Gesture control: investigating new ways of interactions with the hearing instrument

Re-engineering Innovative Technology Based On Electromyogram (EMG) Signal To Help The Hearing Impaired

Devid Bianco, David N. Black, Paolo A. Mesiano, Tue S. Sørensen

Abstract—Currently the most technically advanced means to interact with a hearing device is tactile interaction via an application on a smartphone. Proposing gesture interaction as an additional or alternative method of interacting with hearing devices, we investigated the possibility of improving usability and discreetness of interactions, addressing the problems faced by developers and their end users.

Using the Myo gesture control armband to measure the electric potential of the muscles in the arm (EMG) to detect arm and hand gestures, we applied gesture control to a simulation of a hearing device and then compared it with the smartphone application.

Despite available technologys drawbacks regarding discreetness, faster and more direct interactions emerge when using gestures rather than the mobile application. Therefore, future technological implementations of gesture recognition may allow an alternative way of interacting with hearing devices.

Index Terms—here, we, put, keywords

I. INTRODUCTION

The hearing device industry is constantly facing issues with usability and discreetness of interaction with technology. Therefore, new ways of interacting and controlling hearing devices is a focal point of hearing device developers and their end users.

Hearing instruments settings are currently accessed through several methods: tactile pushbuttons placed on the hearing device or situated on a secondary device normally hung around the neck. This secondary device also acts as an intermediary that allows communication between the hearing instrument and different audio-streaming devices (such as tv, phone, radio, etc.) which can also be controlled with a mobile phone application.

Each interaction type has drawbacks related to discreetness and usability. After time, these lead to the user limiting their interaction to simply turning the hearing device on and off. In detail, buttons have only spring resistance as feedback for the user and the tactile recognition may not be accurate. On the other hand, smartphones require visual feedback that may distract the user from their environment. Furthermore, some users may consider the intermediate device conspicuous to wear even if its location around the neck makes it accessible at all times.

In this research we propose gesture control based on electromyography as an alternative way of interacting with a hearing device with the purpose of overcoming the current impediments for the user and to improve accessibility. A profile of gestures is designed and implemented to control a simulation of currently accessible features. The complexity of the two ways of interaction is measured and compared through real life simulations of hearing instrument usage.

II. RESEARCH - WORK IN PROGRESS

Given the concerns with interaction from current hearing device users and also looking from the perspective of a first time user we explored the possibilities of using sonic interaction, gesture interaction, neural interaction and extrasensory devices to improve usability and discreteness. Whilst sonic interaction has the benefit of not needing the user to be trained it also has the unwanted effect of being audibly indiscreet when giving feedback to the user which can distract from social interaction. Gesture interaction is limited in the number of actions that are available to the user, needs training to be used fluently and can be aesthetically conspicuous to wear. In spite of these points it has the advantages of allowing discreet gestural interaction and permits the user to focus on primary tasks. Despite neural interaction being the most discreet physical method of interaction it is still an underdeveloped technology with regard to precision. Unintentional activation, training times and focus needed to interact with a device are all details that need to be refined before this technology becomes practical for everyday use. Extrasensory devices can be very specific in their ability to

After evaluating the state of the art from these different possibilities, we found that the Myo gesture control armband fulfilled the requirements of both discreteness and usability in terms of faster, more immediate and less evident interaction, especially in situations where the full attention of the user is demanded for more important tasks such as social interaction. In the Myo gesture control armband, medical grade stainless steel EMG sensors measure the electric potential of the muscles around the forearm and permit the classification of 5 hand gestures based on the collected signals. A nine-axis IMU containing a three-axis gyroscope, three-axis accelerometer, three-axis magnetometer permit the detection of position and motion in space. Wireless technology and rechargeable battery make it usable for our purposes.

