Sequential and simultaneous tactile stimulation with multiple actuators on head, neck and back for gaze cuing

Oleg Špakov, Jussi Rantala, Poika Isokoski, University of Tampere, Finland

Abstract—Interacting with the environment using mobile eye-tracking is accompanied with challenges in providing non-visual feedback related to gaze events and monitoring the gaze vector estimation quality. Recent studies point to haptic stimulation as a promising feedback channel in this context. In this work, we focused on applying haptic stimulation to inform users of pointing inaccuracies by cuing their gaze in the direction of nearby interactive objects. To decrease the cognitive load, short repetitive vibrations from four actuators were applied to the head and neck of the user. The head area stimulation was compared to the back that has often been used in earlier studies. The results showed that the haptic stimulation on the head and neck cued users as efficiently as the stimulation of the back, although smaller stimulation signal amplitude would be desirable. Another important implication refers to the design of the stimulation signal pattern: if multiple actuators are used in stimulation, then they should be activated sequentially and not simultaneously.

I. INTRODUCTION

Properly designed feedback in gaze-based applications helps users to get quick and accurate verification of the system response that their actions led to. In stationary environment (i.e., users have a screen in front of them) visual feedback is used widely for this purpose: the interactive on-screen objects change visual appearance as they get focused or selected. The location of the gaze pointer is often displayed too, as it serves as feedback of the gaze vector estimation quality, which still requires monitoring even for the most accurate eye-tracking systems. In addition, audible feedback often complements the visual feedback: for example, both feedback types can be used to confirm letter selection when typing text by gaze [10], or to warn users if gaze was lost when interacting with a computer [1].

When using wearable eye-trackers, a common issue is the missing display, and therefore to the lack of feedback for gaze-related events. A head-based device may move relative to the position it was calibrated in. Thus the need for the feedback of gaze estimation quality is even in higher than in the stationary setup. While audible feedback provided via headphones still could be used in the mobile environment, representing complex information such as the location of the gaze pointer via two audio channels may be challenging. If any coding scheme is applied to represent a variety of possible states, then this scheme will likely be rather complex and require notable efforts from users in learning it (e.g., [15]).

The lack of feedback of gaze estimation quality has been reported in earlier studies, (e.g., [11]). However, instead of representing a gaze pointer, the gaze-based system could intelligently cue users to glance in the direction where interactive objects are located. One could hypothesize that if this directional information would be available, users could apply some sort of gaze vector estimation correction when the tracking accuracy is degraded. For example, they could glance slightly away from the target in the direction that was cued, or even use head motions instead of changing gaze direction, as described by Špakov et al. [13].

Recent studies on haptic stimulation while using eye-trackers [9] showed that skin stimulation by vibrations may serve well in providing feedback for gaze-related events. Since the human haptic perception is not limited in the number of input channels as much as auditory perception (two ears only), it should be possible to find a haptic way to represent directional information without using complex encode schemes, thus escaping a high mental load when using the system. It is known that there is a cross-modal link in spatial attention between touch and vision [5] [14]. In other words, tactile stimulation applied to human body tends to guide visual attention in the direction of the stimulation. The link has been utilized earlier, for example, in studying driver assistance systems that direct the user’s attention to a part of the visual field that is critical for safety. In testing such a scenario, Ho et al. [6] showed that spatial tactile warning signals given in the direction of a dangerous event caused participants to reorient their visual attention and react significantly faster. Jones et al. [8] found that this link is at least partly automatic since participants in an experiment tended to move their gaze according to spatial tactile stimulation even when they knew that the stimulation did not help in the experimental task.

In studying the link between touch and vision, researchers have stimulated mainly the user’s torso [3] [6] [7] [14] [16]. This requires mental processing from users because they need to translate the spatial tactile stimulation locations to corresponding coordinates of the visual field before gaze can be moved. An alternative would be to merge the tactile and visual coordinate systems so that the location of spatial tactile simulation would be the same as the cued visual direction. In practice, this would require presenting the tactile stimulation to the user’s head. Wearable eye trackers could provide a suitable platform for head area haptics. The head area has been used earlier for presenting...
navigational information to deaf users [2], visually impaired users [4], and soldiers [12].

