Doctoral Thesis

Perceptual Analysis of Vibrotactile Stimuli and Its Application to Vibrotactile Rendering of Music

Inwook Hwang (황 인 욱)

Department of Computer Science and Engineering

Pohang University of Science and Technology

2013

진동 자극의 인지적 분석과 이를 응용한 음악의 진동촉감 표현기법

Perceptual Analysis of Vibrotactile Stimuli and Its Application to Vibrotactile Rendering of Music

Perceptual Analysis of Vibrotactile Stimuli and Its Application to Vibrotactile Rendering of Music

by

Inwook Hwang

Department of Computer Science and Engineering

POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

A thesis submitted to the faculty of Pohang University of Science and Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science and Engineering

Pohang, Korea

June 24, 2013

Approved by

_____

Seungmoon Choi, Academic Advisor

Perceptual Analysis of Vibrotactile Stimuli and Its Application to Vibrotactile Rendering of Music

Inwook Hwang

The undersigned have examined this dissertation and hereby certify that it is worthy of acceptance for a doctoral degree from POSTECH.

06/24/2013

Committee Chair	최 승 문	(Seal)
Member	김 정 현	(Seal)
Member	이 승 용	(Seal)
Member	한 성 호	(Seal)
Member	경 기 욱	(Seal)

DCSE 20065160 황 인 욱 Inwook Hwang, Perceptual Analysis of Vibrotactile Stimuli and Its Application to Vibrotactile Rendering of Music. 진동 자극의 인지적 분석과 이를 응용한 음악의 진동촉감 표현기법, Department of Computer Science and Engineering, 2013, 132P, Advisor: Seungmoon Choi. Text in English

Abstract

Multimodal sensory displays have a great potential in enhancing user experience and task performance. Moreover, haptic displays are being applied to many domains, such as user-interface (UI) components in mobile devices, special effects for entertainment and information delivery in vehicles. However, only simple vibration signals were used without fundamental understanding on their perceptual characteristics. This study investigates the perceptual characteristics of vibrotactile signals on mobile devices and introduces development of haptic music player which can enhance the music experience in mobile device. This research is in line with recent research thrusts aiming at user experience improvements for mobile devices with haptic feedback.

To develop a 'perceptually effective' haptic music player, this study was started from revealing perceptual characteristics of simple sinusoidal vibrations on hand. Perceptual intensities and dissimilarities of various simple vibrations were measured in psychophysical experiments. Effects of four factors, amplitude, frequency, direction, and weight were analyzed. Perceived intensity functions for frequency, amplitude and direction were built from the experimental results and Stevens' power law. Also power relationships between stimulus power and the perceived intensity were shown.

Qualitative characteristics of vibrations were investigated via measurement of dissimilarities and adjective ratings. Through the three experiments we could estimate the two-dimensional perceptual space of simple vibrations with axes of 13 adjective pairs. The two perceptual dimensions that spanned a low frequency range (40–100 Hz) and a high fre-

quency range (100–250 Hz) were close to orthogonal. The low frequency vibrations were felt close to the negative adjectives, such as slow, sparse, blunt, vague, bumpy, jagged, dark, and dull. The perceived feeling of bi-frequency vibrations are shown as similar to low frequency (about 80 Hz) simple vibrations in even mixture of two frequency components.

Then the perceptual characteristics of superimposed bi-frequency vibrations were also studied based on those of simple vibrations. From the intensity matching between various superimposition conditions, Pythagorean summation model was suggested to explain the perceived intensity of bi-frequency vibrations. The bi-frequency vibrations were distinguished from the simple vibrations in the estimated perceptual space, especially when the two components have equal intensities. The effects of three structural factors of bi-frequency vibrations were also analyzed.

We utilized the results of perceptual studies on development of our haptic music player. The initial version of our haptic music player was developed with several distinguished features such as, dual-channel playback, haptic equalizer, perception based modality conversion and scaling, and real-time vibration rendering. The haptic music player was improved with a use of wideband actuator and auditory saliency detection algorithm. User study results for the two versions of haptic music player showed feasibility of this application on mobile device.

This study is targeting at practical use of the results in both of industries and academics. The results of this study will fertilize further studies on perception based vibrotactile rendering in mobile device.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Multimodal sensory displays have a great potential in enhancing user experience and task performance. Currently, visual and auditory displays are standard in majority of consumer electronic devices. Moreover, haptic displays are being applied to many domains, such as user-interface (UI) components in mobile devices, special effects for entertainment, and information delivery in vehicles. Haptic feedback is generally regarded as particularly effective in private environments under sensory overload [13] and as ambient interfaces [68][53].

This research is in line with recent research thrusts aiming at user experience improvements for mobile devices with haptic feedback. Nowadays, quite large portions of multimedia contents are delivered to users via mobile devices. Mobile devices have sufficient computation power and battery to operate haptic actuators. Hence, the merits of multimodal interaction are easily achievable by haptic rendering in a mobile device. Also, many users bring their mobile devices in anywhere they go. The mobile device mostly keep in contact with the user's hand during the use for operation or support of it. Human hand is an organ that controls most tools in our life and sensitive to the high frequency vibrations. Therefore, mobile device can be an effective and practical medium for information transfer using haptic rendering.

There were several attempts in delivering multimedia contents via vibrotactile rendering

to mobile devices. However, those attempts were not supported by mature knowledge of human vibrotactile perception and cognition, with the lack of fundamental data. Thus the vibrotactile rendering cannot be transferred to user in a desired way and it evoked incongruence of multimodal feedback. This study investigates perceptual characteristics of simple and complex vibrations in hand via a mobile device. As an application, we present a development of haptic music player for "vibrotactile music" on mobile device, based on the perception model of vibration.

