Bilateral Vibrotactile Feedback Patterns for Accurate Lateralization in Hearing Instrument Body Area Networks

B. Tessendorf, D. Roggen, M. Spuhler, T. Stiefmeier 
G. Tröster 
Wearable Computing Lab., ETH Zurich, Gloriastr. 35 8092 Zurich, Switzerland (lastname)@ife.ee.ethz.ch

T. Grämer 
Laboratory for Experimental Audiology 
University Hospital Frauenklinikstr. 24 8091 Zurich, Switzerland Tobias.Graemer@usz.ch

P. Derleth, M. Feilner 
Phonak AG 
Laubisrütistrasse 28 8712 Stäfa, Switzerland firstname.lastname@phonak.com

ABSTRACT
Hearing Instruments (HIs) have emerged as true body area networks, so called HI-BANs. Besides streaming audio data they connect wirelessly to accessories such as remote controls and Bluetooth devices. Multimodal sensor data from a HI-BAN is a way to adapt the HI behavior to the user’s current hearing situation. As a potential future HI-BAN component we investigate bilateral vibrotactile feedback to support localization of sound sources. As a foundation for integrating vibrotactile cues we investigate which kind of feedback and vibration patterns are most suitable. We implemented two approaches for encoding lateral target angles: Continuous Guidance Feedback (CGF) and 6 variants with evolving complexity of Quantized Absolute Heading (QAH). In a user study with 16 normal hearing participants (7 m, 9 f, age 23–61) we evaluate lateralization error and user response time. For QAH results show a trade off between the minimal quantization error due to the encoding and the number of user errors due to misinterpretation of presented patterns. Moreover, results show a trade off between response time and minimum lateralization error. Choosing the most suitable bilateral vibrotactile encoding schemes is application specific: For QAH a minimal average lateralization error of 27° (σ = 22°) was achieved with eight 45°-segments and an average user response time of 1600 ms (σ = 545 ms). A minimal average user response time of 900 ms (σ = 325 ms) was achieved with four 45°-segments and an average lateralization error of 43° (σ = 29°). CGF guides the user within a given tolerance margin to the target at the cost of higher response time. We find that for complex encoding schemes the overall performance is person-specific.

Categories and Subject Descriptors
H.5.2 [User Interfaces]: Haptic I/O

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. BODYNETS 2011, November 07-08, Beijing, People’s Republic of China Copyright © 2012 ICST 978-1-936968-29-9 DOI 10.4108/icst.bodysnets.2011.246910

1. INTRODUCTION
Recent studies have shown that hearing impairment is increasingly affecting people worldwide, especially in developed countries [6,13]. In particular, localization of sound sources is an essential ability of the human auditory perception. It enables the listener to turn towards an interesting sound source to maximize intelligibility or to identify approaching vehicles to be safe in traffic situations. It is challenging for listeners with a strong asymmetric hearing loss to localize sound sources in the horizontal plane, also referred to as lateralization. The large group of affected people includes unilateral deaf, unilaterally implanted cochlear implants (CIs), and also listeners with bilaterally implanted CIs as well as listeners with a central hearing loss high up in the auditory pathway. Moreover, the human lateralization relies on cues that can end up distorted due to signal processing in HIs. This also leads to decreased lateralization performance [8].

People that experience difficulties with lateralization often wear HIs or CIs and these devices are in the progress of becoming true multimodal interaction interfaces. Therefore, we investigate how to address lateralization difficulties by integrating alternate feedback modalities into HIs or CIs. HIs communicate with a variety of other accessories such as remote controls, Bluetooth, or FM devices as well as the user’s smart phone to form wireless networks, so-called hearing instrument body area networks (HI-BANs) [2]. The rising trend to higher-end HIs motivates and supports our investigation of additional sensor and actuator modalities for HIs that may eventually be included within the HI itself. In particular, the integration of vibrotactile feedback into HIs or CIs is a promising approach to support the hearing impaired with lateralization or at least to provide the side of the most prominent sound source around.

To address the question "how accurate and fast can listeners localise target angles in the lateral plane through bilateral vibrotactile feedback at the ears?" we designed a continuous guidance feedback and 6 bilateral vibrotactile encoding schemes with evolving complexity optimized for encoding the lateral target angle in the complete 360°-range using bilateral vibrotactile feedback behind the ears.

