International Journal of Innovative Computing, Information and Control
Volume **9**, Number **12**, December **2013**
ICIC International ©2013 ISSN 1349-4198
pp. **4647–4663**

3D OBJECT SELECTION WITH MULTIMODAL FEEDBACK IN MOBILE AUTO-STEREOSCOPIC DISPLAY

Euijai Ahn, Hyunseok Yang and Gerard J. Kim*
College of Information and Communications
Korea University
Anam-dong 5-ga, Seongbuk-gu, Seoul 136-713, Korea
{ saintpio; baroeye }@korea.ac.kr; *Corresponding author: gjkim@korea.ac.kr

Received August 2012; revised January 2013

Abstract. *Interacting in a relatively small mobile/hand-held auto-stereoscopic display volume (3D "phone" space) can be difficult because of the lack of accurate tracking of an interaction proxy, and having to maintain a fixed viewpoint and adapt to a different level of depth perception sensitivity. In this work, we first introduce an articulated mechanical stylus with joint sensors for 3D tracking the interaction point in the phone space. We also investigate a way to assist the user in selecting an object in the phone space through supplementary multimodal feedback, such as sound and tactility. We have carried out experiments in two conditions, stationary and moving, comparing the effects of various combinations of multimodal feedback to object selection performance. We have found that multimodal feedback was generally significantly helpful for auto-stereoscopic 3D object selection, and particularly more so when the user is moving, considering the added difficulty.*

Keywords: Auto-stereoscopy, Depth perception, Mobile interaction, Multimodal interaction, Selection, 3D tracking

1. **Introduction.** With the continuing technological innovation and recent keen public interests, stereoscopic displays are becoming more commonplace these days. They are being adopted for TVs, desktop displays and finally smart phones and hand-held devices [1-4], but in most cases, still used for viewing only. An exception is when used for special purpose virtual reality (VR) based interactive applications. In fact, 3D interaction techniques for VR (with relatively large-sized stereoscopic display) have been studied considerably [5]. However, not much attention has been paid to the problem of interacting in a relatively "small-sized" hand-held and mobile stereoscopic display such as that of a smart phone.

One obvious difficulty is the lack of an accurate and robust solution to 3D tracking within the small "phone" space (e.g., small rectangular volume right above the phone display, also see Figure 2). Another possible source of complications is the fact that small mobile 3D displays are invariably auto-stereoscopic (e.g., parallax barrier type), limiting the user to fix and maintain one's view point to feel the 3D effect. This becomes even more difficult when the user is moving, a frequent situation when using a smart phone. This is expected to make the task more difficult than that with large fixed displays. In addition, it is plausible to expect some differences in workings of the human's depth perception in the significantly "small" and "moving" phone space compared with the nominally sized space (e.g., human scale).

Our paper thus starts with a proposal for a practical solution to 3D tracking for a mobile smart phone, using an articulated stylus with joint sensors (see Figure 1). Then we also propose to assist the user in selecting an object in the small moving phone space through supplementary multimodal feedback, such as sound and tactility to overcome the

FIGURE 1. 3D interaction on a stereoscopic phone using an articulated stylus with joint sensors for 3D position tracking. The stylus can be folded for easy stow away (left).

aforementioned projected difficulties. We have carried out an experiment in two conditions, stationary and moving, comparing the effects of various combinations of multimodal feedback to object selection performance.

Our paper is organized as follows. First we provide a review of previous and related research. Then we describe the 3D auto-stereoscopic smart phone and newly proposed tracking system, used for the following experiment. Section 4 and 5 give details of the experiment and results. Finally, we discuss and summarize the findings from our experiments to conclude the paper.

2. Related Work.

3D interaction techniques have been studied much in depth mainly in the virtual reality community. An excellent review of the various techniques and their taxonomy are given in [5]. However, subtle difference in their performance or usability according to different types of 3D displays (e.g., auto-stereoscopic, active or passive type, head mounted display) has not been looked at much, especially for small displays [6]. One noteworthy work by [7] studied interaction for small hand-held stereoscopic (passive chromatic anaglyph) display. In their work, the interaction was indirect or gesture based realized by tracking the user's fingertip on the other side of the display using the back facing camera. To our knowledge, there has not been a research study for directly interacting with stereoscopically rendered object in 3D. This is partly due to the difficult problem of accurate tracking in the "phone" space in a self-contained way (i.e., without any third party sensor). The most prevalent approach is to use the phone camera; however, due to its limited field of view, it is not feasible for the tracking volume to cover the entire "phone" space, especially near the display surface.

The phenomenon of altered depth perception with the use of stereoscopic display also has been reported in [8,9]. For example, humans tend to underestimate depth when head mounted display (HMD) is used [8]. Not much is known about the dynamics of depth perception for auto-stereoscopic displays that use parallax barriers or lenticular sheets, not to mention for small-sized ones. Yang and Kim compensated for the depth underestimation in HMDs by manipulating its geometric field of view and providing additional multimodal feedback [10]. Similarly, multimodal feedback has been regarded as one way to improve 3D task performance (which must be related to depth perception) [11-16]. For example, Mereu and Kazman [14] used multimodal (visual and aural) feedback to help users perceive depth more accurately and carry out 3D spatial tasks. However, to date, the results are not consistent in terms of which modality combination is most helpful due to differences in the task and experimental conditions.

Several researchers have recently investigated interaction performance (e.g., object selection, navigation, reading, text entry) during motion (e.g., walking and running). Generally and obviously, most works have reported degraded task performance (e.g., task completion time and error) [22-25], increased mental load and reverse effects to motion by the task at hand (e.g., slowed walking speed). Consequently, researchers have proposed various interaction techniques to alleviate this problem, e.g., enlarging the button size and providing additional audio feedback [23] and easy-to-use navigation interface such as the "touch-n-go" [24]. While these techniques improved task performance, they were not more particularly so for the moving condition. That is, it was concluded that no relationship between the mobility condition and interaction technique was found.

3. Experimental Platform.

3.1. **3D tracking: articulated stylus.** In order to interact directly with 3D objects, 3D[1] tracking is required. Our proposal is to use an articulated stylus with joint sensors as shown in Figure 1. We believe such a device can be designed almost as compact as the conventional stylus with miniaturized yet highly accurate joint encoding sensors. Such a design consideration is necessary to keep the mobile phone light and "handy" to use. The sensors and feedback devices would be directly connected into the smart phone for end point coordinate computation and feedback control. Note that the articulated stylus is not a haptic device (i.e., no force feedback actuators), but merely a tracking one. When not in use, it can be simply folded and stowed away.

Our actual "lab" implementation was bigger in size than the envisioned version, with four degrees of freedom articulation, and using analog potentiometers at the joints (rather than more accurate and small-sized high resolution digital encoders). Additional external modules were needed for digital conversion and interfacing into the smart phone using Bluetooth (all of which can become self-included in the smart phone in the future). Lego pieces were used for the stylus links (see Figure 2). A small vibrator and button were attached for the interaction purpose.

As for the aforementioned additional module, the Arduino board (with the Atmega328 MCU and an on-board 10 bit A/D converter) [17] was used for converting the potentiometer joint angles into digital values. Based on a standard forward kinematics formulation (which we omit the details in this paper) [18], the positional coordinate of the stylus tip is easily computed. For stable output of the tip position, basic low pass filtering was applied. The orientation (of the tip) was not computed or used in this work. All the computations (for now) were carried out on the MCU and transmitted to the smart phone at a rate of approximately 35Hz. The vibrator motor was controlled by the pulse width modulation signal output from the same board. We re-emphasize that, if professionally built with the state-of-the-art components, the stylus can be as compact and accurate as originally proposed.

