Preliminary Investigation of the Impact of Visual Feedback on a Camera-based Mouse-replacement System

Wenxin Feng, Ming Chen and
Margrit Betke
Department of Computer Science
Boston University
Boston, USA
http://www.bu.edu/cs/ivc
{wenxinf, mingchen, betke} @ bu.edu

ABSTRACT

This paper reports the design, implementation, and results of a carefully designed experiment that examined the performance of a camera-based mouse-replacement interface that was supported with visual feedback. Four different visual feedback modes were tested during the pointing-task experiment. Quantitative results, based on three metrics, do not show statistically significant difference between these modes. Qualitative feedback from the participants of the experiments, however, shows that user experience is improved by static and animated visual feedback during the pointing task.

Categories and Subject Descriptors

H.1.2 [**Models and principles**]: User/Machine Systems – *Human factors*; H.5.2 [**Models and principles**]: User Interfaces – *Evaluation/methodology*.

General Terms

User Interfaces, Experimentation, Human Factors.

Keywords

Assistive Technology, Mouse Replacement, Pointing-task, Visual Feedback, Camera-based Interface, Augmentative and Alternative Communication.

1. INTRODUCTION

Severe paralysis caused by a traffic accident, a stroke, or a degenerative disease can drastically change a person's life. Individuals with severe motion impairments use "augmentative and alternative communication" (AAC) technology to access the computer and have a means to communicate. Video-based mouse-replacement systems are particularly useful AAC technologies for nonverbal users with quadriplegia. The objective of this paper is to assess whether visual feedback provided by a mouse-substitution system can make the interaction between users and the system more effective.

Camera-based mouse-replacement systems [2,7,12], requiring no more than a computer and a webcam, can provide no-contact, un-calibrated and self-initialized interaction for individuals with severe motion disabilities [6]. The mouse-replacement system evaluated in the experiment described here is the Camera Mouse, an ACC mouse-replacement system that is used by nonverbal individuals with quadriplegia in homes, schools and hospitals worldwide. The Camera Mouse is a system that tracks the user's movements with a video camera and translates them into the movements of the mouse pointer on the screen [2]. The interface interprets a pointer that has not moved for a certain period of time (dwell time) as a left mouse click. Camera Mouse is available at no cost at http://www.cameramouse.org [3] and has recently reached the 1.5-million-download mark.

It is well known that visual feedback is an important component of a human-computer interface, since it can help the user understand the state of the interaction while using the system. Manresa-Yee et al. described how users' satisfaction of a vision-based system is affected by its interface design [10]. Recently, Gizatdinova et al. compared text-entry performance under different selection techniques and keyboard size [6]. This kind of investigation of the usability of video-based interaction systems, however, is rare. Research about camera-based mouse-replacement systems has mainly focused on the improvement of the computer-vision component of the system.

Some camera-based interface systems are based on gaze detection. Unlike the Camera Mouse, which supports tracking of different facial features, gaze-based interaction systems can only track eye movement. Gaze-based interfaces must deal with the issue that a selection mechanism must be found for the eyes, which are naturally used only for perception, not actuation. To avoid the "Midas Touch Problem" [8] that everything the user looks at is "turned to gold," i.e., is selected, gaze-based interfaces typically issue a selection command when the user has looked at the desired target for a given dwell period. The dwell-time-based selection mechanism is similar to that of the Camera Mouse. This motivates us to apply methodologies developed for gaze-based interface research to video-based mouse-replacement systems like the Camera Mouse.

This paper describes a pointing task experiment that examines four types of feedback visualizations. The design of the experiment was motivated by the work of Zhang et al. [14], which analyzed the effects of visual feedback on the performance of a gaze-controlled interface.

2. EXPERIMENTAL METHODOLOGY

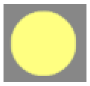
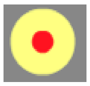
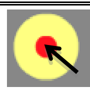
The pointing task experiment examined the performance of the camera-based input system when used with four different visual feedback modes. The feedback mode design was similar to work by Zhang et al. [14]. In our experiment, however, we were

PETRA '13, May 29 - 31 2013, Island of Rhodes, Greece

http://dx.doi.org/10.1145/2504335.2504359

concerned about the different ways of visualizing the focus areas of the mouse pointer target. The visual feedback modes, shown in Table 1, include no feedback, static, and dynamic feedback.

Table 1. Feedback Modes

Name	Target Type	Description
No Feedback		There is no focus area. The yellow target remains unchanged during the dwelling period.
With Focus Area		There is a red focus area. The focus area remains unchanged during the dwelling period.
Shrinking Focus Area		There is a red focus area. The focus area shrinks from the target edge during the dwelling period.
Expanding Focus Area		There is a red focus area. The focus area expands to the target edge during the dwelling period.

