Spectral Discrimination Thresholds Comparing Audio and Haptics for Complex Stimuli

Lorenzo Picinali¹, Christopher Feakes¹, Davide A. Mauro², and Brian F.G. Katz³

¹ De Montfort University, The Gateway, LE1 9BJ Leicester, UK
² LIM-Dipartimento di Informatica, via Comelico 39/41, 20135 Milano, Italia
³ LIMSI-CNRS, Université Paris-Sud, 91403 Orsay, France

Abstract. Individuals with normal hearing are generally able to discriminate auditory stimuli that have the same fundamental frequency but different spectral content. This study concerns to what extent it is possible to perform the same differentiation considering vibratory tactile stimuli. Three perceptual experiments have been carried out in an attempt to compare discrimination thresholds in terms of spectral differences between auditory and vibratory tactile stimulations. The first test consists of assessing the subject's ability in discriminating between three signals with distinct spectral content. The second test focuses on the measurement of the discrimination threshold between a pure tone signal and a signal composed of two pure tones, varying the amplitude and frequency of the second tone. Finally, in the third test the discrimination threshold is measured between a tone with even harmonic components and a tone with odd ones. The results show that it is indeed possible to discriminate between haptic signals having the same fundamental frequency but different spectral. The threshold of sensitivity for detection is markedly less than for audio stimuli.

1 Introduction

Humans are readily able to distinguish between the aural timbre of common notes played by different instruments (for example between an A3 played on a clarinet and the same note played on a flute). The current study investigates whether or not it is possible to observe a similar ability in terms of vibratory tactile sensations.

Vibratory tactile stimulations are often employed to assist in the creation of virtual objects in a computer simulation, or again to support and facilitate specific tasks within a multimodal interactive application. Haptic vibratory actuators can be found in mobile phones, tablet PCs, etc. and are used to enhance the interactivity of the device, and to transfer selected information to the user (e.g. the arrival of an incoming phone call). Typically the types of information transferred through vibration have been of a *boolean* nature, employing single or a series of *on* or *off* signals (e.g. simple alert for an incoming call). With improvements in tractor response and performance, amplitude and frequency modulation of the vibratory stimulus have started to be used in the form of tactons (see Brown, 2006) and, more in general, of vibratory patterns. Nevertheless, variations in the spectral characteristics of the vibratory signal are rarely employed. It is generally assumed that the human sensitivity to such differences, regarding vibratory tactile stimulation, is not particularly high.

C. Magnusson, D. Szymczak, and S. Brewster (Eds.): HAID 2012, LNCS 7468, pp. 131–140, 2012. © Springer-Verlag Berlin Heidelberg 2012

While studies investigating the ability of the human tactile system to discriminate between signals with different frequency have been successfully carried out in the past (see Verrillo, 1963 and Verrillo et al., 1992), similar investigations on the discrimination between signals with different spectral characteristics have only recently started to be performed (see Merchel et al., 2010).

The hypothesis at the base of this study is that the human tactile system can indeed discriminate between vibrations with the same fundamental frequency but different spectral characteristics. The objective is therefore to investigate how spectral variations are perceived through tactile vibratory stimulation, and to compare these results with those measured for auditory stimulations.

2 Related Works

It is well known that differences in the frequency content and spectral envelope of acoustic signals are often perceived as timbre variations, allowing for the differentiation between stimuli with the same fundamental frequency, loudness and duration, while having different spectral characteristics (an overview on studies in this field can be found in Moore, 2003, pp. 105-107 and pp. 270-273). But, is it possible to similarly differentiate complex vibratory tactile stimuli?

Previous studies have investigated the perceptual aspects of frequency and amplitude variations in tactile vibratory stimulations. In Verrillo, 1963, the subjective perceived intensity as a function of the vibration frequency has been studied and compared in relation to the contact area. Results showed that for pure tones detection thresholds improved with frequency, from 25 Hz to 200-300 Hz, at a rate of about 12 dB/octave, then decreasing with the same slope to about 1000 Hz. The larger the contact area (up to 5.1 cm^2), the lower the threshold.