3.2. **Tracking accuracy and calibration.** To measure the accuracy of our device, we built small 3D structures with Lego blocks (see Figure 3) and compared the computed coordinates of the stylus tip with the ground truth values of various points in the structures. Figure 4 illustrates the accuracy of the articulated stylus in the x-y plane (ground truth: red circles, measured and computed: blue diamonds). In all three directions, the errors were on the average within about 2mm. As our focus was more on deriving an effective 3D

---

[1] Or more generally, 6D tracking for position and rotation in all three principal directions are needed. In this paper, we only consider 3D positional tracking.

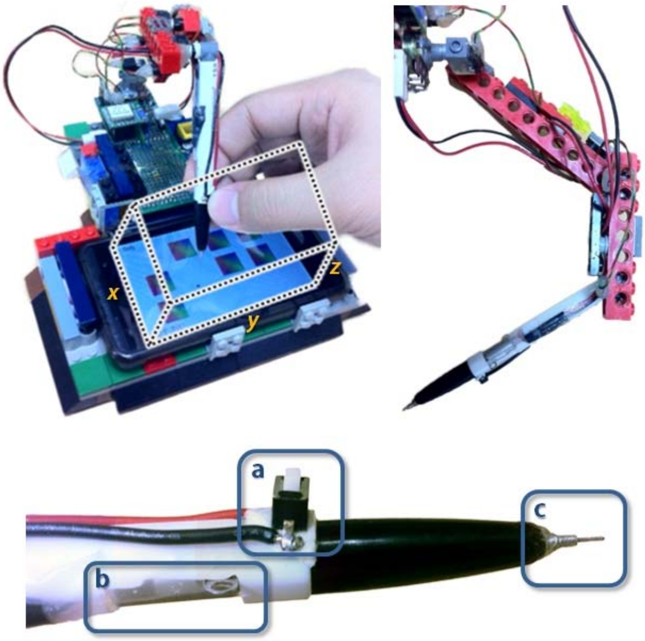


FIGURE 2. Actual "lab" implementation of the articulated stylus using lego pieces, potentiometers and associated circuitries: (a) button, (b) vibrator, (c) stylus tip

FIGURE 3. One of the 3D structures used for accuracy measurement

object selection technique, no further significant effort was made to improve the accuracy. However, due to the personal variations in depth perception, we asked each user, during the experiment, to designate several 3D landmark points (similarly to calibrating a touch screen) to calibrate them against the corresponding ones in the virtual space (see Figure 5).

3.3. **Auto-stereoscopic smart phone.** The auto-stereoscopic phone used in our experiment was an LG Optimus 3D [4] with a 4.3 inch 3D auto-stereoscopic (parallax barrier) LCD display (480 × 800 pixels, 16M colors). Parallax barrier technology refers to creating the 3D effect by using a barrier (layer of material with a series of precision slits) placed on the image source (e.g., LCD) such that each eye sees the respective right or left image (without the need to wear special glasses, see Figure 6) [19]. A disadvantage

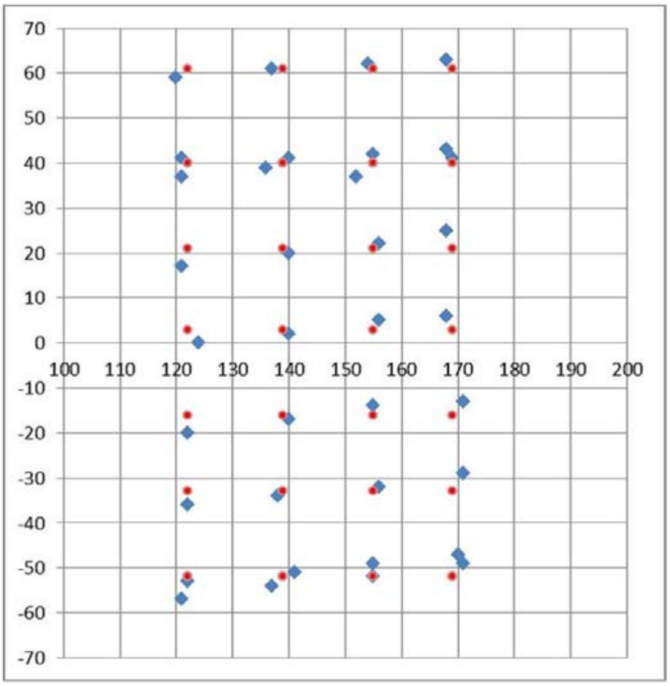


FIGURE 4. Accuracy in the x-y plane (units in mm). Ground truth values are marked in red circles, and the measured and computed in blue diamonds.

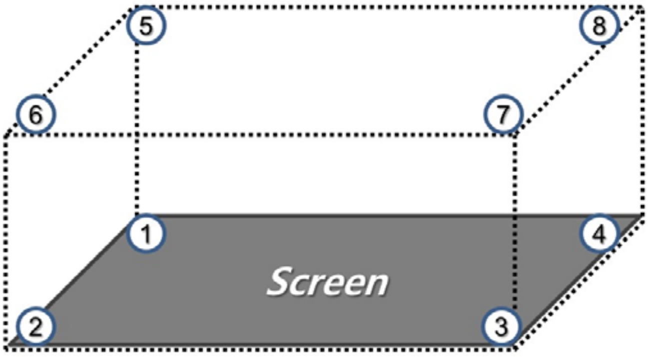


FIGURE 5. The "phone" space and the eight designated points for calibration to the virtual volume

of the technology is that the user must be positioned in a well-defined spot to experience the 3D effect. The exact spot depends on several factors including, e.g., the inter-ocular distance of the user, but for this phone model, it was approximately 30cm perpendicularly above from the center of the screen. Another disadvantage is that the effective horizontal pixel count viewable for each eye is reduced by one half due to the simultaneous alternate rendering of the left and right imagery pixels. The typical operational phone space was assumed to be shaped as a rectangular volume with the physical dimensions of 56mm × 93mm × 40mm, as viewed from the sweet spot. Note that above figures are nominal values only; both the proper viewing position and perceived size of the phone space would be slightly different for different users.

3.4. **Overall system architecture.** Figure 7 illustrates the overall architecture combining the tracking device with the smart phone. The computed stylus tip coordinates are relayed to the smart phone which visually renders the virtual world in stereo, and other modal feedback for interaction (e.g., aural and tactile).

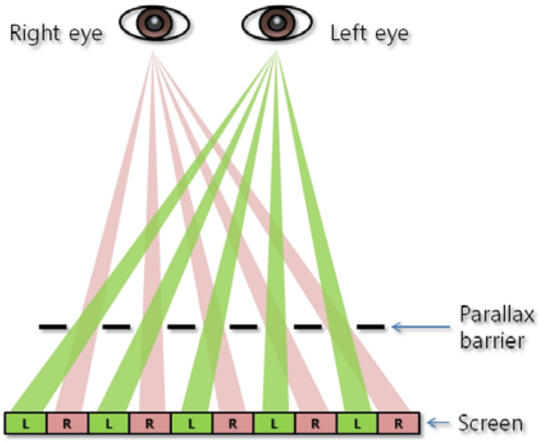


FIGURE 6. The parallax barrier technology used in the LG autostereoscopic phone. Through the slits in the display surface barrier, each eye sees the respective left and right image as from particular spots.