2.1 Setup and Participants

A PC with Intel Xeon® CPU at 2.67 GHz processor, 2.93 GB RAM running Windows™ Windows 7 was used in the experiment, along with a 30-inch LCD display at 2560×1600 resolution. The real-time video was captured by Logitech QuickCam® Pro 9000 webcam with 8 megapixel video resolution, which was fixed on the top-middle of the display. We employed the Camera Mouse 2013 [3] software as the video-based mouse-replacement input system. We selected a "small" clicking radius, a 1-second clicking dwell time and a "low" smoothing configuration.

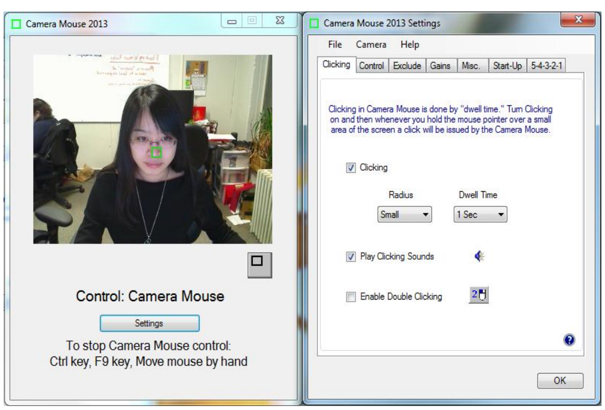


Figure 1. The interface and configuration of Camera Mouse. The green box shows the feature tracked.

Ten able-bodied subjects (4 females and 6 males, with the average age of 25) successfully completed the experiment. All of the subjects had normal or corrected-to-normal eyesight. The subjects had no or little previous experience using Camera Mouse or other camera-based mouse-replacement systems.

2.2 Design and Procedure

The participants sat in front of the display at a 60-80 cm distance. They could adjust the chair to an appropriate height and position. Before the experiment, the experimenter gave a brief introduction of the Camera Mouse and the interaction task. The participants had three minutes to learn the usage of the Camera Mouse and familiarize themselves with the experiment interface and task. There was no formal training phase since the task was simple for the participants to learn. The facial feature that the Camera Mouse tracked in this experiment was the nose, which is a desirable tracking feature for the Camera Mouse [2], as shown in Figure 1. The participants were instructed to sit still during the experiment. Their movements, however, were not limited.

Poulton discovered that range effects exist in within-subject experiments [13]. Therefore, the order of these four modes was counterbalanced using two Latin square patterns in normal and reversed order.[1] In addition, the participants could have a rest between blocks, which can be served as a buffer to reduce fatigue. We could minimize the range effects in this experiment, i.e., skill transfer and fatigue, from one feedback mode to another through these two experimental design components.

The graphical user interface of our experiment is shown in Figure 2. The start button was displayed in the center of the screen, surrounded by targets in eight directions. The red start button was a 100-pixel-diameter disk and the target was a 72-pixel-diameter disk.

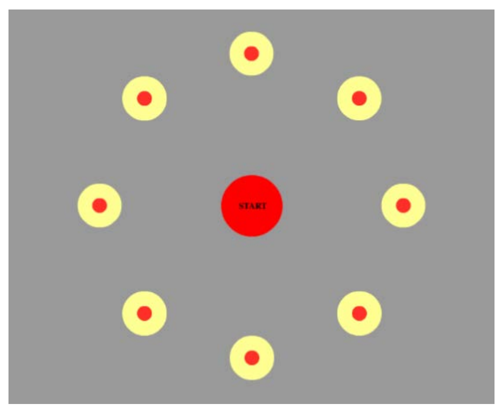


Figure 2. The graphical user interface of the experiment. Only one target will appear in each trial.

The experiment was started by clicking the red start button. Once the start button was clicked, it disappeared and one of the eight target buttons appeared randomly. The participant then needed to move the mouse pointer to the target, using the Camera Mouse, and dwell on the target for 1 second to select the target. When the target was selected, the target disappeared and the start button reappeared.

The participant was instructed to complete a trial as accurately and quickly as possible. If a participant could not click the target successfully within 10 seconds, this trial would not be counted. One block consisted of 20 repeated trials. For each visual

[1] Another non-visual feedback mode was examined in Poulton's experiment as well, so he worked with five feedback modes. This non-visual feedback mode was not included in this paper.

feedback mode, the subject was instructed to complete four blocks.

2.3 Metrics

We took three metrics into consideration when measuring the performance of the camera-based mouse-replacement system with different visual feedback modes.

Average Movement Time (AMT). AMT is the average time that a user successfully finished a trial in one block. This measurement is frequently used in pointing-task experiments.

