The Effect of Frequency Shifting on Audio-Tactile Conversion for Enriching Musical Experience

Ryuta Okazaki[1,2], Hidenori Kuribayashi[3], Hiroyuki Kajimioto[1,4]

[1] The University of Electro-Communications, [2] JSPS Research Fellow,
[3] Nikon Corporation, [4] Japan Science and Technology Agency.
{okazaki,kajimoto}@kaji-lab.jp, Hidenori.Kuribayashi@nikon.com

Abstract. We have applied a frequency shifting method, which was proposed previously in the literature for mixer manipulation, with the aim of generating vibration-based feedback to enrich the listener's musical experience. Experimental results showed that the proposed method significantly increased the listener's evaluation of sound consisting of high-frequency components, while a relatively poor evaluation was observed for sound containing low-frequency components.

Keywords: audio-tactile interaction, audio-tactile conversion, vibrotactile.

1 Introduction

Recent studies have suggested that several cross-modal relationships exist between the tactile and auditory senses. Suzuki et al. reported that tactile roughness perception can be modified by adding a task-irrelevant sound [1]. Yau et al. showed that the subjective auditory intensity is affected by simultaneous presentation of tactile stimuli [2]. Each of the factors of tactile-audio interactions, including phase, synchrony, and frequency, have also been studied [3][4][5][6]. Physiological studies have reported that tactile and auditory sensations share a common neural mechanism [7]. Additionally, similarities between the tactile and auditory senses in a higher order region, e.g. the "consonance" between the two modalities, have also been investigated [8][9][10].

Recently, some audio-tactile conversion methods that focused on the cross-modal relationships were proposed with the aim of improving and enhancing the value of the content or the user experience (UX) using tactile stimulation. Karam et al. developed the Emotichair, which focused on the spatial processing of the frequency at the auricle and then applied it to the tactile presentation [11][12]. Birnbaum and Wanderley proposed a "natural" tactile feedback method for electronic musical instruments based on analysis of the vibration characteristics of real instruments [13]. Lee and Choi proposed an audio-tactile conversion method that focused on the roughness and the loudness [14]. Lim et al. proposed an audio-tactile conversion method that only deals with a specific frequency range to suit the preferences of the users [15].

In this paper, we propose an audio-tactile conversion method that focused on the gap in the perceivable frequency range between the auditory and tactile ranges. It is

adfa, p. 1, 2011.

well known that there is a huge gap between the perceivable frequency ranges of the auditory (20 Hz–20 kHz) and tactile (0–1000 Hz) senses. As a result, the direct conversion of sound is likely to be unperceivable with an increase in frequency.

To address this issue, we tried to “compress” the frequency range of the audio signal to create a tactile vibration with a frequency that was shifted one or two octaves down. Previous studies suggested that a one-octave-shifted vibration created from an audio signal eased the user’s distinction of musical instruments [16][17]. We use a similar method. However, because the auditory frequency range is much wider than the tactile range, a one octave shift might be insufficient for music consisting of higher frequency components. Our research questions are as follows.

(1) To determine whether frequency-shifted vibrations enrich musical experiences.

(2) To assess whether one or two octave shifts are effective, and when this occurs.

2 Conditions

To answer these questions, we prepared a control condition that provides an audio signal as tactile vibration (control) and four different conditions: (1) the tactile vibration is shifted one octave down against the original audio signal (1OT); (2) the tactile vibration is shifted two octaves down against the original audio signal (2OT); (3) combination of the two tactile tracks (1+2OT); (4) combination of the original audio signal, the one-octave-shifted vibration and the two-octave-shifted vibration filtered by a 250 Hz, Q=0.0 band-pass filter (1+2OT+BP). We used the band-pass filter to reduce very low/high frequency components. We found in a preliminary experiment that application of the two-octave shift sometimes produced very low frequency components, which typically leads to an impression that the audio-tactile experience was “muffled”. Similarly, very high frequency components sometimes lead to a tingling feeling.

We used Hayaemon software (http://en.edolfzoku.com/hayaemon2/) to generate the octave-shifted tactile signal.

In the next chapter, the evaluations of the four specified conditions were measured in psychophysical experiments.

