Haptic Cushion: Automatic Generation of Vibro-tactile Feedback Based on Audio Signal for Immersive Interaction with Multimedia

Y.-J. Cho, S.-U. Kim, M.-C. Joung, J.-J. Lee
LG Electronics, Seoul, Korea

Abstract:
This paper presents a haptic display providing audio-based vibrotactile feedback to enhance the immersive feeling of the user who interacts with multimedia content. The newly developed display has two main features, i) an automatic transformation algorithm and ii) a vibrotactile actuator. The proposed algorithm automatically transforms auditory signals into vibrotactile patterns in real-time by extracting principal frequencies from acoustic unit sequences and superposing vibration waves. The actuator was designed based on the structure of the voice coil linear motor to operate effectively over a wide range of vibration frequencies. Experiments were carried out to evaluate characteristics of the implemented system and demonstrate the effectiveness of the proposed approach0

Keywords: auditory-vibrotactile feedback, automatic generation, extraction, haptics, haptic cushion

Introduction

With advanced human computer interaction technologies, many studies related to haptic feedback have been performed to enlarge the degree of realism and create the immersive sensation in interacting with multimedia contents. As one of the attractive interaction methods, various researches focused on the vibrotatile feedback have been proposed due to its advantages regarding low cost and complexity [1,2,4]. Earlier efforts were to stimulate the skin of human and investigate the perception through mechanoreceptors [8,9] and recent works are focused on how to effectively convey sensation and induce users to be immerged in media contents [1,3].

Although these efforts are ongoing and some approaches have been largely contributed on these fields, we concentrate on the significant issue related with the automatic generation of the vibrotactile feedback for making the vibration effect widely used in the multimedia system. The methodologies for the automatic extraction of haptic feedback from the contents can be mainly classified in accordance with the type of sources: *a) visual information, b) audio information* [3]. Traditional methods to extract the signal from the visual data are based on analyzing the frame image [3,5]. For example, specific features such as facial expressions and motions in images can be detected and matched to the defined vibration signal. However, video based extraction techniques have the difficulty to be in common use because of its robustness and reliability problems[5].

The extraction method based on the audio infor-mation can compensate the limitation of the video based technique. Main approaches have mostly used low or band pass filters to isolate specific frequencies, which can be directly matched with the frequency response of the vibration motor, from the audio signal [3,10]. However, they have problems with limited expressions of vibration feedback because the high frequency range, which has meaningful information of the audio feature, is usually removed.

To overcome these problems, this paper introduces the noble approach to generate automatically vibrotactile effects based on auditory signals. We propose the algorithm to extract principal frequencies from the full band of the sound signal and generate vibration patterns with frequencies sensitive to the human body. We design the vibrotactile actuator and implement the haptic cushion in order to convey immersive vibrotactile sensations to the entire body of the human user. Consequently, the proposed methodology is demonstrated through experiments.

Automatic Transformation Algorithm: Audio to Vibrotaction

The standard range of frequency, which is audible to the human, is 20 to 20,000Hz. The sound created by several major sources such as human voice, instruments have the principal frequencies between approximately 50 to 12,000Hz[6,7]. On the other hand, Morioka et al. [9] experimented absolute thresholds for the human perception of vibration frequency at the seat. According to the research, the thresholds for the human perception of vibration frequency at the seat. According to the research, the sensitivity of the body is high at low frequencies of vibration (e.g. < 200-300Hz). The aim of our proposed is to analyze over the meaningful band of the audio source and transform original signals into vibrotactile patterns with discernable frequencies for the human body.
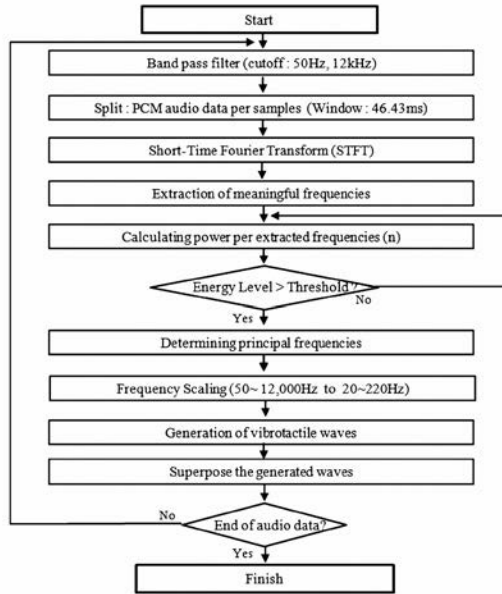
Fig. 1: Flow chart of the control algorithm

