Augmented Reality Target Finding Based on Tactile Cues

Teemu Ahmaniemi

Vuokko Lantz

Nokia Research Center

Itämerenkatu 11-13, 00180, Helsinki, Finland

{teemu.ahmaniemi, vuokko.lantz}@nokia.com

ABSTRACT

This study is based on a user scenario where augmented reality targets could be found by scanning the environment with a mobile device and getting a tactile feedback exactly in the direction of the target. In order to understand how accurately and quickly the targets can be found, we prepared an experiment setup where a sensor-actuator device consisting of orientation tracking hardware and a tactile actuator were used. The targets with widths 5°, 10°, 15°, 20°, and 25° and various distances between each other were rendered in a 90° -wide space successively, and the task of the test participants was to find them as quickly as possible. The experiment consisted of two conditions: the first one provided tactile feedback only when pointing was on the target and the second one included also another cue indicating the proximity of the target. The average target finding time was 1.8 seconds. The closest targets appeared to be not the easiest to find, which was attributed to the adapted scanning velocity causing the missing the closest targets. We also found that our data did not correlate well with Fitts' model, which may have been caused by the non-normal data distribution. After filtering out 30% of the least representative data items, the correlation reached up to 0.71. Overall, the performance between conditions did not differ from each other significantly. The only significant improvement in the performance offered by the close-to-target cue occurred in the tasks where the targets where the furthest from each other.

Categories and Subject Descriptors

H.5.2 User Interfaces: Haptic I/O;

General Terms

Documentation, Design, Experimentation, Human Factors.

Keywords

Augmented reality, Pointing, Haptics, Fitts' law

ICMI-MLMI'09, November 2–4, 2009, Cambridge, MA, USA.

1. INTRODUCTION

Pointing is a natural gesture in situations where indication of direction is needed. If a person is lost and asks for the way to a certain spot in the town, the guidance is nearly impossible without showing the way by pointing. Recently, low-cost mobile navigation systems, such as GPS navigators, have become available for consumers and there is not so much need for asking for help or skill for reading paper maps. The navigators are measuring the accurate location of the user and indicating it with a pointer symbol on top of the map on the screen. Landmarks or other points of interests may be represented on the map by other specific symbols. In addition, some systems are measuring the heading of the device and rendering the map in absolute orientation with respect to the user's orientation. However, these systems still leave the interaction between the virtual objects and the real world objects abstract. It may not be completely intuitive to the user which actual landmark or object corresponds to the virtual marker on the screen, e.g. does 'this' arrow here mean 'that' church over there? Also, although the device heading would be in line with the environment, the user cannot get a picture of both the virtual and real object at the same glance; first she/he has to look at the virtual object on the device screen and then the real object in the assumed direction.

Pointing with the device to the direction of interest would solve the problem of associating the virtual and real object but then the visual rendering of the virtual object would form another problem; the screen cannot be seen if the device is used for pointing. Therefore, speech, audio, and haptic modalities could be solutions for representation. Several studies have shown the usefulness of speech in pointing interaction both as input and output modality [1, 2]. Also, non-speech auditory cues have proven to be accurate in some contexts to indicate the direction of interest [3, 4]. However, the general problems of speech and audio modalities in public contexts are the lack of privacy and interference of ambient noise. Although speech can be used for revealing many details of the virtual object, it can not be heard if the speech is masked by the environmental sounds such as traffic noise and people talking loudly on the street. Furthermore, representing the virtual object by speech or auditory cue takes quite a lot of time, which may have critical impact on user experience [1]. This study aims at investigating the potential of tactile feedback in pointing interaction. An obvious benefit of the tactile feedback is its tolerance to ambient interference. In addition, tactile feedback provides a natural and fast response to pointing. Our main interest is in finding out how accurately the targets can be represented and how fast they can by found. We base our work on research work done on pointing interaction and aim to investigate by experimentation whether the tactile rendering of the targets follows the same principles as the other modalities.

1.1 User Scenario

Our study is based on a user scenario where virtual targets in the environment could be scanned by pointing. An example of such scenario is described in the following paragraph:

John and Tom had decided to grab a beer in downtown around six o'clock. They had agreed to meet each other on the central square, where John arrives a bit earlier. When Tom arrives, he can not locate John since there are so many people on the square. Tom picks his mobile phone from his pocket and selects 'scan'-option under John's name in the phone book. He explores the environment by pointing the phone into the direction of the people. A tactile feedback in the certain direction reveals the exact location of John. Tom starts walking into the direction and soon he recognizes John's blue jacket in the crowd.

A similar scenario of scanning contacts by pointing was presented also by Stratchan and Murray-Smith [5]. Their idea was to use pointing gesture to reveal more details about the contact located in the pointing direction. This kind of 'GeoPoking' could be a private and subtle way to interact with closely related people. The information, such as availability of the poked person was suggested to be represented by audio or tactile feedback. Strachan et al. [6] and Williamson et al. [7] showed that pointing could be used as predictive assistance for navigating in augmented space. In their approach, in addition to presenting some information of the pointed target, the uncertainty of the pointing accuracy could also be presented. Their solution was to use Monte Carlo sampling for approximating the uncertainty of pointing and audio and tactile feedback for representing that to the user. Similarly, Robinson et al. [8] presented a study where pointing was used for discovering geo-tagged information embedded in the environment. The type of the information was represented either by vibrotactile feedback or visual icon when the pointing direction was in parallel to the direction of the target. In their study, the tactile feedback also reflected the amount of information available in the target, which was realized by the spread of the vibration.