General Terms
Hearing Instrument, Vibrotactile Feedback, User Study
2. RELATED WORK

A pair of glasses enhanced with 4 vibrators and 3 microphones with the purpose to locate sound sources for visually and hearing impaired people was presented in [3]. The approach doesn’t focus on integration into HI-BANs but the user is required to wear special goggles.

Research regarding tactile sensitivity measure for the head is presented in [9]. They found frequencies around 32 Hz to be optimal for perception and concluded that the part around the ears are one of the most sensitive head regions for vibrotactile stimulation. The region behind the ears at the mastoid bone is one of the most sensitive head regions for vibrotactile stimulation [9].

In [18] the authors present a device to be placed inside the ear channel to transduce sound intro vibration. The device has two active states with two different vibration intensity levels: one for high and one for low frequency sound.

A pedestrian navigation system using vibration of mobile phones is shown in [11]. The running direction is encoded in the length of vibration and the distance in the pause time between two pulses. A further navigation system is presented in [14]. They use a vibrotactile waist belt to provide vibrotactile feedback at 8 locations to allow hands-free navigation in unfamiliar places for blind people. A direction is represented by a one-to-one mapping of the respective tactor on the waist belt facing this direction.

Audio feedback can be used, as e.g. in [7] where the authors propose a portable, self-contained system that allows visually impaired people to travel through an unfamiliar environment. The information about the navigation is conveyed by earphones to the user. The disadvantage of this overlaying approach is potential interference of the computer voice with the actual real life sound.

A cell phone based approach is presented in [1]. The authors investigate to use one vibrator in a mobile-phone device and let the user use it as a vibrating pointing device, e.g. to find a friend in a crowd. Their approach could be adapted to be based on head movement instead of pointing with a mobile phone to increase intuitiveness and remove the requirement that the user must hold a device in the hand to benefit.

In [15] the authors propose to integrate sensors for body and eye movement into HI-BANs as they were found to improve HI performance, automatic hearing program selection in particular. As this kind of HI-BAN functionality improves the context-awareness of HIs it can help to decide when providing lateralization feedback to the user is adequate.

3. BILATERAL VIBROTACTILE ENCODING OF LATERAL TARGET ANGLES

3.1 Experimental Setup

We designed a bilateral vibrotactile feedback system for lateralization to investigate which kind of feedback and vibration patterns are optimal to provide support for lateralization [16]. We evaluated a wide range of vibrators and selected coin-shaped vibrators with a diameter of 10 mm (310-105 Vibration Motor 2.7 mm Button Type from Precision Microdrives) that have been developed for vibrotactile feedback in handheld applications. We emulate the integration of vibrotactile feedback into HIs by enhancing a pair of goggles with vibrators at the sidepieces as depicted in Figure 2. Instead of glasses we could also enhance headphones as we use them just as an unobtrusive mechanism to press the vibration motors on the mastoid bone to effectively conduct vibrations. We control and power the vibration motors with an USB Bit Whacker [12] module combined with a motor driver chip (TI ULN2803A). Via a serial interface of a desktop computer we use pulse width modulation (PWM) to flexibly control the vibrators’ intensity to implement a variety of different vibration patterns. To track the user’s head angle relative to the target we attached an inertial measurement unit (Xsens MT9 IMU, accuracy ±0.5°) on a baseball cap as depicted in Figure 2 to track the user’s head yaw angle.

3.2 Quantized Absolute Heading

We define vibrotactile encoding schemes as sets of patterns to encode lateral target angles. Based on existing guidelines [4,5,10,17] we designed a set of bilateral vibrotactile encoding schemes specifically suitable for lateralization of target angles in the complete 360°-range. Table 1 gives an overview of the bilateral vibrotactile encoding schemes designed for encoding lateralization information. Considering a potential integration into HIs we focus on bilateral vibrotactile stimulation, i.e. presenting vibrotactile cues behind each user’s ear. We further focus on intuitive vibrotactile encoding schemes that can be learned within a few minutes. Our proposed patterns range from basic encodings with low angular resolution to more complex ones with finer angular resolution. Figure 1 illustrates encoding schemes S1–S6. We divide the lateral 360°-range into four base segments: front (red), left (green), right (purple) and back (black), each with a size of 90°. The schemes S2–S4 consecutively extend S1. They are composed of various combinations of the 4 base

