ENHANCING MUSICAL EXPERIENCE FOR THE HEARING-IMPAIRED USING VISUAL AND HAPTIC DISPLAYS

SURANGA CHANDIMA NANAYAKKARA

NATIONAL UNIVERSITY OF SINGAPORE

2009

ENHANCING MUSICAL EXPERIENCE FOR THE HEARING-IMPAIRED USING VISUAL AND HAPTIC DISPLAYS

SURANGA CHANDIMA NANAYAKKARA

BEng (Hons), NUS

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2009

Acknowledgements

Writing a dissertation is really a collaborative effort. Although there is only one name on the title page, that name stands for a huge network of colleagues and close friends. This work would have been impossible or unbearable without them. So I would forthwith like to thank everyone who has made this thesis possible, bearable, or both.

I was blessed with three outstanding research faculties to supervise and guide me through this research over the past four years—Dr. Elizabeth Taylor (Head of the Marine Mammal Research Laboratory, Tropical Marine Science Institute), Associate Professor Ong Sim Heng (Head of the Biomedical Engineering Group, Department of Electrical and Computer Engineering) and Associate Professor Lonce Wyse (Director of the Arts and Creativity Laboratory, Interactive and Digital Media Institute).

More like a mother than my supervisor, Dr. Taylor, you deserve my heartiest gratitude. Your acute intuitions, expansive experience got me started on the right track and inspired me all along these four years. It is your support and guidance that made this all possible. You were the force behind every success in my PhD. It is your indomitable energy that kept me going. You believed in me, which helped me gain that extra confidence. I am indebted for you for life.

Prof. Ong, you believed in me and agreed to become my supervisor from the Department of Electrical and Computer Engineering (ECE) when I needed to have a supervisor from ECE. You were always there to provide me valuable advice and guidance.

Prof. Wyse, your thoughtful suggestions and great ideas always opened up new avenues to think through. Thanks for being friendly, approachable and supportive all the way through. Although you joined only during the second year of my PhD, your inputs have immensely shaped the research ever since.

In addition to my supervisors, there were others who made noteworthy contributions to this project. I would like to thank Associate Professor Elaine Chew (University of Southern California) for her technical contributions as well as supervising my work during the four month attachment at the Music Computation and Cognition (MuCoaCo) Laboratory, Integrated Media Systems Centre at the University of Southern California; Dr. Nigel Helyer, whom I met during the international Symposium on Electronic Art (ISEA) 2008, for sharing his valuable ideas; John Reading, a local artist, for providing feedback about an early version of the visual display.

The great company and excellent support provided by my colleagues who worked with me at various stages of the project is never forgotten. I am truly grateful to Jolyn Tan and Yeo Kian Peen for their help and technical support at various phases of the project; Dr. Paul Seekings, for his advices and comments; Tim Merritt and Nicolas Escoffier, for discussions on conducting user studies; Norikazu Mitani for assistance in recording and extracting videos; and Hiroki Nishino for help provided in debugging programs.

When it comes to conducting user studies, I would take this opportunity to gratefully acknowledge the support of Mrs. Maud Senaratne from the National Council for the Deaf, Sri Lanka; Mr. Yeo Ando from Singapore Association for the Deaf, Singapore and Mr. Robert Sidansky from the National Centre on Deafness, California State University, USA for assisting with the initial background survey. I would also like to thank the wonderful people at Dr. Reijntje's School for the Deaf in Moratuwa, Sri Lanka for providing me the opportunity to continuously interact with the deaf students throughout such an extensive user studies. I am very grateful to Mrs. Tineke De Silva–Nijkamp, principal of the school, for her exceptional enthusiasm and continuous help without which this research project would not have been possible. I should also mention with gratitude, Mrs. Buddhini Gunasighe, the speech therapist/sign language interpreter, for facilitating the interaction with deaf students and supporting during the user studies. I would also like to thank Muditha De Silva and Pathum Madushan for their support from time to time; the teachers and supporting staff for their cooperation in successfully conducting the studies; and on top of all, the students for continuously taking part in this study and their parents for providing their consent.

My heartiest gratitude goes to the deaf musicians, Mr. Azariah Tan and Miss Lily Goh, for their invaluable feedback; Sebastian Tan and Siddharth Jain for help in measuring frequency response of the Haptic Chair and 3D implementation of visual effects as part of their final year project work. Mr. Chong Wai Lun, administrative officer and projects coordinator of Yong Siew Toh Convervatory of Music, your support is much appreciated.

I wish to mention with gratitude, my friends at NUS Dr. Newton Fernando, Charith Fernando, Roshan Peiris, Sameera Kodagoda and Dr. Namunu Maddage for their valuable inputs to the project from time to time; Janaka Wijesinghe, for the help given in debugging Flash Action script codes; Bhagya Amarasinghe, for helping with logistics in taking electronics equipment to Sri Lanka for user studies; Chamithri Vidyarathne for assistance in drawing 3D models of the Haptic Chair; and Eshani Motha for performing for the videos used in some of the user studies. Avanthi Ratnayake, Udanie Salgado and Chinthaka Perera deserve special thanks for their proof-reading skills. I am sure they now know more about this topic now than they ever wanted to. I am grateful to my good friend, Milinda Dharmadasa for introducing me to the Dr Reijntje's School for the Deaf in Moratuwa, Sri Lanka, without which this research wouldn't have been this fruitful.

My dear mom, this is for you—for the countless times you stood beside me and gave me strength. You were my only hope when I felt like giving up. I hope I have made you proud. Without you this work would never have come into existence. I wish to thank my dad, brother and sister for their love and support throughout the years. I would like to thank my cousin Dimuthu Makawita and his wife Narmada Makawita, who were always my family away from home. Thanks for always being there, especially when times were bad and making Singapore feel like home. My closest and dearest friends, especially Melani Jayasuriya, thanks for bearing up with me all these years and being there for me in good times and bad.

The list goes on and on, as four years of this project has acquired loads of debt and as Dr. Taylor would say *it's not easy to see or feel sound*.

Thanks all!

National University of Singapore Suranga Nanayakkara

31 July 2009

Dedication

To my mother, Manel Nanayakkara

Summary

Music is a multi-sensory experience informed by much more than hearing alone and thus could be made accessible to people with most variations in hearing ability. Little guidance is available in the existing literature since few researchers have explored the specific question of enhancing the musical experience of hearing-impaired people.

This dissertation addressed the broad question of understanding whether and how a combination of haptic and visual information could be used to enhance the experience of music by the hearing-impaired. Initially, a background survey was conducted with deaf people from multi-ethnic backgrounds to find out the techniques they used to "listen" to music and how their listening experience can be enhanced. Information obtained from this survey and feedback received from two profoundly deaf musicians was used to guide the initial concept of exploring haptic and visual channels to augment or convey a musical experience.

The proposed solution had two main components—a vibrating "Haptic Chair" and a computer display of informative visual effects. The Haptic Chair provided sensory input of vibrations via touch by amplifying vibrations produced by music. Although this seemed to be simple, it worked well due to the fact that the hearing-impaired are used to sensing vibrations when listening to music. The visual display initially consisted of abstract animations corresponding to specific features of music such as beat, note onset, tonal context and so forth. Since most of the hearing-impaired place a lot of emphasis on lip-reading and body gestures, their experiences were also explored when they were exposed to human gestures corresponding to musical input.

Rigorous user studies with hearing-impaired participants suggested that the prototype system enhanced their musical experience. Most users preferred watching human gestures synchronised to music rather than watching abstract animations. They were very sensitive to any visual effect asynchronised to music and expressed their dislike of this. All the hearing-impaired users preferred either the Haptic Chair alone or the Haptic Chair accompanied by the visual display. These results were further strengthened by the fact that user satisfaction was maintained even after regular use over a period of three weeks. One of the comments received from one deaf user when the Haptic Chair was taken away (*I am going to be deaf again*), poignantly expressed the level of impact it had made.

During the course of our research, we kept seeing evidence which suggested that people can detect vibrotactile stimuli of higher frequencies. This led us to study the sensory abilities of people with normal hearing and those with hearing impairments using open-hand contact with a flat vibrating surface that represented 'real-world situations'. To explore a more complete range of vibrotactile sensory input we used complex signals in addition to sine tones. Sensitivity to vibrotactile frequencies at least up to 4 kHz, (two octaves higher than previously reported) was demonstrated for all signal types. We also found that complex signals are more easily detected than sine tones, especially for low fundamental frequencies. These findings are applicable to a better understanding of sensory biology, the development of new sensory devices for the hearing-impaired, and to the improvement of human-computer interaction where haptic displays are used.

Apart from enhancing the musical experience of a deaf person, the system described here has the potential to be a valuable aid for speech therapy. A user study is being carried out to explore the effectiveness of the Haptic Chair for speech therapy. It is also expected that the concepts presented in this dissertation would be useful in converting other types of sounds in the environment into a visual display and/or a tactile input device that might, for example, enable a deaf person to hear a doorbell ring, or footsteps approaching from behind, or the fact that a person is calling him, etc. Moreover, the prototype system could be used as an aid in learning to play a musical instrument or to sing in tune.

The findings presented in this dissertation could serve as a valuable knowledge base to researchers in the field of Human Computer Interaction (HCI) in developing systems for the hearing-impaired. This research work has shown great potential in using new technology to significantly change the way the deaf community experiences music.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Consider the kinds of listening behaviours that typical non-musically trained listeners with normal hearing engage in as part of everyday life. Such listeners can tap their foot or otherwise move rhythmically in response to a musical stimulus. They can quickly articulate whether the piece of music is in a familiar style, and whether it is a style they like. If they are familiar with the music, they might be able to identify the composer and/or performers. The listeners can recognise the instruments they hear being played. They can immediately assess stylistic and emotional aspects of the music, including whether it is loud or not, complicated, sad, fast, soothing, or generates a feeling of anxiety. They can also make complicated socio-cultural judgments, such as identifying a friend who would like the music or a social occasion for which it is appropriate and even share a musical experience with a friend.

Now, if the listeners are hearing-impaired[1], what would their listening behaviour be? Partial or profound lack of hearing makes the other ways humans use to sense sound in the environment much more important for the deaf than for people with normal hearing. Sound transmitted through the air and through other physical media such as floors, walls, chairs and machines act on the entire human body, not just the ears, and play an important role in the perception of music and environmental events for almost all people, but in particular for the deaf. Music being a multi-sensory experience, should not keep the hearing-impaired from enjoying it as much as a person with normal hearing does. However, little has been addressed in research so far about the issue of how to optimise a musical experience for a deaf person. This dissertation describes the design and evaluation of a system developed for conceptualising approaches that move us towards understanding how best to provide musical sensory enhancement for the deaf.

[1] The terms "hearing-impaired", "hard of hearing" and "deaf" are used interchangeably in this dissertation. These loosely defined terms are used to refer to a person whose primary mode of accessing sounds is not through the "conventional" hearing route, an air-filled external ear canal. The hearing loss is typically measured by an audiometer which measures the loss in decibels (dB) at different frequencies when listening to sound through external ear canals. Depending on the degree of the hearing loss, hearing-impaired people are typically classified as: mildly deaf, moderately deaf, severely deaf and profoundly deaf. However, in this dissertation, we considered two broad categories:

- Profoundly deaf (hearing loss of 95 dB or more)
- Partially deaf (hearing loss ranging from 25 to 95 dB)

More information about deafness can be found in the Royal National Institute for Deaf People (RNID) website: http://www.rnid.org.uk/information_resources/aboutdeafness/

1.1 Motivation

Conducting doctoral research to find out how best we can assist the hearing-impaired to sense music was a golden opportunity for the author to fulfil a lifelong dream of helping those who perceive the world differently.

Music is the time-based art of sound. Listeners bring a host of cultural and personal musical experience to bear when listening to a piece of music. Statistical regularities among a set of twelve tones are the fundamental blocks on which the structural regularities in western tonal music is based. A chord is defined as the simultaneous playing of three tones, while a subset of seven tones and chords generated from them defines a key [1]. Conventions of melodic patterns, chord sequences and key changes are exploited to create an intellectual and emotional response that we call the musical experience. A question central to this research is whether or not the musical experience can be conveyed by sensory channels other than sound. It is not just musical "information" that we want to convey, but the musical "experience". For example, rhythm is physical and is related to a sense of movement or dance. Can the difference in experience we have when listening to a march compared with listening to a waltz be conveyed using a visual method of presentation? Are there relationships between sound and graphics such that some mappings work better than others? How do we make use of other sensory inputs such as tactile[2] channels and hearing through bone conduction of sound to provide a more satisfying musical experience?

Some previous work has been done on providing awareness of environmental sounds to deaf people [2, 3]. However, no guidance is available to address the

[2] In this dissertation, the term "tactile" refers to information perceptible by touch. Tactile channel refers to the path by which tactile stimulation reaches the central nervous system from receptors in the skin.

challenges encountered at the early stage of designing a system for the deaf to facilitate a deeper experience of music. In order to keep the focus on the musical experience for the deaf and minimise potential bias from assumptions about musical experiences of hearing people, it was imperative to involve hearing-impaired people in the design loop from the beginning. Therefore, as a starting point, a survey was conducted to gather information from deaf people about how and how much they engage in music related activities and how to augment their musical experience. Based on the results of this survey, we implemented a prototype system which had two components: a "Haptic[3] Chair" that vibrates with the music; and a computer display that generates different visual effects based on musical features. Once the initial prototype was ready, possible improvements were explored through extensive feedback from hearing-impaired users.

Because people naturally sense musically derived vibrations throughout the body when experiencing music, any additional "information" delivered through this channel might actually disrupt the musical experience, and this confounding effect is potentially more significant for the deaf. There is so much we still do not know about the brain and its ability to integrate different natural stimuli to replace missing information [4]—for example, using natural occurring tactile stimuli to replace missing auditory stimuli for those who are profoundly deaf. Since we know that the human central nervous system (CNS) is particularly plastic in its intake of various sensory inputs and production of often different sensory output, it is important to support this ability to create new sensory experiences for people with specific sensory impairments. The human CNS is still largely a "black box" in data processing terms and it would be

[3] Haptic refers to the sense of touch. A haptic device typically has a physical contact between the device and the user. Substantial information on haptic interactions can be found in the website of the Multimodal Interaction Group at the University of Glasgow: http://www.dcs.gla.ac.uk/~steven/haptics.htm.

unforgivable to assume we can create a computerised system to replace its many and various abilities. Therefore, it was decided not to alter the natural vibrations caused by musical sounds, but to design a prototype Haptic Chair to deliver the natural vibrations produced by music tactilely via different parts of the Haptic Chair. Preliminary tests suggested that the Haptic Chair was capable of providing not only haptic (or touch) sensory input but also bone conduction of sound.

1.2 An Inter-disciplinary Effort

Human Computer Interaction (HCI), by definition, is an inter-disciplinary study which requires understanding of computer systems, the human user and the task to be performed by the user. Understanding of systems needs knowledge of engineering, programming languages, input/output devices and so on while appreciation of human behaviour, social interaction, environment and attitude are among the few essentials to understand users and their needs. Identification of the task, the reason for performing it and the characteristics of the environment are required to get a better understanding of the task being performed. Faulkner [5] has discussed the variety of disciplines and their contributions to HCI as shown in Figure 1.1. He describes the figure as follows:

> *...each of the areas is suitable for study in its own right, so to take them on board at the same time as studying HCI would require an unacceptable effort on the part of an average person.*

Nevertheless, this dissertation makes a considerable effort to combine the knowledge from different domains (particularly engineering, computing, design, psychology, etc.) to understand and enhance the way the hearing-impaired community experiences music. Thus, shaping of the design process was influenced by the hearing-impaired community. They were constantly kept in the design loop through interviews, on-site observations and questionnaires; and prototypes were built to demonstrate design concepts. This applied approach allowed us to understand the cross-modal[4] interactions between haptic, visual and auditory channels in providing a musical experience to a hearing-impaired person. This understanding would not have been possible to be derived from a purely theoretical perspective.

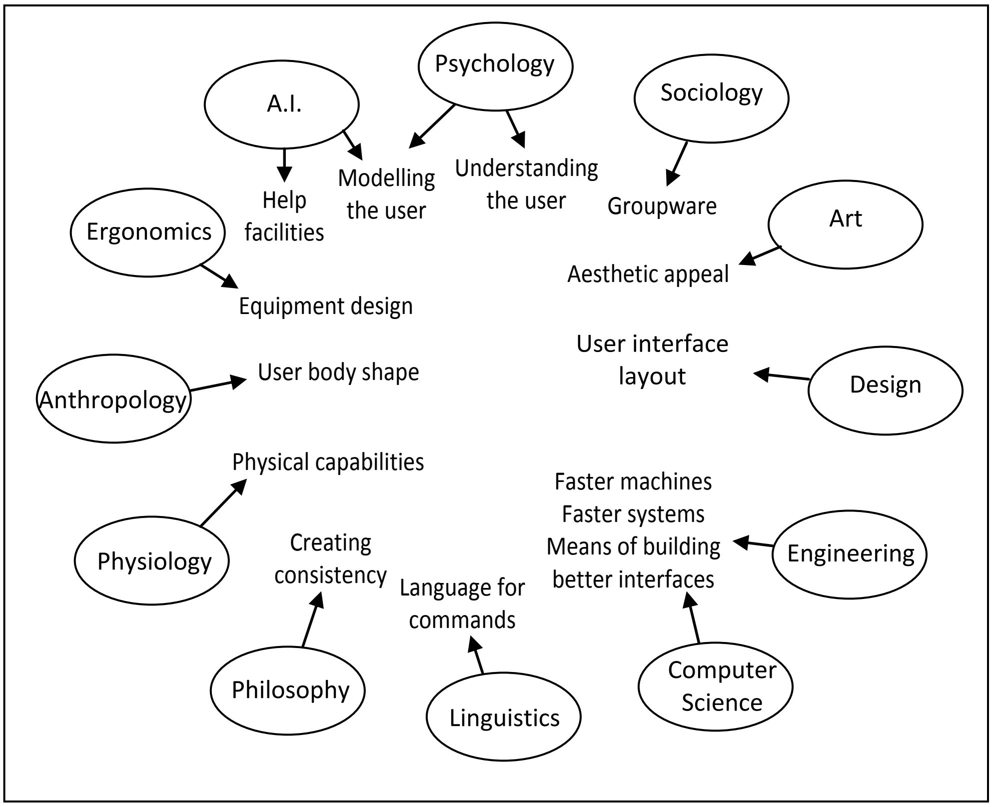


Figure 1.1: Various disciplines and their contributions to HCI (*image taken from [5]*)

1.3 Contributions

We hope this thesis would have contributed to the field of HCI and directly to the hearing-impaired community. More specifically, this thesis describes the design and user evaluation of a Haptic Chair and Visual Display which together provide a

[4] Cross-modal refers to interactions between two or more different sensory modalities. Section 2.2 includes a discussion of cross-modal interactions.

multi-sensory input system aimed at enhancing the musical experience of people with hearing impairments.

The experimental results and conclusions drawn from the extensive surveys done with the hearing-impaired from 3 different groups of multi-ethnic subjects as well as the numerous user studies performed on the system throughout its development stages provided significant understanding of the perceptual, cognitive, and behavioural capabilities and user interactivity of the hearing-impaired. This provided an important knowledge base for HCI researchers to develop more sophisticated systems for the hearing-impaired. Current literature in the field faces a significant shortage of materials to guide the design of a system to enhance the musical experience of the deaf. Thus the author expects that future research in this discipline will benefit from the work presented in this thesis.

Regular interactions with potential users resulted in new methods of using existing technologies for an enhanced experience. We found that both hearing and hearing-impaired participants can feel the vibrations at least up to 4 kHz depending on complexity and amplitude of the signal. This finding made a contribution to the literature where cut-off frequency of 1 kHz has often been accepted as canon, and used for haptic design purposes. Moreover, with rigorous user studies, it was shown that the concepts presented in this dissertation clearly enabled a new user experience, and revealed how it can affect the way the hearing-impaired community can sense music.

Apart from the scientific contributions, this work showed the potential to bridge the gap between the normal hearing and hearing-impaired communities. This was very much evident from the feedback and observations received from the deaf participants throughout the research project.

1.4 Overview of Thesis Content

The contents of this thesis are organised as follows:

Chapter 1 gives a general overview and introduction to the thesis, and describes the motivation, approach and major contributions.

Chapter 2 includes a critical assessment of related work and discusses its relationship to the proposed work.

Chapter 3 contains an analysis of a survey conducted with deaf people to make important design decisions at an appropriately early stage of the research. In addition, this chapter describes the design and implementation of a system that transcodes sequences of musical information into a sequence of visual effects in real-time.

Chapter 4 describes the initial design and evaluation of a prototype system consisting of a visual display and a Haptic Chair which is aimed to provide an enhanced musical experience for the hearing-imapired.

Chapter 5 reviews the initial design of the prototype system, particularly exploring the different methods of presenting visual cues. Details of user studies designed to evaluate the effectiveness of the revised prototype are also given in this chapter.

Chapter 6 concludes the dissertation by summarising the findings and outlining possible future work.

Chapter 2

Background

This chapter reviews appropriate topics from both scientific and other referenced sources of literature pertinent to this research so as to put the work presented in the next few chapters in perspective. Since the main focus is on supporting the abilities of the deaf to experience music, the review starts with a discussion of music and the deaf. A question related to this research is whether or not a musical experience can be conveyed by sensory channels other than sound. To answer this, a basic understanding of human cross-modal sensory perception is essential. Thus, a short review of cross-modal interactions is given. This is followed by a review of related work from a wide range of music visualisations with the focus on different strategies used for audio-to-visual mappings; feeling sound through tactile channels and bone conduction of sound are also reviewed.

2.1 Music and the Deaf

Profoundly deaf musicians and those with less profound hearing problems have demonstrated that deafness is not a barrier to musical participation and creativity. Dame Evelyn Glennie is a world renowned percussionist who has been profoundly deaf since the age of 12 but “feels” the pitch of her concert drums and xylophone and the flow of a piece of music through different parts of her body from fingertips to feet [6]. According to her, feeling vibrations through the body is as good as normal hearing. Following is a quotation taken from one of Glennie’s essays [6]:

> *Hearing is basically a specialized form of touch. Sound is simply vibrating air which the ear picks up and converts to electrical signals, which are then interpreted by the brain. The sense of hearing is not the only sense that can do this, touch can do this too. ... For some reason we tend to make a distinction between hearing a sound and feeling a vibration, in reality they are the same thing. ... Deafness does not mean that you can’t hear, only that there is something wrong with the ears. Even someone who is totally deaf can still hear/feel sounds.*

This suggests that a deaf person might be able to appreciate music if we could find a way to enhance the vibrations produced by musical sounds and let the person feel them. To explore this idea, a structure that could provide vibrotactile feedback to the whole body had to be developed.

On the other hand, Glennie thinks that her brain could create corresponding sounds by seeing things moving;

> *... There is one other element to the equation, sight. We can also see items move and vibrate. If I see a drum head or cymbal vibrate or even see the leaves of a tree moving in the wind then subconsciously my brain creates a corresponding sound.*

This statement is encouraging because it suggests that with a visual display generated by music, a hearing-impaired person might be able to make sense of the music being played.

Other examples of deaf musicians include the profoundly deaf Shawn Dale–the first and only person born completely deaf who achieved a “top ten hit” on Music Television (MTV) [7]; Beethoven, the German composer who gradually lost his hearing in mid-life but who continued to compose music by increasingly concentrating on feeling vibrations from his pianoforte; Azariah Tan, the seventeen-year-old severely deaf (with 15% of hearing remaining) pianist (currently pursuing a bachelor’s degree in music at the National University of Singapore) who has won many prizes at various competitions including the Trinity College London Music Competition; and Lily Goh, a profoundly deaf percussionist who teaches percussion music and other performing styles at Canossian School, School for the Deaf and Mountbatten Vocational School in Singapore. Both Azariah and Lily gave valuable feedback at various stages of this research work. A list of famous deaf musicians can be found in [8].

2.2 Cross-modal Interactions

2.2.1 Integration of visual and auditory information

It is widely accepted that our brain combines information from all available senses to form a coherent perception of the environment. The eyes and ears, which are used as the first stage of human perception, are perceived by the brain in two different regions, i.e., hearing is generally regarded to be processed in the temporal lobe while the occipital lobe is responsible for the sense of vision. We explore how visual input, which we perceive by the sensation of sight, can be used to stimulate the part of the brain that processes the sense of hearing.

Integration of audio and visual information is very much universal; in fact, majority of people combine audio and visual information while having face-to-face conversations. Other examples include watching movies or attending concerts. In addition, as discussed below, a number of phenomena such as the ventriloquism effect [9], McGurk effect [10] and synesthesia [11] show how auditory information can mutually reinforce or modify one another.

The term "ventriloquism" is generally used to describe integration of auditory-visual spatially. Ventriloquism explains the situation in which perceived location of the auditory source is shifted toward a visual cue [9]. For example, when watching television, the visual source captures the sound and it appears as if the voices and other sound effects are coming from the video source itself.

The McGurk effect demonstrates the interaction between hearing and vision in speech perception [10]. For example, visual “ga” (a video shows phoneme “ga”) combined with audio “ba” (sound of phoneme “ba”) is generally perceived as phoneme “da”. This effect shows that visual cues can fundamentally change what the listener hears.

Seeing colours when listening to sounds is one type of “synesthesia”, a condition in which the real information of one sense is accompanied by a perception in another sense [11]. For synesthetes (those who have synesthesia), listening to music is not just an auditory experience. Individuals who have music-to-colour synesthesia experience colours in response to tones or other aspects of musical stimuli (e.g. timbre or key). An excellent review of synaesthesia and music can be found in [12]. This phenomenon, on the surface value, seems to be important for this research because it can give clues to sound-to-visual mappings that are naturally meaningful and therefore potentially useful for conveying musical experiences. If so, these synesthetic experiences might be useful guidelines to map audio data to visual features since specific audio features tend to be associated with specific visual features. However, there is rarely any agreement among synesthetes that a given tone will be of a certain colour. In other words, the perceptions of different synesthetes often differs [13-15]. It has not yet been established how people are in fact able to “see music”.

Taylor [16] suggested that the human perirhinal cortex helps to bind the major aspects of audio-visual features to provide meaningful multi-modal representations. It can thus be established that even though humans perceive senses using distinct sensory pathways, the information is not perceived independently, but put together before being processed. Based on several brain imaging techniques [16], it is evident that visual input can be made to stimulate the auditory cortex of the brain, thus inducing brain responses that are experienced while hearing.

It is particularly important for the purpose of this dissertation to note that audio-visual integration influences many structural aspects of music experience. For example, the perceived duration of a note is affected by the length of the gesture used to create the note when the audience can see the performer while listening [17]. Integration of audio and visual information serves to extend the sense of phrasing and to anticipate changes in emotional content [18]. Thompson *et al.* [19] have shown that facial expressions of a singer can significantly influence the judgement of emotions in music. Moreover, Wanderley [20] has shown that the expressive gestures of a professional clarinetist could influence the perception of structure and affect in musical performance.

Referring to the above with regard to people with hearing impairment, exploring the visual mode may be one of the ways to compensate for the lack of auditory information. This was explored and several methods have been discussed and evaluated to represent music in visual form in order to offer the hearing-impaired community an enhanced mode to enjoy music.

2.2.2 Integration of touch and sound

Shibata [21] found that some deaf people process vibrations sensed via touch in the part of the brain used by most people for hearing. According to Kayser *et al.* [22], tactile sensation stimulates portions of the auditory cortex in addition to the somatosensory cortex. These findings provide one possible explanation for how deaf musicians can sense music, and how deaf people can enjoy concerts and other musical events. In addition they suggest that a mechanism to physically "feel" music might provide an experience to a hearing-impaired person that is qualitatively similar to the experience a person with normal hearing has while listening to music. However, this concept has not been utilised to optimise the musical experience of a deaf person.

Reed [23] demonstrated that with sufficient training, blind and deaf Tadoma method practitioners are able to utilise tactual sensations to support speech and language processing. In the Tadoma method, the hand of the deaf-blind individual is placed over the face and neck of the person who is speaking such that the thumb rests lightly on the lips and the fingers fan out over the cheek and neck. From this position, the deaf-blind user can primarily obtain information about the speech from vibrations from both the neck and jaw, the movement of the lips and jaw and secondarily from the airflow characteristics during the speech. This series of studies by Reed illustrated that naturally occurring tactual sensations produced by sound can provide acoustic information to the hearing-impaired.

Russ Palmer, a hearing and visually impaired person, has worked on a new approach in understanding how people with sensory impairment perceive and interpret music. He called this idea as "Feeling the music philosophy"—a description for being able to "visualise" and "interpret" music by sensory impaired people, to feel music through vibrations instead of listening to music through the ears [24]. In one of Palmer's articles, he describes how people might feel music through vibrations:

> *... it is true to assume that all people with a sensory impairment, without the use of hearing aids, can feel sound vibrations and "tones" through their bodies. This means that the physiological, neurological functions in the body become activated in a stronger sense, compared to those people who have no hearing impairment i.e. a switching of senses. I know that when I switch off my hearing aids there appears to be a "switching over" of senses through to my "tactile" sense.*

Furthermore, Palmer developed a theory in which he claimed that the vibrations produced by low tones can be felt by body sensors in the feet, legs and hips; middle tones can be felt in the stomach, chest and arms; and high tones can be felt in the fingers, head and hair [24]. An extensive review on the tactile modality, which specifically measures the tactile sensitivity for the human body, has been done by an Army Research Laboratory in USA [25].

2.2.3 Cross-modal displays

The main motivation for investigating cross-modal displays is to find out the ways we can enable hearing-impaired users to have a more satisfying musical experience. Tactile displays are one of the most commonly used alternatives for crossing modalities. For example, Hoggan and Brewster have done extensive work on adding visual, audio and tactile feedback to touch-screen widgets to improve the overall perceived quality [26-29]. Non-speech sounds have been mapped onto tactile displays to communicate audio information to deaf users [30]. Similarly, devices such as the Tactuator [31] have been used to present audio information to a deaf user's fingers using vibrotactile cutaneous stimulation. Cross-modal displays have also been used to simulate texture of visual displays objects such as a force feedback mouse that provided feedback to represent the bumps and contours of textile [32]. Vibrotactile feedback has also been used to assist with learning and playing musical instruments [33]. In virtual reality, cross-modal displays have been often used to enforce a sense of immersion [34].

Tactile devices have been employed as alternative displays for crossing modalities. An excellent review of state-of-the-art in tactile display technology can be found in [35]. Vibrotactile feedback is typically provided by devices that are directly attached to the skin [36]. These devices provide a tactile sensation by stimulating the cutaneous receptors [36]. Gunther *et al.* [37] and Mitroo *et al.*[38] have developed ambient displays with the aim of creating an experience that can influence the mood of the users or enhancing the interaction experience of the users.

Previous research has reported that humans can only detect vibrotactile frequencies between 10 Hz to 1000 Hz [36, 39, 40]. Therefore, most tactile display research aimed to encode the information into this frequency range. In our case, we did not aim to present precise information to the users, but instead sought to create an experience to influence overall satisfaction. Since music is a highly expressive and complex art form, we believe that scaling the audio signal onto the vibrotactile spectrum might potentially reduce the complete effect of music. Most of the vibrotactile devices used in tactile displays such as motors, tactors, and piezoelectric elements [41] cannot vibrate at frequencies high enough to cover the frequency range of music. Voice coils may offer a more accurate source of vibrotactile stimuli [41]. There has been little research into ambient music-touch displays. We believe harmonics, timbre and resonant frequencies are integral to the overall experience and thus pose many challengers for researchers in cross-modal display

2.3 Visualising Music

The visual representation of music has a long and colourful history. Among the earliest reported was Kircher, who developed a system of correspondences between musical intervals and colours [12]. Similar work was done by Marin Cureau de la Chambre. Sir Isaac Newton was able to show the parallel between the colour spectrum and notes on the western musical scale [12]. Louis Castel developed a "light-keyboard" which would simultaneously produce both sound and what he believed to be the "correct" associated colour for each note [12]. Alexander Scriabin has done similar work by developing the "Colour Piano" and "Colour Light Music" [12].

Among the earliest researchers to use a computer based approach to visualise music was Mitroo [38], who input musical attributes such as pitch, notes, chords, velocity, loudness, etc., to create colour compositions and moving objects. In the early 20th century, Oskar Fischinger, an animator, created exquisite “visual music” using geometric patterns and shapes choreographed tightly to classical music and jazz [42]. More recent example is Norman McLaren, a Canadian animator and film director who created “animated sound” by hand-drawn interpretations of music for film [43]. There were many others who have created visual music [44]. Evans [45] gives an excellent analysis of visual music. Since then, music visualisation schemes have proliferated to include commercial products like WinAmp™ and iTunes™, as well as visualisations to help train singers. However, the effect of these different music visualisations on the hearing-impaired has not been scientifically investigated and no prior specific application for this purpose is known to the author.

Musical visualisation schemes can be categorised into two groups—augmented score visualisations and performance visualisations [46]. The first group focuses on generating computer graphics that show the relationship among musical pieces or graphical illustration of the compositions; the target audience of this type of music visualisations is people with a musical background. Malinowski [47] has done a significant amount of work on this type of visualisation especially in developing a system called “Music Animation Machine” (MAM). When asked whether MAM is useful for the deaf, Malinowski wrote [47]:

> *I’ve sent tapes to Gallaudet University and to various deaf people I’ve known of, but I’ve received only one or two responses; from those, it appears that the MAM display is no more interesting to deaf people than it is to hearing people when they watch it with the sound turned off, which is to say: not*

> *very. In a way, this result, though disappointing, makes a certain amount of sense: why, after all, should it be any different? The fact that the MAM is not very interesting when unaccompanied by sound suggests that it is deficient: that there are aspects of musical sound which are not well represented in the MAM display. This is something I'm working on: to figure out what aspects of our perception of sound are not present in the MAM, and try to invent ways of showing them.*

Other examples of this type of music visualisations include the work of Hiraga *et al*. [48] and Foote [49]. The raison d'être of the above mentioned music visualisations is that they are informative, for example they can be used in music analysis, rather than conveying a musical experience and hence are not of direct relevance.

The second group of music visualisation schemes, performance visualisations, deal with different musical characteristics like volume, pitch, mood, melody, instruments, tempo etc. that can be extracted from an audio stream. Such features have been mapped to visual properties of a target rendered scene. The scene typically consists of objects with various colours, positions and other attributes. Kubelka *et al*. [50] have produced a basic scheme that can map music characteristics to the parameters of a particle animation scheme to create real-time animations. There is no report on how to determine the most effective musical features so that the visualisation can be made more meaningful. However, the basic idea of a particle animation system can be used to create a sophisticated and meaningful display, provided that we are able to extract the most appropriate audio features and develop a suitable audio-to-visual mapping.

Smith [51] has performed a music visualisation, which maps music data to 3D space. This system reads in Musical Instrument Digital Interface (MIDI) data files, and tones generated by individual instruments are represented by distinct coloured spheres in the visualisation. The characteristics of each sphere are dependent on three properties that describe musical tones: volume, timbre and pitch. Each note is represented as a sphere where the relative size of the sphere corresponds to the loudness, colour corresponds to the timbre of the tone and relative vertical location of the sphere corresponds to the pitch of the tone. Individual instruments are mapped to particular values along the horizontal axis. Although this music display is totally generated by the music, Smith’s aim was to present an alternative method for visualising music instead of conventional music notation. However, this simple approach of visualising music is easy to interpret and thus might be helpful for someone with less musical competency. Similar work has been done by McLeod and Wyvill [52] using real-time audio stream as the input.