The aim of the current study was to compare guiding of overt visual attention using two different tactile stimulation sites that were the user’s back, head and neck. The back was chosen as a reference body site because it has been utilized successfully in previous research (e.g. [14]). Head was also studied as a body site to provide haptic feedback when gaze is occupied [18]. Neck is known as a body site highly sensitive to vibrations [17]. We were mainly interested in finding out whether the different tactile stimulation locations have an effect on the time it takes for a user to look in the cued visual direction. We hypothesized that this reaction time could be faster for the head condition since the stimulation locations are closer to the eyes. In addition, we explored different ways to present the directional cues using spatial tactile stimulation. We also wanted to understand how people perceive head vs. back stimulation and therefore we collected subjective ratings expressed by participants.

We present next the experimental method, procedure and design of the study, and also describe in more detail the body locations we selected to stimulate. This is followed by an analysis of the recorded data, and finally after discussion we present our conclusions.

II. METHOD AND PROCEDURE

A. Participants

We recruited twelve volunteers to participate in the experiment. All were students from the local university, aged 20-37 year (μ = 26.3), mostly (8) with vision corrected. None of the participants had participated previously in an experiment using haptics, although two participants were familiar with eye trackers. We measured their height, chest, head and neck as we planned to study correlations between individual metrics and subjective ratings as well as objective measurement.

B. Equipment

T60 eye tracker from Tobii Technologies (desktop setup, 60 Hz, 17” screen) was used to record the gaze coordinates. Participants sat 60-70cm away from the device.

For the tactile stimulation we used four C2 actuators from Engineering Acoustics, Inc (http://www.atactech.com). They were connected to two GIGA Port HD USB sound cards from ESI with built-in amplifiers, two actuators per device (left and right audio channels).

To display visual targets, emit stimulation signal, and record gaze point we developed a custom software in C# (.NET 4.5) that was running on an up-to-date PC (Intel Core i5-2520M, 8GB, MS Window 7).

C. Design

We conducted a short pilot study to test a large number of possible locations where four actuators could be attached comfortably. All tested setups could be divided into two large groups based on either a rectangle or rhombus orientation (with the actuators located at corners). In the rhombus-oriented setups the locations included spine and the hair covered area of the neck. Often the vibration was unpleasant, too dominant over other locations, or too weak to feel if hairs were long. Therefore, we rejected all rhombus-oriented setups and chose two rectangle-oriented setups, with actuators placed on the back or on the head and neck.

For the conditions with “back” location we attached all four actuators to the back of a chair as shown on Fig. 1a to apply vibrations to the participants’ back. The actuators were placed at the corners of rectangle that was 16x16 cm in size (similarly to [14]), as shown in Fig. 2a (A1-A2-A4-A3). The distance from the seat level was constant, about ~40cm to the lower actuators.

Two ordinary hair bands were used as the placement for two pairs of actuators in the conditions with “head and neck” location (see Fig. 1c). The actuators were attached on
the sides of both bands symmetrically, so that vibrations from the upper band actuators were applied to approximately temple areas that are sensitive to tactile stimulation [12], and vibrations from the lower band actuators were applied to the opposite sides on the neck, as shown in Fig. 1b. We based the design of the stimulation patterns on the assumption that users’ gaze direction is always nearly orthogonal to the plane of the rectangle with actuators located in corners. Then stimulation by vibration should naturally be perceived as occurring on the periphery of the visual field. To find the way to support this direction coding model best, we experimented in the pilot test with several direction encoding schemes (i.e. tactile stimulation parameters) which differed in vibration strength, frequency and duration. The iterative design process led to two modes of stimulation: the “parallel” and the “sequential”. Most of the other characteristics seemed equal in terms of subjective sensation.

