Evaluation of Human-Myo Gesture Control Capabilities in Continuous Select Operations for Assistive Technology

Ádám Csapó
Széchenyi István University
9026 Győr, Egyetem tér 1.
Hungary
csapo.adam@sze.hu

Árni Kristjansson
University of Iceland
Oddi v. Sturlugötu, Reykjavik 101, Iceland.
ak@hi.is

Hunor Nagy
Széchenyi István University
9026 Győr, Egyetem tér 1.
Hungary
nhunorz@gmail.com

György Wersényi
Széchenyi István University
9026 Győr, Egyetem tér 1.
Hungary
wersenyi@sze.hu

ABSTRACT

Tactile and haptic devices can be used to control and interact with a wide range of systems, including games, virtual environments and assistive technologies. Although many psychophysical studies have measured thresholds of human sensory capabilities for interpreting haptic and tactile feedback, relatively little is known about the precision with which we are able to guide the behavior of a system based on kinesthetic and myoelectric gestures. A broad study of the latter problem is important, especially now that a number of devices have appeared – such as the Leap Motion Controller and the Myo armband – which enable humans to use finger, hand and arm gestures to interact with the digital world. This paper provides a broad overview on the topic, and reports a set of preliminary experiments on the extent to which the Myo armband can be used to control auditory feedback in real time. The goal is to investigate ways in which visually impaired users could use the Myo to control the output of an assistive technology.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: *Input devices and strategies*, Haptic I/O

H.1.2 [**User/Machine Systems]**: Human factors

General Terms

Measurement, Design, Economics, Reliability, Human Factors, Verification.

Keywords

haptics, gesture control, acoustics, auditory feedback.

1. INTRODUCTION

Tactile and haptic devices can be used for interaction with and control of a wide range of systems. The term *tactile* refers to all sensations resulting from displacement of the skin. Tactile feedback can extend to the perception of temperature and pain, and is often also thought to encompass internal sensations perceived in muscles, tendons, joints and even body posture [1]. In comparison, *haptic perception* refers to higher level perceptual processing of multiple inputs obtained through the skin, muscles and tendons. It also usually involves active contributions from the subject (e.g. information gathering through movements). When humans explore various properties of an object, different actions are performed to check for spatial qualities (volume and form), surface qualities (roughness / smoothness), material qualities (e.g. softness, temperature) and dynamic qualities (e.g. weight). Participants can recognize objects by exploration through touch only in ~2 seconds [6]. Nevertheless, geometric forms and shapes are easier to detect visually than through haptic feedback only: line drawings or graphs presented through raised contours are often difficult to recognize. At the same time, tactile qualities and temperature of a surface can be detected only by physical contact.

Several different types of experimental measurement have been devised in the literature to quantify human capabilities for using the tactile and haptic modalities in a variety of circumstances. For example, sensitivity to mechanical pressure on the skin is not uniformly distributed on the body surface: in general, sensitivity is highest on the face, arms and fingers, followed by the thigh, calf and foot [2]. Partly related to this is the ability to discover dots on a plane (such as when reading Braille), in which case dots that are at least 1 micrometer high are required for detection at a precision of 75% [7]. During detection of vibrations, frequencies below 5 Hz and up to 400-700 Hz can be detected [8]. In this case, sensitivity is directly proportional to the frequency. For spatial resolution (acuity), the classical two-point touch threshold (also called 2-point discrimination, 2PD) can be measured. This is the smallest separation at which participants can tell whether they are touched at one or two points. Spatial acuity also varies across the body, but the same parts (toes, face, and fingertips) show the highest acuity. The resolution can be as high as 1 mm [4], which is higher than auditory acuity, but lower than the visual acuity of humans. Judging by 2-point discrimination threshold values, the fingertips have the highest spatial acuity for touch (and pain). On the arm and hand, glabrous skin is more sensitive (0.5 - 1 cm thresholds) than hairy skin (1.5-3 cm) [12]. The same applies to the legs, where both the calf and thigh show about 2-3.5 cm threshold levels. Two point thresholds are somewhat higher for successive than for simultaneous stimuli. Finally, sensitivity to temporal changes can be detected when subjects have to decide whether pairs of stimuli are simultaneous or successive. Temporal differences as low as 5 ms can be resolved [3]. This is better than vision (25ms) but worse than audition (0.01 ms). Although accessible to both sighted and visually impaired users, the precision of the tactile sense varies by task [14].