Furthermore, there have been attempts to extract meaningful musical features from live performances and map them to the behaviour of an animated human character in such a way that the musician’s performance elicits a response from the virtual character [53, 54]. DiPaola and Arya have developed a music-driven emotionally expressive face animation system, called “MusicFace” [46], to extract the affective data from a piece of music to control facial expressions. Although the analysis attempted to extract affective features, it was not their intention to elicit emotional responses in the viewers.

Most of the music visualisation schemes reported in the literature have not targeted hearing-impaired people. However, there have been attempts to build displays capable of providing information to the hearing-impaired about sounds in their environment. For example, Matthews *et al*. [3] have proposed a small desktop screen display with icons and spectrographs that can keep the deaf person informed about the sounds in their vicinity. Similar work has been done by Ho-Ching *et al*. [2] where they implemented two prototypes to provide awareness of environmental sounds to deaf people. However when it comes to experiencing sounds, Ho-Ching writes:

> *...there is still a gap between the sound experience of a hearing person and the experience of a deaf person. For example, although there are several methods used to provide awareness of certain notification sounds, there is little effective support for monitoring.*

This quotation seems especially applicable for musical sounds, where experiencing the music is more important than just knowing the acoustic signal attributes.

Matthews *et al.* [3] conducted interviews with deaf people to find out their visual design preferences and functional requirements for peripheral visualisations of non-speech audio that might help improve future applications. They found that deaf people prefer to have displays that are easy to interpret, glance-able and appropriately distracting. As function requirements, users (deaf people) wanted to be able to identify the sounds that occurred, view a history of the displayed sounds, customize the information that is shown and determine the accuracy of displayed information. Matthews *et al.* have reported that, during an interview, one participant expressed her interest in music: *she loved watching live musicians and feeling the vibrations through the floor or speakers*. This comment encourages the idea of employing a combination of haptic and visual effects to provide a better musical experience to a hearing-impaired person.

2.4 Feeling the Vibrations

As mentioned in the previous section, feeling sound vibrations through different parts of the body plays an important role in perceiving music, particularly for the deaf. Based on this concept, Palmer, developed a portable music floor which he called "Tac-Tile Sounds System" (TTSS) [55]. However, we have not been able to find a report of any formal objective evaluation of the TTSS. Recently, Kerwin developed a touch pad that enables deaf people to feel music through vibrations sensed by the fingertips [56]. Kerwin claimed that when music is played each of the five finger pads on a pad designed for one hand vibrates in a different manner that enables the user to feel the difference between notes, rhythms and instrument combinations. As in the previous case (TTSS), not much technical or user test details about this device are available.

Recently, Karam *et al.* described an "EmotiChair" [57, 58], which is an audio-tactile device that represents audio information as tactile stimuli. The EmotiChair had a simplified "Model of Human Cochlea" (MHC) to separate audio signals into discrete vibro-tactile output channels that are presented along the back of a user's body. The EmotiChair has been designed to process sound, including music inputs, according to pre-defined transformations before producing haptic output. In addition, the haptic output was presented only to the back of the user's body. This leaves out the other important parts of the body (such as finger tips, feet, etc...) that deaf people generally used to feel music [6].

Gunther *et al*. introduced the concept of "tactile composition" [37] based on a similar system comprising thirteen transducers worn against the body with the aim of creating music specifically for tactile display. Gunther composed music specifically to produce vibrotactile frequencies below 1000 Hz. Although this approach aims to ensure that the vibrations are within tactually perceivable ranges [40], it also limits the range and type of music that can be presented.

The "Multisensory Sound Lab" (MSL) developed by Oval Window Audio [59] is a special environment consisting of vibrating floors generated by mechanical vibrations and colourful visual displays generated by spectrum analysers intending to educate or enhance the sound experience for the hearing-impaired. The MSL environment uses a transformation of sound to generate vibrations (low frequency sounds as slow vibrations and high frequency sounds as faster vibrations) and the visual display effects are basically representing the different waveforms of the audio signal which is very different from the approach taken in this dissertation.

Many vibrotactile hearing aids are available in the market. They are classified as, a) devices which provide a pulsing action to inform users about particular events such as door bell, telephone ring; and b) devices which convert acoustic information of speech or music to tactile input [60, 61]. Vibrotactile hearing aids uses a single mechanical transducer or an array of transducers placed on a person's body to deliver the tactile signal [60]. Examples of tactile hearing aids include the Little Tactile Device (LTD) and Tactaid7 by Audiological Engineering Corporation [62]. A generic example of the tactile device would be a single-channel system that converts sound into mechanical displacements (vibrotactile) or electrocutaneous (electrotactile) signals that follow the frequency-intensity pattern of the sound over time. The output of such a device is similar to the acoustic signal being filtered by an equivalent low-pass filter; thus the output of these tactile hearing aids generally consists of lower frequency vibrotactile output of 50-300 Hz resulting some loss of information between the actual signal and the resulting output [61]. Therefore, typical tactile hearing aids may not be able to convey information in the high frequency contents of music. Despite these limitations, tactile aids prove to be able to help hearing impaired individuals in speech perception and speech learning, and the results will be more evident when the users use the devices for a long period [63].

There are commercial products advertised to enhance the listening experience by providing tactile information for explosions, gun shots, and other high noise events, for example, the "Tactile Effects System" by Crowson Technology [64]. However, currently, the closest commercially available comparisons to the proposed Haptic Chair include the "Vibrating Bodily Sensation Device" [65] from Kunyoong IBC Co, the "X-chair" [66] by Ogawa World Berhad, the MSL from Oval Window Audio [59], Soundbeam® products [67] (soundchair, soundbed, sound box and minibox) by Soundbeam Project and Snoezelen® vibromusic products from FlagHouse, Inc. The Vibrating Bodily Sensation Device is a vibro-tactile device which is advertised as a mobile device that can be placed on existing chairs to enhance the listening experience of the hearing-impaired [65]. The "X-chair" has been designed for long-hour indulgence while listening to music, playing games or watching movies with two modes of operation—an asynchronous mode where vibrations are generated independent of the sound stream (this mode is more like a massage chair) and synchronous mode where vibrators are synchronised to the audio [66]. Both these commercial products only stimulate one part of the body (the lower lumbar region of the body which is more sensitive to lower frequencies). "Soundchair", one of the Soundbeam products, is a vibro-acoustic device advertised as a comfortable chair with adjustable back and leg-rest extension [67]. This chair has 3 pairs of speakers driving 3 separate resonant cavities attached to the back, seat and leg-rest of the chair. The chair is to be mainly used with "Soundbeam", a device that creates music according to body movements. Developers of the "Soundchair" claim that it enables people, including those with hearing impairment and other disabilities, to experience the physical vibrations of music through their bodies. However, a formal scientific evaluation of this product is still not available in the public domain.

2.5 Bone Conduction of Sound

Bone conduction of sound is likely to be very significant for people with certain hearing impairments and a far greater range of frequencies is transmitted via bone conduction of sound compared to purely tactile stimulation [68]. Bone conduction is the process of transmitting sound energy through vibrations of the skull or neighbouring parts of the body [60]. A comprehensive study about all aspects of bone conduction has been done by an Army Research Laboratory in USA [60]. For a person with normal hearing, bone conduction is a secondary auditory pathway supplementing the air conduction process. In normal situations (when there is no direct stimulation on skull), the contribution of bone conduction to the sense of hearing is very small. This is due to the impedance mis-match between the skull and the air. However, bone conduction technology has been widely used in a variety of commercial products including development of hearing aids and devices for listening to music. Vonia [69] is one of the very few companies that has designed bone conduction headphones for use as a replacement for a standard pair of headphones, to listen to music.

The major difference between hearing through bone conduction and hearing through air conduction is the difference in the perceived frequency range. The frequency range of normal human hearing (through the air conduction pathway) is approximately 20Hz to 20 kHz. In contrast, Deatherage *et al.* [70] demonstrated that humans can perceive ultrasonic frequencies by bone conduction. In fact, this ultrasonic hearing can go up to as high as 100 kHz [68]. Dobie *et al.* [71] has suggested that the bone conduction process is able to demodulate the ultrasonic signal to the perception of a frequency that is within the audible range which represents the fluctuations of the carrier signal. Moreover, bone conduction of ultrasound has been shown to be able to help users with sensorineural and conductive hearing loss [68, 72-74]. Imaizumi *et al.* [74] found that bone-conducted ultrasound can activate the auditory cortex of profoundly deaf subjects. Lenhardt *et al.* [68] have suggested that

> *Bone-conducted ultrasonic stimulation may provide frequency discrimination and speech detection in normal, older hearing impaired and profoundly deaf human subjects.*

They claim that, when speech signals are used to modulate the amplitude of an ultrasonic carrier, the result is a clear perception of the speech stimuli and not a sense of high-frequency vibration. HiSonic™ is one such commercially available ultrasonic bone conduction hearing aid. Staab *et al*. [75] have shown that this device is useful for the people with profound hearing losses.

2.6 Conclusions

This chapter has given an introduction to music and the deaf, cross-modal interactions and displays, types of music visualisations, sensing sound through tactile channels and bone conduction of sound. A study of the state-of-the-art reveals the variety of methods that have been suggested to compensate for hearing disabilities in music appreciation, but as yet little has been done to assess their applicability to the hearing-impaired community. We addressed this gap by proposing a user-centred approach, primarily by employing an extensive series of user studies, to explore how best we could enhance the musical experience of a deaf person. The details are presented in Chapters 3, 4 and 5.

Chapter 3

Design Conceptualisation

The literature review revealed that little guidance is available to address the challenges encountered at the early stages of designing a system for the deaf to facilitate a better experience of music. Therefore, it was important to study what hearing-impaired people might find most useful in helping them enjoy music. The first part of this chapter discusses the findings from a background survey that was conducted with deaf people from multi-ethnic backgrounds. This forms the basis for conceptualising approaches that moved us towards understanding how best to provide musical sensory enhancement for the deaf.

A question central to this research is whether or not a musical experience can be conveyed by sensory channels other than sound. Conventions of rhythmic and melodic patterns, chord sequences and key changes are exploited to create an intellectual and emotional response that we call the musical experience.

As a starting point, a system was developed to transcode musical sequences of information into a visual sequence in real-time. The proposed system architecture supported building different types of displays rapidly thus allowing us to experiment in finding a suitable mapping between musical data and visual data. The development of this real-time music visualiser is described in the second half of this chapter.

3.1 Background Survey

This survey was conducted with hearing-impaired people to investigate the following fundamental issues:

- To what extent do deaf people engage in musical activities?
- What type of music do they listen to?
- What are the strategies they used to listen to music?
- Are they upset by not being able to enjoy music as much as they would like?
- Are they willing to be involved in developing a new assistive device that would enhance their musical experience?

Forty one people (20 male participants and 21 female participants; 36 of them aged 15–30 years and 5 subjects aged 31–45 years) with various degrees of hearing impairment took part in this survey. They were asked to complete a standardised survey form. There were 22 partially deaf and 19 profoundly deaf participants who all had normal eyesight. Teachers proficient in sign language were available for any clarification requested by the participants. In the analysis, we did not exclude any of the responses since there were no obvious outliers. The findings of this survey are summarised in next few sub-sections.

3.1.1 Involvement in musical activities

The respondents were asked whether they had taken part in musical activities: whether they attend concerts or listen to music at home. Seventy seven percent of the partially deaf subjects reported taking part in musical activities, whereas only 32% of the profoundly deaf subjects reported being involved in musical activities (Table 3.1). This observation supports the hypothesis (perhaps widely held concept) that the partially deaf are more likely to have taken part in musical activities than the profoundly deaf. The results are shown in Table 3.1.

Table 3.1: Observed frequencies for profoundly deaf and partially deaf subjects taking part in a musical activity

		Have taken part in a musical activity		
		Yes	No	Total
Level of deafness	Partially Deaf	17	5	22
	Profoundly Deaf	6	13	19
	Total	23	18	41

The value of chi-square, $\chi^2(1, N = 41) = 8.58, p < 0.01$ rejects the null hypothesis of no association between the two variables. In other words, the data suggests that partially deaf subjects are more involved in musical activities than the profoundly deaf. This might seem obvious but needed to be formally tested.

3.1.2 Types of music preferred

In order to decide what music genres we should work with, the subjects who participated in the background study were asked to indicate the types of music or songs they listen to. Figure 3.1 summarises their responses and suggests that most hearing impaired people listen to music with a strong beat.

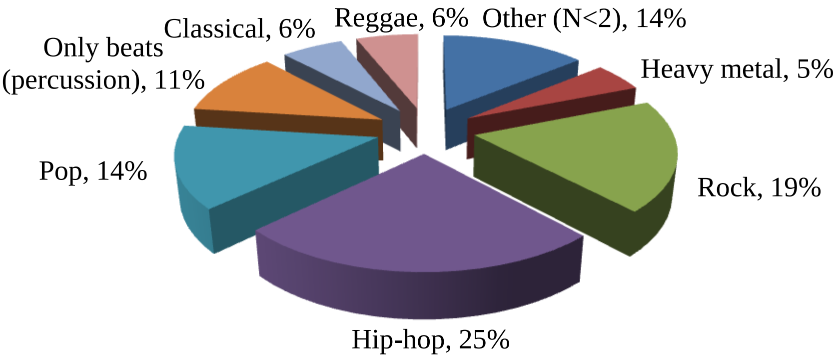


Figure 3.1: Preferred music genres by hearing-impaired participants

3.1.3 Factors that enable enjoyment of music

The respondents were asked to identify the dominating factor that enabled them to enjoy a musical activity and this information was used to decide the type of assistive system we should develop. From the responses shown in Figure 3.2 it appears that most deaf people rely either on feeling vibrations or watching visual displays.

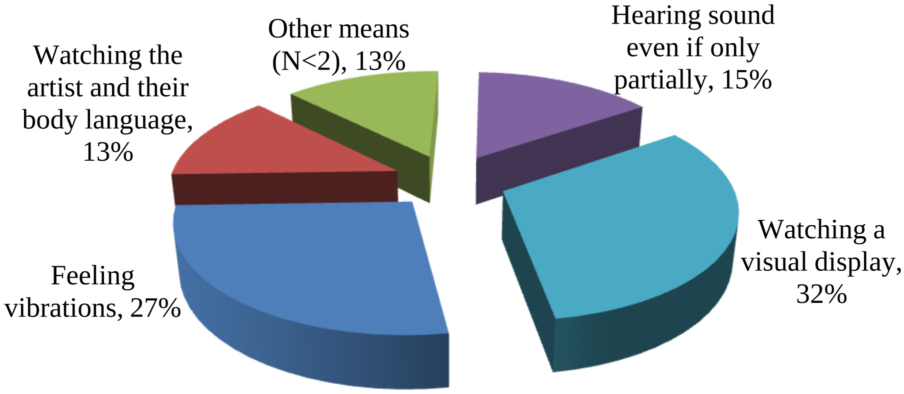


Figure 3.2: Factors that enable the hearing-impaired participants to enjoy music

3.1.4 Regret over the lack of musical accessibility

The subjects who have participated musical activities were asked whether they regret the fact that they were not able to enjoy music as much as they would like to. Sixty five percent of the partially deaf and 67% of the profoundly deaf subjects reported that they feel "upset" about not being able to enjoy music to their potential ability. These observations support the hypothesis that, regardless of their hearing ability, deaf people are likely to express some degree of dissatisfaction over any hindrance to full enjoyment of music. A chi-square test was carried out to verify the hypothesis. Since some cells of the contingency table (Table 3.2) have values less than 5, Yate's correction [76] was applied.

The value of chi-square, $\chi^2(1, N = 23) = 0.27, p > 0.05$ supports the null hypothesis of no association between the two variables—"level of deafness" and "regretting lack of musical accessibility". This implies that, regardless of the level of deafness, people do get upset about not being able to enjoy music.

Table 3.2: Observed frequencies for profoundly deaf and partially deaf subjects reporting "being upset about not being able to enjoy a musical activity as much as they would like to"

		Regret not being able to enjoy a musical activity		
		Yes	No	Total
Level of deafness	Partially Deaf	11	6	17
	Profoundly Deaf	4	2	6
	Total	15	8	23

3.1.5 Assistive devices that might enhance a musical experience

Type of assistive devices used during a musical activity

Figures 3.3 and 3.4 illustrate the types of assistive devices hearing-impaired people have used while engaging in a musical activity and whether they were deemed useful. We categorized these assistive devices into 5 groups:

1. Sound Amplification: Hearing aids or use of powerful amplifiers to increase the volume of music.
2. Subtitle Display: A device that would show information about the music using text-based output; e.g. karaoke type lyrics.
3. Haptic Display: A mechanism that would provide a vibrotactile feedback. This included feeling vibration from the floor, touching the speakers, holding air-filled balloons, etc...
4. Graphical Display: A mechanism that would provide visual effects (as opposed to text-based visual information). Lighting effects, music videos, displays showing live music performances were few examples.
5. Sing Language: A mechanism that shows a sign language interpretation of music. Typically, in these interpretations, the gesturing patterns are more rhythmic compared to standard sign language.

Sign language and sub-title displays were the most commonly used methods during a musical activity. One reason for this could be the fact that these were two of the most easily available options.

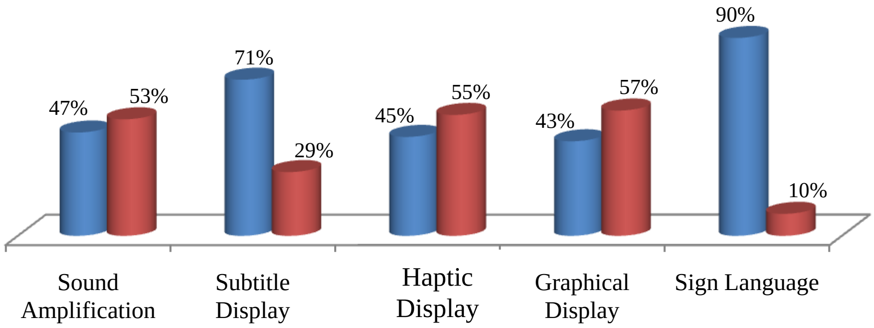


Figure 3.3: Assistive devices that hearing-impaired people have used while engaging in a musical activity (■ Have not used; ■ Have not used)

One of the significant observations for the purpose of this study was that most people (94%) who have used a graphical display or haptic display found these assistive devices contribute significantly to their musical enjoyment.

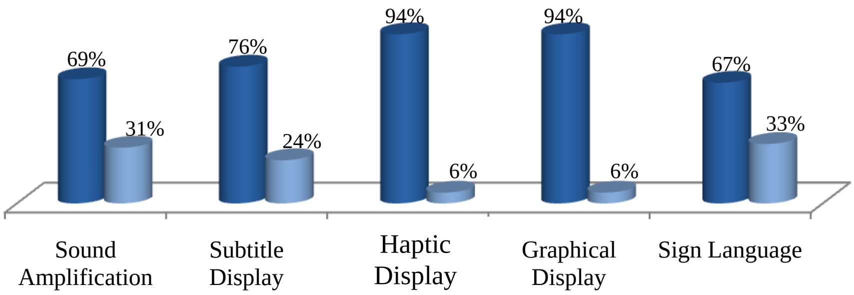


Figure 3.4: Usefulness of different assistive devices (■ Helpful; □ Not helpful)

Willingness to use a new device

Our initial idea was to study the visual and haptic methods to provide a more satisfying musical experience to a deaf person. The plan was to adopt a user centred design approach as described in ISO 13407 [77]. We wanted to make sure that our target users are willing to take part in the iterative process of developing an assistive device consisting of a visual display and a chair.

Visual display

Participants in the study were asked whether they would be willing to use a visual display that reflects basic musical features such as note onset, pitch, loudness, type of instrument and changes in the overall pitch context. It was found that most partially deaf and profoundly deaf people are willing to use such a device (Figure 3.5). The chi-square value, $\chi^2(2, N = 40) = 0.95, p > 0.05$, indicates that there is no association between the level of deafness (whether profoundly deaf or partially deaf) and the willingness to use a visual display.

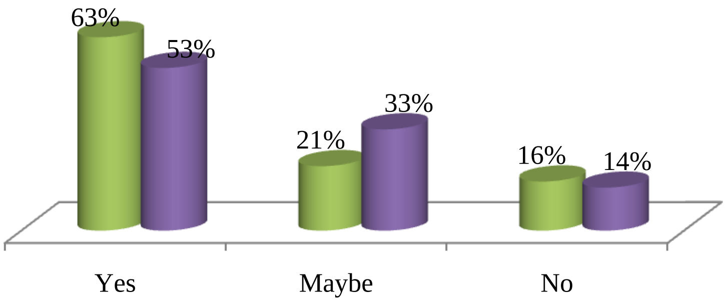


Figure 3.5: Willingness to use a visual display by hearing-impaired participants
(■ partially deaf; ■ profoundly deaf)

Haptic Display

Our initial concept was to use the concept of "whole body stimulation" with vibrotactile feedback. A chair was a simpler, multi-purpose and portable structure to test this concept. In addition, we assumed our participants would be more comfortable watching a visual display (the other intended component) while sitting on a chair. We studied if our intended participants are willing to be involved in the iterative process of developing a haptic chair. When asked whether participants would be willing to use a chair that vibrates to reflect the musical sound signal, most partially deaf and profoundly deaf people said they would use it (Figure 3.6). As in the previous case, the chi-square value, $\chi^2(2, N = 27) = 1.37, p > 0.05$, revealed that there is no association between the level of deafness and the willingness to utilise a haptic input.

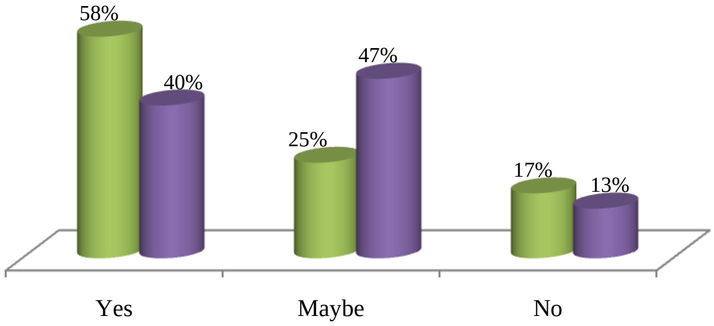


Figure 3.6: Willingness to use a Haptic Chair by hearing-impaired participants (■ partially deaf; ■ profoundly deaf)

We asked if they preferred to use one device at a time or a combined system of visual and haptic displays. The Figure 3.7 shows the responses received.

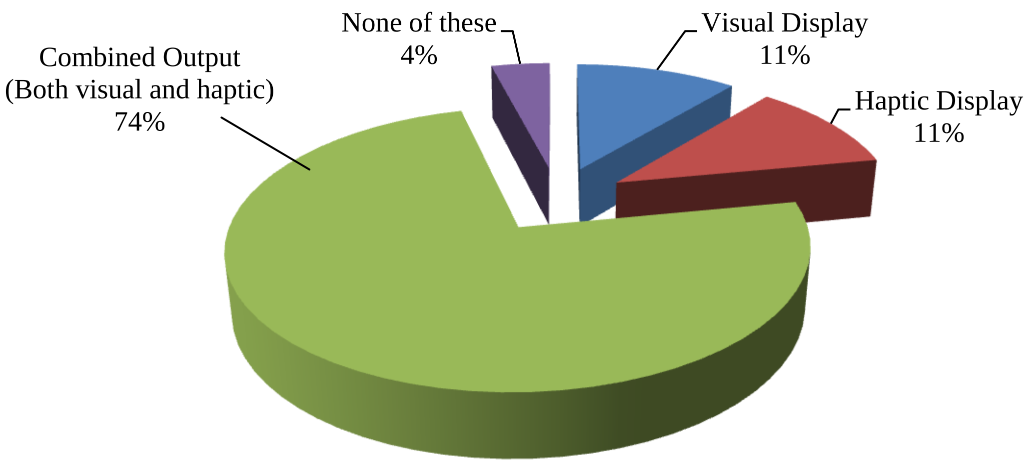


Figure 3.7: Preferred configuration of the new assistive device.

Most of the hearing-impaired subjects said they expected that they would use a combined system of a visual display and a haptic input.

3.1.6 Qualitative findings

The participants were asked to state any additional comments/suggestions that they might think would enhance their ability to enjoy music. Following are some of the significant responses:

More visual and actions that makes sense for hearing loss people.

Visual displays such as hand motions and vibration sounds attachment all would enhance the ability to enjoy music.

Maybe add subtitle for music when it's played.

Read lyrics before listening to music.

Very loud music.

Vibration seat with sound sensors + lighting effect.

Sign language interpretation of music.

In summary, the following are the findings of the background survey:

- Partially deaf subjects were more involved in musical activities than the profoundly deaf.
- Regardless of the level of deafness, the participants expressed the desire to enjoy music.
- Apart from sign language and subtitle displays, most of the participants thought vibrating devices and visual displays would be helpful.
- A combination of haptic and visual displays could be useful in conveying a musical experience to the hearing-impaired.

3.2 Music Visualisation with Abstract Patterns

Jones and Nevile [43] suggest that mapping single musical voices in a composition to single graphical elements would result in a visual experience with a musical meaning. Motivated by this, we developed a system to generate visual effects where each graphical element corresponded to a specific feature of music. As a start, a basic music visualiser was designed to provide a pleasing complementary sensory experience. During the initial exploratory stage, we evaluated the system with normal-hearing people. These hearing subjects were expected to give more informative feedback about the music-to-visual mappings since they could hear the music. However, understanding that the requirements of normal-hearing and hearing-impaired people might be different, our follow-up studies were done with hearing-impaired subjects. The music visualiser was implemented using the mappings and architecture described in next two sections.

3.2.1 Related audio-to-visual mappings

A large number of parameters can be extracted from a music data stream and these can each be mapped to several different visual properties. The number of all possible mappings is too large to be explored fruitfully without some guiding principles. Mappings reported in the literature and results of the studies of human audiovisual perception could be used as a start. If nothing else, we need to convey the beat and overlying percussive elements[43]. However, the musical “key” and key changes are also of great importance since they are normally part of the structural foundation of musical works, and they evoke emotional empathy [78]. For example, major keys are associated with happy and uplifting emotions, whereas minor keys are typically associated with sad emotions [78].

Initially, each note produced by a non-percussive instrument was mapped to a graphical star-like object to emphasise the note onset, and notes were arranged from left-to-right in order of increasing pitch, mainly because this method mirrored the piano keyboard and allowed chord structure to be visualised. The smaller a physical object is, the higher the frequencies it tends to produce when resonating. Hence, higher pitched notes of a piece of music were mapped to smaller sized visual objects [43]. A visualisation that maps high notes to small shapes and low notes to large shapes is expected to be more "natural" and intuitive than the reverse because it is consistent with our experience of the physical and biological world. Similarly, there is a rational basis for amplitude being mapped to brightness because both amplitude and brightness are measures of intensity in the audio and visual domains, respectively, as justified experimentally by Marks [79].

As far as colour is concerned, although a number of "colour scales" have been proposed [80-82]. Kawanobe *et al.* [67] have proposed a mapping table between music and colour combinations and they have also shown that there is a strong similarity between musical affect and colour affect. However, there is no basis for the universality of any of them. It was beyond the scope of this work to analyse and evaluate different colour scales. In this music visualisation scheme, the "colour scale" proposed by Belmont (1944) was used to represent the colours of notes (Figure 3.8).

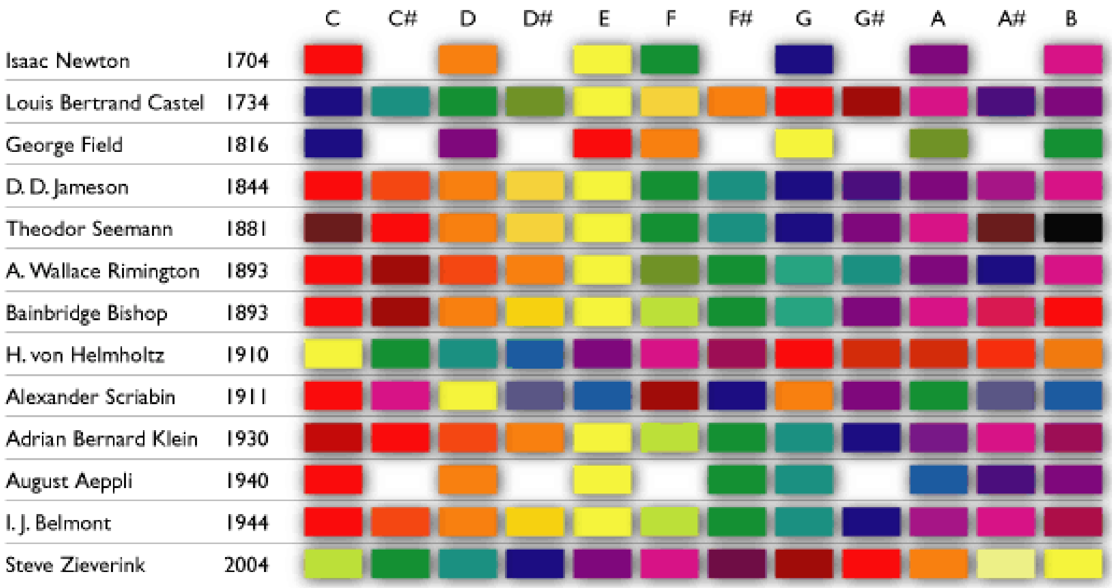
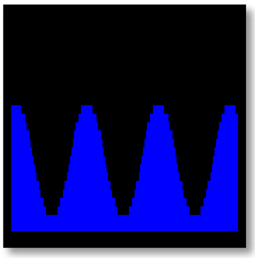
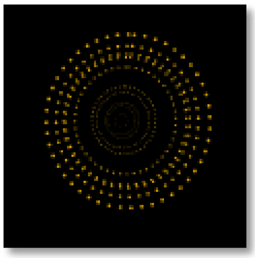
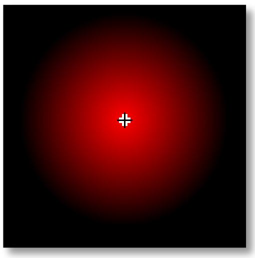
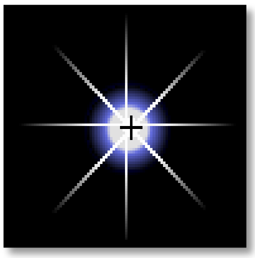
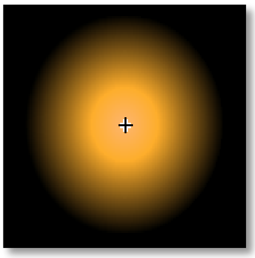
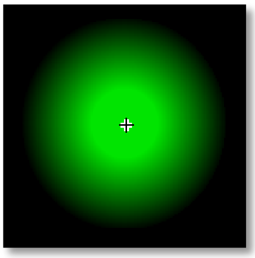
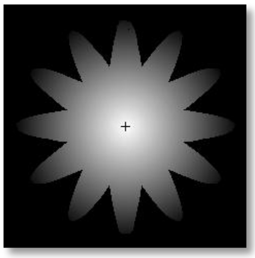


Figure 3.8: Examples of notes-to-colour maps
(*Image taken from http://rhythmiclight.com/archives/ideas/colorscales.html*)

Visual effects were implemented to represent the sounds of different percussion instruments. Most of these initial visual effects were exploratory and experimental. Table 3.3 shows some of the music-to-visual mappings used in this work. In addition, visual effects were introduced to represent the key changes that might occur during a song. Since many synesthetes artists, for example Amy Beach and Nikolai Rimsky-Korsakov, have associated musical key with colour, it was decided to visualise key changes using different backgrounds instead of representing the absolute “value” or precise sound frequencies in a key by the size of an icon as done with notes. Different keys function as a kind of background context for chords and notes without changing the harmonic relationship between them. This also supported the idea of mapping key to the background colour of the visual display. The next section will describe a pilot study carried out to find out a suitable key-to-background colour mapping scheme.

Table 3.3: Audio-visual mapping of different instruments used in the heuristic based music visualiser

Instrument	Visual Effect	Instrument	Visual Effect
Bass Drum Effect: Wave (Pulsating up and down) Colour: Blue Screen position: Bottom edge of the screen		**Closed Hi-Hat** Effect: Bursting effect Colour: Yellow Screen position: Lower-Left of the screen	
Snare Drum Effect: Sphere fading into background Colour: Red Screen position: Lower-middle of the screen		**Ride Cymbal** Effect: Star falling Colour: Silver Screen position: Top-Middle of the screen	
Hi Tom Effect: Sphere fading into background Colour: Gold Screen position: Lower-Right of the screen		**Crash Cymbal** Effect: Firework effect Colour: Red Screen position: High-Middle of the screen	
Hi Bongo Effect: Sphere fading into background Colour: Green Screen position: Lower-left of the screen		**Piano note** Effect: Star-like object into background Colour and screen position: depends on the note class (C, C# ect.)	

3.2.2 Development of the platform

Extracting notes and instrument information from a live audio stream is an extremely difficult problem [83] and is not the main objective of this dissertation. Hence, in this work, Musical Instrument Digital Interface (MIDI) data, a communications protocol representing musical information similar to that contained in a musical score, was used as the main source of information instead of a live audio stream. Using MIDI makes determining note onsets, pitch, duration, loudness and instrument identification straightforward. However, just as with musical scores, key changes are not explicit or trivially extractable from the MIDI note stream. A method developed by Chew [84] based on a mathematical model for tonality called the "Spiral Array Model" was used to accomplish this task.

System architecture

The proposed music visualisation scheme consists of three main components: Processing layer, XML Socket and Flash® AS3.0 application (Figure 3.9). The processing layer takes in a MIDI data stream (from an external MIDI keyboard or from a standard MIDI file) and extracts note onset, pitch, loudness, instrument and key changes. This processing layer is implemented using the Max/MSP™ musical signal and event processing and programming environment.

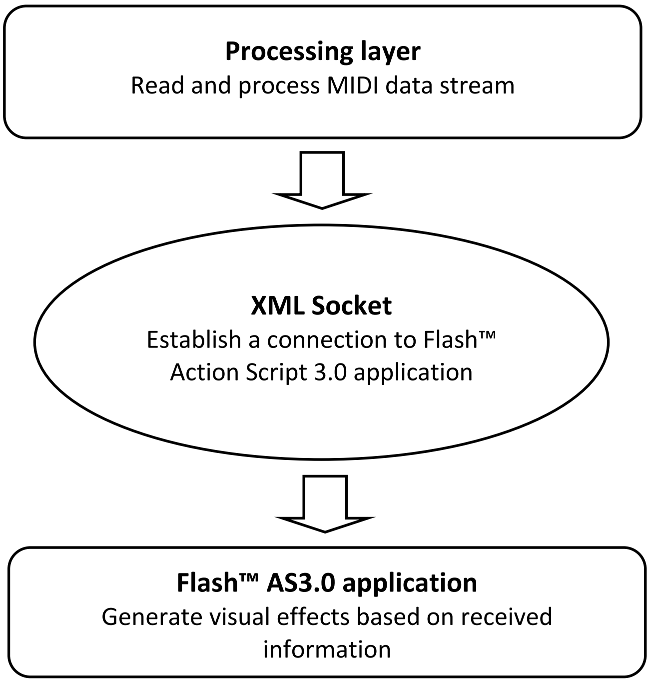


Figure 3.9: System architecture of the proposed music visualiser

The Max *midiin* object was used to capture raw MIDI data coming from a MIDI keyboard and the *seq* object was used to deal with the standard single track MIDI files [85]. Note and velocity of note onset can be read directly from MIDI using the *midiparse* object. Percussive and non-percussive sounds were separated by considering the MIDI channel number [86]. The extracted musical information is passed to a Flash CS3 program written using Action Script 3.0 via a Max *flashserver* external object [87]. The basic functionality of the *flashserver* is to establish a connection between Flash CS3 and Max/MSP. The TCP/IP socket connection that is created enables exchange of data between both programs in either direction thereby enabling two-way Max-controlled animations in Flash CS3. With this architecture, each musical feature extracted by Max/MSP can be programmatically associated with a corresponding visual effect in Flash.

