Fakultät I
Institut für Sprache und Kommunikation
Fachgebiet Audiokommunikation

# A Framework for Audio-Tactile Signal Translation

Master Thesis Proposal

Primary Supervisor: Prof. Dr. Stefan Weinzierl
Secondary Supervisor: Dr. Charalampos Saitis
Queen Mary University of London

Submitted by: Maximilian Weber
Matriculation Number: 385153

Technische Universität Berlin
Fakultät I, Institut für Sprache und Kommunikation
2019-10-07

# Contents

# 1 Abstract

To enable a transition away from primitive, buzzing vibrations towards an new generation of wideband vibrotactile feedback requires strategies and tools for designing, storing and transmitting tactile stimuli signals and patterns. Due to the early sensory integration of auditory and vibrotactile modalities [1] and resulting perceptual similarities between them [2, 3] it appears feasible to translate auditory to vibrotactile signals and therefore enable the usage of the vast amounts of audio material available to the tactile sense. This insight might become useful as auditory perception is well researched — methods for handling, editing and displaying audio material are widespread and might prove to be useful for vibrotactile signals aswell.

This proposal aims to validate the *coherence* of unimodal vibrotactile stimuli towards their (non-musical) auditory sources utilizing signal processing methods discussed in previous works [4, 5]. While these works focused on musical audio sources and used a bimodal stimuli setup this work aims to use non-musical abstract audio material as audio sources. The validation of this process is proposed to be done by an empirical user test using both unimodal and bimodal, non-interactive stimuli presentation to measure the sense of coherence between both the auditory and vibrotactile stimuli.

This work further aims to refine signal processing methods for audio to vibrotactile signal translation by using an intermediate signal decomposition method. This intermediate step allows for a parametric representation of the signal that allows for precise manipulation and better control over the re-synthesis process of the vibrotactile stimuli — while maintaining spectral and temporal coherence towards the audio source. This method is also expected to generate coherent results across various vibrotactile display technologies due to the flexibility afforded by the proposed decomposed model. A second part of the empirical user test will utilize re-synthesized stimuli to see if the coherence is maintained throughout the re-synthesis process.

Validating a reliable method for audio to vibrotactile signal translation will enable the opportunity to research vibrotactile perception with regards to timbre, semantics and validating other, audio-related processing methods for vibrotactile stimuli. This will give rise to an opportunity to apply well-researched topics of audio signal processing and psychophysics to the study of vibrotactile signals and their perception.

After a formal introduction and motivation for this work in Chapter 2 the state of the art of auditory and tactile perception, vibrotactile signal generation and display technologies are introduced in Chapter 3. The proposed signal processing framework, audio source material and methods for the empirical user test are described in Chapter 4. Information on preparatory work and a schedule for the proposed thesis efforts can be found in Chapters 5 and 6 respectively.

# 2 Introduction and Motivation

"Sound is touch at a distance." This quote by Stanford's Anne Fernald comes from a podcast in 2007[1] in light of her research on the cognitive response of infants towards their mothers voice. In this embodiment the word "touch" is attributed metaphorically towards an emotional response induced by sound [6]. Being emotionally "touched by sound", especially by music is a common experience many people can relate to. Explaining this emotional or cognitive response to sound is a core research topic in the field of music psychology [7, 8].

For the efforts described in this proposal "being touched by sound" is meant in a more literal sense: Sound as a propagating vibration can not only be sensed by the human ear, but also by mechanoreceptors in the skin. The resulting, evoked sensation of touch can be induced by various mechanical forces such as pressure fluctuations, shearing forces and vibrations applied to the skin [9]. The sense of touch is commonly referred to as the tactile modality, which together with the proprioceptive (the sense of position) and kinaesthetic (the sense of movement) modalities constitute what is know as haptics or haptic perception. Even though the skin is argued to be the second largest human organ [10] and the skin's largely exposed surface area to the outside world the sensation of touch and it's role in perception is not as well understood as the visual or auditory sense.

The auditory and the vibrotactile sensitivity to vibrations overlap in a frequency range between 20 to 1000 Hz [11, 12]. To give an example on what that means lets consider an acoustic event: Given the sound of said event transports enough energy in this frequency range through vibrations, either air- or structure-borne, to allow for a mechanical deformation of the skin sound can not only be heard (auditory system) but also felt through the skin's mechanoreceptors (somatosensory system). This experience is common at concerts or when perceiving the engine's vibrations while driving a car. The effects of integrated auditory and tactile sensations have been researched in recent publications and cross-modal effects on loudness perception and quality of musical reproduction have been discovered [13, 14].

The role of joint (bimodal) auditory-tactile perception of music reproduction has been explored in recent years [5, 4]. This proposal however wants to take a more generalized approach by enabling research on similarities and differences between both modalities in

[1]WNYC RadioLab Podcast "Sound As Touch", September 24th 2007
https://www.wnycstudios.org/story/91514-sound-as-touch

an *unimodal* fashion by deriving vibrotactile stimuli signals from arbitrary, *non-musical* audio sources. In an empirical user test both auditory and vibrotactile stimuli are therefore presented *consecutively* (unimodal) to investigate the perceived coherence between both modalities. The goal of this work is therefore to revisit and expand on existing signal processing methods for inter-modal stimuli signal translation and validate a method that is able to coherently convert audio signals to vibrotactile stimuli signals — even when the vibrotactile stimuli is presented without the auditory source stimuli simultaneously. We further propose a method for decomposition and re-synthesis to enable a flexible authoring and editing scheme that is also suspected to work well across various vibrotactile display technologies. The same user test will be run on a set of re-synthesized signals to make sure the experienced coherence is maintained throughout this process. The main research questions can be formulated as:

- If and how can audio signals be used for (wideband) vibrotactile stimulation?
- What signal processing design aspects are important for such a cross-modal signal translation?
- How can we validate a proposed signal translation method?

Addressing these questions will help better understand important issues at the intersection of the auditory and tactile modalities:

- Do we use the same language (semantics) to describe a cross-modal translated stimuli?
- What sort of domain knowledge can we translate to the tactile from the well researched audio domain?
- How can these findings inform design choices of audio and vibrotactile related applications in both unimodal (auditory or tactile) and bimodal (auditory and tactile) scenarios?

Beyond scientific research, applications in medicine, entertainment, and mobile communications can be envisioned. A better understanding of audio-tactile signal translation is beneficial for authoring and curating future applications using wideband vibrotactile feedback, since domain knowledge around audio workflows, transmission, processing and reproduction could potentially be transferred to vibrotactile content curation and tactile stimulation technology.

