Gaze-Based Human-SmartHome-Interaction by Augmented Reality Controls

Tim Cottin[✉], Eugen Nordheimer, Achim Wagner,
and Essameddin Badreddin

Institute of Computer Engineering, Automation Laboratory,
University of Heidelberg, Heidelberg, Germany
{tim.cottin,eugen.nordheimer,achim.wagner,
badreddin}@ziti.uni-heidelberg.de

Abstract. The use of eye tracking systems enables people with motor disabilities to interact with computers and thus with their environment. Combined with an optical see-through head-mounted display (OST-HMD) it allows the interaction with virtual objects which are attached to real objects respectively actions which can be performed in the SmartHome environment. This means a user can trigger actions of real SmartHome actuators by gazing on the virtual objects in the OST-HMD. In this paper we propose a mobile system which is a combination of a low cost commercial eye tracker and a commercial OST-HMD. The system is intended for a SmartHome application. For this purpose we proof our concept by controlling a LED strip light using gaze-based augmented reality controls. We show a calibration procedure of the OST-HMD and evaluate the influence of the OST-HMD to the accuracy of the eye tracking.

Keywords: User interfaces · Gaze based SmartHome control · Optical feedback system · Eye-tracking · Augmented reality controls

1 Introduction

People who suffer from motor disabilities, e.g. locked-in-syndrome or amyotrophic lateral sclerosis, cannot use conventional computer input devices. With the use of eye-tracking systems these people are able to generate input commands through their eye movement, which enables them to interact with their environment. For example communication systems can be built up by displaying a virtual keyboard on a monitor. The user selects letters by gazing on them to formulate words and phrases. Such systems depend on the use of monitors which are placed in front of the user. The disadvantage of using conventional computer monitors is that the user's sight is impaired. This problem can be solved by the use of optical see-through head-mounted displays (OST-HMD). Furthermore, OST-HMDs allow the superposition of real objects with virtual objects to generate an augmented reality (AR). In addition, by combining an

eye-tracking system with an OST-HMD AR system, the user is able to interact with SmartHome actuators in the environment by means of gaze on virtual ones representing actions the SmartHome actuator can perform.

There are three main areas of relevant works regarding our research. The first is eye tracking in human-computer-interfaces. The use of eye tracking systems to interact with a computer is the focus of many researchers. In [8] Majaranta and Bulling give a good overview of the recent development in this field.

The second area is the integration of eye tracking systems and HMDs. Toyama et al. present in [13] a wearable integration of an eye-tracker and a see-through head-mounted display. The system tracks the users gaze position during document reading. The OST-HMD provides additional information about specific key words the user is gazing on. The user can interact with the display by gazing on elements shown in the display. The elements are triggered after a gazing dwell time of 2 s. In [12] Sonntag and Toyama present a similar system; they use gaze data and object recognition to register objects for location awareness. The OST-HMD provides navigation instructions. In [9] Nilsson et al. use a video ST-HMD with an integrated eye tracking system; they use marker tracking to highlight virtual information about corresponding real objects. The system is used for instructional applications and allows the user to respond to instructions or questions by gazing on virtual elements. Recently, commercial systems combining eye tracking and HMDs became available [7,10,11].

The last area is the interaction with the environment respectively home appliances using AR. An early work in this area is described in [4]. The user's field of view is determined by a head tracking sensor. Informations and control options of home appliances in the user's field of view are highlighted in a OST-HMD. The user needs to press a mouse button to trigger control options. In [14] Ullah et al. propose a system for elderly or physically disabled people to remote control home appliances using a smartphone. The system tracks QR codes with the smartphone for home appliances recognition. Control options and informations about the recognized home appliances is shown on the smartphone and can be selected using the touchscreen of the smartphone. A similar system is described in [5]. To the best of our knowledge no system exists which combines eye tracking and OST-HMD AR in a home automation application.

