EUROREGIO
BNAM2022
Joint Acoustics Conference
9th-11th of May, Aalborg, Denmark

A comparison of audio-to-tactile conversion algorithms for melody recognition

Razvan Paisa[1,*], **Jesper Andersen**[2], **Niels Christian Nilsson**[1], **Stefania Serafin**[1]

[1]Multisensory Experience Lab, Aalborg University, Copenhagen, Denmark
[2]The Royal Danish Academy of Music, Copenhagen, Denmark
[*]rpa@create.aau.dk

Abstract

Besides language, music is one of the two major acoustic channels for expression of human nature and it is ubiquitous to all cultures. Due to a tight correlation between auditory and haptic stimuli, more and more attention is focused on the importance of the latter sensation in a musical context [1]. For the hearing impaired especially, tactile feedback has been investigated extensively for it's musical applications and hearing assistive devices, as early as 1983 [2]. This study compares three common audio-to-haptic signal processing algorithms designed for full range vibrotactile transducers used for tactile augmentation of music. The focus is on melody discrimination over three instruments: double bass, digital subtractive synthesizer with a sawtooth oscillator and trumpet. The transducer used is a high fidelity Tactuator BM1C, enclosed in a custom anatomical handheld case, inspired by an orthopedic *resting hand splint*. An evaluation was conducted on 34 participants and used a within-group design with three alternative forced choice task assessing the participants ability to match melodies to tactile stimuli. The results indicate that no algorithm performs better than others, which is in line with the literature regarding the overall poor frequency discrimination of the skin. Nevertheless, post experiment interviews suggest that some participants perceived multiple frequencies simultaneously, on different areas of their hand, similar to auditory polyphony.

Keywords: Vibrotactile music, Vibrotactile discrimination, Vibrotactile display, Hearing impaired music, Tactile music perception

1 Introduction

In recent years haptic feedback has received increasing attention from the sound and music community, mainly because of the strong connection between the auditory and haptic experiences. This has given birth to *musical haptics* field of research [1]. The mechanism linking auditory and tactile sensations is called multisensory integration, pioneered by Barry Stein and Alex Meredith. It describes how humans form coherent, valid, and robust perception of reality, by processing sensory stimuli from various modalities [3]. The classical rules for multisensory integration demand that enhancement occurs only for stimuli that are temporally coincident and propose that enhancement is strongest for those stimuli that individually are least effective[3]. For this integration to occur, the input from various sensors must eventually converge on the same neurons. In the specific case of auditory-somatosensory stimuli, recent studies demonstrate that multisensory integration can in fact occur at very early stages of cognition, resulting in supra-additive integration of touch and hearing [4, 5]. This translates to a robust synergy between the two sensory apparatuses, than can be exploited to synthesize experiences impossible to achieve by unisensory means. Furthermore, research within auditory-tactile interactions has shown that tactile stimulus can influence auditory stimulus and vice-versa [6, 7, 8]. It can therefore be observed that auditory and haptic stimuli are capable of modifying or altering the perception of each other when presented in unison [9].

This study is the first in a project that has as long term goal to help partially impaired hearing individuals and cochlear implant users to appreciate music. With that in mind, the aim of this particular study is to compare three signal processing methods that convert full spectrum music into vibrotactile haptic feedback suitable for the proprieties of skin receptors, namely the Meissner's corpuscles and Pacinian ones, while preserving the melodic information encoded in the original signal. The three processing methods were chosen from existing literature [10, 11, 12, 13, 14, 15]. The experiment revolved around a handheld device designed to be comfortable to hold for longer periods of time, and capable of reproducing full spectrum audio signal. The hand was identified as the most sensitive body region for touch, due to a very high density of receptors[16, 17].

This paper describes the device built for the study followed by a detailed presentation of each signal processing technique used to convert music to vibrotactile stimuli. Subsequently, it is presented the experimental study evaluating user performance when tasked to match the haptic stimuli to a coherent auditory one. The aim of the study was twofold: (1) to evaluate the three signal processing methods in terms of their ability to convey the melodic structure existing in the original signal.(2) To evaluate the proposed hardware in terms of it's ergonomics and ease of use, as well as it's ability to produce a satisfying haptic experience. Specifically it was considered relevant to determine if a satisfactory experience can be elicited with a single, high-fidelity actuator.

2 Background

Live concerts, especially amplified ones, as well as movie scores are know to create haptic sensations coupled with the sound, conveying valuable information such as articulation and timing. Several studies have tried to replicate and quantify this phenomena with compelling results [18, 17, 19, 20].