Similarly to the tactile modality, human performance can also be quantified across various dimensions for haptic perception. For example, haptic *object localization* means finding objects without visual help (e.g. reaching for the alarm clock without opening the eyes). Interestingly, in contrast to audition and vison, there is no fixed frame of reference (egocentric middle) for haptics [1]. Actions like left-right index finger touching, setting bars to parallel with left and right hand etc. show low accuracy. Furthermore, tactile attention, like auditory and visual attention, is a limited resource that can be influenced, distracted or strengthened [15]. Detection accuracy in haptics is largely decreased in the presence of distracting vibrations. The use of multiple modalities can also cause distraction if they transmit contradicting information. While touch and vision are usually complementary, audio and vision conflict each other more often [1]. Finally, it is worth noting that depending on the application and the system that is used for feedback, device and subject can influence each other in unforeseen ways.

Another dimension that is central to haptics is the *kinesthetic sense*, i.e. the perception of limb positions and movements. The precision of this sense can be tested by finding the smallest imposed movement that subjects can detect with their muscles relaxed. Usually subjects are asked to identify the direction of the movement as well. For movements about any joint, the size of the smallest detectable movement varies with velocity. Faster movements are more easily detected than slower ones. For the limb joints, head and trunk, very small displacements of 0.1–0.5 degrees can be detected with movement velocities of more than 1 deg/sec [13]. At slower velocities, larger movements of 1–3 degrees are required for detection. When joints are moved at extremely slow velocities (less than 2 deg/min), there is no sensation of movement at all. Surprisingly, for the joints of the fingers and toes, thresholds for detection of movement are higher than for the limb joints. With very slow displacements, subjects can identify changes of position of 2–5 degrees at the ankle, knee, or shoulder. For the fingers, displacements of 5–10 degrees can be detected [13]. Using vibrating devices, special attention has to be paid to the placement and level of vibrations, as vibration of muscle tendons can activate muscle spindle endings and cause illusions of joint movement. In other words, vibration can distort perception of body shape and posture.

In this paper, we consider some of the psychophysical aspects associated with the use of gestural systems, based on kinesthetic and muscle activity, for controlling feedback in computer interfaces. In particular, we report a set of experiments with the Myo armband from Thalmic Labs. We use the following nomenclature to distinguish between various (high-level) gesture types:

- Discrete search: the aim of the gesture is to make discrete steps along an ordered set of values (such as switching between octaves in auditory output)
- Discrete select: the aim of the gesture is to select one among several discrete steps belonging either to an ordered or an unordered (categorical) set. The key difference compared to the previous point is that in this case, jumps can be made among categories, or between non-adjacent members of an ordered set
- Continuous search: the aim of the gesture is to change the value of a continuous parameter in real time (such as changing the frequency of an auditory signal)
- Continuous select: the aim of the gesture is to select one among an infinite number of continuous parameter values.

The distinction between search and select seems to us to be an important one, regardless of the fact that in any given scenario, one might be able to perform a select operation through a series of search operations. The bottom line is that a successful select operation can in some cases be faster, as it allows values in between to be skipped, but it can also be less reliable, when the intended value is approached from a distance in a single shot.

The paper is structured as follows. In section II, we briefly describe the motivation for this research. Sections III, IV and V deal with our experimental setup, an analysis of results and a short discussion of the findings.

2. THE ROLE OF TACTILE / HAPTIC DEVICES IN ASSISTIVE TECHNOLOGIES

The authors participate in the Sound of Vision project, funded by the European Commission, with the goal of developing an assistive technology that supports the capabilities of the visually impaired for orientation in the world. Besides supporting navigation in unstructured environments, the goal of the project is to provide a generic tool for all aspects of visual perception, including high-level object recognition.