This study starts with revealing the basic perceptual characteristics of vibration on mobile device. Perceived intensity is one of the most important characteristic to transfer a desired signal to user without distortion. We measured perceived intensities of vibrations and built a mathematical estimation model. Qualitative characteristics were also investigated via estimation of perceptual spaces for simple and complex vibrations. In this step, the perceptual spaces were configured from perceptual distances between vibrations. Adjective ratings for simple sinusoidal and superimposed vibrations were also conducted and analyzed on the perceptual spaces. Through this process, we achieved psychophysical knowledge for a realistic vibrotactile rendering of music. During the study, the first version of dual-mode haptic music player was developed as an application for mobile device and evaluated in a user experiment. Perceptual characteristics of bi-frequency vibration was utilized as the dual-mode rendering in the haptic music player. After all, we improved the haptic music player with auditory saliency-estimation algorithm and a use of wide-band actuator. Perceptual merits of the improved haptic music player were revealed by a user study of the initial and improved haptic music player.

1.2 Contributions

The major contributions of this dissertation are summarized as follows:

- Acquired perceptual data directly adjustable for transparent and expressive vibrotactile rendering in mobile device,
- Revealed perceptual space of simple and bi-frequency sinusoidal vibrations,

- Found qualitative perceptual effects of superimposition of two sinusoidal vibrations,
- Showed feasibility of perception-based haptic music rendering on mobile device, and
- Suggested haptic music rendering algorithm based on auditory saliency model.

1.3 Organization

In Chapter 2, backgrounds are presented with respect to vibration perception and rendering in both mobile devices and other applications. Established perceived magnitude model for simple sinusoidal vibrations in mobile devices is described in Chapter 3. Chapter 4 presents qualitative characteristics of simple vibration, evaluated via estimation of perceptual space with adjective ratings of simple and bi-frequency superimposed vibrations. Chapter 5 introduces perceptual characteristics of bi-frequency superposed vibration in their perceived intensity and perceptual spaces. In Chapter 6, initial development of a perception-based haptic music rendering system is described for its industrial feasibility from the result of an usability test. The development of the improved haptic music player with auditory saliency-estimation is reported in Chapter 7. The conclusion of this dissertaion is in Chapter 8.

Chapter 2

Background

2.1 Perceptual Intensity of Vibration

Detection and magnitude perception of vibrotactile stimuli is a classical research topic in haptics field. However, the complex processes of human perception induced plenty of different studies until now. In general, many factors affect the perceived intensity of vibrotactile stimuli. These include vibration amplitude and frequency, stimulated body site, contact area, stimulus duration, vibration direction, stimulator weight, and age [81, 72]. In the case of mobile devices, however, some of these factors do not need to be considered. For example, hands are the major contact site for interaction with a mobile device. When the hand encloses a mobile device, the contact area between them is so large that the spatial summation of the Pacinian (PC) channel ceases to have effect [72]. Even though short vibrotactile stimuli are often used in applications, stimulus duration is usually set to a larger value (e.g., over 1 s) in psychophysical studies regarding perceived intensity to avoid the temporal summation effect of the PC channel. In this section, we introduce earlier psychophysical studies with their findings related to the perceived intensities of vibrotactile stimuli. As described in detail below, the knowledge of their effects is integral to good actuator and stimulus design.

2.1.1 Vibration Amplitude and Frequency

The perceived intensity of a vibrotactile stimulus increases with its amplitude. Its functional relationship can be characterized by the detection threshold and the growth rate of perceived intensity, both depending on the stimulus frequency. A detection threshold is the smallest signal intensity that can be reliably perceived, and it serves as the reference point of zero perceived intensity [15]. The detection thresholds of vibrotactile stimuli form a U-shaped function of frequency. The minimum threshold appears between 200 and 300 Hz when vibration magnitude is represented by its displacement [34] and between 80 and 160 Hz by its acceleration [60]. An example of vibrotactile detection thresholds measured for a mobile device can be found in [72]. Like the detection thresholds, the rate at which perceived intensity increases with amplitude depends on frequency [60, 72]. This rate of increase can be represented by the exponent of a power function following Stevens' power law [76]. These exponents also exhibit a U-shaped relation against frequency, with a minimum between 150 and 250 Hz for mobile devices [72].

Hence, a psychophysical magnitude function for vibrotactile stimuli can be defined as a mapping from the frequency and amplitude of a vibration to its perceived intensity [15]. This magnitude function (e.g., see Fig. 3.5) can also be transformed to construct equal sensation contours, each of which represents a set of vibration frequencies and amplitudes that result in the same perceived intensity. These perceptual data are useful for the design of effective vibrotactile actuators and stimuli. They can clearly visualize the perceptual consequences of any decisions made on actuator or stimulus design in terms of the strength of the sensation that users would perceive. For example, linearizing the user's perceived intensity using the inverse of a perceived magnitude function can improve the identification of vibrotactile patterns [73].

To obtain the psychophysical magnitude functions that are instrumental for actuator and application development, the functions must reflect the exact real-use conditions of vibrotactile rendering and cover a wide range of associated physical parameters. Otherwise, their immediate practical utility can be significantly undermined. However, thus far, such psychophysical magnitude functions are rarely published, in part because of the relative lack

of their immediate theoretical needs for psychological research and the difficulty of conducting exhaustive psychophysical measurements. To the best of our knowledge, a magnitude function that we presented in [72] is the only available resource for mobile devices. This function was measured using a shaker-type *grounded* actuator that stimulated the hand along the height direction of a mobile device mockup (see Fig. 2.1) in a fashion that is similar to most of the previous studies related to tool-transmitted vibration [57, 69, 60]. The vibration frequencies covered were 20–320 Hz, and the amplitudes were 6–45 dB SL (sensation level; dB above the detection threshold). The perceived intensities were estimated using a standard procedure of absolute magnitude estimation. From the intensity estimates, we built a perceived magnitude function for the two independent variables of frequency and amplitude. We also compared the data with those obtained from the use of other miniature actuators that could provide *ungrounded* vibrotactile stimulation in a limited parameter range. The two kinds of perceived intensities were consistent in their changing trends, i.e., derivatives, but the perceived intensities for the grounded case were considerably higher than those for the ungrounded case.