Merchel approached the topic from an architectural acoustics point of view, aiming to prove that concert halls with a strong haptic feedback improve the overall quality of the concert experience [19]. His studies proposes several signal processing techniques to be used for the haptic channel, indicating that in music with a rich low end, the audio signal passed through a low pass filter is enough to improve the experience. Furthermore, he suggests that simple sinusoidals with frequencies not related to the audio signal will produce an enhanced listening experience, but the frequency of these haptics oscillators will have an impact in the overall perception [19].

Other group of authors suggested to account for haptic feedback at the composition stage, creating a coherent audio-haptic experience, instead of approaching haptics as an afterthought [21, 20]. Gunther and O'Modhrain coined the term *tactile composition* as a *system that facilitates the composition and permeation of intricate, musically structured spatio-temporal patterns on the surface of the body*, emphasizing the importance of a compositional language for the sense of touch[20]. Their 2001 *Concerts for the skin* experiments surface some important notions like selective haptic attention - the ability to selectively direct attention into different stimuli, if several body areas are actuated at the same time [20]. On top of that they suggest that the music-haptic relationship does not need to produce congruent stimuli at all times, and the composer should engage into a parallel multimodal composition that inter-plays between the two sensory channels.

Listening for pitch is almost always dependant on the frequency of the audio content, while the timbre and amplitude rarely have an impact on pitch perception [22]. In contrast, the perception of frequency from a vibrotactile stimuli is more complicated due to the multi-channel nature of the cutaneous sensing organ - the skin [22]. Moreover, perception of tactile frequency is amplitude and time dependant, and it varies significantly depending on the position on the body. Nevertheless, there is one important similarity between auditory and tactile pitch perception: within certain frequencies, the discrimination fits a critical band model [23]. Specifically, certain frequency ranges are perceived as distinct sensations, indicating that with enough exposure, tactile pitch perception can be interpreted similarly to the auditory one - a fact proven by many hearing impaired people [24]. This does not mean that understanding music through vibrotactile stimuli is equivalent to hearing it, but the experience, while different, could be just as enjoyable.

Music usually uses a wider frequency spectrum than the skin can provide, and the tactile pitch-amplitude coupling only makes understanding music without hearing it more complicated. Unlike the ear, with its single receptor capable of 20Hz-20000Hz frequency range perception, the skin has multiple types of receptors, each with its own frequency and temporal characteristics. For music perception, two of them prove to be useful: the Meissner corpuscles and the Pacinian receptors [25]. The Meissner corpuscles, also known as Rapid Adapting (RA) receptors, have a very high innervation density and have a limited frequency range of 10Hz-100Hz, with a peak sensitivity around 40Hz. The Pacinian receptors are larger than the RA ones, have a low spatial resolution, and a frequency response between 40Hz and 1000Hz, and are most sensitive around 250Hz [25]. In an attempt to describe the tactile music properties, Erp & Spapé conducted an experiment on the perceptual attributes of vibrotactile melodies [26]. Their results indicate that users can perceive and evaluate multiple characteristics from the tactile stimuli(f.ex. aggressive, soft, alarming, bombastic, etc) and that melodies generally land in one of four clusters, on a two dimensional tempo-intrusiveness map [26]. In a similar fashion, Ternes & MacLean designed a large set of distinguishable tactile rhythms, further highlighting the potential of tactile melodies [27].

3 Implementation

3.1. Hardware

The hardware device is an ovoid shape with the following dimensions: 84mm wide, 58mm tall and 89mm deep and can be seen in figure 1.

Figure 1: Side and front view of the haptic device

The shape was inspired by the resting hand position when fixed with an orthopedic splint. This pose should minimize the strain on the wrist, and allow the fingers to relax in their natural rest position. The initial shape was created using modelling clay, aiming to ensure the finger position is anatomic, each digit having its own socket. The clay artefact was 3d scanned using Autodesk ReCap[1], by analyzing 40 still images of the subject, taken from multiple angles with a Fujifilm X-T1 camera and a Fujinon XF 35mm @ f2.0 lens. The artefact was suspended in midair with fishing line, affording visibility from all angles. The 3D scan resulted in a very high fidelity digital model, but in order to improve topology, a new 3d model was created using the scan as an outline.

---

[1] https://www.autodesk.com/products/recap/

The final shape was split in half horizontally, to have access inside where the electronics would eventually lie. The two halves were held together by three M3x16 bolts that have been incorporated in the design show in figure 1.

For the actuator, a socked was created on the bottom half of the device, and a 3.5mm female jack opening has been installed on top half to connect the transducer to the amplifier. The jack was oriented towards the left side of the device, allowing for cable connection that should not interfere with the user while holding it. The haptic device halves were fabricated with 2mm wall thickness using an Ultimaker 3 and PLA material.