In both modes one-actuator vibration was used to encode diagonal directions, and two-actuator vibrations were used to encode horizontal and vertical directions, as shown in Fig. 2a. A 250Hz sine wave signal was used to produce vibrations in both modes. We chose to use 250Hz because it is the resonant frequency of the C2 actuator. The whole one-time stimulation consisted of a 100ms vibration applied 3 times with a 200ms pause in between. The difference between modes was in the order of actuators activation in two-actuator condition. In the “sequential” stimulation mode one actuator was activated 150ms later than the other resulting in a 50ms interval between the successive vibrations, while in the “parallel” mode both actuators were activated simultaneously. The relative activation times of two actuators for the direction “left” in both modes are shown in Fig. 2b.

In our pilot tests we found that if two actuators are active and one of them causes a stronger feeling than the other, then a user may not notice the vibration from the weaker actuator. Similar reports can be found in literature (e.g. [19]). We hypothesized that vibrations applied in a sequence could solve that problem.

Therefore, in our comparative study we had a 2x2 factorial design in which we tested back location versus head and neck location, and the parallel stimulation mode against the sequential.

D. Procedure

Before the test started the participants signed an informed consent form and filled in a background information form (age, vision, experience, etc.). Then they were familiarized with aims of the study and the eye tracker was calibrated for their eyes (regular 5-point calibration was used). Next, the study supervisor put the hair bands on the participants with the actuators attached. The participants spent a few minutes learning what stimulations corresponded to each of the eight directions. The test continued after the participant confirmed that all stimulations were clearly felt.

Next the test procedure was introduced to the participants. At the beginning of every trial they had to look at a small red square on a grey background (home-box) that was the only visible object at this time. After gazing at the home box for one second, eight small blue squares (target boxes) appeared around it at the distance of 350 pixels, one per each direction (see Fig. 3). At the same time the home box disappeared and a vibration signal was played. The target boxes stayed visible for 2.5 seconds, and after that another trial began by displaying only the home box. The participants were instructed to make selections by glancing at a target in the appropriate direction as fast as possible.

Eighty trials (10 trials per direction) were completed for each of the four conditions. The order of the conditions was counterbalanced between the participants, except that the conditions with the same location (i.e., back or head and neck) were always completed successively. After both sessions with the same location of actuators were finished, the participants rated the tested technology in terms of general impression, pleasantness, subjective accuracy and subjective quickness of their reaction using 7-point Likert scales.

For each trial the experimental software recorded the history of targets the participant had looked at. The data analysis was performed using MS Excel and online ANOVA calculation tools.

III. RESULTS

Hereafter, we use abbreviations for condition levels: H – head and neck, B – back, p – parallel, s – sequential (e.g. Bp – parallel stimulation on the back).

A. Objective measures

The main objective measurements were the selection error and the reaction latency. The selection error was recorded when the last zone hit by gaze at the end of a trial was not the target zone. The reaction latency is the time interval between the stimulation start and a moment when
A repeated-measures ANOVA showed that the number of pointing errors made using the sequential stimulation was statistically significantly lower than when using parallel stimulation ($F_{1, 11}=12.4, p=0.005$). The effect of stimulation location and the interaction of type and location were not statistically significant. Further pairwise t-tests showed that using parallel stimulation resulted in lower error rate for both locations: 0.8% vs 2% for head and neck ($p = 0.046$), and 1% vs 4.2% ($p = 0.012$) for back. Notably, the sequential stimulation resulted in the same error level for both back and head, while the parallel stimulation resulted in about twice more errors on the back than on the head and neck.

As seen in Fig. 4 the reaction latency was longest when sequential stimulation was applied on the back (~700ms) while for other conditions it was about 80ms shorter. A two-way repeated-measures ANOVA showed that the effect of stimulation type had a statistically significant effect ($F_{1, 11}=5.0, p=0.047$). The effect of stimulation location and the interaction of type and location were not statistically significant. Further pairwise t-tests showed that $B_p$ was faster than $B_s$ ($p=0.008$). The difference between $H_p$ and $H_s$ was not statistically significant ($p=0.9$).