Here we should also insert a short description of the accessible features with the phone that were replicating. ConnectLine app

III. TEST

A. Pilot studies

Before going into the final test, the handshaking between the C++ script and the PD patch for the MYO-application saw many debug processes and basic usability test. Several opinions were collected about usability and discreeteness, thus adjustments were implemented based on the feedback we obtained through repetitive rehearsals, pilot experiments on fellow students, and from presenting a prototype to professional hearing device developers. In general, a large interest of the public, specially the young public, was shown toward gesture interaction, with regard to a possible future implementation in smaller devices. In these occasions, the MYO showed very poor adaptability to many different arms over a short period of time. As the Myo has a limited number of of recognizable hand gestures, an online test group was given a questionnaire to ascertain which gesture would be most suited to control certain commands from a generic media player. The results informed us of the most appropriate way to design our final test gestures.

B. Purpose of final test

With the aim to evaluate the quality of gesture control as a new approach in controlling hearing devices, we focused on its usability and learnability, together with the perceived discreeteness, considering these as important aspects for the end user of the hearing device. These three features have been evaluated in absolute terms and also in comparison with the actual approach made through the mobile phone application, in order to understand if gestures can effectively improve or give an alternative to the end user experience.

C. Test group

A group of twenty subjects has been tested, with an age ranging from 22 to 38 years old, and a balanced number of male and female participants. All subjects reported having normal hearing and no hearing impairments. 80% of the subjects rated their experience with technology as high or very high and only one evaluated their experience as very low. 40% use a music application on a mobile device daily and only one declared not to use it at all. 10 participants had no familiarity with gesture interaction devices (such as the Wii©controller, Microsoft Kinect©or the Myo©) and 7 participants had good or high familiarity with this kind of technology.

D. Apparatus

To allow the Myo armband to interact with applications, a software development kit (SDK) released by Thalmic Labs has been included in a C++ application. We mention a short description of this SDK provided by the Thalmic Labs:

At its core, the Myo armband provides two kinds of data to an application, spatial data and gestural data. Spatial data informs the application about the orientation and movement of the user's arm. The Myo SDK provides two kinds of spatial data: An orientation represents which way the Myo armband is pointed. In the SDK this orientation is provided as a quaternion that can be converted to other representations, like a rotation matrix or Euler angles. An acceleration vector represents the acceleration the Myo armband is undergoing at any given time. The SDK provides this as a three-dimensional vector. Gestural data tells the application what the user is doing with their hands. The Myo SDK provides gestural data in the form of one of several preset poses, which represent a particular configuration of the user's hand. For example, one pose represents the hand making a fist, while another represents the hand being at rest with an open palm [source]

The C++ script receives both spatial and gestural data from the SDK, then performs a wide range of tasks: convert and filter orientation data to arms rotation data; rescale rotation data to a range that fits the users specific mobility; translate rotation data and gestural data to messages according to OSC protocol [specifications]; create a UDP network over which the OSC messages are sent to Pure Data; based on rotation and gestural data provide feedback to the wearer of the Myo armband by issuing a vibration. This causes the Myo armband to vibrate in a way that is both audible and sensed through touch;

Parallel to the Myos software, a TouchOSC application runs in a smartphone, showing buttons and scrollbars in a layout similar to the Oticons Connect Line app. When the user interacts with the objects on the screen, OSC messages are produced and sent via Wifi connection through an ad-hoc network hosted by the computer containing the software.

All OSC messages are then collected over the network by a patch in Pure Data, which is programmed to: simulate the audios and functionalities of an hearing device; receive, translate and route OSC messages to control the simulation parameters (volume, sources, presets); provide an a GUI feedback showing in which way the user is interacting with the simulation; provide a testing engine. The collected data will be recorded in a *.txt file The user can listen to the audio feedback given by the simulation in real time through headphones. Moreover, they receive vibrotactile feedback while interacting with gesture control, and visual feedback while interacting with the smartphone.

