3D Virtual Hand Selection with EMS and Vibration Feedback

Max Pfeiffer
University of Hannover
Human-Computer Interaction
Hannover, Germany
max@uni-hannover.de

Wolfgang Stuerzlinger
Simon Fraser University
Vancouver, Canada
http://ws.iat.sfu.ca

CHI'15 Extended Abstracts, Apr 18-23, 2015, Seoul, Republic of Korea
ACM 978-1-4503-3146-3/15/04.
http://dx.doi.org/10.1145/2702613.2732763

Abstract

Selection is one of the most basic interaction methods in 3D user interfaces. Previous work has shown that visual feedback improves such actions. However, haptic feedback can increase the realism or also help for occluded targets. Here we investigate if 3D virtual hand selection benefits from electrical muscle stimulation (EMS) and vibration. In our experiment we used a 3D version of a Fitts' task to compare visual, EMS, vibration, and no feedback. The results demonstrate that both EMS and vibration are reasonable alternatives to visual feedback. We also found good user acceptance for both technologies.

Author Keywords

3D pointing, feedback, vibration, EMS.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Haptic I/O, Input devices and strategies.

Introduction

Selection in 2D is well understood. In comparison, 3D selection is both more complex and less well investigated. One of the largest differences to 2D interaction is that moving one's hand/finger to a 3D location requires control of three degrees of freedom (3DOF). Such interaction is typically associated with input devices that track 3 axes. Current stereo displays

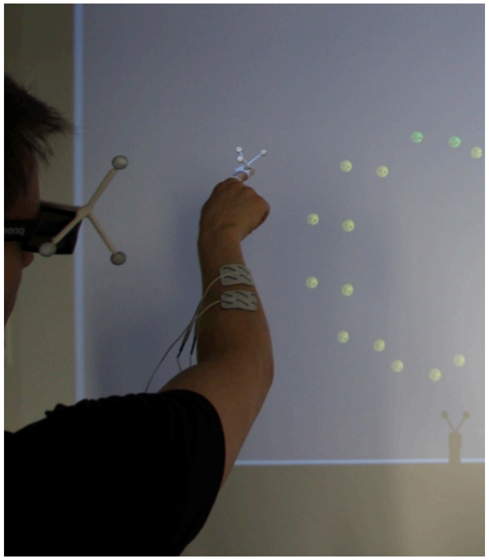

Figure 1 User interacting with a 3D scene. The head and finger trackers are visible, as well as the EMS "pads".

$$MT = a + b \cdot ID \text{, where}$$

$$ID = \log_2\left(\frac{A}{W} + 1\right) \qquad (1)$$

$$TP = \frac{\log_2\left(\frac{A_e}{W_e} + 1\right)}{MT} \text{, where}$$

$$W_e = 4.133 \cdot SD_x \qquad (2)$$

ISO 9241-9 Fitts' law equations [25].

introduce the well-known vergence-accommodation conflict [11]. Consequently, selection of targets in 3D space, e.g., via direct touch, is difficult [4,21], even with the additional depth cues afforded by stereo.
Many 3D selection experiments use highlighting to provide visual feedback when the cursor/finger intersects a potential target object. Another option is haptic feedback, which helps participants "feel" target depths and may improve performance [7]. Yet, its absence may affect one's ability to find the true depth of targets [21]. With haptic feedback a user can feel if they are at the right depth. Yet, human physiology does not transmit haptic input as quickly as visual one, so investigation is needed. Another factor that affects selection is that the finger of the user may occlude the target for small targets. This is the "fat finger" problem. In this case additional feedback methods [24] are necessary and haptics can serve this role.
Most studies in this area use a 3D extension of the ISO 9241-9 methodology [25]. A standardized methodology improves comparability between studies. With this methodology, the benefits of visual feedback have been clearly demonstrated [23]. Yet, and partially due to the lack of standardized experimental methodologies, the effect of haptic feedback with vibration or EMS has not been investigated. Our current work targets this issue.

Related Work

One of two main approaches to 3D selection is virtual finger/hand/3D cursor-based techniques [1,3,16]. The other approach is ray-based. Virtual hand-based techniques rely on the 3D intersection of the finger/hand/3D cursor with a target and thus require that the user picks the correct distance, i.e., visual depth. In such techniques, color change is the most commonly used visual feedback mechanism [1,17].

