

# Acceptable Dwell Time Range for Densely Arranged Object Selection Using Video Mirror Interfaces

Kazuyoshi Murata<sup>1(✉)</sup> and Yu Shibuya<sup>2</sup>

<sup>1</sup> Aoyama Gakuin University, Kanagawa, Japan  
kmurata@si.aoyama.ac.jp

<sup>2</sup> Kyoto Institute of Technology, Kyoto, Japan  
shibuya@kit.ac.jp

**Abstract.** We evaluated an acceptable dwell time range to decrease erroneous selections using a video mirror interface. In this interface, users select a target object by moving his/her palm and dwelling on the object. We focused on a situation wherein objects are densely arranged and users must select a target object from these objects. The effects of dwell time, object size, and distance between objects on the object selection task were experimentally evaluated. The results indicated that a dwell time of 0.3 s is the most appropriate to decrease both erroneous selections and unpleasant experiences. The results of this study can contribute to defining an effective basis for dwell time for selection operations in gesture-based interaction systems.

**Keywords:** Dwell time · Object selection · Gesture interaction system · Video mirror interface

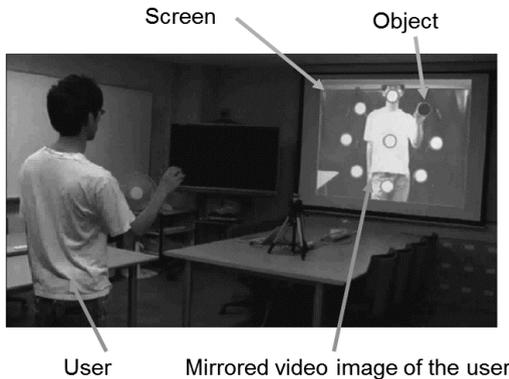
## 1 Introduction

Non-touch interfaces using body gestures are becoming increasingly popular. For example, popular body gesture sensing devices allow users to play games using body gestures. Body gestures are also used as an input method for public displays or large displays. Furthermore, gesture-based interaction has several advantages for interactive public displays [22]. Touch-based interaction is still common for interactive public displays; however, it may be inappropriate for reasons such as hygiene [22].

A video mirror interface is a type of non-touch interface. Figure 1 shows a user selecting an object using a video mirror interface. As shown in Fig. 1, a mirrored video image is displayed on a screen in front of the user, and some objects are superimposed on the video image. Users can select these objects by moving his/her palm over the desired object and dwelling on it. This type of interaction is intuitive and commonly used for selecting actions in gesture-based interaction systems [5, 23].

Objects such as file icons are often densely or closely located in common interfaces. The required dwell time for object selection affects the usability of video mirror interfaces. If the required dwell time is short and objects are densely arranged or are close to the target object, then erroneous selections occur frequently. This is because of the “Midas Touch” problem [7], i.e., users cannot select an object without touching other

objects. An easy way to avoid the “Midas Touch” problem is to increase the required dwell time; however, long dwell time slows down interactions, causing unpleasant experiences for users.



**Fig. 1.** Object selection using a video mirror interface.

Our study evaluates an acceptable dwell time range using a video mirror interface for users in order to decrease erroneous selections. In particular, we focus on a situation wherein objects are densely arranged and users must select a target object from these objects. This study can contribute to the definition of appropriate dwell time for target selection in gesture-based interaction systems such as video mirror interfaces.

## 2 Related Work

Many studies related to defining an appropriate dwell time for target selection have been conducted in various human-computer interaction fields. For example, in eye gaze interaction systems, the dwell time, i.e., the time for looking at objects to be selected, is commonly used as an alternative to mouse clicking. Jacob reported that the minimum dwell time for a simple object selection task was 0.15–0.25 s, while a dwell time over 0.75 s is not beneficial and causes the user to suspect that the system has crashed [7]. However, Stampe reported no difficulty for dwell times as long as 1 s [18]. Penkar et al. conducted experiments to confirm the relation between dwell time and button size [16]. In their case, a 0.2 s dwell time was appropriate for easy and accurate selection of a circle button with a 150-pixel diameter.