The device was for left hand only, as it was intended to have the users navigate the experiment's questionnaire with the computer mouse, which is generally used with the right hand. Initial informal tests showed that people unfamiliar with the device tend to hold it in unintended ways, thus for the experiment finger positioning visual signifies have been painted on. When it comes to the transducer, a Tactuator BM1C vibrotactile actuator manufactured by Tactile Labs[2] was used. Haptuator Mk1 and Mk2 were also tried, but the Tactuator BM1C proved to have the highest amplitude in the current setup, and the distortion (if any) was non disruptive, as would be the case with Mk2 and M1 that rattle rather loud when overdriven. All transducers tested offer full spectrum reproduction. The tactuator requires amplification to achieve desirable amplitude therefore a high gain Behringer HA8000 headphone mixing and distribution amplifier was used.

3.2. Tactile signal processing

In an attempt to improve the perception of pitch through tactile sensing, three signal processing methods that convert arbitrary auditory signal into a tactile one were compared. Each of the processing methods was inspired from existing literature, and was re-implemented to exploited one physical or perceptual trait relevant for music listening.

Method 1: Compression of frequency spectrum

The first method focused on compressing the musically relevant frequency spectrum defined between 40Hz and 2093 Hz into a narrower "tactile range" one up to 1046 Hz, to address the Pacinian receptors exclusively, since these are the most sensitive to vibrotactile stimuli[22, 25]. The lower limit represents the crossover between RA receptors and Pacinian receptors, and the high frequency represents the top range of the Pacinian ones. The frequency compression was implemented as seen in Figure 2 as following:

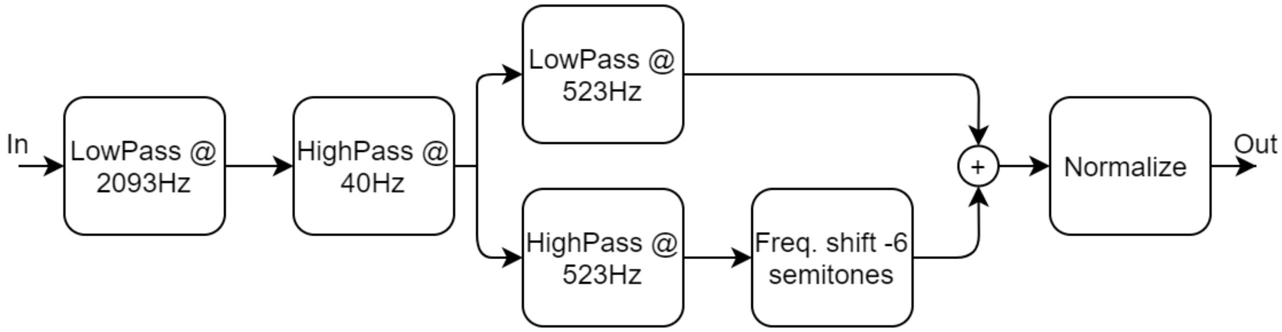

Figure 2: Signal processing for first condition

1. Apply a lowpass FIR filter with 60 dB/octave attenuation at 2093Hz - the corresponding frequency of the fundamental for a C7 note with A4 = 440 Hz tuning. This meant that only the highest octave available on a piano was ignored. Nevertheless, that the majority of instruments, including the human voice, have the high limit considerably lower than C7, Since the upper harmonics are not contributing much to melody perception, the frequency limitation was not consider to be a practical problem [28].

---

[2] www.tactilelabs.com

2. Apply a highpass FIR filter with 60 dB/octave attenuation filter at 40Hz to limit the actuation of RA receptors.
3. Split the frequency band at 523 Hz (C5 note) in two spectra, using a lowpass and a high pass filter. The lower one (40Hz - 523Hz) will be called Spectrum A, and the higher on B. The C5 note was chosen in relationship to the fundamental frequencies of the melodies used, and described in 3.
4. Pitch shift down spectrum B 6 semitones to shift the high frequency content into the tactile sensible range
5. Add the Spectrum A (original) and Spectrum B (pitch shifted)
6. Normalize to 1 to avoid clipping, and ensure equal amplitude throughout the melodies selection.

Method 2: Sinusoidal oscillators

The second method focused on Pacinian receptors as well, and it used sinusoidal with the frequency equal to the fundamental one of the actual tone, instead of the original signal as suggested by Merchel [10, 13, 14, 19]. This was done with the aim of avoiding higher frequency content from masking or diminishing the fundamental harmonic perception, since tactile spectral masking works similar to auditory one [23]. The sinusoidals were generated using the same MIDI information as the auditory signal, using Xfer Serum[3] wavetable synthesizer with *Basic Shapes* table, on position one and a square envelope(0 attack, max sustain, 0 decay). In order to ensure amplitude coherence between the auditory and tactile stimuli, the contour/envelope was extracted from the original file and applied to the haptic one. The last step was to apply normalization, similar to method 1.

