
The Emoti-Chair: An Interactive Tactile Music Exhibit

Maria Karam

Ryerson University
350 Victoria St., Toronto
Ontario, Canada, M5B 2K3
Maria.karam@ryerson.ca

Carmen Branje

Ryerson University
350 Victoria St., Toronto
Ontario, Canada, M5B 2K3
cbranje@gmail.com

Gabe Nespoli

Ryerson University
350 Victoria St., Toronto
Ontario, Canada, M5B 2K3
gabe@psych.ryerson.ca

Norma Thompson

Ryerson University
350 Victoria St., Toronto
Ontario, Canada, M5B 2K3
n3thomps@ryerson.ca

Frank A. Russo

Ryerson University
350 Victoria St., Toronto
Ontario, Canada, M5B 2K3
russo@psych.ryerson.ca

Deborah I. Fels

Ryerson University
350 Victoria St., Toronto
Ontario, Canada, M5B 2K3
dfels@ryerson.ca

Abstract

The Emoti-Chair is a sensory substitution system that brings a high-resolution audio-tactile version of music to the body. The system can be used to improve music accessibility for deaf or hard of hearing people, while offering everyone the chance to experience sounds as tactile sensations. The model human cochlea (MHC) is the sensory substitution system that drives the Emoti-Chair. Music can be experienced as a tactile modality, revealing vibrations that originate from different instruments and sounds spanning the audio frequency spectrum along multiple points of the body. The system uses eight separate audio-tactile channels to deliver sound to the body, and provides an opportunity to experience a broad range of musical elements as physical vibrations.

Keywords

Haptic I/O, crossmodal displays, sensory substitution, assistive technologies

ACM Classification Keywords

H.5.2 Information Interfaces and Presentations: User interfaces – Haptic I/O

General Terms

Human Factors

Copyright is held by the author/owner(s).
CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Introduction

The Emoti-Chair presents an approach to translating sounds, including music, speech, and environmental noises, into physical vibrations. While the notion of 'feeling music' is not new, the concept of using the skin as a robust mechanism for experiencing sound offers exciting possibilities.

The model human cochlea (MHC) is a sensory substitution technique we use to convert sound into vibrations [6]. The main principle behind the MHC draws on the human cochlea as a design metaphor; essentially, the cochlea has thousands of tiny hair cells or *mechanoreceptors* that serve as sensors for specific frequencies when sounds enter the ear. Hair cells are ordered along the basilar membrane of the inner ear according to frequency: high frequencies are positioned near the oval window and low frequencies are positioned near the apex. This *place coding* is supplemented with a *time code* in which the rate of firing in the auditory nerve corresponds to the stimulating frequency. The MHC implements a similar dual code, with multiple vibrating devices (voice coils) ordered with respect to frequency along the body. When a signal is sent through one of the voice coils, it leads to vibrotactile stimulation that contains a place and time code. The place code pertains to the position of the coil and the time code pertains to the rate and intensity of the stimulating frequency. Effectively, this research aims to improve the way we feel music by increasing the available resolution of audio-tactile displays.

The number of hair cells in the cochlea is vastly greater than the number of voice coils used in the MHC —there are over 16 thousand hair cells, but only 16 voice coils

in the MHC —however, the same basic principle applies to both systems: the greater the number of receptor cells, the more fine-grained our sensory discriminations become. Thus, we can theoretically increase the audio-tactile resolution of the MHC by increasing the number of channels in the display. This can potentially lead to improved detection of sounds as tactile stimuli. Presently, the MHC uses eight discrete audio-tactile channels to translate sounds into vibrations, but this can be increased to a current maximum of 16 channels. Research is underway to investigate the effects of increasing the audio-tactile resolution of the MHC to 12 and 16 channels.

The eight-channel system used to date has shown a remarkable increase over a four-channel system in the ability to translate emotional content from sounds as comprehensive physical vibrations.

Audio-Tactile Resolution

Research on audio-tactile displays is relatively new, and many concepts require further exploration before we can more fully comprehend the implications of using the skin for high-resolution audio detection. As an introductory stage, our team is using an eight channel version of the MHC as a tool to support research on improving tactile perception through the simultaneous processing of multiple audio-tactile channels: by increasing the number of channels, we can effectively increase the audio-tactile resolution of the display, which can potentially improve the level of sound information that can be accessed through the skin. To illustrate the type of sensations that the audio-tactile system can provide, we have prepared a demonstration of the Emoti-Chair using the eight-channel MHC configuration.