Progress in audio-to-vibrotactile translation can be utilized to aid people with hearing impairment in various applications by translating acoustic cues to touch and helping in daily navigation [15], transport immediate warnings [16] and other means of affective computing [17]. Many applications in the realm of human-machine interaction, such as teleoperation for industrial, medical or end user purposes benefit from vibrotactile

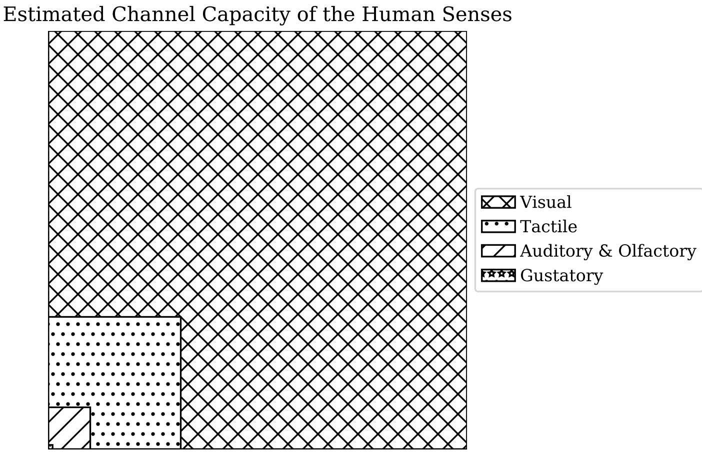


**Figure 2.1:** The estimated channel capacities of the human senses in context of information theory [20]. For the purpose of illustration the equivalent surface area for each sensory channel is displayed.

feedback next to (kinaesthetic) force-feedback to allow a remote operator more precise and immediate feedback without off-loading information to other senses, such as the highly loaded visual or auditory senses [18]. Off-loading information to other sensory modalities is highly desirable in today's information driven world: In his book "The user illusion: Cutting consciousness down to size" Tor Nørretranders estimates the information bandwidth of touch to be ten times higher than for hearing which gives a hint towards the untapped potential of this sense [19] (see figure 2.1 for reference).

Further use cases can be found in the entertainment sector, such as music, movies, telepresence and video games. Various mobile and desktop applications utilizing virtual and augmented reality technologies can benefit from additional modalities by providing a heightened sense of immersion for the user. In the past a majority of tactile hardware implementations used dull, low-bandwidth actuators such as Eccentric-Rotating Mass motors (ERMs) to induce haptic feedback. Due to the need in all industry segments to evolve away from these legacy solutions and to provide truly wideband high fidelity tactile feedback methods for curating vibrotactile signals will become mandatory.

A complexity arising from the industry is the need for a platform- and technology agnostic framework to integrate and display curated stimuli in a reliable and coherent way. The need for a unified format, agnostic to hardware capabilities (described in section 3.3) and maintaining the required information bandwidth across device-platforms (such as desktop, automotive, wearable, mobile) in a fragmented ecosystem bring problems that are yet to be adressed.

Since we are apparently only made consciously aware of an estimated 0.7% of the total sensory information processed by our body [20, 21] making informed design choices for all modalities (including touch) is an important aspect and will certainly be crucial in future user experience design.

# 3 State of the Art

To design a method that derives a vibrotactile stimuli from an auditory signal it is necessary to expose the characteristics of both auditory and tactile perception, as well as vibrotactile stimulation technology. Therefore this chapter provides an introduction to relevant subjects, such as auditory and tactile perception, methods for audio-tactile signal translation and vibrotactile stimulation technology.

## 3.1 Auditory and Tactile Perception

To form a coherent percept of the environment, an object or event, our brain combines information from various senses [22]. For an auditory-tactile experience the combination of both modalities occurs early and close to primary sensory areas, as experiments using functional magnetic resonance imaging (fMRI) scans of primate brains have shown [1]. An interesting, easy to replicate experiment illustrating such a sensory integration is named the "parchment-skin illusion" [23]. It is further, easy to reproduce evidence on how the auditory system can alter our perception of a haptic event, such as rubbing our hands together or tapping on a surface by modifying the corresponding auditory stimuli [24, 25, 26, 27].

Both the term haptic and tactile are often used synonymously, while in this embodiment a clear definition of each term is desirable. While the term haptic is often used as an "informal" umbrella term, it encompasses an *active* component — making the haptic modality an integrated experience of kinaesthetic, proprioceptic and tactile components. A tactile sensation however can be evoked both through active movement by the user, as well as by *passive* vibrotactile stimulation by an external process. This act of *passive* vibrotaction is the main focus of this work.

As early as 1851 it has been argued that haptic encompasses an interactive, exploratory act while tactile is a passive experience of touch [28]. This fortifies the definition of tactile being a subset of sensory information that remains when removing *active* proprioceptic and kinaesthetic attributes.

Even though both the auditory and tactile modality are sensitive to vibration they exhibit anatomical, functional and perceptual differences. The following sections are meant to give a brief overview of these similarities and differences between both sensory modalities.

### 3.1.1 Auditory Anatomy and Perception

Hearing - referred to as the auditory sense - is the ability to perceive sound by detecting vibrations with the ear. The visible part of the ear, composed of the pinna and the ear canal, is referred to as the outer ear. The function of the pinna is primarily to act as a funnel which assists in directing sound further into the ear canal. At the other end of the ear canal the eardrum (a.k.a. the tympanic membrane) separates the outer ear from the middle ear and the so called tympanic cavity. Sound as pressure fluctuations in the air is transmitted through the ear canal and excites the ear drum. Through small bone structures called ossicles, that are attached to the (excited) ear drum, sound is mechanically transmitted to the inner ear via the membrane-covered oval window. The other side of the oval window is a fluid-filled, spiral-shaped cavity called the cochlea. Sound is mechanically transmitted through the oval window and excites the fluid within the cochlea. It is important to note that this is not the only way the fluid in the cochlea can be excited: Structure-borne vibrations through bones and tissue within the body can also excite the fluid of the cochlea and form an auditory percept.

Inside the spiral-shaped cochlea energy is transmitted via traveling pressure waves of the fluid on the basilar membrane. The basilar membrane along the snail-shaped cochlea is wide at the opening and more narrow towards the center. It works similar to a frequency analyzer as the traveling waves in the fluid of the cochlea excite a frequency specific area via sensitive hair cells called Stereocilia along the organ of Corti. The displacement of the hair cells induces a change of electric conductance of the inner hair cell membranes. This change of conductance induces transmitters to be released to nerve endings and information to be transferred to the brain stem. The evoked action potentials travelling towards the central nervous system and the auditory cortex of the brain contain all of the temporally coded acoustical information. Within the auditory cortex and higher-level brain regions a percept of the acoustic event is formed.

Sound captured by the ear can be heard from about 20 Hz to 20 kHz. Next to inherited predisposition the age of a subject affects the upper hearing limit of a human subject. The most sensitive range of hearing is between 300 and 7000 Hz while being less sensitive for increasing or decreasing frequencies.

### 3.1.2 Tactile Anatomy and Perception

A haptic percept is an integrated, multimodal experience that is built up by a variety of proprioceptic, kinaesthetic and other sensory aspects. Various physical properties of an object, such as the shape, orientation, hardness, warmth conductivity and surface roughness can be sensed through exploratory action [29, 9]. These properties can be picked up by an integration of specialized somatosensory receptors, enabling us to sense object attributes, such as location, stiffness, temperature and texture. The integration of these attributes allow us to to form a unified percept of an object or event. The purely tactile components thereof mainly contribute to the sensation of mechanical forces, warmth conductivity and pain reception. Being able to passively sense mechanical forces by the skin's mechanoreceptors (while leaving out pain and thermal properties) is what we call the tactile sense. It allows us to sense perceptual properties, such as contact, shearing, pressure and vibrations. A tactile percept can also be formed by actively probing an object, but since active movement inherently influences the speed, location and pattern of exploration this extended active exploration process is commonly defined as haptics, while tactile remains a subset of haptics. This disambiguation is important due to the fact that once integrated sensory, kinaesthetic and proprioceptic information can not be separated and researched independently.