1.2 Pointing and Rendering Accuracy

Although the above studies suggest using tactile feedback for target or path finding, a detailed investigation of the methods to present the feedback is still missing. The characteristics of tactile feedback are always highly depended on the actuator properties. In the above studies, the tactile feedback has been presented in a mobile device with a traditional vibration motor. This is a practical choice, since this type of actuator is a very low cost component and widely used in mobile devices, such as mobile phones. However, a vibration motor is limited in its capacity of producing a wide range of frequencies, intensities, or waveforms [9]. Linear electromagnetic actuators, such as C2 Tactor [10], provide much more flexibility in the actuation control. C2 provides a decent frequency response of 200-300 Hz and it can be controlled with an audio signal allowing modifications in the waveform and intensity. In addition, actuators of this kind are still relatively low cost and small in size to be integrated in mobile devices.

Similarly, the pointing accuracy depends on the sensor properties. Although the pointing uncertainty can be taken into account in the representation of the pointed objects [6, 7], the simplest way to improve the interaction is to maximize the accuracy of the pointing. Pointing that is based only on the magnetometer signal tends to suffer from long latencies because the earth's magnetic field is rather weak to be measured with portable sensors and thus requires heavy filtering. In order to include vertical dimension into pointing, accelerometers and angular velocity sensors can be considered. Sensor hardware such as SHAKE (http://www.samh-engineering.com/) and Nokia Motion Band [11], widely used in interaction research, provide readings of magnetometer, accelerometer, and gyroscope and enable the calculation of 3-dimensional orientation with respect to earth's magnetic field and center of gravitation. However, due to their small size and mobility, the designers of these devices have been forced to find a compromise between the sensor performance and power consumption. This has caused them not to be able to utilize the most accurate sensors for orientation tracking. With the expense of power consumption, mobility, and size, Xsens MT9 [12] motion tracking system provides accurate orientation data (±0.05°) with reasonably large dynamic range of angular velocity (±900°/s). This system was used in the current study for orientation tracking.

Tactile actuator and the continuous feedback may interfere with the sensor measurements. Especially magnetometer tends to suffer from the proximity of the tactile actuator which is heavily magnetic. The phenomenon is emphasized when the device is moving with respect to the earth's magnetic field. Therefore, the readings of the magnetometer should be taken into account with consideration. In this study, the magnetometer readings were ignored completely. The accelerometers and angular velocity sensors provided orientation readings stable enough for our purposes.

1.3 Tactile Synthesis

In our previous work, we introduced the idea of dynamic tactile synthesis which allows real-time control of tactile output based on dynamic input, such as the angular velocity of gesturing [13, 14]. We used Xsens motion tracking system for measuring the movements of the user's hand and fed the motion data into audio synthesis software, Pure Data (PD). The synthesis was based on the mixing of the optimized resonance signal for the actuator, 250-260Hz sinusoid, and an envelope signal. The dynamic character of the feedback was realized by modifying the envelope signal amplitude and frequency according to the gesture input. The synthesized signal was fed to a tactile actuator (C2) which was housed in the same chassis as the motion tracker hardware. Using the angular velocity of the user's hand as input, we created dynamic virtual textures with various ridge shapes and spatial densities. The system created an illusion of virtual textured canvas that was explored by pointing the device to it.

1.4 Target Representation with Tactile Feedback

One of the potential use cases of the virtual textures was considered to be augmented reality pointing. We stated that the virtual textures could be used for extracting information from virtual objects in augmented reality by pointing to the direction of them and by performing probing gestures. The focus of the previous study was on the probing and texture perception. In this study, we focus on the target finding, i.e. pointing interaction, and investigate this use case in laboratory condition. However, although using exploration velocity to control the virtual texture density is a novel and attractive approach, in our pilot testing it appeared to be problematic. The objects in the augmented space should be relatively small (e.g. <20°) in order to achieve accurate pointing and enabling the user scenario described above. Using exploration velocity as the control of the synthesis assumes the pointing to be on the target in order to get a picture of the texture. If the target is very small, e.g. <5°, the exploration becomes overly difficult. Therefore, more practical approach in this scenario is to use pointing direction, or the deviation from the target, as an input for the tactile synthesis.

There is a number of ways to map the direction of pointing into tactile synthesis. The tactile feedback could become more frequent and intensive when approaching the center of the target as described by Robinson et al. [8] and discussed above. Or there could be a continuous tactile feedback when pointing is on the target. This approach was used in the study by Akamatsu et al. [15], where participants were asked to perform target selection task by mouse pointing on a computer screen. The tactile feedback was provided by a solenoid driven pin attached to the left mouse button. The tactile feedback did not shorten the target selection time but the time between the entering the target and providing the response was significantly shorter with tactile feedback than without it. In addition to feedback-on-the-target, Tähkäpää and Raisamo [16] investigated also conditions where tactile feedback is provided near the target or far from the target in a similar target selection task. The task times within conditions did not differ significantly from each other but the feedback-on-the-target was clearly liked the most compared to other conditions.