This paper presents a new interface for Human-SmartHome-Interaction. The first prototype was realized during a student project of Cottin [1]. The enhanced and embedded version is a first result of the AICASys Project [3]. We present a prototype system consisting of a wearable eye tracker and an AR OST-HMD. The system enables the user to trigger "on-off" actions of real objects in a domestic environment by gazing at dedicated virtual elements which are attached to the real object. Furthermore the OST-HMD gives visual feedback of the available actions of an object and simultaneously displays the measured gaze position in display coordinates. The proposed system consists of commercial hard- and software components which can be easily integrated, are lightweight and allow a mobile setup. For the proof of our concept we used the system to control a LED strip light as an example of a SmartHome application.
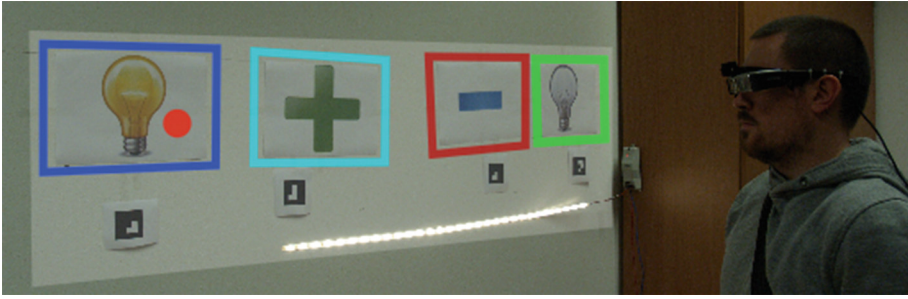
Fig. 1. User wearing the system. The augmented reality controls of the actions are depicted as frames around the real object in the OST-HMD. Additionally the Point of Gaze is displayed in the OST-HMD.

Figure 1 shows a user wearing the system. The user sees augmented reality controls represented as frames and the computed Point of Gaze (PoG) in the OST-HMD. By gazing on the virtual controls the user can control the LED strip light.

2   Implementation Details

At the time of this study there were no commercial systems available which combined an eye tracker and an OST-HMD. Therefore we built up the system with separate eye tracker and OST-HMD and integrated them by ourselves. The *Pupil Pro* eye tracking system [10] is an affordable wearable eye tracker and allows the integration with an OST-HMD because of the open physical design.
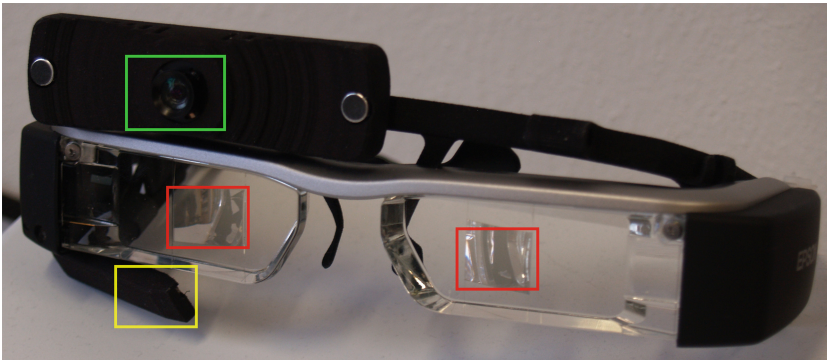

Fig. 2. Integrated Pupil Pro eye tracker with the Epson Moverio BT-200. Green frame: scene camera. Yellow frame: eye camera. Red frames: optical see-through display

For the OST-HMD we chose the *Epson Moverio BT-200* [2], which is a binocular OST-HMD and can be physically integrated with the *Pupil Pro* eye

tracker. Figure 2 shows the physically combined systems. We chose an *ODROID XU3 Lite* single board computer as processing unit for the eye tracker. It is a creditcard sized computer to allow a mobile setup. The home automation application is a LED strip light controlled by a FS20 DC-LED-dimmer. The LED-dimmer can be controlled with a FS20 remote control, connected to the *ODROID XU3 Lite.* With the dimmer one can perform on, off and dimming actions. Figure 3 depicts a block diagram showing the main components of the system. Additionally the information flow with the user and the environment is given.
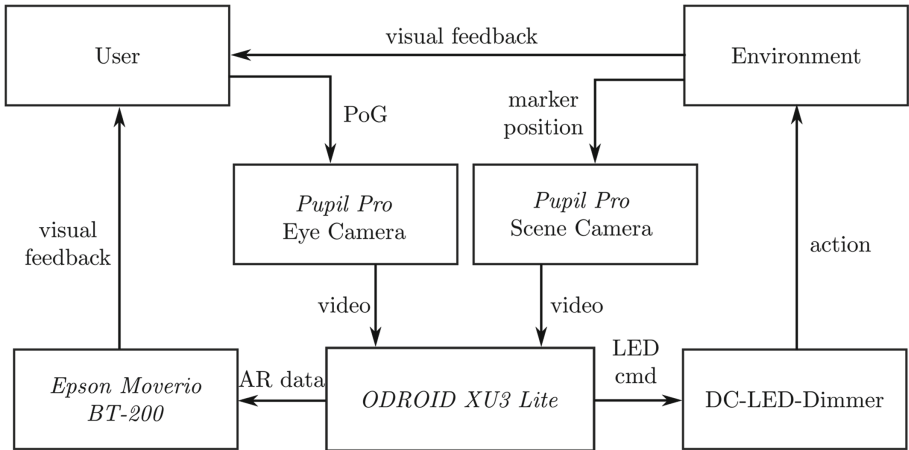