Method 3: Tactile transient reinforcement

The last tactile signal processing tried to make use of both the RA and Pacinian receptors. The tactile signal combined the auditory stimuli with a haptic reinforcement one, aimed at the RA receptors in order to emphasize changes in pitch, practically working as an exciter or *transient emphasizer*. This feedback approach was inspired by the way frets provide guitar players feedback about the note selection, as described it [29]. The haptic signal was created by adding a haptic reinforcement component, to the signal generated with method 2. The haptic reinforcement was generated similarly to the sinusoidal described above, but 3 octaves lower than the auditory signal. This meant that the frequencies lied in the peak frequency response of the RA receptors [25]. An attack-decay (AD) envelope was used for the haptic reinforcement signal, with 10ms attack time, in order to reduce artefacts(clicks), and 500ms logarithmic decay time to avoid temporal masking over the higher frequency signal. The two signals were summed with amplitudes of 0.8 for the haptic reinforcement, and 0.2 for the original, unprocessed signal, followed by normalization. The large difference in volume between the two signals is due to the lower amplitude response of the Tactuator BM1C below 40Hz.

All processing was done in Matlab unless specified otherwise. Highpass and lowpass filters had a steepness of 0.8 (default in Matlab). The amplitude contour was computed as the moving RMS envelope of the unprocessed melody every 5000 samples, in order to avoid artefacts introduced by abrupt changes in loudness. Sampling frequency used was 48kHz, and all files were exported uncompressed (wav).

3.3. Melodies

The 75 musical melodies composed for the project, were all of a duration of three to eight seconds and spread across a randomized selection of different major- and minor keys; figure 3 shows the distribution of notes across all melodic phrases. The melodies were simple and kept in a melodic style easily recognizable for listeners familiar with western music. They all represented a small musical progression with a beginning and an end. Tempo was 120bpm and, rhythmically, there were a mixture of whole- quarter- eighth and sixteenth- notes. For each of the 75 true melodies, two false were added. The false melodies always had the same rhythmic content as the true one, but at least 75 percent of the tones were changed. In the false melodies, the musical progression would not be perceived as natural, since the selection of notes were not following western melodic tradition. Figure 4 shows the average number of semitones deviating from the correct melodies.

---

[3] https://xferrecords.com/

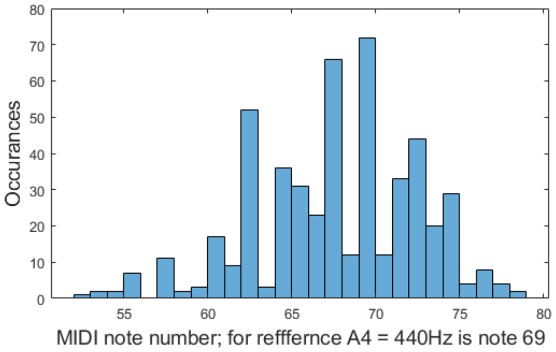

Figure 3: Distribution of notes in the correct melodies

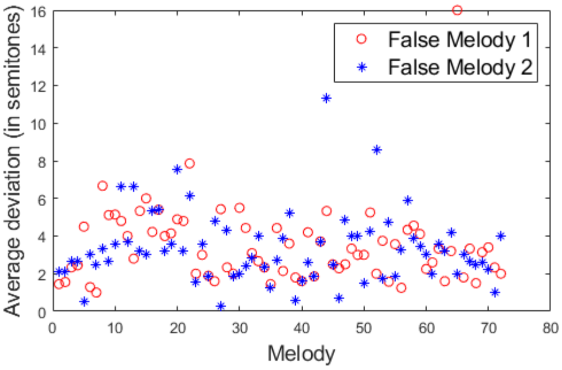

Figure 4: False melodies deviation from the correct ones

4 Evaluation

The aim of the study was to (1) evaluate three signal processing techniques for converting audio material to tactile stimuli, and (2) to evaluate the usability of the haptic device itself. To meet this aim a within-subjects study was performed, comparing four conditions that varied in term of the tactile feedback provided through the haptic device, when listening to melodic phrases.
The hypothesis was: *There is a difference in terms of tactile melody discrimination between an unprocessed signal and a processed one when presenting congruent bi-modal melodic phrases through a single-actuator handheld vibrotactile device and headphones.*

4.1. Task and Stimuli

The task for the participant was to select the haptic feedback that matched the auditory signal played through the headphones. A three alternative forced choice design was used, with only one correct option; for each trial the participants were presented with three types of haptic feedback. The experiment had four conditions, with different singal processing techniques described in section 3: [1]Control condition with no processing, [2]Frequency compression, [3]Sine wave at the fundamental frequency and [4]Tactile reinforcement of transients.