During the project, we are testing 3 different kinds of tactile / haptic devices: tactile bands around the arms and ankles, a tactile vest, and the Myo for control based on myoelectric activity and arm movements. The current work reports results from a task that uses kinesthetic (gestural) input for continuous search operations. The main question is whether the Myo can be used to successfully control the value of a continuous variable. Although a single experiment goes only so far in answering this question, we hope to find key points where further investigations are necessary. In our discussion in section V, we provide a list of such points.

3. EXPERIMENT AND LABORATORY SETUP

The purpose of our experiment was to determine how well the Myo armband can be used to reproduce sinusoidal sounds with changing frequencies in a number of different settings (i.e. with different movement directions and different motion-to-frequency mapping). Each of the 5 stimuli used in the experiment began at a frequency of 500 Hz and contained one or two inflection points between 0.5 and 2.5 seconds. Each stimulus started with a rising change in frequency, with the direction only changing in opposite directions at the inflection points (Figure 1).

The 5 stimuli were to be reproduced by test subjects through kinesthetic movement one by one. There were 4 different test cases, through a combination of two different movement directions and movement-to-frequency mapping approaches. With respect to movement direction (Figure 2), one variation involved the subject moving his or her arm forward and backwards at the side of the body, while in a standing position (this is referred to as the “at-side” case), while the other variation involved the user moving his or her arm towards the left and right in front of the body, while in a standing position (this is referred to as the “in-front” case). In terms of movement-to-frequency mapping, the “direct” case consisted of a mapping between displacement from origin to specific frequency (with pitches rising and falling in alternate directions), while the “acceleration” case consisted of a mapping between degree of displacement and speed at which the frequency changes (for example, in the “at-side” case, moving the arm further to the front caused the frequency to rise more rapidly, while moving the arm further to the back caused the frequency to fall more rapidly).

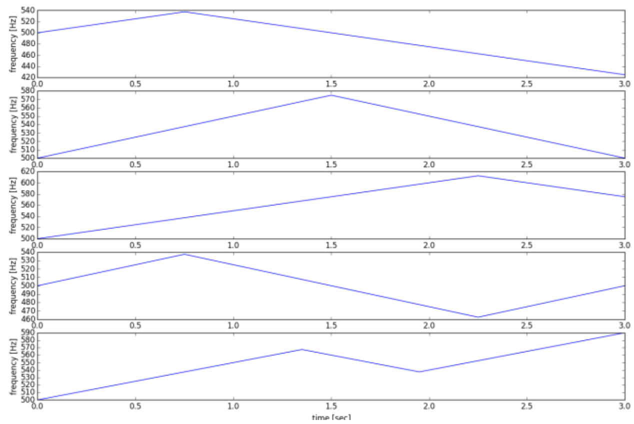

Figure 1. Temporal profiles of pitches as they change between inflection points in the 5 test stimuli.

Figure 2. Illustration of at-side (left) and in-front conditions (right).

Eleven fully sighted university students participated in the experiment. Each subject was asked to reproduce all 5 stimuli in all 4 test cases, but the order of the test cases and the order of the stimuli in each of the test cases was varied randomly to control for learning effects. Prior to each test case, subjects were shown how the movements and movement-to-frequency mappings worked, and were allowed as much time as they needed to prepare for the tests. When they were ready, subjects were asked to alternately listen to and reproduce the 5 stimuli, one after the other. Following each stimulus, there was a period of silence for 3 seconds. Subjects were alerted of the time to start reproducing the stimuli in two ways: through a visual display (the text “GO!!!” was shown), and also through audio (subjects could listen to the sound that was being recorded in real time, so they heard when it started). Prior to each reproduction phase, users could use a fist gesture to reset the state of the device, so that sound reproduction would begin at the starting frequency of 500 Hz.

4. EVALUATION AND RESULTS

Test results were evaluated on a semi-objective scale from 1 to 6. The scoring system used can be credited with some objectivity because each value on the scale had a specific meaning. At the same time, it cannot be seen as fully objective because it was applied to the test stimuli and reproductions through human evaluators prone to occasional errors.