Fig. 3. The systems components and information flow

The software of the *Pupil Pro* eye tracking system is used to calculate the gaze position of the user. The user's pupil respectively its centre is detected by illuminating the eye with an IR LED, finding the dark pupil region in the eye camera image and approximating it with an ellipse [6]. The gaze position in scene camera coordinates is a function of the pupil position in the eye camera image and the function parameters are determined by point correspondence calibration [6].

Furthermore, the *Pupil Pro* software provides a fiducial marker tracking in the scene camera image. To each marker a so called surface can be attached. This surface is a virtual object and can be placed arbitrarily relative to its marker. In our system we use these surfaces and connect them with control inputs of our DC-LED-dimmer. By calculating the PoG relative to a surface we can determine if the user is gazing on it. With this mechanism the user can trigger actions by gazing on the surfaces with a dwell time. We have 4 different surfaces for the 4 actions: LED on, LED off, LED dimm lower, LED dimm higher. The surfaces are displayed in the OST-HMD if the corresponding marker is detected in the scene camera image and the surface overlaps the display (see Fig. 1). The user

is provided with visual feedback of triggered actions in the OST-HMD and by the LED strip light itself.

Figure 4 shows the geometrical relationships between the coordinates of the scene camera $\{S\}$, the display $\{D\}$, the eye camera $\{E\}$, the marker and a surface $\{Su\}$ from the users view.
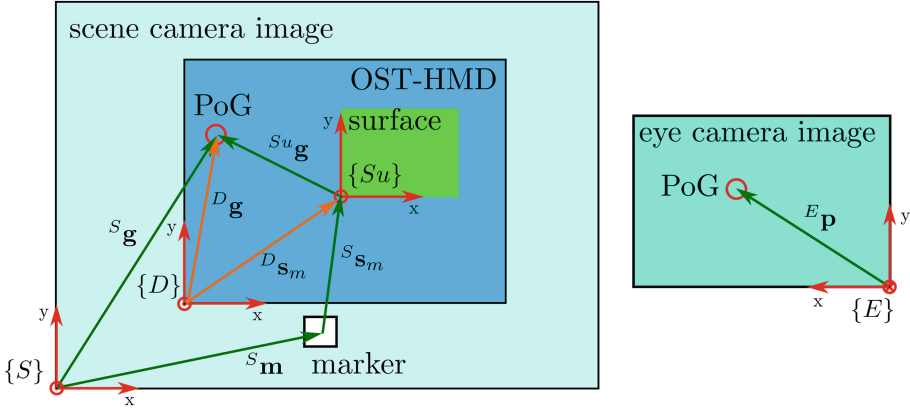

Fig. 4. Geometrical relationships between the scene camera $\{S\}$, display $\{D\}$, eye camera $\{E\}$, marker and surface $\{Su\}$ from the user's view [1]. The green vectors are calculated by the Pupil Pro software. The orange vectors are required.

The green vectors are known and provided by the *Pupil Pro* software. This are the pupil position in the eye camera image $^E\mathbf{p}$, the PoG $^S\mathbf{g}$, the marker position in scene camera coordinates $^S\mathbf{m}$, the surface relative to the marker $^S\mathbf{s}_m$ and the PoG in surface coordinates $^{Su}\mathbf{g}$.