In Verrillo et al., 1992, various studies are reported on frequency and amplitude discrimination for tactile stimuli using pure tones, pulses, and narrow-band noise signals. Comparing the frequency discrimination thresholds between audio and tactile modalities highlights that while the ear can discriminate frequency differences of the order of 0.3%, the performances of the skin were found to be much lower, of the order of 30%. In terms of amplitude discrimination, the threshold for vibratory stimulation was found to be between 0.4 and 2.3 dB, values that are very similar to the threshold for auditory stimulation (between 0.5 and 1 dB, as reported in Moore, 2003, pp. 139-139). Furthermore, the tactile system was found to be capable of processing vibrations within a dynamic range of 55 dB, compared to a sensibly larger range for the auditory system (120 dB, see Moore, 2003, pp. 127-128).

In West and Cutkosky, 1997, the possibility of using a haptic device for displaying fine surface features was investigated. Comparative tests were conducted using both a stylus and fingertips, outlining how the perception of sinusoidal features was qualitatively similar when comparing real and virtual walls. A similar study has been presented in Choi and Tan, 2005, where more sophisticated haptic rendering algorithms were used for rendering complex surface textures of different virtual objects.

In Branje, 2010 a sensory substitution system, aimed at translating music into vibrations presented on the human back, was evaluated in a frequency discrimination task. Findings outlined that vibrotactile stimulation can indeed be used for supporting the experience of music even in absence of sound. Furthermore, voice coils were found to be suitable for transmitting certain sound characteristics in the form of tactile vibrations.

The ability to discriminate between different signals and design parameters for the generation of tactile feedback has been investigated in Merchel et al., 2010. In this study, experiments were conducted in an attempt to determine whether one can distinguish different looped audio signals rendered through an electro-dynamic shaker positioned under a touch-sensitive screen. Stimuli differed in their spectral content and rhythmic characteristics. Results outlined that a distinction was indeed possible. Additional studies have examined the advantage of audio, tactile and coupled audio-haptic stimuli (Altinsoy and Merchel, 2009), although there was no direct relation between the audio and tactile stimuli selected.

Basic studies performed using pure and/or very simple tones for quantifying the discrimination thresholds in terms of spectral variations for tactile stimuli have not been found in literature. The use of an experimental protocol which employs the same stimuli design for both audio and haptic modalities is seen as a fundamental study which is currently absent. The current study proposes an evaluation of the discrimination ability and sensitivity to complex haptic stimuli as compared to auditory stimuli as a reference using a common protocol.

3 Preliminary Study Using Audio-haptic Source Differentiation

A recent study by the authors concerned the exploitation of audio and/or haptic cues for the selection of a desired target in a 3D virtual environment containing multiple distractors (see Menelas et al., 2010). In order to promote coherent identification of audio-haptic targets, four haptic signatures were designed using waveform amplitude modulation. These signatures were defined empirically in order to obtain a clear distinction during the actual targeting task. The equations used for generating the different vibratory stimuli were the following:

$$W_1 = a\sin(2\pi \times 121dt) \qquad \qquad W_3 = a\sin(2\pi \times 3dt)\sin(2\pi \times 121dt) W_2 = a\sin(2\pi \times 0.5dt)\sin(2\pi \times 121dt) \qquad \qquad W_4 = a\sin(2\pi \times 31dt)\sin(2\pi \times 53dt)$$
(1)

 W_1 defines a sinusoidal vibration at 121 Hz. W_2 is an amplitude modulation of W_1 by a 0.5 Hz sinusoid, producing the sensation of a rhythmic pulsing vibration. W_3 is a modulation of W_1 by a 3 Hz sinusoid, producing the sensation of rapid impulse vibration. W_4 is a 53 Hz sinusoid modulated by a 31 Hz whose combination resulted in a rather rough vibration sensation due to the inharmonicity of the two components.

The availability of spectral discrimination thresholds for vibratory stimuli (aim of the current study) would have allowed a more direct, parametric, and precise stimuli definition process, facilitating creation of multiple distinguishable haptic signatures.

4 Experimental Study

Aiming at investigating the differences between audio vibratory tactile perceptions relative to the detection of spectral variations between signals with the same fundamental frequency, three tests have been designed and carried out.