<table>
<thead>
<tr>
<th>Name</th>
<th>#Segments</th>
<th>Segment Size</th>
<th>#Intensities</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1–S4</td>
<td>{4,6,8,12}</td>
<td>{90°,60°,45°,40°}</td>
<td>1</td>
</tr>
<tr>
<td>S5–S6</td>
<td>{8,12}</td>
<td>{45°,30°}</td>
<td>3</td>
</tr>
<tr>
<td>CGF</td>
<td>continuous</td>
<td>–</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 1: Overview of bilateral vibrotactile feedback approaches delivered behind the ear for encoding lateralization information. For QAH (S1–S6) and CGF the number of vibration intensities and segments with their corresponding sizes are given.
elements and represent a refinement of the preceding ones. The segment size is reduced to 60°, 45° and 30°, respectively. This evolving procedure, by means of composing more complex tactile messages of well-known meaningful component was proposed in [17]. It makes learning the vibrotactile encoding schemes easier and more intuitive for the user. As disclosed in [4] it is possible to encode information using different vibration intensities. This has been considered in another encoding approach illustrated in S5 and S6. In [17] the authors recommend a maximum number of four different intensity levels. We used up to three vibration intensities for schemes S5–S6 considering that the vibration motors are not directly attached on the skin but on the sidepieces of the glasses. Schemes S5 and S6 are variants of schemes S3 and S4, introducing three different vibration intensities: low (PWM duty cycle of 0.1), medium (PWM duty cycle of 0.5) and high (PWM duty cycle of 1). The low intensities are used for target angles in front of the user. The high intensities are used to emphasize vibrotactile feedback for sources coming from the back of the user that usually appear more unexpectedly.

3.3 Continuous Guidance Feedback

Schemes S1–S6 that inform the user about the target location. In contrast, continuous guidance feedback (CGF) guides the user until the user has successfully turned the head to the target. The user’s head yaw angle is measured with an inertial measurement unit (IMU) to track the current angle to the target. To eliminate the need of the head tracker for CGF in later III application, the tracking could be performed using sound localization with a microphone array. CGF tells the user to turn the head to the target by vibrating on the corresponding side of the target. This vibrotactile feedback is provided continuously based on the current angle to the target. The encoding is performed with two intensity levels as illustrated in Figure 3: If the target is near, meaning in the range of 20°, the vibration intensity is halved. The vibrotactile feedback stops when the user has moved the head to the target within a tolerance of 5°.

4. USER STUDY

We recruited 16 participants (7 males, 9 females, age 23–61). In this empirical study we recruited normal hearing subjects. In our discussions with a hearing aid manufacturer and a university hospital in charge of hearing impairment pa-
Participants were instructed to localize as fast and accurate as possible. We measured the performance for each participant and on average over all participants. For the CGF method we define the user response time as the time starting from the initiation of the vibrotactile feedback until the mouse click in the GUI (for CGF until the user reaches the target angle). This includes the motor start delay (we neglect it because it is in the range of 60 ms according to data sheets), the pattern duration, the user response time and the time to operate the GUI.

User Response Time We measure the user response time starting from the initiation of the vibrotactile feedback until the mouse click in the GUI (for CGF until the user reached the target angle by moving the head). This includes the motor start delay (we neglect it because it is in the range of 60 ms according to data sheets), the pattern duration, the user response time and the time to operate the GUI. To consider the time to operate the GUI, we also measured user response time based on visual stimuli only. No direct user feedback is required for the CGF approach.

5. PERFORMANCE EVALUATION

We evaluated vibrotactile encoding schemes considering lateralization performance and user response time.