There were 72 trials in total: 18 for each processing technique, plus 18 for unprocessed acting as control condition. Melodies where chosen and presented randomly out of pool of 75 possibilities, distributed equally among the three instruments. A total of 900 possible trials were used for the whole experiment: 75 melodic lines * 4 conditions * 3 instruments, ensuring a high level of validity. For each instrument, condition and melodic line there was one correct haptic stimulus, and two incorrect ones. The order for melodies, order of conditions and choice of instruments were assigned randomly, in real time, for each participant. In order to ensure similar exposure levels for all participants, they were allowed to experience each stimuli/melody combination only once. The experiment took place in the Multisensory Experience Lab at Aalborg University campus in Copenhagen.

4.2. Participants

Participation in the experiment was voluntary and the majority of participants were students of *Sound and Music Computing* and *Medialogy* programs, that are affiliated with the Multisensory Experience Lab, and the Tonmeister program from the Royal Danish Academy of Music. Some participants had musical experience, but this was not a selection criteria. The participants have been encouraged to partake in the experiment at their convenient time and there was no reward for doing it. The data collected has been anonymous, without the possibility of matching answers sets with the participant. There were a total of 34 participants (24 male, 10 female). Although the ultimate goal is to create a device for hearing impaired users, current COVID-19 restrictions prevented us to test on the relevant target group. The experiment was conducted with 3 participants in parallel, in 3 different rooms, that were briefed and debriefed together, by the first author.

4.3. Setup and Equipment

The hardware setup consisted of a Windows computer running the experiment application with a Behringer HA8000 headphone distribution amplifier connected to it. The left audio channel contained the haptic melodies, and the right channel carried the auditory signal. The distribution amplifier routed the auditory signal to both headphones channels. The headphones used were different due to availability, but had similar price and quality level: *Creative Aurvana Live!, AKG K240 and Sennheiser HD240 Pro*. The level balance between the haptic and auditory signal was set by the second author, and calibrated to have a natural balance between the two sensory inputs to assure that not one would overpower the other. First the headphones were adjusted to a comfortable playback-level and then the haptics were added up close to the distortion limit of the transducer. The mix was constant for all participants.

4.4. Procedure

As mentioned, three participants partook in the experiment at the same time. They were welcomed and introduced to the task, emphasizing that it is not required from them to over-analyze the vibrotactile stimuli, but instead they should answer based on their intuition. The participants were then guided to the setup rooms and required to experience Queen's ”Don't stop me now”, as training and accommodation with the system. The song was chosen due to its popularly, but also because it features many combinations of instruments and intensities: from low intensity voice only, to high intensity full band playing. This should provide the users with most of the possible stimuli expected throughout the experiment. After the accommodation phase was finished, the users were required to click on ”Start” button to begin the experiment. Each trial consisted of listening to the same auditory melody three times, with different haptic stimuli for each as described in 2. A visual indicator was signaling what exposure was playing at all times, and the ”Select the haptic stimuli that matches best the melodic phrase you heard” message was permanently displayed in the bottom of the page, followed by the trial number. There was a 2 seconds gap between exposures within the same trial. Since all potential melodies were fairly similar in length, the experiment was completed in 18-19 minutes. After all 3 participants in a series finished, they were gathered for a post-experiment debriefing discussing about their experience, comfort, amplitude and potential suggestions. The setup, similar to an ad-hoc focus group, facilitated interesting discussions between participants, as well as between conductors and participants.

4.5. Data Collection

The data collected has been anonymous, without the possibility of matching answers sets with the participant. The following information was logged: trial number(1-72), melody number(1-72), condition(1-4), instrument(1-3), correct answer, user answer, and inevitably, the log file creation time.

5 Results

The data collected from the 34 participants was treated as nominal and was analyzed using Friedman tests. The main test was run to determine if there were significant differences between the three proposed signal processing techniques with respect to number of correct identification of the matching haptic stimuli. In addition to this analysis, the data was analysed on a per-instrument basis, as well as instrument performances against each-other. The number of correct answers were not statistically significantly different among conditions $\chi^2(3) = 0.885, p = .829$. Similarly, no significant differences were found when the instruments were analyzed independently: bass trials ($\chi^2(3) = 2.590, p = .459$), synth trials ($\chi^2(3) = 4.528, p = .210$), and trumpet trials ($\chi^2(3) = 1.401, p = .701$).
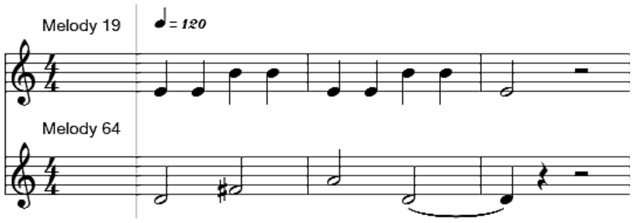
Figure 5: Best(19) and worst(64) performing melodies