We employ a 3D extension of the ISO 9241-9 standard [25] based on Fitts' law [9]. This paradigm has been used in recent 3D pointing studies, e.g., [4,6,21,22]. The movement time (MT) is proportional to the index of difficulty (1) and depends on the size W and distance A of targets. Throughput (2) depends on effective measures, and captures speed-accuracy tradeoff.
The effect of haptic feedback has been evaluated in the past, typically with force feedback devices or vibration [2,8]. The results show that haptic feedback increases performance, but that vibration was slightly slower than the non-feedback condition.
EMS offers a broad application range for haptic feedback, ranging from tactile sensations (titillation) to large physical movements. EMS has been tested as a feedback method in games [12], for controlling finger-joints [19] for learning and for gestures such as touch and grasp [15]. The effect of haptic feedback through EMS for selection tasks has not yet been investigated.

Issues with 3D Selection

3D user interface research has studied various factors that affect performance [18,23]. Here, we review several factors relevant to our context, such as occlusion, the "fat finger" problem, and stereo viewing.
First, the finger/hand of the user can occlude objects shown on the display, even if they are positioned to "float" in front of the user's finger/hand relative to the viewer – even in monoscopic, head-tracked displays. A transparent display in front of the hand just reverses the issue by always occluding the hand with contact.
Second, the tip of the finger can occlude targets of similar size or smaller. This is known as the "fat finger" problem in touch interaction [24], but applies to 3D selection as well. To address this, we displayed a cursor slightly above the tracking sleeve worn on the index

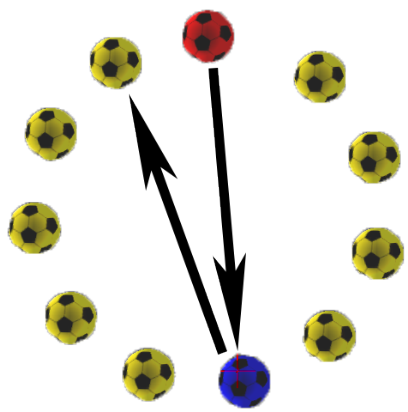

Figure 2 ISO 9241-9 reciprocal selection task with eleven targets. The next target is always highlighted in blue. Targets turn red after they have been hit. Participants start with the top-most one. The arrows indicate the pattern in which the targets advance.

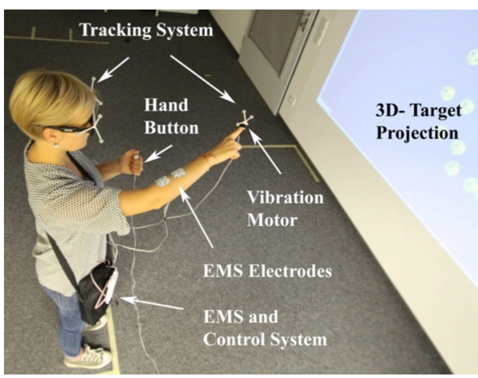

Figure 3 A participant standing in front of 3D projection and performing a task, while performing getting no, EMS, vibrotactile or visual feedback. A button in the other hand indicates selection.

finger. Note that this still leaves the problem that the hand or even the arm of the user can occlude one (or more) targets, especially on downwards motions.
Third, stereo display introduces additional cue conflicts, such as the accommodation-vergence conflict. Also, the human visual system is unable to focus simultaneously on objects at different distances (e.g., a finger and a target displayed on a screen). When focusing on a 3D target displayed on the screen, viewers see a blurred finger or when focusing on the finger, they see a blurred target [5]. This may impact both the ballistic phase of pointing, as the motor program may target the wrong location in space, and also the correction phase, where visual cues are very important [14].
Selection Feedback: In real finger pointing, several cues, including tactile feedback and stereo viewing, indicate when we have touched a target. But when selecting a virtual 3D target using a tracked finger, the finger will typically pass *through* the target, due to the lack of tactile feedback. Stereo cues then indicate that the target is in front of the finger, while occlusion cues indicate the opposite – the finger *always* occludes the screen. Thus, another means of feedback is required. Recent work [23] used target highlighting to address this issue. Haptic feedback is a viable alternative, which can complement or even replace visual cues. It also can increase realism. Haptic feedback can be provided with different devices, including robotic arms. Yet, only vibration and EMS are currently lightweight and mobile enough for most use cases. Both consume only very little power (milli-Watts) and work even for fast motions. Previous work has used vibration as a feedback modality on the lower arm, the hand and the finger tip)[8,13]. EMS, an emergent topic in HCI research, provides another form of haptic feedback. We wanted to compare it to vibration and visual feedback.

