AUDIO GUIDANCE FOR OPTIMAL PLACEMENT OF AN AUDITORY BRAINSTEM IMPLANT WITH MAGNETIC NAVIGATION AND MAXIMUM CLINICAL APPLICATION ACCURACY

Ognjen Miljic, Zoltan Bardosi, Wolfgang Freysinger

4D Visualization Laboratory
University ENT Hospital
Medical University of Innsbruck, Austria
ognjen.miljic@student.i-med.ac.at

ABSTRACT

For patients with ineffective auditory nerve and complete hearing loss, Auditory Brainstem Implant (ABI) [1] presents diversity of hearing sensations to help with sound consciousness and communication.

At present, during the surgical intervention, surgeons use pre-operative patient images to determine optimal position of an ABI on *cochlear nucleus* on brainstem. When found, the optimal position is marked and mentally mapped by the surgeon; Next, the surgeon tries to locate the optimal position in patient's head again and places the ABI. The aim of this project is to provide the surgeon with maximum clinical application accuracy guidance to store the optimal position for the implant, and to provide intuitive audio guidance for positioning the implant at the stored optimal position. By using three audio methods, in combination with visual information on Image-Guided Surgery (IGS), surgeon should spend less time looking at the screen, and more time focused on the patient.

1. INTRODUCTION

This work presents a dynamic audio feedback system for positional guidance in real-time during surgical procedure of ABI placement. ABI is a solution for individuals with hearing loss due to an ineffective auditory nerve, and it is implanted on *cochlear nucleus* which is located on the anterior part of the brainstem. Bypassing both, the inner ear and the auditory nerve ABI stimulates the *cochlear nucleus* and provides the patient with a hearing sensation, which can improve communication and consciousness.

At present, during the surgical intervention, surgeons use pre-operative patient images, usually including Magnetic Resonance Imaging (MRI) and/or Computed Tomography (CT), to determine the optimal position for the ABI on the *cochlear nucleus*.

The main purpose of this project is:

- Providing support for better spatial accuracy in preoperative planning of the ABI implementation

- Remembering the spatial position that is localised with an Electronic Auditory Brainstem Response (E-ABR)

- Final positioning of the ABI with maximum possible accuracy in the appropriate position

During the IGS, location of the surgical probe is visualized by tracking and mapping its location to the pre-operative model of patient anatomy. Nevertheless, the main issue during IGS is that the surgeon often has to divert attention from patient to the screen (navigation system). By using auditory cues, surgeon should spend less time looking at the screen and more time focused on the patient. Combining audio guidance signals with existing IGS visual information, may result in greater accuracy when locating a given target in a 3D volume [2].

2. METHODS

The proposed audio guidance system is comprised of the following elements: (i) *NavABI* software for IGS developed in house - basis of this software is Rhinospider Technology [3] developed by the University Hospital for ENT at the Medical University of Innsbruck (ii) custom NavABI audio plug-in software developed using OpenAL(Open Audio Library) software interface to audio hardware (iii) Electromagnetic Tracking System - Aurora NDI [4] (iv) SoundWear Companion speaker BOSE - 2.0 wearable Bluetooth speaker for presenting audio guidance.

In this work three different types of audio guidances are used. The main difference between following methods lies in complexity and cognitive effort used to understand audio signals. From the simplest, Pulsed Tone sonification distance guidance that most of the participants are familiar with (car parking assistant), over Signal To Noise sonification (participants should recognize it as tuning the old radio) which also represent distance information, to the method designed for guidance in all three axes of Euclidean space by using three different perceptions of the sound Pitch, Loudness and Duration.

To avoid perceptual inaccuracies on the part of the listener, amplitude scale (frequency as a function of amplitude) according to the Fletcher-Munson Curve [5] is taken into account for each tone in all three methods.