The demonstration includes the Emoti-Chair control software, which supports fine tuning of the frequency distribution and volume levels for each of the eight channels. The system can accept audio input from iTunes, iPods, iPhones, or midi devices to enable users to experience different types of sound on the chair. The Emoti-Chair, the hardware components, the software application, and the configuration options are described next.

Background

Imagine placing your hand on an audio speaker when it is playing your favourite song. What you would find is that mainly the low frequency signals, such as the bass and rhythm sections can be detected. For people who are deaf, this type of vibration serves as the main method of accessing musical sound: however, this method also leads to an impoverished representation of music. To illustrate, try placing your hand on your throat as you speak, and notice what happens as you change the tone of your voice; the rate of vibration rises with rising pitch. As the pitch reaches the highest voice register, the vibration gets weaker, until we can no longer feel the sound.

Audio-Tactile Resolution

Sounds can be easily converted into physical sensations, which can also be made available as tactile stimuli. However, unlike feeling voices when we speak, feeling music that is comprised of multiple instruments and voices as individual sets of vibrations is not possible. This is because the audio signal originating from the voice, drums, bass, guitars, and piano are combined into an extremely complex wave-form intended for perception through the hearing organ.

While our ears can discriminate the individual sounds from the combined signal, our skin cannot.

Extending this notion to a stereo speaker system, if we try to feel the sounds with our hands, we can only access the vibrations resulting from the combined stronger, low frequency signals, which mask the weaker, higher tones. By applying the cochlea metaphor to the MHC design, we can isolate more of the signals from the combined audio source and make them available as discrete tactile sensations. Unlike the ear, which uses thousands of sensors to capture the complete spectrum of sound, the skin cannot naturally access the thousands of signals that comprise sound. Thus, to effectively leverage the skin's ability to detect vibration and "hear" with our skin, we must spatialise the frequency regions or sound sources if we expect to approximate a true representation of the sound signal.

This concept has led to the development of the multiple channel audio-tactile display system that translates sounds into physical vibrations. Our first system used only four channels of frequency signals, but the current version, which uses eight discrete audio-tactile channels, has shown that an increase in the number of vibrotactile channels can lead to an improved ability to detect emotional content through the skin [6]. To this end, the MHC represents an effective method of bringing sounds to the sense of touch, which can have profound implications for the deaf and hard of hearing communities. To date, we have only explored some of the basic characteristics of sound and its translation into physical sensations. These include emotional expressiveness [5], and timbre identification [9], while speech prosody has also shown to be detectable in pilot studies [2].

The System

There are three major components that make up the Emoti-Chair. The first component is the vibrotactile display system (the MHC), which is comprised of a two column by eight-row array of voice coils. The voice coil is currently the most effective device we have found for supporting the direct translation of sound onto the tactile display. Because voice coils are directly driven by audio signals, they can effectively represent and control frequency and amplitude information from sound. Unlike motors, which are commonly used for vibrotactile displays and that only offer frequency controls, voice coils are a direct source for creating physical vibrations derived from frequency and amplitude information.

Signal Processing

The second component enables the sensory substitution of the audio signal, which separates the multiple discrete channels of sound displayed on the MHC. An eight channel sound card supports processing and control of the multiple audio signals sent to the amplification system that drives the voice coils. The sound card can process digital or analog signals originating from midi instruments, digital audio, or live audio feeds, which are then divided into multiple frequency bands and allocated to specific voice coil channels. We are using a variety of eight channel sound cards, including the Motu 828MKII [7], the Presonus Firestudio [8], and the ESI Gigaport HD [4]. Audio signals for the system can be accessed through XLR, ¼", ½", RCA, or SPDIF cables, depending on the sound card.

Signal Processing

Signals can be processed using either a frequency model, which separates a single audio source into multiple frequency bands, or a track model, which uses separate audio tracks for each channel. The track model presents individual instrument tracks, voice recordings, or live instruments to different channels of voice coils, creating a clearer sense of what each part of the music feels like [5]. Although the track model was shown to be most effective for translating sound into vibrations in previous studies, the frequency model is more useful as a research tool since it can effectively translate any source of audio onto the tactile display.

Signal Amplification

The third component is the amplification system, which uses the multiple audio signals to drive the voice coils. An amplified signal is required for each of the separate channels of audio output sent to the display. The current prototype uses two Pyramid PB442X Super Blue 4x35W Amplifiers to drive the eight-channel system, but 7.1 surround sound systems can also be used to deliver the signals to the individual channels of the MHC.