3 Experiment

Apparatus: The setup comprised a computer with two audio channels. One audio channel powered the two sides of a set of high-quality headphones (QuietComfort, Bose Inc., USA), with strong active noise cancellation; the other channel was connected to an audio amplifier (RSDA202, Rasteme Systems Inc., Japan) driving a vibrotactile transducer (Haptuator mark 2, Tactile Labs, Canada). The transducer was firmly attached to a mobile device via a plastic cover (iPod Touch, 5th generation, Apple Inc., USA). The total weight of the vibrating device was 113 g (**Fig. 1** left).

Participants: Five participants (three males, two females), aged between 22 and 43 years. Each reported no auditory or tactile impairment.

Stimuli: To observe the effects of the proposed method for various frequency ranges, we prepared three different kinds of music. (A) The sound of a music box containing high-pitched tones (Akaneiro-op.x03). (B) Classical music containing mid-range sound (The Moldau, or Vltava). (C) Jazz music containing low-frequency tones (Sleepin´ Maple Syrup Jazz). We used the first 15–20 s of each musical piece for the experiments.

The experimental stimulation consisted of an audio signal/ tactile vibration pair. The total experimental stimulation included 15 pairs of audio and tactile stimuli. The average amplitudes of the signals were set to be equal. The auditory stimulation amplitude was set at about 55 dB SPL. The maximum acceleration with this setting was approximately 1.5G.

Evaluation: To determine whether evaluation of the content was altered by the combination of tactile and audio track signals, we carried out an evaluation using the VAS (visual analogue scale) with the following five questions. The questions used here were based on the questions used by Lee and Choi [14], but were partly modified: *Q1: Temporal harmony* – "Did the vibrations match with the sound temporally?"; *Q2: Frequency harmony* – "Did the frequency of the vibrations match with the sound?"; *Q3: Comfort* – "Did you feel comfortable with the vibrations, enabling you to enjoy the sound?"; *Q4: Preference* – "Did you enjoy the vibrations when presented with the sound?"; and *Q5: Fun* – "Did the vibrations make the sound fun?"

Procedure: While seated, each participant wore headphones and held the device in both hands. Each of the 15 pairs of auditory and tactile stimuli was presented simultaneously. After experiencing these stimuli, the participants were instructed to answer the five questions using the VAS. No time limit was set for the task, but all tasks were completed within 60 s. All 15 pairs of stimuli were presented twice but at random, giving 30 tasks in total per participant (**Fig. 1** right).

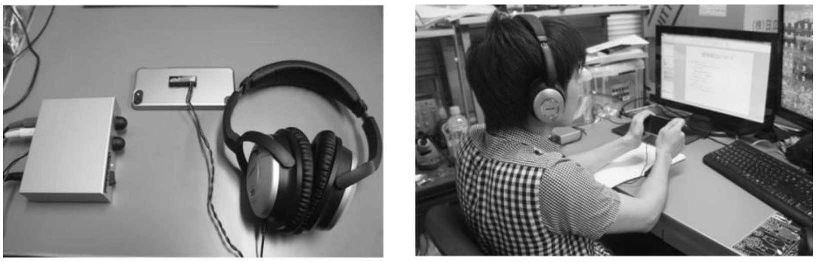

Fig. 1. Left: The main components of the apparatus, consisting of headphones, an amplifier, and the vibrating device. Right: Overview of the experiment.

4 Results and Discussion

The experimental results are shown in **Table 1**. The lines in the table represent each condition, while the columns represent questions. The figures in the table indicate the average of the VAS scaled from 0 to 10. To verify the differences when compared with the control condition, Friedman's Test and repeated Wilcoxon signed-rank tests were performed. The p-values obtained were then corrected by the Holm–Bonferroni method at a significance level of 5% to avoid multiplicity concerns.

Q1: Temporal harmony: In this question, the participants assessed the temporal matching of the sounds and vibrations that they experienced.

In the case of the sound of the music box, the evaluations of the 1OT, 2OT, 1+2OT, and 1+2OT+BP conditions were significantly higher when compared with that of the control condition ($p<0.05$). It is believed that the evaluation of the control condition decreased because the main frequency contained in the original sound of the music box was much higher than the perceivable range of the tactile signal.

In contrast, in the cases of jazz and classical music, no significant differences were found when compared with the control condition.