The role of tactile feedback is rather different when there is no visual feedback at all. This condition was analyzed in the study by Oron-Gilad et al. [17]. They studied the efficiency of different tactile close-to-target cues in a target selection task. The experiment setup consisted of a mouse equipped with two tactile actuators placed on different sides of it. The close-to-target cues where tactile pulses either becoming more frequent or less frequent when getting closer to the target. The on-the-target cue was frequent, infrequent or suppressed. The results showed that targets were selected significantly faster when the difference between the on-the-target cue and close-to-target cues was larger on the border of the target. However, the study did not reveal what was the benefit of the close-to-target cue compared to not having it at all. Also, the effect of target width and close-to-target width in target selection time was not studied.

1.5 Pointing, Finding and Human Control Theory

When studying target selection in spatial interaction, a commonly used theory for human control, Fitts' law [18], cannot be ignored. The law models the human control behavior when a person is aiming to reach a target that is visible. It states the time that is needed for reaching a target as a function of target width and distance to it. Based on experimental data, the model is the form:

$$MT = a + b \log_2(\frac{A}{W} + 1) \qquad (1)$$

Where MT is the movement time starting from target observation to target selection, A is the distance from the initial position to the target and W is the target width. The logarithmic term is often called as index of difficulty (ID) and a and b are system depended constants. The model includes the processes of perception of the target, planning the motor control program to reach the target, execution of the program, updating the control during the execution and finally confirmation when the target has been reached [19]. This model has been widely used in human-computer interaction research. Recently, the model was used by Crossan et al. [20] in gesture based interaction where wrist rotation was used for controlling a pointer on a mobile device screen. In a target selection task the test participants' performance correlated with model of the Fitts' law. Also, Cabral et al. [21] studied a target selection task in a virtual reality environment. Camera tracked gestures were used for pointing and selecting targets. Task time followed Fitts' model although arm pointing was considered unfamiliar and somewhat exhausting.

The movement control is rather different when the user does not see the target or have any prior knowledge of the target location. This is basically the case in our study since the target is rendered only by tactile feedback when pointing is close to or on the target. The user may have an approximate idea in which direction to go but the absolute location of the target is still unknown. Hence, the user has to develop a strategy to find the targets, which in our case was simply the selected scanning velocity. Similar interaction challenge is faced in so called *peephole displays* [22], in which a large visual interface is looked through a smaller window. The user cannot see the interface outside of the window, which makes the target finding similar to our setup. Cao et al. [23] studied the peephole target finding with a setup where window size (S), target width (W) and distance (A) to target were varied. In this setup, the task has basically two phases: first the user has to get the target into the peephole window and then to select the target within the window. Cao et al. [23] developed a model for these two stages following the principle of the Fitts' model. The model was proved to be valid by a series of experiments. Similar findings were presented in a study by Rahs and Oulasvirta [24] where the camera of a mobile phone was used for finding targets on a large display. Although the results were correlating with Fitts' law quite well, the two phase model similar to that of Cao et al. [23] provided even better correlation. Also Andersen [25] studied the case where the target was not initially visible to the user in a display scrolling task. In his data the task time did not correlate well with the index of difficulty but a simpler model where task time was linearly proportional to the target distance seemed to offer much better correlation.

Our aim in this study was to investigate the effect of target width and distance in target selection time related to the user scenario described above. Our experimental setup allowed us to render the targets rather accurately in space being explored which encouraged us to study the validity of Fitts' law in non-visual conditions. The procedure of our experiments was similar to those with visual target selection studies and the tactile cues were selected based on our earlier studies and conclusions of study by Oron-Gilad [17].

2. EXPERIMENT

The experiment documented in this paper was the first investigation of the tactile rendering of various widths of targets in space. Therefore, the experiment setup was aimed to be as simple as possible. To simplify the search task, the targets in this experiment were represented as vertical ribbons that can be found by horizontal scanning, similarly as in the study by Cao et al. [22].

2.1 Experiment Equipment

The sensor hardware and the tactile actuator were similar to those used in our previous studies [13, 14]. The sensor of the type Xsens MT9 [12] was housed in an 85x40x15mm plastic chassis together with a tactile actuator C2 (See Figure 1.). The sensor data was sent to a PC (Acer Travelmate C110) via serial cable. The data acquisition and processing was done on a Windows software developed with the sensor manufacturer toolkit. It provided the orientation values of the sensor within accuracy of 0.1° sampled with 100Hz. The software controlled the test sessions and feedback synthesis and stored all the sensor data into files.

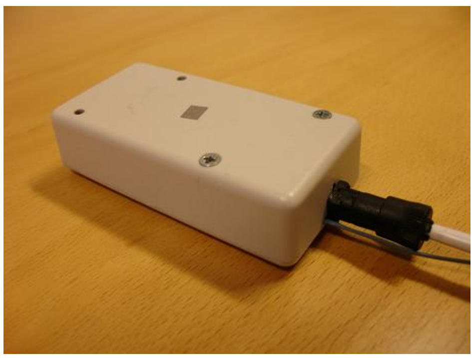

Figure 1. Sensor-actuator device consisting of motion sensor hardware and a vibrotactile actuator C2.

The feedback synthesis was done using Pure Data (PD, http://puredata.info) audio synthesis software. The orientation values, target locations and widths were sent to PD. A PD program (patch) compared the current pointing location to the target values and performed the synthesis accordingly. In this study, there were two different feedback cues: one for indicating that the pointing is on the target and one indicating the proximity of the target.

