S. Boll, S. Maaß & R. Malaka (Hrsg.): Mensch & Computer 2013
München: Oldenbourg Verlag, 2013, S. 333– 336

Interaction Spaces: Interactive Spatial Areas to Control Smart Environments

Stefan Schneegass, Bastian Pfleging, Tilman Dingler, Albrecht Schmidt

Institute for Visualization and Interactive Systems, University of Stuttgart

Abstract

Throughout recent years new input modalities found their way into consumer electronics. Recognizing body posture and gestures in the three dimensional space is now possible using hardware that is available for about 100 EUR. We aim at providing a system to convert any environment into an interactive space. Hence, we created a system that is able to detect the user's body in three dimensions and to determine the presence of body parts at pre-defined/user-defined locations in order to trigger actions of the environment. We built a first Kinect-based prototype where users can define trigger areas and link them to suitable actions. We then conducted a study to evaluate the usability of the system and how size and memorability of spaces affect user performance with regard to trigger area tasks. Results show that with increasing area size the task completion time goes down while error rates go up.

1 Introduction

In recent years we have moved away from looking at computer input as a simple combination of keyboard and mouse interaction. Just like computing devices have changed, modalities of input have been adopted. Multi-Touch input, stylus interaction, speech, and gesture recognition are examples for how input technologies adapted to new demands to interact with computers that have become smaller and ubiquitous. The multitude of devices brings challenges for a comprehensive user interface, which bears the question: How does human input and control look like across the plethora of smart devices?

We started exploring the space that surrounds our home appliances as potentially interactive areas. The idea is to trigger actions, such as control the music or close the blinds by performing gestures in certain spatial regions. This way, appliances can be controlled from remote locations (i.e., while sitting on the couch) and special areas in the room can be mapped to controlling certain appliances by performing gestures with various body parts. Different approaches have been developed in order to enrich things of everyday life, as for instance meeting rooms (e.g., Johanson et al. 2002). Early projects focused at investigating a suitable middleware (Johanson et al. 2002) or features such as tracking people (e.g., Krum et al. 2000). Mobile phones or tablets are often used as input means. Peters et al. (2011) used a mobile phone to control the lights and window blinds with gestures. However, the usage of a mobile phone and its sensors is required, whereas our system works without any user-worn devices. We built a first prototype on top of a Kinect sensor to evaluate the feasibility of our approach and to explore body movements as input means to control home appliances.

2 Interactive Spaces for a Smart Environment

Our system is able to turn any ordinary (physical) space into an interactive, "smart" environment: The main input is provided by a sensor that is able to track the three dimensional position and posture of the human body and limbs in the corresponding environment. Events such as switching on a lamp, controlling the hi-fi system, or starting a program on an embedded system can be activated when selected body parts hit pre-defined *interactive spatial areas* of the environment (cf. Figure 1). We therefore distinguish three different items that define the system behavior: triggers, actions, and mappings.

Figure 1: Overview of our concept. A user is able to control the smart home environment by placing parts of the user´s body into trigger areas to trigger actions such as changing the TV channel.

Trigger areas can be of arbitrary shape, like for instances cubic boxes or more complex shapes, and are used to detect the presence of a user's body parts in order to trigger actions. Each action is a potential response of the smart environment, such as switching on/off a light or controlling the TV. Mappings connect trigger areas, body parts, and corresponding actions, hence the presence of a user's elbow in a dedicated trigger area may result in switching the TV channel. Thus, different user controls are available for the TV, hi-fi system, heating, or window blinds. In a living room it may make sense to create interactive areas in the couch's vicinity for controlling the hi-fi system for example.

We designed and implemented a prototype in C# that is used to evaluate our concept. The functionality of the prototype is threefold. It is used to (1) define new trigger areas, (2) link them to specific actions, and (3) execute the actions. At first it is necessary to define trigger areas with specific parts of the body, which is done by using the Kinect. Secondly, actions need to be created (e.g., increase volume). Last, the actions and trigger areas need to be mapped to each other. These trigger areas are now being observed by our system: as soon as the corresponding part of the body is detected inside the particular area, the corresponding action is executed. Furthermore, we created a prototypical interface to a hi-fi system. Four actions are created: two for changing the volume and two for changing the played track.

3 Evaluation

We conducted a user study to gain insights into the overall feasibility of our concept by analyzing the memorability of trigger areas in 3D spaces (Hypothesis 1: *Users are capable of memorizing spatial positions*). This is one of the main requirements for our concept. Second, we wanted to find out what size fits best for the areas (Hypothesis 2: *The size of the trigger areas is important)*. Thus, two measures are used: task completion time (TCT) for placing the corresponding body part into the intended area and how often areas are hit accidently (error rate – ER). We recruited 18 participants (6 female) aged 22 to 43 years ($M = 26.56$, $SD = 6.30$) and had them define trigger areas, interact with them, and fill out a questionnaire.

