HAPTICS SYMPOSIUM '16, APRIL 8–11, 2016, PHILADELPHIA, PENNSYLVANIA, USA
978-1-5090-0903-9/16/$31.00 ©2016 IEEE

Vibrotactile Rendering of Gunshot Events for 4D Films

Jongman Seo[1], Reza Haghighi Osgouei[1], Soon-Cheol Chung[2], and Seungmoon Choi[1]

***Abstract*— 4D films present diverse physical effects including vibration. A common approach for generating vibrotactile effects is manual authoring. Alternatively, 4D effect designers sometimes rely on direct audio-tactile conversion with no processing. While applying this direct method to gunshot events, we realized that the perceptual intensity difference between the audio and tactile channel is a major problem. In this paper, we propose a compensation method for auditory-to-vibrotactile conversion transparent to both actuator dynamics and perception.**

I. INTRODUCTION

4D films present a variety of physical effects such as motion, vibration, wind, water, and scent. Although the number of 4D theaters has been growing rapidly, 4D effect designers still use in-house manual authoring software to create all the effects [1]. Vibration effects are no exception, but their production load can be reduced by directly sending an audio signal to a tactile device without processing. 4D effect designers often use this direct audio-tactile conversion owing to its simplicity.

In haptics, several attempts have been made to improve the perceptual quality of vibrotactile feedback converted from audio data for different aims. Karam et al. presented Emoti-Chair, which supports translation of sound into tactile vibration with a pitch shift feature [2]. They used eight discrete audio-to-tactile channels to bring music to the sense of touch effectively. Hwang and Choi applied the equalizer concept to the dual-band rendering of vibrotactile music [3] and also improved the algorithm by adopting auditory saliency estimation [4]. Lee and Choi proposed a real-time perception-level audio-to-vibrotactile translation framework that maps the roughness and loudness of audio signal to the roughness and perceived intensity of vibration signal [5]. This framework allows for selective rendering appropriate for a given purpose, e.g., to provide vibrotactile feedback only for rough special effect sounds but not for background music. Lim et al. developed a haptic library that creates tactile feedback by converting audio data in a user-selected frequency band [6]. Okazaki et al. investigated a two-octave frequency-shifting method on audio-tactile conversion for enriching musical experience [7]. This method improved user-assessed quality of the music that included high-frequency components, but did not for the music with low-frequency components.

In this work, we focus on vibrotactile rendering of gunshot events for 4D films. Gunshot is a frequent event in action movies, and emphasizing it with vibrotactile effects can increase immersiveness to a great extent.

A gunshot sound is generally composed of gun blast and its subsequent reflection, and these features clearly appear in its sound signal. However, these features may not be appropriately transformed to vibration with the direct conversion approach that does not take into account the differences between audio and tactile transducers and between auditory and tactile perception. Further, the previous sound-to-vibration methods may not be very adequate since they mostly focused on enhancing musical experience. As such, they do not preserve the time-domain characteristics of gunshot signals carefully when manipulating signals in the frequency domain to shift pitch or extract a particular frequency band. The former is important for short gunshot sound to synchronize the waveforms of sound and vibration.

We propose a compensation method for the direct conversion approach that aims to mitigate the perceived intensity difference problem between audio and tactile signal while retaining the time-domain characteristics of the sound signal for a precise description of the gunshot features.

II. ISSUES OF DIRECT GUNSHOT AUDIO-TACTILE CONVERSION

Direct audio-tactile conversion is a technique to send an audio signal directly to a tactile device without any processing. We applied this method to vibrotactile rendering of gunshot events for 4D films and carried out a small-scale user study to confirm its anticipated disadvantages.

For vibration generation, we used a large linear actuator (12 cm$\times$12 cm$\times$3 cm) with an amplifier (T108SM and C-602, Crowson) controlled by a PC using a data acquisition card (PCIe-6341, National Instruments) at the sampling rate of 44.1 kHz. This actuator was designed to provide vibrotactile motion effects to a chair, a recliner, or a small couch. We also extracted 32 scenes that included gunshots from movies and concatenated them into one clip ($45'3''$).