Q2: Frequency harmony: Here, the perceived frequency harmony between the sound and the vibration was evaluated. The evaluations of the 1OT, 2OT, 1+2OT, and 1+2OT+BP conditions were significantly higher than that of the control condition for the sound of the music box ($p<0.05$). In the classical music case, the only significant difference was observed between the 2OT condition and the control condition.

However, in the jazz music case, the evaluations of all four conditions (1OT, 2OT, 1+2OT, 1+2OT+BP) were significantly lower than that of the control condition ($p<0.05$).

Q3: Comfort, Q4: Preference, and Q5: Fun: The evaluation tendencies for these criteria were quite similar. In the case of the music box sound, the evaluations of the 1+2OT and 1+2OT+BP conditions were significantly higher than that of the control condition at each criterion ($p<0.05$). A significant difference between the 2OT condition and the control condition for jazz music was also observed in the evaluation of fun.

Discussions: From the experimental results, application of octave-shifted tactile vibrations to music will increase the evaluation of that music in terms of temporal matching, frequency harmony, comfort, preference, and fun, particularly when the original music comprises high frequency components, e.g. the sound of a music box. In particular, in the 1+2OT and 1+2OT+BP cases, their evaluations were significantly increased for all criteria when compared with those of the control condition. It is therefore indicated that the proposed method will improve the quality of music that mainly consists of high frequency components.

However, there was no significant difference between the evaluation of the proposed method and the control condition in the case of music that mainly comprises midrange sounds, such as classical music. This may be because the midrange music originally contains a range of frequency components that is perceptible with the tactile senses, regardless of the octave shifting.

In the case of jazz music, which mainly contains low frequency components, the evaluation was partly reduced by octave shifting. There are two possible reasons for this. The first is the effect of the very low frequency components generated by repeated octave shifting. Because jazz music originally consists of low frequency components represented by the bassline, very low frequency components (<10 Hz) were produced by performing the octave shifting once or twice. As a result, this very low frequency component might suggest a muffled quality to the participants. In the 1+2OT+BP condition, the band-pass filter was applied to avoid this very low frequency, however, the participants still reported the muffled sensation in their introspection reports, which suggests that the band-pass filter was not sharp enough. The second possible reason is the frequency response characteristic of the transducer that we used, which exerts its full performance for vibrations at more than 60 Hz.

Table 1. Average results for Q1: Temporal harmony, Q2: Frequency harmony, Q3: Comfort, Q4: Preference, and Q5: Fun. Figures in parentheses represent standard deviations. Significant differences are shown with dark grey (higher than control) and light grey (lower than control) highlighting.

Q1	Control	1OT	2OT	1+2OT	1+2OT+BP
Music box	1.02 (1.30)	4.16 (2.93)	7.15 (1.83)	6.60 (2.70)	7.87 (1.87)
Classic	6.35 (1.90)	6.51 (2.35)	8.28 (1.05)	7.21 (1.20)	7.90 (1.84)
Jazz	8.56 (0.96)	7.81 (1.46)	7.59 (1.17)	8.09 (1.67)	7.97 (0.94)
Q2	**Control**	**1OT**	**2OT**	**1+2OT**	**1+2OT+BP**
Music box	1.14 (1.57)	4.24 (2.47)	4.15 (2.30)	5.13 (2.86)	6.34 (2.17)
Classic	4.71 (1.85)	6.10 (2.38)	7.87 (1.34)	6.32 (1.49)	6.32 (3.33)
Jazz	8.96 (0.71)	6.45 (1.96)	7.14 (1.28)	7.32 (1.57)	6.68 (2.38)
Q3	**Control**	**1OT**	**2OT**	**1+2OT**	**1+2OT+BP**
Music box	2.33 (2.41)	3.87 (2.90)	4.72 (2.00)	5.34 (2.46)	6.78 (1.35)
Classic	4.33 (2.35)	6.42 (2.34)	7.60 (1.44)	5.42 (2.88)	6.85 (2.54)
Jazz	8.31 (1.86)	6.74 (2.21)	5.57 (2.65)	6.87 (2.36)	5.71 (2.62)
Q4	**Control**	**1OT**	**2OT**	**1+2OT**	**1+2OT+BP**
Music box	1.73 (1.92)	3.73 (2.81)	4.44 (2.23)	5.62 (2.49)	5.81 (1.72)
Classic	4.24 (2.78)	6.21 (2.52)	8.09 (1.40)	5.57 (2.61)	6.26 (2.91)
Jazz	8.33 (1.60)	6.73 (1.89)	5.19 (3.29)	6.45 (3.13)	5.64 (2.90)
Q5	**Control**	**1OT**	**2OT**	**1+2OT**	**1+2OT+BP**
Music box	1.64 (1.92)	3.84 (2.78)	3.75 (2.74)	5.42 (2.77)	5.93 (1.89)
Classic	4.16 (2.84)	5.64 (2.45)	7.76 (1.69)	5.96 (2.04)	5.98 (2.71)
Jazz	7.96 (2.10)	6.69 (1.69)	6.10 (2.32)	6.39 (2.95)	5.87 (2.86)