2.2 Target Rendering

Based on our earlier work [13, 14] we selected two feedback synthesis methods for the on-the-target and close-to-target cues. On-the-target cue was created by mixing the actuator resonance frequency sinusoid, 260 Hz, with a smooth sine wave like envelope signal of frequency of 30 Hz. The signal remained the same as long as the pointing was within the target. The close-to-target cue, on the other hand, became more intense when approaching the target. In the synthesis we used a similar approach as in the on-the-target cue but the envelope shape was different. It was based on a rectangular shape wave, where the ridge part of the envelope was longer than in the sine wave shape envelope. The frequency and the amplitude of the envelope signal increased when approaching the target. The cue started always 10 degrees before the border of the target and was switched off when entering the target. The envelope frequency increased linearly from 0 to 50 Hz. The signal amplitude increased linearly towards the border but reaching only 20% of the amplitude of the on-the-target cue. These cues were assumed to be distinguishable from each other. The drive signals of the cues are illustrated in Figure 2. The signal corresponds to a steady horizontal movement of the sensor-actuator device passing the target with constant velocity. The illustration reveals the effect of scanning behavior on the vibration; the higher the scanning velocity is, the shorter time the target is active. Also, since the close-to-target cue basically widens the targets, it was expected to be more noticeable with higher velocities than targets without the close-to-target cue.

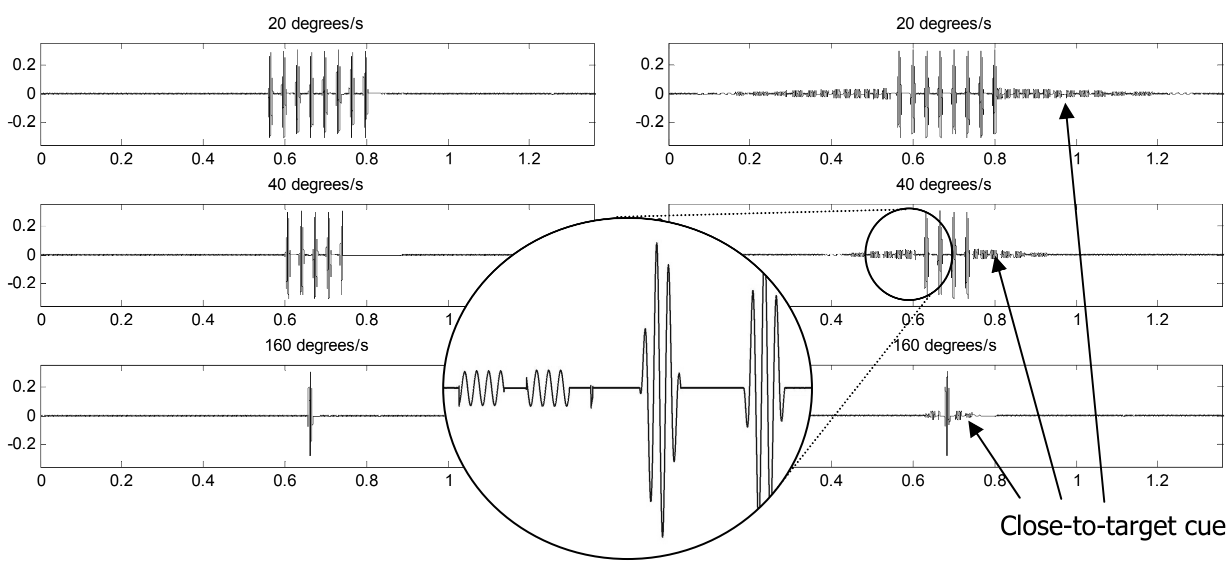

Figure 2. The drive signal of the tactile actuator in both conditions (left: without close-to-target cue, right with the close-to-target cue). The target width is 5 degrees whereas the width of the close-to-target cue is 10 degrees. The signals have been obtained by simulating the pointing interaction with three different scanning velocities, 20, 40 and 160 degrees/s. Since the envelope frequency of the target is constant, fewer envelopes are rendered with higher scanning speed although the target width remains the same. The close-to-target cue is visible in the right part of the figure as a raised contour around the actual targets. The horizontal axis corresponds to time in seconds and vertical axis corresponds to the signal amplitude.

2.3 Experiment Setup and Procedure

Eight subjects participated in the test. The participants were Nokia Research Center employees in the age range from 28 to 38. All the participants were right handed without any disabilities in tactile perception or motor control. The participants were seated in a small office room equipped with the sensor-actuator device, a PC, a monitor and small keypad connected to the PC used for registering the participants' responses. Auditory 1 kHz low-pass filtered white noise was played via headphones masking the sound produced by the tactile actuator. The supervisor of the test was sitting next to the participant in the same room. Figure 3 presents the test setup.

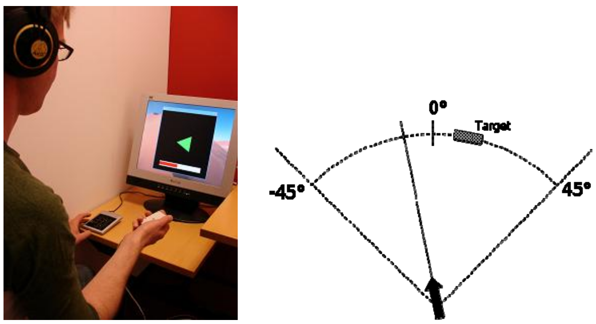

Figure 3. The test setup. Left: the participant held the sensor-actuator device in her/his dominant hand provided the responses by a keypad with the other hand. The arrow on the screen indicated the direction of the next target. It was visible for 500ms after a successful finding of a target. Right: the targets were rendered within -45° to 45° space with respect to the seating position midpoint.