3   Referencing the OST-HMD in the Scene Camera

To display the PoG and the surface in the OST-HMD the vectors $^D\mathbf{g}$ and $^D\mathbf{s}_m$ are required (see Fig. 4). For this purpose the transformation from scene camera coordinates to OST-HMD coordinates need to be calibrated. The display is physically attached to the frame of the *Pupil Pro* eye tracker. Thus the position of the OST-HMD relative to the scene camera coordinate system is constant and the calibration needs to be done only once. We use a simple point correspondence calibration procedure. The user has to sit in front of a computer monitor where 5 reference markers are shown. In the OST-HMD there are 5 markers shown which have to be aligned by the user with the ones on the monitor. Figure 5A shows a user during the calibration process. In B the detected markers in the scene camera image $^S\mathbf{m}_{1,\dots,5}$ (orange markers) and the markers shown in the display $^D\mathbf{m}_{1,\dots,5}$ (green markers) are depicted.

The transformation is given by

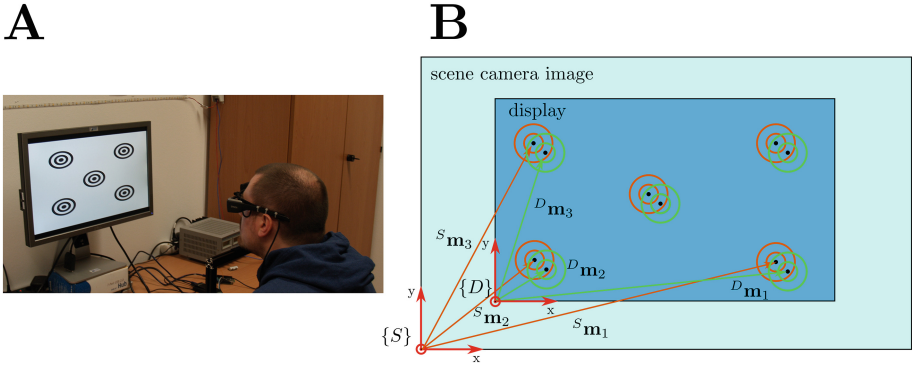$$^D\mathbf{m} = \mathbf{T} \begin{bmatrix} ^S\mathbf{m} \\ 1 \end{bmatrix} \tag{1}$$

A

B

Fig. 5. A: user during the calibration procedure. **B**: scene camera image and display with the corresponding markers. [1]

with **T** the 3×4 transformation matrix. Combining the known points in this equation yields

$$\begin{bmatrix} ^D\mathbf{m}_1 \ldots {}^D\mathbf{m}_5 \end{bmatrix} = {}^D\mathbf{M} = \mathbf{T} \begin{bmatrix} ^S\mathbf{m}_1 & \ldots & ^S\mathbf{m}_5 \\ 1 & \ldots & 1 \end{bmatrix} = \mathbf{T}\,{}^S\mathbf{M} \qquad (2)$$

This can be solved by computing the pseudo inverse $^S\mathbf{M}^+$ of $^S\mathbf{M}$. Now **T** can be determined

$$\mathbf{T} = {}^D\mathbf{M}\,{}^S\mathbf{M}^+ \qquad (3)$$

4   Experiments and Results

To evaluate the system we performed two experiments. The first one is an accuracy test for the eye tracker with real markers. The user sits in front of a computer monitor at a distance of $50\,\mathrm{cm}$. After the calibration the user is asked to gaze on random markers shown on the computer screen. The second experiment is an accuracy test using virtual markers which superimpose real markers shown on the monitor. In this experiment the OST-HMD is covered so that the user does not see the real marker. 5 users performed the experiments.

The angular error between the calculated gaze point and the detected point on the screen can be used as a measure for the accuracy of an eye tracking system. The calculation of the accuracy is taken from the *Pupil Pro* software. The mean angular accuracy $\bar{\alpha}$ of n samples of PoG respectively $^S\mathbf{g}_{1,\ldots,n}$ to the marker positions $^S\mathbf{m}_{1,\ldots,n}$ is given by

$$\bar{\alpha} = \frac{d_{fov}}{n\sqrt{w^2 + h^2}} \sum_{i=1}^{n} \left\| ^S\mathbf{g}_i - {}^S\mathbf{m}_i \right\| \qquad (4)$$