The use of a laser pointer as an interaction device for large displays has also been studied. Dwelling is also a popular selection method for laser pointer interaction systems, and several studies have reported that a 1–2 s dwell time is required to perform accurate object selection [14, 15].

Pen-based or touch-based interactions have become increasingly popular with the development of touch screen devices. In such interactions, when an object is touched for a period with a stylus or a user’s finger, a context-aware popup menu or help menu is displayed. Bau et al. applied a dwell time of 0.25 s to their pen-based interaction

system to display its dynamic guidance function [1]. Freeman et al. proposed Shadow-Guides, which is a system for learning multi-touch and whole-hand gestures using an interactive surface that requires a dwell time of 1 s [4].

The dwelling technique for item selection is also used in cascading menu selection. Some menu selection implementations such as Java Swing and several commercial products require a certain dwell time after a mouse cursor enters a parent cascade menu item before displaying its associated menu [3].

There are alternatives to mouse clicking for gesture-based object selection other than dwelling, such as pushing, drawing, grabbing, and enclosing [5]. Vogel et al. used finger air-tapping [21], and Bolt proposed combinations of gestures and user speech, such as “put that there.” [2] However, these alternatives require that users learn how to use the system. Users who pass by a public display can interact with the display inadvertently. In this case, dwelling is more appropriate than the above alternatives because dwelling is more intuitive and provides greater availability without requiring help or guidance. However, there are few clear guidelines for selection operation dwell time with gesture-based interaction systems such as video mirror interfaces. We expect that our research will contribute to defining an appropriate dwell time for such systems.

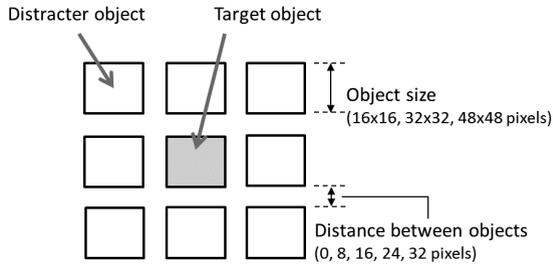
Various interaction systems with a mirrored user image have been studied. In ALIVE [10], users can interact with a virtual 3D environment wherein the user’s mirrored video image is integrated. The Mirror Metaphor Interaction system [6] lets users interact with CG objects and real world objects in a projected video image. Yolcu et al. proposed a virtual mirror system to select clothes in an online shopping system [25]. Taylor et al. proposed a posture training system [19]. Rather than a live video image of the user, several other systems have used computer graphics images such as a silhouette [8, 13, 22] or a virtual character [11, 20]. However, object selection operations are still required for such systems, and our work can contribute to improving the usability of such systems.

### 3 Experiment

The effects of dwell time, object size, and distance between objects in an object selection task were experimentally evaluated. Figure 2 shows the object arrangement in the experiment. Densely arranged objects often appear in a typical window system, e.g., items on a menu bar and file icons in a folder window. In this experiment, the selection of a target object that is surrounded by distracter objects was performed to mimic the abovementioned object selection tasks.

#### 3.1 Dwell Time

Our pilot study [12] showed that dwell time  $> 0.5$  s caused participant fatigue. Therefore, the dwell time conditions in this experiment were set to 0.0, 0.1, 0.2, 0.3, 0.4, and 0.5 s.



**Fig. 2.** Experimental object arrangement

### 3.2 Object Size

A user's palm often passes over a distracter object while selecting a target object. In this case, the dwell time must be longer than the time elapsed when passing over a distracter object. In this experiment, both the target and distracter objects are squares. The object sizes were  $16 \times 16$ ,  $32 \times 32$ , and  $48 \times 48$  pixels<sup>1</sup>. These sizes are in accordance with standard Windows application icon sizes<sup>2</sup>.