One test session consisted of a sequence of consecutive targets which were randomly distributed in the space ranging from -45 to 45 degrees with respect to the midpoint defined by the sitting position. The target positions were 0°, ±15°, ±30°, ±45° and target widths were 5°, 10°, 15°, 20° and 25°. The positions and distances were measured from the midpoint of each target. This yielded 60 combinations of consecutive target distances and target widths including both left and right movements. The distances between successive targets and the targets widths were randomized. The targets were always on different side with respect to the previous target but not necessary on different side of the midpoint. Each distance/width -pair was presented once in the session.

The task of the user was to find the target as rapidly as possible by pointing the sensor-actuator device on the target and pressing the dedicated button on the keypad. The participant proceeded in the test only by being on the target when pressing the button. The successful finding was indicated by a short beep which was played through the headphones and a green triangle on the screen indicating the direction of the next target. The triangle was visible only for 500ms and its purpose was to eliminate the cases where the participant starts to move to wrong direction after finding a target. The necessity of such a visual cue was discovered in a pilot test of the study. A red progress bar on the monitor screen indicated the phase of the session.

Each participant performed both feedback conditions (with and without close-to-target cue) twice. The four sessions were arranged so that the two first sessions were for rehearsal and the two latter ones were taken into analysis. The order of the rehearsal sessions and the actual test session were mixed and balanced yielding 4 different combinations. Before the start of the test, the participants were instructed to find a comfortable sitting position and grip of the sensor-actuator device. They were familiarized with the on-the-target cue and the close-to-target cue. Also, in the beginning of each session, the participants were told the condition of the session. The duration of each session was 2-4 minutes and thus overall test took approximately 20 minutes including the instructions and breaks.

The Windows software used in the experiment collected the participants' responses and sensor data of the movement. We were mostly interested in the effects of target width and distance on the task time. The distance was the angular distance between two consecutive targets and the task time was the duration from a successful target finding to the next one.

3. RESULTS

The task time as a function of target width is presented in Figure 4. The target finding was significantly more difficult with smallest and second smallest target widths (5° and 10°) compared to the other target widths 15°, 20° and 25°. The difference between the conditions was not significant in any of the target widths.

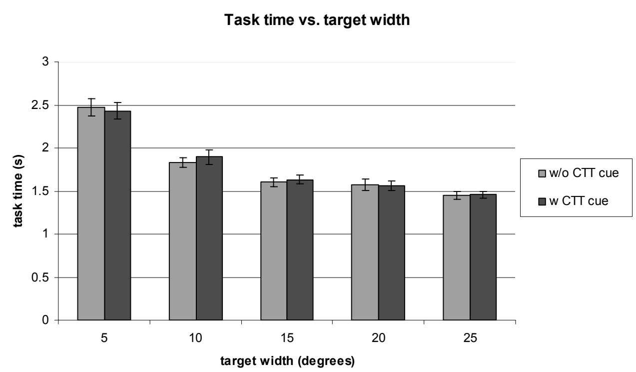

Figure 4. Task time as a function of target width in both conditions, with and without close-to-target (CTT) cue. The error bars reflect the 95% confidence interval.

Distance had significant effect on task time. Task time as a function of distance is presented in Figure 5. Finding the nearest targets, i.e. with the shortest distance, (15°) appeared to take more time than finding the next nearest targets (30°). Otherwise, the task time increased along with the distance. Also, the nearest target was slightly easier to find without the close-to-target cue (t-test $p=0.11$) and furthest target was significantly easier with the close-to-target cue (t-test $p=0.02$).

Overall average task time was 1.79 seconds without significant difference between the conditions (1.79 vs. 1.80, t-test $p>0.05$). The average scanning velocity was 45.1°/s and there was no significant difference between the conditions.

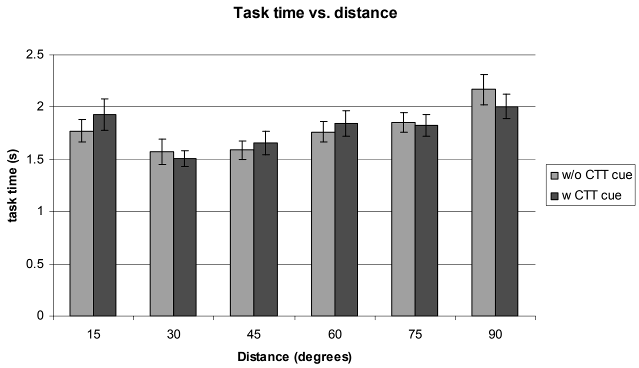

Figure 5. Task time as a function of target distance with and without close-to-target (CTT) cue. The error bars reflect the 95% confidence interval.