2.1. Pitch, Loudness and Duration Sonification (PLD)

This solution is based on the idea that three different perceptions of the sound (Pitch, Loudness and Duration) guides the surgeon along the *X, Y* and *Z* axes of the operating table.

t marks the desired target location in the operating table space, t_x, t_y and t_z mark the projections of the target to the *X*, *Y*, *Z* axes. Similarly, p_x, p_y and p_z marks the projection of the probe to the table axes. The projected signed distance between the target and probe along axis ● is defined (1) as:

$$d_\bullet = t_\bullet - p_\bullet \qquad (1)$$

These distances are used as inputs to control the sonification. This method consists of:

Loudness: *X - Axis* - encoded in stereo channels of the speaker. The left $l(t)$ and the right $r(t)$ (2) channels are encoded as:

$$l(t) = \tilde{l}(d) \cdot o(t, d_y, d_z)$$
$$r(t) = \tilde{r}(d) \cdot o(t, d_y, d_z)$$
$$\tilde{l}(d_x) = \begin{cases} d_x \leq 0, & 1 \\ d_x > 0, & 1 - \frac{d_x}{d_{max}} \end{cases} \qquad (2)$$
$$\tilde{r}(d_x) = \begin{cases} d_x \geq 0, & 1 \\ d_x < 0, & 1 + \frac{d_x}{d_{max}} \end{cases}$$
$$d_x = t_x - p_x$$

Where d_{max} represent the maximum distance of surgical tip from the target point.

Pitch: *Y - Axis* - represented by two sine tones base and alternate frequency. As long as the user keeps pointer at the target, only one tone can be heard (440Hz). As soon as user start moving away from the target along the *Y* axis (either up or down), variations in frequency of one sine tone happens.
$S(t)$ signal (3) is represented as combination of base and alternate signals:

$$S(t) = \text{base(t)} + \text{alternate(t)}$$
$$\text{base(t)} = \sin(t \cdot f)$$
$$\text{alternate(t)} = \sin(t \cdot f_{act}(d_y))$$
$$f_{act}(d_y) = \begin{cases} f, & |d_y| < d_{min} \\ f + sgn(d_y) \cdot f_{max}, & |d_y| > d_{max} \\ f + sgn(d_y) \cdot f_d(d_y), & d_{min} \leq |d_y| \leq d_{max} \end{cases}$$
$$f_d(d_y) = (f_{max} - f_{min}) \cdot \frac{|d_y| - d_{min}}{d_{max} - d_{min}} + f_{min}$$
$$(3)$$

Where f_{act} represents actual frequency for a given distance, and f_{max} and f_{min} represent maximum (500Hz) and minimum (380Hz) frequency. d_{min} and d_{max} represent minimum and maximum distance of surgical tip, from the target along *Y* axis.

Duration: *Z - Axis* - is encoded with pulsed tone (duration of the sound) - Sine (440Hz) tone that pulses gradually faster as the tip of the surgical tool is closer to the target.

$S(t)$ pulsed tone signal (4) is represented as:

$$S(t) = \text{Step(t)} \cdot x(t)$$
$$f_{act}(d_z) = f_{max} - (f_{max} - f_{min})r_{act}(d_z)$$
$$r_{act}(d_z) = \frac{d_z - d_{min}}{d_{max} - d_{min}}$$
$$d_z(\tilde{d}) = \begin{cases} d_{min}, & \tilde{d} < d_{min} \\ \tilde{d}, & d_{min} \leq \tilde{d} \leq d_{max} \\ d_{max}, & \tilde{d} > d_{max} \end{cases} \qquad (4)$$
$$\text{Step(t)} = \left\lfloor t \cdot f_{act}(d(\tilde{d}(t))) \right\rfloor mod\ 2$$

$\lfloor \cdot \rfloor$ represents the number truncated to the closest integer.

2.2. Signal to Noise Sonification(SNR)

This audio guidance, which showed best results in the work of J. Plazak [6] consists of two sounds, white noise and pure sine tone at 440Hz, with the volume mixture of the sounds being controlled (linearly) by distance information:

- Distance $\geq$ 600mm - Results with 100% white noise
- Distance = 300mm - Results with 50% white noise and 50% sine tone
- Distance = 000mm - Results with 100% sine tone

The signal $S_{snr}(t)$ (5) is given by:

$$S_{snr}(t) = r_{act}(d) \cdot n(t) + (1 - r_{act}(d)) \cdot x(t)$$
$$n(t) \sim \mathcal{N}(0,1) \cdot A$$
$$t \in \mathbb{R}$$
$$A \in \mathbb{R} \qquad (5)$$
$$x(t) = \sin(f \cdot t) \cdot A(f)$$
$$f \in \mathbb{R}$$

White noise $n(t)$ is generated from standard normal distribution $\mathcal{N}(0,1)$. A represents the amplitude multiplier and the f denotes the frequency of the tone.