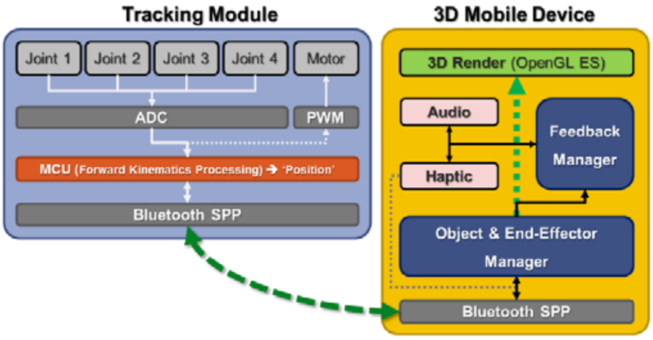


FIGURE 7. Overall system architecture: tracking module on the left and smart phone on the right.

4. Experiment I: Object Selection Technique while Stationary.

4.1. Experiment design.

With the experimental platform in place, rudimentary object selection in the 3D phone space has become possible. Nevertheless due to the factors mentioned previously (e.g., fixed spot viewing, user motion, reduced resolution, unknown dynamics of the depth perception in small sized volume), we expect some difficulties in fluid interaction. As such, we propose to take advantage of supplementary modal feedback, namely, aural and tactile, and carry out an experiment to explore the feedback design space. In Experiment I, we consider four different feedback conditions as an aid to make object selection in a stationary (seated) position. The feedback conditions are (1) visual only (V, the reference), (2) visual and aural (VA), (3) visual and tactile (VT), and (4) visual, aural and tactile (VAT).

Since the correct depth perception is the matter of importance in this work, as for the experimental task, we presented two objects (cubes) of similar depth and asked the user to disambiguate the depth between them. More experimental details follow in the subsequent subsections. In summary, Experiment I had one factor, the type of multimodal feedback, with four levels (1 × 4 within subject repeated measure) and the task performance and usability were measured as major dependent variables. Our main hypothesis was that higher degree of multimodal feedback would generally improve the object selection or depth perception performance.

4.2. **Multimodal feedback.** The visual feedback merely consisted of rendering of the cubes. To remove any external bias, we rendered the cubes with orthogonal projection and minimal lighting effects (Figure 8). It is well known that perspective projection alone is a very strong psychological depth cue. We eliminated all depth cues except for the binocular disparity for stereoscopic rendering itself. All tested objects were rendered with negative parallax for the "floating above the screen" effect. Note that objects seemingly further inside the screen (positive parallax) are not directly selectable.

As for the aural and tactile feedback, they were generated when the 3D cursor (drawn at the tip of the articulated stylus) came into some proximal distance with a nearby object (Figure 9).

A sine tone was generated as aural feedback and its frequency was determined by the depth (or discrete depth level) of the object to whom the 3D cursor was most proximal (i.e., interaction object). The higher the object was (i.e., distant from the screen, closer to the user), likewise the tone frequency. A reasonable audible frequency range was mapped to the depth range of the stereoscopic display (see Table 1).

Similarly, the amplitude/frequency of the vibratory tactile feedback was inversely proportional to the depth (from the user) of the object as well. The vibration motor, when controlled by the PWM signal, varies the vibration frequency between (200∼350Hz) and its amplitude at the same time. It has been reported that 250Hz is the most discernible vibration frequency for the human skin [20]. In this experiment, both feedbacks lasted

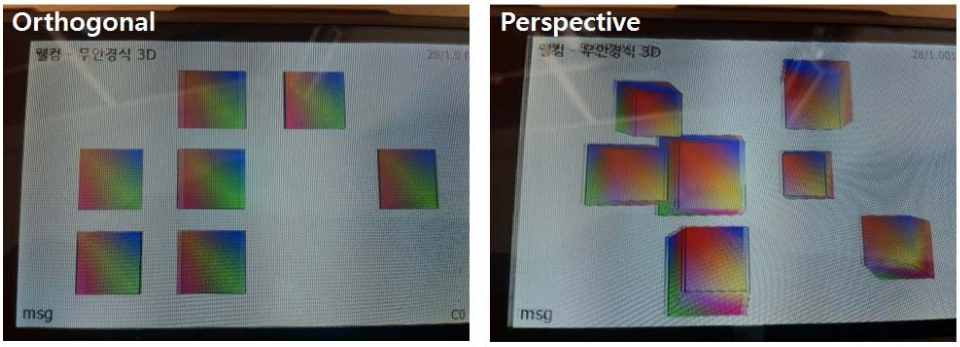


FIGURE 8. Orthogonal vs. perspective rendering. When the depth is similar, it is difficult to tell the difference solely from the appearance (thus stereoscopy is needed), whereas with perspective rendering it is somewhat possible to judge the depth more correctly even without stereoscopy.

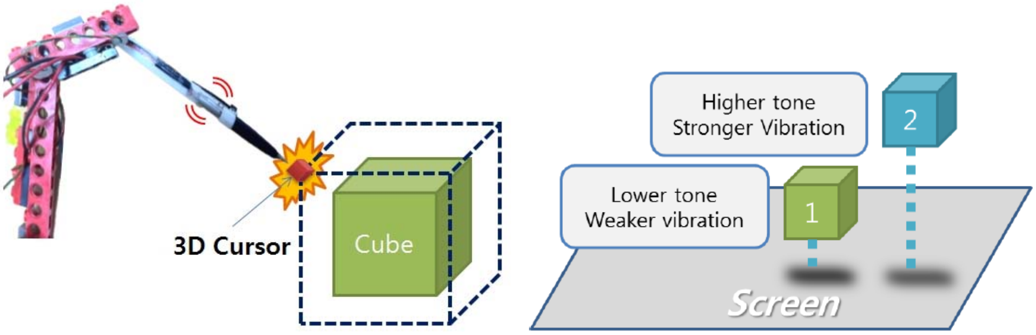


FIGURE 9. Indication of proximity/depth of the 3D cursor to an object. The indication is expressed aurally and/or tactically based on their depth relative to screen.

Table 1. Stimulation parameter values used for aural and vibratory tactile feedback

Depth from user (cm)	*Depth from screen (cm)*	*Tone Frequency (Aural Feedback)*	*Vibration Duty Rate (Tactile Feedback)*
29.0	1.0	730Hz	31%
28.5	1.5	1360Hz	43%
28.0	2.0	1990Hz	54%
27.5	2.5	2610Hz	66%
27.0	3.0	3240Hz	77%
26.5	3.5	3870Hz	89%
26.0	4.0	4500Hz	100%

for one second when generated. Table 1 shows the stimulation parameters used for the respective modal feedback.

4.3. **Experimental task.** The experimental task involved the user to determine the relative depth between two cubes. The subject was to carry out a series of these depth determination tasks as fast and correctly as possible. Two cubes with different (randomly chosen) depth values appeared in the 3D phone space (but with an equal planar distance), and the user was to choose the deeper object using the articulated stylus under different treatment or feedback conditions (see Figure 10).