### 3.3 Distance Between Objects

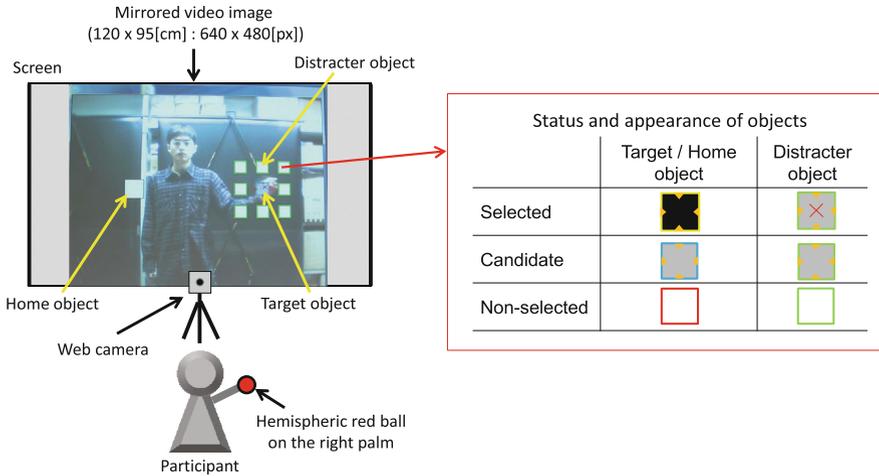
Selecting a target object without touching distracter objects is difficult when the distracter objects are in closely proximity with the target object. In this case, a certain amount of dwell time is needed to avoid erroneous selection. In this experiment, the distances between objects were 0, 8, 16, 24, and 32 pixels. The 0-pixel condition means that the target object adjoins a distracter object.

### 3.4 Experimental Settings

Our experimental system comprised a web camera, PC, projector, and screen. A video image of a participant was captured by the web camera, which was positioned in front of the participant. The captured video image was sent to the PC, which transformed the video to a mirrored video image. Square objects were generated by the PC, which were then superimposed over the mirrored video image. Finally, the processed video image was projected on the screen. The distance between the screen and participant was 5 m. As shown in Fig. 3, nine objects were arranged on the screen to the right of the participant, and one object was arranged to the left of the participant. The object at the center of the nine objects was considered the target object. The other objects were considered distracter objects. The object on the left side of the participant was a home object. The distance between the target object and home object was 320 pixels. All objects were arranged within reach of the participant's right hand.

<sup>1</sup> In this experiment, the real object size on the screen was  $95 \times 95$  mm when the object size was  $48 \times 48$  pixels.

<sup>2</sup> <http://msdn.microsoft.com/en-us/library/dn742485.aspx>.



**Fig. 3.** Experimental settings (object size: 32 pixels; distance between objects: 24 pixels)

When the participants selected the home object with his/her right palm, they were required to twist his/her body and stretch his/her right arm to the left side of his/her body. In this case, it was possible for these postures to cause detection errors with gesture input devices such as Kinect. Therefore, a red hemispheric ball placed on his/her palm was used to detect the palm position using image recognition. The diameter of the ball was 70 mm, and the size of the ball on the screen was 32 pixels. The experimental system recognized that the user's palm was over an object when the red ball overlapped the object.

Figure 3 also shows the status and appearance of the objects. Note that the appearance of the target/home objects and the distracter objects differ.

- Selected object: an object currently selected by the participant.
- Candidate object: an object on which the participant's palm is located within the dwell time. When this status is maintained and the dwell time passes, the status of the object changes to the "Selected object."
- Non-selected object: an object that is not selected. This is the initial status of all objects.

### 3.5 Procedure

A simple target selection task was used in this experiment.

1. The participant stood in front of the screen. Only the home object was displayed on the screen. The participant was asked to select the home object. After the participant selected the home object, the target object and distracter objects were displayed.
2. The participant selected the target object. The participant was asked to perform the task as quickly as possible while avoiding erroneous selections. The participant

continued attempting to select the target object even if a distracter object was selected.

3. The participant selected the home object after selecting the target object.
4. The participant repeated steps 2 and 3 fifteen times because 15 combinations of object size and distance between objects were used in the experiment.

The participants were asked to perform five sets of the above tasks for each dwell time condition. The results of the first set were not used for analysis because this was considered a practice set. The order of the combinations of object size and distance between objects was counterbalanced. The order of dwell time was also counterbalanced. The experiment was video recorded to confirm participant behavior after the experiment.

### 3.6 Design

Ten volunteers (age 22–25) were recruited from the Kyoto Institute of Technology to participate in this experiment. All participants were right hand dominant, who had limited or no experience with gesture input systems such as video mirror interface systems.

