Evaluating Somatosensory Interactions: Designing a Handheld Tactile Acoustic Device for Mobile Phones

Maria Karam[1(✉)] and Patrick M. Langdon[2]

[1] Kings College London, London, UK
maria.karam@kcl.ac.uk
[2] Department of Engineering, University of Cambridge, Cambridge, UK
pml24@cam.ac.uk

Abstract. We present the details of preliminary development efforts to create a tactile acoustic device (TAD) for the hands. The Model Human Cochlea (MHC) is a method for conveying sound to the body in a chair form factor, and originally developed as a sensory substitution system to provide some access to sounds from movies or music to deaf and hard of hearing people. We present initial design research on modifying the MHC system from a chair to a mobile handheld tactile device, towards improving mobile-phone speech comprehension in noisy environments. Scaling the MHC from the back, which has the least sensitive skin on the body, to the highly sensitive skin on the hands, requires an understanding of the physiology, psychology, electronics, and software applicable to this kind of sensation. This research addresses factors critical to expanding the design of tactile acoustic devices for somatosensory system interactions to different areas of the body.

Keywords: Interactive displays · Somatosensory systems · Crossmodal systems · Mobile phone interaction · Multimodal HCI systems · Sensory substitution · Tactile acoustic devices

1 Introduction and Background

The system described in this paper is a sensory augmentation approach to displaying sound on the body, modeled on the workings of the human cochlea and the sense of hearing. The Model Human Cochlea (MHC) (Karam et al., 2009) is applied to the body using the back of the torso and thighs as contact points for an array of transducers that register sound on the body. This work focuses on extending the MHC to other areas of the body, specifically the hands, where speech sounds may be experienced as tactile signals that could enhance mobile phone communications.

1.1 The MHC

The MHC is a chair-based system that provides an alternative means of experiencing entertainment-related sound vibrations, focusing on the enigmatic characteristics such as emotion, prosody, and even timbre, through the back (Russo et al., 2012). The MHC

emulates certain functions of the human cochlea on the body by using sounds as the vibrations mapped onto a tactile display, rather than using representations of sounds based on haptic stimulation (Gemperle et al., 2001, Wall and Brewster, 2006). The MHC was developed for use on a chair as an effective form factor and delivery system for watching movies or entertainment forms requiring seats. Because the back is one of the least sensitive areas of the body (non-glabrous or hairy skin), we hypothesize that applying a TAD on more sensitive area of the body can potentially improve tactile acoustic perception towards the identification of speech sounds. Mobile phones represent an obvious application for delivering sound to the body, where tactile acoustics can improve speech access for deaf and hard of hearing people, or even hearing people in noisy environments. To begin modifying the tactile sound system for use as a handheld device, we explored some of the critical factors identified in previous work on human somatosensory interactions (HSI) to help inform and extend the design of the existing entertainment seating MHC device for the hands (Karam and Langdon, 2015).

2   Human Somatosensory System Interactions (HSI)

The skin is an organ that can detect many types of sensations, including but not limited to heat, pressure, stretch, pain, and vibrations from sound. Research on the substitution and augmentation of sounds as vibrations dates back to the 1920s (Gault, 1927) towards developing a new way for deaf people to access and comprehend speech. There are two main types of layouts used in these kind of systems based on the literature: grid formation, or linear (spectral) array. The grid approach is very effective at communicating spatial information, while the spectral approach separates sound frequencies into bands, expressed on the body in a linear arrangement as vibrations. Different tactile devices can also lead to different effects to create the the desired sensations. Such approaches have been slow in uptake as everyday devices, as they often require extensive training and large, expensive equipment to process and power the signals. Further considerations that influence the design of HSI systems include the sensitivity of different parts of the skin to vibration, and the need to explore more than just haptic vibrations when designing tactile displays for the body.

Mechanical haptic vibrations stimulate more than just the cutaneous receptors, and are commonly used as notifications, semaphoric messages, spatial vibrational patterns (Gemperle et al., 2001, Back-y-Rita et al., 1969, Bach Y Rita et al., 1987, Wall and Brewster, 2006), and tactile speech communications (Brooks and Frost, 1983). This type of discrete haptic signal may not however be sufficiently complex to exploit the full potential of the somatosensory system in speech comprehension. While there has been some success in mapping word elements to vibrations (Brooks and Frost, 1983, Gault, 1927), sound vibrations may potentially be more effective at stimulating cutaneous sensors with more complex yet subtle vibrations (Russo et al., 2012).