Figure1 shows the flow chart of the transformation algorithm. As a first step, The digital band pass filter with two cutoff frequencies 50Hz,12kHz is used to remove noises or meaningless signals. Then, we apply the Short-Time Fourier Transform(STFT), which is well known technique in signal processing to analyze signals as it varies over times, to the filtered signal. There are two reasons why we use this signal processing technique: a) to extract principal frequencies effectively from the unit sequence, b) to progress the algorithm in the real-time. Mathematically, the STFT can be expressed by

$$X(\tau, f) = \int x(t) W(t - \tau) e^{-j2\pi ft} \, dt \qquad (1)$$

where $x(t)$ is the original audio signal in the time domain, $W(t)$ is the window function. In order to satisfy the real time condition and make the generated vibrotactile signal to be perceptibly synchronized with the auditory signal, the length of window function should be sufficiently short. We use the hamming function as the window and set the time duration of the window function with 46.43ms. Next, we define the decision function for extracting principal frequencies, which have the higher power level than the fixed threshold, from the every analyzed unit sequences. The function is designed as

$$r(f_{cp}) = A_f(f_{cp}) * \Delta f, \quad 0 \le f_{cp} \le n \qquad (2)$$
$$f_p = f_{cp} \qquad , \quad r(f_{cp}) > C_{th} \qquad (3)$$

where f_{cp} is the candidate of the principal frequency and f_p is the principal frequency. A_f is the magnitude of f_{cp}. Δf is the half-power bandwidth of f_{cp}, n is the number of extracted frequencies. Determined principal frequencies need to be changed to the suitable frequencies that are easily perceptible by human. For this, we introduce the frequency scaling function as

$$s(x) = B_c \ln\left(\frac{x}{fm}\right) + F_b \qquad (4)$$

where B_c is the scaling coefficient, x is the original frequency in the audio signal, F_b is the frequency offset for the vibrotactile actuator and f_m is the first cutoff frequency used in the bandpass filter. We easily get F_b from the length of the window used in the STFT because the vibration frequency has to be expressed in the unit time. The vibration frequency $s(x)$ is mapped to the original frequency with the logarithmic scale for reflecting the characteristics of the auditory perception.

Finally, we generate sinusoidal waves with scaled vibration frequencies $s(f_p)$ as numerous as m, the number of f_p, in the unit sequence of the original sound. These waves are superposed to produce an output vibrotactile signal $g(t)$ as

$$g(t) = \sum_{k}^{m} P_k z(s, t) \qquad (5)$$
$$z(s(f_{p,k}), t) = r(f_{p,k}) \cos(2\pi s(f_{p,k})t) \qquad (6)$$
$$P_k = \frac{r(f_{p,k})}{\sum r(f_{p,k})} \qquad (7)$$

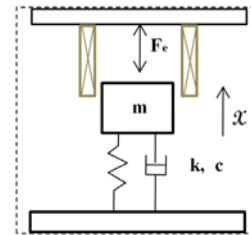**Design of Vibrotactile Actuator**

Fig. 2: Schematic diagram of the actuator system

We designed the vibrotactile actuator based on the structure of the voice coil linear motor. The schematic diagram of the actuator system shown in Figure 2 can be modeled as

$$m\ddot{x} + m\omega^2 x + c\dot{x} = F_e \qquad (8)$$
$$F_e = \frac{\partial W'(x,i)}{\partial x} \qquad (9)$$

where ω is the system fundamental natural frequency given as

$$\omega = \sqrt{\frac{k}{m}} \qquad (10)$$

and W is the co-energy generated by the solenoid coil. The stiffness of the spring and the weight of the mass were decided in order to make the actuator to have the low resonant frequency in the light of the perception level and the sensitivity of the human skin. The overall structure of the designed actuator is assembled with a total dimension of 100 mm x 100 mm x 30mm (W x L x H) (Figure 3).
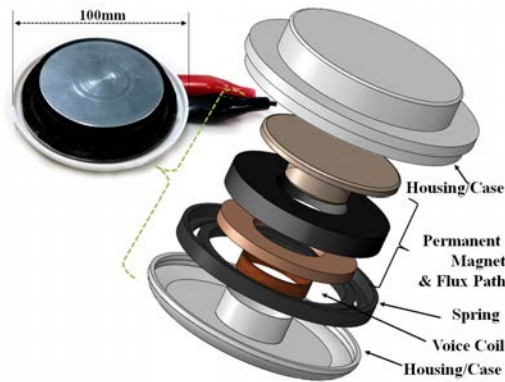
Fig. 3: *Structure of the designed vibrotactile actuator*

Implementation of Haptic Cushion and Experimental setup

Figure 4 shows the implemented haptic cushion that can be operated with the external display for playing audiovisual contents. The designed actuator was installed in the haptic cushion. For the sake of compactness and convenience, we used an Optimus 2X (LG Electronics, Korea) mobile phone as the controller and inserted our algorithm into the "*audio flinger*" in the Android framework to generate vibrotactile signals. Finally, the actuator was controlled by created signals via a Type 2718(Bruel and Kjaer, Denmark) power amplifier.