Fig. 1. (left) Haptic test stimuli protocol, including the use of noise-suppression headphones. (left driver) Cone removed and added dome for haptic stimuli, (right drive) unmodified for audio stimuli. (right) Hand of a participant, showing the wrist resting on the wooden board and the three mid-fingers placed on the plastic dome.

The order and complexity of the tests is related to the two aims of the study: (1) to identify any difference between the two modalities, and (2) to quantify this difference. In the first test, the ability to discriminate between three signals with different spectral characteristics is assessed for the auditory and haptic modalities. This is a relatively trivial auditory task, and the audio component is included just as a reference. The second and third tests aim at quantifying the discrimination thresholds between signals with the same fundamental frequency but different spectral content; tone/2-tones discrimination, odd and even harmonics components discrimination. The use of the same hardware and software for the delivery of the auditory and haptic feedback was made in order to facilitate a consistent comparison of results between the two modalities.

Participants consisted of 12 subjects, 8 males/4 females, aged 19 - 52 years. Each test session took approximately one hr: 5 min for the calibration, 15 min for the first test, 30 min for the second test, and 10 min for the third.

4.1 Experimental Apparatus

The experimental apparatus is composed of a software component, a computer, an audio interface, an audio amplifier and two 8 inch loudspeaker woofer drivers, mounted on a wooden board (see Fig. 1): one of these has been modified by removing the speaker cone. A coupling system (a rigid 10 cm diameter plastic dome, on which the fingers of the subjects are placed) was installed in order to transfer the vibrations of the coil to the hand. The subjects are instructed to rest their dominant hand on the wooden board surrounding the driver, and to position the last phalanx of their middle three fingers (index, middle, and ring fingers) on the plastic dome, without applying any pressure (see Fig. 1). These choices (hand position, parts of the finger to be in contact with the vibratory actuator, etc.) have been made based on Verrillo, 1962 and Verrillo, 1963. Considering the audio rendering modality, the subjects have been asked to place their head at 1 m from the driver. All subjects completed tests for both modalities. The generation and processing of the signals, the testing procedures and the data collection have all been implemented in a Max/MSP ¹ platform/patch. The digital signals are sent to a MOTU Traveler FireWire audio interface, converted to analogue signals, and sent to

¹ http://www.cycling74.com

an Oniphonics Footprint 150 amplifier, and then to one of the two drivers, depending on the testing modality. Frequencies could be generated between 35 Hz and 1400 Hz without any notable resonance outside the dynamic range of ± 10 dB (un-weighted).

4.2 Calibration

After a series of informal trials and evaluations, and reviewing previous literature in the field of auditory and vibratory-tactile perception (see Sec. 2) as well as the limitations of the playback system, the frequency ranges for the tests have been set at 35-250 Hz for haptic, and 200-1400 Hz for audio.

The audio modality signal amplification has been calibrated in order to generate a SPL value of 70 dB A-weighted for a 1000 Hz pure tone at 1 m distance from the loudspeaker driver (head location during audio tests). This value has been chosen considering the standard levels used in audiological evaluations (Penrod, 1985).

Due to the sensitivity of thresholds to contact area, the haptic feedback level calibration stage is performed for each subject individually. The participants are asked to place their fingers on the plastic dome (see Sec. 4.1) while a 100 Hz sinusoidal stimulus is reproduced. The level is increased until the subjects can just perceive a vibration. The signal gain is then increased by 20 dB, in order to have a clear presentation level and to assure consistency in the haptic presentation stimuli across subjects. The rendered level is therefore calibrated at *Threshold of Perceptibility* (100 Hz)+20 dB.

During haptic testing, the SPL produced by the cone-less driver was between 50 dB and 57 dB A-weighted (measured at 100 Hz), depending on the level calibration described in the previous paragraph, while the background noise in the testing environment was 32 dB A-weighted. In order to avoid auditory stimulation from the haptic driver, a pair of passive noise-suppression headphones were used (see Fig. 1) which provided a sound level reduction of 20 dB (manufacture's statement). No subject reported hearing the audio signal produced by the haptic device during the test.