Entering Target Times (ETT). ETT is the total number of times that the mouse pointer entered a target region in one block. Since the positioning of the pointer is not as smooth with the camera-based input system as with the regular mouse, the pointer may leave and re-enter the target region unintentionally. This measurement is important for evaluation, not only for the Camera Mouse, but also for other mouse-replacement input devices, for example, eye-tracking devices [14].

Average Distance (AD). AD is the average distance between the clicking point and the center of the target in one block. The use of this metric helps examine which visual feedback type may support a user to focus on the target center.

3. EXPERIMENTAL RESULTS

The statistical analysis of the performance of the different visual feedback modes was done by repeated measures ANOVA.

3.1 Average Movement Time Metric

No significant difference in average movement time was found between different feedback modes ($F_{3,36}$=0.27, p=0.8457). The difference between blocks is not significant ($F_{3,36}$=0.43, p=0.7307). However, inspection of Figure 3 reveals that the target with a focus area has a lower total average movement time AMT than the other modes.

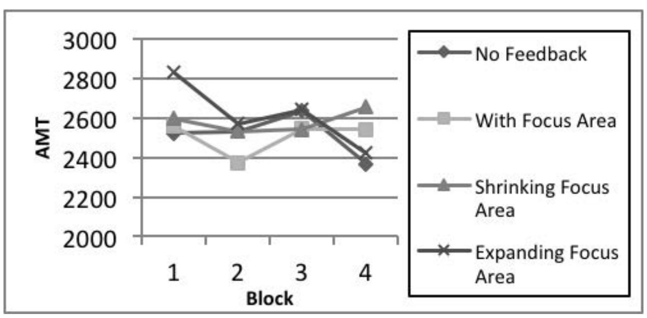


Figure 3. AMT by block under different feedback modes

3.2 Entering Target Times Metric

No significant difference in the number of times that the mouse pointer entered a target region was found between different feedback modes ($F_{3,36}$=0.75, p=0.5308). The difference between blocks is also not significant ($F_{3,36}$=0.61, p=0.6149), as can be seen in Figure 4.

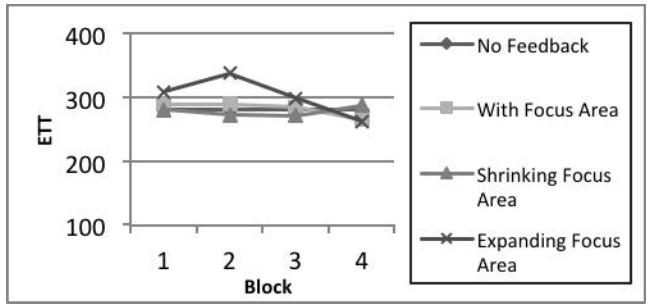


Figure 4. ETT by block under different feedback modes

3.3 Average Distance Metric

No significant difference in the average distance between the clicking point and the center of the target between different feedback modes was found ($F_{3,36}$=0.3, p=0.8221). The difference between blocks is also not significant ($F_{3,36}$=0.34, p=0.7942). Inspection of Figure 5 however reveals that the target without visual feedback has lower total AD than the other modes.

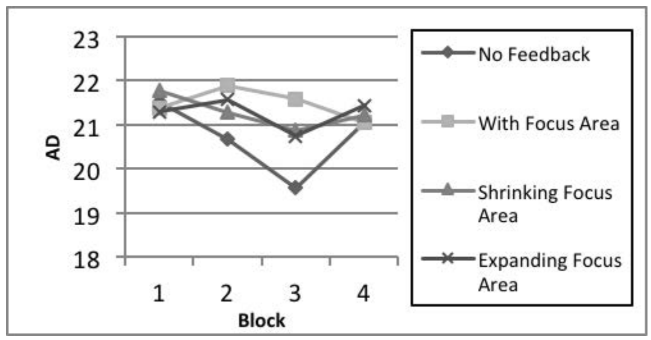


Figure 5. AD by block under different feedback modes

3.4 Subjective Satisfaction Metrics

Nine out of ten participants mentioned that they preferred a target with visual feedback. They believe that their performance is better when there is visual feedback. We found that 50% of the participants preferred the target with a focus area, *"The focus area makes focusing on the target easier."* And 40% of the participants preferred a target with animated visual feedback (shrinking or expanding). *"The animation of the red area lets me know that the cursor is in the target,"* one participant said.

4. DISCUSSIONS AND CONCLUSIONS

The experiment reveals that the specific feedback mode does not have statistically significant effects on the pointing task with the camera-based mouse-replacement system. However, we can learn from Figure 3 and Figure 5 that the static feedback modes have positive effects on the performance of camera-based mouse-replacement system. In addition, from Figure 3, the AMT in dynamic feedback modes is improved by 14% over blocks, which indicates that subjects take time to adapt to dynamic visual feedback.