The measurement system for the evaluation of the proposed algorithm and the developed device is illustrated in Figure 5. The movement of the actuator was monitored by a Type 4374(Bruel and Kjaer, Denmark) 1-axis accelerometer and measured signals were amplified by a Type 2692(Bruel and Kjaer, Denmark) conditioning amplifier. As mentioned previously, automatically generated vibration signals were also amplified by the power amplifier. Each amplified signals were collected through a NI USB 6259 (National Instruments, USA) data acqusition board connected to a host PC.

Results

The performance of the haptic cushion with the automatic generation algorithm was evaluated based on measured informations by the data acquisition board. Figure 6 shows the frequency response plot from the designed actuator for all specimens range of 0-300Hz. The resonant frequency was found at about 90 Hz as we intended to effectively express the vibrotactile feedback with low frequencies.

Figure 7 shows signals from original audio sources in both the time domain and the frequency domain. The result of the STFT analysis is depicted as the center plot. Although many traditional extraction approaches ignored the range of high frequencies, we could observe principal frequencies in the low frequency range(100Hz, 350Hz) as well as the high frequency

range (950 Hz, 1350Hz, 4200Hz) from the unit sequence of the audio at 1.811s - 1.857s as illustrated in the bottom plot.

Figure 8 depicts automatically generated vibro-tactile signals based on original audio signals. We could compare original informations and generated datas in top plots of Figure 7 and 8. In addition, extracted audio frequencies from the original source were effectively transformed into vibration frequen-cies(20 – 220Hz) which are suitable to stimulate the human at the seat (the center plot in Figure 8). The bottom plot shows the acceleration signal of the actuator which is operated by the generated pattern. Although residual vibrations were included in the signal due to the damping of the actuator, the act-uator could convey abundant vibrotactile sensations with various frequencies to the user.

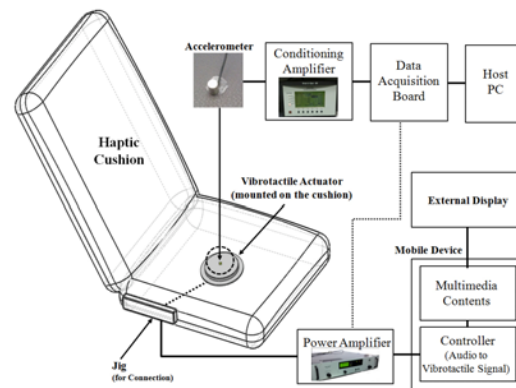**Fig. 4:** *Implementation of the Haptic Cushion*

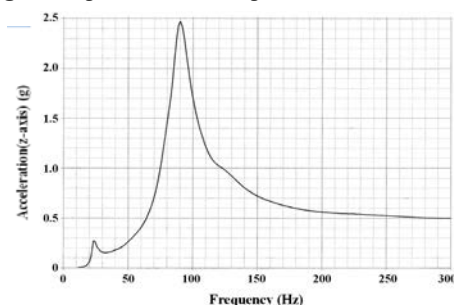**Fig. 5:** *Experimental setup*

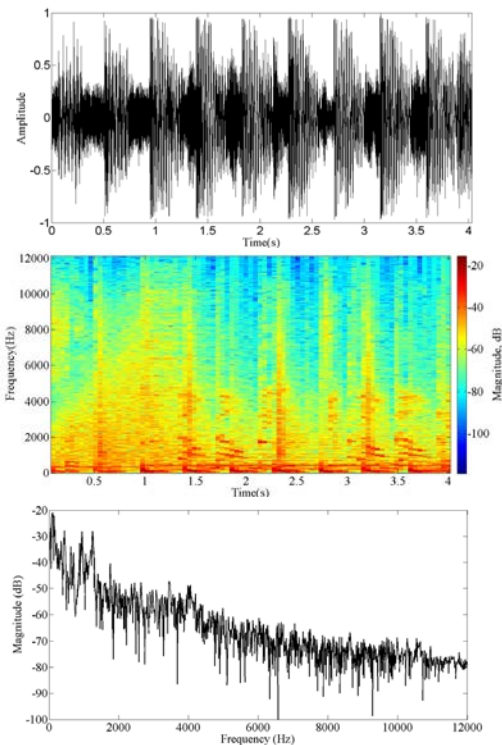**Fig. 6:** *Frequency response of the designed actuator*