Four students participated in the evaluation. They watched the movie clip with vibrotactile feedback provided by direct conversion from the sound file of the movie clip. They were asked to comment on the quality of vibrotactile feedback, especially in terms of harmony with sound. Most of their comments (78 of 98 comments) were related to the perceptual intensity difference between audio and vibrotactile gunshot signals; the tactile intensity they perceived was weaker

This work was supported by the Mid-Career Researcher Grant NRF-2013R1A2A2A01016907 and the Pioneer Research Center Program 2011-0027994 from the National Research Foundation (NRF) of Korea.

[1]J. Seo, R. H. Osgouei, and S. Choi are with the Haptics and Virtual Reality Laboratory, Department of Computer Science and Engineering, Pohang University of Science and Technology (POSTECH), Republic of Korea. E-mail: {cjman224, haghighi, choism}@postech.ac.kr.

[2]S.-C. Chung is with the Department of Biomedical Engineering, Konkuk University, Republic of Korea. E-mail: scchung@kku.ac.kr.

Fig. 1. Vibration actuator used in this study (on a chair).

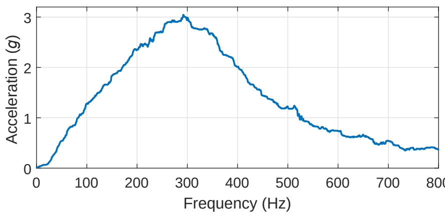

Fig. 2. Frequency response of the linear actuator, $H_d(s)$.

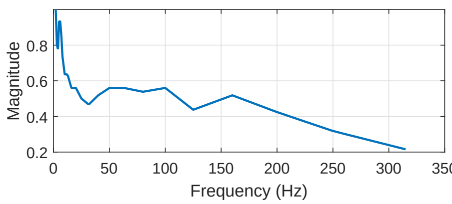

Fig. 3. Critical band filter for the Pacinian channel, $H_m(s)$.

than the auditory intensity. Note that this perceptual intensity difference cannot be easily cured by simple adjustment of the conversion gain due to the actuator's limited input range.

III. Compensator Design

The source of this problem can be sought by a close systematic look at the direct conversion process. Two systems, linear actuator and human tactile perception, are involved in the process. First, an audio signal passes through the vibrotactile actuator, which has a limited bandwidth with a resonance peak at 292 Hz (Fig. 2), eliminating the higher frequency components of the audio signal. Second, the signal proceeds to the human mechanoreception system mainly through the RA (Rapidly Adaptive) 1 and PC (Pacinian) channel. The output of the linear actuator is transformed further by our mechanoreceptor system.

As a solution, we propose a feed-forward compensator to compensate for the effects of both aforementioned systems. Designing a compensator requires to know the overall transfer function from audio signal (input) and perceived vibration intensity (output). This transfer function is a cascade of two individual transfer functions, $H_d(s)$ (tactile device) and $H_m(s)$ (human mechanoreception system).

$H_d(s)$ was determined by the empirical frequency response[1] of the tactile device (Fig. 2).

[1] A frequency response identified from a physical system via an experiment. We used a pseudo-random noise as input and a nonparametric identification method [8].

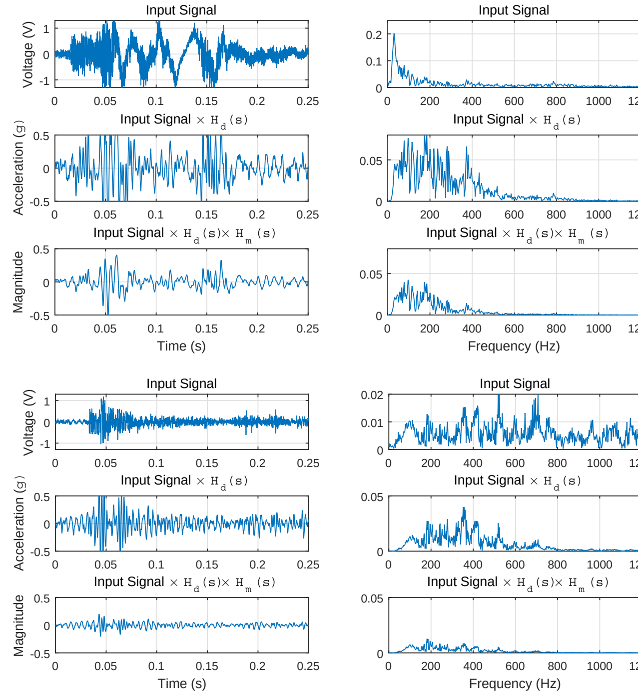

Fig. 4. Conversion process of a gunshot sound. Top: Good vibrotactile effect. Bottom: Insufficient vibrotactile effect.