Lateralization Angle Error Figure 5 illustrates the flow of information: The target angle is encoded into a vibrotactile pattern and presented to the user. The user interprets the vibrotactile pattern and indicates the perceived angle. The resulting error is influenced by both the vibrotactile encoding due to quantization and user errors due to misinterpretation of patterns. We investigate both influencing factors separately as well as the resulting combined influence on the overall performance. As an error measure for the angular deviation we use root mean square (RMS) to consider greater deviation with greater error values. The performance for CGF is intrinsically guaranteed to the given tolerance of 5° because the feedback continues until the target is found. Therefore also no user errors occur with this scheme. Figure 6 shows the user error due to misinterpretation of the presented pattern for each participant as share of the overall number of presented stimuli. Figure 7 shows user response time for each participant and on average over all participants. For the CGF method we define the average duration of the pattern as user response time in a wider sense. The user needs more time to locate far targets than near targets. On average the users needed a time of 902 ms ($\sigma = 161$ ms) to operate the GUI with visual stimuli only as described in section 4 as shown with the bars for the case ‘visual’ in Figure 7. Figure 8 shows the RMS lateralization error for each participant and on average over all participants.

This work was carried out in collaboration with a hearing instrument company and an university hospital. In a planned extended study we will investigate combined acoustic and vibrotactile feedback to further evaluate the benefit of the system for hearing impaired listeners.
6. USER QUESTIONNAIRE

The participants assessed 24 statements about intuitiveness of the different encoding schemes, the source of error for confusing patterns and the comfort of the vibrotactile glasses. A continuous rating scale from 0 (‘not true’) to 10 (‘very true’) has been used and the participants had to put a mark on the scale, depending how true the statement was. At the end of the questionnaire the study participants had the opportunity to answer some open questions about improvements and general comments. As a result of the rating scale the CGF approach was evaluated as the most intuitive scheme, before the patterns S1–S4. The schemes S5 and S6 that featured three different vibration intensities were rated as the least intuitive ones. Most participants stated to have remembered the coding schemes during the test phase after a few minutes of memorizing. As the main cause of error, the participants answered that some vibration patterns – especially the three intensities in S5 and S6, i.e. right front, right and right back – were hard to distinguish. But overall the intensity of the vibration has been felt as appropriate and comfortable. The feedback we got concerning the comfort of the vibrating glasses was without exception positive: All of the participants stated not to feel tense, physically different or strange wearing the device and that the vibration wasn’t perceived as painful or disturbing even after one hour of wearing it.

7. DISCUSSION

7.1 Lateralization Performance and User response time

Results show a trade off between user response time and minimum angular lateralization error. Choosing the most suitable bilateral vibrotactile encoding schemes depends on the requirements of the application:

- To minimize the RMS lateralization error, scheme S3 is the most suitable scheme out of schemes S1–S5. It encodes 8 segments of size 45° and achieves an average RMS lateralization error of $27^{\circ} (\sigma = 22^\circ)$ an average user response time of 1600 ms ($\sigma = 545\, \text{ms}$). For participants P11, P13 and P15 scheme S4 that encodes 12 segments of size 30°, and for participants P8, P10 and P14 one of the schemes S5 or S6 that encodes three vibration intensities show the lowest error. With the CGF approach maximum lateralization performance was achieved by guiding the user to the given 5° tolerance margin to the target. However, CGF shows an increased average user response time of $3140\, \text{ms} (\sigma = 2468\, \text{ms})$ and a four times higher average energy budget than scheme S3.

- To minimize response time, e.g. alerting in emergency situations like an approaching car from behind, vibrotactile encoding scheme S1 is most suitable. It encodes 4 segments of size 90° and achieves an average user response time of $900\, \text{ms} (\sigma = 325\, \text{ms})$ with an average RMS lateralization error of $43^\circ (\sigma = 29^\circ)$.

The lateralization performance required depends on the application: E.g. for emergency situations, it might be desired to inform the user quickly and reduce the risk of potential harm is coming from the sound source. For different applications the goal might be a maximization of lateral accuracy, e.g. to find targets that are far way best possible. The resulting RMS lateralization error is influenced by both the vibrotactile encoding due to quantization and user errors due to misinterpretation of patterns as illustrated in Figure 5. The influence of the two error sources are shown with Figures 6 and 8. An example to illustrate the influence the quantization effects is the case of participant P9 who made almost no user errors. The corresponding RMS lateralization errors still show a lower bound due to quantization during vibrotactile encoding. This is due to the quantization of the lateral 360°-range into segments and the fact, that target angles do occur continuously inside the segments.