A within-subject design was used with the dwell time, object size, and distance between objects factors. The dependent variables were erroneous selection rate and selection time. Furthermore, participants were asked to answer a questionnaire to facilitate a subjective evaluation.

**Erroneous Selection Rate.** When a participant selected a distracter object rather than the target object, the selection was defined as erroneous. The erroneous selection rate was defined as the rate of erroneous selection to the number of times a distracter object was passed over. It is assumed that the erroneous selection rate would be approximately 0 if there was a certain amount of dwell time.

**Selection Time.** The selection time was defined as the total time required to execute steps 2 and 3, i.e., the total time required for selecting the target object and then selecting the home object. The participant must move his/her palm carefully to avoid touching distracter objects if the distance between objects is short. Therefore, it is assumed that the time required for selection increases as the distance between objects reduces.

**Subjective Evaluation.** The participants were asked to answer a question after they completed the task for each dwell time condition. The question was “Did you feel unpleasant during the task?”

## 4 Result and Discussion

Data from this experiment were analyzed in a three-way repeated measures analysis of variance for the following factors: dwell time, object size, and distance between objects. Furthermore, a post hoc test with the Bonferroni procedure was performed after significant primary effects were determined.

### 4.1 Erroneous Selection Rate

The results of the erroneous selection rate are shown in Fig. 4. Passing over a distracter object always resulted in erroneous selection when the dwell time was 0.0 s. Approximately one-half of the times a distracter object was passed over resulted in erroneous selection even when dwell time was 0.1 s.