$H_m(s)$ is estimated based on the critical band theory [9], which suggests that the Pacinian channel filters and integrates the power of stimuli within "critical bands". We are not aware of such readily-applicable theories for the RA 1 channel and do not consider its effects. We designed the critical band filter using the absolute thresholds measured with vertical vibrations applied to participants sitting on a seat [10]:

$$H_m(s) = \left\{ \frac{1}{T(s)} \right\}^2, \tag{1}$$

where $T(s)$ is the absolute threshold as a function of vibration frequency. The peak sensitivity of $H_m(s)$ was normalized to 1 (Fig. 3).

Examples for this conversion process are shown in Fig. 4. The top panel corresponds to a gunshot sound that delivers good vibrotactile effect with direct audio-to-tactile conversion. However, the gunshot sound given in the bottom panel is an example that needs improvement.

The estimated overall transfer function $H(s) = H_d(s)H_m(s)$ is approximated by

$$H(s) = \frac{k(s + \omega_{cut_1})}{(s + \omega_{cut_2})^3}, \tag{2}$$

with $k = 2.44 \times 10^6$, ω_{cut_1} = 20 rad/s, and ω_{cut_2} = 1400 rad/s (Fig. 5). The compensator $C(s)$ is obtained by inverting the overall transfer function H(s):

$$C(s) = \frac{c(s + \omega_{cut_2})^3}{(s + \omega_{cut_1})}, \tag{3}$$

with $c = 4.10 \times 10^{-7}$ (Fig. 5).

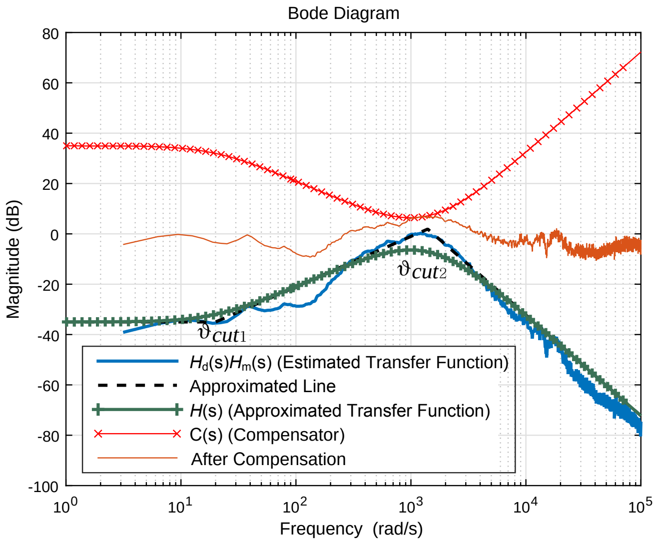

Fig. 5. Bode plots of the software compensator for transparent vibrotactile gunshot rendering of sound.

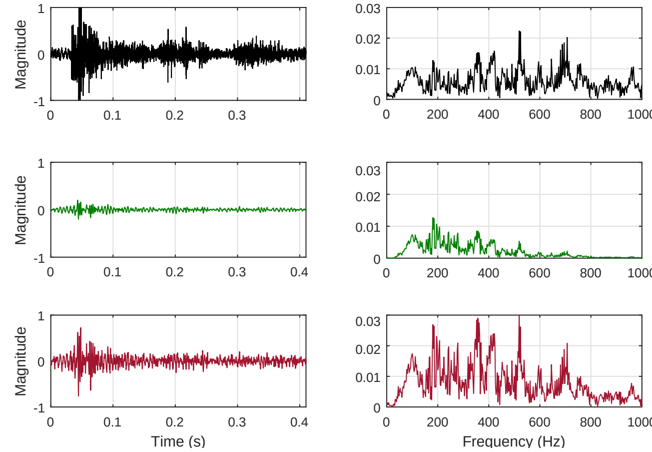

Fig. 6. Example of audio-tactile conversion. Upper panel: original audio signal; middle panel: perceived intensity after conversion without compensation; lower panel: perceived intensity after conversion with compensation.