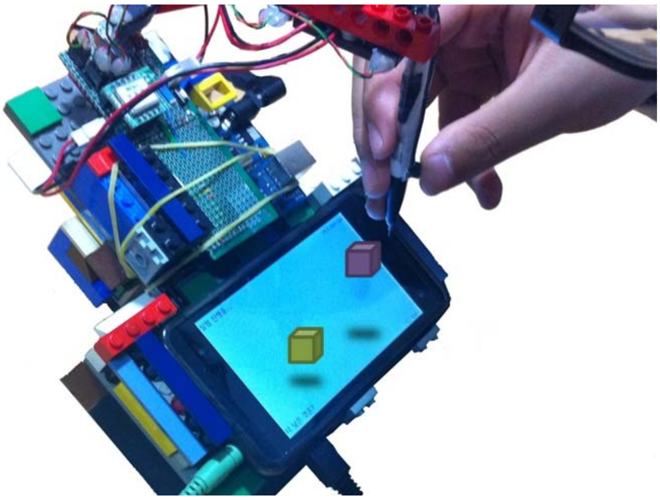


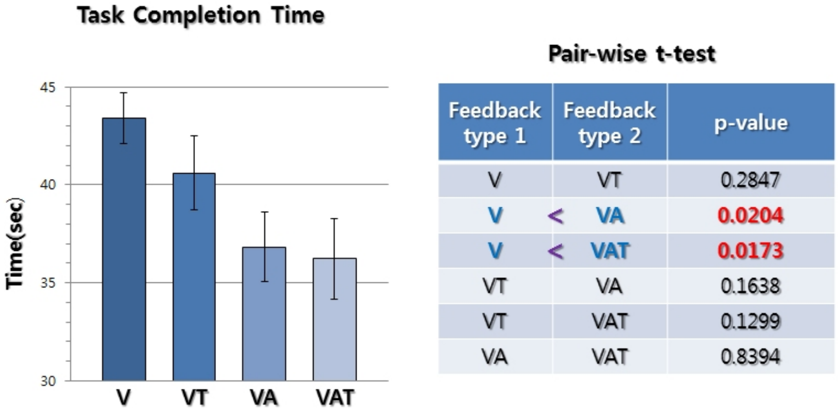
Figure 10. Experimental task: depth determination

4.4. **Experimental procedure.** Twenty one paid subjects (15 men and 6 women) participated in the experiment with the mean age of 24.5. After collecting one's basic background information, the subject was briefed about the purpose of the experiment and instructions for the experimental task. A short training (15 to 20 minutes) was given for the subject to get familiarized to the experimental process and use the stylus.

The subject tried out each treatment combination in a balanced order. For each treatment, the depth test was conducted five times. The task was carried out in a seated position. The task completion time and correctness data were captured and after all the treatments were tried, a general usability questionnaire was filled out (answered in 7 scale Likert scale).

4.5. **Results.**

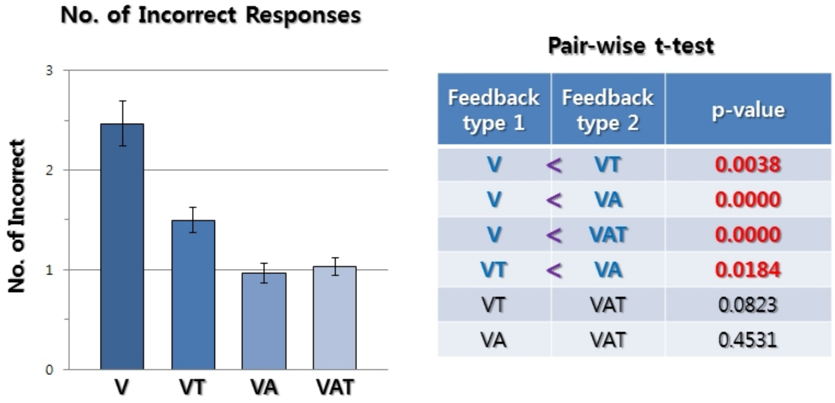
4.5.1. *Task performance.* ANOVA has reaffirmed the effect of the multimodal feedback. VAT exhibited the fastest task performance but with no statistical difference from VA. In addition, VT was neither differentiated from V. Thus, in our experiment, only the aural feedback was meaningfully effective (Figure 11).



Pair-wise t-test

Feedback type 1	Feedback type 2	p-value
V	VT	0.2847
V <	VA	0.0204
V <	VAT	0.0173
VT	VA	0.1638
VT	VAT	0.1299
VA	VAT	0.8394

FIGURE 11. Task completion times for the four feedback conditions (stationary)



Pair-wise t-test

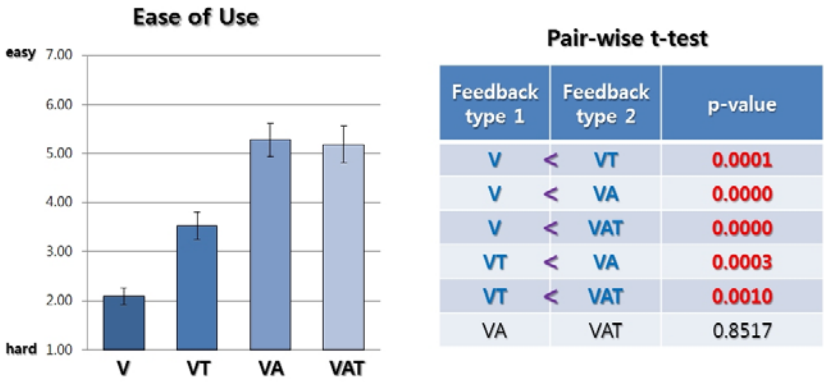
Feedback type 1	Feedback type 2	p-value
V <	VT	0.0038
V <	VA	0.0000
V <	VAT	0.0000
VT <	VA	0.0184
VT	VAT	0.0823
VA	VAT	0.4531

FIGURE 12. Number of incorrect responses for the four feedback conditions (stationary)

Figure 12 shows the number of incorrect answers among the four feedback conditions. In this case, while tactile feedback was effective, it was not as effective as that by the aural (VT $>$ V, VA $>$ VT, VAT $>$ V, VAT $>$ VT).

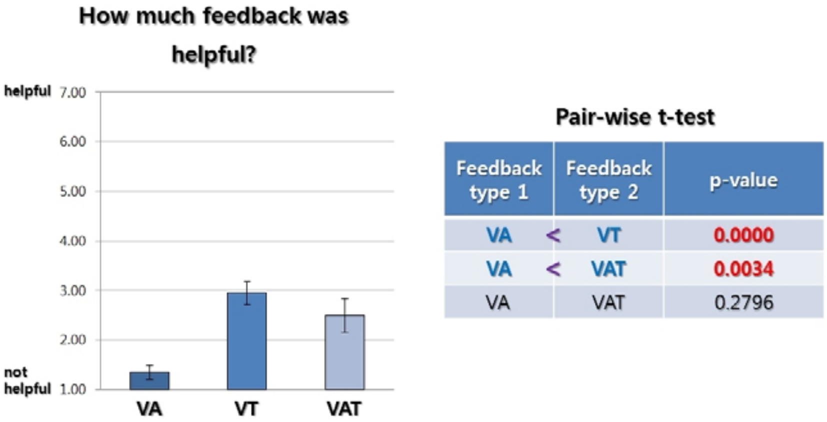
4.5.2. *Usability.* The usability questionnaire asked the subject to comparatively rate the four selection (or feedback) techniques in terms of ease of use, degree to which feedback was helpful in recognizing the depth, ease of learning, interaction naturalness and the level of fatigue. Figures 13-17 illustrate the results.