We also analyzed the average number of button presses per target, the on-the-target time before the button press, and the number of overshoots i.e. passes of the target before the successful finding. There was no significant difference between the conditions in any of these measures. On average, extra button presses occurred only in 7% of the tasks. But on the other hand, 62% of the extra presses were done when the target width was the smallest (5°). Slightly more extra presses were detected in condition with the close-to-target cue (30 vs. 38, t-test p=0.23). There were no significant differences in the on-the-target times between the conditions or target widths. However, the target width influenced significantly in the number of overshoots in the target finding. The smallest target width had significantly more overshoots than the other target widths. The average number of overshoots by target width is presented in Figure 6.

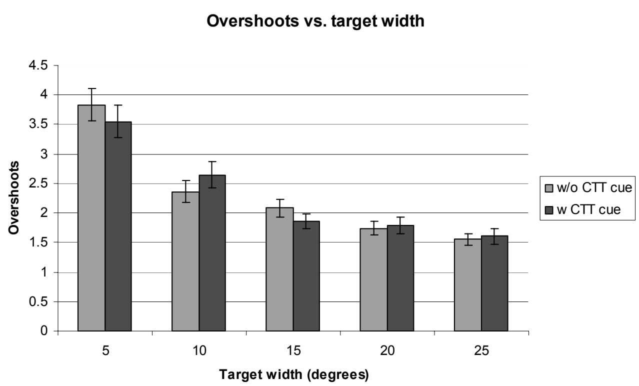

Figure 6. Average number of overshoots by target width with and without close-to-target (CTT) cue including the 95% confidence interval.

The data of the experiment outcome was also analyzed with respect to the model of the Fitts' law (Equation 1). In both of the conditions, the fitting to the model yielded poor results. The Pearson correlation coefficient between the task time and index of difficulty (ID) were 0.45 and 0.36 for the conditions without and with the close-to-target cue, respectively. In Figure 7, the upper plots present the task time of the experiment data as the function of index of difficulty (ID).

Although the participants had time to familiarize themselves with the system and the task, there were lots of data items where the task time was way above the baseline. This was caused by either taking the wrong direction after a successful target finding despite the guiding triangle or passing the target without noticing it in the first place. These data items are visible as ripples above the main cluster in upper part of the Figure 7. If these potential outliers are removed, the correlation can be improved as the Fitts' model parameter estimation assumes Gaussian distribution of fitting errors. Outliers were removed by filtering out 30% of the slowest task times of the values within all the indexes of difficulty (ID). This modification increased the correlation up to 0.71 and 0.58 in conditions without and with the close-to-target cue, respectively. The better fit can be seen in the lower plots of the Figure 7.

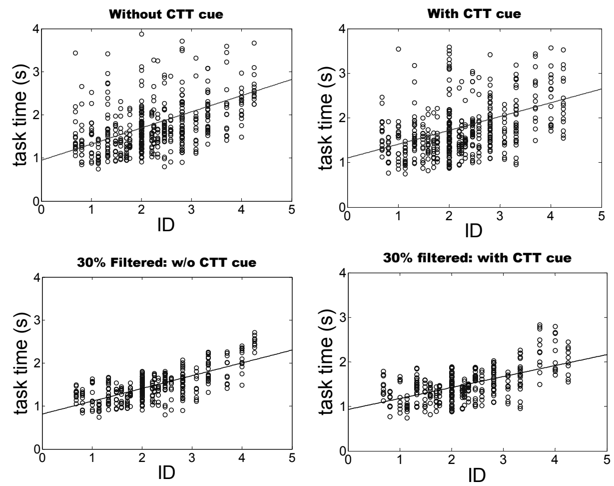

Figure 7. Task time plotted against index of difficulty with both conditions. Left: without close-to-target cue and right: with close-to-target cue. The upper figures represent all of the data while the lower figures illustrate the data where 30% of the least representative items have been filtered out.

Target finding strategy varied between the participants. Some participants selected faster scanning speed which caused more overshooting with the smaller targets. Figure 8 represents two different strategies of the same targets. The illustration displays pointing trace of 6 consecutive targets. In some of the cases, the initial scanning direction appeared to be wrong which caused a change in direction soon after the start. Also, passing the target with too high speed caused some delay due to direction change.

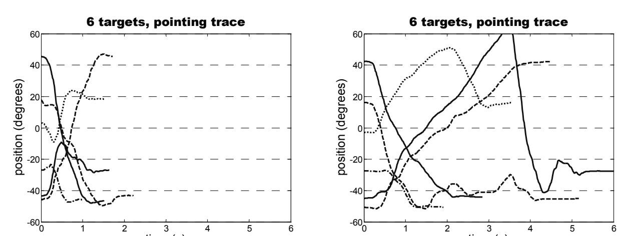

Figure 8. Pointing traces of two different strategies in target finding. The plots are of the same targets but of different participant and condition. The vertical axis is the pointing direction and horizontal axis represents time.

4. DISCUSSION

Our study on tactile target rendering appeared to lead into rather different results than the corresponding studies with visual feedback. Although the target width and distance to the target had significant effect on the task time, the data correlated poorly with the commonly used Fitts' model. This can be explained to some extent with the relationship between the average scanning velocity and the width and distance of the targets. When the participants obtained a regular pace in the target finding task, the closest targets were often passed with such a high speed that the participants were not able to react to them within the target width. Participants were then forced to change the direction and search the target again. This was most likely to happen with smaller targets. This phenomenon is clearly visible in Figure 5, where the task time is presented against the target distance. However, if only 70% of the most representative data were taken into the analysis, the correlation between the index of difficulty and task time yields more significant correlation and thus better fit to the Fitts' model. The filtering basically removed the potential outliers from the data i.e. the cases where the participants have headed accidentally to the wrong initial direction or when the target has been passed several times with too high velocity. This filtering was justified because the experiment setup was unfamiliar to the participants and caused the data not to be normally distributed. We could claim that reasonable amount of training would have removed the outliers from the data, yielded normally distributed data, and provided better fit to the Fitts' model.