For QAH encoding schemes results show a trade off between the minimal quantization error due to vibrotactile encoding and the average number of user errors due to misinterpretation of presented patterns. The simpler the schemes are concerning the number of encoded segments, the lower are the user error and response time, but the higher the induced quantization error due to vibrotactile encoding. On the other hand, reducing the quantization error with encodings of smaller segments induced a higher number of user errors and user response time. The smaller the segments, the less impact has the type of user error that confuses an adjacent segment on the resulting RMS lateralization error. For the complex encoding schemes the overall performance of the vibrotactile feedback for lateralization is person-specific. E.g., participant P9 would maybe cope with more complex schemes, however other participants didn’t show as low user error rates for all encoding schemes. The standard deviation for the depicted in Figure 6 as errorbars for the average case is an indicator to characterize how person-specific the performance turns out for each encoding scheme. Figure 6 shows the relative number of times the user confused a pattern. The values are not directly comparable between schemes, because the difficulty to recognize patterns is different for the different vibrotactile encoding schemes. The outlier for participant P14 and scheme S4 in Figure 6 was explained by the corresponding participant by not having remembered the scheme.

Figure 7 shows the time for the baseline response time based on visual stimuli online. The standard deviation when averaging over all participants seems sufficiently low to skip the measurement of the baseline in further experiments and to assume a common constant offset for all participants. Average response times for scheme S1 are in the range of and for some participants even below the baseline. This confirms previous research that the haptic sense can lead to quicker response times than for the visual baseline. CGF shows highest response time, because it does not inform the user, but provides continuous feedback until the target is found. The time for lateralization increases with the initial distance of the target. To compare schemes S1–S6 and CGF a main difference needs to be considered: For schemes S1–S6 the user is just informed about the target location and still needs to move the head into the correct direction if desired.
Figure 6: User error rate: The share of user errors due to misinterpretation of patterns for each participant. Also average values and the standard deviation across all participants are given. All confusions are counted as one error. For CGF this measure is not applicable.

Figure 7: User response time for each participant and standard deviation. Also average values and the standard deviation across all participants are given. Also the baseline based on visual stimuli is showing. For CGF the response time correspond to the time needed to find the target.

Figure 8: RMS lateralization error: This measure including the quantization effect due to vibrotactile encoding where applicable and the effect of the user error due to misinterpretation of patterns. Also average values and the standard deviation across all participants are given. For CGF this measure is not applicable as it guarantees error values below a given tolerance margin.
7.2 Limitations

We used glasses to attach the vibration motors to the user. This provides good contact to the mastoid bone to conduct vibrations. In a later HI application the fitting might be not that tight requiring higher vibration intensities.

Moreover, we focused on the feedback component of the overall system. We assume the feasibility of calculating the target angle of relevant sound sources occurring in daily life situations. To assess the relevance of sound sources, context information needs to be considered, e.g. by context recognition based on HI-BANs.

To show the effective benefit of the system, a larger user study with hearing impaired participants that rate the complete processing chain in mobile daily life settings would be required. However, as a first step our focus in this explorative work was to systematically evaluate vibrotactile encoding schemes for lateralization to assess the potential benefit.

8. CONCLUSION AND OUTLOOK

Our results encourage to use bilateral vibrotactile feedback to support localization of sound sources, especially for users of HIs and cochlear implants that have difficulties with lateralization. Results show a trade off between user response time and minimum angular lateralization error. Choosing the most suitable bilateral vibrotactile encoding schemes depends on the requirements of the application: Bilateral vibrotactile encoding can be optimized for low RMS lateralization error or short response time while considering the available energy budget. For QAH encoding schemes results also show a trade off between the minimal quantization error due to vibrotactile encoding and the average number of user errors due to misinterpretation of presented patterns. The simpler the schemes are concerning the number of encoded segments, the lower are the user error and response time, but the higher the induced quantization error due to vibrotactile encoding. For the complex encoding schemes the overall performance of the vibrotactile feedback for lateralization is person-specific. The presented results are applicable to other domains of human computer interaction such as navigation or gaming.

We plan to deploy the system in a study with hearing impaired people to evaluate the benefit of the system with combined acoustic and vibrotactile feedback.

Acknowledgements

This work was part funded by CTI project 10698.1 PFLS-LS "Context Recognition for Hearing Instruments Using Additional Sensor Modalities". The authors gratefully thank all participants of the experiment.

9. REFERENCES

[12] B. Schmalz. USB Bit Whacker - An inexpensive, simple input/output device to connect your computer to the real world, 2011.