The meaning of the different scores was determined as follows:

- 1 point: Instead of beginning with a rising pitch, the reproduction began with a falling pitch
- 2 points: The number of inflection points in the reproduction were incorrect, and there was no sign of any hesitation at the inflection point(s)
- 3 points: The number of inflection points in the reproduction were incorrect, however a slowing down in the change of frequency was perceptible at the right time (e.g. the test subject clearly intended to alter the direction of the change in frequency, but could not achieve this because of time constraints)
- 4 points: The number of inflection points in the reproduction were correct, but the frequency profile of the reproduction was clearly different from the original stimulus.
- 5 points: The number of inflection points in the reproduction were correct, and the profile of the reproduction was quite similar to the original stimulus, but the reproduction began at the wrong pitch, and a consistent offset was maintained.
- 6 points: Reproduction was perfect.

Evaluation was performed by two evaluators who are musically trained and are among the authors of the paper. While the evaluators were in agreement for the majority of assessments, in some cases there was a difference of one (and very rarely, more than one) point between the assessments of the two evaluators. Therefore, an average score was used for evaluation, resulting in a set of 11 different scores (from 0 to 6, with increments of 0.5). The distribution of scores and offsets between them can be seen on Figure 3.

Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion, Dec. 1–3, 2016, Vila Real, Portugal.

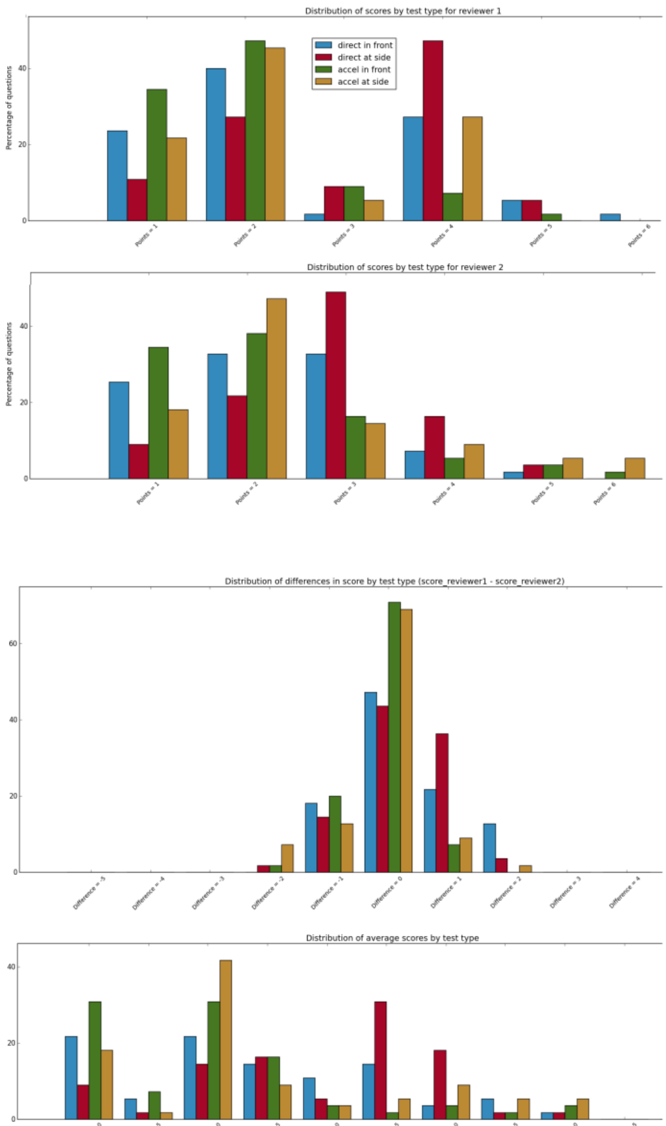

Figure 3. Distribution of scores and disparity between them for the two evaluators.