2.1 HSI Framework

The somatosensory system is a complex network of neural mechanisms, cognitive processes, and responses that are connected to all human physical perceptions (Gault 1927; Gallace and Spence 2014) and represent an abundance of potential sensations to explore for HSI. The HSI framework is next applied to help expand the design of the MHC for mobile phone interactions on the hands (Karam and Langdon 2015).

Interaction Scenario. The HSI framework led us to initially consider a mobile phone case as the housing for the new design. While headphones and hands free interactions are commonly used for phone communications, this work considers a mobile phone interface as a practical approach to initial investigations of the MHC for the hands. Size and power constraints influenced the design of this project, which influenced our use of the existing processing system designed for the theatre to drive this research, which will primarily explore different transducer array layout, sizes, and positioning.

Physiology. The front part of the hand is one of the most sensitive areas of the body, (glabrous or non-hairy skin) capable of detecting fine details of texture and Vibratory discrimination thresholds for touch sensors are dependent on both frequency and amplitude of stimulation. Sensations are dependent on rapidly adapting and slow adapting mechanoreceptors (RA I, SA I, RA II, SA II), embedded in skin sensitive to frequencies between 0.50 and 1000 Hz and with receptive field sizes varying from 1–1000mm2. Touch sense has been identified with Pacinian Corpuscles, the RAII receptors, but in fact is known to be associated with all the glabrous skin afferents including Meissner corpuscles, Ruffini corpuscles and the Merkel complex cells. Amplitude in such studies is measured in mm of displacement rather than work done (power), and thresholds vary from 0,01 to 40 mm. The contact discrimination threshold (receptive field size) on the body varies from 0.7 to 100 mm, but this is undoubtedly modified by amplitude or vibratory power in psychophysical functions, as has been described (Gallace and Spence 2014; Karam and Langdon 2015).

Cognition. Using sound as vibrations may improve comprehension and detection of speech on the hands using the multimodal integration of speech and vibration, which may be easier to identify than haptic signals that represent sounds. Based on the familiarity of sound vibrations, we hypothesize that speech comprehension on the MHC will improve tactile sound detection and identification when it supports an audio signal that is distorted or masked. The higher levels of tactile sensitivity of the glabrous skin on the hand may also reveal additional signals that could not be detected on the non-glabrous parts of the body, potentially increasing comprehension of tactile sound.

Technology. Characteristics of the transducer size, number, arrangement, power requirements, form factors, materials, drivers, and processing characteristics pose technological challenges when scaling the MHC. Several prototypes were developed during this design process to help us explore the different transducer sizes and layout options, however, we did not develop new processing hardware for this work, but use the existing processing drivers and algorithms to evaluate the mobile devices.

3   Design Considerations

The smaller area of skin on the hands represents a challenge to designing multi-user form factors that support multi-channel transducer arrays, which must not disrupt the hand. Additionally, the variation in size and shape of different hand size limits the layout of spectral array, as does the number of channels we explore in this work. The original TAD system uses 16 voice coils, aligned along 2 rows on either side of the spine. Placement of the transducers aims to maximize the contact points to increase tactile acoustic resolution, while avoiding bone conduction or deep tissue vibrations used in haptic displays.

3.1   Interaction Design

The initial 8-transducer design was first considered using 1 cm diameter contact point transducers to fit into the phone case (Fig. 1a). A breakdown of the MHC chair suggests that the 8 × 2 spectral layout represents left and right sides of the spine, with upper and lower segments. This suggested 4 discrete segment mappings to consider when translating the system from the back to the hand, where the segments are closer together, (fingers, palm, wrist palm…), limiting the linear placement of the array of transducers. Additional layouts considered are shown below (Fig. 1b, c).

Fig. 1.   a: 8-channel layout; b: 4-channel case; c: Right hand edge design

3.2   Physiology: Sizing the Transducers

Transducers that support multi channel arrays for the hand were optimal at 1 cm to 2 cm in contactor diameter, leveraging the higher sensitivity of the hand, but using less power, but posing interesting questions in determining an optimal layout.

Body Segments.   Early studies on the MHC suggested that sound 'chunks' or perceptual units of sounds, like a musical composition, had to be located on the same linear segment of the body for users to easily map sounds to the vibrations, for example, arms, legs, torso (Karam et al., 2009). The current work further identifies segments based on skin

type: glabrous or non-glabrous skin, and joint separations. The segments - left/right, skin type, and position - are critical factors in determining optimal configurations of transducers to body segment. Further breakdown of the body can reveal additional design challenges, as with the fingers, where the segments are small, and required freedom of movement, unlike the larger, less functional arms and torso.