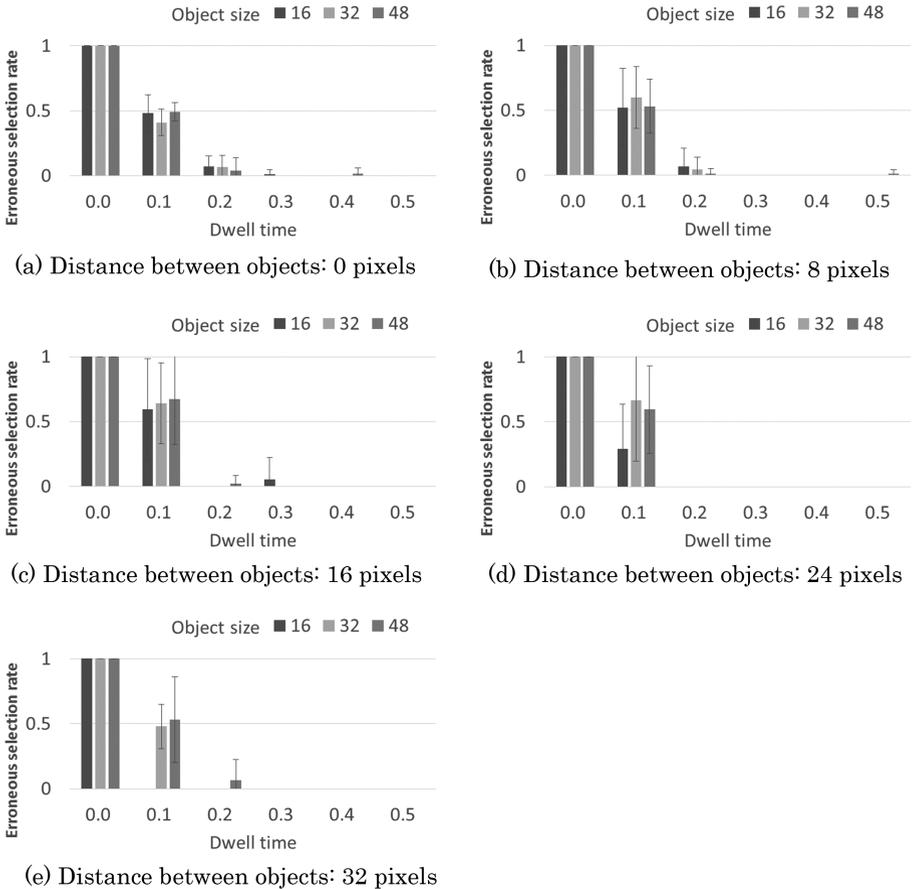


Fig. 4. Erroneous selection rate results

However, the erroneous selection rate decreased precipitously to approximate 0 when the dwell time was greater than 0.2 s. When the dwell time  $\geq 0.3$  s, the erroneous selection rate was 0.04 or less for all combinations of object size and distance between objects, except one. Studies of Fitts' law [9, 17, 24] indicate that a 4 % error rate can be predicted in a simple pointing task if the participant is instructed to perform as quickly and accurately as possible. Therefore, in this experiment, it was assumed that the erroneous selection was acceptable when it was less than 0.04. Consequently, these results

indicate that a dwell time of at least 0.3 s is necessary for selecting a target object without erroneous selection.

Furthermore, the highest erroneous selection rate was 0.07 when the dwell time was 0.2 s. Therefore, it seems that a dwell time of 0.2 s is acceptable if the user can tolerate a few errors, e.g., in a situation wherein the user can perform an undo action easily.

### 4.2 Selection Time

The results for selection time are shown in Fig. 5. Significant effects on the dwell time ( $F(5, 45) = 22.395, p < .01$ ) and object size ( $F(2, 18) = 192.670, p < .01$ ) were observed. The time for selection increased gradually with dwell time. A post hoc test revealed that selection times for dwell times of 0.4 and 0.5 s were longer than those of 0.0, 0.1, and 0.2 s dwell time. Furthermore, the selection times for 0.5 s dwell time were longer than that for 0.3 s dwell time. A certain amount of dwell time was required to decrease