4.3 Test One: Signal Comparison

The subjects were asked to discriminate between three types of simple signals:

- (a) Single pure tone (Sine): frequency chosen in the middle of the ranges for the two modalities, therefore $f_1^h = 100 \text{ Hz}$ for haptic and $f_1^a = 600 \text{ Hz}$ for audio.
- (b) *Narrow-band noise* (Noise): white noise processed with a band-pass filter centred at f_1^h for the haptic stimulus and at f_1^a for the audio stimulus.
- (c) Two concurrent pure tones (2Sine): the first with frequency f_1^h or f_1^a for the two modalities, and the second with frequency $f_2 = f_1 \times 1.7$. This value has been chosen in order to generate an overtone with no harmonic relation to the fundamental.

The test consists of three separate trials for each rendering modality. For the first trial signals a and b are compared: 20 repetitions of two randomly selected signals (a&b, b&a, a&a, or b&b) are presented to the participant, who is asked to state, after each repetition, if the two signals are the same or if they are different. For the second trial, signals b and c are compared, for the third trial signals a and c. The two signals are reproduced in the following sequence: first signal for 1000 ms, 200 ms of silence, second



Fig. 2. (left) Mean and std of identification errors for test one. (right) Distribution of identification error for test one; — median, \circ mean. Values are displayed according to rendering modality and signal pairs for stimuli *a*: Sine, *b*: Noise, *c*: 2Sine.

signal for 1000 ms, with each signal processed with a 5 ms fade in and fade out. All signals are calibrated in order to generate the same dB RMS levels.

The percentage of identification errors in the discrimination task between the different pairs of signals are reported in Fig. 2. As expected, the audio modality error rate is quite low (mean of 0.7% over the three pairs, std 2.1), while for the haptic modality error rates are higher (mean of 11.7%, std 13.1). Furthermore, it can be noted that for the haptic modality the error rates for the first two trials are similar, while the third (Sine-2Sine) are notably higher. A similar relative increase is noted in the results for the audio modality. This can be justified considering that in the first two trials a periodic signal (*a* or *c*) is compared with a narrow-band noise (*b*), while in the third trial two sinusoidal signals are compared (*a* and *c*), making the discrimination task more complex.

The initial outcome of this test is that it is indeed possible to discriminate through the tactile sense between vibrations with equal (or similar) fundamental frequency and different spectral content.

4.4 Test Two: Two Tones Detection

Using a simple up-down 1 dB step adaptive procedure (Levitt, 1978), the discrimination threshold is measured between a pure tone signal and a stimulus composed of two concurrent pure tones, changing the amplitude and frequency of the second tone.

The participants are presented with groups of two stimuli in the following sequence: first signal for 1000 ms, 200 ms of silence, second signal for 1000 ms, with each signal processed with a 5 ms fade in and fade out. Initially, the two stimuli are the same (a pure tone *a* with frequency f_a). The second stimulus is then iteratively modified by adding to *a* another pure tone *b* with frequency f_b , increasing adaptively the amplitude of *b* and decreasing the one *a*, in order to maintain the same RMS level for both signals. The participants are then asked to determine when a difference can be heard between the first and the second stimulus. The test is then carried out adaptively until a threshold value is found (after 5 up-down direction changes).



Fig. 3. (a) Waveform of the pure tone audio signal used for test two. (b) Waveform of the twotones audio signal ($f_2 = f_1 \times 1.7$) used for test two. For the haptic rendering, the same signals have been used, with $f_a^h = 100$ Hz instead of $f_a^a = 500$ Hz.

The values of f_1 are set at 500 Hz for the audio modality, and 100 Hz for haptic. Six values have been chosen for the signal *b*, where f_b is a multiple of f_a defined by the multiplier factor *m* (for both modalities): m = 0.5, 0.7, 1.7, 2.0, 2.7, 3.0. These values are chosen in order to allow various combinations of two concurrent tones at different frequencies, with and without harmonic relations. Example waveforms of both signals are displayed in Fig. 3.

The discrimination threshold values, expressed in terms of dB difference between the a and b components in the second signal, are reported for each modality and for each value of m in Fig. 4.

There is a notable difference between the mean discrimination threshold values for the haptic modality (mean of -17.7 dB, std 8.3) and for audio (mean of -46.4 dB, std 11.9), the latter being distinctly lower (higher sensitivity). This quantifies the fact that the human hearing system is more sensitive in discriminating between a pure tone and a complex tone composed of two pure tones if compared with the tactile system.