3.3   Form Factor

The mobile phone was chosen to support real time mobile phone interactions with the transducers. However, while we aimed to provide both sounds and vibrations to a user, this was not easily accomplished in a working mobile phone, where the headphone jack does not permit the sound to be split to another channel within the phone. A workaround was to use headphones and a signal splitter to allow us to use the sound for both audio and tactile displays.

3.4   Technology: Signal Processing and Power Requirements

Our unique design opportunities lie in understanding and leveraging the different sensitivities of the skin, enabling us to explore new designs of transducers that could be very low profile. However, for this work, we are focusing on supporting form factor, layout, and comprehension testing rather than designing new hardware. With the glabrous skin's higher sensitivity to vibrations, we can leverage the hands to support decreasing the power levels for the transducers, while maintaining enough vibration to effectively stimulate the skin. This will be explored next.

3.5   Cognition: Evaluating User Perceptions

While strong vibrations may be easily detected on the body, the finer vibrational information that relate to speech sounds are not detectable to all parts of the skin. Perceptual effects that are also found in audio perception appear to be present in the tactile acoustic domain: masking, individual tastes, and amplitude preferences still occur in the perception of tactile sound. Different frequencies may also have to be placed on the body in an optimal location to achieve maximum detection and perception.

To conduct a first set of tests, we selected a 4-channel setup, allowing us to use the minimal configuration of transducers, while still delivering the multi-channel MHC signal to the hand. We tested this initial prototypes for signals strength, layout, and resolution, to determine if there any feature of speech could be perceived on the hand. 14 volunteers were asked to try the system, and to provide initial user feedback using a mobile phone in a TAD case (Fig. 1b). Participants were asked to place one hand on the mobile TAD, which was initially set up on a table. An audio signal from the phone was redirected to the tactile transducers, which was not audible.

A radio talk show segment was used as the sound sample, with the vibrations divided up along the four-transducer array along the side of the phone case. The show featured a calm male host, speaking with an excited female caller. Each participant re-ported that they could feel voices, with most indicating that they could detect the sex of the voice,

and many indicating that they could the emotion of argument or persuasion. Some observed behavior we observed during the sessions included moving the phone around their hands to feel all the transducers, and placing the device by the ears to try and hear the vibrations. All participants also identified that the signal was speech.

4   Design Challenges

Participants tended to hold the phone prototype along the edges, possibly a reflection of the poorly designed form factor that requires the fingers to ensure the device doesn't slip out of the hands: this suggested that moving the transducers to the edges of the case would be more in consonance with the natural way that people hold their phones. Three transducers were moved to the one side, with a fourth set on the opposite side to improve contact with the fingers in a new prototype design (see Fig. 1c).

We had to consider handedness in this case, resulting in versions of the new device. This was deemed impractical, and the prototype was abandoned for the remainder of this study. The next prototype was a block of wood, which was modeled as a phone with both audio and tactile signal outputs (Fig. 2).

Fig. 2. Early prototyping showing non-finger handed design

For this version, we used a stereo signal to drive the transducers, which was somewhat limited in its use, but it did reveal that two channels alone have some effect, but constraining the transducers to different form factors did not allow us to properly gauge tactile perception of speech, nor to evaluate individual transducers and their effects on different areas of the hand.

Several other two-transducer versions were developed to support left-right hand interactions, and to isolate transducer size, number, and placement on the body without the physical constraints of a mobile phone form factor. These prototypes were secured in putty or silicone to protect the connections from the stress that mobile interactions placed on the transducers (Fig. 3a, b). Further studies can be run, but the system requires a more effective design to support a more functional form factor that won't constrain the hand movements or functionality.

Fig. 3 a and b. Early prototyping showing transducer pair in silicone

5   Discussion

After considering the current form factors and transducer sizes, the mobile phone form factor was abandoned for several reasons: First, people increasingly use headsets with Bluetooth connections to interact with their phones, reducing the amount of time holding the phone while in use. Second, it is impractical, from an engineering perspective, to embed transducers into phones as they already struggle with power consumption and additional hardware would further reduce battery life. Third, although the smaller sized transducers were somewhat effective at communicating some speech information to the hands, individual differences in behavior and approaches to mobile phone interactions suggested that a more universally accessible form factor would need to be developed to support multi-channel vibrations for the hand while on the move. Fourth, the shapes and sizes of mobile phones are not designed to ergonomically fit in the hand, and it became apparent that we would have to find an alternative form factor to better support tactile perception. Fifth, implementing a commercial TAD into a mobile phone would require a drastic modification in the hardware, and sound sources to be distributed to multiple channels to support the multi-modal interactions required by the TAD.