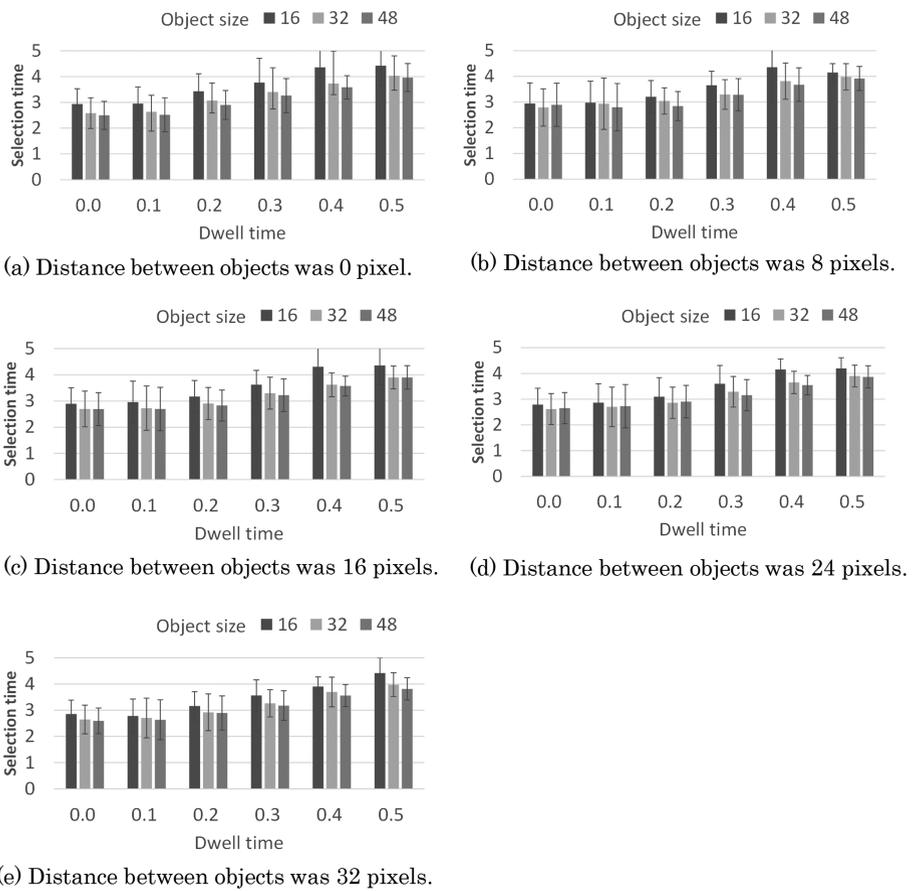


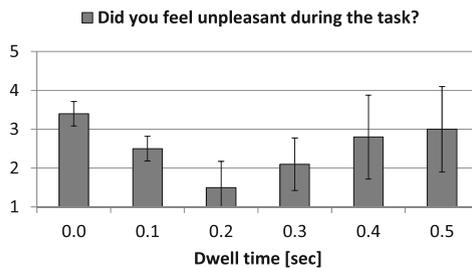
Fig. 5. Selection time results

erroneous selection. However, overly long dwell time, e.g., 0.4 or 0.5 s, increased selection time. From the video of the experiment, it was observed that participants moved his/her palms from the target object region before the dwell time passed, i.e., participants could not wait until the dwell time expired, when the dwell time was 0.4 or 0.5 s. This was also one of the reasons why the selection time increased.

The distance between objects was also observed to have significant effect. However, a post hoc test revealed that there was no difference among each object distance condition. Therefore, it appears that the effect of object distance was small in this experiment.

### 4.3 Questionnaire

Figure 6 shows the results of the questionnaire. Participants felt unpleasant if dwell time was too short or too long. When the dwell time  $\leq 0.1$  s, frequent erroneous selection caused unpleasant experiences. Similarly, unpleasant experiences increased with the dwell time because maintaining the palm in the air proved to be troublesome for the participants. Unpleasant experiences were felt least often with a dwell time of 0.2 s. However, there was no significant effect on dwell time.



**Fig. 6.** Questionnaire results (1–5: “No,” “a little unpleasant,” “unpleasant,” “strongly unpleasant,” and “very strongly unpleasant”)

### 4.4 Recommendation

Our experimental results indicate that dwell time should be 0.3–0.5 s to decrease both erroneous selections and unpleasant experiences. Erroneous selection does not occur when dwell time is within this range even if objects adjoin each other. However, when the dwell time  $\geq 0.4$  s, the selection time increases. This can increase the total time for a selection operation if repeated selection operations are required, e.g., selecting files in deep hierarchical folders. Therefore, under our experimental conditions, a dwell time of 0.3 s was most appropriate. A dwell time of 0.2 s was acceptable if the user could tolerate a few errors, e.g., if the user could easily perform an undo operation.

It is possible that the acceptable dwell time range is affected by the given application or environment. Nonetheless, our experimental results can be considered as an effective basis for finding an appropriate dwell time for selection operations in gesture-based interaction systems.

## 5 Conclusion

Dwelling is commonly used to select an object in gesture-based interaction systems. In this study, an experiment was conducted to evaluate an appropriate dwell time for a selection gesture using a video mirror interface. The experimental results suggested that a dwell time of 0.3 s was the most appropriate to decrease both erroneous selections and unpleasant experiences. Furthermore, a dwell time of 0.2 s was acceptable if users could tolerate a few errors, e.g., if the user could easily perform an undo operation. We expect that the results of this study will contribute to defining an effective basis for dwell time for selection operations in gesture-based interaction systems.