Furthermore, it can be observed that for the haptic modality the values are generally lower (better performance) when f_2 is not in harmonic relation with f_1 . For m = 0.7, 1.7, 2.7, the mean discrimination threshold is -20.2 dB (std 8.1), while for m = 0.5, 2.0, 3.0 it is -15.2 dB (std 9.4), a difference of 5.0 dB. A similar tendency can be observed for the audio modality, but in this case the difference is only of 2.3 dB.

Inferential statistics have been performed to identify whether the differences between groups are statistically significant. As the data sets are small and not normally distributed, exact Mann-Whitney tests were used. For the tactile stimulation, results with the in-harmonic f_2 components differ significantly from the harmonic ones (MW U = 431, p = 0.015), while for the auditory stimulation the differences between the two groups are not significant (MW U = 569, p = 0.374). The effect size can be considered very small for both modalities, with r = -0.706 for tactile and r = -0.26 for audio.

An explanation could be attempted considering the fact that non-harmonic overtones are more likely to generate amplitude beats with the fundamental component, and these could be used to discriminate between different stimuli, offering a further cue for this experimental task. Nevertheless, this cue should be available for both modalities, and not only for the tactile one as outlined by the results of the Mann-Whitney test. At the moment, there does not seem to be a precise and clear explanation of this result.



Fig. 4. (left) Mean and std of the discrimination threshold results for test two. (right) Distribution of discrimination threshold results for test two; — median, \circ mean. Values are displayed according to rendering modality and f_2 multiplier *m*.

Other tendencies can also be observed (e.g. the audio modality seems to be more sensitive to m,), though not enough data has yet been collected for allowing further statistical analysis (see Sec. 5).

4.5 Test Three: Spectral Detection

Using a simple up-down adaptive procedure (see Levitt, 1978), the discrimination threshold is measured between a tone with even harmonic components (such as the *typical* organ pipe) and a tone with odd harmonic components (such as the sound of the clarinet). The testing procedure is the same as for test two (see Sec. 4.4).

Initially, participants are presented with the same stimulus repeated twice (a pure tone with the first 3 even harmonic components, 2^{nd} , 4^{th} , and 6^{th} , each one with an amplitude 3 dB lower than the previous). The second repetition is then changed adaptively, reducing the amplitude of the even harmonic components and increasing that of three odd harmonic components (3^{rd} , 5^{th} , and 7^{th}). In order to maintain the same overall levels for both repetitions, the reduction and amplification of the harmonic components is performed using a cosine function on a 128-step linear scale from 0 to $\pi/2$ for the even harmonic components, and from $\pi/2$ to 0 for the odd ones. In Fig. 5 the waveform and spectrum of both signals are displayed.

The participants are asked to identify when a difference can be heard between the two stimuli. The test is carried out adaptively until a threshold value is found (after 5 up-down direction changes). The values of the fundamental frequency for the complex tone have been set at 200 Hz for the audio modality, and 35 Hz for the tactile one.

The discrimination threshold values, expressed as dB difference between the evenharmonics and odd-harmonics stimuli, are reported for each modality in Fig. 6. It is possible to observe that there is a sensible difference between the mean values for the haptic modality (mean of -15.3 dB, std 13.2) and for the audio one (mean of -27.7 dB, std 9.2). For the audio modality, it is therefore possible to discriminate between a tone with just even harmonic components and a tone with also odd harmonic components when the level difference between these is 27.7 dB, while this value decreases to 15.3 dB for the tactile rendering, a difference of 12.4 dB in harmonic detection sensitivity.



Fig. 5. (a) Waveform of the even-harmonic audio signal used for test three. (b) Waveform of the odd-harmonic audio signal used for test three. For the haptic rendering, the same signals have been used, but with a fundamental frequency of 35 Hz instead of 200 Hz, and the harmonics consequently scaled.



Fig. 6. Distribution of discrimination threshold results for test three by rendering modality; — median, \circ mean. Values are expressed as the difference in dB between the even-harmonics and odd-harmonics stimuli.