The pointing performance, which is measured by the ETT, is stable with the camera-based system. Unlike typical gaze-based interaction systems, the Camera Mouse can track the movement of other features of face and body [2]. Thus, it could avoid gaze jitter and other gaze-input problems. From the experimental results we can find out that the average ETT is 1.5 per trial, which is significantly lower than the time measured in a similar experiment conducted with a gaze-based interaction system [14].

Another interesting result is that the subjective user feedback was different from the statistical analysis. Visual feedback during the pointing task could improve user experience even if the targeting performance was not improved.

Our experiment also evaluated the performance of Camera Mouse system in a pointing task. Since we have no formal training phase before the experiment and there is no significant difference in the results between blocks, we can conclude that the Camera Mouse is intuitive and simple for able-bodied people to use.

5. FUTURE WORK

Future work will extend the described experiment to include users with motion impairments. Also, different visual and non-visual feedback modes will be included in further experiments. Subjective research (semi-structured interviews, questionnaires, etc.) will be conducted to investigate the difference between subjective and objective results observed in this work.

We will test other mouse-replacement systems as well, for example, the beta version of the Camera Mouse, called "Camera Mouse Suite" [4], which is available for free download on the internet and includes the Camera Mouse team's most recent work: (1) an optical-flow-based version of the Camera Mouse [1] (2) a version that uses kernels as the tracking mechanism [5], (3) an interface that interprets eyebrow raises for input of scan-based text-entry systems [9] and (4) an interface that provides full pointer control via blink and wink detection [11].

6. ACKNOWLEDGMENTS

We would like to thank the participants of the experiment for their efforts. Funding from the National Science Foundation IIS-0910908 and IIS-1329119 are gratefully acknowledged.

7. REFERENCES

[1] Akram, W., Tiberii, L., and Betke, M. 2006. A customizable camera-based human computer interaction system allowing people with disabilities autonomous hands-free navigation of multiple computing tasks. In *Proceedings of the 9th conference on User interfaces for all* (ERCIM'06). Springer-Verlag, Berlin, Heidelberg, pp. 28-42.

[2] Betke, M., Gips, J., and Fleming, P. 2002. The Camera Mouse: visual tracking of body features to provide computer access for people with severe disabilities. *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* 10(1):1-10, March 2002.

[3] Camera Mouse 2013 website. Camera-based mouse replacement interface for users with severe motion impairments. http://www.cameramouse.org. Accessed March 2013.

[4] Camera Mouse Suite website. http://cameramouse.bu.edu. Accessed March 2013.

[5] Epstein, S., Missimer, E., and Betke, M. 2012. Using kernels for a video-based mouse-replacement interface. *Personal and Ubiquitous Computing*. Online access, pp. 1-14.

[6] Gizatdinova, Y., Špakov, O., and Surakka, S. 2012. Comparison of video-based pointing and selection techniques for hands-free text entry. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (AVI '12). ACM, New York, USA, 132-139.

[7] Gorodnichy, D. and Roth, G. 2004. Nouse 'use your nose as a mouse' perceptual vision technology for hands-free games and interfaces. *Image and Vision Computing,* 22(12):931-942, October 2004.

[8] Jacob, R. J. K. 1990. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '90). ACM, New York, NY, USA, pp. 11-18.

[9] Lombardi, J. and Betke, M. 2002. A camera-based eyebrow tracker for hands-free computer control via a binary switch. In Proceedings of the *7th ERCIM Workshop on User Interfaces for All,* Paris, France, pp.199–200.

[10] Manresa-Yee, C., Ponsa, P., Varona, J., and Perales, F.J. 2010. User experience to improve the usability of a vision-based interface. *Interacting with Computers,* 22(6):594-605, November 2010.

[11] Missimer, M. and Betke, M. 2010. Blink and wink detection for mouse pointer control. In *Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments* (PETRA '10). ACM, New York, NY, USA, Article 23, 8 pages.

[12] Morris, T. and Chauhan, V. 2006. Facial feature tracking for cursor control. *J. Network and Computer Applications.* 29(1):62-80, January 2006.

[13] Poulton, E. C. Range Effects in Experiments on People. In The American Journal of Psychology , 88(1):3-32, March 1975.

[14] Zhang, X., Feng, W. and Zha, H. 2011. Effects of different visual feedback forms on eye cursor's stabilities. In *Proceedings of the 4th international conference on Internationalization, design and global development* (IDGD'11). Springer-Verlag, Berlin, Heidelberg, pp. 273-282.