Another fundamental display decision concerns the window of time to be visualised. Two distinct types of visualisation can be identified; a “piano roll” and a “movie roll”-type (See Figure 3.10).

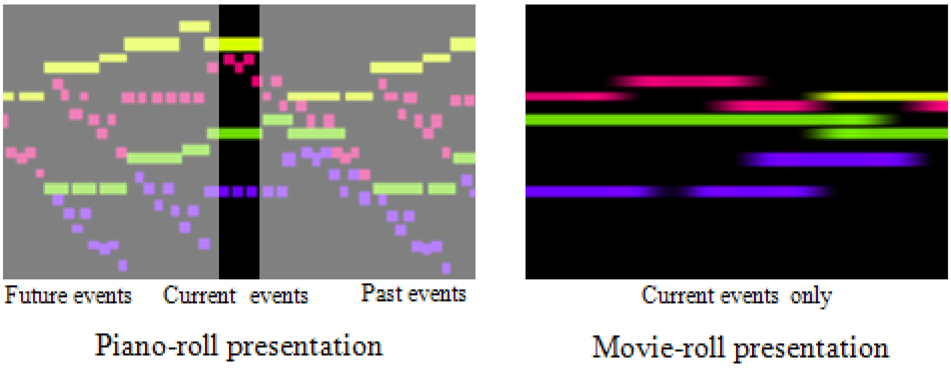


Figure 3.10: Example of piano roll and movie roll.

The “piano roll” presentation refers to a display that scrolls from left to right in which events corresponding to a given time window are displayed in a single column, and past events and future events are displayed on the left side and right side of the current time respectively. In contrast, in a “movie roll”-type presentation, the entire display is used to show instantaneous events which also allows more freedom of expression. The visual effect for a particular audio feature is visible on screen for as long as that audio feature is audible, and fades away into the screen as the audio feature fades. When listening, people tend to hear instantaneous events: future events are not known (although they might be anticipated); and past events are not heard (although they might be remembered). Thus, a “movie roll”-type visual presentation is more directly analogous to the audio musical listening process than the “piano roll” depiction.

Automatic tracking of changes in the tonal context

As mentioned previously, key changes are not trivially extractable from a MIDI note stream because a) musical key is only implicit in a stream of notes, and b) MIDI note numbers do not carry quite the same amount of key information as note names. The musical key was estimated using the "Spiral Array Model", a method developed by Chew [88] based on a mathematical model for tonality. The spiral array model, proposed by Chew [88], is a three-dimensional model that represents pitches in 3D space. Any pitch-based objects that can be described by a collection of pitches, such as intervals, chords, and keys, can be represented in the same three-dimensional space for ease of comparison. In the spiral array model, pitches are represented as points on a helix, and adjacent pitches are related by intervals of perfect fifths, while vertical neighbours are related by major thirds. The pitch spiral is shown in Figure 3.11(a) [89]. Central to the spiral array is the idea of the Centre of Effect (CE), the representing of pitch-based objects as the weighted sum of their lower-level components. The CE of a key is shown on Figure 3.11(b). Further details of the spiral array model are given in [84, 88]. By utilising the Spiral Array Model, Chew developed an algorithm that determines the keys in a musical passage. Comparing the performance of this algorithm to previous methods by others, Chew has shown that her algorithm is able to predict the correct key with much greater accuracy than previous computational methods [84].

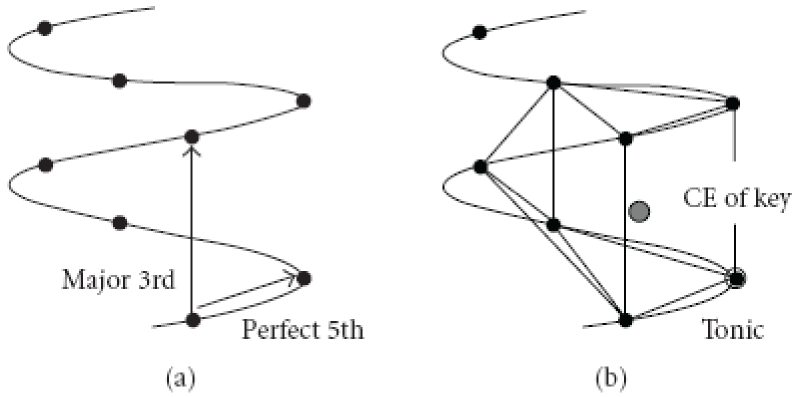


Figure 3.11: The Spiral Array Model: (a) Pitch spiral in the spiral array model;
(b) Centre of Effect (CE) as a proximity to represent the key (*Image taken from[89]*)

For key analysis using the spiral array, one needs to map any numeric representation of pitch to its letter name. For example, MIDI note number 60 could be spelled as “C”, “B#” or “Dbb” depending on the key context. The pitch spelling algorithm, described in [90, 91], was applied to assign letter names to the pitches so that they can be mapped to their corresponding representations in the spiral array for key finding. The pitch spelling algorithm uses the current CE as a proxy for the key context, and assigns pitch names through a nearest-neighbour search for the closest pitch-class representation. To initialise the process, all pitches in the first time window are spelt closest to the pitch class “D” in the spiral array, then the CE of these pitches is generated, and they were re-spelt using this CE.

Implementation of the real time pitch spelling algorithm was done in Java programming language. Performance of this pitch spelling algorithm was tested with a range of parameter values and reported an overall accuracy of 99.37% [91]. A short segment of the classical music piece “2-Part Invention #1” by Johann Sebastian Bach was used to ensure the correct Max/MSP implementation of the pitch spelling algorithm. Two hundred and fifty nine notes from the test piece were decoded manually and compared with the output of the pitch spelling algorithm. With the optimum parameter combination, our implementation gave 100% accurate results, which confirms the correct implementation of pitch spelling algorithm. A Max external patch was then developed for pitch spelling and generation of CE. Evolving CE was calculated as follows and used as a proxy for the key context:

$$\text{CE}(t) = \alpha.\text{CE}(t) + (1-\alpha).\text{CE}(t-1)$$

where t is the time-stamp. The value of α can be adjusted to emphasise local or global key context (higher α value results in more local CE, lower α value results in more global CE).

User Study: Key-to-colour mappings

The most conventional and simplest was to map the x, y, z coordinates of CE to RGB of the background colour. However, this mapping is low level. Therefore, attempts were made to map higher level meanings. For example, distance from the key is typically an important factor in perceiving energy. Thus, mapping the large movements of CE in the space to colours that convey high energy and mapping stable movements to colours that convey low energy was considered. Red was used to represent high energy and blue was used to represent low energy [92]. Movement of CE in the space was calculated in two different ways—movement with respect to the initial CE (absolute movement) and movement with respect to CE at the previous time instant (differential movement). Eight different mappings, as shown in Table 3.4 were created and a web based evaluation was carried out to determine the best mapping strategy.

Table 3.4: Different key-to-colour mapping strategies

Trial ID	Mapping strategy
Random	Random colours
XYZ_jet	XYZ of CE to jet[5] colours
Rel_rb	Relative distance of CE to mixture of Red and Blue
Rel_jet	Relative distance of CE to jet colour map
Rel_gray	Relative distance of CE to gray colour map
Dif_rb	Differential movement of CE to mixture of Red and Blue
Dif_jet	Differential movement of CE to jet colour map
Dif_gray	Differential movement of CE to gray colour map

A preliminary study was designed to evaluate the different key-to-colour mapping strategies. The main objective of the survey was to find the answers to following questions:

- Is content based visualisation better than random?
- What is the best mapping: XYZ, relative distance or differential movement?
- What is the most appropriate colour map: red/blue, jet or gray colour mapping?

Participants

Thirteen participants (5 female and 8 male) took part in this study. They were from the age group between 18-30 years, except for one participant who was from the age group 31-45 years. All had normal vision and hearing and 8 of them had taken music as an academic subject. The majority of them were familiar with Sri Lankan music samples (69%) and classical music samples (77%).

[5] Continuous colour spectrum that begins with blue, and passes through cyan, yellow, orange, to red

Apparatus

Two classical pieces (MIDI renditions of Bach's Prelude No. 1 in C from Book 1 of the Well-Tempered Clavier and Chopin's Prelude Op. 28 No. 24) and two Sri Lankan pieces (MIDI renditions of "Mango kalu nande" by Annesley Malewana and "Maliniye" by Clarence Wijewardana) were chosen as the test cases. Video clips of key-to-colour mapping (shown in Table 3.4) corresponding to the tonal context of the music samples were generated. In order to make the evaluation more rigorous, music pieces were selected such that some had stable tonal change and some had a high rate of fluctuations. Participants were presented with video clips (each consisted of three colour boxes where the colour of each of the boxes change with the music according to a particular mapping shown in Table 3.4). Subjects were asked to rank the choices according to how well the colour changed in relation to changes in the pitch context. A screen shot of a typical video clip used in this experiment is shown in Figure 3.12.

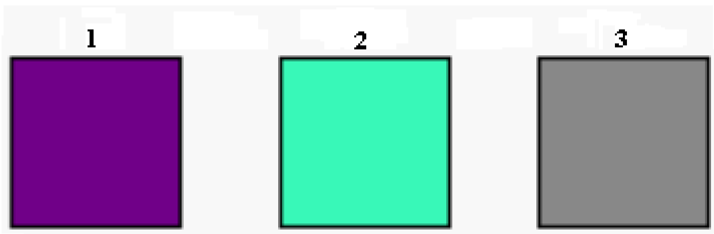


Figure 3.12: Screen shot of a video clip used during the survey (Colour of each of the boxes changes with the music according to a particular mapping shown in Table 3.4)

The focus of this survey was to evaluate representation of changes in pitch and not other aspects of music which may be more immediately appreciated, such as beat, tempo, and rhythm. Therefore, participants were given the opportunity to listen to two music samples—one with a stable tonal context and one with a fluctuating tonal context—to make the concept clearer. A web-based interface [93] was developed to carry out the experiment.

Procedure

In order to find out the most preferred key-to-colour mapping strategy, participants were presented with and asked to rate the key-to-background colour generated using mappings given in Random, XYZ_jet, Rel_jet and Dif_jet (Table 3.4). All these trials used the same colour map and the only variable was the rule used to pick the colour depending on CE. As mentioned earlier, out of the four cases (Random, XYZ_jet, Rel_jet and Dif_jet), three background colour animations were shown at a given time and depending on the user rank, the least liked visualisation was replaced by the remaining (4th) visualisation and re-tested. If newly added visualisation happened to be the least preferred one, one more round of testing was conducted to compare the two least liked visualisations. With this procedure, it was possible to rank all four cases according to the participants' preference.

In order to determine the most preferred colour scheme, two sets of videos were presented to the participants to rate. First video clip consisted of three colour animation boxes corresponding to mappings Rel_rb, Rel_jet and Rel_gray; second video clip consisted of three colour animation boxes corresponding to mappings Dif_rb, Dif_jet and Dif_gray (Table 3.4). Each video used the same set of rules except for the colour map (jet vs. gray scale vs. red/blue) to generate the colour animation.

For a given music sample, participants were asked to evaluate either 4 or 5 video clips (each video clip had 3 colour boxes, where colours were changed according to the changes in the tonal context of the music; participants rated each colour box). Since there were 4 different music samples, each participant ended up evaluating 16 to 20 video clips (i.e. ranking 48 to 60 colour boxes). Each participant took between 45 minutes to 1 hour to complete the test.

Results

Schulze's sequential dropping method [94] was used to find the winner of the most preferred mapping. With this method, it was possible to select a single winner using votes that expressed preferences. Out of the many different heuristics for computing the Schulze method proposed, the Schwartz set heuristic was used to determine the most preferred mapping. Table 3.5 summarises the winner of the mapping strategy. As shown in Table 3.5(b), the Schwartz set is Dif_jet (Differential movement of CE to Jet colour map) which outscored the others 3 to 0.

Table 3.5: Most preferred key-to-colour mapping strategy: (a) Pair-wise winners; (b) Schwartz set is Dif_jet

(a)

Pair	Winner	Winning Margin (%)
Random vs. XYZ_jet	XYZ_jet	15.69
Random vs. Rel_jet	Rel_jet	50.98
Random vs. Dif_jet	Dif_jet	60.79
XYZ_jet vs. Rel_jet	Rel_jet	45.10
XYZ_jet vs. Dif_jet	Dif_jet	54.90
Rel_jet vs. Dif_jet	Dif_jet	33.33

(b)

Mapping	Wins	Defeats
Dif_jet	3	0
Rel_jet	2	1
XYZ_jet	1	2
Random	0	3

Table 3.6: Most preferred colour scheme for key-to-colour mappings: (a), (c) Pair-wise winners; (b), (d) Schwartz sets are Rel_jet and Dif_jet

(a)

Pair	Winner	Winning Margin (%)
Rel_rb vs. Rel_jet	Rel_jet	11.76
Rel_rb vs. Rel_gray	Rel_rb	3.92
Rel_jet vs. Rel_gray	Rel_jet	19.61

(b)

Mapping	Wins	Defeats
Rel_jet	2	0
Rel_rb	1	1
Rel_gray	0	2

(c)

Pair	Winner	Winning Margin (%)
Dif_rb vs. Dif_jet	Dif_jet	36.00
Dif_rb vs. Dif_gray	Dif_rb	12.00
Dif_jet vs. Dif_gray	Dif_jet	18.00

(d)

Mapping	Wins	Defeats
Dif_jet	2	0
Dif_gray	1	1
Dif_rb	0	2

Table 3.6 summarises the winner of the colour schemes. As shown in Table 3.6 (b) and Table 3.6 (d), the Schwartz sets are Rel_jet (Relative movement of CE to Jet colour map) and Dif_jet (Differential movement of CE to Jet colour map) as they beat all the others 2 to 0. Both Rel_jet and Dif_jet mapping uses the jet colour scheme. Therefore, it was concluded that the Jet colour map is the most preferred by our subjects.

Based on the results of this preliminary study, it was decided to select a colour from the Jet colour scheme (a continuous colour spectrum that begins with blue, and passes through cyan, yellow, orange, and red) corresponding to the differential movement of the CE. The selected colour was used as the background colour of the music visualiser.

Visual display

As described earlier, a MIDI music stream was analysed by a Max/MSP patch and data corresponding to different musical features were sent to an Adobe Flash graphics application. Each received data point corresponding to a musical feature triggered a spatially localised animation sequence in the display area. The shape, colour and screen position of the visual effects depended on the music-to-visual mappings described earlier in this chapter. A screen capture of the proposed music visualiser is shown in Figure 3.13 but obviously cannot convey the effect of a visual display corresponding with a piece of music; this must be left to the imagination of the reader.

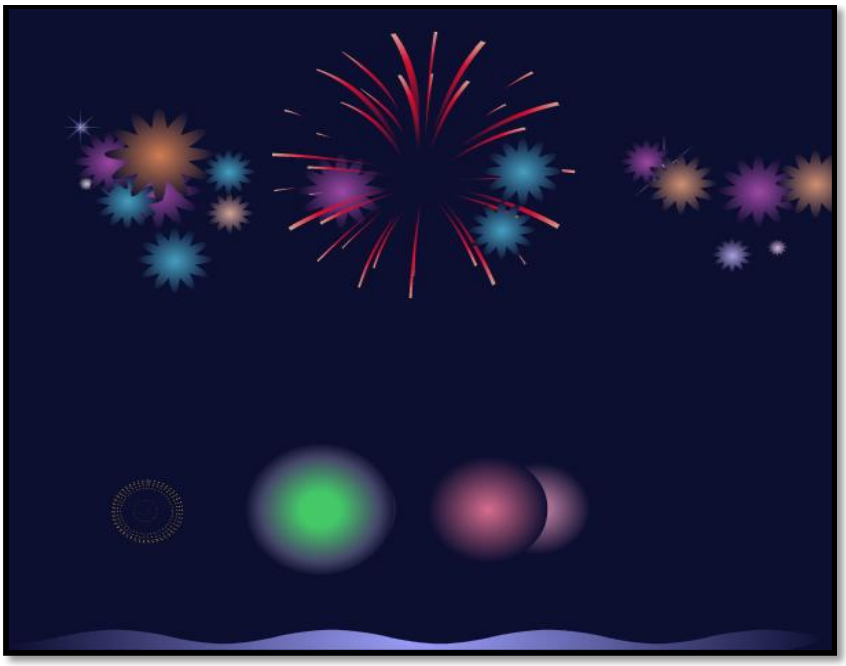


Figure 3.13: Screen capture of the heuristic based music visualiser

.

3.2.3 Preliminary user study of the music visualiser based on heuristics

A user study was conducted to evaluate the users' enjoyment of the proposed music visualiser based on Csikszentmihalyi's "theory of flow" [95]. There were several reasons for conducting this preliminary user study. The primary objective was to obtain more user feedback to improve the design. In order to find out whether the rule-based music visualisation alone is able to enhance the listening experience, rule-based visuals alone and rule-based visuals with music were tested. In addition, random visuals (the same visuals effects with no correspondence to music) were used as a control. Following components of the theory of flow was used to determine if a person is really having an optimal listening experience.

- Concentration: Users will not be distracted by things happening around them while listening/watching.
- Centring of attention: Users would feel deeply involved in listening/watching without any effort.
- Timeliness: Users would focus on present and not notice time passing while watching/listening the visualisation/music.
- Control of actions: Users would feel that they can understand the relationship between music and the visuals.
- Intrinsic motivations: User will continue to watch/listen due to the enjoyment an experience it provides.

Participants

Eighteen participants of this study were recruited from the university community. There were 9 males and 9 females and all of them were aged 18-30 years. All the participants had normal vision and hearing. Seventeen participants had some formal education in music. Participants were asked to complete a written consent form before taking part in the study and the study was conducted in accordance with the ethical research guidelines provided by the Internal Review Board (IRB) of the National University of Singapore and with IRB approval.

Apparatus

The study was carried out in a quiet room. A Dell™ Inspiron 530 Desktop with the following specifications was used to present the music and visual displays.

- 256MB NVIDIA® Geforce® 8600GT graphics card
- Dell 22" Widescreen Flat Panel LCD Monitor
- Dell A525 Stereo Speakers with Sub-woofer

Since this was a preliminary study, only one music sample was used. (A MIDI rendition of the song "Words" by popular Irish boy band Boyzone of the 1990s). Participants were given the chance to be familiar with the experimental settings. Once they were ready, a stimulus was presented and they were given a questionnaire after each trial. The questionnaire was developed based on the five components of the "flow" described in the previous section. There were statements corresponding to each of the components. Participants had to indicate (on a scale of 1 to 5) the level of agreement to each statement. Strong agreement to a statement (with a value of 5) corresponds to an optimal flow experience. This method is different from the Flow State Scale (FSS) method which was used in the later studies.

Procedure

The experiment was designed to comparatively evaluate four different conditions: music alone, music with rule-based visuals, rule-based visuals without music and music with random visuals as shown in Table 3.7. The sound levels of the speakers were calibrated to the participant's comfort level. The duration of the music sample was approximately 1 minute. For each participant, tests A, B, C and D were presented in a random order. Each participant took about 15-20 minutes to complete the experiment. It took 2 days to collect all the responses.

Table 3.7: Description of the tests used to evaluate the heuristic based music visualiser

Test ID	Music	Visual Effects	Description
A	ON	No visual effect	Follow the music
B	OFF	Visuals are generated by the music using the rule based mapping where there is a direct correspondence between music and visual patterns	Pay attention to the visual patterns.
C	ON	Visuals are generated by the music using the rule based mapping where there is a direct correspondence between music and visual patterns	Follow the music while paying attention to the visual patterns.
D	ON	Random visual effects where there is no correspondence between the music and the visual patterns.	Follow the music while paying attention to the visual patterns.

Results

The mean "flow" score was obtained combining the response for 10 rating scale questions (each rating scale having 5 categories). Therefore, the number of possible values for the mean "flow" score included all possible values from 0 to 1. Although rating scales are technically ordinal scales, in this situation we can treat "flow" score as a continuous variable and apply parametric methods [96, 97].

Figure 3.14 shows the mean "flow" score for the four different conditions tested in this experiment. One-way ANOVA test was used to analyse the data. The analysis showed that listening condition has a significant effect on the mean "flow" score $(F(3,68) = 11.77, p = 0.00001)$.

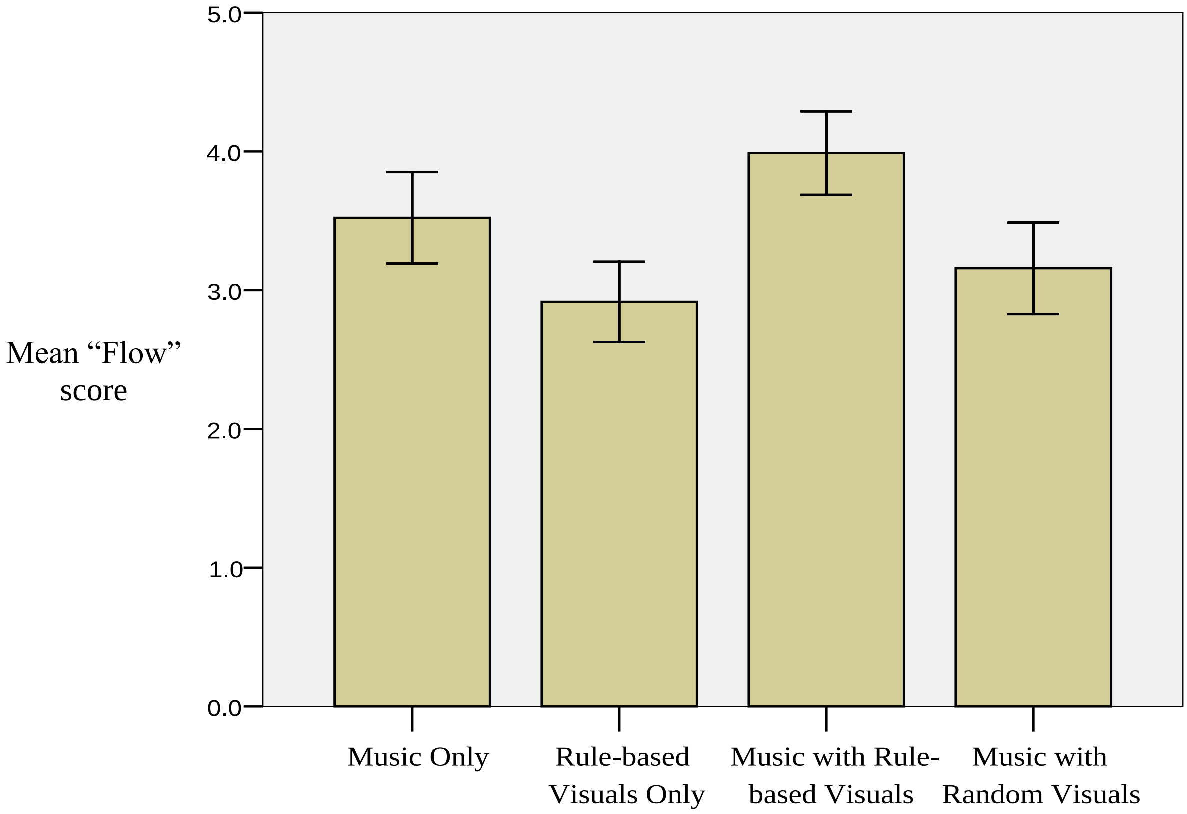


Figure 3.14: Plot of the mean "flow" score for the four different listening conditions tested in the experiment. The error bars show the 95% confidence interval.

Tukey HSD tests (Appendix A2) were performed to compare the mean "flow" score for different conditions. In this situation, Tukey's HSD test was preferred over the multiple pair-wise compassions using t-tests since multiple t-tests would have introduced probability inflation.

Tukey's post-hoc comparisons of the four conditions indicated that the proposed heuristic based music visualiser gives a higher flow experience to the participants when watching visual effects while listing to music compared to watching visuals without the music ($p = 0.00001$).

The rule-based visuals-only condition received the lowest "flow" score. Since all participants had normal hearing, they preferred to hear music over watching mute visual display.

To control for the possibility that graphics enhanced the flow experience independent of their relationship to the music, we presented a combination of music and visuals where visuals were random and not triggered by specific musical features. Tukey's post-hoc comparisons showed that participants were able to experience a higher flow when watching visual effects that has direct correspondence to the features of the music than watching random visuals ($p = 0.001$). The random visuals had the same visual effects as presented in the heuristic-based music visualiser; the only difference was that the visual effects were presented in a random manner so that the visual effects did not correspond to the music. Therefore, this result further emphasises that a direct correspondence (synchronisation) between music and visuals is an important factor in having an optimal listening experience.

3.3 Conclusions

The first part of this chapter discussed the findings from a preliminary survey that was conducted with deaf people from multiethnic backgrounds. We found that regardless of the level of deafness, our participants expressed the desire to enjoy music more. In addition, deaf people reported that a device that provided visual effects or vibrotactile stimuli were more helpful than the other assistive devices they have used.

The central question of this research was whether or not a musical experience can be conveyed by sensory channels other than sound received via the external ear canal. We developed a system that codes sequences of information about a piece of music into various visual sequences in real-time. Using this system, we initially created a basic music visualisation scheme that was exploratory and experimental. We evaluated this system with normal-hearing people since we could compare the effect of the visuals with and without audible music. Although a basic visualisation scheme was used, and it was evaluated with people with normal hearing, it proved to be a reasonable start. From user feedback it was evident that music driven visual effects directly corresponding to the important features of the music being played could enhance the listening experience. Findings from the user study with normal-hearing people and the requirements of hearing-impaired people might be different. Thus, two deaf musicians were consulted to obtain more informed feedback. They suggested that too much visual information is hard to follow and therefore much simpler mappings were preferred. The use of a tactile input was also strongly suggested by the musicians. Furthermore, they expected that, with a haptic input, the visual display could play a complementary role.

Based on the findings summarised in this chapter, and with regular feedback from the hearing-impaired, an integrated system consisting of a visual graphic component and a vibrating wooden chair was designed and evaluated. Details of this prototype system are given in the next two chapters.

Chapter 4

System Description

This chapter describes a prototype system designed to enrich the experience of music for the deaf by enhancing sensory input of information via channels other than in-air audio reception by the external ear canal. The prototype system has two main components—an informative visual display and a Haptic Chair. The first part of this chapter describes the visual display that codes sequences of information in a piece of music into various visual sequences in real-time. The second part of the chapter describes the development of a "Haptic Chair" that provides sensory input of vibrations via touch and possibly also via bone conduction of sound. This prototype system was developed based on an initial concept guided by information obtained from a background survey conducted with a cohort of hearing-impaired students.

4.1 Visual Display

The system that codes sequences of information derived from a piece of music into a visual sequence in real-time was described in Chapter 3. Since that design was formerly evaluated by people with normal-hearing, two deaf musicians (a pianist and a percussionist) were consulted to get more insight into the requirements of hearing-impaired people. Based on their feedback, the former system was modified. The visual display described in this chapter has visual effects corresponding to note onset, note duration, pitch of a note, loudness, instrument type, and key changes.

4.1.1 Music-to-visual mapping

As suggested by Jones and Nevile [43], high notes were mapped to small shapes and low notes were mapped to large shapes. This mapping was approved by the two deaf musicians who were involved in the development phase. Similarly, following the findings from psychophysical experiments [79], the amplitude of a particular note was mapped to the brightness of the corresponding visual object. The informal interviews with deaf musicians suggested that they would like to be able to differentiate between the various instruments being played. Timbre is too complex to be represented by a small number of changing values and there have been many approaches to characterise musical timbre [98]. However, the two deaf musicians who worked with us suggested using different colours to represent different instruments such that each instrument being played at a given time is mapped to a unique colour. Both deaf musicians were comfortable with the key-to-background colour mapping proposed in Chapter 3. Moreover, a “movie roll”-type of visual presentation (where the entire display is used to show instantaneous events) was used to provide a more natural "musical listening process". In a pilot study, our deaf musicians confirmed the more natural feel of the "movie roll"-type presentation.

4.1.2 Implementation

The same system architecture (Max controlled Flash animations) described in Chapter 3 was used to implement the visual display. Flint Particle System—a set of Actionscript 3.0 classes—was used to create the visual effects [99]. Flint Particle System offers a wide range of functionality to create common particle behaviours. Using these classes, a music animator was developed which received the real-time music data from Max/MSP application and generated corresponding visual effects on the screen.

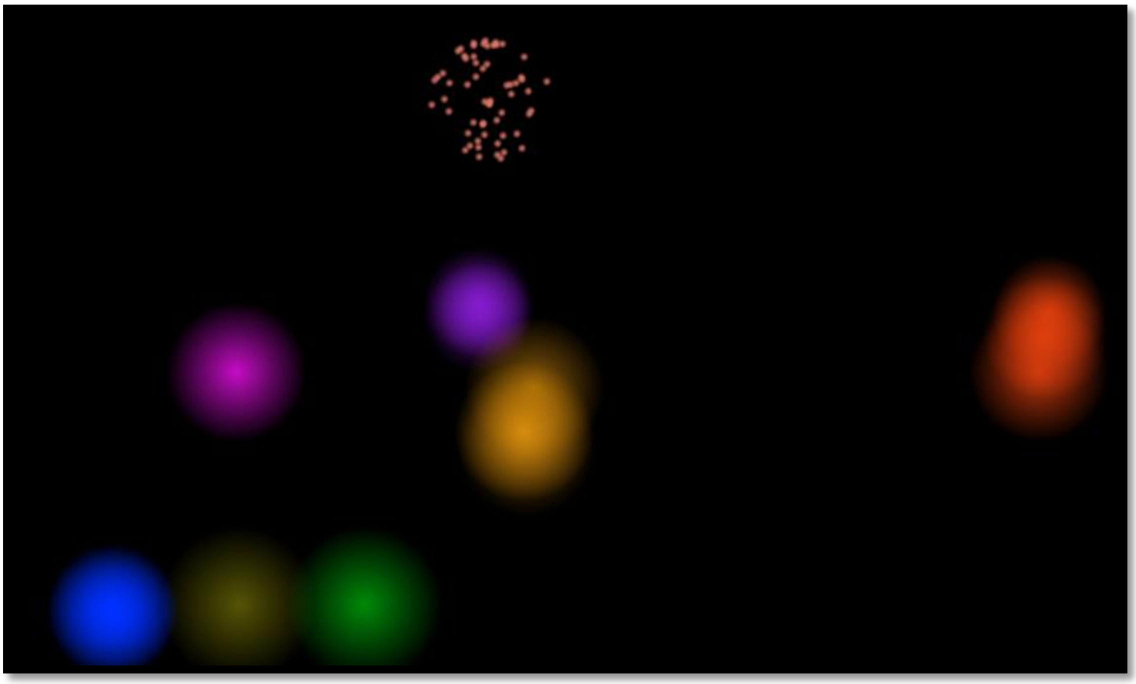


Figure 4.1: Screen capture of the visual display with real-time abstract animations corresponding to music

- The percussive instruments were separated from the non-percussive ones such a way that corresponding visual effects appeared at the bottom of the screen.
- Particles corresponding to non-percussive instruments appeared at the centre of the screen.
- Based on the number of instruments being played, the screen was divided into columns, and particles corresponding to one instrument appeared in one column.
- Particles were assigned specific colours based on the type of instrument. For every song, a specific colour was assigned to a particular instrument and remained consistent for that instrument for the entire piece.
- The position, size and brightness of the particle were depended on the pitch and velocity of the musical note.
- The differential movement of the centre of effect (CE) was used to generate the background colour of the screen.

Based on these factors, the visual display produced graphics corresponding to the music data fed into it. Figure 4.1 shows a screen capture of the visual display but obviously cannot convey the effect of a visual display corresponding with a piece of music; this must be left to the imagination of the reader.

4.2 The Haptic Chair

The literature review, background survey results and informal interviews with deaf musicians suggested that if vibrations caused by sound could be amplified and sensed through the body as they are in natural environmental conditions, this might increase the experience and enjoyment of music over a mute visual presentation or by simply increasing the volume of sound. This led to the development of a device to achieve the above which was called the “Haptic Chair”. Since hearing-impaired people are used to sensing vibrations from their fingertips to the soles of their feet [6], it was important to provide a vibrotactile feedback to the entire body. A chair was a simpler, multi-purpose structure to test this concept of “whole body stimulation”. O'Modhrain has suggested that a haptic feedback provided through a computer controlled actuator embedded in a sofa would “create a greater sense of immersion” [100]. In addition, Martens has shown that users (with normal-hearing) generally preferred to watch action-oriented DVD films while sitting in a chair mounted on a haptic platform [101]. In fact, our initial informal tests with hearing-impaired musicians suggested that the Haptic Chair enabled listeners to be comfortably seated while being enveloped in an enriched sensation created by the received stimuli.

4.2.1 Structure of the chair

The current concept underlying the Haptic Chair is to amplify vibrations produced by musical sounds, delivering them to different parts of the body through the chair without adding any additional artificial effects into this communications channel, such an approach might be used in future if it is shown to produce better results. But the most basic approach needed to be formally tested before making major changes to the acoustic signal. The first version of the chair used contact speakers (SolidDrive™ SD1 and Nimzy™ Vibro Max—Figure 4.2) which are designed to make most surfaces they are attached to vibrate and produce sound. The quality and frequency response of the sound these contact speakers produce is similar to that of conventional diaphragm speakers. This is important since many partially deaf people can hear some sounds via in-air conduction through the “conventional” hearing route: an air-filled external ear canal.

Figure 4.2: Contact speakers: (a) SolidDrive™ SD1; (b) Nimzy™ Vibro Max

After exploring many different materials including solid wood, laminated wood, glass, metal, plastic and configurations for the chair frame, a densely laminated wooden chair that was widely available at relatively low cost ("Poäng" made by IKEA®) was chosen. The frame was made of layer-glued bent beech wood which provided flexibility, and solid beech cross-struts that provided rigidity. The chair was able to vibrate relatively freely and could also be rocked by the subjects. A contact speaker was mounted under each arm-rest, one under a similar rigid, laminated wooden foot-rest (also "Poäng" by IKEA), and one on the back-rest at the level of the lumbar spine (Figure 4.3).