5 Conclusion

The initial outcomes of this series of perceptual evaluations comparing audio and hapticvibratory senses is that for both modalities spectral differences between different stimuli with the same fundamental frequency can be perceived, with auditory perception being more sensitive when compared to tactile one.

In terms of discrimination thresholds between a pure tone and a stimulus composed of two pure tones, the difference between the two modalities was 28.7 dB, with haptic sensitivity being distinctly lower. Furthermore, a lower discrimination threshold (5 dB) for the haptic modality is found when the two tones composing the second stimulus are not in harmonic relation. A similar tendency, but with reduced magnitude, is also observed for the audio modality, but cannot be considered statistically significant for the current results. In terms of discrimination threshold between a tone with even harmonic components and a tone with odd harmonic components, the audio modality exhibited a sensitivity 12.4 dB greater than the haptic modality.

Considering the results reported in Merchel et al., 2010, where it was shown how haptic stimuli with different spectral and rhythmic characteristics could be discriminated, the results from the current study expand those outcomes outlining that discrimination is possible even between stimuli with no rhythmic features at all, and with solely spectral differences. Furthermore, referring to the preliminary study described in Sec. 3, the initial results of the current study could be employed for allowing the creation of a larger number of distinguishable tactile-vibratory signals. Considering for example the outcomes of test three, six distinguishable haptic signatures could be created changing the levels of the overtones between a signal with only even harmonic components and a signal with only odd ones. Moreover, the use of in-harmonic overtones could be exploited even further for allowing a higher discrimination factor.

This preliminary study is currently being repeated with a larger test population. Furthermore, these subsequent tests will be carried out on additional groups of visually and hearing impaired subjects, in an attempt to consider the case of individuals with and without sensory deprivations.

References

- Altinsoy, M.E., Merchel, S.: Audiotactile Feedback Design for Touch Screens. In: Altinsoy, M.E., Jekosch, U., Brewster, S. (eds.) HAID 2009. LNCS, vol. 5763, pp. 136–144. Springer, Heidelberg (2009)
- Branje, C., Maksimouski, M., Karam, M., Fels, D.I., Russo, F.: Vibrotactile Display of Music on the Human Back. In: ACHI 2010: International Conference on Advances in Computer-Human Interactions, St. Maarten, Netherlands, pp. 154–159 (2010)
- Brown, L.M., Kaaresoja, T.: Feel who's talking: using tactons for mobile phone alerts. In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems, Montreal, Canada, pp. 604–609 (2006)
- Choi, S., Tan, H.Z.: Toward realistic haptic rendering of surface textures. IEEE Computer Graphics and Applications 24(2), 40–47 (2005)
- Levitt, H.: Adaptive testing in audiology. Scand. Audiol. Suppl. 6, 241-291 (1978)
- Menelas, B., Picinali, L., Bourdot, P., Katz, B.F.G.: Audio Haptic Feedbacks for an Acquisition Task in a Multi-Target Context. In: IEEE 3DUI User Interfaces Conf., Waltham, MA, US (2009)
- Merchel, S., Altinsoy, E., Stamm, M.: Tactile Music Instrument Recognition for Audio Mixers. In: 128th Conv. of the Audio Egineering Society, London, UK (2010)
- Moore, C.J.B.: An Introduction to the Psychology of Hearing. Academic Press, London (2003)
- Penrod, J.I.: Speech Discrimination Testing. In: Katz, J. (ed.) Handbook of Clinical Audiology, Baltimore, MD, pp. 235–255
- Verrillo, R.T.: Investigation of Some Parameters of the Cutaneous Threshold for Vibration. J. Acoust. Soc. Am. 34(11), 1768–1773 (1962)
- Verrillo, R.T.: Effect of Contactor Area on the Vibrotactile Threshold. J. Acoust. Soc. Am. 37, 843–846 (1963)
- Verrillo, R.T., Gescheider, G.A.: Perception via the Sense of Touch. Tactile Aids for the Hearing Impaired (1992); edited by Summers, I.R.: Whurr Publishers, London
- West, A.M., Cutkosky, M.R.: Detection of Real and Virtual Fine Surface Features with a Haptic Interface and Stylus. In: 6th Annual Symposium on Haptic Interfaces, Dallas, TX (1997)