We suspected that these absolute level differences may have resulted from the difference in *mechanical ground*. A handheld object can be supported mechanically by only the user's hand (ungrounded) or by any other connection to the external environment (grounded). The potential effect of mechanical ground was also suggested by Yao et al. who used a custom-made miniature actuator (now commercialized as Haptuator by Tactile Labs) to configure a mechanically ungrounded condition for mobile devices [84]. Thus, we further investigated this issue and found that the perceived intensity of a mechanically grounded vibration is significantly higher than that of an ungrounded one when all the other conditions are identical [28]. Its exact cause is still unknown, but we presume that tactile suppression, degradation of tactile sensitivity during active movements of our body parts [66], is responsible for it (at least partially). This means that we need a new set of psychophysical magnitude functions measured when the hand supports a mobile device freely in space without any external mechanical connection. This is one of the primary motivations of this study. We also extend our previous study [72] by considering the effects of vibration direction and

device weight.

2.1.2 Vibration Direction

The vibration *direction* of a mobile device is determined by the vibration direction of the actuator and its orientation relative to the device in which it is installed. The three cardinal directions along the width, height, and depth of a mobile device are depicted in Fig. 2.1. We use this definition in the rest of this paper.

Several previous studies examined the effects of vibration direction on detection thresholds and perceived magnitudes. In an early study by Miwa [57], participants pressed their palm onto a vibration table with a large force (5 kg $\simeq$ 49 N). The vibrations were found to produce similar detection thresholds and perceived intensities regardless of their direction (normal or lateral to the palm). Using handles instead of the bare hand did not affect these results. Reynolds et al. also measured detection thresholds and equal sensation curves for a 19-mm diameter aluminum handle held in the hand [69]. Two grip postures (finger and palm grip), two grip forces (8.9 and 35.6 N), and three cardinal vibration directions were considered as independent factors. The detection thresholds showed some dependency on all of the three factors, but no statistical tests were reported. To obtain the equal sensation curves, they used an individual standard stimulus for each experimental configuration. This design did not allow for the comparison of sensation magnitudes between different configurations, therefore no concrete conclusions could be drawn as to the effect of vibration direction. Brisben et al. used a cylindrical handle supported by wires in the air (so close to ungrounded holding) and measured vibrotactile detection thresholds. The results showed lower (but statistically insignificant) detection thresholds for the height direction than for the width direction [6]. Recently, Morioka and Griffin reported a large set of vibrotactile perceived intensities measured using a 30-mm diameter wooden cylindrical tool to study the roles of the four neural channels responsible for tactile perception in magnitude perception [60]. They also used an individual reference stimulus within each experimental configuration, precluding comparisons between absolute intensities measured under different vibration directions.

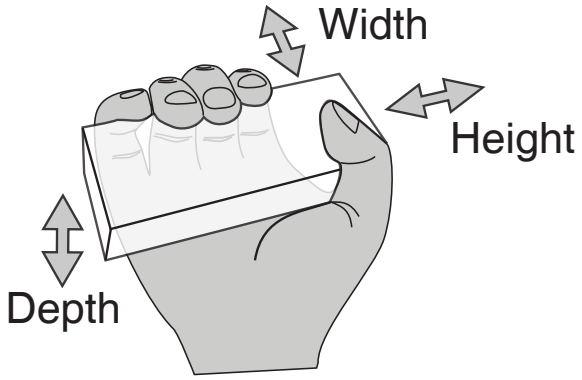

Fig. 2.1 The three vibration directions of a mobile device. In the given grip, the width, height, and depth directions correspond to the proximal-distal, medial-lateral, and ventral-dorsal directions, respectively, relative to the skin in contact with the mobile device.

In summary, these previous studies provided some evidence implicating the dependence of vibrotactile perceived intensity on vibration direction, but the quantitative effect of vibration direction has not yet been measured in terms of absolute measures, which is necessary for the optimal design of vibrotactile actuators and stimuli for mobile applications.

2.1.3 Device Weight

The heavier *weight* of a mobile device applies more pressure onto the user's palm and requires increased physical energy to vibrate the device at the same amplitude. Recently, Yao et al. investigated this issue using ungrounded mobile device mockups with three weights ranging from 50 to 200 g [84]. They observed relative differences in perceived magnitude via pairwise matching with a 110 g mockup as the reference and found a significant influence of device weight.

This important scientific finding needs further attention to determine its implications for actual applications. First, the range of contemporary mobile device weights is much narrower (about 80–150 g). Second, in actuator and device development, the first priority is on the absolute strength of vibration, and this is what users assess more frequently. These two issues were taken into consideration while designing our experiments.