The current tests suggested that the vibrations provide some level of recognizable speech information to the hands, but did not work out as an effective form factor to support and effectively test the multichannel TAD for mobile phone interactions.

Shifts in the zeitgeist of end-user devices are suggesting that mobile phones may soon be replaced by more practical devices, in support of wireless, wearable devices that provide better ergonomics and usability for interactions with the somatosensory system. Further electronic processors and drivers will also be developed and designed into the system to improve HSI, and the design of these devices must be reimagined to enable them to utilize skin contact as a way to help offload some of the attention demands placed on the ears and eyes to the body without interfering with primary tasks. Some new models we are developing are shown in Fig. 4.

Fig. 4. 1 New form factors that are being developed to support future experiments

6   Conclusions and Future Work

We have begun to explore the extension of the MHC to include the hands, and other areas on the body, where the higher sensitivity of cutaneous receptors could increase tactile acoustic perception and offer more of the body's surface as locations for tactile information displays. The design exercise presented in this paper suggests that the MHC has the potential to communicate similar information to the hands, using smaller and fewer transducers than is required for the back. Adding more channels can also potentially lead to an increase in tactile acoustic resolution, further expanding on the principles behind the MHC. New prototypes will be developed to help explore different interaction paradigms and devices that are now being designed into fashion, jewelry, and other accessories, towards increasing access to and availability of the somatosensory system and the continued development of the study of HSI in new applications including way-finding, navigation, communication, and intimacy.

References

Bach-y-Rita, P., Webster, J.G., Tompkins, W.J., Crabb, T.: Sensory substitution for space gloves and space robots. In: Space Telerobotics Workshop, Jet Propulsion Laboratory, Pasadena, CA, 20–22 1987, pp. 51–57 (1987)

Bach-Y-Rita, P., Collins, C.C., Saunders, F.A., White, B., Scadden, L.: Vision substitution by tactile image projection. Nature **221**(5184), 963–964 (1969)

Brooks, P.L., Frost, B.J.: Evaluation of a tactile Vocoder for word recognition. J. Acoust. Soc. Am. **74**(1), 34–39 (1983)

Gallace, A., Spence, C.: In touch with the future: The sense of touch from cognitive neuroscience to virtual reality. OUP, Oxford (2014)

Gault, R.H.: Hearing through the sense organs of touch and vibration. J. Franklin Inst. **204**(3), 329–358 (1927)

Gemperle, F., Nathan O., Siewiorek, D.: Design of a wearable tactile display. In: Proceedings of the 5th IEEE International Symposium on Wearable Computers ISWC 2001. IEEE Computer Society, Washington (2001)

Karam, M., Langdon, P.: Seeing, hearing and feeling through the body: the emerging science of human-somatosensory interactions. In: Antona, M., Stephanidis, C. (eds.) UAHCI 2015. LNCS, vol. 9176, pp. 205–216. Springer, Heidelberg (2015)

Karam, M., Russo, F.A., Fels, D.I.: Designing the model human cochlea: an ambient crossmodal audio-tactile display. IEEE Trans. Haptics **2**(3), 160–169 (2009)

Meenowa, J., Hameed, M., Furner, S., Langdon, P.M.: Using haptic displays for workload reduction and inclusive control. In: 11th International Conference on Human Computer Interaction (HCII 2005), 2005-7-22 to 2005-7-27, Las Vegas, Nevada, USA (2005)

Russo, F.A., Ammirante, P., Fels, D.I.: Vibrotactile discrimination of musical timbre. J. Exp. Psychol. Hum. Percept. Perform. **38**, 822–826 (2012)

Wall, B., Brewster, S.: Feeling what you hear: tactile feedback for navigation of audio graphs. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2006), pp. 1123–1132. ACM, New York (2006)

Withana, A., Koyama, S., Saakes, D., Minamizawa, K., Inami, M., Nanayakkara, S.: RippleTouch: initial exploration of a wave resonant based full body haptic interface. In: Proceedings of the 6th Augmented Human International Conference (AH 2015), pp. 61–68. ACM, New York (2015)