Memorability

Analyzing the memorability of the different areas, we measured TCT and error rate. The TCT decreases slightly from the first trial ($M_{Trial\ 1} = 24.33$ s, $SD_{Trial\ 1} = 17.39$ s) to the second trial ($M_{Trial\ 2} = 20.00$ s, $SD_{Trial\ 2} = 13.11$ s). The ER decreases even more from $M_{Trial\ 1} = 2.44$ ($SD_{Trial\ 1} = 2.25$) to $M_{Trial\ 2} = 0.94$ ($SD_{Trial\ 2} = 1.11$). We performed Wilcoxon signed rank tests because of non-normal distribution of the data. The Wilcoxon test shows no significance for the TCT, $Z = -0.308$, $p = .758$, $r = -.051$, however, the ER is statistically significant different, $Z = -2.362$, $p = .018$, $r = -.394$. This result suggests that there is a significant difference between both trials. With regards to these results, we cannot accept Hypothesis 1. The results indicate that users are able to remember spatial, however, some further investigation needs to be done to finally accept Hypothesis 1.

Trigger area size

We analyzed how participants perform using trigger areas with different sizes. In terms of TCT, they perform best with 40 cm edge length areas, followed by 20 cm and 10 cm (cf. Figure 2). In contrast, the performance in terms of ER for these sizes is inverted. Participants performed worst in 40 cm condition, followed by 20 cm and 10 cm (cf. Figure 2).

We performed Friedman-tests to analyze the data because of non-normal distribution. For the TCT, the test shows a statistically significant result, χ^2 (2) = 19.972, $p = .000$. Analyzing the data more deeply, we used Wilcoxon-Signed-Rank tests with a Bonferroni correction applied, resulting in a significance level set at .016. The test shows statistically significant results for 10 cm compared to 20 cm, $Z = -2.548$, $p = .011$, and 40 cm, $Z = -3.724$, $p = .000$. The difference between 20 cm and 40 cm is not statistically significant, $Z = -2.297$, $p = .022$.

Investigating the ER, a Friedman test shows statistically significant difference between the edge lengths 10 cm, 20 cm, and 40 cm, χ^2 (2) = 10.361, $p = .006$. Exploring these differences more deeply, we performed a series Wilcoxon-Signed-Rank tests with a Bonferroni correction setting the significance level at .016. The result shows that the difference between 10 cm and 40 cm is statistically significant, Z = -2,917, p = .004, r = .486. The other differences are not statistically significance. Based on our findings, we can accept Hypothesis 2. The size of the trigger areas is crucial with regards to TCT and ER. The larger areas are the more errors users make. In contrast, the smaller areas are the longer it takes to trigger them.

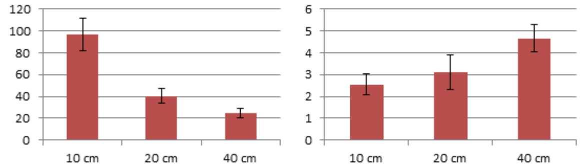

Figure 2: The mean task completion time in seconds (left) and mean error rates (right) separated by the different trigger area edge length. The error bars indicate the standard error.

Questionnaire

In the questionnaire, participants were asked about their likes and dislikes of this concept and the prototype. Most participants liked the concept and stated that it is a "*natural way of interaction*" (p12) and that it is "*fun to interact*" (p4). However, some improvements were suggested, such as that "*no feedback is given*" (p4) and that "*slips are made too easy*" (p3). Furthermore, the participants filled out a System Usability Scale questionnaire. Our system scores 70 points.

4 Conclusion

We presented a prototype that allows users to control appliances by moving certain body parts in pre-defined trigger areas. These trigger areas can be used to control a variety of appliances using body movements. We used a Kinect to track users and conducted a study to evaluate appropriate sizes for trigger areas and memorability of pre-defined areas. Results show that spaces are well remembered and targets are better hit when trigger area size gets bigger, which on the other hand increases error rates.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme ([FP7/2007-2013) under grant agreement no 600851. We thank Peter Vollmer, Niklas Schnelle, and Verena Käfer for their support.

References

Johanson, B., Fox, A. & Winograd, T. (2002). *The Interactive Workspaces Project: Experiences with Ubiquitous Computing Rooms. IEEE Pervasive Computing, Vol. 1 (2). p. 67–74.*

Krumm, J., Harris, S., Meyers, B., Brumitt, B., Hale, M., & Shafer, S. (2000). *Multi-camera multi-person tracking for EasyLiving*. In *Proc. IEEE Int. Workshop on Visual Surveillance. p. 3–10.*

Peters, S., Loftness, V. & Hartkopf. V. (2011). The intuitive control of smart home and office environments. In *Proc. ONWARD '11*. New York, NY, USA: ACM, p. 113-114.

Contact Information

Stefan Schneegass, Bastian Pfleging, Tilman Dingler, Albrecht Schmidt
{firstname.lastname}@vis.uni-stuttgart.de