2.2 Percepual Space of Vibration

Perceptual space is one of the most effective ways for visualizing the variations of percepts resulted from changes in the physical parameters of associated proximal stimuli. In the haptics literature, Hollins et al. found a perceptual space for texture perception using dissimilarity rating followed by Multi-Dimensional Scaling (MDS)[1] [25]. They showed that the perceptual space consists of three dimensions, and the pronounced dimensions were hard-soft, rough-smooth, and sticky-slippery, with the last dimension less weighted than the first two [24]. Also using the MDS, MacLean et al. investigated the perceptual characteristics of haptic icons with amplitude, waveform, frequency, and rhythm as design parameters. [54][64][80]. In [54], a custom-made haptic knob stimulating two fingers was used for stimulus generation. Their study demonstrated that stimulus frequency played a dominant perceptual role among a set of time-invariant parameters and that the dissimilarity was maximized in 5–20 Hz vibrations. In addition, the effects of waveform, duration, and amplitude were studied using a laterotactile piezoelectric pin array in [64]. An MDS analysis showed that vibrations were grouped to three clusters depending on their waveforms. Dissimilarities between haptic icons of various rhythms were examined using a piezo-mounted handheld touch screen [80]. In this study, two axes of 'even-uneven' and 'short notes-long notes' were found as prominent perceptual dimensions Kim et al. evaluated the feelings of vibratory stimuli generated by a tactile pin array that stimulated the fingertip in a small contact area [40]. Using adjective rating, they found that the frequency and amplitude of vibration monotonically changed the ratings of such adjectives as rough and prickly. Except this one, there have been few studies as to the qualitative assessment of vibrotactile sensations using a large set of adjectives. Recently, Okamoto et al. found five dimensions on tactile perception through a meta-analysis of the previous studies conducted by other researchers [62]. The dimensions are macro roughness, fine roughness, friction, warmness, and hardness. However, they analyzed studies about real textures and only the macro and

[1] MDS is a statistical technique for the analysis of dissimilarities in data. It determines the coordinates of data points in an N-dimensional Euclidean space while matching the distances between the point coordinates to the original dissimilarities in the data [86].

fine roughness can be accounted for vibration stimuli. The previous study presents some clues for understanding the percepts of vibrations on the perceptual spaces. However, more studies are needed to unveil the entire characteristics of the vibratory perception.

2.3 Perceptual Characteristics of Complex Vibration

Recently, researchers started to extend their attention from the simple sinusoidal vibrations to the complex vibrations. A superimposed signal has multiple spectral components can be made from addition of two or more sine signals with different frequencies. Perceptual characteristics of the superimposed vibrations were recently studied by several researchers [2, 63, 49]. Bensmaia et al. proposed a perceived intensity model for Pacinian mediated vibrations which is the arithmetic sum of the perceived intensity of each component [2]. Dissimilarity between two vibration stimuli was estimated from the difference in their perceived intensities based on the critical band filter theory [55]. With the free parameters introduced in the model, the estimated dissimilarity showed correlation coefficients (R^2) of 0.77–0.79. Muniak et al. was also interested in perceptual intensity of superimposed vibrations [61]. They found that the intensity is proportional to the firing rate of perceptual channel.

Our research group published several studies about qualitative perception of complex vibration. Park and Choi reported the perceptual space of amplitude modulated vibrations [63]. Seven modulation frequencies were tested in 1–80 Hz at 150 Hz carrier frequency. In results, vibrations with a few hertz of modulation frequency were perceived very differently from the pure sinusoidal vibration. The dissimilarity decreased with increment of modulation frequency. Yoo et al. measured consonance for superimposed vibrations in musical octave scale [85]. In a user experiment, the score for consonance increased with the base component frequency and the frequency ratio between two frequency components in a vibrotactile chord. A recent study investigated on the perceived roughness and intensity of bi-frequency vibration [44]. In mixtures of 175 Hz and 210 Hz components, perceived roughness tended to increase as the mixing ratio became closer to 1:1. On contrary, the perceived intensity was minimized at the equal mixing ratio. From another research group, Lim

et al. studied perceptual effect of beat phenomenon occurred by within spectral difference of a superposed signal [49].

In the most of these studies, the intensities of superposed vibrations were calibrated in physical scales. Hence, the vibrations had different perceived intensities and the differences might affected on the percepual dissimilarities. These studies partially unveiled perceptual characteristics of superposed vibration which are distinct from those of simple sinusoidal vibrations. In the current study, we suspect several structural factors that account on percepts of superposed vibrations from the results of these related works, such as intensity ratio between spectral components, component frequencies, within frequency difference in components, and within frequency ratio in components.

2.4 Vibration Actuators for Mobile Devices

During the past decade, many types of vibration actuators have been developed for mobile devices. However, only two types, ERM and LRA, are widely used in consumer mobile devices. ERM is a DC motor that has a rotor with an eccentric mass distribution to induce large centrifugal acceleration. This structure produces 2D vibration in all outgoing directions on the rotational plane of the rotor. ERMs are small and inexpensive. A drawback is that their vibration frequency and amplitude are both determined by input voltage level, so they cannot be controlled independently. This is a serious obstacle against expressive vibrotactile rendering. They also have a slow and nonlinear response with large actuation delays. These problems confine the use of ERMs mostly for alerts. See [72] for further details.

An LRA is a voice-coil actuator with mass and spring components that are connected linearly along the same axis. Vibration is produced by the mechanical resonance of the two components. The vibration direction is parallel to the axis. Since LRA has a fast and linear response, it has been regarded as an adequate choice for touchscreen interaction. However, it has a very narrow frequency bandwidth (only a few hertz wide) centered at the resonance frequency. Thus, the perceptual impression of LRA vibrations is monotone, which precludes rendering of diverse vibrotactile pitches for music playback.