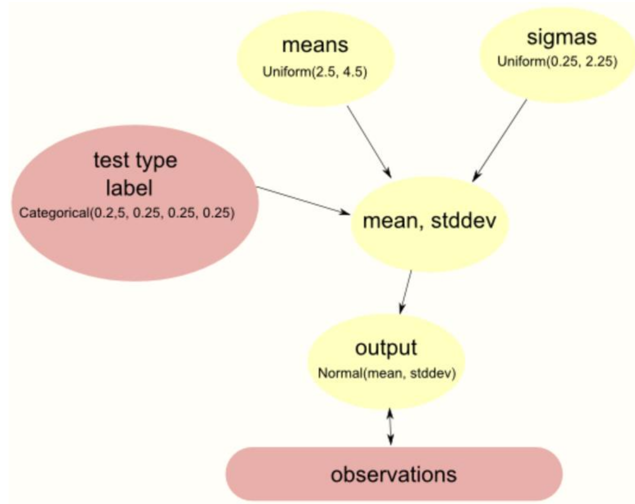

Figure 4. Configuration and initialization of MCMC based approach towards determining the true mean and standard deviation of the obtained scores.

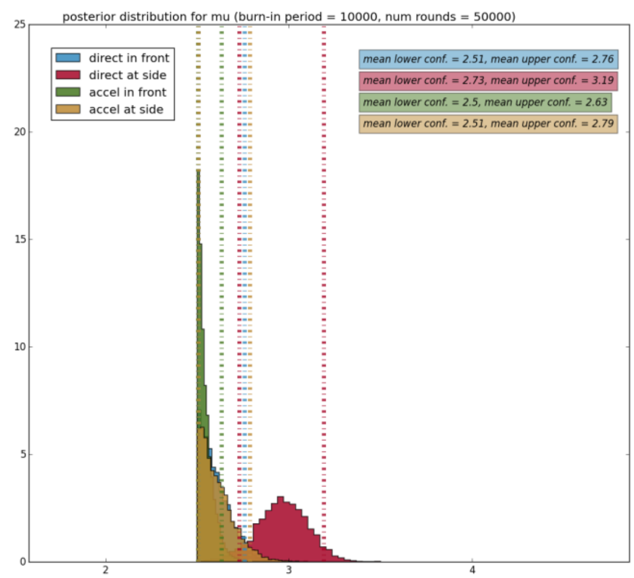

Figure 5. Posterior distribution of mean values for the 4 different test categories. 95% confidence intervals are indicated in dotted vertical lines for all 4 categories.

To better understand the meaning and significance of these average scores, a Markov Chain Monte Carlo simulation based Bayesian approach was applied to the categorical data. Figure 4 shows the configuration of the hidden and observed variables. As a prior, it was conjectured that the results obtained for the 4 test types originate from 4 different normal distributions with different means and variances. The PyMC library was used to implement our inference setup in Python.

The results are displayed in figures 5-7. Figure 5 shows the posterior distribution of the means for the 4 test types. Although 3 of the means have very similar distributions, the distribution of the mean for the "direct-at-side" test was markedly different from the others. At the same time, Figure 6 shows that the standard deviations were practically the same.

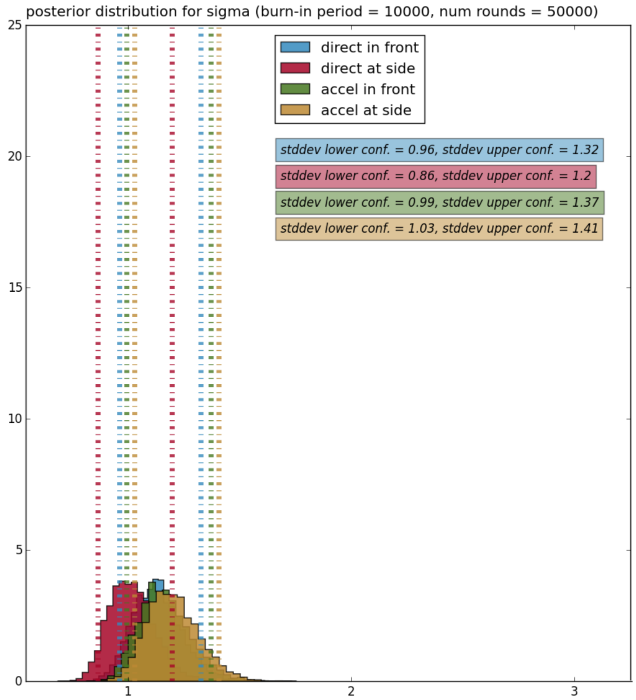

Figure 6. Posterior distribution of standard deviations for the 4 different test categories. 95% confidence intervals are indicated in dotted vertical lines for all 4 categories.

