An Empiric Evaluation of Confirmation Methods for Optical See-Through Head-Mounted Display Calibration

[1]Patrick Maier* [2]Arindam Dey† [1]Christian A.L. Waechter‡ [2]Christian Sandor§ [1]Marcus Tönnis¶
[1]Gudrun Klinker‖

[1]Fachgebiet Augmented Reality (FAR)
Technische Universität München
Fakultät für Informatik
Garching b. München, Germany

[2]Magic Vision Lab
University of South Australia
School of Computer and Information Science
Adelaide, Australia

(a) Keyboard (b) Hand-held (c) Voice (d) Waiting

Figure 1: We compared four confirmation methods: (a) *Keyboard*, (b) *Hand-held*, (c) *Voice*, and (d) *Waiting*. *Waiting* was the most accurate in data collection for optical see-through head mounted display calibration. Averaging over time frames further improved the calibration result.

Abstract

The calibration of optical see-through head-mounted displays is an important fundament for correct object alignment in augmented reality. Any calibration process for OSTHMDs requires users to align 2D points in screen space with 3D points in the real world and to confirm each alignment. In this poster, we present the results of our empiric evaluation where we compared four confirmation methods: *Keyboard*, *Hand-held*, *Voice*, and *Waiting*. The *Waiting* method, designed to reduce head motion during confirmation, showed a significantly higher accuracy than all other methods. Averaging over a time frame for sampling user input before the time of confirmation improved the accuracy of all methods in addition. We conducted a further expert study proving that the results achieved with a video see-through head-mounted display showed valid for optical see-through head-mounted display calibration, too.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities H.5.2 [Information Systems]: User Interface; H.1.2 [User/Machine Systems]: Human Factors

1 Introduction

A big challenge in Augmented Reality (AR) is to achieve a seamless integration of virtual objects into the real world. Optical see-through head mounted displays (OSTHMD) require a high level of user interaction for calibration. Researchers have presented calibration mechanisms for OSTHMD such as the Single Point Active Alignment Method (SPAAM) [5] and the Display Relative Calibration (DRC) [4] which use 2D-3D correspondences to create the projection matrix. After the alignment of a 2D point on the display with a 3D point in the real world, users have to acknowledge the correspondence by usually pressing a key on the keyboard. We assert this acknowledgment method plays a major role on the human performance during the process, as it requires some degree of hand coordination forcing a misalignment of the 2D point with the 3D point, increasing the probability to capture inaccurate data.

In this poster, we present three alternative acknowledgment methods, *Hand-held* (pressing a button on a hand-held device), *Voice* (verbally reporting) and *Waiting* (keeping the head steady for 0.5 second) and the most often used method *Keyboard* (pressing a key on the keyboard) and evaluated them in a user study.

Our results showed that the *Waiting* method outperformed all presented acknowledgment methods (see Figure 2). We also investigated different time frames in which we average the collected data to get better results. The different time frames end at the time of confirmation $[t]$ with increasing time frame sizes in steps of $0.1sec$ until $2sec$ before the time of confirmation. We found that averaging the input data in the window of $[t-0.6sec,t]$ resulted in the most accurate alignment while the current approach of taking the correspondence at the time of acknowledgment $[t]$ resulted in the worst accuracy.

2 Background

Previous works investigate aspects of calibration procedures that are related to human behavior. The user's inability to maintain a stable pose mainly is the most relevant parameter when collecting point to point correspondences. Such effects have already been mentioned right after algorithms like SPAAM have been published.

*e-mail: maierp@in.tum.de
†e-mail: arindam.dey@unisa.edu.au
‡e-mail: waechter@in.tum.de
§e-mail: christian@sandor.com
¶e-mail: toennis@in.tum.de
‖e-mail: klinker@in.tum.de

IEEE International Symposium on Mixed and Augmented Reality 2011
Science and Technology Proceedings
26 -29 October, Basel, Switzerland
978-1-4577-2184-7/10/$26.00 ©2011 IEEE

McGarrity et al. [3] stated that the user must be factored in when calibrating an HMD. Errors are induced by users because calibration procedures involve manual steps. Earlier work by Axholt et al. [1] investigated postural stability during the calibration process.

The work by Axholt et al. [2] lets users aim at a correlation point for at least $2s$ with low differences in the HMD rotation. All data was collected over this time period. If the rotation of the HMD changed more than $0.19°$ per sample, the data was discarded.

3 Experiment

The experiment aimed at evaluating the influence of the different acknowledgment methods on the accuracy of the input data (correspondence points) for the OSTHMD calibration process.

To measure differences in the confirmation methods, we set up a scenario where the users should collect 2D-3D point correspondences. To be able to calculate aiming errors w.r.t. baseline data, we simulate the OSTHMD calibration process by incorporating a VSTHMD. To analyze the aiming error, the camera recorded 3D target positions on the display are estimated using a computer vision algorithm. These estimations provide us with reliable information about the 2D display positions at which the users should have aimed. This baseline data is used to calculate the residual error in comparison to the user acknowledged 2D positions.