E. Method

Before the test, the participant was asked to wear the armband in a specific (but comfortable) position on the right arm, which we found giving the best results in terms of reliability. In this way, while they were introduced to the experiment, the armband could have enough time to warm up and adapt to the their muscle activity. For each subject two different calibrations followed: the first was needed to set the subjects individual arm rotation range; the second one was the one requested by the Myo for a proper gesture recognition. In the former, we were asking every user to rotate repeatedly the arm (resting on the side) from the minimum possible angle to the maximum, without forcing the extremes. This calibration let the script know what was the comfortable rotation range for the users arm to be used for interaction. The latter calibration is done through the Myo Connect utility from Thalmic Labs, and record the users muscle activity for every gesture, in

order to properly classify them later. The test consisted of two sections, one with the gesture interaction (using the Myo) and one with the haptic interaction (using the smartphone). In each section the subject was asked to complete 20 tasks to control the features of the hearing device simulation. The tasks were randomly chosen from a list of five possible tasks, hence each of the five tasks were repeated four times. Each single task consisted in adjusting three different settings: volume regulation, source selection and preset selection (see appendix for detailed description of the list of tasks).

Since the subject could get used to the tasks during the first section, and thus affecting the pureness of the second one, the order of the sections was inverted for every participant. The subject was given oral instructions and was allowed to ask for help at any time. In addition, they were provided with headphones, which were reproducing the sound feedback from the hearing device simulation. Furthermore they were given a poster containing graphical instructions of the HA simulation, thus showing clearly which gesture was needed for every action. Before every section, the participant went through a short preliminary training to memorize and practice the actions to perform. After the training part, one at a time, the number of the task to be completed was communicated to the participant. A reference list of the actions to be performed for every task was also provided. The tester supervised the process to check over the correctness of the tasks and the integrity of the recorded data. After the 20 iterations were completed, next interaction was introduced for the next section, which followed the same procedure. The average overall time for the test was around 45 minutes. After the test, the user was asked to compile a short questionnaire.

F. Data collection

Usability and learnability of the control method can be expressed in terms of complexity of the interaction with the device. For a given action on the hearing device, we measured its complexity by combining three types of parameters: the duration, the number of physical interactions with the device and number of attempts needed by the user to complete the action. These three parameters are measured as follows. Regarding the test on the mobile phone, the time of interaction is intended as the time to take the phone, unlock it, open the application, interact with it, lock the phone, and put it away. The physical interactions counted are the unlock action, each subsequent graphical object tapped on the screen and the lock action. The number of attempts is intended to be the number of times the user has unlocked the phone in order to have access to the settings of the hearing device. For the test with gestures, the time of interaction is intended as the unlock state duration of the Myo device. The number of interactions is measured as the number of gestures made by the user during the unlock state while completing the task (included the unlock gesture). As with the phone, the number of attempts is meant as the number of times the Myo is unlocked. In the final questionnaire, participants were asked to rate the ease and comfort of the two approaches used in the experiment together with their discreteness in a 7-point scale and to indicate whether they would use gestures or the smartphone in a list of everyday situations and environments. Additionally, a written report of the test was taken, in order to be able to subsequently account possible bugs or problems with the data in the analysis process.

IV. RESULTS ANALYSIS

Due to the relatively low number of participants the data of every single participant was checked against the readings of the Myo functioning without error. This led to discarding one subjects data, which was corrupted by too much noise caused by an imperfect calibration of the Myo. This was made evident by the unjustified trend of high time, number of interactions and attempts over the test proceeding, plus a signaled low accuracy of the gesture detection on that occasion. This did not have a meaningful effect on the final data analysis.

A first analysis of the gesture-controlled section shows that the group of participants can be divided into three subgroups based on the total time of their test and the total number of interactions needed to complete all twenty tasks iterations.