The usability results are quite consistent with that of the quantitative performance results. For instance, with multimodal feedback, the user felt the task was generally easier, and the easiest for VA and VAT, which were statistically not different, again showing the reduced role of the tactile feedback (Figure 13). Users also responded that the aural feedback was the most helpful, and less so when only tactile feedback was present or mixed (Figure 14). Figure 15 shows, as we have hypothesized, that it was



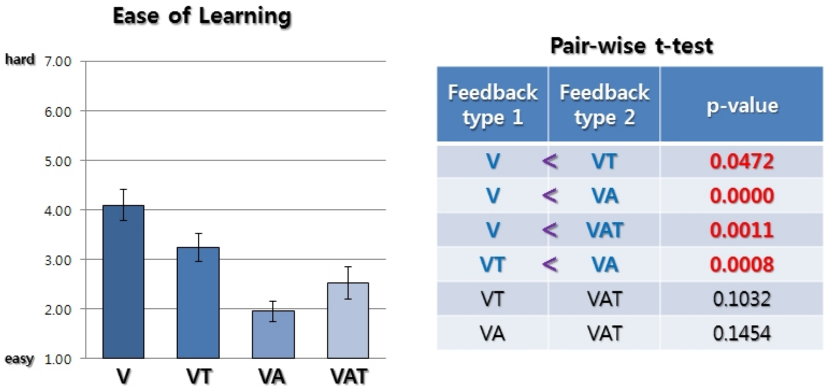
Feedback type 1	Feedback type 2	p-value
V <	VT	0.0001
V <	VA	0.0000
V <	VAT	0.0000
VT <	VA	0.0003
VT <	VAT	0.0010
VA	VAT	0.8517

FIGURE 13. Usability result (ease of use) among the four feedback conditions (stationary)



Feedback type 1	Feedback type 2	p-value
VA <	VT	0.0000
VA <	VAT	0.0034
VA	VAT	0.2796

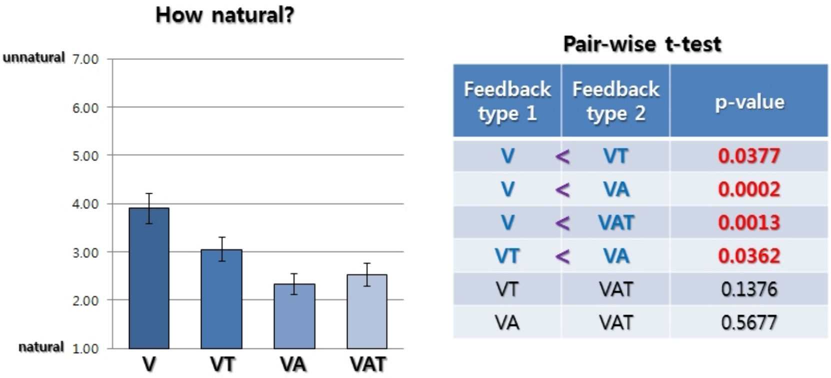
FIGURE 14. Usability result (feedback helpfulness) among the four feedback conditions (stationary)



Feedback type 1	Feedback type 2	p-value
V <	VT	0.0472
V <	VA	0.0000
V <	VAT	0.0011
VT <	VA	0.0008
VT	VAT	0.1032
VA	VAT	0.1454

FIGURE 15. Usability result (ease of learning) among the four feedback conditions (stationary)

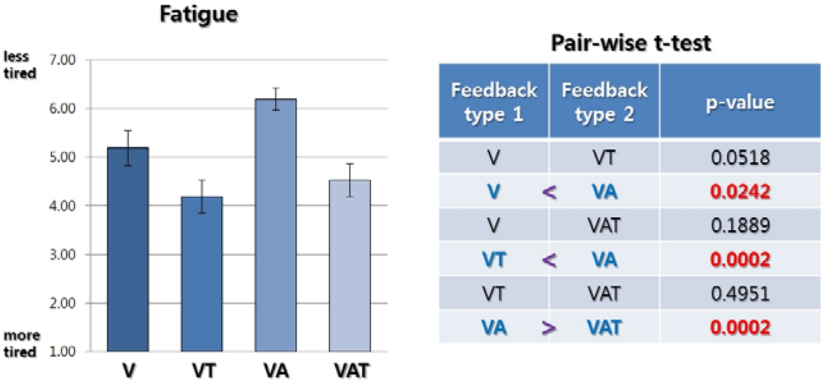
difficult for the users to determine depth solely from visual feedback. Again, the subjects felt the selection technique was easiest to learn, most natural and least tiring with the aural feedback only. We observe in general that when aural and tactile feedback are both presented, the usability and task performance were lower than those when only aural feedback is presented. Thus there seems to be an interaction among these two elements.



Pair-wise t-test

Feedback type 1	Feedback type 2	p-value
V <	VT	0.0377
V <	VA	0.0002
V <	VAT	0.0013
VT <	VA	0.0362
VT	VAT	0.1376
VA	VAT	0.5677

FIGURE 16. Usability result (interaction naturalness) among the four feedback conditions (stationary)



Pair-wise t-test

Feedback type 1	Feedback type 2	p-value
V	VT	0.0518
V <	VA	0.0242
V	VAT	0.1889
VT <	VA	0.0002
VT	VAT	0.4951
VA >	VAT	0.0002

FIGURE 17. Usability result (level of fatigue) among the four feedback conditions (stationary)

In fact, it is reported that redundant feedback may degrade task performance [21] and this result is also consistent with cases when object selection is carried out in larger interaction space [13]. However, it is also quite possible that the vibratory tactile feedback we devised was not ideally designed to human perception.

5. Experiment II: Object Selection Technique while Moving.

5.1. **Experiment design.** The main purpose of Experiment II is to assess the same multimodal feedback effect for the case when the user is moving. It is expected that user motion will make it further more difficult for the user to accurately select the auto-stereoscopically rendered object. The experimental design is mostly the same as Experiment I except that it was conducted for both stationary and moving cases. Thus the stationary case was repeated for the sake of comparison to the case of moving. We compared three different feedback conditions as an aid to make object selection while moving. The feedback conditions are (1) visual only (V, the reference), (2) visual and aural (VA), and (3) visual, aural and tactile (VAT). Differently from Experiment I, VT was not tested, as the effect of tactile feedback was deemed as marginal. We have used the same task and multimodal feedbacks as in Experiment I.

In summary, Experiment II had two factors, the user mobility (with two levels: stationary or moving) and the type of multimodal feedback (with three levels) resulting in 2×3 within subject repeated measure. The same dependent variables (e.g., task performance and usability) were measured. Our main hypothesis was that higher degree of multimodal feedback would generally improve the object selection or depth perception performance even more when the user is moving.

5.2. **Experimental task while moving.** The same depth determination was used as described in the previous section for Experiment I. For the moving case, the user was asked to stand up, hold the device, walk freely (mostly in a circular path) in an empty space (i.e., no worries for collision) at one's comfortable speed, and carry out the designated task (see Figure 18). The user continued to walk and carry out the task until it is finished.