Methology

To compare the four different forms of feedback we consider here – none, visual, vibrational, and EMS – we built an appropriate apparatus and designed a Fitts' law study based on ISO 9241-9 [25].
Participants: We recruited 12 participants (3 female) with ages from 21 to 32 (mean = 25.5, SD =3.1). Except for one, all had used 3D technology before and had watched at least one 3D movie at the cinema in the last year. Seven participants had used haptic feedback devices in 2D and 3D games before. Six of the 12 participants had experienced EMS before, four of them for physiotherapy and massage purpose.
Hardware: For stereo display we used a BenQ W1080 ST 3D projector at 1280x800 and 120Hz (projection size 3.3 x 1.9 m). Ten Naturalpoint Optitrack Flex13 cameras provided a 3D tracking accuracy of 0.32mm. The tracking targets were mounted onto a custom, 3D printed finger sleeve. For head-tracking, the tracking targets were attached to the stereo glasses, again via 3D printed mounts (Figure 3). The user wore a small bag, which contained the control electronics for the vibration motor and the EMS, driven by an Arduino Uno for access via WiFi (Figure 4). To enable participants to indicate selection with the other hand, we inserted a Logitech mouse button into a 3D printed handle. We created a custom application in Unity 4 with the iminVR MiddleVR 1.4 plugin. The application was connected to the vibration and the EMS device through the WiFi interface of the Arduino. The virtual reality simulation has an end-to-end the latency of 54.6 ms (SD = 5.24), the EMS condition 61.8 ms (SD = 4.76), and the vibration condition 66.6 ms (SD = 6.39).
We mounted a KF2353 vertical vibration motor (9,000 rpm, 90 mA at 2.3V) within the tracking sleeve below the fingertip, with hook-and-loop fasteners. With this

Figure 4 EMS feedback unit, Arduino Uno, WiFi unit and control board.

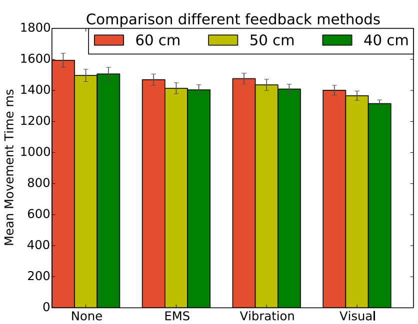

Figure 5 The average movement time for all conditions and three depths level.

mounting method, the sound of the vibration motor is very low, too small to be easily audible in the lab environment. For EMS feedback we use an off-the-shelf device (Beurer SEM43 EMS/TENS) for EMS signal generation (Figure 4). In previous studies we found that impulses with 50μs duration and a frequency of 80 Hz are suitable for a large range of users. We calibrated the intensity of EMS for each user individually to account for different skin resistance and the variance of the contraction effect. We placed 40x40mm self-sticking electrodes on the arm surface above the musculus extensor digitorum. When the user holds the index finger in a pointing position as shown in Figure 1, this muscle lifts the index finger up, which simulates the sensation of hitting a (light) physical object. We scale the EMS intensity during calibration down to a level where the finger itself visual does not move. Across our participants this happened at an average current of 21.41 mA (SD 3.4 mA) and an average voltage of 79.8 V (SD 23.77 V).

Feedback: In all conditions the user sees a 1x1x1 cm cross as cursor about 1cm above the finger sleeve. When the cursor is inside the target in the visual feedback condition, the target is highlighted red. In the vibration and EMS conditions the user gets haptic feedback as long as they are inside the target.

Study Design: Our study had two independent variables: 4 types of feedback and 3 target depths, for a 4x3 design. The four feedback types were: none, EMS, vibration and visual feedback. Target depths were 40, 50 and 60 cm from the user. We used 3D balls (Figure 2) as targets, with sizes of 1.5, 2, and 3 cm. We arranged them in circles with 20, 25 and 30 cm diameter. Similar to previous work [20], we positioned targets within the circle at the same target depth. The order of all of the above conditions and factors was determined by Latin squares to minimize learning effects. In total, our experiment had thus 4x3x3x3 = 108 target circles with 11 targets each.

Procedure: At the start we introduced the participants to the context of the study and asked them to fill a consent form and a questionnaire for background and demographic information. Then we demonstrated vibration feedback. Subsequently, we attached the electrodes and calibrated the EMS step-by-step to a level where the finger just stopped moving. After the calibration we measured the current and voltage for the EMS. We placed participants 2 m in front of the screen. We then equipped them with the 3D glasses and made sure that the finger sleeve was placed correctly. If the user clicked the button while the cursor was in the target, we registered a "hit". Otherwise a selection error was recorded. At the end, participants were asked to fill a second five point Likert scale based questionnaire to gather qualitative feedback.