5 Conclusion

In this paper, we aimed to create a tactile vibration to increase the evaluation of music corresponding to that vibration, with the idea of compressing the frequency range by octave shifting that was previously used for tactile musical instrument distinction. Experimental results showed that proposed method significantly increased the evaluation of sounds that consisted of high frequency components in terms of temporal harmony, frequency harmony, comfort, preference, and fun. However, relatively poor evaluations were obtained for sounds with low frequency components.

References

1. Suzuki, Y., et al.: Selective Effects of Auditory Stimuli on Tactile Roughness Perception. Brain Res. 1242, 87-94 (2008)
2. Yau, J.M., et al.: Separate Mechanisms for Audio-Tactile Pitch and Loudness Interactions. Front. Psychol. 1, 160 (2010)
3. Yau, J.M. et al.: Temporal Frequency Channels are Linked across Audition and Touch. Curr. Biol. 19, 561-566 (2009)
4. Wilson, E.C., et al.: Integration of Auditory and Vibrotactile Stimuli: Effects of Phase and Stimulus-onset Asynchrony. J. Acoust. Soc. Am. 126(4), 1960-1974 (2009)
5. Wilson, E.C., Reed, C.M., Braida, L.D.: Integration of Auditory and Vibrotactile Stimuli: Effects of Frequency. J. Acoust. Soc. Am. 127(5), 3044-3059 (2010)
6. Wilson, E.C., et al.: Perceptual Interactions in the Loudness of Combined Auditory and Vibrotactile Stimuli. J. Acoust. Soc. Am. 127(5), 3038-3043 (2010)
7. Kayser, C., et al.: Integration of Touch and Sound in Auditory Cortex. Neuron, 48(2), 373-384 (2005)
8. Altinsoy, E. M, et al.: Cross-Modal Frequency Matching: Sound and Whole-Body Vibration. In: Proc. HAID2010. LNCS vol. 6306, pp. 37-45. Springer (2010)
9. Yoo, Y., et al.: Consonance Perception of Vibrotactile Chords: A Feasibility Study. In: Proc. HAID2011. LNCS vol. 6851, pp. 42-51. Springer (2011)
10. Okazaki, R., et al.: Judged Consonance of Tactile and Auditory Frequencies. In: Proc. WHC2013, 663-666 (2013)
11. Karam, M., et al.: Designing the Model Human Cochlea: An Ambient Crossmodal Audio-Tactile Display. IEEE Trans. Haptics. 2(3), 160-169 (2009)
12. Karam, M., et al.: The Emoti-chair: an Interactive Tactile Music Exhibit. In: Proc. CHI2010, 3069-3074. ACM (2010)
13. Birnbaum, D.M., Wanderley M.M.: A Systematic Approach to Musical Vibrotactile Feedback. In: Proc. ICMC2007, 397-404 (2007)
14. Lee, J., Choi, S.: Real-time Perception-level Translation from Audio Signals to Vibrotactile Effects. In: Proc. CHI2013, 2567-2576. ACM (2013)
15. Lim, J.M., et al.: An Audio-Haptic Feedback for Enhancing User Experience in Mobile Devices. In: Proc. ICCE2013, 49-50 (2013)
16. Merchel, S., et al.: Tactile Identification of Non-Percussive Music Instruments. In: Proc. ForumAcusticum2011, 1257-1261 (2011)
17. Merchel, S., et al.: Touch the Sound: Audio-Driven Tactile Feedback for Audio Mixing Applications. J. Audio Eng. Soc. 60(1), 47-53 (2012)