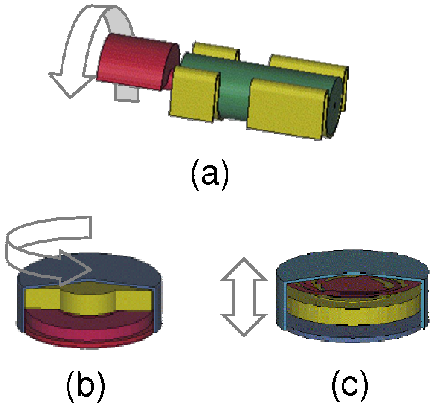

Fig. 2.2 Three types of vibration actuators widely used in commercial mobile devices. The arrows indicate vibration directions. (a) Bar-type ERM. (b) Coin-type ERM. (c) LRA.

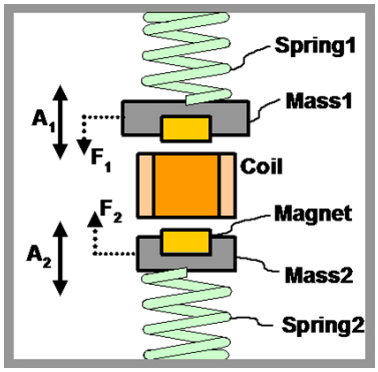

Fig. 2.3 Internal structure of DMA.

Efforts to embed more sophisticated actuators, e.g., piezoelectric and electro-active polymer actuators, into mobile devices are ongoing. However, their industrial adoption has been rare because of stringent industry requirements, such as size, reliability, and durability to external shock.

2.4.1 Dual-mode Actuator for Mobile Devices

DMA is a new vibration actuator developed by LG Electronics [46, 47]. DMA is based on the same working principle with LRA, but it has a more advanced design as shown

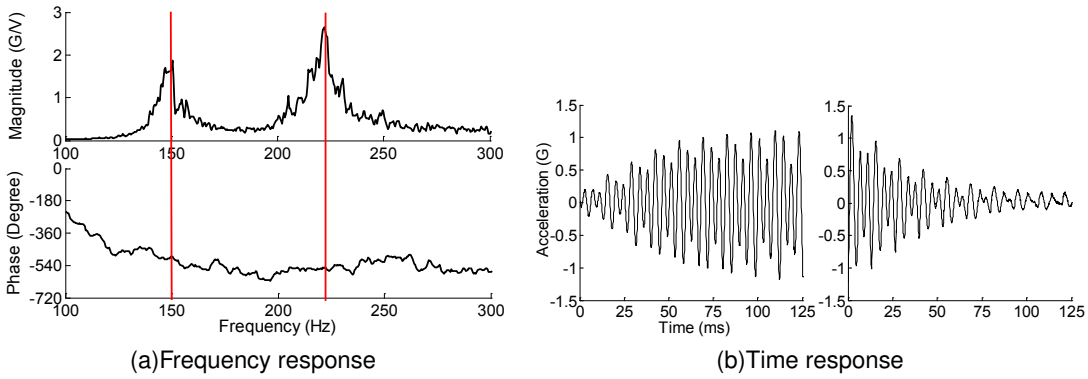

Fig. 2.4 Example responses of DMA.

in Fig. 2.3. DMA includes two built-in mass-spring elements with different resonance frequencies. The two elements share the magnetic field of a common coil located in the center. Each element responds only when the common input to the coil includes spectral energy around its resonance frequency. Thus, a single voltage input with superimposed frequencies can control both frequency components independently. This use of the common coil enables compact size and small power consumption ($10\times10\times3$ mm; 0.1 W max) suitable for mobile devices.

Given the two resonance frequencies, f_1 and f_2, all the mechanical parameters can be determined. As such, DMA has a frequency response with two distinct peaks at the two resonant frequencies. An example response with $f_1 = 150$ Hz and $f_2 = 223$ Hz is shown in Fig. 2.4a. f_2 was selected to be within the frequency range to which humans are most sensitive [34]. To elicit distinctive sensations from f_2, f_1 was chosen as the lowest frequency from f_2 while satisfying the requirements for actuator size and vibration strength. The DMA responses are fairly linear at both f_1 and f_2. Fig. 2.4b shows the time-domain response of DMA for a superimposed sinusoidal input. The rising time and falling time of DMA are around 50 ms and 100 ms, respectively, similar to that of LRA. This prototype model was used throughout this study.

We used the following voltage input $V(t)$ to drive the DMA:

$$V(t) = V_1 \sin(2\pi f_1 t) + V_2 \sin(2\pi f_2 t), \quad (2.1)$$

$$V_1 + V_2 \leq V_{rated}, \quad (2.2)$$

where V_{rated} is the rated voltage of DMA. This superposition rule creates a vibration output that contains the two frequency components, f_1 and f_2. This dual-frequency stimulation is a unique advantage of DMA. Further details on the mechanical design of DMA are available in its patent specifications [46, 47].

2.5 Vibration Rendering in Mobile Devices

In order to create vibrotactile effects, the mobile device makes use of a miniature actuator. Examples that have been commercially adopted include a vibration motor, a Linear Resonant Actuator (LRA), and a piezoelectric actuator. After the early work of Poupyrev [68], persistent research efforts realized embedding a piezoelectric actuator into a commercial mobile phone very recently. Furthermore, multiple actuators can be used together for more diverse vibrotactile effects [39].

Using the actuators, a number of studies have attempted to develop theories and applications for vibrotactile feedback in the mobile device. For instance, Poupyrev and Maruyama designed vibrotactile signals for finger interactions with a touch screen, such as touching down, holding, dragging, and lifting off [67]. Hall et al. suggested the need of GUIs specialized for the mobile device, instead of those adapted from desktop GUIs, and proposed one such GUI named as T-Bar. The T-Bar can prevent the user's finger from occluding visual icons displayed on the touch screen, and the finger movement is guided for icon selection using vibrotactile feedback [21]. In addition, the concept of "haptic icon" [11] or "tacton" [5] has been formalized as a general means for conveying abstract meanings via well discriminable haptic signals. In mobile applications, tactons are usually encoded in vibrotactile signals, and can deliver various information such as the type and priority of a mobile phone alert [8] and the emotion of a sender. See [52] for a recent review on how

to design tactons with high discriminability, learnability, and retainability, along with a low attentional requirement.