24 participants aged between 23 to 53 years (M=28.9, SD=6.05) had to align a cross hair with the 3D target by moving the head and body to get to an alignment. Once this alignment is established, they had to use one of our four methods to acknowledge the alignment. Each user had to generate 180 correspondence point pairs.

4 Results

We did not only examine the residual errors at the time of confirmation t but also the average residual error in 20 time frames (each ending at the time of confirmation $[t]$ and each with an increasing time span by $0.1sec$).

The ANOVA reported a significant main effect of the acknowledgment methods $F(3,428) = 22.07; p < .001; \eta_p{}^2 = .13$ on residual error (see Figure 2). A Tukey's HSD post-hoc test revealed that among all methods *Waiting* outperformed and *Keyboard* was worst.

ANOVA found a significant main effect of time frames $F(1.087,465.047) = 127.14; p < .001; \eta_p{}^2 = .23$ on the residual error. The time frame $[t-0.6sec,t]$ was best overall and a significantly better window than all time frames except $[t-0.5sec,t]$ and $[t-0.4sec,t]$. The error calculated at the point of acknowledgment is significantly ($p < .001$) worse than the time frames ranging from $[t-0.9sec,t]$ to $[t-0.1sec,t]$ (see Figure 2).

We can see that the *Waiting* method produces higher accuracy than currently used keyboard based methods. We have found, OSTHMD calibration data must be collected in longer time frames to achieve higher accuracy. Collecting data just at the point of calibration is an inaccurate practice.

5 Discussion and Conclusion

We focused on the human factor in collecting more accurate input data for the calibration process and found that averaging the collected data over time frames results in better calibrations. This behavior can be explained by a oscillation movement of the users head while aiming at the target. When averaging over the points where the user aimed, we get a mean point that lies closer to the target spot.

We explain the good results of the *Waiting* method in the following way. The users had to concentrate on the target spot trying not to move a lot. This reduces the ability to do sloppy confirmations and also calms down the user.

To validate the gathered results we use two different methods. In a first validation we proofed the results by a numerical analysis of the user generated calibrations with the ground truth calibration

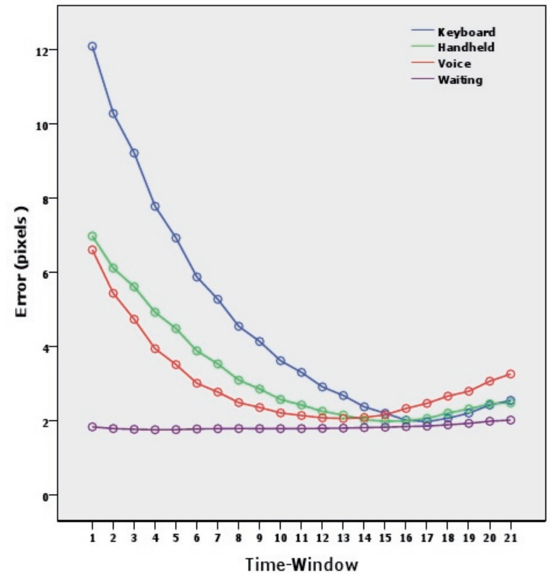

Figure 2: Time frame 1 on X-axis above represents $[t-2.0sec,t]$, successive windows are at 0.1 seconds intervals thereafter. Time frame 21 represents $[t]$.

generated from the computer vision data of the previous experiment. We projected a point cloud on the one hand with the ground truth projection matrix and on the other hand with the user generated projection matrices. When comparing the deviations to the ground truth projections, we found that *Waiting* was significantly better than *Keyboard* ($p < .001$) and *Voice* ($p = .025$).

The second validation procedure proved that the achieved results can be transferred from our VSTHMD experiment to OSTHMDs. Five experienced users did the same calibration procedure as described earlier. The users then had to compare the quality of four augmentations generated by the above created projection matrices. From user feedback we found that *Waiting* was the best method for everyone and *Voice* was the worst. This confirms that results of our experiment using a VSTHMD are applicable to OSTHMDs.

Implementing the presented contributions for calibration processes for OSTHMD will improve quality of object alignment in AR systems.

References

[1] M. Axholt, S. Peterson, S. Ellis, et al. Visual Alignment Accuracy in Head Mounted Optical See-Through AR Displays: Distribution of Head Orientation Noise. In *Human Factors and Ergonomics Society Annual Meeting Proceedings*, volume 53, pages 2024–2028. Human Factors and Ergonomics Society, 2009.

[2] M. Axholt, M. A. Skoglund, S. D. O'Connell, M. D. Cooper, S. R. Ellis, and A. Ynnerman. Parameter Estimation Variance of the Single Point Active Alignment Method in Optical See-Through Head Mounted Display Calibration. In *Proceedings of the IEEE Virtual Reality Conference*, 2011.

[3] E. Mcgarrity, Y. Genc, N. Navab, M. Tuceryan, and C. Owen. Evaluation of calibration for optical see-through augmented reality systems. 2001.

[4] C. B. Owen, J. Zhou, A. Tang, and F. Xiao. Display-relative calibration for optical see-through head-mounted displays. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '04, pages 70–78, Washington, DC, USA, 2004. IEEE Computer Society.

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