2.3. Pulsed Tone Sonification(PT)

This audio guidance consists of a short (t = 0.1s) sine tone (440Hz) that pulses continuously faster rates as the user approaches the target, up to a point at which pulses linked to form a continuous tone (on target). This type of audio guidance can be found in variety of other commercial applications (e.g. car parking system), and users are familiar with this type, which was the main reason why this method was included in this research. The rate at which the tone pulses is controlled by the distance information, having a range from a slow pulse at 1Hz (60 beats/min) until it reaches a continuous sine tone 20Hz(1200beats/min) on target. The pulsed tone

signal PT(t) (6) is given by:

$$PT(t) = Step(t) \cdot x(t)$$
$$f_{act}(d) = f_{max} - (f_{max} - f_{min})r_{act}(d)$$
$$r_{act}(d) = \frac{d - d_{min}}{d_{max} - d_{min}}$$
$$d(\tilde{d}) = \begin{cases} d_{min}, & \tilde{d} < d_{min} \\ \tilde{d}, & d_{min} \leq \tilde{d} \leq d_{max} \\ d_{max}, & \tilde{d} > d_{max} \end{cases} \quad (6)$$

Step function Step(t) is defined in (4).

3. EVALUATION OF METHODS

An experimental study has been designed with 20 planned participants in controlled experiment, 10 participants with IGS experience (surgeons and researchers), and 10 participants without experience with IGS. There are total of 7 experimental conditions and each condition should be repeated 10 times by each participant (a sum of 1400 trials):

- 3 audio only conditions (3 methods listed above)
- IGS only
- 3 audio and IGS combined conditions

In each trial, the main goal is to navigate the surgical pointer to randomly placed target points within the CT and/or MRI volume as quickly as possible. Each trial should last for 20 seconds. Following metrics will be used to evaluate results. Experiment setup is presented in Fig. 1. Several metrics are evaluated on the trajec-
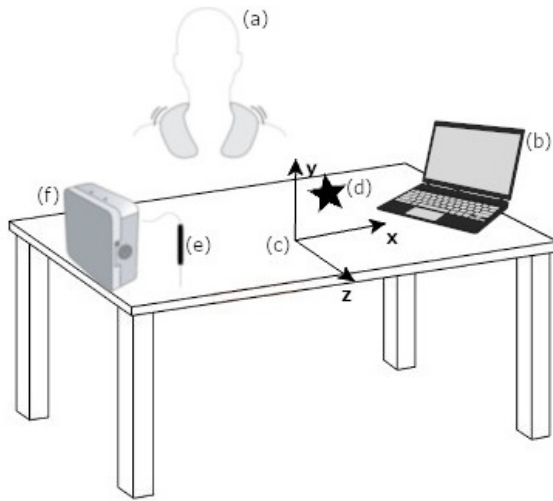

Figure 1: Experiment setup (a) Participant wearing a BOSE audio hardware; (b) Computer containing IGS and audio guidance plugin; (c) Euclidean coordinate system of the table ; (d) Target point in 3D space, which need to be found by participant; (e) Surgical probe tool, which is tracked by magnetic navigation; (f) Field emitter of the NDI Aurora magnetic tracker

tories followed by the participants during the navigation sessions.

3.1. Length of Trajectory

Main hypothesis states that adding audio information to visual display would improve performance within 3D navigation task by shortening the length of trajectory between starting and ending point. For each 20s trial, we can calculate the total length of trajectory by summing the length values between the surgical pointer and the randomly placed 3D target location, and then dividing by the number of samples recorded within the trial. These trial lengths of trajectories were then used as data points to investigate the average trial efficiency as a function of both stimulus modality (audio, visual, audio-visual) and sonification type. Trial lengths of trajectories for the different sonification types that are used in the study are also analyzed. This metric is useful only for cases when user find the target point in given time, otherwise, length of the unsuccessful trajectories will not be considered. Formally, trajectory length $l(x(t))$, for trajectory $x(t)$ is given by (7):

$$l(x(t)) = \sum_{t=0}^{T-1} \|x(t+1) - x(t)\|^2$$
$$\widehat{x}(t) = \frac{1}{2k-1} \sum_{s=t-k}^{t+k} x(s) \quad (7)$$
$$l(\widehat{x}(t)) = \sum_{t=0}^{T-1} \|\widehat{x}(t+1) - \widehat{x}(t)\|^2$$