The mechanoreceptors responsible for the sense of touch are commonly classified according to their adaptation properties and morphology. Hairless (glabrous) parts of the skin contain four different types of receptors: Merkel's receptors (SA–I); Ruffini's corpuscles (SA–II); Meissner's corpuscles (RA–I); and Pacinian corpuscles (RA–II). Slowly adapting receptors (SA-I and SA-II) evoke action potentials as long as pressure on the skin is present. The firing rate of these receptors is proportional to the intensity of the applied force. The rapidly adapting receptors (RA-I and RA-II) mainly react to movement of the skin, for example in form of sheering or vibrational forces. The numerals I and II indicate the sizes of the corresponding receptive fields. Receptors marked with the numeral I lie close to the surface of the skin and have small receptive fields. Receptors deeper in the tissue have larger receptive fields and are labeled with the numeral II. An overview of the different properties of the mechanoreceptors mentioned above is provided in Table 3.1. In addition to the adaptation characteristics and the sizes of the receptive fields, mechanoreceptors differ regarding the minimum amount of force that is necessary to evoke a sensation, the density of the receptors, and the sensitive frequency range.

Each of the listed mechanoreceptors has a specific function for the sense of touch. The slowly adapting Merkel's receptors are active when applying static pressure (indentation) to the skin. Due to their small receptive fields, they are able to detect fine contours, such as borders and edges. Ruffini's corpuscles are specialized in detecting sheering forces, such as stretching of the skin. Meissner's corpuscles, which are only present in hairless skin areas, detect the speed of skin deformation at comparably slow rates. This enables them to detect low frequency vibrations. Last but not least, the Pacinian corpuscles are specialized in detecting the speed of skin deformation. Compared to the Meissner's corpuscles they are able to detect a larger frequency range and encompass the largest receptive fields.

Due to the ability of the Pacinian corpuscles (RA-II) of detecting frequencies between 40 Hz and 1000 Hz (peak sensitivity between 225 and 275 Hz) they are considered to be the most important receptors for this proposal, next to the Meissner corpuscles (RA-I)

| Receptor | Type | Frequency Range (peak sensitivity) | Threshold skin deformation on hand (median) | Receptive field (median) | Receptor density at fingertip (palm) |
|---|---|---|---|---|---|
| Merkel's receptors | SA-I | | $7 - 600\mu m$ ($56.5\mu m$) | $2 - 100mm^2$ ($11mm^2$) | $70/mm^2$ ($8/mm^2$) |
| Ruffini's corpuscles | SA-II | | $40 - 1500\mu m$ ($331\mu m$) | $10 - 500mm^2$ ($59mm^2$) | $9/mm^2$ ($15/mm^2$) |
| Meissner's corpuscles | RA-I | 5 - 200 Hz (25 - 40 Hz) | $4 - 500\mu m$ ($13.8\mu m$) | $1 - 100mm^2$ ($12.5mm^2$) | $140/mm^2$ ($25/mm^2$) |
| Pacinian corpuscles | RA-II | 40 - 1000 Hz (225 - 257 Hz) | $3 - 20\mu m$ ($9.2\mu m$) | $10 - 1000mm^2$ ($101mm^2$) | $21/mm^2$ ($9/mm^2$) |

**Table 3.1:** Properties of the hairless (glabrous) skin mechanoreceptors. Information derived from Treede and Russo et al. [33, 2]

with peak sensitivity between 25 and 40 Hz [30, 31]. The sensitivity may vary due to inheritance, sex and usually decreases when we get older towards higher frequencies — similar to decreased hearing sensitivity with age.

Similar to the auditory pathway present frequency and magnitude information is coded into time-varying patterns of action potentials before being transmitted to the sensory nervous system. High intensity vibrations evoke multiple action potentials for each cycle, whereas for low intensity vibrations not every cycle period results in the release of an action potential [32].

Early integration of both auditory and vibrotactile stimuli is evident through brain imaging [1, 3] and empirical studies: F.e. based solely on vibrotactile feedback participants (including a group of deaf participants) were able to discriminate between musical timbres and also identify timbre features, such as dull and bright varying only with regard to spectral centroid. These observations lead to a proposal that, as with auditory discrimination of musical timbre, vibrotactile discrimination may involve the cortical integration of filtered output from frequency-tuned mechanoreceptors functioning as critical bands [2].

## 3.2 Audio-Tactile Signal Translation

Given the ubiquity of high quality audio samples from recordings and digital synthesis, as well as the already present implementation of such material in various applications it is desirable to create vibrotactile stimuli from these readily available, information rich sources. Using an existing audio assets to derive a vibrotactile stimuli is especially desirable given the possibility for the resulting tactile stimuli to intrinsically *match* the temporal, dynamical and spectral progression of the audio asset to form a *coherent* percept.

Early experiments in auditory-tactile translation were conducted in the 1920s by Gault in an experiment using a 14 feet long tube that was pressed against the palm of a subjects hand. Through multiple training sessions a subject was able to correctly identify up to 34 words together with sentences constructed by those words in various combinations, hinting at the potential of acoustic vibration cues for tactile stimulation [34]. Given the complexity in temporal and spectral modulation of human speech understanding entire sentences through mere sound-induced vibrotactile stimuli is rather impressive and hints at the capabilities of the human skin in reliably differentiating vibrational patterns.

Efforts in unifying audio-tactile synthesis methods for virtual object interaction have been experimented with using physical model based synthesis approaches [35, 36]. These methods show that the fundamental principal of audio-tactile stimulation has been explored for interactive (haptic) applications, requiring active user participation. In a similar embodiment physical modelling based synthesis was used to investigate the role of physics-based auditory and haptic feedback provided at feet level to enhance realism in a virtual environment by simulating the audio-tactile sensation of walking on ground surfaces by providing both auditory cues through loudspeakers and vibrotactile stimuli through actuator-enabled footwear [37].

For joint audio-tactile music reproduction Merchel has proposed and tested various signal translation methods. In his work various methods ranging from low-pass filtering, octave shifting the audio signal, as well as sub-harmonic synthesis and signal substitution have been explored. Uniformly all of the experiments showed higher ratings in overall quality of the reproduction compared to using no additional tactile stimuli [5]. The method of pitch-shifting and low-passing the audio signal to render a vibrotactile stimuli was also utilized in a study examining the possible influences of vibrotactile stimulation on melody and rhythm perception of children and adults with cochlear implants [4]. Both studies mentioned above provide auditory and vibrotactile simultaneously (bimodal) whereas this embodiment aims to validate the possibility of using vibrotactile stimuli derived from audio without providing additional sensory modalities (unimodal).