Initially, a thin but rigid plastic dome was mounted over each hand-rest to amplify vibrations produced by high frequency sounds for detected by the hands and fingers. The domes also provided an ergonomic hand rest that brought fingertips, hand bones and wrist bones in contact with the vibrating structures in the main body of the chair. The arm rests also served to conduct sound vibrations to the core of the body. The sound signal was presented in conventional stereo output to the right and left arm rests. A textured cotton cushion with a thin foam filling supplied with the IKEA

Poäng” chair was used to increase physical comfort but not significantly interfere with haptic perception of the music. It might have reduced bone conduction of sound particularly to the skull but since this was not the specific focus of the present study, the cushion was used because it increased the overall comfort of the user.

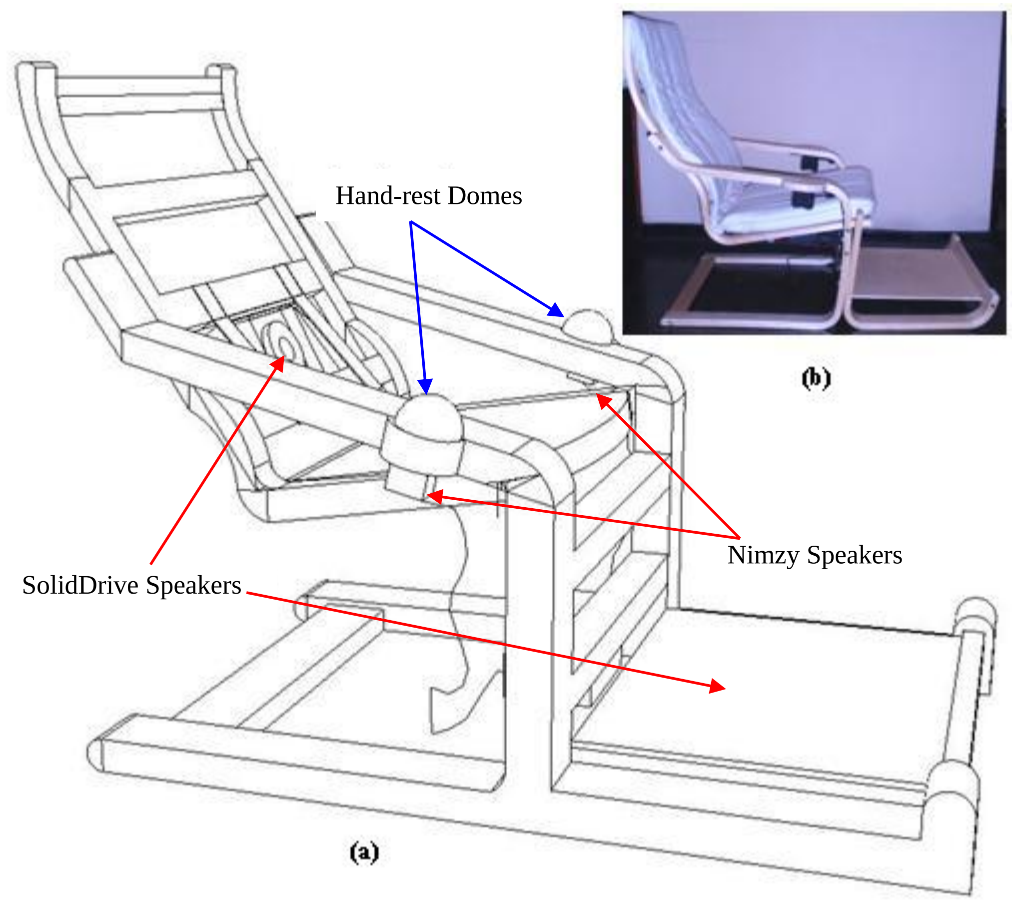


Figure 4.3: Haptic Chair: (a) Sketch; (b) Actual chair

4.2.2 Frequency response characteristics of the chair

To study the frequency response of the Haptic Chair, vibrations were measured in different parts of it in response to different input frequencies in the range of 50-5000Hz. The lower frequency was limited by the response of the contact speakers and upper limit was chosen such that it effectively covered the range of most musical instruments [58]. A notebook computer running Windows® XP and Adobe® Audition® 2.0 was used to send the input signal which was created offline using MATLAB™. To measure the strength of vibrations, an accelerometer (3041A4, Dytran Instruments, Inc.) was mounted at various locations on significant parts of the chair as indicated in Figure 4.5. The accelerometer has a flat (±2dB) frequency response in the rage 50Hz to 10 kHz. The output of the accelerometer was connected to a signal conditioner and output of the conditioner was collected by a data acquisition module (USB-6251, National Instruments). The data was then processed and collected by a notebook computer running LabVIEW™ 8.2. A block diagram of the process is shown in Figure 4.4.

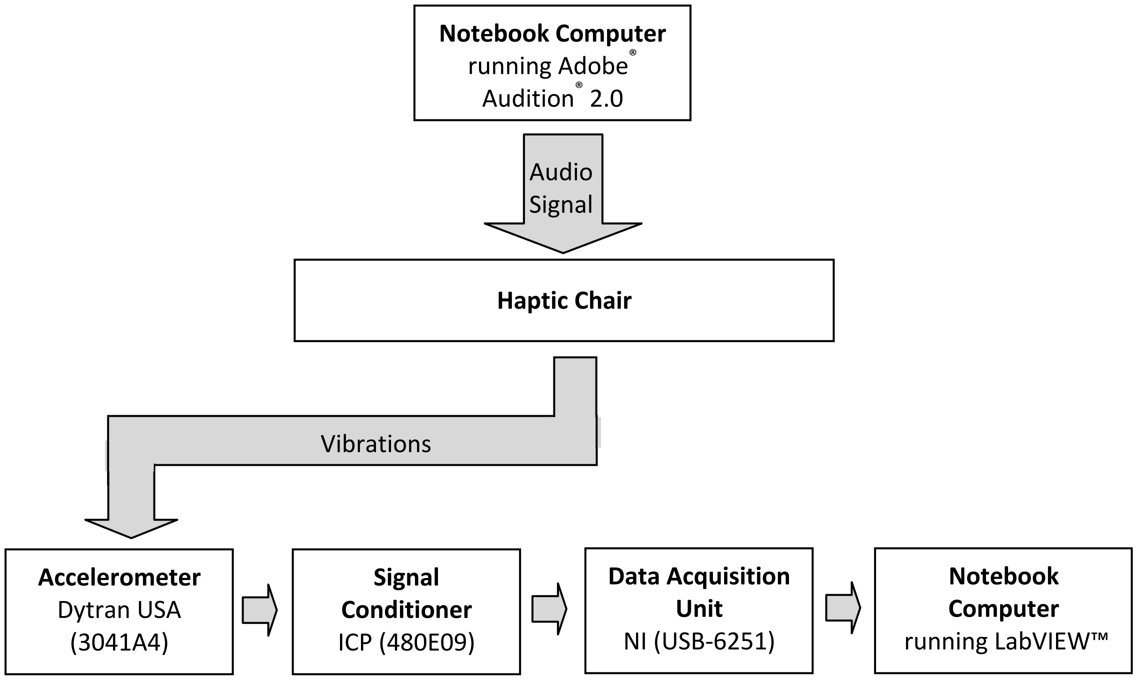


Figure 4.4: Block diagram of the equipment used in the data acquisition process

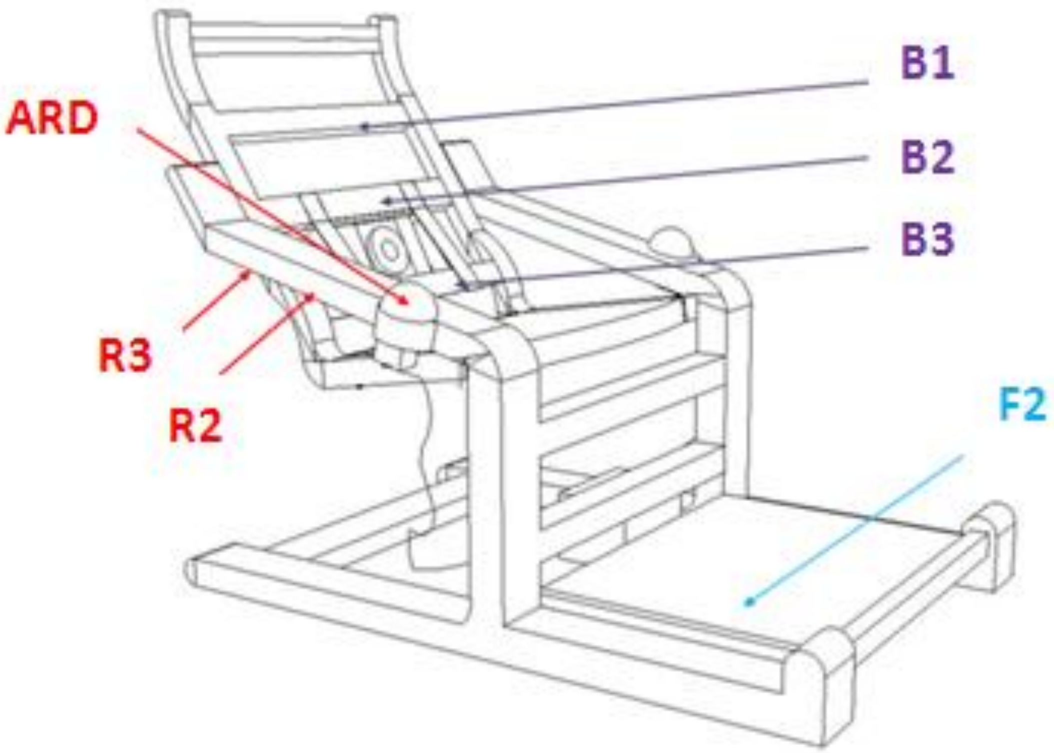


Figure 4.5 Schematic drawing of the Haptic Chair showing the locations where the accelerometer was mounted (F2–foot rest; B1 to B3–back rest; R2, R3–arm rest; ARD–plastic dome)

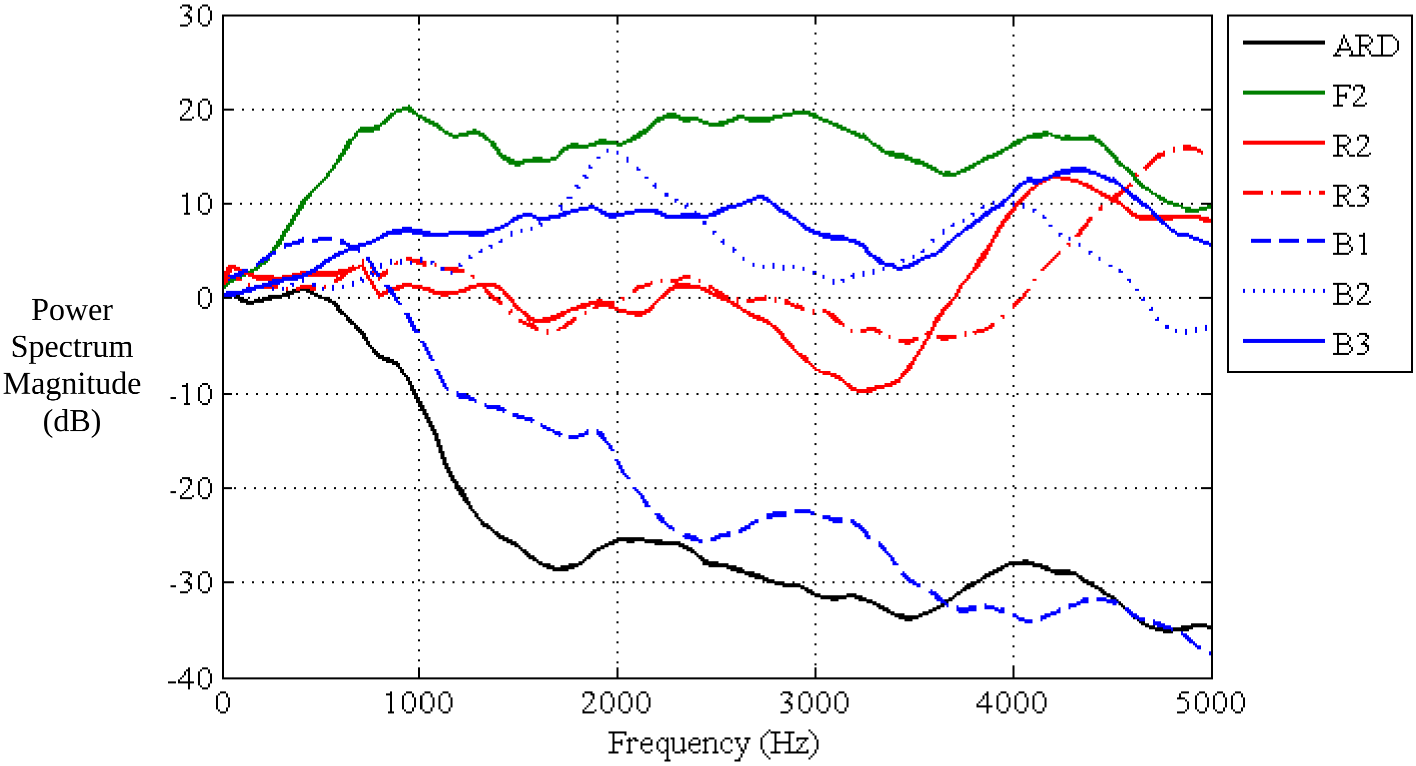


Figure 4.6: Power spectrum of vibrations measured at various positions on the chair: F2–foot rest; B1 to B3–back rest; R2, R3–arm rest; ARD–plastic dome (Power of the input is taken as the reference; i.e. 0 dB)

The frequency response (in response to a frequency chirp from 50Hz to 5000Hz that effectively excite all the frequencies in the range) was measured at various points of the chair as indicated in Figure 4.5. This response was measured under the loading condition (i.e. a person was sitting in the chair). Since both contact speakers and accelerometer were directly attached to the structure of the chair, we did not observe a significant damping due to loading effect. The obtained frequency response is shown in Figure 4.6. The response measured from the foot-rest produced the strongest vibrations compared to the back-rest and the arm-rests.

Measurements recorded from the wooden frame of the back-rest of the chair were different from the actual vibrations felt by the user. This is because the SolidDrive speaker mounted on the back-rest was not directly in contact with the frame of the chair but had a direct contact with the lumbar spine of the user. The user was likely to experience stronger vibrations compared to mechanically recorded responses at the back-rest locations, B1 to B3 (Figure 4.7).

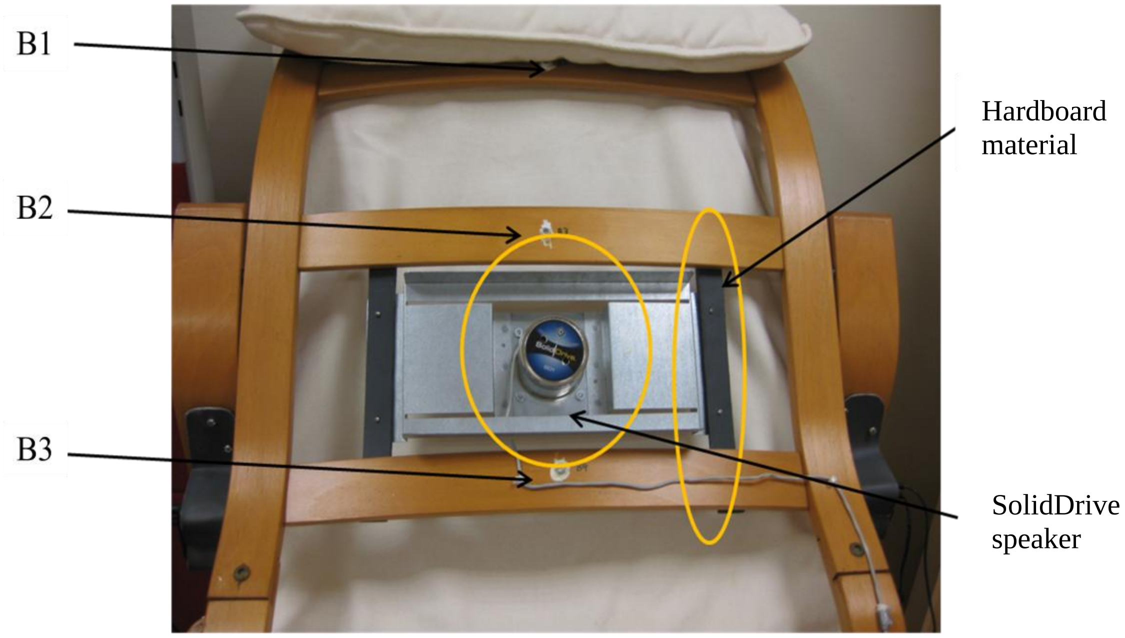


Figure 4.7: Picture of the back of the chair showing how the back speaker is attached to the chair (B1 to B3–locations where the vibration measurements were taken)

The vibrations measured from the arm-rest domes were very weak. This issue was addressed in the second prototype. To provide a qualitative analysis, a short music sample was played through the Haptic Chair and the vibration pattern was recorded. When the recorded vibration pattern was played back as an audio stream, the resulting sound had the same quality as the original signal (amplitude was dependent on the location from which the vibrations were recorded). This was a casual observation, however it supported the idea that the vibrations produced by the chair did not have any significant distortions.

4.3 User Evaluation of the System

4.3.1 Measuring musical experience

The notion of "musical experience" is often used in everyday life. However there is no widely accepted definition that adequately captures this complex concept. This phrase is often used in the ordinary everyday sense of how people psychologically experience and respond to music. In a personal email to the author of this dissertation, Professor Daniel Levitin [102] provided a simpler definition for musical experience:

> *The aesthetic and emotional appreciation of variations in pitch and rhythm over time, which typically cause a desire to synchronise one's body movements to the underlying (rhythmic) pulse.*

Even this simplified definition (which only talks about synchronising body movements to the rhythm) does not lead itself to quantifying the musical experience.

The aim of the present research project is to provide a more satisfying musical experience to a hearing-impaired person. Qualities of musical engagement have been systematically studied using self reporting methods and interviews [103, 104]. These investigations have contributed to our understanding of how people make meaning from their interactions with music in social contexts. Sloboda, O'Neill and Ivaldi [105] have studied musical engagement in everyday life using conventional "flow" methodology. Thus, we proposed to evaluate musical engagement based on Csikszentmihalyi's "theory of flow" [95]. The timelessness, effortlessness and lack of self-consciousness one experiences are what Csikszentmihalyi would describe as being "in flow". He describes "flow" as a state in which people are so involved in an activity that nothing else matters: the experience itself is so enjoyable that people will do it even at a high cost, for the sheer "joy" of doing it. Flow has been described as having nine main components [106, 107]:

- No worry of failure—a feeling of being "in control"
- Clear goals—a feeling of certainty about what one is going to do
- Immediate feedback—feedback confirming a feeling that everything is going according to plan
- Complete involvement—a feeling of being entirely focused
- Balance of challenge and skill—a feeling of balance between the demands of the situation and personal skills
- No self-consciousness—no worries about self
- Unaware of time—thoroughly focused on present and not noticing time passing
- Merger of action and awareness—a feeling of automaticity about one's actions
- Autotelic experience—a feeling of doing something for its own sake

Although "flow theory" has been widely used to analyse interactive experiences such as theatrical plays, sports or gaming, among the passive activities that can result in flow is relaxing while listening to music [108]. This explains the link between enjoying a musical performance and optimal experience—when someone is really enjoying a musical performance, he or she is said to be in “flow state”. It has also been suggested [106], [109] that the flow model could be used as a reflective tool for monitoring, regulating and assessing the learning of music. However, not all of the nine dimensions of flow described by Csikszentmihalyi are applicable for a passive activity such as listening to music. For example, when listening to music, there is no immediate feedback confirming that everything is proceeding according to plan.

Researchers have applied various methods to assess flow. Csikszentmihalyi [107] used a method called “Experience Sample Method” (ESM) which requires participants to answer a short questionnaire when they receive a random signal from an electronic beeper. Typically, participants wear this beeper for a period of one week. Jackson and Marsh [110] pointed out the need for further research to examine the validity and reliability of this method in field research. Bakker [111] suggested that a simple questionnaire may offer a reasonable alternative.

Jackson and Marsh [110] have developed and validated a questionnaire to measure the flow—a questionnaire with 36 items which they called the “Flow State Scale” (FSS). FSS evaluates the nine dimensions of flow described by Csikszentmihalyi [107]. A similar approach was used in this dissertation to quantify an optimal experience. Most of the questionnaires used in this research to measure optimal experience were developed based on the FSS. The original FSS instrument was modified in such a way that only the questions applicable to a scenario of listening to music were used. The modified FSS consisted of six statements where the participants were asked to rate a statement (of modified FSS) on a 5 point scale, ranging from 1 (strongly disagree) to 5 (strongly agree). In addition to the above quantitative evaluation, we used "observable indicators of the flow" [112, 113] to obtain a qualitative measurement. Custodero [112] originally used this method to analyze video data of young children participating in early childhood music activities. According to Custodero, the following observations indicate that a person might be experiencing a flow state [112, 113]:

- Focused and immersed in the musical activity
- Desire to repeat the task
- Visible enjoyment once the activity has finished
- Expresses feelings of high competency

On the other hand, following behaviours indicates that a person is not in flow:

- Frustration
- Anxiety
- Off task behaviour

Quantitative evaluation using "FSS score" and qualitative observation using "observable indicators of the flow" were used to cross-validate the data and thus provided a reliable method to assess whether a person is having an enjoyable experience. Nevertheless, the fact remains that a musical experience is much more than the measures of enjoyment, and complete characterisation of musical experience is still an open question.

4.3.2 The formal user study

A user evaluation was carried out with hearing-impaired people to examine the effectiveness of the proposed system. The objective of the study was to find the answers to the following:

- Whether visual display alone could provide a more satisfying musical experience to a deaf person.

- Whether the Haptic Chair alone could provide a more satisfying musical experience to a deaf person.
- Whether a combined output (visual display together with the Haptic Chair) could provide a more satisfying musical experience to a deaf person.
- What is the effect of different music genres?
- How does the level of deafness affect the musical experience?

Participants

Forty three participants (28 male participants and 15 female participants) took part in the study. Their median age was 16 years ranging from 12 to 20 years. All participants had normal vision. The participants in this study were not from the same group of subjects who took part in the background survey and informal design interviews and therefore provided us with a fresh perspective. Communication with the participants was done with the help of an expert sign language interpreter. The study was conducted in accordance with the ethical research guidelines provided by the Internal Review Board (IRB) of the National University of Singapore and with IRB approval

Apparatus

The study was carried out in a quiet room resembling a calm home environment. A notebook computer with a 17-inch LCD display was used to present the visual effects. We did not include the size of the LCD display as a variable in this study, and chose the commonly available 17-inch monitor that was both easily portable and widely available in homes and workplaces. During the various study blocks, participants were asked to sit on the Haptic Chair (keeping their feet flat on the foot rest and arms on the armrests), and/or to watch the visual effects while listening to the music, or simply listen to the music. The subjects did not report any discomfort to follow this procedure. The visual display was placed at a constant horizontal distance (approximately 150 cm) and constant elevation (approximately 80 cm) from the floor. Participants switched off their hearing aids during the study.

Procedure

The experiment was a within-subject 4×3 factorial design. The two independent variables were:

- Music sample (classical, rock, or beat only)
- Test condition (neither visual display nor Haptic Chair; visual display only; Haptic Chair only; and visual display and Haptic Chair)

The musical test samples were chosen based on the background survey results. MIDI renditions of Mozart's Symphony No. 41, "It's my life" (a song by Bon Jovi), and a hip-hop beat pattern were used as classical, rock, and beat only examples, respectively. The duration of each of the three musical test pieces was approximately one minute.

Table 4.1: Conditions tested with a piece of music in the formal user study

Trial	Visual Display	Haptic Chair	Task
A	OFF	OFF	Follow the music
B	ON	OFF	Follow the music while paying attention to the visual display
C	OFF	ON	Follow the music while paying attention to the vibrations provided via the Haptic Chair
D	ON	ON	Follow the music while paying attention to the visual display and vibrations provided via the Haptic Chair

For each musical test piece, there were four trial blocks as shown in Table 4.1. In all four blocks, in addition to the prototype system, the music was played through a normal diaphragm speaker system (Creative™ 5.1 Sound System). Before starting the blocks, each participant was told that the purpose of the experiment was to study the effect of the Haptic Chair and the visual display. In addition, they were given the opportunity to become comfortable with the Haptic Chair and the display. Also, the sound levels of the speakers were calibrated to each participant's comfort level. Once the subject was ready, sets of stimuli were presented in random order.

After each block, each participant was asked to rate their experience by answering a questionnaire. The questions were designed based on the FSS instrument [110]. Each question was rated on a 5-point scale, ranging from 1 (strongly disagree) to 5 (strongly agree). Upon completion of the four trials for a given piece of music, the participants were asked to rank these four configurations (A, B, C and D as shown in Table 4.1) according to their preference. This procedure was repeated for the 3 different music samples. Each participant took approximately 45 minutes to complete the experiment. It took 8 days to collect responses from 43 participants.

Results

Responses were analysed to address the questions presented at the beginning of this section. The overall FSS score was used as a measure of the experience. The FSS score was calculated as a weighted average of the ratings given for each question, and ranged from 0 to 1 where a FSS score of 1 corresponded to an optimal experience. Strictly speaking, individual rating scales are ordinal scales. However, in this case, the composite variable, FSS score, could be considered as a continuous variable [96, 97]. This is because the FSS score was obtained by combining the response for 6 rating scale questions (each rating scale having 5 categories). Therefore, the FSS score could potentially get any value in the range of 0 to 1.

Overall results

The mean FSS score was plotted across all experimental conditions. From the results shown in Figure 4.8 it appeared that the Haptic Chair had a dominant effect on the FSS score. Also, as might be expected, the FSS score was minimal for the control situation in which both the visual display and Haptic Chair were turned off

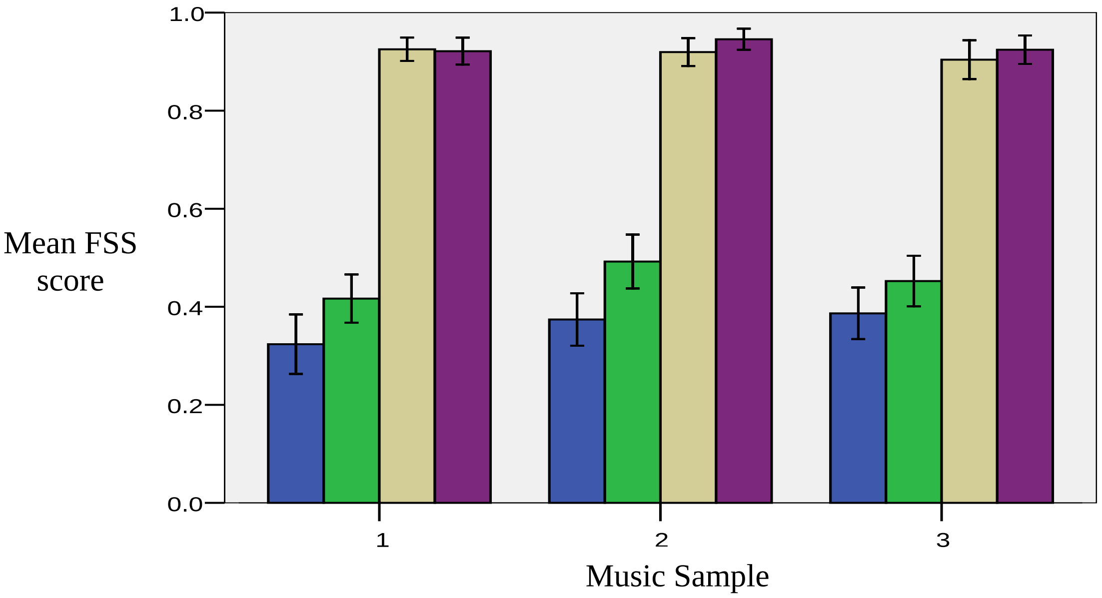


Figure 4.8: Overall FSS score for 3 music samples under all experimental conditions with error bars showing 95% confident interval. (A–music alone, B–music & visual display, C–music & Haptic Chair, D–music, visual display & Haptic Chair)

ANOVA analysis showed no main effect for music genres $(F(2,504) = 2.85, p = 0.16)$. Also, there was no interaction between music genres and the four conditions $(F(6,504) = 1.11, p = 0.35)$. However, there was a main effect for the four conditions $(F(3,504) = 589.18, p = 0.000001)$. This suggested that music genres did not significantly affect the FSS score, but the situation (whether they were watching the Visual Display, whether they were sitting on the Haptic Chair, etc..) did have a significant effect on the FSS score.

Level of deafness vs Experience

A 2-way repeated measures ANOVA analysis was carried out to study whether the level of deafness (partially deaf or profoundly deaf) has a significant effect on the FSS score (Figure 4.9). Analysis showed no main effect for level of deafness $(F(1,508) = 0.33, p = 0.56)$ and no interaction between the level of deafness and the four conditions $(F(3,508) = 1.38, p = 0.25)$. This suggested that the level of deafness did not significantly affect the FSS score.

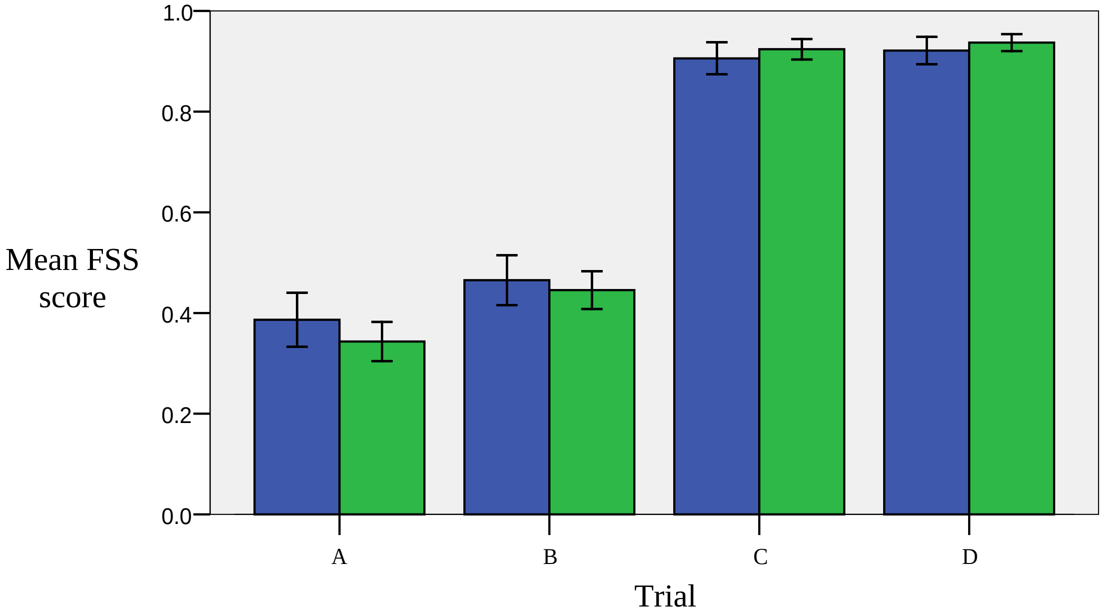


Figure 4.9: Mean FSS score of ■ partially deaf and ■ profoundly deaf subjects for four different conditions with error bars showing 95% confident interval. (A–music alone, B–music & visual display, C–music & Haptic Chair, D–music, visual display & Haptic Chair)

Comparison of the four conditions

Since there was no significant effect for music genres and level of deafness, the mean FSS score was compared across the four different experimental combinations: music only; music and visual display; music and Haptic Chair; music, visual display and Haptic Chair. As seen from Figure 4.10, conditions where subject were feeling the vibrotactile stimulation from the Haptic Chair had a dominant effect on the FSS score.

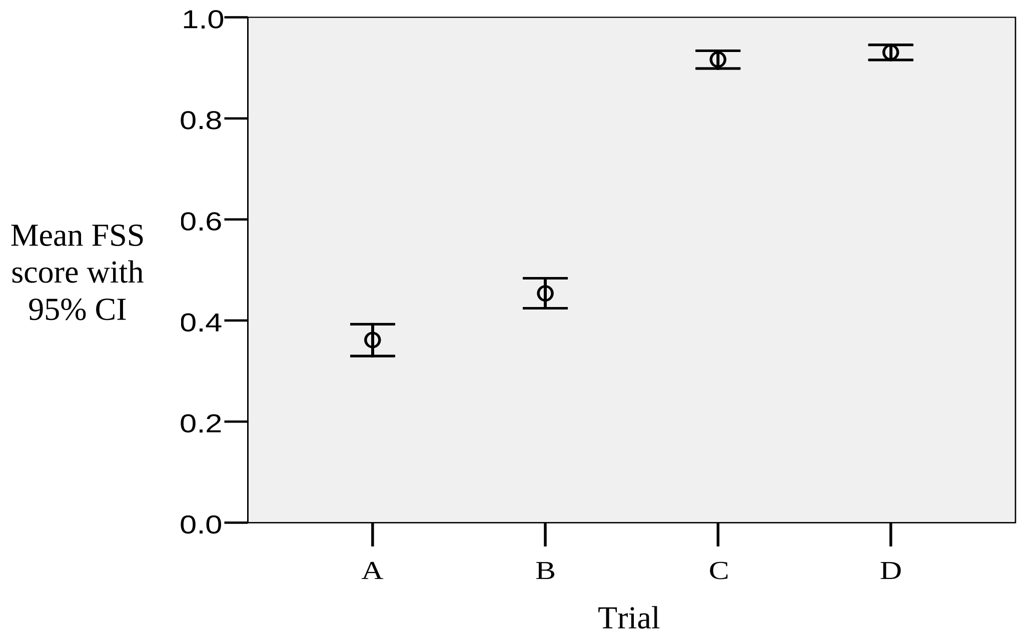


Figure 4.10: Plot of mean FSS score with 95% confidence interval (CI) for four different combinations (A–music alone, B–music & visual display, C–music & Haptic Chair, D–music, visual display & Haptic Chair)

Tukey's Honestly Significant Difference (HSD) test was used to compare the means. The outcome of this test was as follows:

- Mean FSS score of music with visuals (Trial B) was significantly higher $(p < 0.01)$ than music alone (Trial A).

- Mean FSS score of music with Haptic Chair (Trial C) was significantly higher $(p < 0.01)$ than music alone (Trial A).
- Mean FSS score of music, visuals and Haptic Chair together (Trial D) was significantly higher $(p < 0.01)$ than music alone (Trial A).
- Mean FSS scores of music, visuals and Haptic Chair together (Trial D) and music with Haptic Chair (Trial C) were significantly higher $(p < 0.01)$ than music and visuals (Trial B).
- The difference between the mean FSS score of music with Haptic Chair (Trial C) and music, visuals and Haptic Chair (Trial D) was not significant $(p = 0.38)$.

When the participants were asked to rank the most preferred configuration, 54% chose music together with the Haptic Chair. 46% ranked music and visuals together with the Haptic Chair as their first choice. None of the participants preferred other possible options (music alone, or music with a visual display).

The low FSS scores for the "music alone" and "music and visuals" options can be explained by some of the comments received from the participants. One said:

> *I can't hear with the visuals alone, but when I get the vibrations* (from the Haptic Chair)*, there is a meaning to the visuals*.