We further applied a low pass filter with a 1-kHz cutoff frequency to the compensator. This was to remove the high-gain behavior of the compensator in the high-frequency region, which tends to make the final output exceed the actuator limit. Spectral components over 1 kHz are also insignificant for the perception of short gunshot vibrations.

IV. RESULTS AND DISCUSSION

Thirteen gunshot audio signals, which become very weak when they are converted to vibration, were extracted from different movies. Then we measured the output accelerations from the vibrotactile actuator when the audio signals were used as input with or without compensation. Then the measured acceleration spectra were multiplied with $H_m(s)$ to assess their perceived intensity in the frequency domain.

A representative example is shown in Fig. 6. The top panel shows the original audio signal of a gunshot. The middle panel suggests that the perceived intensity of the vibrotactile signal rendered without compensation is quite weak in both the time and frequency domains. In contrast, the bottom panel demonstrates that the perceived intensity of the vibrotactile signal rendered with compensation is much stronger, resembling the original audio signal in both the time and frequency domains. We were able to observe similar results for most of the 13 gunshots.

We are currently evaluating the perceptual performance of the compensation algorithm. Preliminary results indicate that the perceptual magnitude of compensated gunshot vibration is generally comparable to that of sound even for those sound-to-tactile signals that are very weak before compensation. This appears to more clearly express gun blasts via vibrotactile effects.

V. CONCLUSIONS

This work is concerned with direct audio-tactile conversion for gunshot rendering to be used for 4D films. Whereas this method is easy and fast to implement, it may lead to an audio-tactile intensity mismatch that degrades the sense of immersion. As a solution, we studied a vibrotactile rendering method that compensates for the distorted signal passing through the tactile device and the human mechanoreceptive system. Preliminary results demonstrated the effectiveness of the compensation algorithm. Our compensator design might be generally effective for sound-to-tactile automatic conversion in other applications, but its extent needs further investigation. At present, we are working on a perceptual evaluation and plan to further improve the algorithm based on its results.

REFERENCES

[1] J. Lee, B. Han, and S. Choi, “Motion effects synthesis for 4D films,” *IEEE Transactions on Visualization and Computer Graphics*, 2016, accepted for Publication.

[2] M. Karam, C. Branje, G. Nespoli, N. Thompson, F. A. Russo, and D. I. Fels, “The emoti-chair: an interactive tactile music exhibit,” in *Proc. of CHI Extended Abstracts*. ACM, 2010, pp. 3069–3074.

[3] I. Hwang, H. Lee, and S. Choi, “Real-time dual-band haptic music player for mobile devices,” *IEEE Transactions on Haptics*, vol. 3, no. 3, pp. 340–351, 2013.

[4] I. Hwang and S. Choi, “Improved haptic music player with auditory saliency estimation,” *Lecture Notes in Computer Science (EuroHaptics 2014, Part I)*, vol. LNCS 8618, pp. 232–240, 2014.

[5] J. Lee and S. Choi, “Real-time perception-level translation from audio signals to vibrotactile effects,” in *Proc. of CHI*. ACM, 2013, pp. 2567–2576.

[6] J.-M. Lim, J.-U. Lee, K.-U. Kyung, and J.-C. Ryou, “An audio-haptic feedbacks for enhancing user experience in mobile devices,” in *Proc. of ICCE*. IEEE, 2013, pp. 49–50.

[7] R. Okazaki, H. Kuribayashi, and H. Kajimoto, “The effect of frequency shifting on audio–tactile conversion for enriching musical experience,” in *Haptic Interaction*. Springer, 2015, pp. 45–51.

[8] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 1999.

[9] J. C. Makous, R. M. Friedman, and C. J. Vierck, “A critical band filter in touch,” *The Journal of Neuroscience*, vol. 15, no. 4, pp. 2808–2818, 1995.

[10] M. Morioka and M. J. Griffin, “Absolute thresholds for the perception of fore-and-aft, lateral, and vertical vibration at the hand, the seat, and the foot,” *Sound and Vibration*, vol. 314, no. 1, pp. 357–370, 2008.