No effect on the number of correct answer was observed, but while inspecting the descriptive statistics, it seemed that one instrument (synthesizer) stands out therefore an exploratory analysis was ran, comparing the sets of trials for each instrument against each other. The results are $\chi^2(2) = 1.746, p = .418$. Furthermore, when looking at the best and worst performing three melodies in terms of correct answers, it was discovered that the ones with multiple short notes had a slightly higher average number of correct answers, while the ones with longer, sustained notes had a lower number of correct answers, even when harmonic content is very similar. The correlation between average note length in the melody and number of correct answers is ρ = -0.19 with p = 0.1. Figure 5 show the best performing melody with 14 correct answers out of 34 (41.1%) and worst performing one with 2 correct answers (5.8%). Lastly, there was no correlation found between the average deviation from correct melody as show in 4 and the number of correct notes: $\rho < 0.07$.

5.1. Post-experiment interview

The post experiment interview highlighted some interesting facts regarding the physical design, the experiment as well as potential directions for further experimentation. Several participants remarked that the experiment is too long and repetitive, loosing focus towards the end. Regarding the physical properties, some subjects reported that they experimented with different arm resting positions (arm resting on the knee facing up/down, arm resting on the table, crossed arms) noticing that each position will produce slightly different results. Out of those who mentioned position, there seemed to be a consensus that palm facing up feels the best, with one mention that it felt stronger. Probably the most interesting feedback was that some participants felt different frequencies in different areas of the hand, one participant mentioning that sometimes it could sense two frequencies at the same time. This has been expressed in various forms, some claiming that higher frequencies feel too strong, especially for the fingers, but the lower frequencies feel good.

Regarding the hardware, the feedback has been generally good, but some participants suggested that the device was either too big or too small for their hands. At the same time, few participants reported that it is a slightly uncomfortable to hold for long time while most mentioned it was comfortable.

6 Discussion

The results related to *number of correct answers* suggest that the proposed processing techniques do not result in statistically significant different performances. Furthermore, looking at the nature of the haptic stimuli, it is observed that the *bass, synthesizer and trumpet* perform similarly among the 4 conditions. Even though the performed tests does not permit us to conclude that the results are statistically equivalent, the descriptive statistics does seem to indicate that the effect of processing methods was negligible, and possibly non-existing. This can be seen in figure 6 showing the median for all conditions is 1 and the variance is consistent, regardless of the instrument presented. A potential explanation for this similarity can be found in the fact that a higher number of shorter notes are easier to recognize, opposite to longer, sustain ones as seen in Figure 5. A similar behavior was observed by Tommerdahl et. al in 2005 in his studies on the vibrotactile discrimination capacity of skin for various stimuli lengths, concluding that the cerebral cortex undergoes a profound inhibition withing 1-2 seconds from the start of a 200Hz stimuli [30]. That being said, this study did not investigate the impact of legato or staccato as all harmonic events were rhythmically independent, therefore no conclusion can be reached

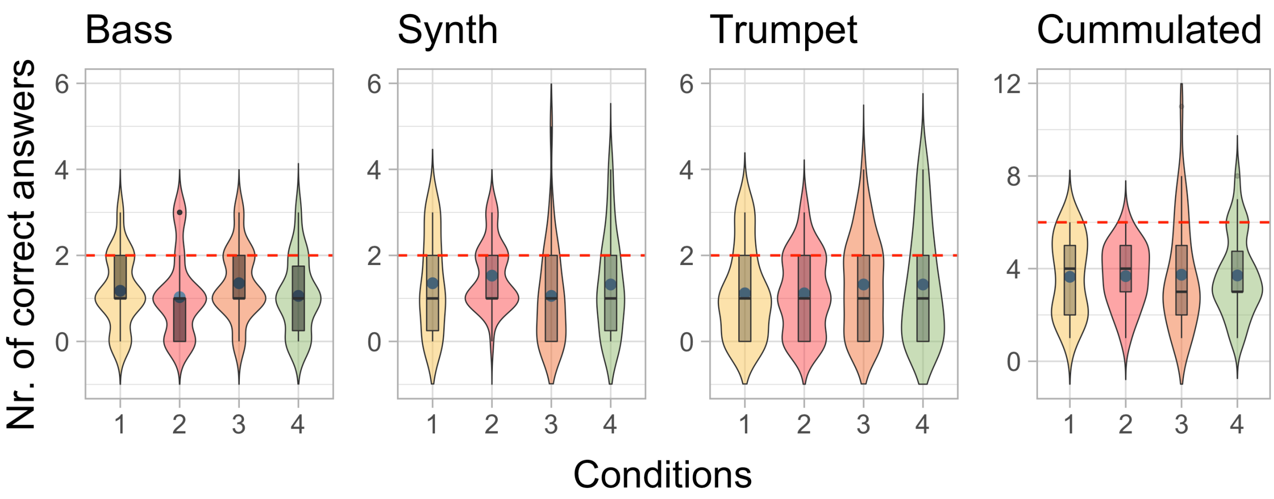

Figure 6: Distribution of number correct answers (max. possible = 6/instrument/condition),for each instrument, for each of the four conditions: 1 = control condition; 2 = frequency compression; 3 = sine wave at the fundamental frequency; 4 = haptic reinforcing of transients. Box plots inside the violin plots represent interquartile range, blue circles represent the mean and red line represents the expected chance level.