Qualitative Observations

After the formal studies were completed, we interacted with our deaf participants in a more informal way that provided insight into how our system worked in a more natural environment. A sub-group of 11 subjects were selected to listen to songs of their choice. They were asked to imagine the Haptic Chair was their own and use it in whatever way they wanted. They were also given a demonstration of how to connect an audio device (mobile phone, CD player, Apple iPod, or notebook computer) to the Haptic Chair, and they were free to choose whether or not to use their normal hearing aids. Their behaviour was observed and, after the session, questions were asked about their experience.

One very excited participant said that it was an amazing experience unlike anything she had experienced before. She reported that now she feels like there is no difference between herself and a person with normal hearing. She preferred combination of the Haptic Chair and visual display the most. She said, if she could see the lyrics (karaoke-style) and if she had the opportunity to change the properties of this visual display (colour, objects, how they move, etc.) however she wishes, that would make the system even more effective.

Many of the participants told us that they could identify the rhythm of the song and could "hear" the song much better compared to when using standard hearing aids. Another user mentioned that he wanted to use headphones together with the chair so that he could detect the sound through the headphones as well. A few participants who were born with profound deafness said that this was the first time they actually "heard" a song and they were extremely happy about it. They expressed a wish to buy a similar Haptic Chair and connect it to the radio and television at home. It was observed that majority of the profoundly deaf participants appeared to be "hearing" something when they were sitting on the chair. The following comments were encouraging:

Yes, I can hear from my legs!

I will ask my father to buy me a similar chair.

Now there is no difference between me and a normal hearing person. I feel proud.

Discussion

The statistical analyses in the previous sections showed that the Haptic Chair has the potential to significantly enhance the musical experience of person with hearing impairment. The music samples we used were only one minute long. This was because we limit the total time per participant to be 45 minutes since a longer time would have reduced the concentration and interest of the participants. Within 45 minutes we could test 3 music samples (of duration 1 minute) as each sample was played 4 times under 4 different conditions and the subjects had to answer several questions after each trial. At the experimental design stage, we thought a 1 minute of listening would be adequate for our hearing-impaired participants to answer the questionnaire.

Our observations during the formal study were similar to the "observable indicators of flow experience" reported by Custodero [113]. This indicates that duration of 1 minute was long enough for them to have an enjoyable experience. For example, many participants told us that the haptic chair experience was an "amazing experience". One minute of listing (while receiving an enhanced feedback) might be long enough for a hearing-impaired person to get into a flow state. Moreover, we had an informal session with our participants where they were allowed to use the chair for a longer period (more than 10 minutes). The observed reactions were similar to what we saw during the formal study where they listen to 1 minute long music samples. These observations supported the reliability of the results based on FSS scores.

4.4 Conclusions

This chapter has described the design and evaluation of a prototype system consisting of a visual display and a Haptic Chair. The display shows abstract visual effects corresponding to note onset, note duration, pitch of a note, loudness, instrument type, and key changes. The Haptic Chair provides vibrotactile feedback to the whole body including fingertips, hand, feet and lumbar-spine area. The system was evaluated based on FSS instrument [110] and qualitative observations. Forty-three deaf students with varying degrees of deafness took part in the formal study. All three different music genres used in the study received highly positive feedback. In addition, the level of deafness did not have a significant effect on the level of enjoyment. The overall results of the user evaluation studies supported the hypothesis that a combination of haptic and visual input would enhance the music experience of the hearing-impaired

.

Chapter 5

Refinements to the System

In the Chapter 4, it was observed that encouraging results were largely due to the influence of the Haptic Chair. The strategy of using abstract animations driven by music did not make as much impact as expected. Therefore, alternative means of using the visual channel were explored. A formal user study was conducted to evaluate the effectiveness of different visualisation approaches. On the other hand, since the initial version of the Haptic Chair received very positive feedback, only minor design modifications were made to it. The highly positive feedback received for Haptic Chair was further verified through another user study conducted over an extended period of time. In addition, we conducted a study to compare the vibrotactile thresholds for people with normal hearing and those with hearing impairments.

5.1 Visual Display

One of the significant observations made during the formal user study of the prototype system described in the previous chapter was that the impact of the abstract visual patterns were low compared to the responses received for the Haptic Chair. Therefore, several attempts were made to improve the visual display. One obvious extension to the visual display introduced in Chapter 4 was to incorporate 3D effects. Apart from that, a completely new approach was taken where human gestures synchronised with music were used to convey a better musical experience.

5.1.1 Implementation of 3D abstract patterns

It can be argued that a 3D visual display might provide more options to display. The additional dimension of 3D implementation allowed us to represent the time in a more intuitive manner. For example, as audible sound faded away, the corresponding sound also faded away into the screen. This was clearly visible in 3D version compared to 2D version. The Flint Particle Library (version 2.0) [99] was used to implement the 3D effects into the visual display described in previous chapters.

Another particular improvement made using the 3D capabilities was making the particles corresponding to non-percussive instruments appear in the centre of the screen with an initial velocity towards the user then accelerating away from the user (into the screen). As a result, it appeared to the user that the particle first came closer for a short instant, and then receded, slowly fading away as the corresponding note “died out” in the music piece. This movement greatly improved the appearance of the animator. The colouring and presentation of particles were kept consistent with that of the 2D implementation described in the previous chapter. As for the percussive instrument based particles, the positions were still kept at the bottom of the screen in the 3D view. However, the behaviour was changed so that when such a particle appeared on screen, it made a rapid upward movement before disappearing. This upward movement simulates a pulsating effect corresponding to the beat of the music. Our deaf musicians reported that they usually feel the sounds of percussive instruments through their legs as if the sound came up from the ground. Therefore, the pulsating up-and-down movements corresponding to percussive instruments were expected to enhance the visual effect of the percussive instruments in the music flow.

In general, 3D visuals have the potential to increase the richness of mappings as well as the aesthetics of the display over a 2D design. However, the overall responses obtained from the deaf users for both 2D and 3D abstract patterns were not significantly different from each other.

5.1.2 Music visualisation with human gestures

It has often been noted that hearing-impaired people employ lip-reading as part of the effort to understand what is being said to them. One possible explanation for this comes from the hypothesis of "motor theory of speech perception" which suggests people perceive speech by identifying the vocal gestures rather than identifying the sound patterns [114]. This effect could be even more significant for people with hearing difficulties. The McGurk effect [10, 115] suggests that watching human lip-movements might substantially influence the auditory perception. McGurk and MacDonald (1976) found that seeing lip-movements corresponding to "ga" results in the audible sound "ba" being perceived as "da".

Moreover, Davidson (1993), Boone and Cunningham (2001) have shown that body movements contain important information about the accompanying music [116, 117]. This could be one of the possible explanations as to why many people tend to enjoy live performances of music, even though a quiet room at home seems to be a more intimate and pristine listening environment. Combining these factors, the effects and experiences of hearing-impaired people were explored when they are exposed to simple series of "ba" "ba" lip-movements corresponding to the beat of the music. A preliminary user study was conducted with hearing-impaired people to assess if these simple gestures have the potential to be more effective in conveying a musical experience than animated abstract patterns.

Participants

Ten participants (5 profoundly deaf and 5 partially deaf) from Dr. Reijntjes School for the Deaf, Sri Lanka took part in this experiment. They were from the age group of 12-18 years. All the participants had taken part in the previous study described in Chapter 4. However, they were new to this experiment and were unaware of the information that might have biased or skewed the results. The study was conducted in accordance with the ethical research guidelines provided by the Internal Review Board (IRB) of the National University of Singapore (NUS) and with IRB approval.

Apparatus

A 17-inch LCD display was used to present the visual effects; the visual display was placed at a fixed horizontal distance (approximately 150 cm) and fixed elevation (approximately 80 cm) from the floor. MIDI rendition of the song "It's my life" (by Bon Jovi) was used as the test sample. Since this was a preliminary study, we worked with only one music sample. The duration of the sample music was approximately 1 minute. Three different trials with different visual effects were created for testing.

- The first trial consisted of a video of a young woman making lip movements (making the phoneme "ba" "ba") and minimal facial expressions corresponding to the beats of the song (See Figure 5.2.)
- The second trial had abstract visual patterns corresponding to the features of the music as described in the Chapter 4.
- There were no visual effects for the last trial; the display was switched off while the music was played.

Procedure

The participants were asked to listen to a music piece while watching visual effects on the computer screen. Each participant experienced all three stimuli where each of them had different visual effects—abstract particle animation, human face animation and just blank computer screen with no visuals. After each trial, the participants were asked to answer questions designed to find the intensity of the "flow" experience. Trials were presented to the participants in a random order.

Results

As in the Chapter 4, user experience was measured using the Flow State Scale (FSS) score. FSS is rated from 0 to 1 where a FSS score of 1 corresponded to an optimal experience. Figure 5.1 shows the average FSS score for the three different modes—music alone, music with a visual display showing face animation, music with a visual display showing abstract patterns.

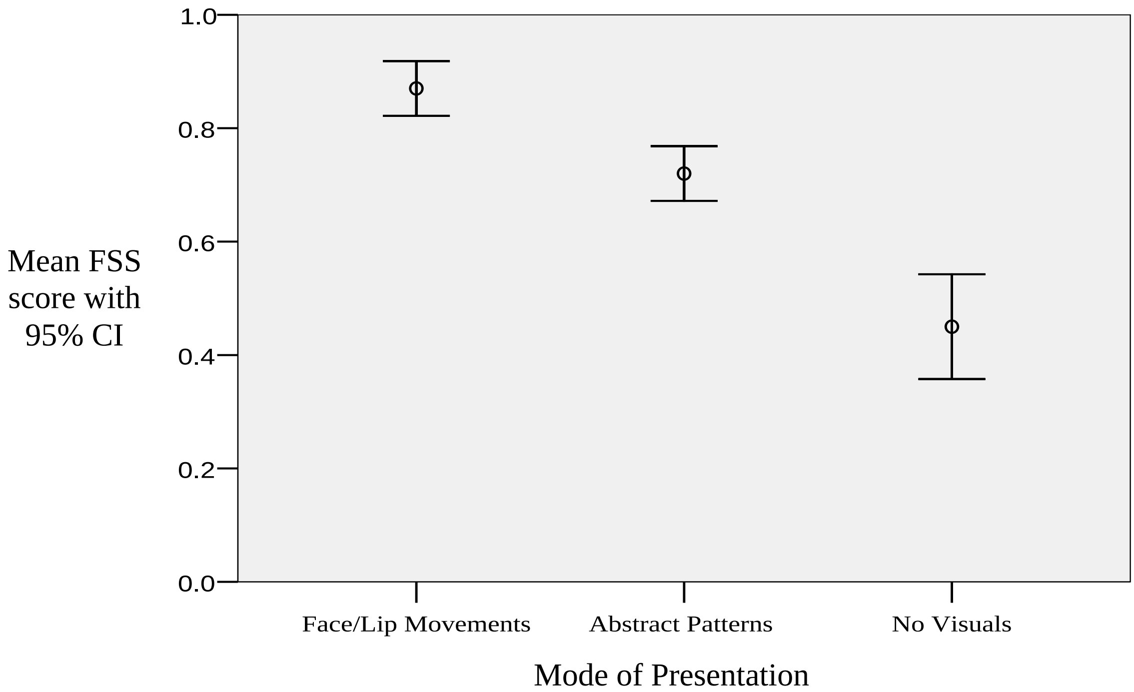


Figure 5.1: Plot of mean FSS score with 95% confidence interval (CI) for different modes of presentation

One way ANOVA analysis confirmed that type of visualisation had a significant effect on the enjoyment $(F(2,27) = 57.72, p = 0.000001)$. Tukey's honestly significant difference (HSD) test was used to compare the mean FSS scores. The outcome of this test was as follows:

- Mean FSS score of music with abstract patterns was significantly higher $(p < 0.01)$ than music alone (no visual).
- Mean FSS score of music with facial animation was significantly higher $(p < 0.01)$ than music alone (no visual).
- Mean FSS score of music with facial animation was significantly higher $(p < 0.01)$ than music with abstract patterns.

When asked to rank their preference, 8 participants (out of 10) ranked watching human face animation while listing to music as their most preferred choice while 2 ranked watching abstract particle animation as their most preferred choice. Listening to music without visuals was the least preferred mode by all the participants. Some of the significant comments received from the participants who preferred a visual display showing lip movements/facial animation, were as follows:

> *Easier to tap with the beat when I see the rhythmic facial/lip movements.*
>
> *I like the lip movements.*
>
> *It's easy to lip read and follow the body movements.*
>
> *I like to see a different girl; the body/face movement makes it look like a song.*
>
> *I couldn't understand the particle animation but I can follow the lip movements; abstract patterns don't look like a song but lip movements look like a song.*

Two participants who preferred watching a display with abstract patterns said:

> *Louder sounds can be seen with the particle animation; "ba" "ba" movement gets boring after a while.*
>
> *Abstract patterns have more colours; facial animation is boring.*

Overall the results of this user study indicate that expressive human gestures may be more effective than abstract animations in conveying a musical experience to a deaf person.

Lip/Face animation

The results from the preliminary user study showed that a facial movement involved in saying the syllable "ba" with the beat of the song might be helpful. This was assumed to be particularly true for songs with a strong beat. The closing and opening of lips while making a "ba" movement, was something deaf people were likely to understand easily as verified by the preliminary user study.

As a result, the visual display with abstract animations was replaced with a video recording of a young woman making the "ba" "ba" lip movements, even though this meant leaving aside a substantial amount of work done on realistic human motion synthesis [118].

An undergraduate student of NUS was chosen to make the facial/lip movements with the music. Apart from making the lip movements, she was given specific instructions to make other facial changes to complement the lip movement. As the lips come together, the eyelids close a bit and the eyebrows come down. Also, the head tilts slightly to the front like it would when a person listening to music is beginning to get into the rhythm of it. As soon as the lips are released to move apart, the eyes open more, eye brows move upwards and the head gives a slight jerk to move backwards, keeping the lip movement in sync with the rest of the face. In addition, she was instructed not to “dance” with the music since that would introduce additional variables. Figure 5.2 shows some screen shots of a video recording where the human character makes lip/facial movements corresponding to the music being played.

Figure 5.2: Screen captures of a young woman making lip/facial movements

Conductor's expressive gestures

The facial/lip movement strategy described in the previous section is more suitable to express music with a strong beat. However, a different approach was required to express the richness of a classical music piece. During a typical orchestral performance, an experienced conductor would transmit his/her musical intentions with highly expressive visual information through gestures. In fact, Rudolf [119] reported that conductor's left arm indicates features such as dynamics or playing style while the right arm indicates the beat. Therefore, to convey a better listening experience while listening to classical music, we showed the conductor's expressive gestures on a visual display while sitting on the Haptic Chair.

Figure 5.3: Screen captures of orchestra conductor making expressive gestures

Wöllner and Auhagen [120] have shown that watching the conductor from positions of woodwind players and first violists is perceptually more informative compared to cello/double bass position. Therefore, a video camera was positioned next to the woodwind players, from which the conductor's expressive gestures were recorded. This was done when Wang Ya-Hui, a music director of the conservatory orchestra at the Yong Siew Toh Conservatory of Music, NUS was conducting the Mendelssohn's Symphony No. 4. Figure 5.3 shows some screen shots of a video recording where Wang Ya-Hui conducts the NUS symphony orchestra.

The proposed approach of showing lip/facial movements and conductor's expressive gestures synchronised to music was compared with the previously found best case (showing abstract animations synchronised with the music). The results are summarised in Section 5.3—user evaluation of the system.

5.2 The Haptic Chair

As mentioned in the previous chapter, the original version of the Haptic Chair received very positive feedback from all of the hearing impaired users. Therefore, major modifications were not introduced. However, it was observed that the strength of the vibrations felt through the hand-rest domes was considerably weaker compared to those at other locations of the chair (especially back-rest and foot-rest). In fact, it was a possibility that the strong vibrations from the back-rest and foot-rest of the chair might override the vibrations felt at the hand-rest domes. Therefore, the rigid plastic domes that were mounted on the hand-rests were replaced by a set of flat panel speakers (NXT™ Flat Panels Xa-10 from TDK) to improve the vibrations felt by the finger tips, a particularly important channel for sensing higher frequencies. Flat panel speakers were found to be a cheaper alternative to produce stronger vibrations at the hand-rest compared to vibrations felt by the plastic dome structure that was originally on the hand-rest. These audio speakers (contact speakers) are designed to support the entire audio frequency spectrum, whereas typical vibrotactile actuators (inertial shakers, linear actuators, etc...) do not have such a broad frequency response [121]. With this modification, the location of the Nimzy contact speakers was shifted further back along the arm-rest. The purpose of this was to maintain the vibrations felt via the wooden arm-rest. These modifications are shown in Figure 5.4.

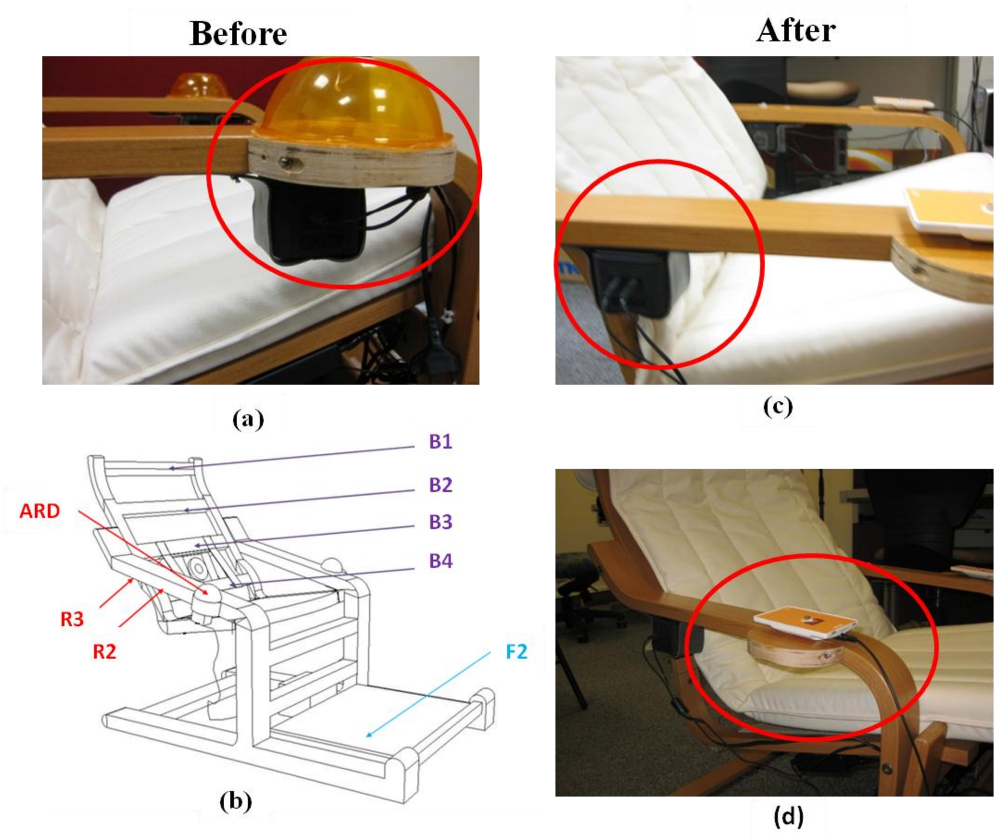


Figure 5.4: Photographs showing the modifications to the chair: (a) original version; (b) schematic drawing of chair (F2–foot rest; B1 to B4–back rest; R2, R3–arm rest; ARD–plastic dome); (c) new location of the Nimzy speakers; (d) plastic dome was replaced by Xa-10 flat panel speakers from TDK

After the modification, the frequency response of the chair at position "ARD" (position on the hand-rest as shown in Figure 5.4(b)) was compared with that of the previous prototype. The flat panel speakers attached at the "ARD" position did not have a significant effect on the response from the other positions of the chair. This is because the flat panel speakers do not operate in the same way as the contacts speakers. Since the flat panel speakers operate similarly to conventional diaphragm speakers, they do not directly vibrate the structure they are in contact with. Hence, the flat panel speakers did not introduce significant additional vibration to the chair structure.

The frequency responses of the position "ARD" in the original and revised prototype are shown in Figure 5.5. From Figure 5.5, it appears that the frequency response of "new ARD" is much higher than the frequency response of the former "ARD". In other words, the introduction of the flat panel speakers provides better haptic input to the fingertips.

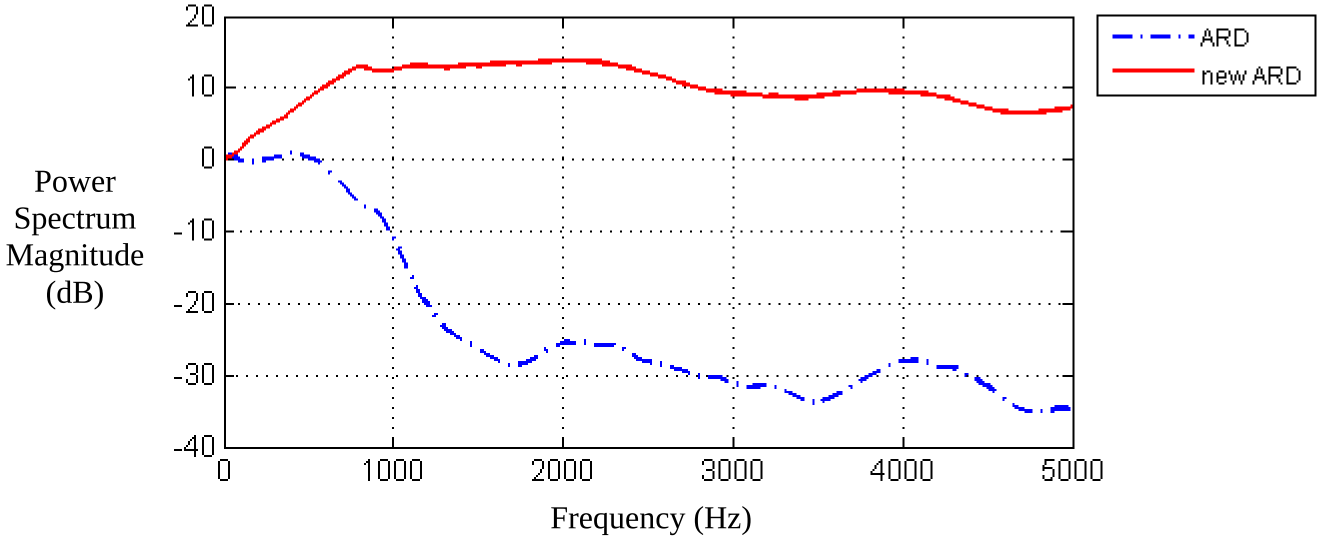


Figure 5.5: Comparison of the frequency response at the "ARD" position of initial and revised prototype

5.3 User Evaluation of the Revised System

Three different user studies were carried out to evaluate the revised version of the visual display and the Haptic Chair. The following sections include a summary of the experimental procedures, results and discussion for each study. The studies were conducted in accordance with the ethical research guidelines provided by the IRB of NUS and with IRB approval.

5.3.1 User Study: Comparison of the proposed music visualisation strategies

The objective of this study was to compare the performance of the two new visualision strategies proposed in this chapter. We hypothesized that showing human gestures directly corresponding to music might provide a more satisfying musical experience than watching previously developed abstract animations. In addition, we hypothesized that watching abstract visual effects in 2D space and 3D space might be different since 3D space offered an additional dimension to represent time. To test our hypotheses, we compared effect of watching 2D abstract animations, 3D abstract animations and human gestures while listing to music with the Haptic Chair.

Participants

Thirty six participants (21 male and 15 female) took part in the study. All had normal vision. All the participants had taken part in our previous studies. However, they were new to this study, and thus, were unaware of the information that might have biased or skewed the results. An expert sign language interpreter's service was used to communicate with the participants.

Apparatus

The study was carried out in a quiet room resembling a home environment. As previous studies, a notebook computer with a 17-inch LCD display was used to present the visual effects and was placed at a constant horizontal distance (approximately 170 cm) and constant elevation (approximately 80 cm) from the floor. During the various study blocks, participants were asked to sit on the revised version of the Haptic Chair (keeping their feet flat on the foot rest, arms on the armrests and finger tips on the flat panel speakers), and to watch the visual effects while listening to the music. Participants were asked to switch off their hearing aids during the study.

Procedure

The experiment was a within-subject 3×2 factorial design. The two independent variables were: musical genres (classical and rock) and type of visuals (2D abstract patterns; 3D abstract patterns; and human gestures synchronised with the music). MIDI renditions of Mendelssohn's Symphony No. 4 and "It's my life" (by Bon Jovi) were used as classical and rock examples, respectively. The duration of each of the two musical test pieces was approximately one minute.

For each musical test piece, there were three blocks of trials as shown in Table 5.1. In all three blocks, in addition to the visual effects, music was played through the Haptic Chair to provide a tactile input. Before starting the blocks, the participants were given the opportunity to become comfortable with the Haptic Chair and the display. The sound levels of the speakers were calibrated to the participant's comfortable level. Once each participant was ready, stimuli were presented randomly.

Table 5.1: Three different trials for a piece of music used to compare different music visualisation strategies

Trial	Visual Display	Haptic Chair	Remark
A	2D	ON	Best known condition (Discussed in Chapter 4)
B	3D	ON	Implementation of the visual effects
C	Human gestures	ON	"ba" "ba" lip/facial movement for the rock song; Orchestral conductor's expressive gestures for the classical piece

The FSS instrument described in Chapter 4 was used to measure the experience of the participants. Each participant took approximately 25 minutes to complete the experiment. The experiment was conducted over 7 days to collect responses from 36 participants.

Results

Figure 5.6 shows the mean FSS score across for each experimental condition. From the figure, it appears that watching human gestures with music has a dominant effect on the FSS score.

A 2-way repeated measures ANOVA analysis showed no main effect for music genres $(F(1,210) = 0.51, p = 0.48)$. However, there was a significant main effect for visual type $(F(2,210) = 90.29, p = 0.000001)$. There was no interaction between music genres and three visualisation strategies $(F(2,210) = 0.19, p = 0.83)$.

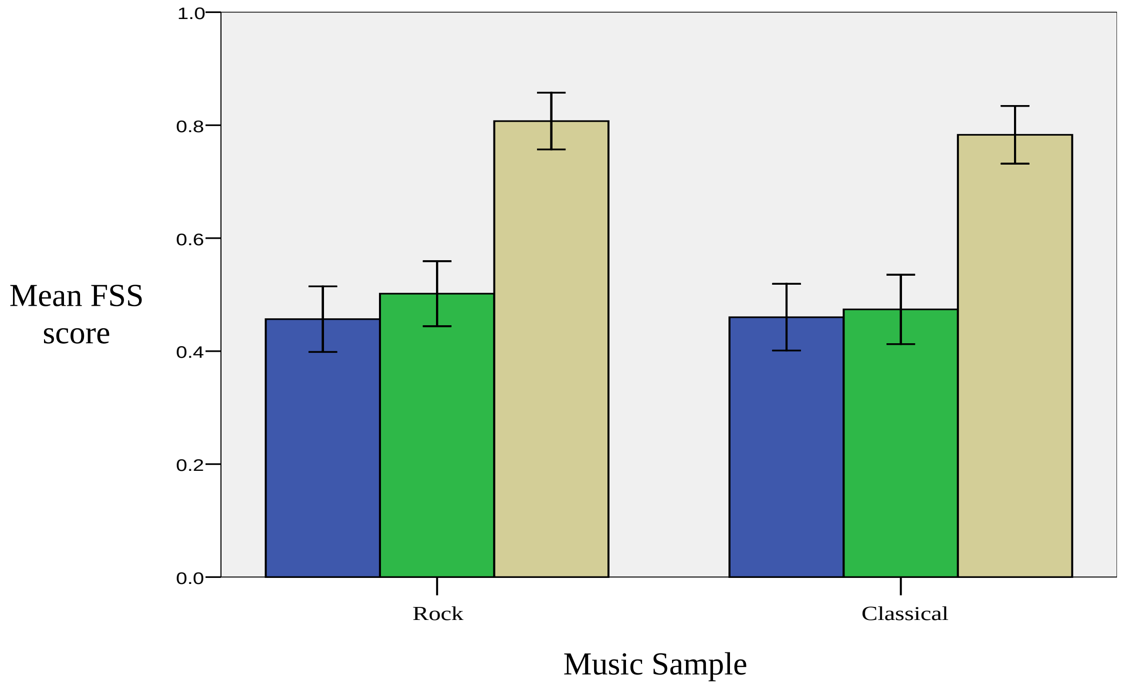


Figure 5.6: Overall FSS score for all experimental conditions with error bars showing 95% confidence interval. (■ A–2D abstract patterns, ■ B–3D abstract patterns, □ C–Human gestures)

ANOVA analysis suggested that the music genre did not significantly affect the FSS score. Therefore, results obtained from different music genres were combined and the means were compared using Tukey's HSD test. As seen from Figure 5.7, listening to music while watching synchronised human gestures and feeling the vibration through Haptic Chair (Trial C) was found to be the most effective way to convey a musical experience to a hearing-impaired person

Turkey' HSD revealed following additional results:

- Mean FSS score of watching human gestures (Trial C) was significantly higher $(p < 0.01)$ than watching 2D abstract patterns (Trial A—best case in Chapter 4) or watching 3D abstract patterns (Trial B).

- The difference between the Mean FSS scores of watching 2D abstract patterns (Trial A—best case in Chapter 4) and watching 3D abstract patterns (Trial B) is not statically significant $(p > 0.05)$.

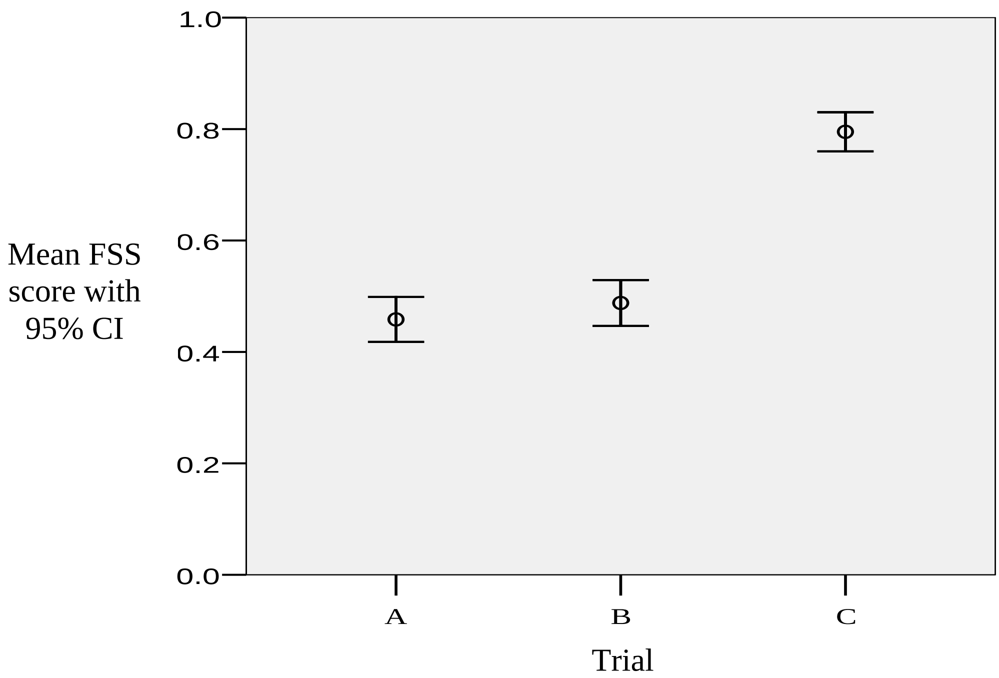


Figure 5.7: Plot of mean FSS score with 95% confidence interval (CI) for three different combinations (A–2D abstract patterns, B–3D abstract patterns, C–Human gestures)

Many participants reported that they could “hear” better when watching human gestures while listening to music sitting on the Haptic Chair. Referring to face/lip movements and conductor’s gestures, some participants said *these* (gestures) *are more musical*. Only one participant commented that the conductor’s gestures were *difficult to understand.* Perhaps this was because conductor’s gestures were particularly subtle. Overall, most of the participants liked to watch human gestures synchronised to music. From the statistical analysis, comments received from the participants and their level of excitement observed, it appeared that the use of human gestures might be the right approach for enhancing the musical experience through visuals.

5.3.2 User Study: Synchronised gestures vs asynchronised gestures

One possibility for preferring watching human gestures as opposed to abstract patterns could be the presence of a real human character. However we hypothesized that the tight synchronisation between music and gestures was an important factor to provide a satisfying musical experience. To test for this, a comparison of three different scenarios—human gestures synchronised with music, human gestures asynchronised with music and music without any visuals—was carried out. Asynchronised gestures and synchronized gestures contained the same gesturing patterns; the only difference in asynchronised gestures was that gesturing patterns and music had no correspondence.

Participants and apparatus

Twelve participants (7 male and 5 female students) took part in this study. All of them had taken part in the previous study but new to this study. They were unaware of the information that might have biased or skewed the results. As previously, an expert sign language interpreter's service was available to communicate with the participants. Same set up—a 17-inch LCD display placed at a constant horizontal distance (approximately 170 cm) and constant elevation (approximately 80 cm) from the floor in a quiet room resembling a home environment—was used to present the visual effects.

Procedure

The experiment was a within-subject 3×2 factorial design. The two independent variables were: musical genres (classical and rock); type of visuals (no visuals; music with synchronised human gestures; and music with asynchronised human gestures). Same music samples used in the previous experiment (Mendelssohn's Symphony No. 4 and "It's my life" by Bon Jovi) were used.

Table 5.2: Three different trials for a piece of music were conducted to compare the effectiveness of synchronised and asynchronised human gestures

Trial	Visual Display	Haptic Chair	Remark
A	No visuals	ON	Control case
B	Music with synchronised human gestures	ON	Gestures correspond to the music being played
C	Music with asynchronised human gestures	ON	Gestures do not correspond to the music being played

Results

Figure 5.8 shows the overall results across all experimental conditions. As might be expected, music with synchronised gestures had the maximum score, music alone was the second best and music with asynchronised gestures had the lowest FSS score.