In spite of the performance differences of miniature vibrotactile actuators, their output vibrations are commonly of a sinusoidal waveform. Therefore, it is essential to understand the perceptual characteristics of sinusoidal vibrations when they are transmitted to the hand via a mobile device that has a large contact area. Compared to the dynamic research activity oriented to the mobile applications of vibrotactile feedback, studies for the fundamental aspects of perception have been rather lacking.

Recently, our group published preliminary data for the case of a vibration motor attached on the thenar eminence and presented the perceived intensity model of mobile device vibration [72]. The model was set on two control variables: physical vibration amplitude and vibration frequency. Based on this model, perceptually transparent haptic rendering was designed to improve identification performance of haptic patterns [73]. Yao et al. also represented their measured perceived intensities that relatively increases with mobile device weight and decreased with increase of vibration frequency on unvaried physical intensity levels in an acceleration unit [84].

The previous studies used pairwise comparison methods of test and reference stimulus on testing equal sensation level of vibration. Direct comparisons among different conditions were difficult due to the individual reference stimuli for direction or weight factors. In the present article, we extend our previous work by reporting the perceived intensities with two additional factors, including direction and weight. We used absolute magnitude estimation method which relies on a constant scale for a participant's evaluation. Different to other previous studies, closed-loop control of physical vibration magnitude was introduced to guarantee the desired level without the effect of damping on human skin. We also revised the experiment designs and ranges of conditions for practical uses of the results in mobile devices. For instance, who design vibration pattern for notification or alarm on UI components, or who design a new miniature vibration actuator for mobile device are main beneficiaries of this research. They would eager to apply results of this research to improve their perceptional performance.

2.6 Audio-Haptic Rendering

The haptic music player in this study can be classified into multimodal UIs where multiple sensory channels are stimulated for information presentation. Crossmodal icons that combine intuitively similar earcons and tactons are good examples [23]. Another topic, which our work pertains to, is enhancing the user experience of music listening by simultaneous playback of music and cutaneous vibration. Even though the aesthetic values of vibrotactile music have not been elucidated [20], the industry took a rapid step forwards and released several products with this functionality onto the consumer market (e.g., the mobile phone Galaxy S3 and the MP3 player YP-P3 from Samsung Electronics). The TouchSense player for mobile devices by Immersion Corp. is another notable commercial solution. In particular, it has a convenient authoring feature of automatic vibration generation from parsed sound information in a MIDI file or directly from a wave file. However, the basic operations of these commercial products are the same as that of Chang and O'Sullivan [9], where vibrotactile patterns are created from sound signals in a low-frequency bass band. This approach was reported to amplify the sense of beat and improve the perception of sound quality. Recently, Li et al. developed an interesting system, PeopleTones, which notifies the presence of friends in the vicinity via musical and vibrotactile cues played by a mobile phone [48]. A vibrotactile pattern for a song was made by the use of amplitude thresholding and bandpass filtering on the wave file of the song.

Overall, the current research status for the simultaneous playback of audio and vibrotactile music via a mobile device can be regarded as immature. High-performance commercial vibration actuators suitable for this purpose have not been available. We also need signal conversion algorithms that adequately take into account the human perception of both music and tactile vibration. Their feasibility of real-time operation with a mobile device CPU needs to be validated as well. More importantly, the subjective responses of users to this new functionality are largely unexplored.

According to many general haptics literature, transfer of speech into vibrotactile stimuli has been actively studied since the pioneering work of Gault [14] in the early 20th cen-

tury. Bernstein et al. investigated the effectiveness of six speech-to-tactile transformation methods to convey intonation and stress information contained in speech [3]. Brooks and Frost found that lipreading correctness improved from 39% to 88% by the use of a tactile vocoder [7]. Their system was equipped with 16 voicecoil actuators, and each actuator was in charge of conveying the speech information contained in one of the 16 logarithmic equidistant frequency intervals in 200–8,000 Hz. This approach was also taken in [36] to directly transfer music via multiple tactile stimulation sites on the user's torso for the hearing impaired.

2.7 Auditory Feature Extraction

Lastly, we point out that research on computer music has continued to develop algorithms for automatic feature extraction from music. For instance, Schierer proposed a method for extracting the tempo and beat of music by iteratively matching original and reconfigured signals [74]. Several bandpass and comb filters were used to estimate the energy of each frequency band, demonstrating 68% correctness for 60 songs in various genres. To find the beat information of music, Mayor used a filter bank and extracted the energy of each band followed by maximum correlation search with multiple hypotheses of beat speed [56]. Zils et al. extracted percussive sounds from polyphonic music using pattern matching to simulated drum sounds [87]. Jiang et al. suggested octave scale based spectral contrast as an auditory feature for music classification [33].

There were several studies to build human's auditory saliency model. Kayser et al. adopted main structure of visual saliency model by Itti [32] on their auditory model [37]. Three features of intensity, frequency contrast, and temporal contrast were used in their model. They also evaluated their model by detecting salient sounds in noisy environments. Ma et al. implemented a user attention model that included an aural saliency model and speech and music attention models for video summarization [51]. Evangelopoulos et al. also developed an auditory saliency model based on energy separation algorithm for event detection in movie [12]. To calculate the auditory saliency, energy from changes in amplitude and frequency were considered for AM-FM audio signal. They used three features of

mean amplitude, mean frequency, and energy, Several forms of the auditory saliency model were tested and compared for audio event detection [88]. The above auditory saliency models mostly focused on detection of salient events which does not require an exact psychophysical scale.