## 3.3 Vibrotactile Stimulation Technology

In the last two decades various types of actuator technologies have been brought to market while other methods are still being researched. A review of various rendering techniques can be found in the work of Salisbury Et Al. "Haptic rendering: introductory concepts" [38].

The most widely spread actuator type today is the Eccentric Rotating Mass (ERM)[1] actuator. It is composed of an eccentric mass attached to the axis of a DC controlled motor. The speed of the DC motor controls both strength and frequency simultaneously, restricting the motor from creating a truly wideband frequency response. The centrifugal force of the mass can be felt as a buzzing vibrational force by the mechanoreceptors. ERM actuators are cheap and can be found in most mobile phones and peripheral hardware, such as game controllers today.

In more recent times Linear Resonant Actuators (LRA)[2] have become more commonplace in modern smart phones due to their enclosed form factor, efficiency and easier control. The efficiency of the LRA is due to the high Q-factor, which reduces the amount of power required to run the actuator close to the inherent resonance frequency. The high Q-factor also restricts these actuators from being truly wideband. Compared to ERM motors they have shorter rise- and fall times which makes them useful for recreating short, impulse-like stimuli like clicks for a User Interface (UI) application.

As of today the most promising actuator category for providing high-fidelity tactile stimuli are Voice Coil Actuators (VCA). Mechanically the VCA actuators share a high level of similarity to loudspeaker drivers. Tuning the mass, spring stiffness and coil properties allows these actuators to have a wider frequency response when compared to ERMs and LRAs. Using such an actuator for this thesis is highly likely due to their desirable frequency response and reasonable rise- and fall time. E.g. the Berlin-based company "Lofelt"[3] has constructed such advanced VCA technology.

When it comes to interactive (haptic) vibrotactile reproduction there are methods that enable a modulation of the already existing friction between the fingers and a surface. This method of "friction modulation" can be achieved by electrostatic modulation of a display or by inducing a ultrasonic vibration into the interaction surface. Through this method vibrotactile actuated touchscreens have been prototyped [39, 40, 41] and companies like "Hap2U"[4] are trying to bring this technology to market. From personal

[1] Precision Microdrives: Eccentric Rotating Mass Vibration Motors, last visited 24th May 2019
https://www.precisionmicrodrives.com/vibration-motors/eccentric-rotating-mass-vibration-motors-erms/

[2] Precision Microdrives: Linear Resonant Actuators, last visited 24th May 2019
https://www.precisionmicrodrives.com/vibration-motors/linear-resonant-actuators-lras/

[3] Lofelt GmbH: L5 Actuator, last visited 24th May 2019
https://lofelt.com/technology

[4] Hap2U: Ultrasonic Friction Modulation Technology, last visited 24th May 2019
http://www.hap2u.net/

experience these devices are still in research and don't provide the type of high fidelity response that would be required to validate audio-tactile translation methods, while keeping the structure of the stimuli coherent between both modalities. These displays also require additional interaction by the user which introduces further kinaesthetic and proprioceptic modalities to the audio-tactile reproduction which is not desirable for this research.

The most experimental method for vibrotactile reproduction are ultrasonic displays [42]. This technology is composed of a grid array of ultrasonic transducers and software allowing for gesture tracking and phase array control — comparable to wave field synthesis technology. Current limitations in the precision and resolution of these displays are due to the low density of transducers in the array and the physical limit of their individual size to pack them closer together. Applications of this technology reach from rendering 3D displays using Styrofoam pebbles that are levitated by concentrated pressure zones in the wave field, simulating textures in mid-air and also modulating the ultrasonic carrier with audio to enable object-based sound positioning. Next to the company "Ultrahaptics"[5] [43] many publications in this field come from a joint EU-funded research project called "Levitate"[6].

Due to systematic differences between the aforementioned vibrotactile display techniques finding a universal data-format, that allows for a coherent experience across all technologies poses a current industry problem. This is because no single company wants to rely on an end-to-end solution that only works for a single component supplier.

Further, differences in available bandwidth, system resonance and on-body localization should ideally be solved by software and not pose a problem that needs to be addressed by content creators and product managers. Finding a universally applicable container format, that provides each technology with the necessary information is a key component for enabling a market shift towards the next generation in haptic feedback.

[5] Ultrahaptics: Tactile sensations in mid-air, last visited 24th May 2019
https://www.ultrahaptics.com/

[6] "Levitate" publications, last visited on 18th June 2019
https://www.levitateproject.org/publications

# 4 Method

## 4.1 Audio Source Material

To cover the requirement of non-musical, “abstract” audio source material a corpus of audio samples from CNRS’s PRISM laboratories[1] is proposed. These abstract sounds lack readily available source/cause associations, such as those found in instrumental and vocal timbre classes and therefore reduce the role of linguistic mediation when describing the evoked sensation. A useful way of collecting such sounds is through electroacoustic music, where sound identities appear intentionally obscured or unconnected to their source.

Researchers at the PRISM laboratory in Marseille have previously compiled the aforementioned collection of 200 electroacoustic sounds representative of the nine balanced sound classes of Schaeffer’s typology of *acousmatic* sounds—sounds experienced by attending to their intrinsic morphology and not to their physical cause [44]. These are based on three profiles of temporal energy envelope (continuous, impulse, iterative) and three profiles of spectral content (tonal, complex pitch, varying pitch) [45]. This classification system offers an objective tool to obtain a sound corpus representative of most sound morphologies.

However, sounds with iterative envelopes and/or varying pitch may not be suitable to study intrinsic qualia aspects of timbre. For example, the second type tends to evoke motion [44]. Accordingly, starting from these 200 samples (made available courtesy of PRISM), only those sounds that are representative of the four {continuous, impulse}×{tonal, complex pitch} classes of Schaeffer’s typology (see figure 4.1) will be selected through an informal listening protocol as the generic sound material for the proposed auditory-tactile translation and empirical user test.

## 4.2 Audio-Tactile Signal Translation Method

There are various methods to synthesize vibrotactile stimuli from audio sources found in previous research. The simplest method consists of merely low-passing an existing audio signal within the range of tactile sensitivity (ranging from 5 Hz to 1000 Hz). This

[1] Laboratoire PRISM: Perception Représentations Image Son Musique, last sighted: 11. June 2019 https://www.prism.cnrs.fr/

**FACTURE / SUSTAINMENT**

| MASS | *continuous* | | | *impulse* | *iterative* | | |
|---|---|---|---|---|---|---|---|
| | *unpredictable* | *nonexistent* | *formed* | | *formed* | *nonexistent* | *unpredictable* |
| *tonal* | En | Hn | **N** | **N’** | N” | Zn | An |
| *complex* | Ex | Hx | **X** | **X’** | X” | Zx | Ax |
| *varying* | Ey | Tx/Tn | Y | Y | Y” | Zy | Ay |
| *unpredictable* | E | T | W | Φ | K | P | A |

**Table 4.1:** Schaeffer’s typology of sound objects. The highlighted cells illustrate the desired types used for this study.
Table derived from *Des Objets Sonores* [46].

method is feasible, given sufficient information is contained in the low frequency band of the original audio signal.

