes the projection of the virtual element (filled blue star) with a virtual camera. The hollow red star denotes the projection position of virtual element which users can see. (a) In the initial state, the real hand and the virtual element (filled blue star) are well registered. (b) Once the eye moves, the hand should move in order to ensure the correct registration between the real hand and the virtual element for human eyes. But this will lead to a projection error between the hand and the virtual element. (c) The projection error can be detected, and the center of the virtual camera should be translated according to this error. Notice that the intrinsic parameters of this virtual camera are changed as well. Without the synchronous adjustment of intrinsic parameters, the rendering offset of virtual element can affect the registration. (d) Modify the intrinsic parameters of the virtual camera, then the new position and intrinsic parameters have been changed to be consistent with the actual state. After the correction process, the projection of virtual elements will encounter a little translation, not impairing the interaction obviously. (e) When the next operation occurs, the hand can be brought back to an accurate interaction state.

intrinsic parameters of the eye-HMD model should be modified synchronously.

2.2 Solution for Real-Time Correction

Taking the finger-based interaction for instance, if an eye movement occurs, the finger should be moved to ensure that the fingertip and the marker are aligned from the user's perspective. However, this movement of finger may lead to some spatial registration error for the virtual object and the virtual hand model, which may cause certain failures when the error is accumulated enough. To decrease this kind of error, the proposed LAC should be applied. The workflow of correction is shown in Fig. 2.

Some researches use an external module to capture the pose of human eyes, then compute the real-time error for correction. However, the external module increases the amount of work required for calibration as well as the system complexity. We instead propose a method that utilizes the user's feedback when they are performing tasks. Users will not be asked to stop and do recalibration, which ensures the continuity of tasks.

3 Latent Active Correction

In a task based on hand interaction, the hand-eye calibration should be finished according to SPAAM. Then the online process of LAC is shown in Fig 3. The proposed correction is latent and active for two reasons. The first reason is that the user's operations are implemented to input information, not especially for correction, which is absolutely different from some other applications with independent correction (or recalibration) steps. The system collects errors of every operation and automatically determines whether to trigger a correction action. All correction actions are performed without users' awareness, so it is treated as a latent method. The second reason is that the error of every operation is measured based on the user's manual operations in tasks. The errors of operations, which are computed with the ideal positions on screen and the actual pressed positions on screen, serve as a kind of input data of LAC algorithm. Apparently, it requires many times of confirmations in tasks, such as pressing buttons and selecting small components.

4 Experiments

First, we have constructed a fixed system, as shown in Fig. 4 (b) and (c), to perform all the experiments, so we can control variables

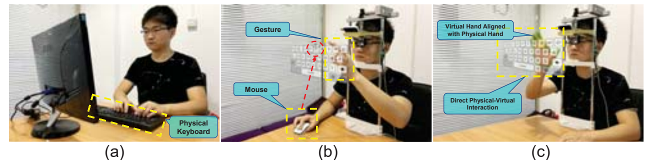

Figure 4: Three input methods. (a) Physical keyboard. (b) Mouse for selection and gesture for confirmation. (c) Direct physics-inspired interaction.

easily. The system contains a monocular OST-HMD to display virtual images, and a Leap Motion controller to track the hands. Second, we use a binocular OST-HMD (NED+ X1), as shown in Fig 1 (d), to verify the application about virtual keyboards.

We have designed a task to test the input methods. There are three input methods for this task, as shown in Fig. 4. In this task, 100 characters, including uppercase letters, lowercase letters and numbers, are randomly shown on the screen. Users are asked to type all the characters with three different methods. First, a real keyboard and a desktop monitor are used to construct the traditional input method. Second, a gesture-based input method is performed. The user should move the cursor using a mouse to select the target button and perform an air-tap gesture to confirm the button. Third, the user should use their real hand to press the virtual buttons, in which LAC is integrated and only one virtual button is triggered in every pressing operation.

We ask 12 users to finish the task using three different ways and record the finish time. The result shows that the mean finish time is 30.50 seconds for the physical keyboard, 230.92 seconds for the gesture-based interaction, and 258.42 seconds for direct physics-inspired interaction. LAC can help the direct physics-inspired interaction get a similar performance with Gesture-Based Interaction, which indicates its value in practice.

5 Conclusion

In this paper, we have proposed a method based on active feedback to correct projection errors in real-time without any additional devices. This method is useful because it does not need any additional device (like eye tracker) to provide assistance. The experimental result shows that the direct physics-inspired interaction has a similar performance with that using gestures, but provides a lower learning cost and appears to be more natural.

Acknowledgments

This work has been supported by the National Key R&D Program of China (No. 2016YFB0401202) and the National Natural Science Foundation of China (No. U1605254).

References

[1] Y. Genc, F. Sauer, F. Wenzel, M. Tuceryan, and N. Navab. Optical see-through hmd calibration: A stereo method validated with a video see-through system. In *Augmented Reality, 2000.(ISAR 2000). Proceedings. IEEE and ACM International Symposium on*, pp. 165–174. IEEE, 2000.

[2] Y. Itoh and G. Klinker. Interaction-free calibration for optical see-through head-mounted displays based on 3d eye localization. In *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, pp. 75–82. IEEE, 2014.

[3] K. R. Moser and J. E. Swan. Evaluation of user-centric optical see-through head-mounted display calibration using a leap motion controller. In *3D User Interfaces (3DUI), 2016 IEEE Symposium on*, pp. 159–167. IEEE, 2016.

[4] M. Tuceryan, Y. Genc, and N. Navab. Single-point active alignment method (spaam) for optical see-through hmd calibration for augmented reality. *Presence: Teleoperators and Virtual Environments*, 11(3):259–276, 2002.

[5] Z. Zhang, D. Weng, Y. Liu, Y. Wang, and X. Zhao. Ride: Region-induced data enhancement method for dynamic calibration of optical see-through head-mounted displays. In *Virtual Reality (VR), 2017 IEEE*, pp. 245–246. IEEE, 2017.