A recent study conducted by our group has similar approach to the haptic music player in this study [44]. The roughness of sounds was extracted from multimedia contents to present vibration effects in mobile device. Sounds for special effects were detected by the estimated roughness and the vibration were rendered selectively for the special effects. The applicability of computer music algorithms to haptic music, however, is not evident at the moment, and is subject to further investigation.

Chapter 3

Perceived Intensity of Simple Vibration

This chapter presents the methods and the results of two psychophysical experiments that were carried out to find the psychophysical magnitude functions of vibrotactile stimuli transmitted to the user's hand through a mobile device. The magnitude function represents a mapping from the amplitude and frequency of a vibrotactile stimulus to its resulting perceived intensity. We also considered the effects of vibration direction relative to the hand and mobile device weight. In this research, emphasis was placed on providing the perceptual data that can be readily used by actuator and application developers to understand the perceptual strength characteristics of their designs.

In order to remove all the confounding factors of the previous studies presented earlier, our experimental designs were improved in the following three aspects: (1) A miniature actuator was used to produce vibrations in a mechanically ungrounded condition; (2) The absolute magnitude estimation procedure, which does not require the use of a standard stimulus and a modulus, was used to allow for comparisons between different configurations; and (3) A closed-loop amplitude control—robust to individual hand-arm impedance differences and grip force changes— was used to ensure the precise delivery of the target vibration amplitude.

In Exp. I, we examined the effects of vibration direction and device weight on perceived

intensity. The tested vibration directions (height, width, and depth) and the device weights (90–130 g) were determined based on those used in recent mobile devices. Only the vibration direction was found to be a statistically significant factor. Next, in Exp. II, we measured the perceived intensities of vibrations with various amplitudes and frequencies along the three vibration directions. Then, for each direction, a psychophysical magnitude function and equal sensation contours were constructed based on Stevens' power law, which can visualize the consequences of vibration parameter changes on the resulting perceptual strength.

The remainder of this paper is organized as follows. In Section 3.1, general methods common to both experiments are presented. The experimental methods used and the results obtained from Exps. I and II are reported in Sections 3.2 and 3.3, respectively, followed by a general discussion in Section 3.4.

3.1 General Methods

In this section, we describe the experimental methods used commonly in Exps. I and II.

3.1.1 Apparatus

A miniature linear vibration actuator (Tactile Labs; Haptuator TL002-14-A; weight 12.5 g) was used to produce vibrations in both experiments. This actuator has a much wider bandwidth (50–500 Hz) with a weak resonance around 60 Hz and a stronger output range than standard mobile device actuators. It was controlled by a computer via a 16-bit data acquisition (DAQ) board (National Instruments; model PCI-6251) at a 20-kHz sampling rate. Output commands from the DAQ board were passed through a custom current amplifier and then relayed to the Haptuator. The Haptuator was attached to a mobile device mockup with by means of an adhesive rubber tape. A three-axis accelerometer (Kistler; model 7894A500; weight 7.5 g) was fastened to the top of a mockup, and its output for each vibration axis was sent to the computer through the DAQ board.

We prepared three mobile device mockups made of acrylic resin. They all had the same width (45 mm) and thickness (15 mm), but different heights (90, 115, and 141 mm) to make

different weights (70, 90, and 110 g, respectively). Even the height of the shortest mockup exceeded the palm width of our participants, thus the height differences did not alter the contact site or area. In each mockup, the Haptuator could be attached to each of its three faces for vibration production along each of the three directions shown in Fig. 2.1. Including the Haptuator and the accelerometer, the total moving masses of the three mockups were 90, 110, and 130 g, respectively. This weight range covers the majority of recent mobile devices.

To avoid for the Haptuator to directly contact the hand, the Haptuator could not be attached at the center of each face. We confirmed via pilot experiments that this eccentricity did not incur noticeable changes in the vibrotactile perceived intensity.

3.1.2 Stimuli

As stimuli, 1.2 s long sinusoidal vibrations reconstructed at 20 kHz were used. In this paper, their amplitudes are represented in terms of an industry standard acceleration unit ($G \approx 9.8\,m/s^2$; the gravitational constant) because of the practical motivations of this work, unless specified otherwise. The acceleration amplitude of a simple sinusoidal vibration can be easily converted to a displacement or velocity unit that is more frequently used in academic literature. In addition, a closed-loop PD control was used for the accurate generation of vibration amplitude, despite the individual-dependent hand-arm mechanical impedance and other time-varying error sources such as grip force changes. Through the control, the overshoot and the steady-state error was bounded by $\pm 1.5\%$. See Appendix in supplementary materials for further details.

3.1.3 Procedures

In both experiments, participants placed a mockup on their right palm while resting their wrist on a sponge support block attached onto a table, as shown in Figs. 3.1 and 3.3. Their right palm was held upwards while perceiving the stimuli.

In each trial, participants reported the perceived intensity of a vibration by an absolute magnitude [15], i.e., a positive number in their own internal scale without a reference stim-

ulus. This method of absolute magnitude estimation is free from the problem of modulus bias [15]. It also allows us to compare perceived intensities collected with different vibration directions and device weights. The participants entered the absolute magnitudes onto a keypad using their left hand. The next trial began 3 s after the response.

Before the experiments, the participants were given the standard instructions of absolute magnitude estimation [15]. During the experiments, their view of the mockup was blocked by a cloth curtain so that the height of the mockup would not cause a response bias. Sound cues were also prevented by earplugs and headphones that played white noise. The level of the white noise was as low as the auditory noise we experience in our daily life, so that the participants could remain comfortable and concentrate on the experimental tasks.