We have also developed a dedicated circuit board that replaces the sound card and amplifiers. The circuit divides the audio signal into eight frequency bands and amplifies each channel for presentation on the tactile display. It accepts only a single source of sound input and eliminates the need for software or external hardware components, offering only volume control for each channel. While the circuit model greatly simplifies the system for end user access, the highly configurable complete system, including hardware and software, is required for research purposes.

The Software

The Emoti-Chair control software was developed using Cycling 74's Max/MSP platform for audio processing [3] and Adobe Flex for the user interface components [1]. The software was originally intended for research purposes, with support for manual frequency distribution, octave step pitch shifting, and gain controls for each of the eight channels. The interface has been simplified for end-user access, with more complex functions included as advanced features.

Interface Design

The software interface displays two large buttons, a green START and a red STOP, and a main volume controller, which includes a slider and a row of buttons that jump to pre-set volume levels. An advanced volume control can be expanded using a check box, which reveals volume sliders for each of the eight channels: this functions like a graphic equalizer, supporting individual volume settings for each channel, and maintaining these levels even when the main volume is adjusted. The gain control panel can also be expanded to reveal independent volume level controls for each channel to avoid peak clipping, which can lead to uncomfortable vibrations.

A pitch shift feature is also included in this panel, which can lower the signal of a high frequency channel by a single-octave. This makes very high frequency ranges, (those over 2 kHz), more easily detected as physical vibrations without drastically altering the melody or flow of the music. Another advanced feature of the software supports configuration of the frequency bands that are sent to each channel.

Several frequency distribution options are included as pre-set buttons on the main screen. These include a *manual* button that reveals a new window where each channel can be assigned a unique frequency band, saved, and then reloaded for future use.

The manual frequency distribution option enables customization of the way that an audio signal is distributed across the multiple channels, and is essential for exploring different genres of music on the MHC. For example, signals from the Drum'N'Bass genre tend to be in the lower frequency range, while a flute concerto maintains an overall higher frequency range: this requires a different frequency distribution for each genre if we wish to maintain a proportional distribution of signals across the channels.

The default frequency distribution may be effective at providing a general translation of sounds to the MHC, however, for research purposes, we are interested in exercising greater control over the way the signals are distributed to the multiple channels.

Conclusion

The Emoti-Chair is a high-resolution audio-tactile display that supports the translation of sound into tactile vibrations. The software extension is highly configurable, enabling researchers from different disciplines to use the Emoti-Chair as a tool for investigating sensory substitution of sound as tactile vibrations. Although the Emoti-Chair is still a prototype, the underlying functionality of the MHC can serve as an effective tool for investigating sound as vibrations. We have also obtained interesting reports from end users, who have suggested that higher quality sounds are noticeably distinguishable from those of lower quality.

Interference such as static, clipping, or other distortions have proved to be as disrupting to the tactile enjoyment of the music as they are to the audio experience. This suggests that there are additional characteristics of sound that are being transferred directly onto the tactile display and detected by the skin. In addition, deaf and hard of hearing individuals who have been using the Emoti-Chair for longer periods of time have also begun to develop a preference for

specific music genres based on their experience with the tactile vibrations. These and other findings are motivating future research on the MHC, which aims to continue explorations and evaluations that address increasing the audio-tactile resolution by using more channels, improving the quality of the audio-tactile signals, and developing our skin as a pseudo-hearing organ.

References

- [1] Adobe. Flex. Web Resource, 2010. <http://www.adobe.com/products/flex/>
- [2] C. Branje, M. Maksimowski, G. Nespoli, M. Karam, D. I. Fels, F.A. Russo. Development and validation of a sensory-substitution technology for music. In *Canadian Acoustics* 37 (2009), 186-187.
- [3] Cycling 74. Web resource, 2010. <http://cycling74.com/>
- [4] ESI. Gigaport HD. Web Resource, 2010. <http://www.esi-audio.com/products/gigaportag/>
- [5] M. Karam, F. A. Russo, C. Branje, E. Price, D.I.fels. Towards a model human cochlea: Sensory substitution for crossmodal audio-tactile displays. In *Proc. GI'08* (2008), 267-274.
- [6] M. Karam, F.A. Russo, D.I. Fels. Designing the model human cochlea: An ambient crossmodal audio-tactile display. *IEEE Trans. Haptics*, 2(3) (2009), 160-169.
- [7] MOTU. Motu 828mkII. Web Resource,2010. <http://www.motu.com/products/motuaudio/828mkII>
- [8] Presonus. Firestudio. Web Resource, 2010. <http://www.presonus.com/products/>
- [9] F.A. Russo, M. Maksimowski, M. Karam, D.I. Fels. Vibrotactile discrimination of musical timbre. Presentation at Society for Music Perception and Cognition (2009)