Based on these results, Figure 7 shows the median posterior distribution for the scores in different categories. It is clear based on the figure that most data points are not significantly different in the 4 categories, but given a sufficiently large sample size from the same category, the average score in the "direct-at-side" category would be significantly different from others.

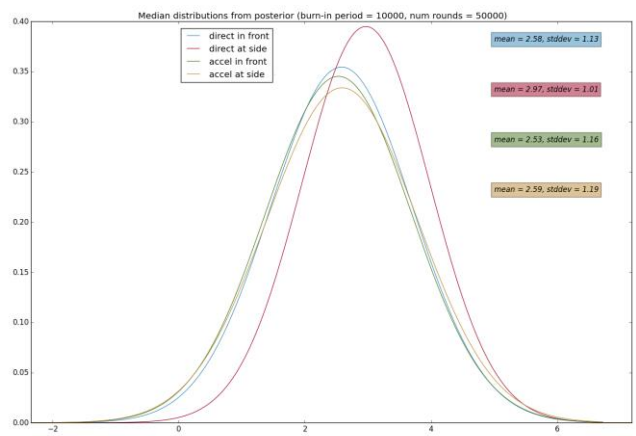

Figure 7. Median posterior distributions for the 4 different test categories.

5. DISCUSSION

Based on the results, we can conclude with some certainty that the "at-side" body posture with direct displacement-to-frequency mapping was on average more comfortable and / or more suitable for continuous search operations than the other alternatives. However, it is important to bear in mind several caveats with respect to this assessment:

- Being more comfortable and / or suitable on average does not mean that individual measurements would always be better (in fact, the median posterior distributions show that individual measurements would be quite similar)
- Regardless of the previous conclusion, both the results and our experience show that the chosen task was difficult: there were very few cases where test subjects were able to reproduce the original samples without a glitch. There can be several reasons for this:
    - The task inherently involved the use of auditory working memory, which may have been difficult in itself for some of the test subjects. It is possible that continuous search operations in a different domain (for example, moving a visual cursor on the screen) would have been easier to perform.
    - Calibration of the same Myo device among different users was difficult. Although each session began with a recalibration, even the stability with which certain gestures were recognized (in particular, the fist gesture used to re-initialize the sound frequency) varied noticeably.
    - The acceleration sensor built into the Myo was used to sense the displacement of the users' arm, however the measurements used were quite noisy and also depended on such factors as the direction in which users were facing, and possibly the arm length of the user. Although a simple relationship was set up between yaw, pitch and roll measurements and their mapping onto frequency, the order in which orientation is converted into the yaw, pitch and roll coordinates turned out to have a strong influence on the output (for example, while standing and stretching the arm to the front, a small displacement would have been converted to a yaw value quite predictably, larger displacements resulted in strong non-linearities that were also different from user to user).
    - Often users forgot to use the fist gesture to reset the output to the starting frequency of 500 Hz prior to reproducing the stimuli, and in an important number of cases, the fist gesture was not registered by the device. In still other cases, a large displacement of the arm from the resting position (especially when moving backwards at the side) resulted in a certain muscle tension that was registered by the device as a fist gesture, causing the frequency to return to 500 Hz in unintended cases.
- Finally, the inadequacy of the scoring system used may have also contributed to the difficulty of evaluating the results. Although each score value had a different meaning, it is for example possible that the distance between 4 and 6 points is too large, as a minor glitch

would have reduced the value of a reproduction from 6 to 4, even if it started at the correct frequency.

- Although differences between the scores given by the two evaluators were rarely greater than 1, the fact that many differences exist may undermine the objectivity of the assessment.