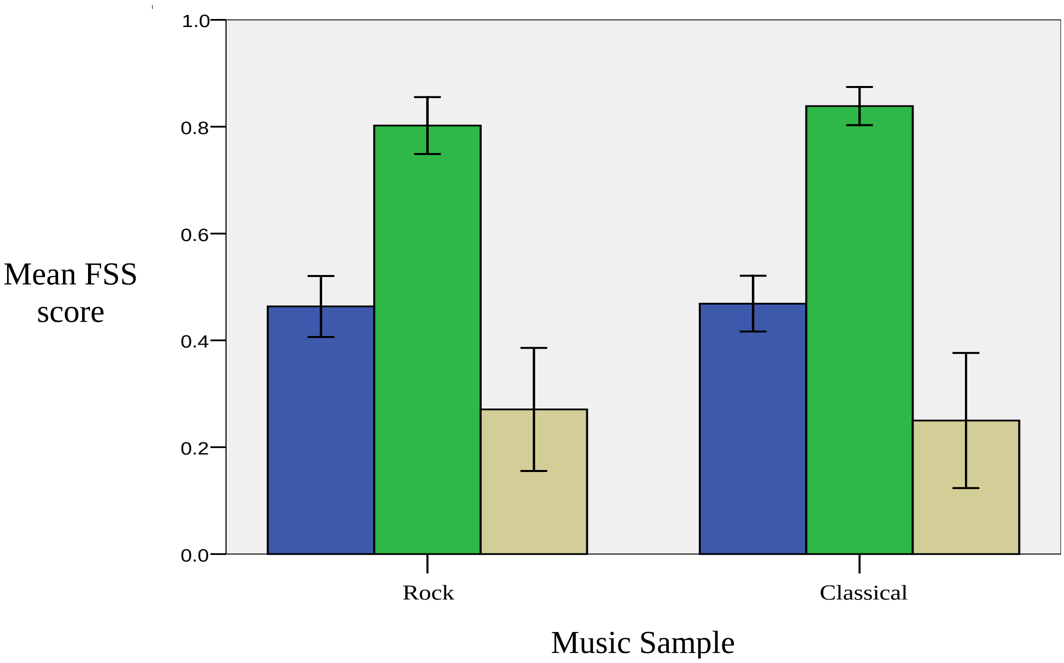


Figure 5.8: Overall FSS score for all experimental conditions with error bars showing the 95% confidence interval (A–no visuals, B–music with synchronised gestures, C–music with asynchronised gestures)

A 2-way repeated measures ANOVA analysis showed no main effect for music genres $(F(1,66)=0.53, p=0.81)$. However, there was a significant main effect for visual type $(F(2,66)=118.19, p=0.000001)$. There was no interaction between music genres and three visualisation strategies $(F(2,66)=0.303, p=0.74)$. Since music genres did not significantly affect the FSS score, FSS score was averaged across the different music samples. The results are shown in Figure 5.9.

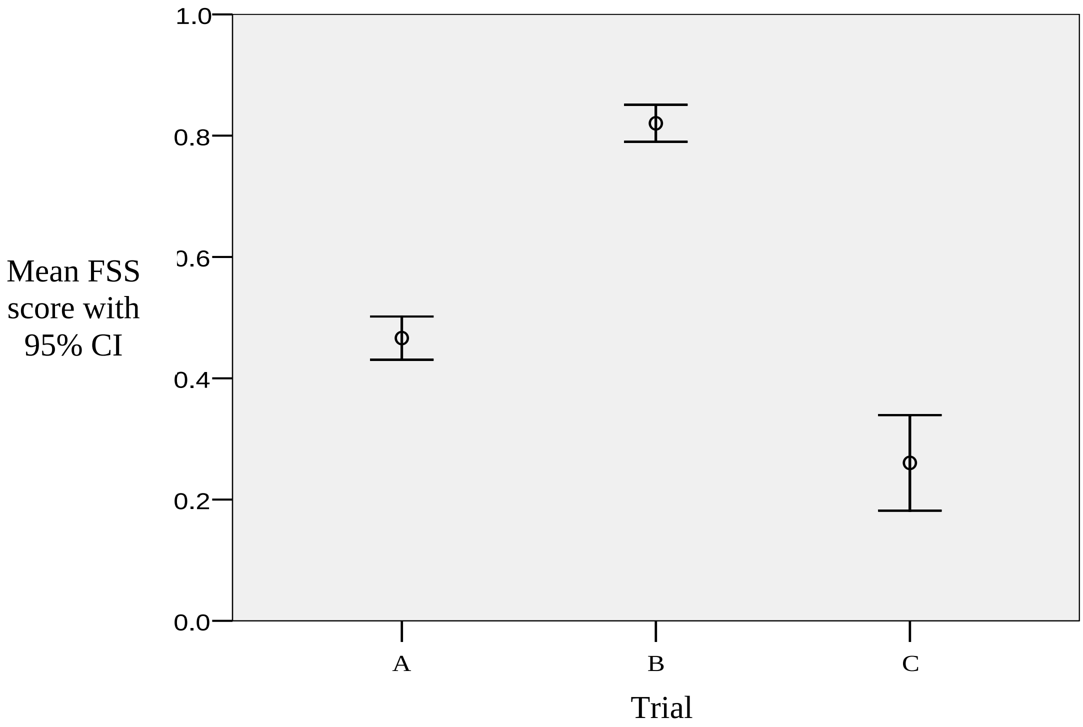


Figure 5.9: Plot of mean FSS score with 95% confidence interval (CI) for three different combinations (A–No visuals, B– music with synchronised gestures, C– music with asynchronised gestures).

Tukey's HSD test was used to compare the means. The outcome of this test was as follows:

- Mean FSS score of music with synchronised gestures (Trial B) was significantly higher $(p < 0.01)$ than music alone (Trial A) and had the best outcome.
- Mean FSS score of music with synchronised gestures (Trial B) was significantly higher $(p < 0.01)$ than music with asynchronised gestures (Trial C).
- Mean FSS score of music alone (Trial A) was significantly higher $(p < 0.01)$ than music with asynchronised gestures (Trial C).

Observations

Many participants reported that *the visuals are wrong*, when they listened to music with asynchronised gestures. Only one participant could not tell the difference between synchronised and asynchronised gestures for the rock song (the "ba" "ba" movements). She could still differentiate between synchronised and asynchronised gestures for the classical music (the orchestral conductor's gestures).

Following are some comments received after watching the asynchronised gestures:

This is wrong.

I can't understand this.

I'd rather listen to music alone.

Doesn't make sense.

All the participants preferred to watch human body movements synchronised with music. When asked the reason for this, some of the participants said they could "hear" better; however, they were unable to clarify this further.

From the statistical analysis given in the previous section and from the observations above, it appeared that most participants preferred watching human gestures synchronised with music when listening to music. When the music and gestures were asynchronised, the participants preferred just listening to music without any visual display.

5.3.3 User Study: Regular monitoring of response to Haptic Chair

Although the feedback about the Haptic Chair was uniformly positive, it is possible that what we were measuring was due to novelty rather than anything specific about listening to music haptically. Therefore, the objective of this experiment was to further explore the validity of the 100% positive feedback received for the initial prototype of the Haptic Chair. If the positive feedback was not due to the initial excitement of a novel technology, then the user response should continue to be positive even after using the Haptic Chair for a longer period of time. To study this effect, the user satisfaction of the Haptic Chair was regularly monitored over a period of 3 weeks.

The ISO 9241-11 defines satisfaction as "freedom from discomfort and positive attitudes to the use of the product" [122]. Satisfaction can be specified and measured by subjective ratings on scales such as discomfort experienced, liking for the product and many other methods of evaluating user satisfaction [123-126]. In this work, satisfaction was measured using a questionnaire derived from the "Usefulness, Satisfaction, and Ease of use" (USE) questionnaire [126]. The modified USE questionnaire consisted of five statements where the participants were asked to rate a statement (of modified USE) on a 5 point scale, ranging from 1 (strongly disagree) to 5 (strongly agree). Overall satisfaction was calculated as a weighted average of the ratings given for the questions, and ranged from 0 to 1 where a score of 1 corresponded to optimal satisfaction.

Participants

Six participants (3 male, 3 female) took part in this study. They were randomly chosen from the 36 participants who took part in the user study 1 described in Section 5.3.1. However, the participants were unaware of the information that might have biased or skewed the results.

Procedure

The idea of this experiment was to continuously monitor the user's satisfaction with the Haptic Chair. Each participant was given 10 minutes to listen to music while sitting on the Haptic Chair. They were allowed to choose songs from a large collection of MP3 songs the including British rock songs, Sri Lankan Sinhalese songs and Indian Hindi songs. This procedure was repeated everyday over a period of 22 days. Each day, after the sessions, participants asked to comment on their experience. On days 1, 8, 15 and 22 (Monday of each week over 4 weeks), after 10 minutes of informal listening, each of the participants were given the chance to listen to 2 test music samples— Mendelssohn's Symphony No. 4 and "It's my life" by Bon Jovi (the same samples used in the previous experiment). After listening to 2 test music samples, they were asked to answer a few questions derived from the USE questionnaire [126]. User satisfaction was calculated from the responses. In addition, their preferences for the test music samples were recorded.

Results

It appeared that all six participants very much enjoyed the experience of the Haptic Chair. In fact, after two weeks of regular use, all of them requested to increase the time (10 minutes) they were provided within a session. Therefore, the duration for each participant was increased and each participant provided the opportunity to "listen" to music for 15 minutes per day during the last week of the study. Figure 5.10 shows the overall satisfaction of the users measured on days 1, 8, 15 and 22 (Monday of every week over 4 weeks) of the experiment. A Higher value for the USE score corresponds to higher satisfaction. As seen from Figure 5.10, the participants were very satisfied with the Haptic Chair. Moreover, the satisfaction level was sustained over the entire duration of the experiment.

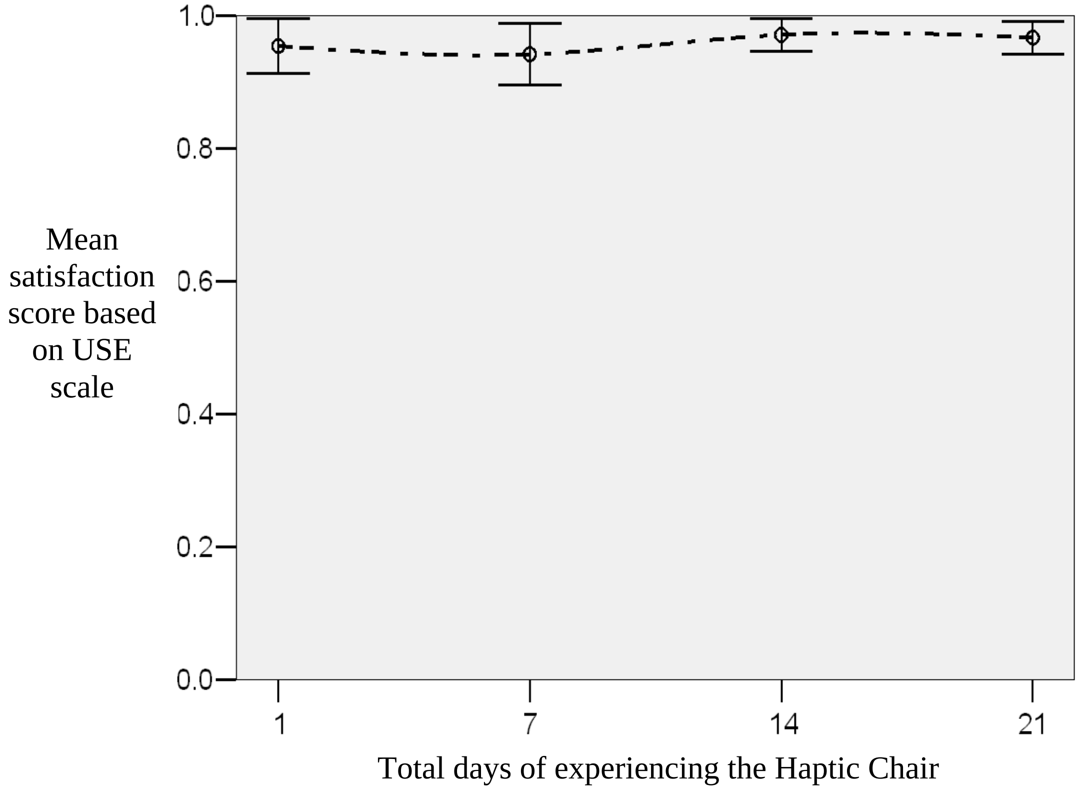


Figure 5.10: Mean USE score of satisfaction monitored from six percipients over a period of three weeks. (Bars at each data point shows the 95% confidence intervals)

Since the USE score was calculated as a average of five rating scale values, the composite variable was treated as a continuous variable. One way ANOVA analysis confirmed that there was no significant difference in the observed level of satisfaction $(F(3,44)=0.64, p=0.59)$. In other words, a participant's satisfaction with the Haptic Chair remained unchanged even after using it 10 minutes every day for a period of more than 3 weeks. The initial response was very positive and there was little room to improve.

Observations made

The participants' reactions to the Haptic Chair were continuously monitored as a way of controlling for a possible novelty effect in our previous data. The level of enthusiasm was maintained throughout the extended experiment. There were times when some participants were unhappy when they were told that his/her session was over. After two weeks, the 6 participants were told that they did not have to come every day to take part in the experiment (to "listen" to music for 10 minutes) if they were not willing to. However, all the participants reported that they looked forward to the listening session. In fact, as mentioned in the previous section, all participants wanted to listen to music using the Haptic Chair for a longer duration. None seemed to get bored with the Haptic Chair. Some of the important comments received were:

I am really happy.

This is very good.

I feel like taking this home.

Can I sit for 5 more mins?

10 mins is not enough.

I couldn't hear the lyrics.

So much better than listening to radio at home.

Since all the participants were making positive comments all the time and not criticising the Haptic Chair; they were specifically asked to make at least one negative comment. This was done on the 18th day of the experiment. However, none of the participants made any negative comments other than reporting that they could not hear the lyrics.

On the 16th day of the experiment, one of the participants (a profoundly deaf student) was listening to music, a recording of a speech was played through the Haptic Chair and he was asked whether he could hear the "Song". He reported that it was not a song!

Another important observation was made on the 15th day of the experiment. Usually, when the six participants came to use the Haptic Chair one student sat on the chair and the rest sat by the laptop that was used to play the music. The music was played through the Windows Media player and apparently the Media Player visualisations were switched ON and visible on the computer screen. It was noticed that the students who were looking at the display were commenting about it to the sign language interpreter. According to the sign language interpreter, some comments of the students were:

I feel sleepy.

Looking at these patterns makes me dizzy.

I am tired of looking at these.

Most of the participants were asking whether it is possible to play facial animations (that they had seen before during other experiments) with the songs.

Overall it appeared that everyone who used the Haptic Chair liked the experience very much. This positive response was not due to the fact that it was a completely new experience for them. If it was due to initial excitement, the response would have gone down as they used the Haptic Chair for more than 3 weeks. The response received at the end of the last day was as good as or even better than the response received on the first day. On the last day of the experiment, when the participants were told that the experiment is over, one of them said *I am going be deaf again* thinking that she would not get the chance to experience the Haptic Chair again.

The combination of human gestures synchronised with music was preferred by the participants over abstract patterns that changed corresponding to music. This could have been due to the presence of a human character. Silent dance can often be very entertaining. However, when the human gestures and music were not synchronised, almost all the participants spotted that and expressed their dislike. This shows that there is little to be gained by showing human gestures with music unless the gesturing patterns and music are tightly synchronised. The approach of using human gestures to convey a musical experience proved to be much more effective than abstract animations. With this modification the overall system became more effective. Deaf people generally take many cues from watching other people move and react to sounds and music in the environment. This could be one explanation for strong preference observed for human gestures over abstract graphics. Brain imaging techniques may provide a stronger explanation for the preference of watching human gestures, though the approach was not within the scope of this dissertation.

5.4 Supporting Experiments

5.4.1 Perception of vibrotactile frequencies above 1 kHz

The current study delivered the entire frequency range of the music through the Haptic Chair as potential tactile stimulation, even though most studies report that tactile system is only responsive up to approximately 1000 Hz [25]. During the course of our studies with the Haptic Chair, however, we kept seeing evidence which suggested that people could detect higher frequencies.

Most of the research on this topic was conducted using simple sine tones as test stimuli, and it is possible that responses to more complex and dynamic signals characteristic of natural environmental stimuli are not predictable from responses to sine tones alone. In the auditory system, for example, Evans found that 20% of cortical neurons in unanesthetized cats respond only to complex signals such as clicks and noise bursts [127]. Whitfield and Evans found cortical neurons they called ‘frequency sweep detectors’ for their preferential responses to frequency changes in particular directions [128]. The visual system is also well known to contain many types of complex ‘feature detector’ neurons [129]. Complex signals are qualitatively more than the sum of their parts. For example, harmonic components with properly constructed amplitude and phase relationships can create signals with instantaneous pressure variations (approaching square waves) with steeper slopes than those in any of the constituent sine wave components alone, and these fast pressure variations could conceivably play a role in signal detection. In the haptic domain, Verillo has also noted that intensity discrimination is better for pulsed and amplitude modulated tones than for pure tones [40]. With the dearth of relevant literature on the topic, we believe that the role played by higher frequencies in tactile perception is still an open question.

The delivery mechanisms in the majority of laboratory experiments have also been limited. Frequency response curves for sine tones reported in the literature have typically been measured using point-like stimulators on the skin [40, 130]. In addition, the interaction with vibrotactile stimuli in everyday life are very different from that used in these controlled experiments. Therefore, to determine thresholds of detection for a variety of stimuli, we used a mechanism that produced a source of vibration and a variety of signal types more typical of those found in everyday life, and particularly in vibrotactile systems designed to support music and speech communication

We conducted a single-blind controlled experiment with both hearing and hearing-impaired subjects to study whether they could detect vibrations above 1000Hz. In addition, we investigated the relationship between vibrotactile sensitivity thresholds and the complexity of signals. Each subject's hand was free to rest on a flat vibrating surface such that signals could be detected via the skin of the palm and fingers of the hand, thereby offering the subjects the opportunity to detect vibration in a natural and biologically meaningful way.

The reason that both people with normal hearing and hearing-impairments were studied was two-fold. We were interested to find out whether people with hearing impairments were more sensitive to environmental vibrations than people with normal hearing, since the former might be expected to rely more on vibrotactile sensory perception during their everyday activities. Second, to support the design of our Haptic Chair—we were interested to learn more about the frequency range of vibrotactile perception to ensure that a significant range of frequencies produced by music was indeed sensed through haptic input.

User study

Participants

Twelve participants with normal hearing (6 male subjects and 6 female subjects; median age 22.5 years ranging from 20 to 35 years) and twelve hearing-impaired participants (5 male subjects and 7 female subjects; median age 24 years ranging from 16 to 31 years) took part in the study. All the participants were recruited from Singapore and none of them had taken part in our previous studies. All participants had normal vision. A person experienced in using and interpreting sign language for the Deaf was present to help explain, when necessary, the purpose of the study, the procedure, and to answer any questions subjects might have. All subjects were able to stop participating in the study at any time if they did not want to continue. The study was conducted in accordance with the ethical research guidelines provided by the Internal Review Board (IRB) of the National University of Singapore (NUS) and with IRB approval.

Apparatus

The same mechanism that was used to generate vibrations in the Haptic Chair (contact speakers mounted on a densely laminated wooden frame) was used to present the vibrotactile stimuli. In the current experiments, a vibrating wooden surface consisting of a SolidDrive™ speaker directly mounted on the under-surface of a densely laminated rectangular wooden board (33 x 23.5 cm) as shown in Figure 5.11 was used to present the vibrotactile stimuli. The surface of the board was fairly smooth with light, regularly distributed surface texture, made of a similar kind of wood and with similar composition to the densely laminated wooden frame of the ‘Haptic Chair’. Subjects were seated throughout the experiment and the vibrating surface was positioned so that subjects could rest their hands comfortably without the need to tense the muscles of the arm. Before each experiment began, subjects were given the opportunity to ensure they were seated comfortably and the vibrating wooden test surface was at an appropriate height and position relative to the subject’s body. This enabled the findings of this research study to be directly applied to our work with the ‘Haptic Chair’.

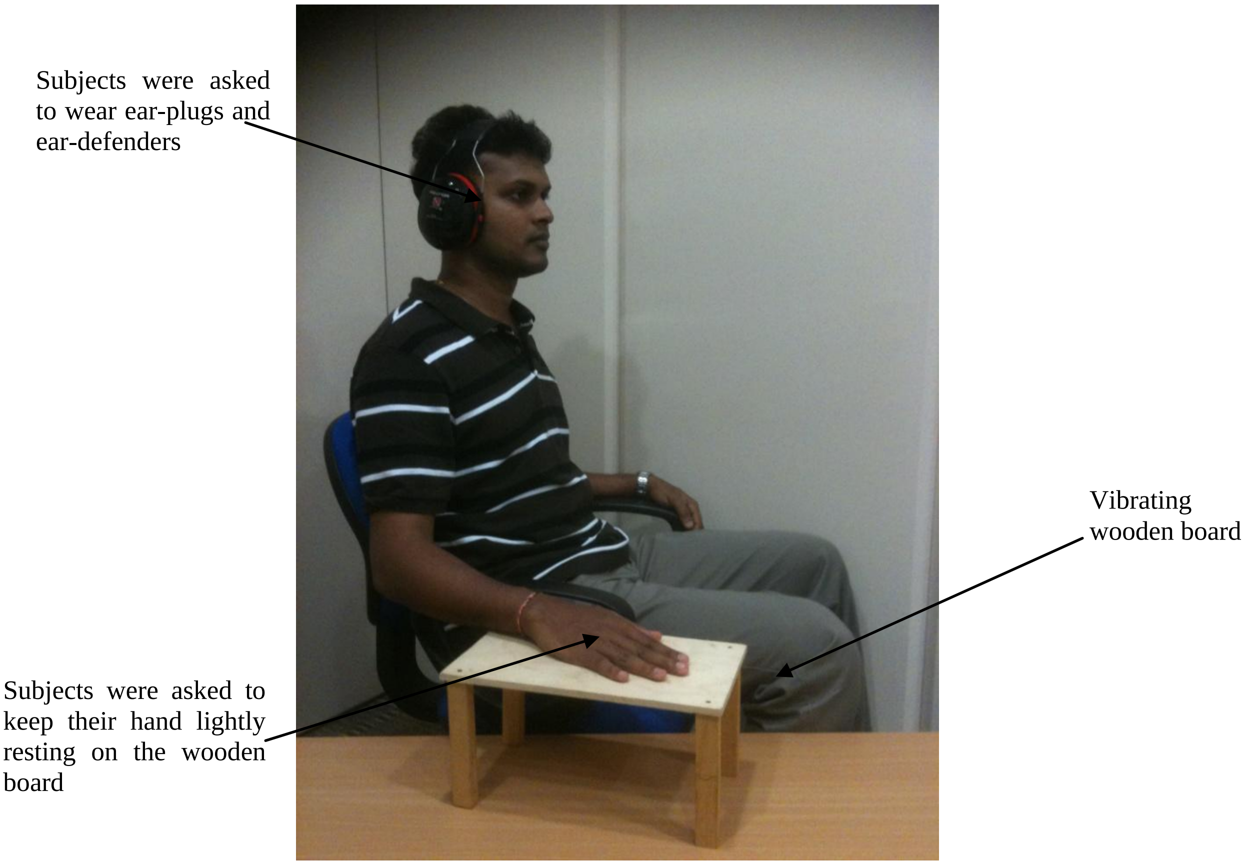


Figure 5.11: Experimental set-up showing the relaxed placement of the hand and position of the arm assumed by the subjects

A notebook computer running LabVIEW™ 8.2 was used to control the intensity of the input signals, which were created offline using MATLAB™. To measure the strength of vibrations at the contact surface, an accelerometer (3041A4, Dytran Instruments, Inc., U.S.A.) was mounted underneath the vibrating wooden board. The output of the accelerometer was connected to a signal conditioner and the output of this device was collected by a data acquisition module (USB-6251, National Instruments). The data was then processed and collected by the same notebook computer running customised software written in LabVIEW™ 8.2

We carried out the following procedure to validate the recorded data. Firstly, the power spectral density (PSD) of each recorded data was examined to verify that the signal was not distorted. Secondly, for each subject, the PSD of every detected tone was compared against all the detected lower frequency tones by the same subject. Any high frequency tone found to have detectable low frequency distortion components (when there was less than a 10dB difference between the detectable level of a lower frequency tone and the energy in the distortion component at that frequency) were discarded from the further analysis. This ensured that detection of any tone above 1 kHz was not due to the presence of detectable low frequency components.

Stimuli

Five different tones were used: sine waves, square waves, amplitude-modulated (AM) tones, frequency-modulated (FM) tones, and frequency up-sweeps. For each tone type, 5 stimuli were created for 5 different frequencies (250, 500, 1000, 2000, 4000 Hz). The specifications of the tones are given in Table 5.3.

Table 5.3: Specifications of the tones used in the haptic sensitivity experiment

Tone Type	Specifications
Sine wave	Pure tones at frequencies 250, 500, 1000, 2000 and 4000 Hz
Square wave	Square waves at frequencies 250, 500, 1000, 2000 and 4000 Hz
Frequency sweep	Frequency sweeps at starting frequencies 250, 500, 1000, 2000 and 4000 Hz Upward frequency sweeps based on $f*2^{2t}$; for t in [0, 1] seconds where f is the starting frequency (f = 250, 500, 1000, 2000 and 4000 Hz)
FM tone	FM tones at carrier frequencies 250, 500, 1000, 2000 and 4000 Hz Frequency varies between ±10% of the carrier frequency at a rate of 2 Hz
AM tone	AM tones at carrier frequencies 250, 500, 1000, 2000 and 4000 Hz Modulated with a 2 Hz tone Modulation depth = 100%

- Duration of each of the 25 tones (5 tone types and 5 frequencies from each type) is 1 second
- A 10 ms ramp up at the beginning and 10ms ramp down at the end was added to avoid clicks and distortion at the endpoints
- All the tones were normalised to have equal amount of power.

Procedure

On arrival, each subject was given at least 10-15 minutes to rest and become accustomed to their new surroundings before they took part in the experiment. In other words, this settling time ensured that the heart rates of the subjects were not significantly different from their normal values and their skin temperature was adjusted to the ambient environmental conditions. The experiment was conducted in a research room that was carefully designed to be sound-proof and also anechoic; room temperature was maintained at 24°C. During the initial rest period, subjects read the information sheet prepared for participants and were given the opportunity to sign the consent form or to not take part, as they wished. They were also allowed to ask questions about the study. The help of a proficient sign-language interpreter was available to facilitate communication between the research team and hearing-impaired participants.

Subjects (both hearing and hearing-impaired) were asked to wear soft foam ear-plugs (3M™ Foam Ear Plug 1100) and were instructed on how to use them correctly since this was necessary for a good fit and effective performance; and ear-defenders (H540A-411-SV by Peltor), to reduce sound conduction via the external ear canal and minimise the possibility of confounding the results by detecting any audible sound generated by the test stimuli. This combination attenuated air-borne sound by over 40 dB (tested by members of the research team who were not subjects in the experiment). Each subject was asked to place their dominant hand resting lightly on the vibrating wooden board and indicate when they feel the surface vibrating. They were clearly instructed to concentrate on the vibrations. A few test trials were conducted before the actual experiment to make sure the subjects understood the instructions and to familiarize them with the experimental procedure. A standard up-down staircase method [131] was used to determine the threshold of detection. For a given stimulus, intensity level was decreased by a step of 1dB after a positive response or increased by a step of 1dB after a negative response. This procedure was carried out until six reversals were obtained. Two members of the research team handled stimulus intensity control and data recording independently. Figure 5.12 shows a typical data set recorded for a given stimulus. Following Levitt [131], we used the midpoints of run 2, 4 and 6 to calculate the threshold. Participants were given short breaks between trials to avoid adaptation to the various stimuli.

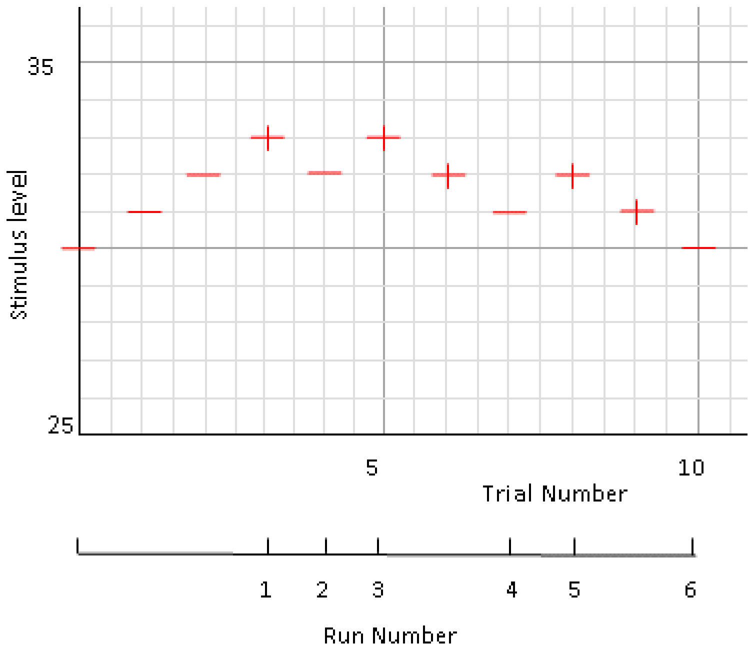


Figure 5.12: Typical record of data during the experiment. The initial value was determined based on our preliminary studies. Threshold was calculated as average of runs 2, 4 and 6.

We considered that having to concentrate during the presentation of 25 stimuli of 5 tone types and 5 frequencies from each type might reduce the discriminatory ability and interest of the participants. Therefore, to avoid this, the number of tested stimuli per subject was reduced to 15 per subject: 5 tone types and 3 base frequencies from each type. What was most important for the purpose of this study was to test the different signal types with frequencies above 1000Hz on a given participant. We were more interested in studying the relationship between signal types and whether subjects could detect frequencies above 1000Hz than in the threshold curves across the frequency spectrum. In addition, the experimental time was reduced by representing only three of the base frequencies to each participant. To achieve this, the participants were divided into 4 groups and each group was tested with different frequencies: Group (1) 250, 500, 2000 Hz; Group (2) 250, 500, 4000 Hz; Group (3) 250, 1000, 2000 Hz; Group (4) 250, 1000, 4000 Hz.

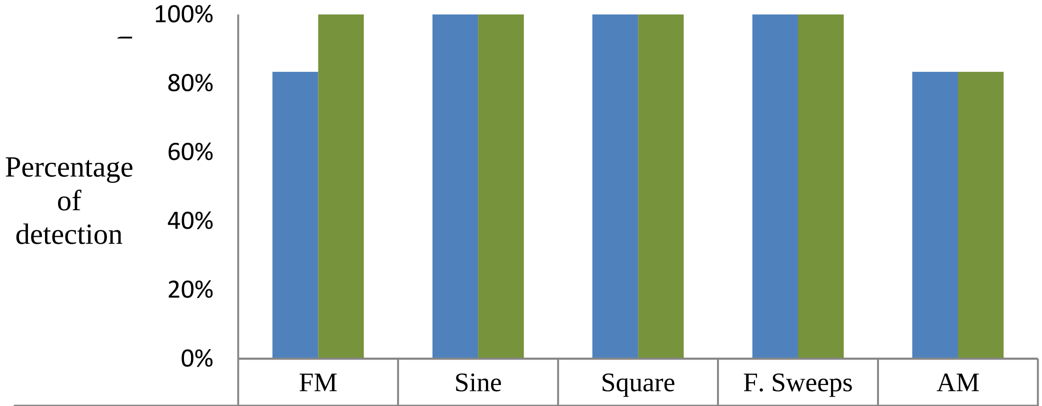
Results

Detection of high frequency tones

When the responses to the 15 test stimuli presented had been collected from each subject, the power spectral densities of the signals recorded from the vibrating board were examined to ensure there was no detectable low frequency noise or sub-harmonics. Any stimulus found to be distorted with low frequency noise was discarded from the analysis. Based on this, 17 data points out of 360 data points were removed.

We defined a given stimulus as 'undetected' if a subject could not detect it before the maximum intensity level was reached, or if the stimulus was found to be corrupted with low frequency noise or sub-harmonics when the subjects reported detection. Otherwise, the stimulus was considered to be "detected" and the threshold of detection was calculated using the standard up-down procedure [131].

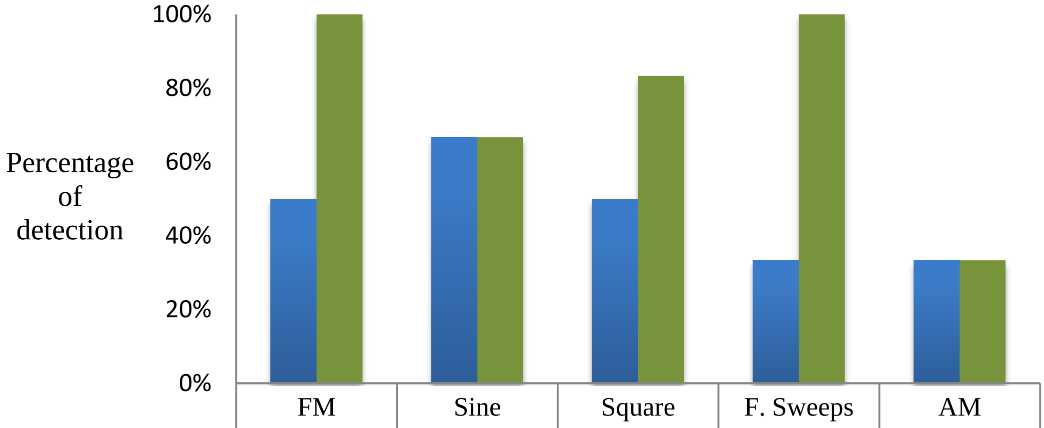
It was observed that all 5 tone types at frequencies 250, 500 and 1000 Hz were detected by all the hearing and hearing-impaired subjects. More importantly, we observed that our subjects were able to detect tones at 2000Hz and 4000Hz. As described previously, the possibility that detection could have been due to the presence of low frequency noise or sub-harmonic distortion of the signal was ruled out. Figure 5.13 shows the percentage of detection for each tone type at 2000 Hz. The results show that all of the hearing and hearing-impaired subjects were able to detect sine, square tones, and frequency sweeps at 2000 Hz. Eighty three percent of both hearing and hearing-impaired subjects detected AM tone at 2000 Hz. The hearing-impaired subjects appeared to be somewhat better at detecting FM tones at 2000Hz.



	FM	Sine	Square	F. Sweeps	AM
Hearing	83%	100%	100%	100%	83%
Hearing-impaired	100%	100%	100%	100%	83%

Figure 5.13: Comparison of the ability to detect tones at 2000 Hz by the total of 12 hearing and hearing-impaired subjects.

Figure 5.14 shows the percentage of subjects who detected each tone type at 4000 Hz. All the hearing-impaired subjects were able to detect FM and frequency sweeps at 4000 Hz. The majority of hearing-impaired subjects were also able to detect sine and square tones at 4000 Hz. The results suggested that hearing-impaired subjects were better able to detect FM, square waves and frequency sweeps at 4000 Hz than the hearing subjects.



	FM	Sine	Square	F. Sweeps	AM
Hearing	50%	67%	50%	33%	33%
Hearing-impaired	100%	67%	83%	100%	33%

Figure 5.14: Comparison of the detectability of tones at 4000 Hz by hearing and hearing-impaired subjects

Thresholds of detection at different frequencies

Figure 5.15 and Figure 5.16 show the thresholds of tactile sensitivity for the different tone types at different frequencies. As expected, the subjects were more sensitive to tones at 250 Hz. More importantly, tones with complex waveforms (AM, FM, square and frequency sweeps) generally have lower thresholds as compared to a static sine tone at same frequency. This result holds true for both hearing and hearing-impaired subjects.