If the amount of information in the vibrotactile sensitive band is too low (or non-existent) the signal requires artificial augmentation or substitution to ensure a coherent reproduction. Previous studies experimented with various methods to allow for such augmentation, such as frequency-shifting the signal or substituting the vibrotactile signal by a synthesized low-frequency sinusoidal following the temporal amplitude progression (envelope) of the audio signal [13, 4].

The amount of information (entropy) present in a frequency band can be estimated using an energy distribution approach, such as comparing the (weighted) Power Spectral Density (PSD) of the low-passed tactile-band in relation to the spectral energy content beyond that range or by estimating the Power Spectral Entropy of each respective frequency band. Other (spectral) audio features, such as Spectral Centroid, Spectral Spread and Spectral Slope [47] are considered to inform the processing chain if signal augmentation is required or not (see figure 4.1). Tuning the analysis to achieve a reliable automation on the decision of the signal processing might require work extending the efforts of this thesis. To ensure a practical method for the empirical user test and to negate interference of the validation of the aforementioned signal translation method this decision will most likely be made on a manual, subjective basis and use the filtering, frequency-shifting or signal-substitution methods discussed above depending on the intermediate results that seem most promising and useful across the chosen audio source material.

Another approach to make aforementioned signal analysis and translation more transparent and versatile is to decompose the signal into individual components and therefore enable a higher level of control in design and reconstruction of a vibrotactile stimuli. Such methods of decomposing and reconstructing a signal can be found in the domain of audio *re-synthesis*. The term re-synthesis generally describes a method comprised of an analytical process, in which a signal is decomposed to a sparse representation —

and a synthesis process, in which decomposed components are used to synthesize an approximation of the target signal. The resulting sparse representation of such a process can allow for more control over individual frequency components, spectral shape and temporal evolution of the target signal [48].

A known method for the decomposition of a signal and spectral modeling re-synthesis was made famous by a publication of Xavier Serra in 1990 [49]. This method retrieves modal components (deterministic sinusoidals) and residual noise components (filtered stationary stochastic noise). Sinusoidal components are traced in the frequency spectrum of a signal and separated into individual tracks. The spectral shape of the residual components can be approximated by using either a spectral approximation with random phase information, or by approximating the spectral shape with a set of filters on additive white Gaussian noise. Finally, both the deterministic and stochastic components are combined to retrieve an approximation of the original signal.

To make the initial separation easier a pre-filtering using Harmonic-Percussive Source Separation (HPSS) is proposed. This method applies median filtering on the target spectrogram $S$ along the time axis to emphasize sustained tonal components $H$ (harmonic) and applies the filtering again on the frequency axis to emphasize broadband components associated with transients $P$ (percussive) [50]. The residual (stochastic) component $R$ can be isolated by subtracting the harmonic components $H$ and percussive components $P$ from the original spectrogram by $R = S - (H + P)$. The "signal decomposition" method described here (also found in figure 4.1) is further illustrated in figure 4.2.

To enable an efficient trade-off between quality and the amount of data stored an approximation routine for the retrieved envelopes using the Douglas-Peucker Algorithm is proposed. This algorithm allows parametric control over a decimation of a curve composed of line segments to a similar curve with fewer points [51]. An alternative, computationally more intensive optimization method utilizes a Least-Square approach to fit a desired amount of breakpoints to the derived envelope curve data [52].

The overall decomposition method is interesting, as the intermediate information on pitch, spectral shape and the temporal evolution of a signal can be saved and individual components can be modified to optimize the signal for any connected vibrotactile display technology. The benefits of such a decomposed format would be to allow for a technology- and hardware-agnostic specification of the desired vibrotactile stimuli and allow high-level control on modal, stochastic and temporal components of the stimuli signal. A similar, but more simple tactile-stimuli-defining file format called "AHAP" (Apple Haptic and Audio Pattern) can be found in the aforementioned implementation of Apple's recent publication of their CoreHaptics API.

Another key benefit of a re-synthesis method is the capability of introducing variance to repeating events, such as game and user interface assets, by randomly altering modal component amplitudes and applying random filters to the residual noise. This allows for

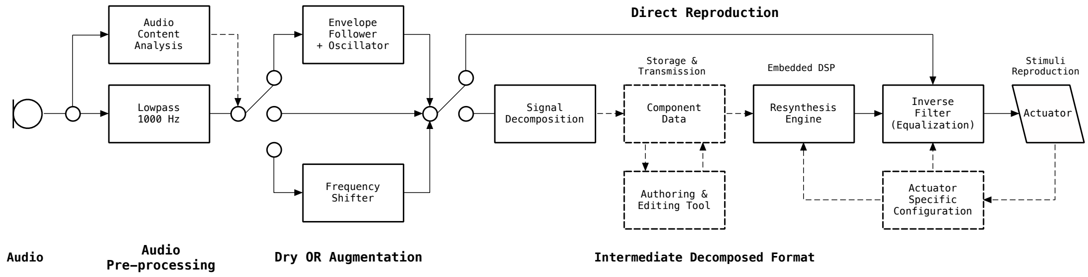


**Figure 4.1:** Flowchart of the proposed audio-tactile signal translation and data handling framework. Signal paths are illustrated as solid lines, control signals and data are illustrated using dashed lines. The "signal decomposition" block is further described in figure 4.2.

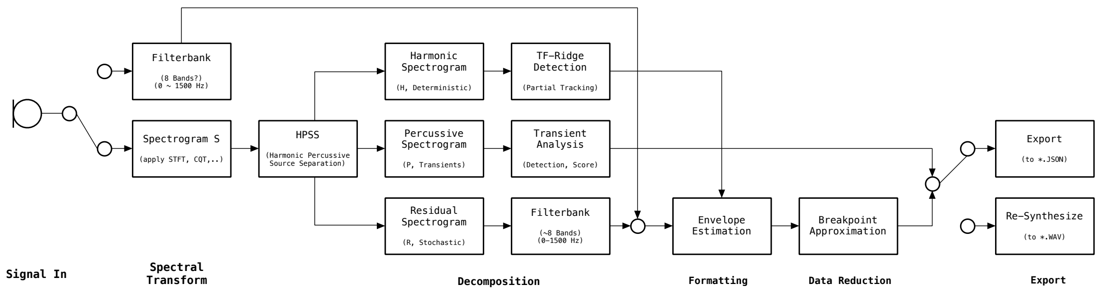


**Figure 4.2:** Flowchart of the proposed signal decomposition method. First the spectrogram of the incoming signal is computed. Next, Harmonic-Percussive Source Separation is applied to isolate tonal and transient components. The residual component is then computed using the retrieved harmonic and percussive spectrogram. The individual spectral components are then further analyzed to retrieve a parametric representation of each.

a significant reduction of the required bandwidth to transmit and/or the amount of data needed to store vibrotactile assets: Variations of the same interaction don't have to be stored and transmitted as separate files but can be synthesized on the fly, if desired [53].

Using a sparse or model-based representation to render vibrotactile stimuli has been explored in various embodiments, ranging from approximating recorded material texture signals by source-filter models [54, 55] and approximating texture signals from sinusoidal components [56]. Given that the decomposition method described by Serra results in both additive sinusoidals (deterministic components) and residuals estimated using subtractive synthesis (stochastic components) utilizing this method appears to be feasible while maintaining the desired level of control and ensuring a coherent translation between both the auditory and tactile modality across display technologies.