In light of the above points, further experiments could be useful to improve the precision of the mapping between displacement to frequency (some kind of data-driven calibration technique may be useful, i.e. a more intelligent interface between the user and the Myo), and a smoother scale of scoring (perhaps through a multidimensional scoring system with a weighted outcome, or an automated approach using temporal-spectral analysis). As long as the issues with calibration are not suitably addressed, we conclude that the setup investigated is not ready to be applied for continuous search operations in application settings.

6. SUMMARY

Many psychophysical studies have measured thresholds of human sensory capabilities when interpreting haptic and tactile feedback, however, relatively little is known about the precision with which we are able to guide the behavior of a system based on kinesthetic and myoelectric gestures. In this paper, we distinguished between discrete search / select and continuous search / select tasks, and described an experiment for determining the accuracy of continuous search operations with the Myo controller. Although the task turned out to be quite difficult, and the methodology used in some respects lacking in rigor, we can with some confidence conclude that a certain posture and movement-to-signal mapping approach was, on average, better than other alternatives. We also conclude that the methodology for calibration would have to be improved for the Myo to provide convincing results when used for continuous search operations.

7. ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643636 "Sound of Vision".

8. REFERENCES

[1] Wolfe, J. M. 2014. *Sensation & Perception*, 4th edition, Sinauer Ass. Inc., USA, 405-413.

[2] Weinstein, S. 1968. Intensive and extensive aspects of tactile sensitivity a function of body part, sex, and laterality. In D.R. Kenshalo (Ed), *The Skin Senses*, Springfield Ass., USA, 195-222.

[3] Gescheider, G. A. 1974. Temporal Realations in Cutaneous Stimualtion. in *Proc. of Conf. on Cutaneous Communication Systems and Devices*, Oxford.

[4] Loomis, J. M. 2008. On the tangibility of letters and braille. *Percept. Psychophysics* 29, 37-46.

[5] Legge, G. E., Madison C., Vaughn, B. N., Cheong, A. M. Y. and Miller J. C. 2008. Retention of high tactile acuity throughout the life span in blindness. *Percept. Psychophysics* 70, 1471-1488.

[6] Klatzky, R. L., Ledermann, S. J. and Metzger V. 1985. Identifying objects by touch: an "expert system". *Percept. Psychophysics* 37, 299-302.

[7] Skedung, L., Arvidsson M., Chung, J. Y., Stafford, C. M., Berglund, B. and Rutland W. M. Feeling small: Exploring the tactile perception limits. *Sci Rep* 3, doi: 10.1038/srep02617

[8] Verrillo, R. T. 1963. Effect of contractor area on the vibrotactile threshold. *J. Acoust. Soc. Am.* 35, 1962-1966.

[9] Deb, A. 2015. Phantom vibration and phantom ringing among mobile phone users: A systematic review of literature. *Asia Pac Psychiatry* 7(3), 231-239.

[10] Kruger, D. J. and Djerf, J. M. 2016. High Ringxiety: Attachment Anxiety Predicts Experiences of Phantom Cell Phone Ringing. *Cyberpsychology, Behavior, and Social Networking* 19(1), 56-59.

[11] http://www.intechopen.com/books/advances-in-modern-woven-fabrics-technology/sensory-and-physiological-issues

[12] Mancini, F., Bauleo, A., Cole, J., Lui, F., Porro, C. A., Haggard, P. and Iannetti, G. D. 2014. Whole-Body Mapping of Spatial Acuity for Pain and Touch. Annals of Neurology 75(6), 917-924. http://onlinelibrary.wiley.com/doi/10.1002/ana.24179/pdf

[13] Taylor, J. L. 2013. Kinesthesic Inputs. in D.W. Pfaff (ed.) - *Neuroscience in the 21st Century*, Springer Science+Business Media, 931-964.

[14] Alary, F., Duquette, M., Goldstein, R., Chapman, C. E., Voss, P., La Buissonnière-Ariza, V. and Lepore, F. 2009. Tactile acuity in the blind: a closer look reveals superiority over the sighted in some but not all cutaneous tasks. *Neuropsychologia* 47(10), 2037-2043.

[15] Gallace, A., Tan, H. Z. and Spence, C. 2006. The failure to detect tactile change: A tactile analogue of visual change blindness. *Psychonomic Bulletin & Review* 13(2), 300-303.