where d_{fov} is the diagonal field of view of the scene camera, w and h the pixel width and the pixel height of the scene camera.
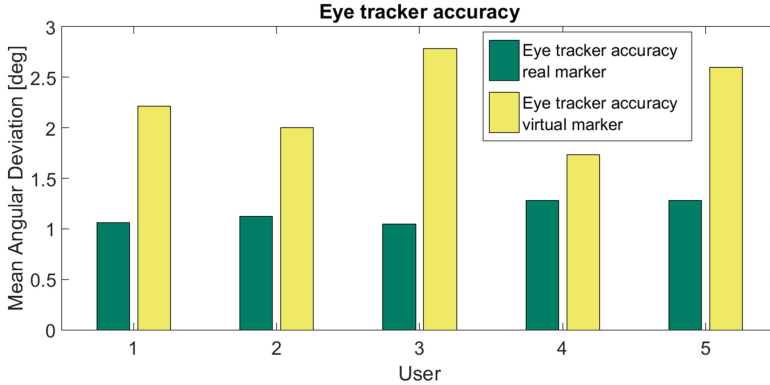
Fig. 6. Bar plot of the accuracy test results.

Figure 6 shows the results of the experiments. The angular accuracy of the eye tracker itself during the first experiments (green) is for all users about $1°$. The second experiments (yellow) shows the influence of the OST-HMD calibration. It varies widely for the individual users which shows the disadvantage of user intervention during the calibration procedure.

5   Conclusion

Our hardware setup shows that the combination of an eye tracker and an OST-HMD is possible with commercial and low cost components. We trigger the actions by evaluating the gaze point relative to the surfaces and use a dwell time to infere the users intend. The usage of the OST-HMD in this application brings an important advantage. The surfaces are virtual objects and can be positioned arbitrarily relative to the interactive objects which are used as augmented controls. Because they are displayed in the OST-HMD the user must not necessarily know the position of the surfaces in his environment.

The calibration procedure of the display needs to be done only once. But it is unfavourable because it is time consuming and requires a highly focused user to align the markers in the display with the ones shown on the monitor manually. We see this as the main drawback of the proposed system. Future works need to tackle this problem and propose a calibration procedure which does not need user intervention.

Another goal for future works is the use of a object recognition algorithm based on natural features like SIFT instead of using markers. Furthermore to encounter the Midas Touch problem there has to be a focus on alternative trigger mechanisms by user gaze.

References

1. Cottin, T.: Aufbau und Implementierung eines tragbaren, blickbasierten Steuerungssystems zur Interaktion mit der Umgebung (2014)
2. Epson: Epson moverio bt-200. http://www.epson.de/de/de/viewcon/corporatesite/products/mainunits/overview/12411
3. German Ministery of Education, Research: AICASys. http://www.mtidw.de/ueberblick-bekanntmachungen/ALS/aicasys
4. Hammond, J.C., Sharkey, P.M., Foster, G.T.: Integrating augmented reality with home systems. In: Proceedings of the 1st International Conference on Disability, Virtual Reality and Associated Technologies (ECDVRAT), pp. 57–66 (1996)
5. Heun, V.M.J.: Smarter objects: Programming physical objects with AR technology (2013)
6. Kassner, M., Patera, W., Bulling, A.: Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. ArXiv e-prints, April 2014
7. Lusovu: Lusovu eyespeak. http://www.myeyespeak.com/
8. Majaranta, P., Bulling, A.: Eye tracking and eye-based human-computer interaction. In: Fairclough, S.H., Gilleade, K. (eds.) Advances in Physiological Computing. Human–Computer Interaction Series, pp. 39–65. Springer, London (2014)
9. Nilsson, S., Gustafsson, T., Carleberg, P.: Hands free interaction with virtual information in a real environment: Eye gaze as an interaction tool in an augmented reality system. PsychNology J. **7**(2), 175–196 (2009)
10. Pupil Labs: Pupil labs eye-tracking systems. https://pupil-labs.com/
11. SMI: Smi eye tracking upgrade for ar glasses. http://www.smivision.com/augmented-reality-eyetracking-glasses/
12. Sonntag, D., Toyama, T.: Vision-based location-awareness in augmented reality applications. In: 3rd International Workshop on Location Awareness for Mixed and Dual Reality (LAMDa 2013), pp. 5–8 (2013)
13. Toyama, T., Dengel, A., Suzuki, W., Kise, K.: Wearable reading assist system: Augmented reality document combining document retrieval and eye tracking. In: 12th International Conference on Document Analysis and Recognition (ICDAR), pp. 30–34, August 2013
14. Ullah, A.M., Islam, M.R., Aktar, S.F., Hossain, S.K.A.: Remote-touch: Augmented reality based marker tracking for smart home control. In: 15th International Conference on Computer and Information Technology (ICCIT), pp. 473–477, December 2012