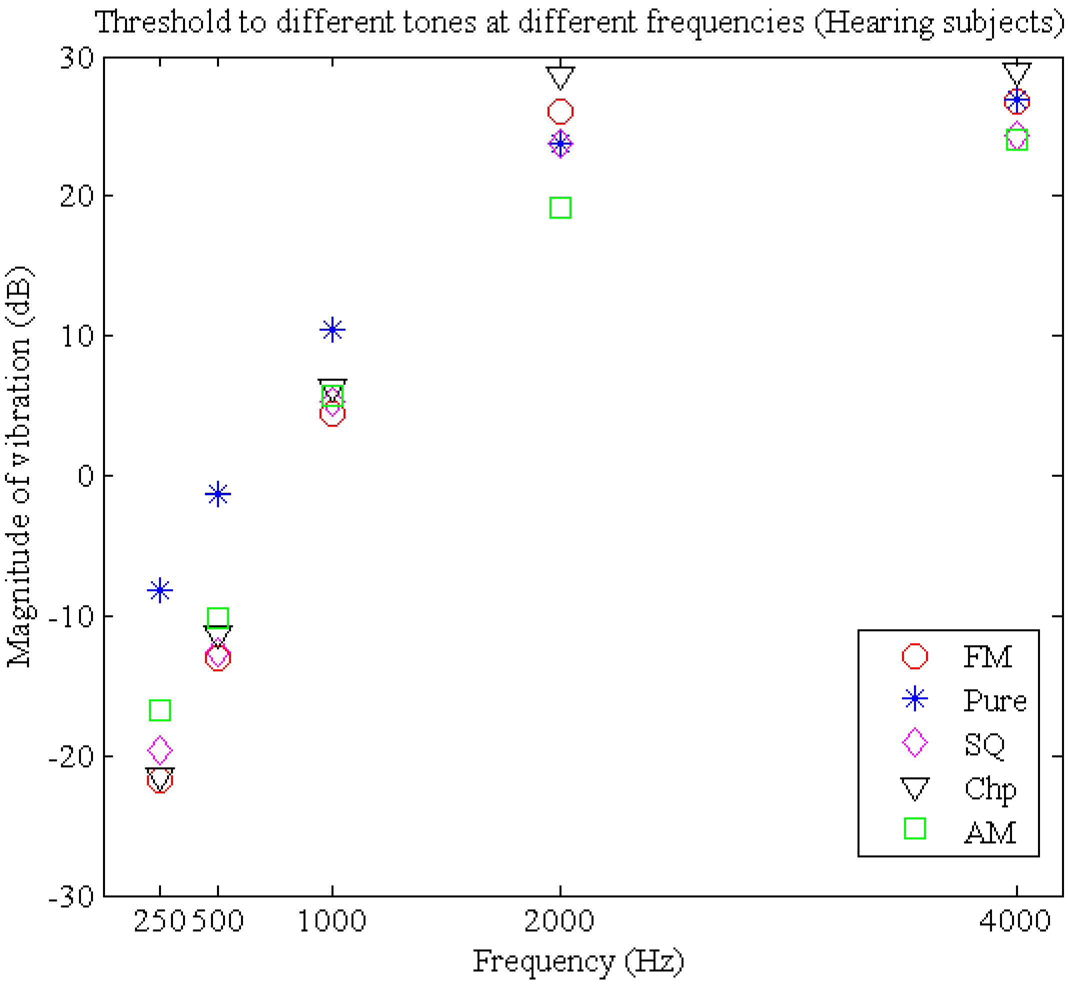


Figure 5.15: Threshold of vibration sensitivity to different tones at different frequencies (threshold values are given in dB with respect to 1g vibrations. The symbol 'g' is often used as the unit of acceleration; it is approximately 9.8 ms^{-2})

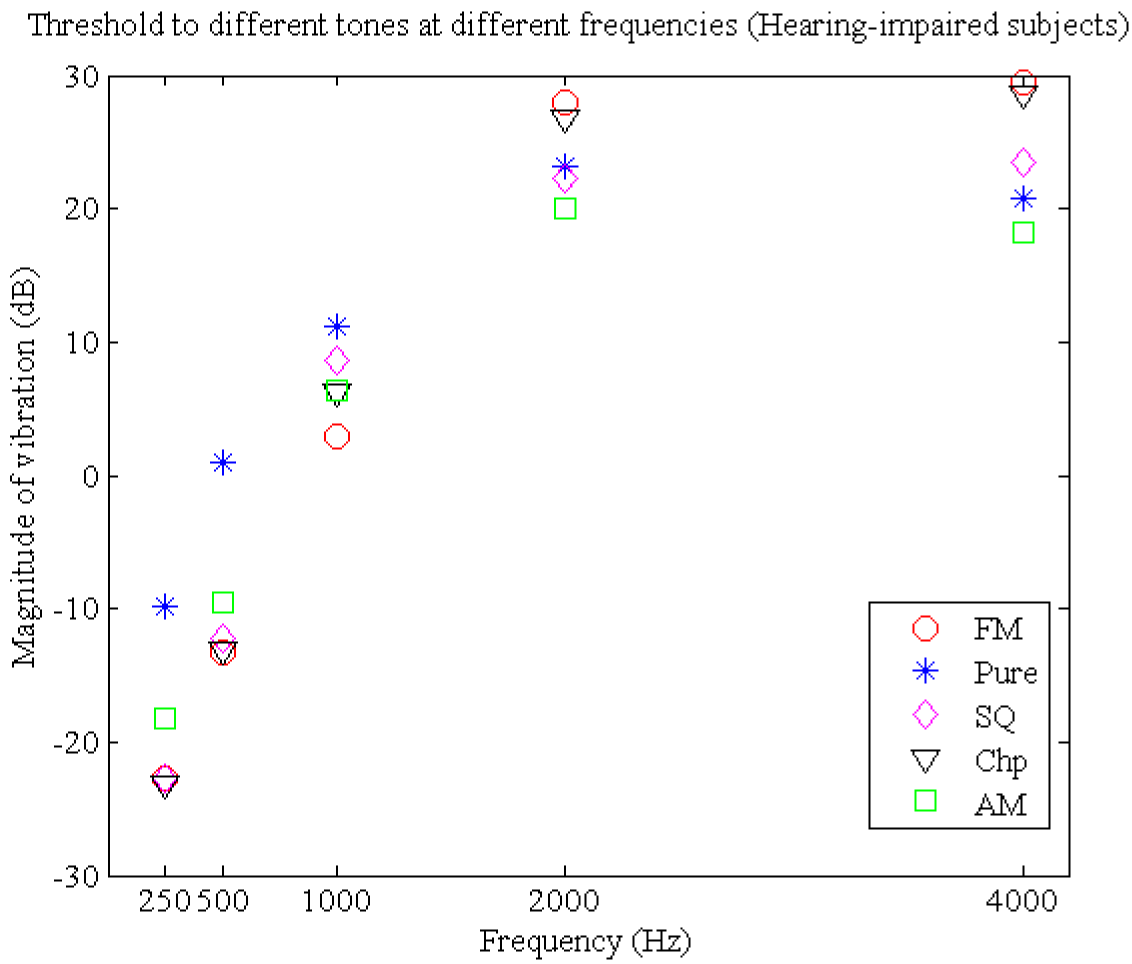


Figure 5.16: Threshold of vibration sensitivity for different tones at different frequencies (Threshold values are given in dB with respect to 1g vibration)

Effect of different tones

A one way repeated measures ANOVA analysis was carried out to compare the sensitivity to different signal types at a given frequency. For hearing subjects, ANOVA reveals a significant difference between the thresholds to different signal types at 250 Hz $(F(4,55) = 5.90, p = 0.001)$, 500 Hz $(F(4,25) = 4.26, p = 0.009)$ and 2000 Hz $(F(4,23) = 4.82, p = 0.006)$. Post-hoc analysis with Tukey's honestly significant difference (HSD) test showed that the sine tone at 250 Hz and 500 Hz has a significantly higher threshold compared to all other tones at the same frequencies

$(p < 0.05)$. At 2000 Hz, frequency sweep has significantly different threshold compared to sine, square and AM tones $(p < 0.05)$.

For hearing-impaired subjects, ANOVA reveals a significant difference between the thresholds to different signal types at all the frequency points we tested: 250 Hz $(F(4,55) = 21.73, p < 0.00001)$, 500 Hz $(F(4,25) = 20.31, p < 0.00001)$, 1000 Hz $(F(4,25) = 6.37, p = 0.001)$, 2000Hz $(F(4,24) = 5.37, p = 0.003)$ and 4000 Hz $(F(4,18) = 7.32, p = 0.001)$. Tukey's HSD test showed that sine tone has a significantly higher threshold compared to all other tones at frequencies 250 Hz and 500 Hz $(p < 0.01)$. At 2000 Hz and 4000 Hz, both FM tones and frequency sweeps showed significantly higher thresholds $(p < 0.01)$.

Discussion

The results of this experiment suggest that both hearing and hearing-impaired participants are able to detect vibrations at base frequencies of 2000Hz and 4000 Hz, though the signal amplitude was 30-40 dB higher than a detectable tone at 250 Hz. This frequency is two octaves higher than the limiting frequency of 1000Hz for tactile sensitivity reported in other studies [40]. Noise and distortion were considered to be potential confounding factors given the relatively large dimensions of the vibrating structure we used. However, analysis of the power spectral densities of the recorded signals measured of the vibrating wooden board (shown in Figure 5.17) verified that they did not have a significant effect on the observations.

The mechanoreceptors in human skin are believed to integrate energy spatially [132, 133] . Therefore, it is reasonable to assume that the relatively large contact area of the whole hand (palm and fingers) used in our experiments would have facilitated spatial integration of vibrotactile stimuli leading to lower detection thresholds for higher frequencies. The contact area of the entire ventral surface of the hand is larger than the contact area used in research work previously reported (0.01-10 cm2) [133]. Therefore absolute detection thresholds obtained in the study described here cannot be directly compared with those found in previous literature but we believe they are more applicable to the sensation of vibration in everyday life. A point-source stimulus applied to a very small area of glabrous skin might by-pass or increase de-sensitisation of important channels of vibrotactile stimuli.

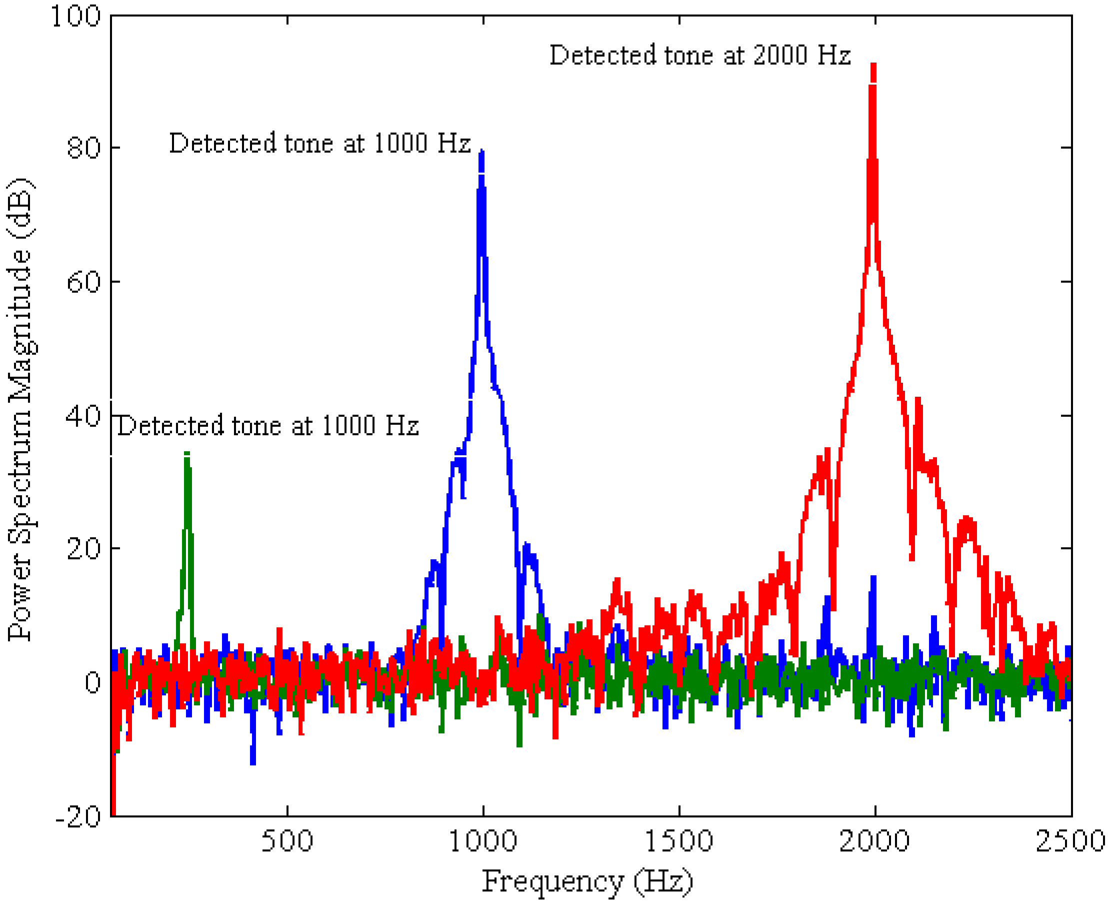


Figure 5.17: The dotted line illustrates the Power Spectral Density (PSD) of 'detected' square tone at 2000 Hz by one representative subject. PSDs of the most sensitive tones (with lowest thresholds) at 1000 Hz and 250 Hz for the same subject were shown by solid-gray and dashed-gray lines respectively. It is clear from this figure that the detection of the tone at 2000Hz could not have been due to the presence of any low frequency noise or sub-harmonics, since the PSD of the detected signal was more than 30dB below the threshold of detectability for other tones at any given frequency.

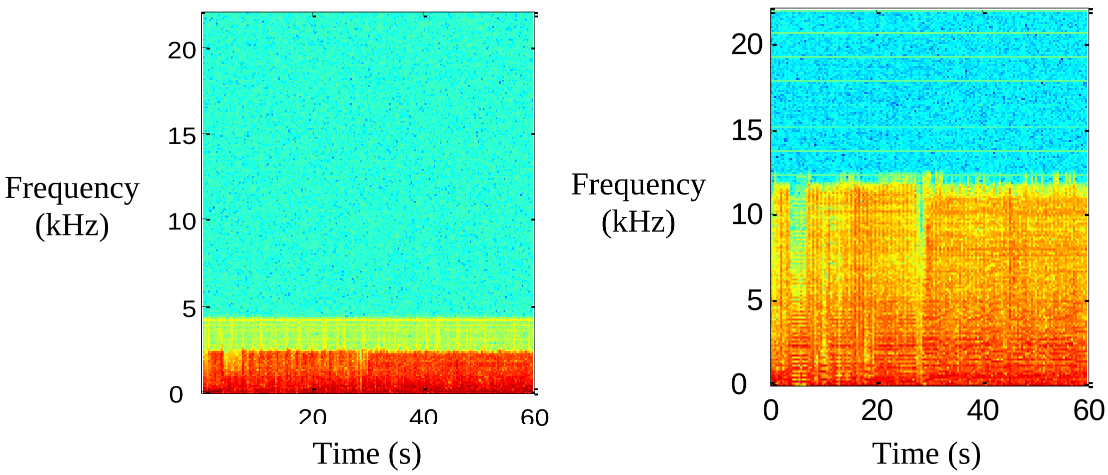
The second finding is that the complex signals we used have lower thresholds of detection than pure static sine tones. Some of the participants reported that they could 'feel something moving' when feeling the vibrations corresponding to a FM tone and thus easily detected the stimulus compared to a static sine tone. This was similar to what we observe in the context of visual perception where a flashing light is more easily detectable than a static light. We observed that increased sensitivity to complex signals is most significant at 250Hz: at this frequency both FM and square tones resulted in a decrease of approximately 5 dB in threshold over a pure sine tone. This occurred despite the normalization of energy which produced a signal with less energy at the fundamental frequency for the complex tones than for the sine tone. Since energy was normalized across signals, this suggests an explanation beyond simple integration across frequency, and points to the possibility that the temporal dynamics of the complex signals play a role in detection.

5.4.2 Unaltered Audio versus Frequency Scaled Audio

The Haptic Chair described in this dissertation, deliberately makes no attempt to pre-process the music but delivers the entire audio stream to each of the contact speakers positioned targeting the feet, back, arms and hands. Most of the related works mentioned in Chapter 2.4 have pre-processed the audio signal before producing a tactile feedback. Generally the tactile output for these devices was less than 1 kHz, which has been accepted as the limiting frequency for vibrotactile sensation [39, 40, 133].

User Study

In addition to the strategic motivation of not manipulating the source signal for tactile music perception, we believe that the role played by higher frequencies in tactile perception is still an open question. In fact, Karam et al. [58] have suggested that altering frequency of a music stream onto the vibrotactile spectrum might potentially reduce the accuracy of the representation of the information expressed in music. We therefore, conducted a preliminary study to compare the response to an unaltered and frequency-scaled-down music played through the Haptic Chair. In the case of frequency scaled music, the frequency range was scaled-down by a factor of 5 using "Acoustic Bandwidth Compression" (ABC) algorithm [134]. The ABC algorithm divides the original time series into segments, compresses their bandwidth and concatenates the segments to resemble a new but representative time series with lower frequency components.



(a) **(b)**

Figure 5.18: Example of the unaltered and altered audio streams: (a) Unaltered version of the song "It's my life"; (b) Altered (frequency scaled using ABC algorithm) version of the same song

The spectrograms of the unaltered and frequency scaled version are shown in Figure 5.18. According to Figure 5.18(b), the unaltered audio stream has frequency content at least up to 10 kHz; in contrast the frequency scaled version, shown in Figure 5.18(a), contained frequency content only up to 2 kHz. Therefore, unaltered audio stream produced vibrotactile feedback with much higher frequency content than the vibrotactile feedback generated from frequency scaled audio stream. It should be noted that ABC algorithm scaled the absolute frequencies by a factor of 5 but kept the same relative harmonic relationships.

Participants, apparatus and procedure

Twelve participants (7 male and 5 female) took part in this study. They were from the age group of 12-18 years. All of them had taken part in our previous studies. However, they were new to this experiment and were unaware of the information that might have biased or skewed the results. The study was conducted in accordance with the ethical research guidelines provided by the IRB of NUS and with IRB approval.

One minute recordings of a classical music piece (Mendelssohn's Symphony No. 4) and a rock song ("It's my life"— by Bon Jovi) were used as test cases. Each of the participants was presented with four trials—unaltered versions and frequency scaled versions of the two music pieces. Trials were presented in a randomised order. The task of the participants was to "listen" to music while sitting on the Haptic Chair. Visual effects were not used for this experiment. After each trial, the participants were asked to answer a questionnaire which was used to calculate the FSS score.

Results

Figure 5.19 shows the mean FSS score for all experimental conditions. It can be seen from the Figure 5.19 that the mean FSS score for unaltered music is much higher than the mean FSS score for frequency scaled music. This observation is similar for both rock and classical music samples. A 2-way repeated measures ANOVA suggested no main effect for music genres $(F(1,44)=1.01, p=0.32)$; a significant main effect for frequency scaling of audio $(F(1,44)=98.15, p=0.000001)$; and no interaction between music genres and frequency scaling of audio $(F(1,44)=1.23, p=0.27)$.

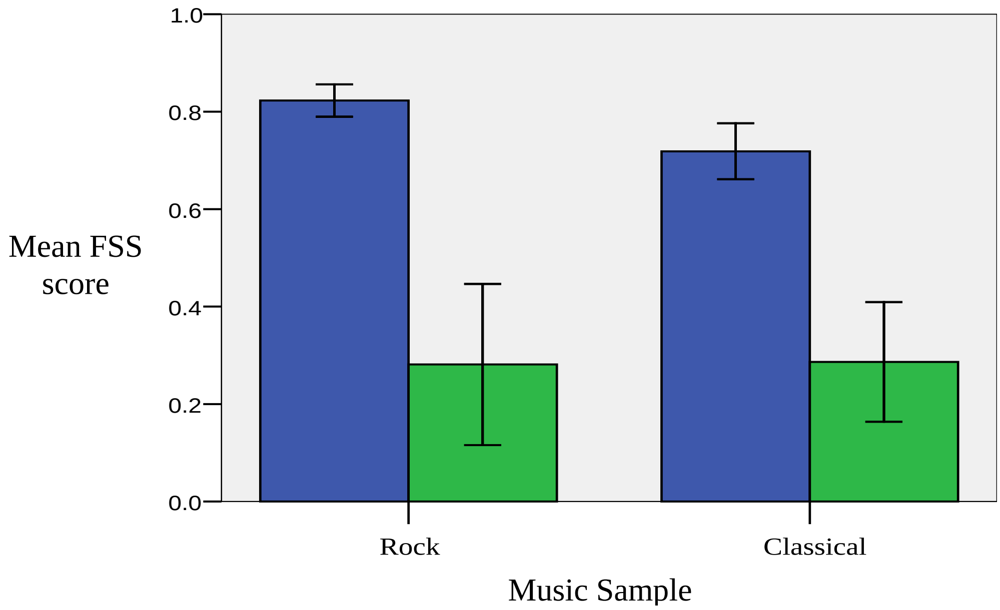


Figure 5.19: Mean FSS score for all trials with error bars showing the 95% confidence interval (■ unaltered audio, ■ altered audio).

Therefore, the FSS score was averaged across music genres and compared across two experimental conditions—altered music and unaltered music. Figure 5.20 shows that the 99% confidence intervals for the mean FSS scores of frequency scaled and unaltered audio do not overlap. This verifies that the difference between theses mean FSS scores is significant $(p < 0.01)$. In other words, mean FSS score for unaltered music stream was significantly higher than the mean FSS score for frequency-scaled-down version of the same music.

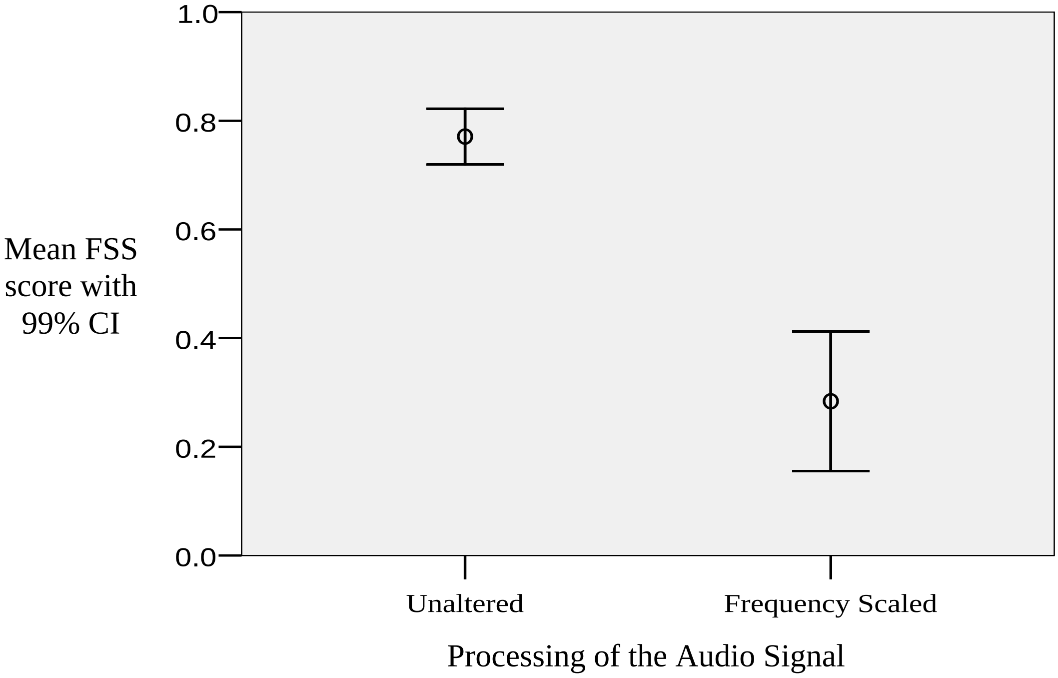


Figure 5.20: Mean FSS scores for unaltered and frequency scaled music sample with 99% confidence interval (CI).

The observations showed that the participants did not like the frequency scaled version of the music samples. Some of the participants made strong comments to show their dislike for the altered (frequency scaled) versions of the audio stream:

I can't recognize a song.

This is ugly.

This is wrong.

I hate it.

This is useless.

Although frequency scaling effectively generates low frequency vibrations (which might be more easily felt than higher frequency vibrations [40]), the variations in the music were diminished and the richness of musical content was lower in the frequency scaled version. This reduction in quality is easily detected by people with normal hearing. It was important to note that even the hearing-impaired could still feel this effect through the Haptic Chair. It is possible that the role of higher frequencies in more realistic audio signals for instance, in creating sharp transients, could still be important. Moreover, as mentioned in the introduction, any additional "information" delivered through the haptic channel might actually disrupt the musical experience, and this confounding effect is potentially more significant for the deaf. This is because deaf people have used to sense the vibrations that occur naturally in objects existing in an acoustic environment through their bodies.

5.5 Discussion

5.5.1 Speaker Listening vs Sensory Input via Haptic Chair

The mechanism of providing a tactile sensation through the Haptic Chair is quite similar to the common technique deaf people use called "speaker listening". In speaker listening, deaf people place their hands or foot directly on audio speakers to feel the vibrations [57, 58]. However, the Haptic Chair provides a tactile stimulation to various parts of the body simultaneously in contrast to normal speaker listening where only one part of the body is stimulated at any particular instant. This is important since as mentioned in the introduction and the background, feeling sound vibrations through different parts of the body plays an important role in perceiving music.

It is also possible that in addition to tactile sensory input, the Haptic Chair might be providing an additional avenue for enhanced sensory input through bone conduction of sound. Bone conduction of sound is likely to be very significant for people with certain hearing impairments. Bone conduction also has the advantage of transmitting a greater range of frequencies of sound compared to purely tactile stimulation [68].

In these regards, the Haptic Chair provides much more than simple speaker listening. The teachers at the deaf school, where most of the user studies were conducted, reported that, as is typical of deaf listeners, some of the deaf participants place their hands on the normal audio speakers available at the school main auditorium and listen to music. Nevertheless, from the observations made throughout this research work, it appeared that even those who had already experienced speaker listening preferred listening to music while sitting on the Haptic Chair.

5.5.2 User Feedback

Feedback received from the two deaf musicians was very valuable although they typically perform for hearing audiences and thus might not have had any special insight into deaf audiences with either limited or no musical training. However, one of the musicians also teaches deaf children and therefore offered a more balanced opinion. In fact, different categories of deaf listeners may need to be addressed differently with the kinds of assistive technologies we are exploring. Of course, musical backgrounds and tastes differ for the deaf as widely as for people with normal hearing. The deaf community can be categorised based on their music competency and involvement—deaf people who have a music background and able to enjoy music; deaf people who have no music knowledge, but appreciate music someway;

deaf people who do not have any music background or experience, but want to know more about music; and deaf people who have no musical background at all and are not able to appreciate music and/or are not bothered about music. This research was not focused on any specific category in terms of different skill levels or musical tastes. Nevertheless, the deaf musicians provided valuable insight at the early stage of the system design. Also the feedback received from the participants at the sequential user evaluation stages regarding their specific needs provided additional insight to a better understanding of the target audience.

Most of the user studies were conducted at the Dr. Reijntjes School for the Deaf, Sri Lanka with a group of deaf children aged 12-20. Therefore, the findings of this research work may have been confounded by a host of cultural differences between the Sri Lankan population who took part in several of the user studies and other cultural group, or between different age groups. However, the two deaf musicians and a few hearing-impaired members of the Singapore Deaf Association who used the chair in Singapore provided feedback similar to what we received from the Sri Lankan users.

5.6 Conclusions

In Chapter 4, the results showed that Haptic Chair could be used to provide a more satisfying musical experience to a deaf person. A problem still remained that the impact of the abstract visual patterns was low compared to the unanimous approval of the Haptic Chair. Therefore, a completely new approach was taken in which human gestures synchronised with music were used to convey a better musical experience.

This was partly motivated by the fact that the body movements contain important information about the accompanying music [116, 117].

From the results of the user studies, it was found that human gestures synchronised with music is capable of providing a better sense of music compared to the abstract animations used in Chapter 4. The Haptic Chair received high level of satisfaction and sustained for an extended period when it was regularly used on a daily basis.

Comparing the response to feeling unaltered audio and frequency scaled down audio through the Haptic Chair, it appeared that the hearing-impaired preferred to feel the vibrations generated by the music without frequency scaling down. Maybe the frequency scaling down would have reduced the accuracy of the representation of the information expressed in the music.

Even though it was not a research goal at the beginning of this research, during the course of our previous research, however, we kept seeing evidence which suggested that people could detect higher frequencies. This led to the study that compared the sensory abilities of people with normal hearing and those with hearing impairments using open-hand contact with a flat vibrating surface that represented ‘real-world situations’. We found that both hearing and hearing-impaired could feel vibrotactile frequencies at least up to 4 kHz, which was about 2 octaves higher than limiting value reported in the literature [39, 40]. We also found that complex signals are more easily detected than sine tones, especially for low fundamental frequencies. Our findings are applicable to a better understanding of sensory biology, the development of new sensory devices for the hearing-impaired, and to the improvement of human-computer interaction where haptic displays are used.

The system proposed in this chapter has a few ways of conveying music to the hearing-impaired—through feeling the vibrations via touch, via bone conduction of sound and seeing human gestures. These different modalities individually or in combination provided a very satisfactory experience to almost all the hearing-impaired people who used the system.

Chapter 6

Conclusion

This dissertation has presented the systematic development of a prototype system to enhance the musical experience of the hearing-impaired using human centred design techniques. A summary of the findings is given in the first part of this chapter. A discussion of future directions that warrant more study concludes the chapter and the dissertation.

6.1 Summary of Results

One of the main objectives of this research project was to convey a musical experience to a deaf person using sensory inputs other than the ear. The findings from the initial survey described in Chapter 3, guided the decision to use a combination of visual and tactile information to help convey a musical experience.

As described in the second part of Chapter 3, a novel system architecture was developed to produce real-time abstract visual effects corresponding to the music being played. This architecture allows the rapid prototyping of real-time music visualisations and was used as the keystone to develop music-to-visual abstract animations. For the tactile input, as mentioned in Chapter 4, a wooden chair that amplifies audio vibrations of music was developed. The main concept behind the development of the Haptic Chair was to enhance the natural vibrations produced by music. The first working prototype consisted of two main components—an informative visual display and a Haptic Chair.

This initial system was evaluated by a formal user study with 43 deaf participants from the Dr. Reijntjes School for the Deaf, Sri Lanka. This study suggested that the Haptic Chair was capable of substantially enhancing the musical experience of deaf people, both children and adults. In fact, the Haptic Chair received 100% positive feedback. On the other hand, many participants reported that our display of abstract animation alone was not very effective. However, they reported that abstract animations conveyed additional musical meaning when presented together with the Haptic Chair. Some of the deaf participants were able to sing along with the music while sitting on the Haptic Chair while watching karaoke videos.

In the revised prototype, the effect of showing human gestures corresponding to music was compared with the abstract animations. From the results of the user studies, it was found that human gestures synchronised with the music we showed was capable of providing a more satisfying musical experience compared to our abstract animations. Many participants reported that gestures were *more musical*. It was found that this preference of watching gestures was not merely due to the visual presence of a human character. When the human gestures and music were not synchronised, many hearing-impaired participants said that the *gestures doesn't make sense*. It can be concluded that showing human gestures synchronised with music might be an effective way of conveying a musical experience using visuals.

We found that both hearing and hearing-impaired people can feel vibrations at least up to 4 kHz depending on the complexity and amplitude of the audio signal. This is significantly higher than the previously reported cut-off frequency of 1 kHz, which has been accepted as canon, and used for haptic design purposes. It was also found that, frequency scaling-down of an audio stream (using ABC algorithm) which produced lower frequency vibrations may not produce results as good as unaltered audio. From the feedback collected, we concluded that shifting musical signals down in frequency might feel unnatural as surely as it sounds unnatural to the ear. Furthermore, a pre-processed audio stream might change the natural listening experience because of the difference between audible and haptically perceived sound. Hearing-impaired people can hear some sounds with residual hearing and therefore the quality of the audible sound should not be ignored.

The 100% positive feedback received for the Haptic Chair during the first round of user study called for further investigation. It was possible that this positive response could have been due to the initial excitement of using a new technology. However, a group of deaf participants used the Haptic Chair every day over a period of 3 weeks and, during this period, none of them got bored with the chair. In fact, they kept asking for longer durations to "listen" to music. Their level of satisfaction was sustained during this extended period. The above results show that the Haptic Chair might significantly change the way the deaf community experiences music.

6.2 Future Work

This section explores some of the possible directions for future research. Contents are ordered from the most specific and direct to the more large-scaled and abstract.

Speech therapy

During the first formal user study, one of the sign language interpreters (a qualified speech therapist) wanted to use the Haptic Chair when training deaf people to speak. Upon conducting her speech therapy programme with and without the Haptic Chair, she expressed confidence that the Haptic Chair would be a valuable aid in this kind of learning. The Haptic Chair was modified so that a user was able to hear/feel the vibrations produced by voice of the speech therapist and his/her own voice. With this modification, the Haptic Chair is currently being tested to enhance its effectiveness for speech therapy. The speech therapist is currently conducting her regular speech therapy program with 3 groups of students under 3 different conditions.

a. Haptic chair with no sound/vibration output
b. Haptic chair with complete sound/vibration output
c. Normal chair

Each student's ability of speech will be assessed (before and after every two weeks) by an independent person or several people to avoid any intentional or unintentional bias. The preliminary improvements displayed by the deaf users indicate the possibility of significantly improving their competence in pronouncing words with the usage of the haptic chair.

Amplitude modulated ultrasound

One of the limitations of experiencing music through the Haptic Chair was the fact that hearing-impaired people could not hear the exact lyrics of a song. One possible solution for this is to use Amplitude Modulated (AM) ultrasound. Staab *et al*. [75] found that when speech signals are used to modulate the amplitude of an ultrasonic carrier signal, the result was clear perception of the speech stimuli and not a sense of high-frequency vibration. It might be possible to use this technology to modulate a music signal using an ultrasonic carrier signal which might result in clear perception of lyrics in a song or simply music. This concept is currently being developed/tested and our preliminary tests showed that hearing is possible via ultrasonic bone conduction. One profoundly deaf participant was able to differentiate AM music and speech. He preferred the sensation when music was presented through AM ultrasound over speech presented through AM ultrasound, could not explain what he heard but simple reported he preferred the "feeling" of music through AM ultrasound. These observations open up an entirely new field to explore.

Detecting multiple vibrotactile stimuli by touch

The work by Karem *et al*.[58], have shown that the emotional responses are stronger when different parts of the musical signal (separated by frequency regions or by instrumental part) are delivered through different vibration elements to different locations on a user's back. One explanation for the improved enjoyment is that there might be masking of some portion of the audio signal that is eliminated by the spatial separation of musical or frequency components. Another explanation has to do with the difference between the nature of the signals typically processed by the skin and the ear. Multiple sound sources excite overlapping regions of the cochlea, and the auditory brain has evolved to perform source segregation under such conditions, whereas multiple sources of tactile stimuli sensed through touch are typically represented by distinct spatial separation. One possible future study would be to determine whether multiple sources can be detected when delivered through a single channel of vibrotactile stimulation. If not, it would significantly enhance the musical information available to spatially segregate sources from each other.

Display size

The size of the LCD display was not considered as a variable in this study, and we chose a commonly available 17 inch size that was both easily portable and widely available in homes and workplaces. However, the question of whether screen size is a significant factor is an important issue to study in our future work.