Fig. 7: *Analysis of the original audio signal. Top figure: the time-domain signal; Center figure: the result of STFT; Bottom figure: the frequency-domain signal of the unit sequence (1.811s-1.857s).*
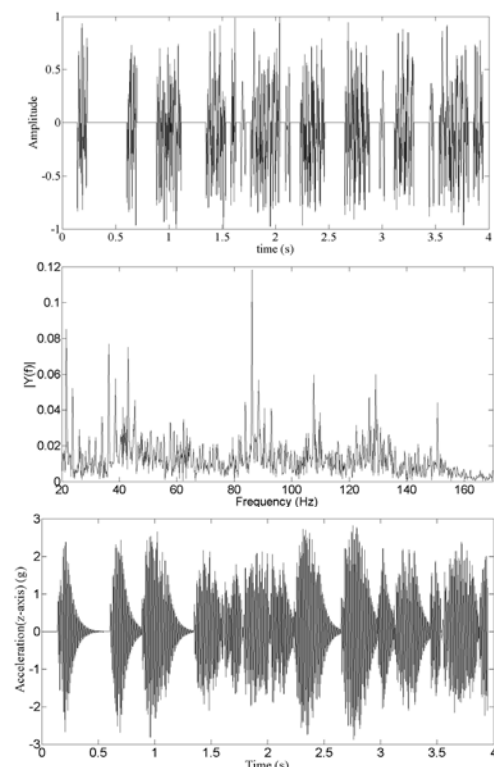

Fig. 8: *Analysis of the automatically generated vibrotactile effects. Top figure: the time-domain signal; Center figure: the frequency-domain signal of the continual sequence(1.45s~2.05s); Bottom figure: the acceleration signal in the time domain*

Conclusion and Future Work

In this paper, we describe the new methodology for the haptic display, which automatically generates the vibrotactile feedback based on the audio signal. The proposed algorithm and the vibrotactile actuator had worked successfully and the implemented haptic cushion could provide abundant vibrotactile effects based on the real-time analysis over the entire audio band of the original source.

In future, we are going to improve the vibration control method for reducing unwanted residual vibrations from the haptic cushion and make our system to convey the much clearer sensation to the user. Furthermore, we will focus on the user study to investigate the human perception with the proposed approach.

References

[1] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf, "Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review", *Psychonomic bulletin & review*, vol.20, pp.21-53, 2013.

[2] D.-S. Kwon, T.-H. Yang, and Y.-J. Cho, "Mechatronics Technology in Mobile Device", *IEEE Industrial Electronics Magazine*, pp.36-41, 2010.

[3] F. Danieaum, A. Lecuyer, P. Guillotel, J. Fleureau, N. Mollet, and M. Christie, "Enhancing Audiovisual Experience with Haptic Feedback: A Survey on HAV", IEEE Trans. Haptics, vol.6, no.2, pp.193-205,2013

[4] Allison M. Okamura, Jack T. Dennerlein, and Robert D. Howe. "Vibration feedback models for virtual environments", In Proc. of the IEEE Int'l Conference on Robotics and Automation, 1998

[5] T.L. McDaniel, K. Kahol, "A Methodology to Establish Ground Truth for Computer Vision Algorithms to Estimate Haptic Features from Visual Image", In Proc. of the IEEE International Conference on Haptic Audio Visual Environments and Their Applications, pp.94-99, 2005

[6] W.B.Snow, "Audible Frequency, speech and noise", Bell System Technical Journal, vo.10, pp.616-627,1931

[7] CE Seashore, "Psychology of music", Music Educators Journal, vo.26, pp31-33, 1938.

[8] Verrillo, AJ Fraioli, RL Smith, "Sensation magnitude of vibrotactile stimuli", Perception & Psychophysics, vol.6,pp.366-372, 1969.

[9] M Morioka, MJ Griffin, "Absolute thresholds for the perception of fore and aft, lateral, and vertical vibration at the hand, the seat, and the foot", *Journal of sound and vibration*, 314(1-2), pp.357-370, 2008

[10] Angela Chang, Conor O'Sullivan, "Audio-Haptic Feedback in Mobile Phones", *CHI 2005*, pp.1264-1267, 2005