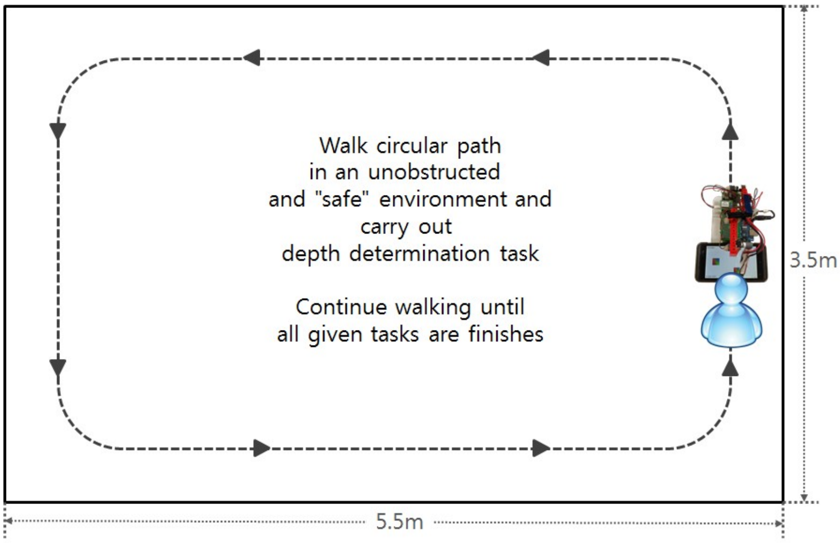


FIGURE 18. A user carrying out the depth determination task while moving

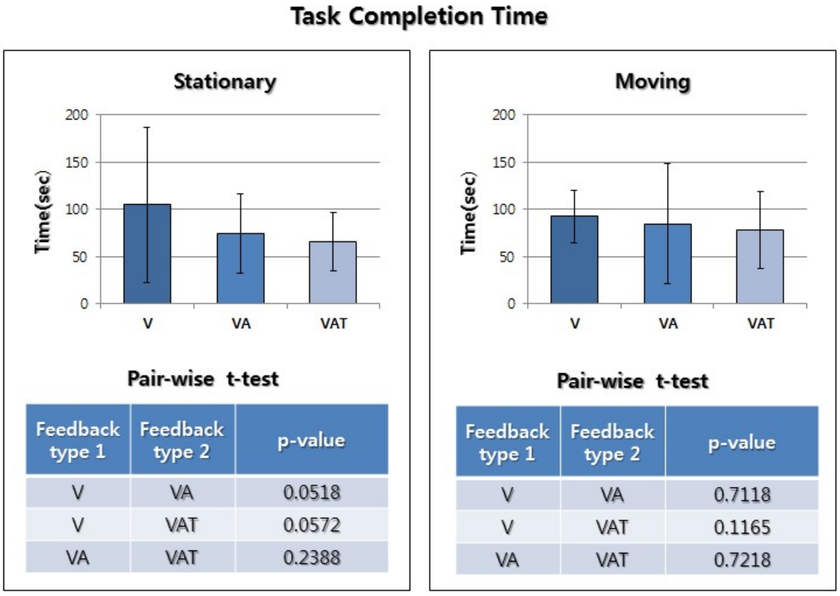
5.3. **Experimental procedure.** Ten paid subjects (9 men and 1 woman), different from Experiment I, participated in the experiment with the mean age of 29.2. After collecting one's basic background information, the subject was briefed about the purpose of the experiment and instructions for the experimental task. A short training (15 to 20 minutes) was given for the subject to get familiarized to the experimental process and use the stylus.

The subject tried out each treatment combination in a balanced order (a total of $2 \times 3 = 6$ treatments). For each treatment, the depth test was conducted five times. The task was carried out in both stationary (sitting) and moving fashion as explained. The task completion time and correctness data were captured and after all the treatments were tried, a general usability questionnaire was filled out (answered in 7 scale Likert scale).

5.4. **Results.**

5.4.1. *Task performance.* Somewhat inconsistently from Experiment I, ANOVA has revealed that there were no statistically significant differences in the task completion time across user mobility and different multimodal feedback conditions (Figure 19). However, VA and VAT did register faster task performance in a trend similar to that of Experiment I (more multimodality, higher performance). In addition, with outliers removed, a statistically significant difference was found (although not shown with the graph).

Figure 20 shows the number of incorrect answers among the different conditions. In this case, while aural and aural/tactile feedbacks were effective for both stationary and moving conditions but with no statistically significant differences between the two, a result that is consistent with Experiment I. However, contrary to our initial hypothesis, no interaction between the two factors was observed. In other words, the multimodal feedback effect was not particularly stronger or weaker for either stationary or moving cases.



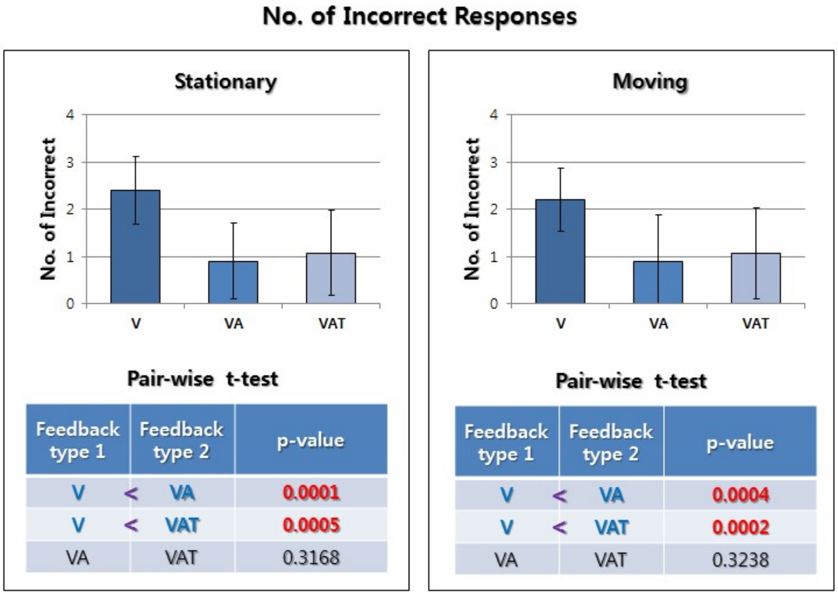
Stationary

Feedback type 1	Feedback type 2	p-value
V	VA	0.0518
V	VAT	0.0572
VA	VAT	0.2388

Moving

Feedback type 1	Feedback type 2	p-value
V	VA	0.7118
V	VAT	0.1165
VA	VAT	0.7218

FIGURE 19. Task completion times for the three feedback conditions (stationary and moving)



Stationary

Feedback type 1	Feedback type 2	p-value
V <	VA	0.0001
V <	VAT	0.0005
VA	VAT	0.3168

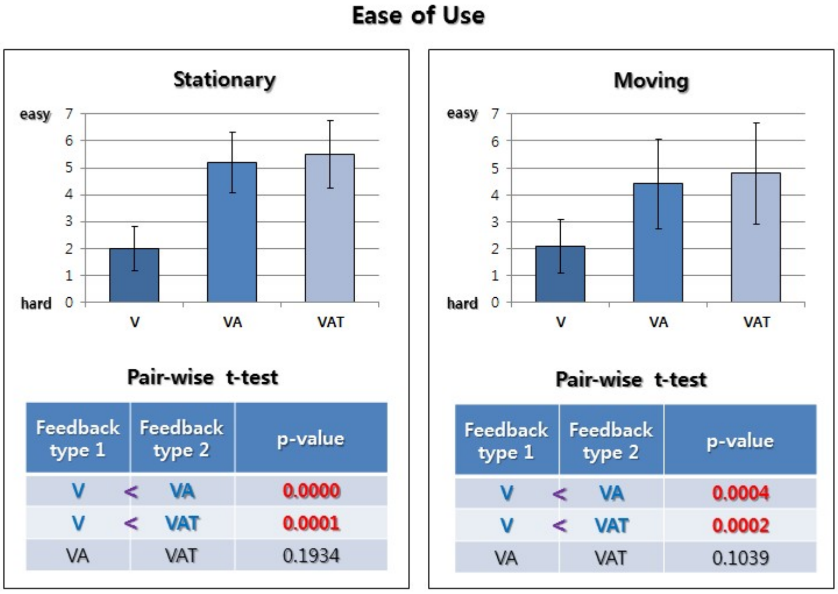
Moving

Feedback type 1	Feedback type 2	p-value
V <	VA	0.0004
V <	VAT	0.0002
VA	VAT	0.3238

FIGURE 20. Number of incorrect responses for the three feedback conditions (stationary and moving)

5.4.2. *Usability.* The usability questionnaire asked the subject to comparatively rate the three selection (or feedback) techniques in terms of ease of use, degree to which feedback was helpful in recognizing the depth, ease of learning, interaction naturalness and the level of fatigue. Figures 21-25 illustrate the results.