Other possible modifications

The system can be modified to capture specific ambient warnings and alerts including boiling kettle, phone rings, door bell that are important to most people. The safety of the deaf should not be compromised while they are enjoying their favourite music. Moreover, the system can also be modified to connect to a microphone array instead of a recorded multi-audio source. Multiple vibrating speakers could be rearranged and configured to indicate the direction of a sound source respective to the listener. This will be useful for the hearing-impaired in assisting them to judge the direction of a sound source which might be a warning of impending danger or required action on the part of the listener. The proposed system could also be used as an aid in learning to play a musical instrument or sing in tune. These possibilities need to be formally tested.

Another possible future study would be to determine the amount of information could be delivered using vibrotactile stimuli over 1000 Hz. This includes an investigation of the ability to identify multiple instruments in a given vibrotactile stimulus (timbre discrimination). In addition, we plan to study whether multiple sources can be detected when delivered through a single channel of vibrotactile stimulation. We hope that these future works will lead to other novel uses of the vibrotactile channel for more effective communication among a broader population.

Finally, the methodology outlined in this dissertation might enhance the enjoyment of music for people with normal hearing and those with narrow sound frequency band "drop-outs". The latter is a relatively common form of hearing loss that is often not severe enough to classify the person as deaf but might cause annoying interruptions in their enjoyment of music or conversation. The Haptic Chair has the potential to bridge these gaps to support music or other types of acoustic enjoyment for this community.

The end of this dissertation is certainly a new beginning to a multitude of work which could make use of this system for the betterment of the hearing-impaired community. May this work be a "prelude" for the many more tunes the deaf community are to hear!

References

[1] B. Tillmann, S. Koelsch, N. Escoffier, E. Bigand, P. Lalitte, A. Friederici, and D. V. Cramon, "Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex," *NeuroImage*, vol. 31, pp. 1771-1782, 2006.

[2] F. W. Ho-Ching, J. Mankoff, and J. A. Landay, "Can you see what I hear? The design and evaluation of a peripheral sound display for the deaf," in *Proc.* CHI '03: Proceedings of the 21st annual SIGCHI conference on Human factors in computing systems, 2003, pp. 161-168.

[3] T. Matthews, J. Fong, and J. Mankoff, "Visualizing non-speech sounds for the deaf," in *Proc.* 7th international ACM SIGACCESS conference on Computers and accessibility, Baltimore, MD, USA, 2005, pp. 52-59.

[4] M. A. Meredith, "On the neuronal basis for multisensory convergence: a brief overview," *Cognitive Brain Research*, vol. 14, pp. 31-40, 2002.

[5] C. Faulkner, *The Essence of Human-Computer Interaction*: Prentice-Hall, 1998.

[6] D. E. Glennie.Hearing Essay.
Internet: http://www.evelyn.co.uk/live/hearing_essay.htm., [Jul. 8, 2009].

[7] S. D. Barnett.SDB Entertainment.
Internet: http://shawndalebarnett.50megs.com, [Jul. 8, 2009].

[8] Musicians with a hearing loss. Internet: http://en.wikipedia.org/wiki/List_of_deaf_people#Musicians_with_a_hearing_loss, Jun. 16, 2009 [Jul. 8, 2009].

[9] I. P. Howard, *Human Spatial Orientation*. London, England: Wiley, 1966.

[10] H. McGurk and J. MacDonald, "Hearing lips and seeing voices," *Nature*, vol. 264, pp. 746-748, 1976.

[11] R. E. Cytowic, *Synaesthesia: a Union of the Senses*. New York: Springer-Verlag, 1989.

[12] S. A. Day.(2001).Brief History of Synaesthesia and Music. Internet: http://www.thereminvox.com/article/articleprint/33/-1/5/, Feb. 21, 2001 [Jul. 8, 2009].

[13] A. Ingerl and N. Döring, "Visualization of Music - Reception and Semiotics," in *Proc.* Audio Mostly, 2007, pp. 9-12.

[14] A. Ione and C. Tyler, "Neuroscience, History and the Arts Synesthesia: Is F-Sharp Colored Violet?," *Journal of the History of the Neurosciences*, vol. 13, pp. 58-65, 2004.

[15] M. J. Dixon, D. Smilek, and P. M. Merikle, "Not all synaesthetes are created equal: Projector versus associator synaesthetes," *Cognitive, Affective, & Behavioral Neuroscience*, vol. 4, pp. 335-343, 2004.

[16] K. I. Taylor, H. E. Moss, E. A. Stamatakis, and L. K. Tyler, "Binding crossmodal object features in perirhinal cortex," *National Academy of Sciences*, vol. 103, pp. 8239-8244, 2006.

[17] M. Schutz and S. Lipscomb, "Hearing gestures, seeing music: Vision influences perceived tone duration," *Perception*, vol. 36, pp. 888-897, 2007.

[18] B. W. Vines, C. L. Krumhansl, M. M. Wanderley, and D. J. Levitin, "Cross modal interactions in the perception of musical performance," *Cognition*, vol. 101, pp. 80-113, 2006.

[19] W. F. Thompson, F. A. Russo, and L. Quinto, "Audio-visual integration of emotional cues in song," *Cognition and Emotion*, vol. 22, pp. 1457-1470, 2008.

[20] M. M. Wanderley and B. Vines, "Ancillary Gestures of Clarinettists," in *Music and Gesture,* A. Gritten and E. King, Eds.: Ashgate Publishing, 2006.

[21] D. Shibata, "Brains of Deaf People "Hear" Music," vol. 16: International Arts-Medicine Association Newsletter, 2001.

[22] C. Kayser, C. I. Petkov, M. Augath, and Nikos K. Logothetis, "Integration of Touch and Sound in Auditory Cortex," *Neuron* vol. 48, pp. 373-384, 2005.

[23] C. M. Reed, "The implication of the Tadoma Method of speechreading for spoken language processing," in *Proc.* 4th International Conference on Spoken Language, Philadelphia, PA, USA, 1996, pp. 1489-1492.

[24] R. Palmer, "Feeling Music," Based on the paper presented at the 3rd Nordic Conference of music therapy, Finland, 1997.

[25] K. Myles and M. S. Binseel, "The Tactile Modality: A Review of Tactile Sensitivity and Human Tactile Interfaces," MD USA 2007.

[26] E. Hoggan and S. A. Brewster, "Designing Audio and Tactile Crossmodal Icons for Mobile Devices," in *Proc.* ACM International Conference on Multimodal Interfaces, Nagoya, Japan

[27] E. Hoggan and S. A. Brewster, "Crossmodal Icons for Information Display," in *Proc.* CHI '06: CHI '06 extended abstracts on Human factors in computing systems, 2006, pp. 857-862.

[28] E. Hoggan, S. A. Brewster, and J. Johnston, "Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens," in *Proc.* CHI '08: Proceedings of the26th annual SIGCHI conference on Human factors in computing systems, 2008, pp. 1573-1582.

[29] S. A. Brewster, F. Chohan, and L. M. Brown, "Tactile Feedback for Mobile Interactions," in *Proc.* CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, CA, USA, 2007, pp. 159-162.

[30] J. Pickett, "Tactual communication of speech sounds to the deaf," *Journal of Speech and Hearing Research*, vol. 6, pp. 207-222, 1963.

[31] A. Israr, H. Z. Tan, and C. M. Reed, "Frequency and amplitude discrimination along the kinesthetic-cutaneous continuum in the presence of masking stimuli," *Journal of the Acoustical Society of America*, vol. 120, pp. 2789-2800, 2006.

[32] D. Holten, J J V Wijk, and J. B. Martens, "A perceptually based spectral model for isotropic textures," *ACM Transactions on Applied Perception*, vol. 3, pp. 376-398, 2006.

[33] D. Birnbaum and M. Wanderley, "A systematic approach to musical vibrotactile feedback," in *Proc.* 2007 International Computer Music Conference (ICMC-07), 2007

[34] Y. Yokokohji, R. L. Hollis, and T. Kanade, "WYSIWYF Display: A Visual/Haptic Interface to Virtual Environment," *Presence: Teleoperators and Virtual Environments*, vol. 8, pp. 412-434, 1999.

[35] V. G. Chouvardas, A. N. Miliou, and M. K. Hatalis, "Tactile displays: Overview and recent advances," *Displays*, vol. 29, pp. 185-194, 2008.

[36] S. A. Brewster and L. M. Brown, "Tactons: Structured Tactile Messages for Non-Visual Information Display," in *Proc.* Australasian User Interface Conference, Dunedin, New Zealand, 2004, pp. 15-23.

[37] E. Gunther, G. Davenport, and S. O'Modhrain, "Cutaneous grooves: Composing for the sense of touch," in *Proc.* Conference on New interfaces for musical expression (NIME'02), 2002, pp. 1-6.

[38] J. B. Mitroo, N. Herman, and N. I. Badler, "Movies from music: Visualizing musical compositions," *SIGGRAPH Comput. Graph*, vol. 13, pp. 218-225, 1979.

[39] R. T. Verillo, "Investigation of some parameters of the cutaneous threshold for vibration," *Journal of the Acoustical Society of America*, vol. 34, pp. 1768-1773, 1962.

[40] R. T. Verillo, "Vibration sensing in humans," *Music Perception*, vol. 9, pp. 281-302, 1992.

[41] M. T. Marshall and M. M. Wanderley, "Vibrotactile feedback in digital musical instruments," in *Proc.* NIME ’06: Proceedings of the 2006 conference on New interfaces for musical expression, Paris, France, 2006, pp. 226-229.

[42] O. Fischinger, " Ten Films." Center for Visual Music CVM, Los Angeles, 2006. Internet: http://www.centerforvisualmusic.org/DVD.htm [Jul. 8, 2009].

[43] R. Jones and B. Nevile, "Creating Visual Music in Jitter: Approaches and Techniques," *Computer Music Journal*, vol. 29, pp. 55-70, 2005.

[44] R. Russet and C. Starr, *Experimental Animation: an Illustrated Anthology*. New York: Van Nostrand Reinhold, 1976.

[45] B. Evans, "Foundations of a Visual Music," *Computer Music Journal*, vol. 29, pp. 11–24., 2005.

[46] S. DiPaola and A. Arya, "Emotional remapping of music to facial animation," in *Proc.* Proceedings of the 2006 ACM SIGGRAPH symposium on Video games, 2006.

[47] S. A. Malinowski.Music animation machine.
Internet: http://www.musanim.com/mam/mamhist.htm., [Jul. 8, 2009].

[48] R. Hiraga, F. Watanabe, and I. Fujishiro, "Music Learning through Visualization," in *Proc.* 1st International Symposium on Cyber Worlds (CW'02), 2002, pp. 101.

[49] J. Foote, "Visualizing music and audio using self-similarity," in *Proc.* 7th ACM international conference on Multimedia, Orlando, Florida, United States, 1999, pp. 77-80.

[50] O. Kubelka, "Interactive music visualization," Czech Technical University.

[51] S. Smith and G. Williams, "A visuaization of music," in *Proc.* 8th conference on Visualization '97, Phoenix, Arizona, United States, 1997, pp. 499-503.

[52] P. McLeod and G. Wyvill, "Visualization of musical pitch," in *Proc.* Proceedings of the Computer Graphics International IEEE, 2003, pp. 300-303.

[53] R. Taylor, P. Boulanger, and D. Torres, "Visualizing emotion in musical performance using a virtual character," in *Proc.* 5th International Symposium on Smart Graphics, Munich, Germany, 2005, pp. 13-24.

[54] R. Taylor, D. Torres, and P. Boulanger, "Using music to interact with a virtual character," in *Proc.* International Conference on New Interfaces for Musical Expression (NIME'05), Vancouver, Canada, 2005, pp. 220-223.

[55] R. Palmer.(1994).Tac-tile sounds system (TTSS).
Internet: http://www.kolumbus.fi/riitta.lahtinen/tactile.html., [Jul. 8, 2009].

[56] S. Kerwin.(2005).Can you feel it? speaker allows deaf musicians to feel music. Internet: http://www.brunel.ac.uk/news/pressoffice/pressreleases/2005/cdata/october/vibrato, Oct. 22, 2005 [Jul. 8, 2009].

[57] M. Karam, G. Nespoli, F. Russo, and D. I. Fels, "Modelling perceptual elements of music in a vibrotactile display for deaf users: A field study," in *Proc*. 2nd International Conferences on Advances in Computer-Human Interactions, 2009, pp. 249-254.

[58] M. Karam, F. A. Russo, C. Branje, E. Price, and D. Fels, "Towards a model human cochlea," in *Proc.* Graphics Interface, 2008, pp. 267-274.

[59] Oval Window Audio. Internet: http://www.ovalwindowaudio.com/, Jan. 28, 2009 [Jul. 8, 2009].

[60] P. Henry and T. R. Letowski, "Bone conduction: Anatomy, Physiology, and Communication," Army Research Laboratory Aberdeen Proving Ground, MD USA 2007.

[61] C. E. Sherrick, "Basic and applied research on tactile aid for deaf people: Progress and prospects," *journal of the Acoustical Society of America*, vol. 75, pp. 1325-1342, 1984.

[62] Site of Tactaid and Tactilator. Internet: http://www.tactaid.com/, [Jul. 8, 2009].

[63] T. Ifukube, "Discrimination of Synthetic Vowels by Using Tactile Vocoder and a Comparison to that of an Eight-channel Cochlear Implant," *IEEE Transactions on Biomedical Engineering*, vol. 36, pp. 1085-1091, 1989.

[64] Stereo Tactile Motion System. Internet: http://crowsontech.com/go/crowsontech/3343/en-US/DesktopDefault.aspx, [Jul. 8, 2009].

[65] Vibrating bodily sensation device. Internet: http://www.kunyoong.com/product.html?grp=GC00570837&cid=CA00570900, [Jul. 8, 2009].

[66] Ogawa X-Chair Internet: http://www.ogawaworld.net/ourproducts/relaxation/xchair/xchair.php, [Jul. 8, 2009].

[67] Soundbeam. Internet: http://www.soundbeam.co.uk/, [Jul. 8, 2009].

[68] M. Lenhardt, R. Skellett, P. Wang, and A. Clarke, "Human ultrasonic speech perception," *Science*, vol. 253, pp. 82-85, 1991.

[69] The Vonia Corporation. Internet: http://www.dowumi.com/eng_index.php, [Jul. 8, 2009].

[70] B. Deatherage, L. Jeffress, and H. Blodgett, "A Note on the Audibility of Intense Ultrasonic Sound," *Journal of the Acoustical Society of America* vol. 26, pp. 582, 1954.

[71] R. Dobie, M. Wiederhold, and M. Lenhardt, "Ultrasonic Hearing," *Science*, vol. 255, pp. 1584-1585, 1992.

[72] H. Hosoi, S. Imaizumi, T. Sakaguchi, M. Tonoike, and K. Murata, "Activation of the auditory cortex by ultrasound," *Lancet*, vol. 351, pp. 496-497, 1998.

[73] S. J. Abramovich, "Auditory perception of ultrasound in patients with sensorineural and conductive hearing loss," *The Journal of Laryngology & Otology* vol. 92, pp. 861-867, 1978.

[74] S. Imaizumi, H. Hosoi, T. Sakaguchi, Y. Watanabe, N. Sadato, S. Nakamura, A. Waki, and Y. Yonekura, "Ultrasound activates the auditory cortex of profoundly deaf subjects," *Neuroreport*, vol. 12, pp. 583-586, 2001.

[75] W. Staab, T. Polashek, J. Nunley, R. Green, A. Brisken, R. Dojan, C. Taylor, and R. Katz, "Audible Ultrasound for Profound Losses," *The Hearing Review*, vol. 36, pp. 28-32, 1998.

[76] F. Yates, "Contingency table involving small numbers and the χ2 test," *Supplement to the Journal of the Royal Statistical Society*, vol. 1, pp. 217-235, 1934.

[77] "ISO 13407. Human-centred design processes for interactive systems," Switzerland 1999.

[78] K. Hevner, "The affective character of the major and minor mode in music," *American Journal of Psychology*, vol. 47, pp. 103-118, 1935.

[79] L. E. Marks, "On associations of light and sound: the mediation of brightness, pitch, and loudness," *American Journal of Psychology*, vol. 87, pp. 173-188, 1974.

[80] Three Centuries of Color Scales. Internet: http://rhythmiclight.com/archives/ideas/colorscales.html., Oct. 19, 2004 [Jul. 8, 2009].

[81] I. C. Firth.(2009).Music and Colour: a new approach to the relationship Internet: http://www.musicandcolour.net/, [Jul. 8, 2009].

[82] M. Kawanobe, M. Kameda, and M. Miyahara, "Corresponding affect between music and color," in *Proc.* IEEE International Conference on Systems, Man and Cybernetics, 2003, pp. 4190-4197.

[83] E. Scheirer, "Music listening systems," in *Media Arts and Sciences, School of Architecture and Planning* vol. Ph.D. dissertation: Massachusetts Institute of Technology, 2000.

[84] E. Chew, "Modeling Tonality: Applications to Music Cognition," in *Proc.* 23rd Annual Meeting of the Cognitive Science Society, Edinburgh, Scotland, UK, 2001, pp. 206-211.

[85] D. Zicarelli, G. Taylor, J. K. Clayton, R. D. Jhno, and B. Nevile, "Max reference manual," Cycling ’74, 2005.

[86] General MIDI 1, 2 and Lite Specifications.
Internet: http://www.midi.org/techspecs/gm.php, [Jul. 8, 2009].

[87] O. Matthes.(2002).Flashserver external for Max/MSP.
Internet: http://www.nullmedium.de/dev/flashserver, [Jul. 8, 2009].

[88] E. Chew, "Towards a mathematical model of tonality," in *Operations Research Center*, vol. Ph.D. dissertation. Cambridge, MA: Massachusetts Institute of Technology, 2000.

[89] C. Chuan and E. Chew, "Audio key finding: considerations in system design and case studies on Chopin's 24 preludes," *EURASIP J. Appl. Signal Process.*, vol. 2007, pp. 156-156, 2007.

[90] E. Chew and Y. C. Chen, "Mapping MIDI to the spiral array: disambiguating pitch spellings," in *Proc.* 8th INFORMS Computing Society Conference (ICS ’03), Chandler, Ariz, USA, 2003, pp. 259-275.

[91] E. Chew and Y. C. Chen, "Real-time pitch spelling using the spiral array," *Computer Music Journal*, vol. 29, pp. 61-76, 2005.

[92] The concept of colour energy. Internet: http://www.colourenergy.com/what-is.html, [Jul. 8, 2009].

[93] S. C. Nanayakkara.(2007).Survey: Evaluating the representation of tonal changes using colors. Internet: http://www-rcf.usc.edu/~mucoaco/COLOR/en/index.php, Sept. 12, 2007 [Jul. 8, 2009].

[94] M. Schulze, "A New Monotonic and Clone-Independent Single-Winner election Method," vol. 17: Voting Matters 2003, pp. 9-19.

[95] M. Csikszentmihalyi, *Beyond Boredom and Anxiety*. San Francisco, CA, USA: Jossey-bass, 1975.

[96] B. D. Zumbo and D. W. Zimmerman, "Is the selection of statistical methods governed by level of measurement?," *Canadian Psychology*, vol. 34, pp. 390-400, 1993.

[97] D. R. Johnson and J. C. Creech, "Ordinal measures in multiple indicator models: A simulation study of categorization error," *American Sociological Review*, vol. 48, pp. 398-407, 1983.

[98] S. Donnadieu, S. McAdams, and S. Winsberg., "Context Effects in 'Timbre Space," in *Proc.* Proceedings of the 3rd International Conference on Music Perception and Cognition, 1994.

[99] FLINT Particle System. Internet: http://flintparticles.org/, Jun. 23, 2008 [Jul. 8, 2008].

[100] S. O'Modhrain and I. Oakley, "Touch TV: Adding Feeling to Broadcast Media," in *In proceedings of the European Conference on Interactive Television: from Viewers to Actors?* Brighton, UK., 2003.

[101] K. Walker and L. M. William, "Perception of Audio-Generated and Custom Motion Programs in Multimedia Display of Action-Oriented DVD Films," in *Proc.* HAID 2006 - Haptic and Audio Interaction Design - First International Workshop, 2006, pp. 1-11.

[102] D. Levitin, "Re: Music Through vision and Haptics," Personal E-mail (Aug. 5, 2008).

[103] T. DeNora, *Music in everyday life*. Cambridge: Cambridge University Press, 2000.

[104] A. Gabrielsson and S. Lindstrom, "Strong experiences of and with music," in *Musicology and sister disciplines; past, present, future*, D. Greer, Ed. Oxford: Oxford University Press, 2000, pp. 100-108.

[105] J. A. Sloboda, S. A. O'Neill, and A. Ivaldi, "Functions of music in everyday life: an exploratory study using the Experience Sampling Method," *Musicae Scientiae*, vol. 5, pp. 9-29, 2001.

[106] M. Sheridan and C. Byrne, "Ebb and flow of assessment in music," *British Journal of Music Education*, vol. 19, pp. 135-143, 2002.

[107] M. Csikszentmihalyi, *Flow: The psychology of optimal experience*. New York: HarperCollins, 1990.

[108] M. J. Lowis, "Music as a Trigger for Peak Experiences Among a College Staff Population," *Creativity Research Journal*, vol. 14, pp. 351-359, 2002.

[109] C. Byrne, R. MacDonald, and L. Carlton, "Assessing creativity in musical compositions: flow as an assessment tool," *British Journal of Music Education*, vol. 20 pp. 277-290, 2003.

[110] S. A. Jackson and H. W. Marsh, "Development and validation of a scale to measure optimal experience: The Flow State scale," *Journal of Sport and Exercise Psychology*, vol. 18 pp. 17-35, 1996.

[111] A. B. Bakker, "Flow among music teachers and their students: The crossover of peak experiences," *Journal of Vocational Behavior* vol. 66, pp. 26-44, 2005.

[112] L. A. Custodero, "Construction of musical understandings: The cognition-flow interface," *Bulletin for the Council of Research in Music Education*, vol. 142, pp. 79-80, 1999.

[113] L. A. Custodero, "Observable indicators of flow experience: a developmental perspective on musical engagement in young children from infancy to school age," *Music Education Research*, vol. 7, pp. 185-209, 2005.

[114] A. Liberman and D. Whalen, "On the relation of speech to language," *Trends in Cognitive Sciences*, vol. 4, pp. 187-196, 2000.

[115] L. D. Rosenblum, "Perceiving articulatory events: Lessons for an ecological psychoacoustics," in *Ecological Psychoacoustics*, J. G. Neuhoff, Ed. San Diego, CA: Elsevier, 2004, pp. 219-248.

[116] J. Davidson, "Visual perception of performance manner in the movements of solo musicians," *Psychology of Music*, vol. 21, pp. 103-113, 1993.

[117] R. T. Boone and J. G. Cunningham, "Children's expression of emotional meaning in music through expressive body movement," *Journal of Nonverbal Behavior*, vol. 25, pp. 21-41, 2001.

[118] S. H. Xia and Z. Q. Wang, "Recent advances on virtual human synthesis," *Science in China Series F: Information Sciences*, vol. 52, pp. 741-757, 2009.

[119] M. Rudolf, *The grammar of conducting: A comprehensive guide to baton technique and interpretation* 3ed. New York: Schirmer Books, 1995.

[120] C. Wöllner and W. Auhagen, "Perceiving conductors' expressive gestures from different visual perspectives. An exploratory continuous response study," *Music Perception*, vol. 26, pp. 143-157, 2008.

[121] B. J. P. Mortimer, G. A. Zets, and R. W. Cholewiak, "Vibrotactile transduction and transducers," *The Journal of the Acoustical Society of America*, vol. 121, pp. 2970-2977, 2007.

[122] "ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs)–Part 11: Guidance on usability," International Organization for Standardization (ISO), Switzerland 1998.

[123] N. Bevan, "Measuring usability as quality of use," *Journal of Software Quality*, vol. 4, pp. 115-130, 1995.

[124] J. P. Chin, V. A. Diehl, and K. L. Norman, "Development of an instrument measuring user satisfaction of the human-computer interface," in *Proc.* ACM Conference on Human factors in computing systems (CHI'88), Washington, D.C., United States, 1988, pp. 213-218.

[125] J. R. Lewis, "IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use," *International journal of Human-Computer Interaction*, vol. 7, pp. 57-78, 1995.

[126] A. M. Lund, "Measuring Usability with the USE Questionnaire," vol. 8. STC Usability SIG Newsletter, 2001.

[127] E. F. Evans, "Corical representation," in *Hearing Mechanisms in Vertebrates*, A. V. S. d. Reuck and J. Knight, Eds. Churchill, London, 1968.

[128] I. C. Whitfield and E. F. Evans, "Responses of auditory cortical neurones to stimuli of changing frequency," *Journal Sound and Vibration*, vol. 21, pp. 431-448, 1965.

[129] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *Journal of Physiology*, vol. 195, pp. 215-243, 1968.

[130] S. J. Bolanowski, G. A. Gescheider, R. T. Verrillo, and C. M. Checkosky, "Four channels mediate the mechanical aspects of touch," *Journal of the Acoustical Society of America*, vol. 84, pp. 1680-1694, 1988.

[131] H. Levitt, "Transformed Up-Down Methods in Psychoacoustics," *Journal of the Acoustical Society of America*, vol. 49, pp. 467-477, 1971.

[132] G. A. Gescheider, B. GüÇlü, J. L. Sexton, S. Karalunas, and A. Fontana, "Spatial summation in the tactile sensory system: Probability summation and neural integration," *Somatosensory and Motor Research*, vol. 22, pp. 255-268, 2005.

[133] R. T. Verillo, "Effect of Contactor Area on the Vibrotactile Threshold," *Journal of the Acoustical Society of America*, vol. 12, pp. 1962-1966, 1963.

[134] T. B. Koay, J. R. Potter, M. Chitre, S. Ruiz, and E. Delory, "A compact real-time acoustic bandwidth compression system for real-time monitoring of ultrasound," in *Proc.* Oceans 2004, Kobe, Japan, 2004, pp. 2323-2329.

[135] D. C. Howell, *Statistical Methods for Psychology*, 4th ed. London: Duxbury Press, 1997.

[136] M. Norusis, *SPSS 13.0 Statistical Procedures Companion*. Upper Saddle-River, NJ: Prentice Hall, 2004.

[137] J. Neter, W. Wasserman, and M. H. Kutner, *Applied Linear Statistical Models*. Homewood, IL: Irwin, 1990.

[138] P. G. Hoel, *Elementary Statistics*, 4th ed. New York: Wiley, 1976.

[139] N. Tideman, *Collective Decisions and Voting: The Potential for Public Choice*. Burlington: Ashgate, 2006.

Appendix A: Statistical Methods

This appendix provides a very brief overview of some of the statistical methods used in this dissertation. The statistical tests were run using either SPSS® or Microsoft Excel® software packages. For more comprehensive descriptions, derivations, and examples please refer to the citations.

Most of these statistical methods (except Schwartz Sequential Dropping method) were used to determine whether the difference in observed data at different conditions is "significant" or "not significant". "Significant" implies that the difference in observed data is due to the test conditions. "Not significant" implies that the difference in observed data is more likely due to a chance. Following statistical methods have been used at different stages depending on the type of data observed and type of analysis required.

A1.T-test (Paired sample T-test)

The paired sample t-test was used to compare the means of the same participants in two different experimental conditions. The hypotheses are:

H_0: $\mu_1=\mu_2$ (means of the two groups are equal)

H_1: $\mu_1\neq \mu_2$ (means of the two groups are different)

The test statistic is t with $n-1$ degrees of freedom, where n is the sample size. The p-value associated with t indicates the level of confidence. Lower the p-value, higher the confidence. For example, if the p-value is less than 0.05, there is evidence to reject the null hypothesis with 95% confidence. Thus, this would confirm that there is a difference in means across the paired observations (H_1 is valid with 95% confidence). Comprehensive description of T-test can be found in [135, 136].

A2. Analysis of Variance (ANOVA)

Even though this test is called analysis of variance, it is used to determine if there is a significant difference between the means. ANOVA was used to test for differences among two or more independent groups. Typically, one-way ANOVA is used to test for differences among at least three groups, since the two-group case can be covered by a t-test. When there are only two groups to compare, the t-test and the one-way ANOVA are equivalent.

One-way ANOVA was used to study the effects of one variable. One-way ANOVA for repeated measures was used when the participants are subjected to repeated measures; this means that the same group of participants are used for each test condition as followed in the experiments described in this dissertation. One-way ANOVA makes the following assumptions:

- The population from which samples are drawn is normally distributed.
- Sample cases are independent of each other.
- Variance between the groups is approximately equal.

Hypotheses for the comparison of independent groups are:

H_0: $\mu_1=\mu_2 ... =\mu_k$. (means of the all groups are equal)

H_1: $\mu_i \neq \mu_j$ (means of the two or more groups are not equal)

The test statistic is an *F* test with *k-1* and *N-k* degrees of freedom, where *N* is the total number of subjects and *k* is the number of groups. A low *p*-value for this test indicates evidence to reject the null hypothesis. In other words, lower the *p*-value the higher the chance of at least one pair of means is not equal.

However, the rejection of the null hypothesis does not specifically indicate which means are different. Therefore, post-hoc analysis was used to determine which means are significantly different from which other means. In this dissertation, a graphical comparison and **Tukey's Honestly Significant Difference (HSD)** was used for post-hoc analysis. Tukey's HSD is essentially a t-test, except that it corrects for experiment-wise error rate. This test provides *p*-values for all possible pair-wise compassions to make quantitative judgement of which means are significantly different from which other means. A plot of error bars was used to make a graphical comparison of means of the groups. If the *p*-value is low, chances are there will be little overlap between the two or more groups. If the *p*-value is not low, there will be a fair amount of overlap between all of the groups.

When there were two variables, one-way ANOVA tests would be able to assess only one variable at a time. Therefore, two-way ANOVA was used to study the effects of the two variables simultaneously. Two-way ANOVA would not only be able to assess both the variables at the same time, but also would show whether there is an interaction between the variables. A two-way test generates three *p*-values, one for each variable independently, and one for the interaction between the two variables. More information on ANOVA can be found in [136, 137].

A3. Chi-square Test

Both t-test and ANOVA was used when the observed data was truly quantitative and continuous. However, when the data is categorical, chi-square test was used to determine whether observed data from two groups are independent from each other.

The hypotheses were:

H_0: The variables are independent of each other. (There is no association between them).

H_1: The variables are not independent of each other.

A contingency table was created to calculate *p*-values to test the hypothesis. Frequency tables of two variables presented simultaneously are called contingency tables (see the Table 2.1 and 2.1). Contingency tables are constructed by listing all the levels of one variable as rows in a table and the levels of the other variables as columns, then finding the joint or cell frequency for each cell. The cell frequencies are then summed across both rows and columns. The sums are placed in the margins, the values of which are called marginal frequencies. The lower right hand corner value contains the sum of either the row or column marginal frequencies, both of which are equal to the number of observations made.

The next step in computing the chi-square statistic is the computation of the expected cell frequency for each cell. This is accomplished by multiplying the marginal frequencies for the row and column (row and column totals) of the desired cell and then dividing by the total number of observations. The value of the chi-square, χ^2, can be calculated by substituting the observed and expected frequencies into the standard chi-square formula.

The degrees of freedom is computed by multiplying one minus the number of rows, times one minus the number of columns. Once the degrees of freedom is known, critical values can be read from chi-square distribution for given *p*-value (typically 0.1, 0.05 or 0.01). This critical value can be compared with the calculated χ^2. If the calculated χ^2 is greater than the critical value, the null hypothesis can be rejected and concluded that there is an association between the variables. More details on chi-square test can be found in [135, 136, 138].

A4. Schwartz Sequential Dropping Method (SSD)

This method is one of the most widely used methods to select a single winner using votes that express preferences. This method was used in dissertation (Chapter 3) to find the most preferred key-to-colour mapping strategy. The Schwartz set is defined as follows:

- An unbeaten set is a set of candidates of whom none is beaten by anyone outside that set.
- An innermost unbeaten set is an unbeaten set that does not contain a smaller unbeaten set.
- The Schwartz set is the set of candidates who are in innermost unbeaten set.

The rules used to calculate the winner based on SSD were:

1. If there is a candidate who is not beaten by any other candidate, then that candidate wins.
2. Otherwise, calculate the Schwartz set, based only on un-dropped defeats.
3. Drop the weakest defeat among the candidates of that set. Go to step 1.

A comprehensive description of SSD can be found in [94, 139].

Appendix B: List of Publications

Significant amount of materials, ideas, and results from this dissertation have previously appeared in the following peer-reviewed publications.

Journal and Conference Papers

1. Human Computer Interaction
   - E. Taylor, **S. C. Nanayakkara**, L. Wyse, and S. H. Ong. "Enhancing Musical Experience for the Hearing-impaired using Visual and Haptic Inputs", *Human-Computer Interaction*, Nov. 2009. (Submitted)

2. ACM Conference on Human Factors in Computing Systems (SIG CHI'09)
   - **S. C. Nanayakkara**, E. Taylor, L. Wyse, and S. H. Ong. "An enhanced musical experience for the deaf: Design and evaluation of a music display and a haptic chair", in *Proc. 27th ACM Conference on Human Factors in Computing Systems (CHI'09)*, pp.337-346, Apr. 2009.

     ** This was the first ever full-paper by an all-NUS team accepted for this tier 1 conference*

3. INNOVATION: The magazine of Research & Technology
   - **S. C. Nanayakkara**, E. Taylor, L. Wyse, and S. H. Ong. "Music made richer: Stimulating the senses of touch and sight for an enhanced musical experience", *INNOVATION: The magazine of Research & Technology*, vol. 8, no. 2, pp.28-29, Dec. 2008.

4. IEEE Region 10 Student Paper Contest
   - **S. C. Nanayakkara**, A. K. Mishra, and D. Mahapatra. "Visual attention while watching movies", *IEEE Region 10 Student Paper Contest*, Mar. 2007.

5. International Conference on Information, Communications and Signal Processing (ICICS'07)
   - **S. C. Nanayakkara**, E. Taylor, L. Wyse, and S. H. Ong. "Towards building an experiential music visualizer", in *Proc. 6th International Conference on Information, Communications and Signal Processing (ICICS'07)*, pp.1-5, Dec. 2007.

Patents

1. Haptic Chair with Audiovisual Input
   - E. Taylor, **S.C. Nanayakkara**, L.L. Wyse, S.H. Ong, K.P. Yeo and G.H. Tan. "Haptic Chair with Audiovisual Input", US Patent No. WO 2010/033086 A1, Mar. 25, 2010.

NEWSPAPER ARTICLES

1. National News Paper in Singapore

   - "A chair that's music to deaf ears" *Straits Times, Singapore* (Jul. 4, 2009), sec. D pp. 8.

2. National News Paper in Sri Lanka

   - "*Haptic Chair* hearing for the deaf" *The Island, Sri Lanka* (Dec. 4, 2008), sec. LL pp. 3.

UNIVERSITY RESEARCH GALLERY ARTICLES

1. Hearing through sight

   - "Hearing through sight", National University of Singapore–Research gallery. Internet: http://www.nus.edu.sg/research/rg99.php. [Jun. 8, 2009].

2. New technology to help the deaf enjoy music

   - "New technology to help the deaf enjoy music", National University of Singapore–Research gallery. Internet: http://www.nus.edu.sg/research/rg163.php. [Jul. 25, 2009].