Finding out how much manual curation on a vibrotactile signal derived from an audio signal needs to be done, or how well this process can be automated through an expert system informed by audio content analysis and decomposition techniques will be a core discussion in this research effort, while keeping focus on validating the fundamental signal translation techniques.

## 4.3 Empirical User Test

To validate the coherence of the sensations derived from the inter-modal audio-tactile translation method an empirical user test in three groups, yielding direct and indirect, qualitative coherence measures is proposed. For this experiment we aim for at least 10 participants per group, resulting in a minimum of 30 individual test subjects. A total of ~10 audio-tactile stimuli pairs, selected from the aforementioned corpus (section 4.1) will be used.

The test procedure for group A is conducted in two separate tasks:

1. The participants will be asked to listen to the (~10) sound stimuli consecutively and to write down the first words that come to mind for each stimuli. The participants will be specifically asked to focus on the associations (i.e., concepts) evoked by the sounds without trying to identify the physical sources or events that produced them.
2. The participants will be presented audio and vibrotactile stimuli consecutively (unimodal) and will be asked to rate the "coherence" between both stimuli on an unipolar, continuous scale.

Group B will carry out the same two tasks, but task 1 will instead involve the corresponding vibrotactile stimuli.

Finally, participants in group C will be presented with audio-tactile stimuli pairs simultaneously (bimodal) and be asked to freely describe the combined sensation (i.e. similar to the first task of groups A and B) and rate the perceived coherence on a unimodal continuous scale.

*Coherence* here is meant to describe the level at which temporal and spectral (i.e. timbral) characteristics of the original (audio) signal are maintained throughout the signal translation process—while providing a comparable sensation across both modalities that we aim to measure through the proposed user test. The methods described above are further proposed to be used for the re-synthesized stimuli derived from the decomposed audio source material.

After collecting all test data we will then look at qualitative comparisons across the three group conditions as a measure of *coherence* for the inter-modal audio-tactile sensations based on the proposed signal translation method. To analyse the collected qualitative attributes we will utilize a psycholinguistic approach [57]. A similar experiment for word-sound relations instead of tactile-sound relations has been reported and results were discussed in terms of similar conceptual processing networks [58]. The quantitative data based on the continuous *coherence* scale will undergo a variance analysis to determine the overall rating, finding corner cases and to validate the statistical significance.

To display the vibrotactile stimuli to the subjects a wristband form factor is proposed. The proposed location for applying the stimuli is to the anterior of the underarm (similar to wearing a watch the wrong way around). This location allows for stimuli exposure to hairless (glabrous) skin and therefore stimulation of the underlying Meissner (RA-I) and Pacinian (RA-II) corpuscles [33]. Using a wearable device is beneficial compared to a (finger) touch display as the applied pressure, position and exposed surface area are easier to control between test subjects. This method also reduces the variability that is introduced by active user participation (i.e. trying to hold the finger in the same position with the same pressure) during the entirety of the test procedure.

To further reduce variability of the stimuli representation induced by non-linear actuator behaviour and varying body contact the response of the vibrotactile display is linearized (equalized) ad-hoc for each test subject after the armband has been attached to the subjects arm. Ad-hoc linearization (equalization) is important as the overall system behaviour and therefore the frequency response of the actuator is altered if attached to a body or device. In such a configuration the body/device acts as an additional, connected system by introducing changes to mass, stiffness and dampening through the skin tissues mass, elasticity and viscosity.

# 5 Preparatory work

This chapter describes all work that went towards the proposed efforts and has already been conducted previous to formulating this proposal.

Through my internship and student work at Lofelt I already have experience with using Voice Coil Actuators (VCAs), especially the model L5, for vibrotactile stimuli reproduction built into headphones and handheld devices. Setting up a suitable reproduction device for this thesis is therefore not the main body of work. Most likely little to no additional electrical engineering or efforts towards actuator design will be necessary for the proposed research.

Efforts towards researching sparse representations and signal coding methods for vibrotactile stimuli have already been conducted in an effort to propose a thesis topic on the feasibility of deriving vibrotactile codecs from audio codecs prior to the proposed topic. The main motivation here was to find data reduction methods due to the limited bandwidth available in target vibrotactile display systems and industry products. At the time no suitable supervisor could be found and the idea was shelved for the time being.

Self-experiments and ad-hoc user tests, like initial feasibility tests on audio-tactile signal translation and resynthesis methods have been conducted with subjectively promising results - but without proper, empirical verification and without refining the processing techniques. Products containing similar (real-time) audio-tactile translation methods are already on the market[1][2][3].

Finding a suitable audio corpus that covers the requirements of being non-musical but complex enough for the proposed user tests has been conducted in coordination with the thesis co-supervisor Dr. Charalampos Saitis. The work of categorizing about 200 samples along the proposed typological map (explained in section 4.1) has already been conducted by an informal listentening test. The next step would be to have one or two other persons do the same task and check then for coherence in the classification.

[1] The “Basslet” by Lofelt, last sighted 8th July 2019
https://lofelt.com/the-basslet

[2] Razer Nari Ultimate using a Lofelt L5 actuator, last sighted 8th July 2019
https://www.razer.com/de-de/gaming-audio/razer-nari-ultimate

[3] Teenage Engineering “Rumble Module” using a Lofelt L5 actuator, last sighted 19th September 2019
https://teenage.engineering/products/op-z/modules/rumble

# 6 Schedule

The work going into the proposed thesis topic is arranged in four phases (see figure 6.1):

1. As a first phase, thorough research in auditory-tactile perception, audio-tactile signal translation methods and related topics has been conducted to formulate this proposal. A suitable audio corpus was investigated and labeled according to a typological model.
2. The second phase is composed of a prototype implementation of a the proposed signal translation method and rendering the vibrotactile stimuli for the user test to validate the chosen method.
3. The third phase will require the setup of the proposed user test and a reliable way to display the auditory and tactile stimuli as a preparation for conducting the empirical user test. Further the collected data is processed and analyzed.
4. The fourth phase will be utilized to write the thesis by documenting the procedures and discussing the literature, research methods and user test results.

If any milestone is reached earlier, the schedule can shift accordingly as there are no blocking dependencies. If any milestone takes longer then anticipated it will be important to reschedule the following phases to avoid conflicts.

To maintain progress within each phase an (Agile) Scrum sprint methodology of two weeks is targeted, allowing each of the following phases approximately four sprints of time.