The usability results are similar from Experiment I, generally consistent with that of the quantitative performance results. Multimodal feedback was subjectively rated to be easier to use, more helpful, easier to learn, natural and less tiring. In most cases, there was no statistically meaningful difference between VA and VAT, marginalizing the effect of the tactile feedback, as was in Experiment I. However, again contrary to our hypothesis,



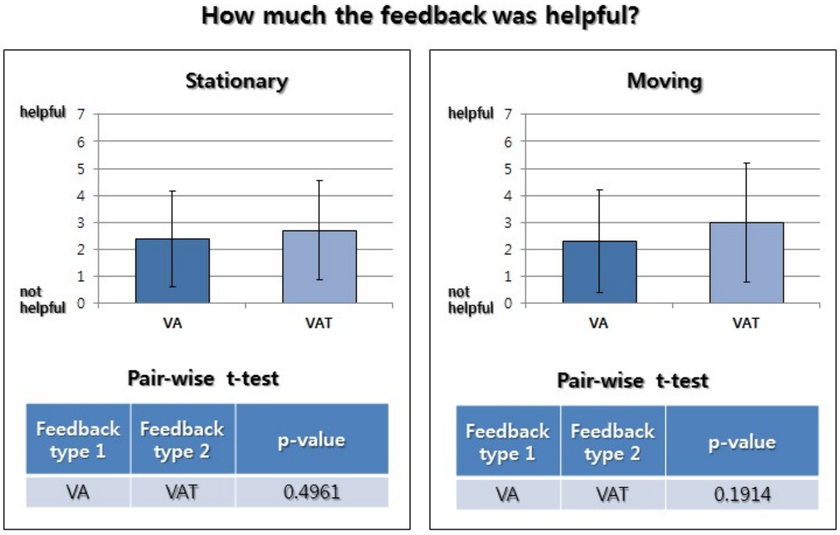
Stationary – Pair-wise t-test

Feedback type 1		Feedback type 2	p-value
V	<	VA	0.0000
V	<	VAT	0.0001
VA		VAT	0.1934

Moving – Pair-wise t-test

Feedback type 1		Feedback type 2	p-value
V	<	VA	0.0004
V	<	VAT	0.0002
VA		VAT	0.1039

Figure 21. Usability result (ease of use) among the three feedback conditions (stationary and moving)



Stationary – Pair-wise t-test

Feedback type 1	Feedback type 2	p-value
VA	VAT	0.4961

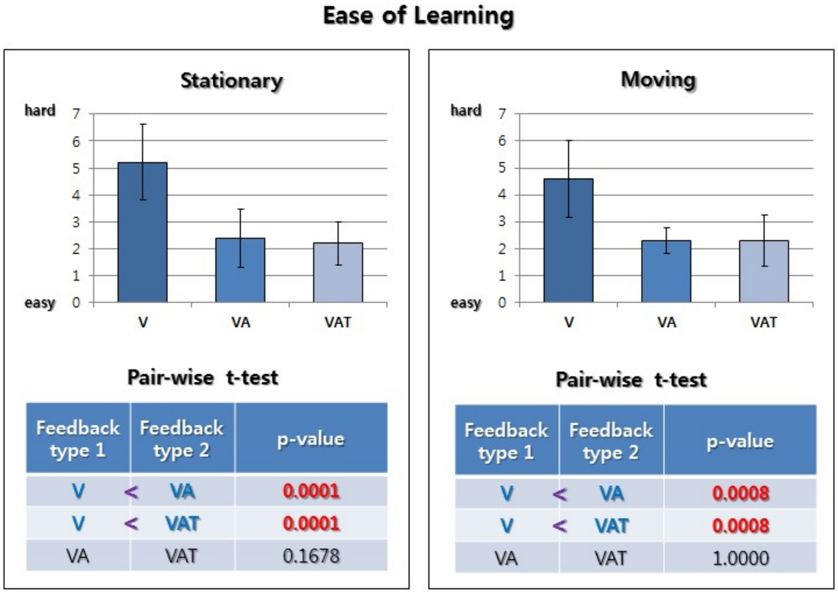
Moving – Pair-wise t-test

Feedback type 1	Feedback type 2	p-value
VA	VAT	0.1914

Figure 22. Usability result (feedback helpfulness) among the three feedback conditions (stationary and moving)

no particular interaction was observed between the two factors, user mobility and types of multimodal feedback.

In the post-briefings, subjects did report added difficulty in object selection during movement compared with the sitting condition. They also reported particularly increased fatigue when they had to walk. We can posit that this difficulty of the task (moving condition being significantly more difficult than the stationary) affected the effects of the multimodal feedback. While as indicated in the related work section, there has been a lot of research work that surprisingly point to virtually no interaction task performance degradation during motion; this was due to the increased task difficulty when moving. That is, for our case too, the positive effect of the multimodal feedback did not manifest itself strongly due to the difference in the task difficulty.



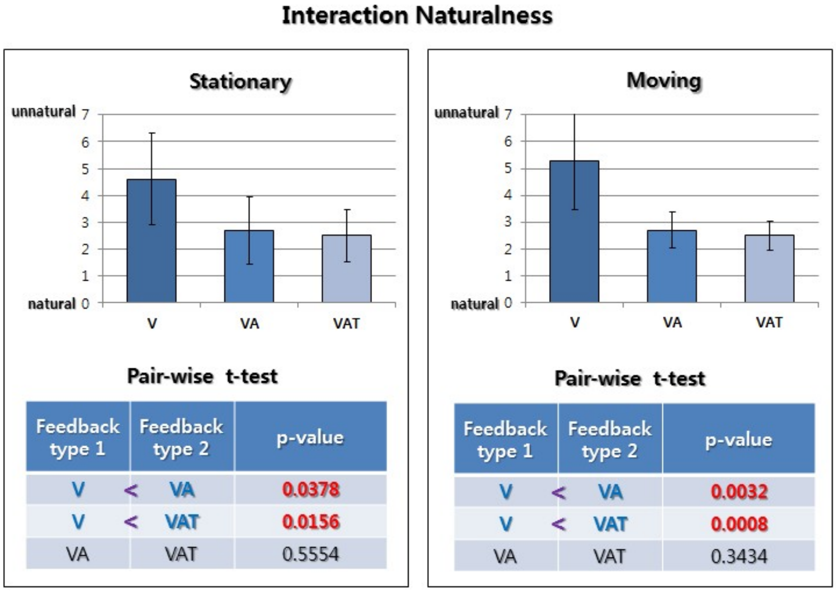
Stationary

Feedback type 1	Feedback type 2	p-value
V <	VA	0.0001
V <	VAT	0.0001
VA	VAT	0.1678

Moving

Feedback type 1	Feedback type 2	p-value
V <	VA	0.0008
V <	VAT	0.0008
VA	VAT	1.0000

FIGURE 23. Usability result (ease of learning) among the three feedback conditions (stationary and moving)



