Schau genau! A Gaze-Controlled 3D Game for Entertainment and Education

Raphael Menges[1] Chandan Kumar[1] Ulrich Wechselberger[1]

Christoph Schaefer[2] Tina Walber[2] Steffen Staab[1]

[1]Institute of Web Science and Technologies
University of Koblenz-Landau, Germany

[2] EYEVIDO GmbH
Koblenz, Germany

Eye tracking devices have become affordable. However, they are still not very much present in everyday lives. To explore the feasibility of modern low-cost hardware in terms of reliability and usability for broad user groups, we present a gaze-controlled game in a standalone arcade box with a single physical buzzer for activation. The player controls an avatar in appearance of a butterfly, which flies over a meadow towards the horizon. Goal of the game is to collect spawning flowers by hitting them with the avatar, which increases the score. Three mappings of gaze on screen to world position of the avatar, featuring different levels of intelligence, have been defined and were randomly assigned to players. Both a survey after a session and the high score distribution are considered for evaluation of these control styles. An additional serious part of the game educates the players in flower species, who are rewarded with a point-multiplier for prior knowledge. During this part, gaze data on images is collected, which can be used for saliency calculations. Nearly 3000 completed game sessions were recorded on a state horticulture show in Germany, which demonstrates the impact and acceptability of this novel input technique among lay users.

Keywords: eye tracking, eye movement, gaze, new media, serious game, user interface, low-cost eye tracking devices

Introduction

As of today, low-cost eye tracking hardware like the *Tobii EyeX* (Tobii Gaming, 2017) is available for end users. These affordable devices offer easy to use and stable eye tracking performance. However, most games utilize gaze data only as additional input source, while still relying on mouse and/or keyboard or gamepad input (Isokoski & Martin, 2006; Isokoski, Joos, Spakov, & Martin, 2009; Nacke, Stellmach, Sasse, & Lindley, 2010), or perform gaze-based emulation of the traditional devices (Istance, Bates, Hyrskykari, & Vickers, 2008; Istance, Hyrskykari, Immonen, and Mansikkamaa, & Vickers, 2010). In this work, we primarily focus on two questions: How is a game solely controlled by gaze accepted and performed by lay users; and how to interpret the gaze on the screen within the gaming environment. For this purpose, we developed a gaze-controlled game and placed it in a standalone arcade box on the state horticulture exhibition 2015 in Landau, Germany, for a period of six month. Multiple disciplines of eye tracking research (Kumar, Menges, & Staab, 2016) were combined to provide a wholesome user experience. Players were introduced to eye tracking, fulfilled a 5-point calibration, selected the language via dwell time buttons and entered a nickname for the high score list using a gaze-based text input mechanism – inspired by Dasher (Ward, Blackwell, & MacKay, 2000). Key features of this work:

- Interface and avatar controlled solely by user's gaze
- Three approaches of mapping gaze to avatar position
- Educational part including flower species information, and collection of users' fixation data on images

The 2017 COGAIN Symposium: Communication by Gaze Interaction Wuppertal, Germany. August 19th and 21st, 2017

Methods

The game is subdivided into five consecutive states, as shown in Figure 1: After an idle screen and introduction about eye-based input (a), which are controlled with a single physical buzzer, the player is guided through a 5-point calibration process. Then, the tutorial screen (b) presents the game mechanics, where either German or English localization can be selected by the fixation of dwell time based buttons. The actual game state loads and the avatar is presented (c). After a session, the player is asked to enter her nickname for the high score table, using the onscreen keyboard (d). This keyboard consists of the Latin alphabet including German umlauts in uppercase and is displayed at top of the screen. On fixation, the chosen letter starts to move downwards while increasing in size. When a certain vertical offset is reached, the letter is typed and the keyboard is reset. After the gaming session, a qualitative survey is presented to some players (e).

Game State

The game state itself is split into the *flower mode*, which is the initial mode, and the *picture mode*. During flower

Figure 1: States of the proposed game. Apart the idle screen in (a), all states are controlled solely by gaze input.

mode, the avatar is flying over a meadow towards the horizon. While the movement towards the horizon is handled automatically by the game, the player defines the position of the avatar on a screen-parallel plane by gaze. Flowers spawn at the horizon line. If both avatar and flower intersect when the flower cuts the interaction plane of the avatar, the flower is collected and points are added to the score. The game is ended when the avatar hits a spider web. They are spawned analogously to the flowers.