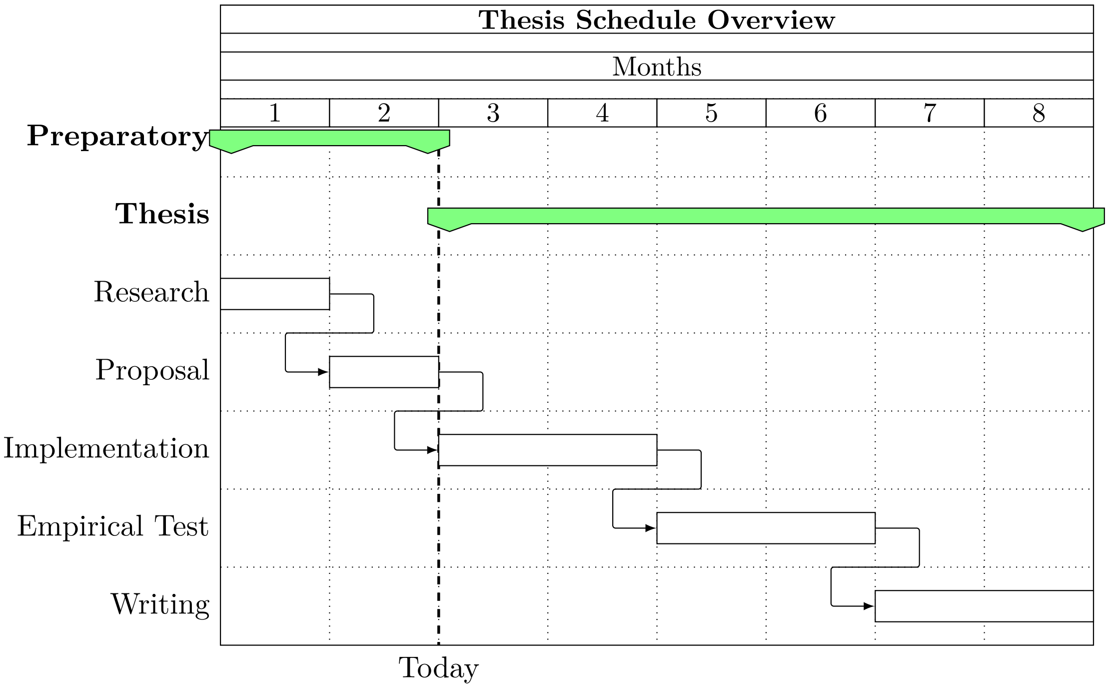


**Figure 6.1:** Proposed time line for the thesis illustrated as a Gantt chart.

# References


[1] Christoph Kayser, Christopher I Petkov, Mark Augath, and Nikos K Logothetis, "Integration of touch and sound in auditory cortex," *Neuron*, vol. 48, no. 2, pp. 373–384, 2005.

[2] Frank A Russo, Paolo Ammirante, and Deborah I Fels, "Vibrotactile discrimination of musical timbre.," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 38, no. 4, pp. 822, 2012.

[3] John J Foxe, "Multisensory integration: frequency tuning of audio-tactile integration," *Current biology*, vol. 19, no. 9, pp. R373–R375, 2009.

[4] Martha Papadogianni-Kouranti, Hauke Egermann, and Stefan Weinzierl, "Auditive and audiotactile music perception of cochlear implant users," *A Thesis presented for the degree of Master of Science*, pp. 1–28, 2014.

[5] Sebastian Merchel and Mehmet Ercan Altinsoy, "Auditory-tactile music perception," in *Proceedings of Meetings on Acoustics ICA2013*. ASA, 2013, vol. 19, p. 015030.

[6] Anne Fernald, "Intonation and communicative intent in mothers' speech to infants: Is the melody the message?," *Child Development*, vol. 60, no. 6, pp. 1497–1510, 1989.

[7] Kurt Kumler, *Being Touched by Music: A Phenomenological-hermeneutical Approach to Understanding Transformational Music Experience*, Ph.D. thesis, Duquesne University, 2006.

[8] William Forde Thompson, *Music, thought, and feeling: Understanding the psychology of music*, Oxford university press, 2015.

[9] Martin Grunwald and Lothar Beyer, *Der bewegte Sinn: Grundlagen und Anwendungen zur haptischen Wahrnehmung*, Springer-Verlag, 2013.

[10] Richard D Sontheimer, "Skin is not the largest organ," *The Journal of investigative dermatology*, vol. 134, no. 2, pp. 581, 2014.

[11] Ronald T Verrillo, Anthony J Fraioli, and Robert L Smith, "Sensation magnitude of vibrotactile stimuli," *Perception & Psychophysics*, vol. 6, no. 6, pp. 366–372, 1969.

[12] ISO/TC 43 Committee, "ISO 226:2003: Acoustics - normal equal-loudness-level contours," Standard, International Organization for Standardization, 2003.

[13] Sebastian Merchel and Mehmet Ercan Altinsoy, "Music-induced vibrations in a concert hall and a church," *Archives of Acoustics*, vol. 38, no. 1, pp. 13–18, 2013.

[14] Sebastian Merchel, Anna Schwendicke, and Mehmet Ercan Altinsoy, "Feeling the sound: audio-tactile intensity perception," in *Proceedings of 2nd Polish-German Structured Conference on Acoustics, The 58th Open Seminar on Acoustics (Jurata, Poland)*, 2011.

[15] Lise A Johnson and Charles M Higgins, "A navigation aid for the blind using tactile-visual sensory substitution," in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2006, pp. 6289–6292.

[16] Surapol Vorapatratorn and Kowit Nambunmee, "iSonar: an obstacle warning device for the totally blind," *Journal of Assistive, Rehabilitative & Therapeutic Technologies*, vol. 2, no. 1, pp. 23114, 2014.

[17] Mohamad A Eid and Hussein Al Osman, "Affective haptics: Current research and future directions," *IEEE Access*, vol. 4, pp. 26–40, 2015.

[18] Karla Allan, Timothy White, Lynette Jones, James Merlo, Ellen Haas, Gary Zets, and Angus Rupert, "Getting the buzz: What's next for tactile information delivery?," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. SAGE Publications Sage CA: Los Angeles, CA, 2010, vol. 54, pp. 1331–1334.

[19] Tor Nørretranders, *The User Illusion: Cutting consciousness down to size.*, Viking, 1991.

[20] Manfred Zimmermann, "The nervous system in the context of information theory," in *Human physiology*, pp. 166–173. Springer, 1989.

[21] Richard Epworth, *Bottleneck: Our Human interface with Reality*, Goforich Publications, 2013.

[22] Barry E Stein and M Alex Meredith, *The merging of the senses.*, The MIT Press, 1993.

[23] Veikko Jousmäki and Riitta Hari, "Parchment-skin illusion: sound-biased touch," *Current biology*, vol. 8, no. 6, pp. R190–R191, 1998.

[24] Helge Gillmeister and Martin Eimer, "Tactile enhancement of auditory detection and perceived loudness," *Brain research*, vol. 1160, pp. 58–68, 2007.

[25] Steve Guest, Caroline Catmur, Donna Lloyd, and Charles Spence, "Audiotactile interactions in roughness perception," *Experimental Brain Research*, vol. 146, no. 2, pp. 161–171, 2002.

[26] Jeffrey M Yau, Jonathon B Olenczak, John F Dammann, and Sliman J Bensmaia, "Temporal frequency channels are linked across audition and touch," *Current biology*, vol. 19, no. 7, pp. 561–566, 2009.

[27] Ana Tajadura-Jiménez, Aleksander Väljamäe, Iwaki Toshima, Toshitaka Kimura, Manos Tsakiris, and Norimichi Kitagawa, "Action sounds recalibrate perceived tactile distance," *Current Biology*, vol. 22, no. 13, pp. R516–R517, 2012.