on whether the rapid melodic changes or the short duration explain this phenomena. Nonetheless, if condition 1 - control and condition 3 - sinusoidal oscillators, are analysed in isolation, the results do not align with the findings of Merchel, that suggest using sine signals with the frequency matching the fundamental one from the auditory signal produces a better tactile experience - at least in terms of melodic content identification [19].

When it comes to the comfort and performance of the physical device, the results are mixed. Some users claimed it was comfortable and provided appropriately strong vibrotactile stimuli, while other complained that it can be too strong at times or that it becomes uncomfortable to use for longer periods of time. These findings indicate a preferences for individual customization of device size as well as control over the haptic intensity.

Lastly, an interesting phenomena was describe by several users, claiming that different frequencies are felt in different areas of the hand. This is a direction worth exploring further, since it can indicate that single actuator devices can address different areas of the hand, providing an extra dimension for communication.

7 Conclusion

In this paper it was proposed a system that allows musical signals to be converted to vibrotactile stimuli. The system was evaluated in a user experiment exploring impact of three signal processing techniques used for audio to haptic conversion, in terms of user's ability to identify the melodic information. The stimuli used for the experiment was composed of short melodies played on double bass, synthesizer and trumpet. The results indicated that there was no significant difference between the 3 proposed techniques and no processing at all, when it comes to melody identification, underlining the need for new algorithms that can be empirically validated. However, there was an indication that users do perform better at the identification task when the haptic stimuli contains shorter notes, regardless of processing algorithm or instrument played. Finally, it was surfaced that different areas of the hand can sense separate frequencies, but further research needs to be conducted in order to fully understand the phenomena.

Acknowledgements

This work is supported by NordForsk's Nordic University Hub - Nordic Sound and Music Computing Network.

References

[1] Stefano Papetti and Charalampos Saitis. *Musical Haptics: Introduction*. 2018. doi: 10.1007/978-3-319-58316-7_1.

[2] P. L. Brooks and B. J. Frost. Evaluation of a tactile vocoder for word recognition. *The Journal of the Acoustical Society of America*, 74, 1983. ISSN 0001-4966. doi: 10.1121/1.389685.

[3] B.E. Stein, P.C.B.E. Stein, and M.A. Meredith. *The Merging of the Senses*. A Bradford book. ISBN 9780262193313.

[4] John Foxe, Glenn Wylie, Antigona Martinez, Charles Schroeder, Daniel Javitt, David Guilfoyle, Walter Ritter, and Micah Murray. Auditory-somatosensory multisensory processing in auditory association cortex: An fmri study. *Journal of neurophysiology*, 88:540–3, 08 2002. doi: 10.1152/jn.00694.2001.

[5] Christoph Kayser, Christopher Petkov, Mark Augath, and Nikos Logothetis. Integration of touch and sound in auditory cortex. *Neuron*, 48:373–84, 11 2005. doi: 10.1016/j.neuron.2005.09.018.

[6] Ryuta Okazaki, Hiroyuki Kajimoto, and Vincent Hayward. Vibrotactile stimulation can affect auditory loudness: A pilot study. pages 103–108, 06 2012. doi: 10.1007/978-3-642-31404-9_18.

[7] Ryuta Okazaki, Taku Hachisu, Michi Sato, Shogo Fukushima, Vincent Hayward, and Hiroyuki Kajimoto. Judged consonance of tactile and auditory frequencies. pages 663–666, 04 2013. ISBN 978-1-4799-0087-9. doi: 10.1109/WHC.2013.6548487.

[8] Tony Ro, Johanan Hsu, Nafi Yasar, Caitlin Elmore, and Michael Beauchamp. Sound enhances touch perception. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 195:135–43, 04 2009. doi: 10.1007/s00221-009-1759-8.

[9] Gareth Young, David Murphy, and Jeffrey Weeter. Haptics in music: The effects of vibrotactile stimulus in low frequency auditory difference detection tasks. *IEEE Transactions on Haptics*, PP:1–1, 12 2016. doi: 10.1109/TOH.2016.2646370.

[10] S. Shin, C. Oh, and H. Shin. Tactile tone system: A wearable device to assist accuracy of vocal pitch in cochlear implant users. 2020. ISBN 9781450371032. doi: 10.1145/3373625.3418008.