3.2. Questionnaire responses

With the questionnaire, we would like to see personal opinion of the participants, and how do they rate the difficulty of each condition, and do they find it useful for specific audio methods. After the experiment, each participant completes a short questionnaire regarding the types of feedback that they find to be most useful. 7-point Likert scale is used to rate difficulty of the three conditions, and also the utility of the three different types of sonification.

3.3. Forward/Backward steps

Main hypothesis states that by adding audio cues, user will need less time to achieve the goal, and by that, most of the time using audio guidance, user will move in the direction of the target. Using this metric, we measure either participant understood audio cues and moves probe towards the target point, or participant misunderstood cues and moves probe in wrong/opposite direction. Logging the coordinates every second, from the beginning until the end of the experiment, information about the tendency (whether the participant is getting closer or further from the target over time) will be provided. The forward fwr(x(t)) and backward bwr(x(t)) steps for trajectory $x(t)$ are represented as (8):

$$bwr(x(t)) = \frac{\sum_T bw(t)}{T}$$
$$fwr(x(t)) = \frac{\sum_T fw(t)}{T} \quad (8)$$

T represents overall length of the trajectory.

$$d(t) = \|x(t) - t*\|^2$$
$$\text{bw(t)} = \begin{cases} d(t+1) > d(t), & 1 \\ d(t+1) \leq d(t), & 0 \end{cases}$$
$$\text{fw(t)} = \begin{cases} d(t+1) > d(t), & 0 \\ d(t+1) \leq d(t), & 1 \end{cases} \tag{9}$$
$$\text{fwr(x(t))} = 1 - \text{bwr(x(t))}$$

Where $t*$ represents the target in 3D space. $d(t)$ represents Euclidean distance between trajectory at time t and target point $t*$ (9).

3.4. Eye Tracking

Main hypothesis states that by using audio combined with IGS, user will spend less time looking in to the screen, and more time focused *in situ*. By setting up web camera that will record participants during experiment, we can measure time that participants spend looking at the screen or looking in to the virtual patient (using only audio guidance). This metric will be used on Audio-Visual combined condition only. Where e (10) represents experiment and $_i$ stands for iteration, $_r$ for repetition.

$$\text{e}_{i,r}(t) = \begin{cases} 0, & \text{participant looking at the screen} \\ 1, & \text{participant looking at the model} \end{cases}$$
$$\text{mr}_{i,r} = \frac{1}{T_{i,r}} \int_0^{T_{i,r}} \text{e}_{i,r}(t)dt \tag{10}$$

mr stands for model ratio of experiment. T - overall time of the experiment (20s).

4. CONCLUSION

The experiment is ongoing and results are expected to be presented. Metrics mentioned above, will be used for evaluation of results.

5. FUNDING

6. REFERENCES

[1] "Mi1000 CONCERTO ABI, Mi1000 CONCERTO PIN ABI surgical guide," MED-EL GmbH.

[2] D. Black, C. Hansen, A. Nabavi, and *ET AL*, "A survey of auditory display in image-guided interventions," *Int. J. Comput. Assist. Radiol. Surg.*, pp. 1–12, 2017.

[3] Z. Bardosi, Y. Özbek, C. Plattner, and W. Freysinger, "Auf dem Weg zum Heiligen Gral der 3D-Navigation: submillimetrische Anwendungsgenauigkeit im Felsenbein," *CURAC*, pp. 155–158, 2013.

[4] "Aurora V2 User Guide," NDI Europe GmbH.

[5] H. Fletcher and W. Munson, "Loudness, Its Definition, Measurement, and Calculation," *B.T.S.J.*, vol. 12, no. 5, pp. 377–430, October 1933.

[6] J. Plazak, S. Drouin, L. Collins, and M. Kersten-Oertel, "Distance sonification in image-guided surgery," *Healthcare Technology Letters*, vol. 4, pp. 199–203, 2017.