The participants' performance was clearly influenced by the search strategy, i.e. the scanning velocity that they adopted during the four experiment sessions. They may have adopted a strategy where the scanning speed is maximized at the expense of overshooting and extra button presses. This strategy reminds a shooting scenario where the target is moving e.g. duck shooting. The shooter pulls the trigger as soon as possible when the sight and the target are overlapping. Alternatively, participants may have minimized the number of extra button presses and very carefully ensure that the pointing is on the target. The illustration of pointing traces in Figure 8 follows this division. The left part of the figure represents the fast shooting strategy whereas the right corresponds to the careful strategy. The existence of these strategies is also supported when analyzing the correlation between the scanning speed and extra button presses. Although the comparison does not yield high correlation (0.53) the participant with the highest speed produced the most of the extra button presses whereas the slowest speed participant produced the least. These strategies remind of the target searching behaviors presented by Robinson et al. [3]. They observed that most of the participants in their tactile target finding task obtained either a 'directly to the target' or 'probing around the target' -behavior.

Overall, the close-to-target cue did not show significant improvement in the task time. The only significant influence was detected when the inter-target distance was the longest (90°). The close-to-target cue had nearly significant negative effect on task time with the nearest targets. This finding suggests that the role of the close-to-target cue was somewhat confusing. It produced slightly more extra button presses, which hints that it was not completely obvious whether the pointing was on the target or near the target. The general explanation for these findings could be the difficulty to discriminate of the target from the surroundings. If the target is presented without the close-to-target cue, the edges are more noticeable and thus the targets are easier to detect. The finding reminds of a scenario where elevated targets are manually explored on a plane surface. The targets can be found more easily if they are not surrounded by any additional elevations since the edges of the target are more distinguishable. The confusing effect of close-to-target cue challenges also the findings of Oron-Gilad et al. [17]. In their study about tactile guidance, there was no experiment condition where close-to-target cue was not present at all. However, the positive impact in finding the furthest targets suggests that more careful design of the close-to-target cue could improve also finding of the closer targets. The close-to-target cue selected for this study was a result of very short piloting.

The overall target finding time was 1.8 seconds which is encouraging with respect to the user scenario that was described in the beginning of the paper. The pointing and scanning must be a short phase in the presented scenario to obtain benefit over looking at the device screen. The findings of this study suggest that even relatively narrow targets, of 5 or 10 degrees, can be found robustly within a short period of time (<2.5s). These results could be taken into account when designing haptically enhanced navigation systems [2, 3, 4] or spatially aware social networking applications [1].

5. CONCLUSIONS AND FURTHER WORK

We presented a study where augmented reality targets were presented to the user by tactile feedback. The width and distance of the target had significant effect on the task time and significantly more probing around the target occurred in the case of the smallest target width. With the most representative fraction of the data, the task time did correlate with the index of difficulty (ID) and provided rather good fit to the Fitts' model. This finding indicates that the interaction with tactile virtual targets to some extent follows the same principles as that of visual and auditory targets and tactile rendering is a considerable option for indicating the directions in augmented space.

However, our findings also hint that Fitts' law is not adequate model for this kind of interaction, because it does not take into account the strategy in finding the targets. Most probably the selected strategy is related to the scale of the target distances. A new model with more emphasis on the trade-off between the task time and missed targets could be developed. This would presume more extensive experimentation with more participants and larger scale of target widths and distances. The model could give an explanation to the non-linear relationship between the target distance and task time in our data (Figure 5) and provide insight on what are the effects of the individual differences in the strategy selection.

Further studies could take the design of the close-to-target cue into more detailed analysis. In this study, the cue was just a smooth vibration increasing towards the border of the target. The target itself was rendered as a rough and strong vibration. The roles of these cues could be exchanged to make participants to pay more attention to the close-to-target cue. A rough vibration could indicate more clearly the proximity of the target because the frequency of a rough vibration can be perceived more clearly. Also, the effect of distance of the close-to-target cue could be studied more carefully. In this study, the cue was always of the same width, 10 degrees, and obviously assisted only the finding of the furthest targets. The width of the close-to-target cue could correspond to the window size in the peephole pointing studies [13, 14, 15] and thus more complex models of the interaction could be applied.

The target finding based on tactile cue could be more challenging in real mobile contexts. The further studies should clarify the usefulness of the tactile rendering in the navigation scenario when used in outdoor contexts and integrated into real mobile applications. These studies will be obliged to take into account the noises and interferences of the environment as well as the limitations of the orientation tracking and tactile rendering in mobile situations.

6. REFERENCES

[1] Kahl, G., Wasinger, R., Schwartz, T. and Spassova, L., Three Output Planning Strategies for Use in Context-aware Computing Scenarios, Proc. of the AISB Symposium on Multimodal Output Generation (MOG), Aberdeen, Scotland, UK, 2008, pp. 46-49.