The analysis of these two measures applied to each tested direction separately revealed that 2 out of 3 errors were made when the target direction was left (28%) or right (39%), while in trials with diagonal directions only 9.4% of all errors were made. The strong dominance of the errors made in trials with vertical or horizontal direction targets is presented in every condition. No direction-specific peculiarities were found in the reaction latency: the $B_s$ condition was always slowest to react and other conditions were about equal regardless of the requested direction.

No strong correlations were found between participants’ individual metrics and objective measures.

B. Subjective ratings

The questionnaire scales were “General” (common sense of the technology in use), “Pleasantness”, “Quickness” and “Accuracy”. The averaged ratings are shown in Fig. 5.

The ratings appear to be rather similar when averaged. Only the differences in pairs $H_p / B_p$ and $H_p / B_q$ for quickness exceed one unit on the Likert scale. However, the variation between the participants’ ratings is high and the statistical analysis (Kruskal-Wallis test) did not reveal significant differences for any pair of ratings ($p > 0.05$).

The variation in ratings was so high, that some questionnaire characteristics received all values from the scale. This applies to the ratings of $H_p$, $B_p$ and $B_s$ for general evaluation, $H_p$ and $H_s$ for pleasantness, and $B_p$ and $B_s$ for quickness and accuracy. On the other hand, ratings of $H_p$ for quickness and $B_p$ and $B_s$ for pleasantness were much less spread.

For Fig. 6 we clustered the rating scale to “negative” (1-3), “neutral” (4-5) and “positive” (6-7) values to create a readable visualization of the cumulative ratings. The “negative” value accumulated more values from the Likert scale than others, since participants generally tend not to rate anything with the most negative values (although few such rates were recorded in this study).

The clustered scale allows faster inspection of the choices made by the participants. For example, the neutrality of rates for pleasantness when stimulating head with the sequential vibrations and back with parallel vibrations are well visible, as well as positive rates for quickness when stimulating head with the parallel vibrations. Many other ratings still appear as quite spread, while some consist of mostly neutral and positive values.

The next method of the analysis of the subjective ratings we applied is related to calculation of the cumulative differences in ratings between two conditions. This method allows deeper inspection of which condition was more preferable, leaving visible the variability of the ratings. When comparing ratings of two conditions, the rates were paired, and each pair is labelled as “first” if the first rate is greater, “equal” if rates are equal, or “second” if the second rate is greater.
Six sets of paired rates were created: “parallel versus sequential” separately for head/neck and back locations, “head/neck versus back” separately for parallel and sequential stimulations, “parallel vs. sequential” for any stimulation location, and “head/neck versus back” for any stimulation type. The comparison results are shown in Table 1. As one can see, most of the differences in ratings were opposite (labelled as “opp”): some participants rated one condition higher than another, while others – vice-versa. Another big set of comparisons covered about equally all possible values on the given scale (labelled as “var”). Only three comparisons showed clear support of the only condition by the majority of participants (labelled correspondingly): 1) the pleasantness of sequential stimulation on the back rather than on the head, 2) the quickness of parallel stimulation on the head rather than on the back, and 3) the accuracy of stimulation on the back by the sequential vibrations rather than by the parallel vibrations. Also the sequential stimulation on the head (and also in general) was never in favor compared to the parallel stimulation in terms of pleasantness (labelled as “−paral”). Similarly, the parallel stimulation on the back was never in favor compared to the sequential stimulation in terms of general evaluation (labelled as “−seq”). The sequential stimulation is perceived in general equally on the head and back (labelled as “eq”).

Finally, we searched for correlations between subjective ratings and participants’ individual metrics. Only strong correlations (>0.7) were taken into account. The following correlation were found: 1) neck size correlated positively with preferring Hs over Bs in general evaluation, 2) head and neck sizes correlated positively with Bp when evaluating own quickness and accuracy, 3) height was negatively correlated with Hs when evaluating own quickness and accuracy.