## References

1. Bau, O., Mackay, W.E.: OctoPocus: a dynamic guide for learning gesture-based command sets. In: *UIST 2008 Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology*, pp. 37–46 (2008)
2. Bolt, R.A.: Put-That-There: voice and gesture at the graphics interface. In: *SIGGRAPH 1980 Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 262–270 (1980)
3. Cockburn, A., Gin, A.: Faster cascading menu selections with enlarged activation areas. In: *GI 2006 Proceedings of Graphics Interface 2006*, pp. 65–71 (2006)
4. Freeman, D., Benko, H., Morris, M.R., Wigdor, D.: ShadowGuides: visualizations for in-situ learning of multi-touch and whole-hand gestures. In: *ITS 2009 Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, pp. 165–172 (2009)
5. Hespanhol, L., Tomitsch, M., Grace, K., Collins, A., Kay, J.: Investigating intuitiveness and effectiveness of gestures for free spatial interaction with large displays. In: *PerDis 2012 Proceedings of the 2012 International Symposium on Pervasive Displays*, Article No. 6 (2012)
6. Hosoya, E., Kitabata, M., Sato, H., Harada, I., Nojima, H., Morisawa, F., Mutoh, S., Onozawa, A.: A mirror metaphor interaction system: touching remote real objects in an augmented reality environment. In: *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, p. 350 (2003)
7. Jacob, R.J.K.: What you look at is what you get: eye movement-based interaction techniques. In: *CHI 1990 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 11–18 (1990)
8. Krueger, M.W., Gionfriddo, T., Hinrichsen, K.: VIDEOPLACE—an artificial reality. In: *CHI 1985 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 35–40 (1985)
9. MacKenzie, I.S.: Fitts' law as a research and design tool in human-computer interaction. *Hum. Comput. Interact.* **7**(1), 91–139 (1992)
10. Maes, P., Darrell, T., Blumberg, B., Pentland, A.: The ALIVE system: wireless, full-body interaction with autonomous agents. *Multimedia Syst. Spec. Issue Multimedia Multisensory Virtual Worlds* **5**(2), 105–112 (1997)
11. Marquardt, Z., Beira, J., Em, N., Paiva, I., Kox, S.: Super mirror: a kinect interface for ballet dancers. In: *CHI 2012 Extended Abstracts on Human Factors in Computing Systems*, pp. 1619–1624 (2012)

12. Murata, K., Hattori, M., Shibuya, Yu.: Effect of Unresponsive Time for User's Touch Action of Selecting an Icon on the Video Mirror Interface. In: Kurosu, M. (ed.) HCII/HCI 2013, Part IV. LNCS, vol. 8007, pp. 462–468. Springer, Heidelberg (2013)
13. Müller, J., Walter, R., Bailly, G., Nischt, M., Alt, F.: Looking glass: a field study on noticing interactivity of a shop wind. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 297–306 (2012)
14. Myers, B.A., Bhatnagar, R., Nichols, J., Peck, C.H., Kong, D., Miller, R., Long, A. C.: Interacting at a distance: measuring the performance of laser pointers and other devices. CHI 2002 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 33–40 (2002)
15. Olsen Jr., D.R., Nielsen, T.: Laser pointer interaction. In: CHI 2001 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 17–22 (2001)
16. Penkar, A.M., Lutteroth, C., Weber, G.: Designing for the eye—design parameters for dwell in gaze interaction. In: OzCHI 2012 Proceedings of the 24th Australian Computer-Human Interaction Conference, pp. 479–488 (2012)
17. Soukoreff, R.W., Mackenzie, I.S.: Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in HCI. *Int. J. Hum Comput Stud.* **61**, 751–789 (2004)
18. Stampe, D.M., Reingold, E.M.: Selection by looking: a novel computer interface and its application to psychological research. *Stud. Vis. Inf. Proc.* **6**, 467–478 (1995)
19. Taylor, B., Birk, M., Mandryk, R. L., Ivkovic, Z.: Posture training with real-time visual feedback. In: CHI 2013 Extended Abstracts on Human Factors in Computing Systems, pp. 3135–3138 (2013)
20. Vera, L., Gimeno, J., Coma, I., Fernández, M.: Augmented mirror: interactive augmented reality system based on kinect. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011, Part IV. LNCS, vol. 6949, pp. 483–486. Springer, Heidelberg (2011)
21. Vogel, D., Balakrishnan, R.: Distant freehand pointing and clicking on very large, high resolution displays. In: UIST 2005 Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology, pp. 33–42 (2005)
22. Walter, R., Bailly, G., Müller, J.: StrikeAPose: revealing mid-air gestures on public displays. In: CHI 2013 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 841–850 (2013)
23. Walter, R., Bailly, G., Valkanova, N., Müller, J.: Cuenesics: using mid-air gestures to select items on interactive public displays. In: MobileHCI 2014 Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services, pp. 299–308 (2014)
24. Wobbrock, J.O., Cutrell, E., Harada, S., MacKenzie, I.S.: An error model for pointing based on fitts' law. In: CHI 2008 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1613–1622 (2008)
25. Yolcu, G., Kazan, S., Oz, C.: Real time virtual mirror using kinect. *Balkan J. Electr. Comput. Eng.* **2**(2), 75–78 (2014)