Results

As the data for movement time was not normally distributed, we log-transformed the data before analysis. Also, we filtered outliers beyond 3 standard deviations from the mean in terms of time and target position. This typically corresponded to erroneous double-selection episodes and removed 350 trials (2.5%). Subsequently we analyzed movement time, error rate and throughput each with a repeated measures ANOVA test.

Movement Time: The ANOVA identified a significant effect for movement time, $F_{3,33}=5.9$, $p<0.005$. According to a Tukey-Kramer test, only the none and visual feedback conditions were significantly different. The average movement times for the none, EMS, vibration and visual feedback conditions were 1522ms,

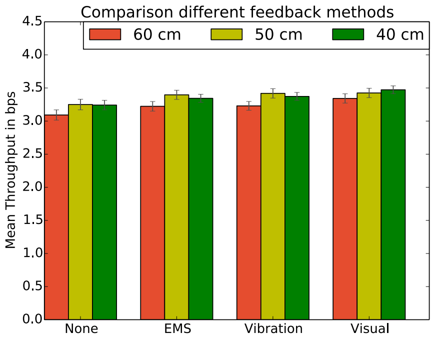

Figure 6 The throughput of all conditions and three different depths level.

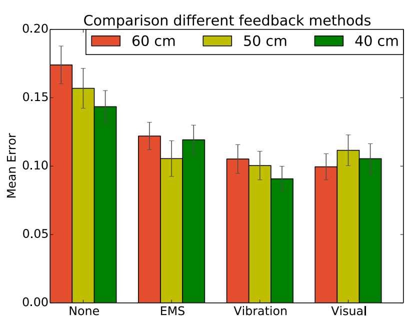

Figure 7 The error rate of all conditions and three depths level.

1449ms, 1465ms, and 1387ms, respectively (Figure 5). In terms of target depths, there was also a significant effect, $F_{2,22}=10.86$, $p<0.001$, with the two levels closest to the user being significantly faster to select than the "deep" level.

Error Rate: An ANOVA identified a significant effect for error rate, $F_{3,33}=6.05$, $p<0.005$. According to a Tukey-Kramer test, the none condition was significantly worse than all others. The average error rates for the none, EMS, vibration and visual feedback conditions were 15.3%, 11.3%, 9.8%, and 10.5%, respectively (Figure 6). For target depth, there was no significant effect on errors, $F_{2,22}<1$, $p>0.5$.

Throughput: The ANOVA identified a significant effect for throughput, $F_{3,33}=3.58$, $p<0.05$. According to a Tukey-Kramer test, only the none and visual feedback conditions were significantly different. The average throughput values for the none, EMS, vibration and visual feedback conditions were 3.19, 3.28, 3.29 and 3.37, respectively (Figure 7). Likely, this throughput result is mostly due to the differences in movement times. For target depth, there was a significant effect on throughput, $F_{2,22}=6.73$, $p<0.01$. The 40 and 50cm levels had again significantly more throughput than the farther 60cm level.

Subjective Results: The participants could differentiate between the haptic feedback methods with a median rating of 1, (where 1 = very well and 5 = not at all) and median absolute deviation (MAD) is 0. All three feedback methods were ranked as reasonable realistic with a median rating of 2 (1 = very realistic to 5 = very unrealistic, MAD = 1). When we asked for the perception on delay in the feedback (where 1 = very low delay and 5 very high delay), the EMS feedback and visual feedback were ranked best (median = 1, MAD = 0), followed by vibration feedback (median = 1.5, MAD = 0.5). Also the position of the feedback was ranked as well fitting (where 1 = very good and 5 = very bad). The EMS feedback at the lower arm was ranked with a median of 2 (MAD = 1) and for vibration at the fingertip with a median of 1 (MAD = 0). While most of the participants were very comfortable with the EMS during the selection task, four reported at the end of the study that the EMS impulses were too strong and sometimes moved the finger out of the targets.

Discussion

Overall, the visual feedback condition performs better than the haptic feedback conditions, but not significantly so. This matches previous work, which identified visual feedback as faster than haptic feedback [10]. The results for vibration and EMS feedback are not significantly different from those of visual feedback, nor from the none condition. In contrast to another study [8], we found that vibration was more effective that no feedback, but again not significantly so. Although the lack of a significant difference does not "prove" equality, our results still indicate that vibration and EMS both could be viable alternatives for feedback in 3D virtual hand selection without a significant cost in terms of throughput. Moreover, participants ranked both conditions very positive. Based on the indications from the experiment results and the positive qualitative results, we see both haptic modalities as reasonable alternatives to visual feedback. Also, some users mentioned that they might like EMS feedback in games.