IV. DISCUSSION

The objective measurements have confirmed our hypothesis that sequential stimulation results in fewer errors than the parallel stimulation, especially if the stimulation covers distinct body locations on the vertical scale. As human’s body has symmetry in the horizontal plane, the sensibility to vibrations changes in vertical plane mostly, and when actuators used for stimulation are located on different vertical levels, one can dominate over another making an effect that only one actuator is vibrating at the moment.

The participants correctly noticed that they made more errors when parallel stimulation was applied on the back rather than the sequential stimulation, but their subjective feeling that the parallel stimulation on the back caused slower reaction than on the head was not correct. Actually, only sequential stimulation on the back caused slower reaction than in all other conditions, but we do not have a clear explanation for why this was so.

Despite that the measurements revealed the advantage of sequential stimulation over the parallel one, most of the participants noted that the sequential stimulation was less pleasant than the parallel stimulation, especially if applied on the head. Thus, we can reject the hypothesis that the sequential stimulation could be preferred over parallel stimulation by the users. One possible explanation for this is that the vibrations were perceived as stronger in the temple areas as compared to the other locations. It is possible that vibrations of smaller amplitude could decrease or remove completely the differences in pleasantness compared to other location and stimulation modes, while preserving the performance advantages that this condition holds.

The analysis of differences in subjective ratings revealed that the participants usually provided opposite opinions to the superiority of one condition over another, rather than supporting only one point of view. The wide range of ratings collected leads to several interpretations. Firstly, the duration of the experiment may have been too short. If opinions change over time, the questionnaires may have been answered before all had formed a final opinion.
Secondly, the same stimulation condition could affect the participants differently due to the variability in size of their bodies and heads. Thirdly, people naturally have different levels of haptic sensitivity in various location, thus some variability in subjective evaluations is likely to show up for any stimulation.

The correlations found between subjective ratings and participants’ individual metrics were hard to explain: we see no reason why participants with thick necks could prefer sequential stimulation on the head rather than on their back, or why this kind of stimulation could feel slower and less accurate for tall participants. We lean towards treating these two correlations as products of random variability (and, probably, to the low number of participants) rather than two correlations as products of random variability (and, thus, not accurate for tall participants. We lean towards treating these two correlations as products of random variability (and, probably, to the low number of participants) rather than two correlations as products of random variability (and, thus, not accurate for tall participants.

V. CONCLUSIONS

We have tested a method of cuing users to glance into a certain direction with the help of four actuators producing short vibrations. We experimented with two actuators locations and the way they stimulate skin when more than one actuator is used at a time to encode a visual direction. A gaze tracker was utilized to evaluate how the different stimulation parameters affected gaze movement.

The major finding is that sequential stimulation is more robust to use in two-actuator stimulation. There is a small penalty in time (80ms) if compared to the parallel stimulation which results in more recognition errors, but only when applying vibration on the back. The sequential stimulation on the head and neck work as well as on the back, but because of constant signal amplitudes used in this study it feels less pleasant in the head and neck. However, the variability of participant ratings cannot conclude that certain stimulation locations and modes were preferable over other. Thus, we conclude that 1) using head and neck stimulation for cuing to glance in one of eight direction is as good as using back stimulation (especially if the vibration on the head will be set to somewhat weaker intensity level), and 2) the sequential stimulation is preferable over the parallel when using several actuators.

As we learned already from the pilot study, the variety of possible locations and vibration characteristics used for stimulation may be rather large and may hide combinations that will appear more comfortable and easier-to-recall. We plan to run more comparative studies of this kind in future, exploring the efficiency of rhombus-oriented encoding model and utilizing vibrations with other temporal characteristics. For example, it would be possible to utilize tactile apparent motion [20] with sequential stimulation. Instead of having to localize separate actuators, users could perceive illusory motion between the physical locations. This could make cuing of gaze more intuitive.