[11] Mark Fletcher, Nour Thini, and Samuel Perry. Enhanced pitch discrimination for cochlear implant users with a new haptic neuroprosthetic. *Scientific Reports*, 10:1–10, 2020. URL `https://eprints.soton.ac.uk/441275/`.

[12] M.J. Luciá, P. Revuelta, Á. Garciá, B. Ruiz, R. Vergaz, V. Cerdán, and T. Ortiz. Vibrotactile captioning of musical effects in audio-visual media as an alternative for deaf and hard of hearing people: An eeg study. *IEEE Access*, 8:190873–190881, 2020. doi: 10.1109/ACCESS.2020.3032229.

[13] U. Trivedi, R. Alqasemi, and R. Dubey. Wearable musical haptic sleeves for people with hearing impairment. pages 146–151, 2019. ISBN 9781450362320. doi: 10.1145/3316782.3316796.

[14] M.D. Fletcher, S.R. Mills, and T. Goehring. Vibro-tactile enhancement of speech intelligibility in multi-talker noise for simulated cochlear implant listening. *Trends in Hearing*, 22, 2018. doi: 10.1177/2331216518797838.

[15] G.W. Young, D. Murphy, and J. Weeter. Vibrotactile discrimination of pure and complex waveforms. pages 359–362, 2015. ISBN 9780992746629.

[16] ALVAR WILSKA. On the vibrational sensitivity in different regions of the body surface. *Acta Physiologica Scandinavica*, 31, 1954. ISSN 1365201X. doi: 10.1111/j.1748-1716.1954.tb01139.x.

[17] Antonella Mazzoni and Nick Bryan-Kinns. How does it feel like? an exploratory study of a prototype system to convey emotion through haptic wearable devices. 2015. doi: 10.4108/icst.intetain.2015.259625.

[18] Marcello Giordano, John Sullivan, and Marcelo M. Wanderley. *Design of Vibrotactile Feedback and Stimulation for Music Performance*. 2018. doi: 10.1007/978-3-319-58316-7_10.

[19] Sebastian Merchel and M. Ercan Altinsoy. Auditory-tactile music perception. volume 19, 2013. doi: 10.1121/1.4799137.

[20] Eric Gunther and Sile O'Modhrain. Cutaneous grooves: Composing for the sense of touch. *International Journal of Phytoremediation*, 21, 2003. ISSN 15497879. doi: 10.1076/jnmr.32.4.369.18856.

[21] Joanne Armitage and Kia Ng. doi: 10.1007/978-3-319-46282-0_9.

[22] David M. Birnbaum and Marcelo M. Wanderley. A systematic approach to musical vibrotactile feedback. 2007.

[23] JC Makous, RM Friedman, and CJ Vierck. A critical band filter in touch. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 15(4):2808—2818, April 1995. ISSN 0270-6474. doi: 10.1523/jneurosci.15-04-02808.1995.

[24] Lisa Bruns, Dirk Mürbe, and Anja Hahne. Understanding music with cochlear implants. *Scientific Reports*, 6, August 2016.

[25] Ki-Uk Kyung, Minseung Ahn, Dong-Soo Kwon, and Mandayam A. Srinivasan. Perceptual and biomechanical frequency response of human skin implication for design of tactile displays. In *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, WHC '05, page 96–101, USA, 2005. IEEE Computer Society. ISBN 0769523102. doi: 10.1109/WHC.2005.105. URL `https://doi.org/10.1109/WHC.2005.105`.

[26] Jbf Van Erp and Mma Spapé. Distilling the underlying dimensions of tactile melodies. *Proceedings of Eurohaptics 2003*, 2003.

[27] David Ternes and Karon E. MacLean. Designing large sets of haptic icons with rhythm. volume 5024 LNCS, 2008. doi: 10.1007/978-3-540-69057-3_24.

[28] John M. Eargle. *Frequency Ranges of Musical Instruments and the Human Voice*, pages 324–325. Springer US, Boston, MA, 2002. ISBN 978-1-4615-2027-6. doi: 10.1007/978-1-4615-2027-6_156. URL `https://doi.org/10.1007/978-1-4615-2027-6_156`.

[29] Håkon Knutzen, Tellef Kvifte, and Marcelo M. Wanderley. Vibrotactile feedback for an open air music controller. volume 8905, 2014. doi: 10.1007/978-3-319-12976-1_3.

[30] M Tommerdahl, KD Hester, ER Felix, M Hollins, OV Favorov, PM Quibrera, and BL Whitsel. Human vibrotactile frequency discriminative capacity after adaptation to 25 hz or 200 hz stimulation. *Brain research*, 1057(1-2), September 2005. ISSN 0006-8993. doi: 10.1016/j.brainres.2005.04.031.