Control Styles

We have defined three interpretations of gaze to position the avatar: First approach is *direct* interpretation of gaze coordinates as avatar position (like emulation of a mouse pointer); second approach features a *grid*-based positioning (similar to fixation smoothing while emulating the mouse pointer); third approach is a mechanism that supports the player by positioning the avatar *indirectly* on a calculated position if available, otherwise uses direct interpretation. This calculated position is the future collision point with the flower currently fixated by the player (adding intelligence to eye-controlled interfaces). The styles were randomly assigned at the beginning of a player's session.

Serious Game

After a certain interval of time, the game state switches into picture mode, where the player is presented with two images containing each a unique flower species. The player is asked to fixate the image displaying the flower species that corresponds to the species' name shown on top of the screen. For a correct selection, the multiplier for further points is increased, otherwise it is reset. Through an instant feedback after final selection, the player's choice is either confirmed or declined, as the incorrect picture fades away and the correct one is displayed for a short amount of time. During the selection process, fixation data is recorded for later saliency computations on the images included in the game.

See the online video for a demonstration of one session on the standalone box: https://youtu.be/WOcb94t6BaQ

Results

The setup ran stable, was self-explanatory and no technical problems were reported during the six months' exhibition period. Furthermore, the exhibition volunteers reported exceedingly positive feedback from the visitors. Nearly 3000 sessions with nickname input were performed. The game seemed to be fun, as various players did multiple runs to improve their ranking. For analysis, 976 different players were randomly selected for evaluation. They were asked for their age (categories included "0–16", "17–30" and "30+") and if they wore visual aid (response options were "no", "glasses" and "contact lenses"). Together with the randomly assigned control style and the achieved high score (as an indicator for controllability of the game) the responses were saved for later evaluation. A one-way analysis of variance with SPSS 23 (α=0.05) revealed that the chance of earning a high score with eye tracking control was in no way significantly affected by visual aid ($F(2)=0.39$; $p=0.68$). The age group of the 17–30 year old achieved the highest score and the 31+ generation the lowest – a highly significant, yet small effect ($F(2)=16.39$; $p<0.001$; $\eta^2=0.03$). The control style had a statistically significant, yet very small effect on high scores: games with direct positioning led to the lowest, games with indirect positioning to the highest scores ($F(2)=4.18$; $p=0.02$; $\eta^2=0.01$).

Discussion and Conclusion

This initial analysis of sessions and survey feedback indicate high acceptance of eye tracking as input mechanism and an advantage for the intelligent indirect control style. In particular, visual aid does not seem to be any obstacle for eye tracking based game control and the player's age does not make a practically relevant difference for eye gaze control. In the future, we will report the detailed investigations of the collected data from scores, questionnaires, and image fixations including a larger data set and additional variables such as perceived control and game enjoyment. The usage and feedback showcase the significance and feasibility of eye tracking as an input device for a broad range of end users, as our experiment demonstrates an overall good acceptance and performance. The complete procedure of calibration, interface control, gaming and nickname entering was performed by hundreds of lay users.

Acknowledgements

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

Isokoski, P., & Martin, B. (2006). Eye tracker input in first person shooter games. *Proceedings of the 2nd Conference on Communication by Gaze Interaction: Communication by Gaze Interaction-COGAIN 2006: Gazing into the Future*, (pp. 78-81).

Isokoski, P., Joos, M., Spakov, O., & Martin, B. (2009). Gaze controlled games. *Universal Access in the Information Society*, 323.

Istance, H., Bates, R., Hyrskykari, A., & Vickers, S. (2008). Snap Clutch, a Moded Approach to Solving the Midas Touch Problem. *ETRA* (pp. 221-228). Savannah, Georgia: ACM. doi:10.1145/1344471.1344523

Istance, H., Hyrskykari, A., Immonen, L., and Mansikkamaa, S., & Vickers, S. (2010). Designing Gaze Gestures for Gaming: An Investigation of Performance. *ETRA* (pp. 323-330). Austin, Texas: ACM. doi:10.1145/1743666.1743740

Kumar, C., Menges, R., & Staab, S. (2016). Eye-Controlled Interfaces for Multimedia Interaction. (B. Susanne, Ed.) *IEEE Multimedia*.

Nacke, L. E., Stellmach, S., Sasse, D., & Lindley, C. A. (2010). Gameplay experience in a gaze interaction game. *CoRR*, 49-54.

Tobii Gaming. (2017, April). Retrieved from Tobii: https://tobiigaming.com

Ward, D. J., Blackwell, A. F., & MacKay, D. J. (2000). Dasher - a data entry interface using continuous gestures and language models. In M. S. Ackerman, & W. K. Edwards (Ed.), *UIST* (pp. 129-137). ACM. doi:10.1145/354401.354427

.