As shown in fig (****fig scatterplot), a first group can be identified as those subjects who needed a short amount of time and a very low number of gestural interactions and who are therefore confined to an area close to the origin of the scatter plot. A cross check with the questionnaire results revealed this to be the group of subjects who rated themselves as very experienced with gesture interaction (6 or 7 on a 7-point scale). A second group of participant can be defined as subjects with normal learning and a congruent functionality of the armband during the test. They present a longer time and more interactions if compared to the very experienced subjects, yet still maintaining a direct proportionality trend in the graph. Linearity between time and number of interactions is less evident in a third group of subjects whomthat exhibitpresents the highest difficulty of interaction with the HA simulation. This couldmay be due to issuesproblems with gesture detection (incorrect detection of gestures results in a higher number of recorded interactions), or to the subjects ability to recognize and interact with the audio feedback given by the simulation (resulting in a longer time), or to a slower learning process. For a restricted A limited number of participants conveyedit has been noticed a difficulty in identifyingdistinguishing a specific audio preset or a specific source simulation: this could be also be attributedascribed to the design of the HA simulation assince some audio sources could be misinterpreted (e.g. radio and tv sources). Despite thisthat, the congruency in the 85Further investigations regarding the self-rated experience (from the questionnaire) have not lead to conclusive results. Thus, because of the questions generality and ratings subjectivity, we can not state solid hypothesis on the qualitative data from the questionnaire.

All the 5 tasks were randomized through the iterations in such a manner that they would have 4 repetitions each one. This was done in order to remove from the results a possible bias due to the different nature of every task.

Learning curves have been used to measure the trend of learnability for the two types of interaction. Complexity has

been averaged over all subjects and evaluated for the twenty task iterations. This produced three curves that quantify the learning for increasing experience: in terms of time, number of interactions, number of attempts.

Due to the design of the HA simulation and to the design of gesture interaction, different tasks were characterized by different overall difficulty. While interacting with gestures, the lack of a visual feedback and the nature of the required hand pose, determined higher difficulty for the selection of some settings among the others.

Another cause could be an unsuccessful gesture design, causing the user to struggle to reach the pose or the rotational angle required. This lead the subject to struggle to reach the required hand pose or rotational angle, in particular when selecting the right audio preset. Moreover the features of some audio feedbacks turned out to be more difficult than others to recognize.

As shown in fig (****plot task difficulty), task labeled as 1 required a bigger amount of interactions, attempts and time for the average of the participants while task labeled 3 turned out to be the less complex. This task is characterized by a very recognizable tone (beep from phone tapping and phone ringing), as well as a preset very easy to select. To take this into account, the reciprocal of task difficulty has been used as a weighting factor for the corresponding measure of time, number of interactions and number of attempts. For example, the time-based complexity has been calculated as a weighted average,

$$\bar{t} = \frac{\sum_1^N w_i t_{task}}{N \sum_1^N w_i} \tag{1}$$

where weights are the reciprocal of normalized task times:

$$w_i = \frac{max(t_{task})}{t_{task}} \geq 1 \tag{2}$$

This issue was not present in the haptic interaction where all tasks had a similar complexity.

The resulting learning curve is smoother but doesnt differ considerably from the one obtained with normal average. This suggests that single task difficulty doesnt affect the measures since the complexity of interaction is self-averaging due to randomization.

The learning curves show a large spread of the data among subjects for the first five tasks: as described in fig (*****plot with percentiles - MYO only) at the beginning of the test some subjects showed difficulty in gesture interaction. These difficulties are averaged down with increasing iterations, as shown by the decreasing spread of outliers, demonstrating a fast learning process for increasing experience. In terms of time and number of attempts the usability of gestures is lower than that of the smartphone, and this can be trivially explained by the fact that the latter is a widespread and well known technology, while gesture interaction and the Myo in particular can be considered a novel technology for the average user. Regarding the number of interactions, the smartphone shows a higher complexity, but the duration of each single interaction must be considered: the single interaction on the phone is a fast tap on the screen while a gesture can be hold for as long as needed by the user.

The three curves show a possible saturation of the learning process for what concerns gestures, but a closer analysis to the subgroup of experienced users demonstrate that with a longer usage the average user can reach a usability comparable to or even higher than the one on the smartphone. Experienced users of the gesture armband device were both faster and more accurate since they needed less interactions than the average population of phone users and a maximum of two attempts to complete an action.

V. CONCLUSION

The conclusion goes here.

APPENDIX A
APPENDIX 1

Appendix one text goes here.

APPENDIX B
APPENDIX 2

Appendix two text goes here.

ACKNOWLEDGMENT

REFERENCES

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.