[2] Wasinger, R., Stahl, C., and Kruger, A., Robust Speech Interaction in a Mobile Environment through the use of Multiple and Different Media Input Types, in Proceedings of the 8th European Conference on Speech Communication and Technology (Eurospeech), pp. 1049–1052, (2003).

[3] Marentakis, G. and Brewster, S.A. Gesture Interaction with Spatial Audio Displays: Effects of Target Size and Inter-Target Separation. In Proceedings of ICAD2005 (Limerick, Ireland), July 2005. ICAD, pp77-84.

[4] Marentakis, G.N. and Brewster, S.A. Effects of Feedback, Mobility and Index of Difficulty on Deictic Spatial Audio Target Acquisition in the Horizontal Plane. In Proceedings of ACM CHI 2006 (Montreal, Canada), ACM Press Addison-Wesley, pp 359-368

[5] Strachan, S. and Murray-Smith, R., GeoPoke: Rotational Mechanical Systems Metaphor for Embodied Geosocial Interaction, NordiCHI 2008: Using Bridges, 18-22 October, Lund, Sweden, 2008.

[6] Strachan, S., Williamson, J. and Murray-Smith, R., Show me the way to Monte Carlo: density-based trajectory navigation, Proceedings of ACM SIG CHI Conference, San Jose, 2007, pp. 1245-1248

[7] Williamson, J., Strachan, S. and Murray-Smith, R., It’s a Long Way to Monte-Carlo: Probabilistic GPS Navigation, Proceedings of Mobile HCI 2006, Helsinki, 2006.

[8] Robinson, S., Eslambolchilar, P. and Jones, M., Evaluating Haptics for Information Discovery While Walking, to appear in Proceedings of BCS HCI 2009, Cambridge, UK, September 2009.

[9] Pesqueux, L. and Rouaud, M. "Vibration level of mobile phones' silent alerts," in Department of Acoustics. Aalborg: Aalborg University, 2005.

[10] http://www.eaiinfo.com/Tactor%20Products.htm, 29.5.2009

[11] Laurila, K., Pylvänäinen, T., Silanto, S., and Virolainen, A. "Wireless Motion Bands", position paper at UbiComp’05 Workshop on "Ubiquitous computing to support monitoring, measuring and motivating exercise", Tokyo, Japan, September 11-14, 2005

[12] http://www.inition.co.uk/inition/pdf/ymocap_XSens_mt9.pdf, 29.5.2009.

[13] Ahmaniemi, T., Lantz, V. and Marila, J.: Dynamic Audiotactile Feedback in Gesture Interaction. In Proceedings of the Mobile HCI 2008 September 2-5, 2008, Amsterdam, Netherlands. pp. 339-342.

[14] Ahmaniemi, T., Lantz, V. and Marila, J.: Perception of Dynamic Audiotactile Feedback to Gesture Input. Proceedings of the 10th International Conference on Multimodal Interfaces, October 20-22, 2008, Chania, Crete, Greece. pp. 85-92.

[15] Akamatsu, M., MacKenzie, I. S. and Hasbrouq, T. (1995). A comparison of tactile, auditory, and visual feedback in a pointing task using a mouse-type device. Ergonomics, 38, 816-827.

[16] Tähkäpää, E. and Raisamo, R. Evaluating Tactile Feedback in Graphical User Interfaces. In proceedings of Eurohaptics (Edinburgh, UK) 2002.

[17] Oron-Gilad, T.,Downs, J.L., Gilson, R.D. and Hancock, P.A.: Vibrotactile Guidance Cues for Target Acquisition. IEEE Transactions on Systems, Man, and Cybernetics, Part C 37(5): 993-1004 (2007).

[18] Fitts, P.M., The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47(6): p. 381-391 (1954).

[19] Jagacinski, R.J. and Flach, J.M., Control Theory for Humans, Quantitative Approaches to Modeling Human Performance. Lawrence Erlbaum Associates, Inc. 2003, pp. 17-22, 75-76.

[20] Crossan, A., Williamson, J., Brewster, S.A. and Murray-Smith, R. Wrist Rotation for Interaction in Mobile Contexts. In Proceedings of MobileHCI 2008 (Amsterdam, Holland), ACM Press, pp 435-438.

[21] Cabral, M.C., Morimoto, C.H., and Zuffo, M.K. On the usability of gesture interfaces in virtual reality environments. Proceedings of the 2005 Latin American conference on Human-computer interaction. Cuernavaca, Mexico. pp. 100-108.

[22] Yee, K-P., Peephole displays: Pen interaction on spatially aware handheld computers. In Proc. CHI 2003, ACM Press (2003), 1–8.

[23] Cao, X., Li, J.J. and Balakrishnan, R., Peephole Pointing: Modeling Acquisition of Dynamically Revealed Targets, Proceeding of the 26th SIGCHI conference on Human factors in computing systems. Florence, Italy, pp. 1699-1708.

[24] Rohs, M. and Oulasvirta, A., Target Acquisition with Camera Phones when used as Magic Lenses, Proceeding of the 26th SIGCHI conference on Human factors in computing systems. Florence, Italy, pp 1409-1418.

[25] Andersen, T. H., A Simple Movement Time Model for Scrolling, CHI 2005 extended abstracts on Human factors in computing systems. Portland, OR, USA, pp. 1180-1183.