Stationary

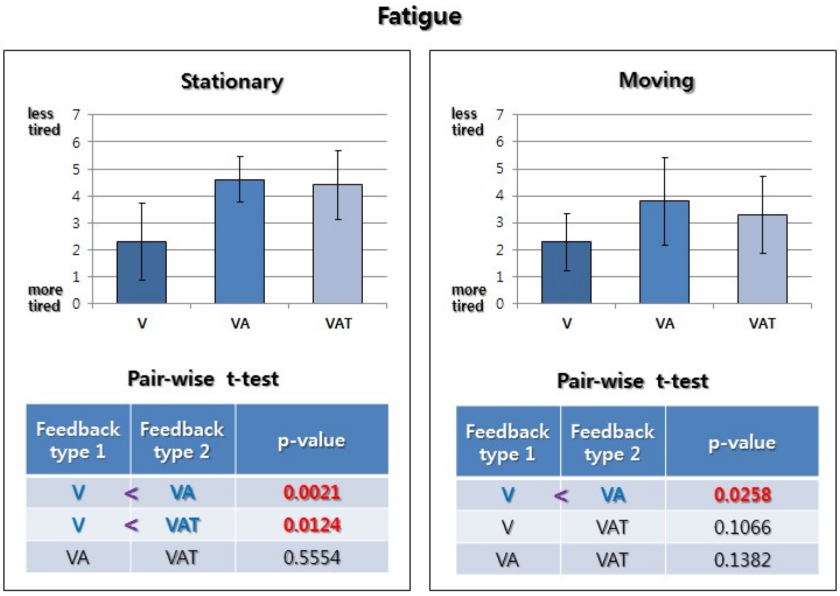
Feedback type 1	Feedback type 2	p-value
V <	VA	0.0378
V <	VAT	0.0156
VA	VAT	0.5554

Moving

Feedback type 1	Feedback type 2	p-value
V <	VA	0.0032
V <	VAT	0.0008
VA	VAT	0.3434

FIGURE 24. Usability result (interaction naturalness) among the three feedback conditions (stationary and moving)

6. **Discussion and Conclusion.** Based on our experiment, we reaffirmed that multimodal feedback helped users select objects better. Post-briefing with the subjects further confirmed this deduction. Many complained of dizziness and blurred imagery in trying to perceive 3D, particularly when moving. This was more apparent with the higher degree of negative parallax (object felt to hover further out of the screen). Despite the possibility of non-ideally designed vibratory tactile feedback method, we converge to a conclusion that only one supplementary (aural) feedback was the most effective object selection method in both stationary and moving cases. Many subjects indicated that they were confused when both aural and tactile feedback were given and preferred only one of the two. They also reported the difficulty to sense the depth with the “vibrating” stylus and due to the



Stationary

Feedback type 1	Feedback type 2	p-value
V <	VA	0.0021
V <	VAT	0.0124
VA	VAT	0.5554

Moving

Feedback type 1	Feedback type 2	p-value
V <	VA	0.0258
V	VAT	0.1066
VA	VAT	0.1382

FIGURE 25. Usability result (level of fatigue) among the three feedback conditions (stationary and moving)

low disambiguating power (relative to the depth range) of the tactile feedback itself. Some even consciously tried to block tactile feedback when presented together with the aural feedback.

Contrary to our expectation, there was no interaction "observed" between the user mobility and multimodal feedback effect. For example, it is plausible to think that the multimodal feedback would be particularly more helpful when the user is moving. However, we realized that the selection task during motion was already a more challenging task; thus, the performance measure between the two cannot be compared directly. All the previous references and subject post-briefing clearly show that the moving case presents a much more challenging task. Unfortunately, the notable cited related works do not consider this aspect in deriving their conclusions. We acknowledge, however, that it would still be difficult to somehow equalize the task difficulty in the experimentation.

Our future work includes investigating in other interaction techniques such as object manipulation and menu selection. We will also apply other methods of aural and tactile feedback.

Acknowledgment. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012-0009232).

REFERENCES

[1] Samsung, *Samsung 3D TV*, http://www.samsung.com/uk/3d/, 2011.
[2] LG, *LG 3D Notebook*, http://www.lg.com/in/computer-products/notebook/LG-A520-D.jsp, 2011.
[3] Nintendo, *Nintendo 3DS*, http://www.nitendo.com/3ds, 2011.
[4] LG, *LG Optimus 3D*, http://www.lge.com, 2011.
[5] D. A. Bowman, E. Kruijff, J. J. LaVioloa and I. Poupyrev, *3D User Interfaces: Theory and Practice*, Addison-Wesley, 2004.
[6] A. Henrysson, J. Marshally and M. Billinghurst, Experiments in 3D interaction for mobile phone AR, *Proc. of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia*, pp.187-194, 2007.

[7] S. Yousefi, F. A. Kondori and H. Li, 3D gestural interaction for stereoscopic visualization on mobile devices, *Proc. of the 14th International Conference on Computer Analysis of Images and Patterns, Part II, LNCS*, vol.6855, pp.555-562, 2011.

[8] J. Postka, S. A. Lewis and D. King, Effects of field of view on judgments of self-location: Distortions in distance estimations even when the image geometry exactly fits the field of view, *Presence: Teleoperators and Virtual Environments*, vol.7, no.4, pp.352-369, 1998.

[9] D. Waller, Factors affecting the perception of interobject distances in virtual environments, *Presence: Teleoperators and Virtual Environments*, vol.8, no.6, pp.657-670, 1999.

[10] U. Yang and G. J. Kim, Increasing the effective egocentric field of view with proprioceptive and tactile feedback, *Proc. of IEEE Virtual Reality Conference*, pp.27-34, 2004.

[11] L. Bouguila, M. Ishii and M. Sato, Effect of coupling haptics and stereopsis on depth perception in virtual environment, *Proc. of the 1st Workshop on Haptic Human Computer Interaction*, pp.54-62, 2000.

[12] E. Murphy, C. Moussette, C. Verron and C. Guastavino, Design and evaluation of an audio-haptic interface, *Proc. of the 4th International Summer Workshop on Multi-Modal Interfaces*, pp.34-38, 2008.

[13] P. Richard, G. Burdea, D. Gomez and P. Coiffet, A comparison of haptic, visual and auditive force feedback for deformable virtual objects, *Proc. of the 4th International Conference on Artificial Reality and Tele-Existence*, vol.94, pp.49-62, 1994.

[14] S. W. Mereu and R. Kazman, Audio enhanced 3D interfaces for visually impaired users, *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, pp.72-78, 1996.

[15] A. Chang and C. O'Sullivan, An audio-haptic aesthetic framework influenced by visual theory correct, *Proc. of the 3rd International Workshop on Haptic and Audio Interaction Design*, pp.70-80, 2008.

[16] C. Müller-Tomfelde, Interaction sound feedback in a haptic virtual environment to improve motor skill acquisition, *Proc. of the 10th International Conference on Auditory Display*, 2004.

[17] Arduino, *Arduino UNO*, http://www.arduino.cc, 2011.

[18] J. Denavit and R. Hartenberg, A kinematic notation for lower-pair mechanisms based on matrices, *Trans. of the ASME. Journal of Applied Mechanics*, vol.22, pp.215-221, 1955.

[19] Wikipedia, *Parallax-Barrier Auto-Stereoscopic Display*, http://en.wikipedia.org/wiki/Autostereoscopy, 2011.

[20] J. Jung and S. Choi, Perceived magnitude and power consumption of vibration feedback in mobile devices, *Proc. of HCI International, Part II, LNCS*, vol.4551, pp.354-363, 2007.

[21] H. Kim and G. J. Kim, Designing of multimodal feedback for enhanced multitasking performance, *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, pp.3113-3122, 2011.

[22] K. Yatani and K. N. Truong, An evaluation of stylus-based text entry methods on handheld devices in stationary and mobile settings, *Proc. of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services*, pp.487-494, 2007.

[23] S. Brewster, Overcoming the lack of screen space on mobile computers, *Personal and Ubiquitous Computing*, vol.6, no.3, pp.188-205, 2002.

[24] B. MacKay, D. Dearman, K. Inkpen and C. Watters, Walk 'n scroll: A comparison of software-based navigation techniques for different levels of mobility, *Proc. of the 7th International Conference on Human Computer Interaction with Mobile Devices and Services*, pp.183-190, 2005.

[25] B. Schildbach and E. Rukzio, Investigating selection and reading performance on a mobile phone while walking, *Proc. of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, pp.93-102, 2010.