Conclusion and Future Work

This work presents a first evaluation of a lightweight, low-energy haptic feedback system to assist 3D virtual hand selection. Overall we found that both vibration and EMS are reasonable alternatives to visual feedback.

There are still several open questions, such as how haptic feedback performs in selection tasks where the targets have different visual depths or if targets are straight behind each other. In the future we will investigate how haptics perform together with visual feedback and how different feedback strength or locations affect performance. We will also investigate how a technique that attracts the finger to the target can be realized by using more than one muscle group.

References

[1] Achibet, M., et al. The Virtual Mitten: A novel interaction paradigm for visuo-haptic manipulation of objects using grip force. IEEE 3DUI 2014, 59–66.

[2] Argelaguet, F. and Andujar, C. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 2013, 121–136.

[3] Bowman, D.A., Kruijff, E., et al. *3D User Interfaces: Theory and Practice.* Addison-Wesley, 2004.

[4] Bruder, G., et al. To Touch or not to Touch? Comparing 2D Touch and 3D Mid-Air Interaction on Stereoscopic Tabletop Surfaces. *ACM SUI 2013*, 9–16

[5] Bruder, G., Steinicke, F., et al. Touching the Void Revisited: Analyses of Touch Behavior on and above Tabletop Surfaces. *INTERACT 2013*, 278–296.

[6] Bruder, G., et al. Effects of visual conflicts on 3D selection task performance in stereoscopic display environments. *IEEE 3DUI* 2013, 115–118.

[7] Chun, K., et al. Evaluating haptics and 3D stereo displays using Fitts' law. *IEEE Creating, Connecting and Collaborating through Computing* 2004, 53–58.

[8] Corbett, B., Yamaguchi, T., and Liu, S. Influence of haptic feedback on a pointing task in a haptically enhanced 3D virtual environment. *Human-Computer Interaction International*, 1954 (2013), 561–567.

[9] Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 1954, 381–391.

[10] Gerovichev, O., et al. The Effect of Visual and Haptic Feedback on Manual and Teleoperated Needle Insertion. *MICCAI 2002-Part I*, 147–154.

[11] Hoffman, D.M., et al. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision* 2008, 33.1–30.

[12] Kruijff, E., Schmalstieg, D., and Beckhaus, S. Using neuromuscular electrical stimulation for pseudo-haptic feedback. *ACM VRST* 2006, 316.

[13] Lehtinen, V., et al. Dynamic tactile guidance for visual search tasks. *ACM UIST* 2012, 445-452.

[14] Nieuwenhuizen, K., Martens, J., Liu, L., et al. Insights from Dividing 3D Goal-Directed Movements into Meaningful Phases. *IEEE CG&A* 2009, 44–53.

[15] Pfeiffer, M., Alt, F., and Rohs, M. Let Me Grab This: A Comparison of EMS and Vibration for Haptic Feedback in Free-Hand Interaction. *Augm. Human 2014*, 1–8.

[16] Poupyrev, I., et al. Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. *CGF* 1998, 41–52.

[17] Prachyabrued, M., Borst, C.W., Borstt, C.W. Visual feedback for virtual grasping. *IEEE 3DUI* 2014, 19–26.

[18] Stuerzlinger, W. Considerations for targets in 3D pointing experiments. HCI Korea 2015, 162-168.

[19] Tamaki, E., Miyaki, T., and Rekimoto, J. Possessed-Hand: Techniques for controlling human hands using electrical muscles stimuli. *ACM CHI* 2011, 543.

[20] Teather, R.J., Pavlovych, A., et al. Effects of tracking technology, latency, and spatial jitter on object movement. *IEEE 3DUI* 2009, 43–50.

[21] Teather, R.J. and Stuerzlinger, W. Pointing at 3D targets in a stereo head-tracked virtual environment. *IEEE 3DUI* 2011, 87–94.

[22] Teather, R.J. and Stuerzlinger, W. Pointing at 3d target projections with one-eyed and stereo cursors. *ACM CHI* 2013, 159-168.

[23] Teather, R.J. and Stuerzlinger, W. Visual aids in 3D point selection experiments. *Proceedings of the 2nd ACM SUI* 2014, 127–136.

[24] Vogel, D. and Balakrishnan, R. Occlusion-aware interfaces. *ACM CHI* 2010, 263.

[25] ISO 9241-9:2000 - Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 9: Requirements for non-keyboard input devices. *International Organization for Standardization*, 2000.