[28] Ernst Heinrich Weber, *Die Lehre vom Tastsinne und Gemeingefühle auf Versuche gegründet*, Friedrich Vieweg und Sohn, 1851.

[29] Gabriel Robles-De-La-Torre and Vincent Hayward, "Force can overcome object geometry in the perception of shape through active touch," *Nature*, vol. 412, no. 6845, pp. 445, 2001.

[30] David Birnbaum and Marcelo M Wanderley, "A systematic approach to musical vibrotactile feedback.," in *ICMC*, 2007.

[31] Stanley J Bolanowski Jr, George A Gescheider, Ronald T Verrillo, and Christin M Checkosky, "Four channels mediate the mechanical aspects of touch," *The Journal of the Acoustical society of America*, vol. 84, no. 5, pp. 1680–1694, 1988.

[32] Robert F Schmidt, Florian Lang, and Manfred Heckmann, *Physiologie des menschen: mit pathophysiologie*, Springer-Verlag, 2007.

[33] Rolf-Detlef Treede, "Das somatosensorische system," in *Physiologie des Menschen*, pp. 272–297. Springer, 2010.

[34] Robert H. Gault, "Progress in experiments on tactual interpretation of oral speech.," *The Journal of Abnormal Psychology and Social Psychology*, vol. 19, no. 2, pp. 155–159, 1924.

[35] Hiroaki Yano, Hiromi Igawa, Toshihiro Kameda, Koichi Muzutani, and Hiroo Iwata, "Audiohaptics: audio and haptic rendering based on a physical model," in *12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2004. HAPTICS'04. Proceedings.* IEEE, 2004, pp. 250–257.

[36] Derek DiFilippo and Dinesh K Pai, "The AHI: An audio and haptic interface for contact interactions," in *ACM SIGGRAPH 2005 Courses.* ACM, 2005, p. 164.

[37] Rolf Nordahl, Stefania Serafin, Niels C Nilsson, and Luca Turchet, "Enhancing realism in virtual environments by simulating the audio-haptic sensation of walking on ground surfaces," in *2012 IEEE Virtual Reality Workshops (VRW).* IEEE, 2012, pp. 73–74.

[38] Kenneth Salisbury, Francois Conti, and Federico Barbagli, "Haptic rendering: introductory concepts," *IEEE computer graphics and applications*, vol. 24, no. 2, pp. 24–32, 2004.

[39] Corentin Bernard, Jocelyn Monnoyer, Sebastien Denjean, Michael Wiertlewski, and Sølvi Ystad, "Sound and texture synthesizer," in *International Workshop on Haptic and Audio Interaction Design (HAID)*, 2019.

[40] David J Meyer, Michael A Peshkin, and J Edward Colgate, "Fingertip friction modulation due to electrostatic attraction," in *2013 world haptics conference (WHC).* IEEE, 2013, pp. 43–48.

[41] Laura Winfield, John Glassmire, J Edward Colgate, and Michael Peshkin, "T-pad: Tactile pattern display through variable friction reduction," in *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07)*. IEEE, 2007, pp. 421–426.

[42] Yasuaki Monnai, Keisuke Hasegawa, Masahiro Fujiwara, Kazuma Yoshino, Seki Inoue, and Hiroyuki Shinoda, "Haptomime: mid-air haptic interaction with a floating virtual screen," in *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 2014, pp. 663–667.

[43] Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, and Sriram Subramanian, "Ultrahaptics: multi-point mid-air haptic feedback for touch surfaces," in *Proceedings of the 26th annual ACM symposium on User interface software and technology*. ACM, 2013, pp. 505–514.

[44] Adrien Merer, Mitsuko Aramaki, Sølvi Ystad, and Richard Kronland-Martinet, "Perceptual characterization of motion evoked by sounds for synthesis control purposes," *ACM Trans. Appl. Percept.*, vol. 10, pp. 1–24, 2013.

[45] Pierre Schaeffer, *Traité des objets musicaux: essai interdisciplines*, Editions du Seuil, Paris, 1966, English edition: Schaeffer, P. (2017). *Treatise on musical objects: an essay across disciplines* (trans: North, C., Dack, J.; University of California Press, Oakland).

[46] Michel Chion, "Guide to sound objects: Pierre schaeffer and musical research," *Trans. John Dack and Christine North), http://www.ears.dmu.ac.uk*, 2009.

[47] Alexander Lerch, *An introduction to audio content analysis: Applications in signal processing and music informatics*, Wiley-IEEE Press, 2012.

[48] Michael Kateley Klingbeil, "Spectral analysis, editing, and resynthesis: Methods and applications," *Unpublished doctoral dissertation, Columbia University*, 2009.

[49] Xavier Serra and Julius Smith, "Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition," *Computer Music Journal*, vol. 14, no. 4, pp. 12–24, 1990.

[50] Derry Fitzgerald, "Harmonic/percussive source separation using median filtering," in *13th International Conference on Digital Audio Effects (DAFX10)*. Graz, Austria, 2010.

[51] David H Douglas and Thomas K Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature," *Cartographica: the international journal for geographic information and geovisualization*, vol. 10, no. 2, pp. 112–122, 1973.

[52] Charles F Jekel and Gerhard Venter, "pwlf: A python library for fitting 1d continuous piecewise linear functions," https://github.com/cjekel/piecewise_linear_fit_py, 2019.

[53] D Brandon Lloyd, Nikunj Raghuvanshi, and Naga K Govindaraju, "Sound synthesis for impact sounds in video games," in *Symposium on Interactive 3D Graphics and Games*. ACM, 2011, pp. PAGE–7.

[54] Vijaya Lakshmi Guruswamy, Jochen Lang, and Won-Sook Lee, "Iir filter models of haptic vibration textures," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 1, pp. 93–103, 2010.

[55] Rahul Chaudhari, Burak Çizmeci, Katherine J Kuchenbecker, Seungmoon Choi, and Eckehard Steinbach, "Low bitrate source-filter model based compression of vibrotactile texture signals in haptic teleoperation," in *Proceedings of the 20th ACM international conference on Multimedia*. ACM, 2012, pp. 409–418.

[56] Dennis Allerkamp, Guido Böttcher, Franz-Erich Wolter, Alan C Brady, Jianguo Qu, and Ian R Summers, "A vibrotactile approach to tactile rendering," *The Visual Computer*, vol. 23, no. 2, pp. 97–108, 2007.

[57] Charalampos Saitis, Claudia Fritz, Gary P Scavone, Catherine Guastavino, and Danièle Dubois, "Perceptual evaluation of violins: A psycholinguistic analysis of preference verbal descriptions by experienced musicians," *The Journal of the Acoustical Society of America*, vol. 141, no. 4, pp. 2746–2757, 2017.

[58] Daniele Schön, Sølvi Ystad, Richard Kronland-Martinet, and Mireille Besson, "The evocative power of sounds: Conceptual priming between words and nonverbal sounds," *J. Cogn. Neurosci.*, vol. 22, pp. 1026–1035, 2009.