3.1.4 Data Analysis

The absolute magnitude estimates collected in each experiment were standardized separately using the standard procedures [22]. For each participant, multiple magnitude responses were averaged per experimental condition using a geometric mean. Following which, a subjective geometric mean M_p over all of the conditions was calculated for each participant, as well as a grand geometric mean M_g across all conditions and participants. Finally, a scaling constant per participant, $M_n = M_g/M_p$, was multiplied to each participant's scores. All further analyses were performed using these normalized perceived intensities.

3.2 Exp. I: Effects of Vibration Direction and Device Weight

This experiment evaluated the significance of vibration direction and device weight on the perceived intensity of mobile device vibration. The results obtained were subsequently taken into account for the design of Exp. II.

Table 3.1 Experimental conditions of Exp. I.

Weight (g)	Direction	Frequency (Hz)	Acceleration Amplitude (G)
90	Depth	60, 120, 250	0.2, 0.8
130			
110	Width		
	Height		

3.2.1 Methods

Participants

Twenty-four participants (20 males and 4 females; 18–28 years old with a mean of 20.9) participated in this experiment. They were everyday users of mobile devices and had no known sensorimotor impairments, both by self-report. Each participant was paid 10,000 KRW (about 9 USD) after the experiment.

Experimental Conditions

The experiment comprised five sessions, as shown in Table 3.1. The baseline session used vibrations transmitted through the 110 g mockup along its depth direction. Two other sessions were to evaluate the effects of vibration direction, and their vibrations were directed along the width and height of the 110 g mockup, respectively. The other two sessions were to compare the effects of device weight. Vibrations were rendered using the other two mockups (90 and 130 g) along the depth direction. In each of the five sessions, the perceived vibratory magnitudes were collected at two amplitudes of peak acceleration (0.2 and 0.8 G) and three frequencies (60, 120, and 250 Hz), resulting in a total of 30 experimental conditions.

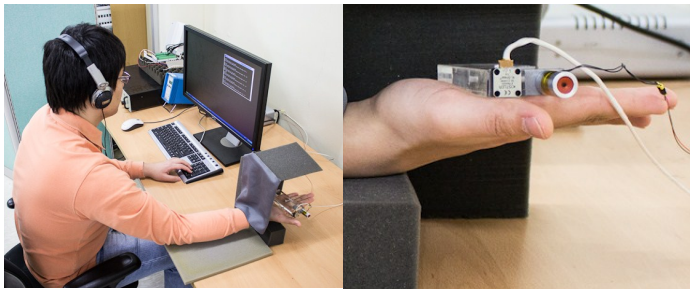

Fig. 3.1 Participant's posture used in Exp. I.

Procedures

Each of the five sessions consisted of five blocks of trials. In each block, the 6 vibrations (3 frequencies $\times$ 2 amplitudes) were presented in random order. The first block was used for training, and its results were discarded. Thus, the perceived intensity of each vibration used for data analysis was measured four times per experimental condition. The order of the five sessions was randomized per participant. Each session lasted for about 4 min. The participants were given 3 min of rest at the end of each session and were allowed to take additional rests whenever needed. The entire experiment took approximately 40 min for each participant.

During this experiment, the participants' right hand was fully open to eliminate any potential effect of grip force. In general, grip force can significantly affect the sensitivity and intensity of vibrotactile perception [6]. In our previous study, the individual grip forces for a mobile device showed a large variance (1.75 ± 1.42 N) [72, Appendix]. Moreover, the user's grip force can change over time. Hence, the open grip posture was used to isolate the effects of vibration direction and device weight without the influence of grip force, even though this posture is not natural for grasping a mobile device. In the main experiment of this work, Exp. II, we did use a closed grip to obtain more practically meaningful data. Note that Morioka and Griffin showed similar vibration sensitivities between the open and grasping postures in the depth direction [59].

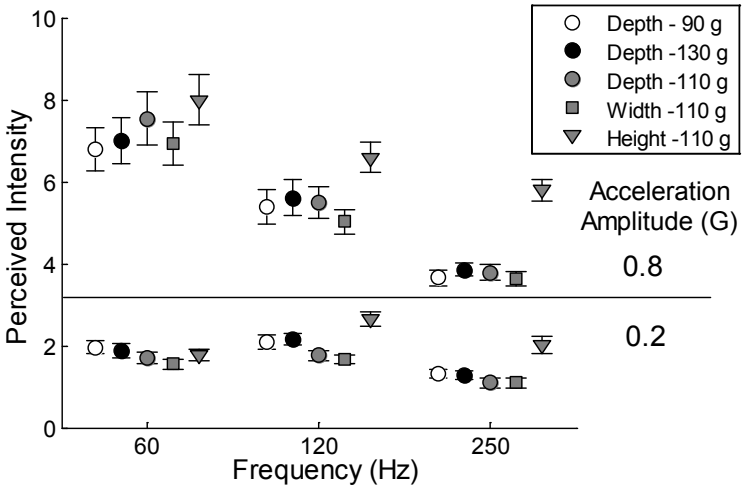

Fig. 3.2 Mean perceived intensities measured in Exp. I. Error bars represent standard errors.

3.2.2 Results

Fig. 3.2 presents the mean perceived intensities of the five direction-weight conditions obtained in Exp. I. To assess the statistical significance of vibration direction and device weight, a three-way within-subject ANOVA with independent variables of direction-weight, amplitude, and frequency was performed on the perceived intensities measured in the five sessions. Combining vibration direction and device weight into one factor was necessary because of their unbalanced design. The combined direction-weight factor showed strong significance ($F(4,92) = 11.2$, $p < 0.0001$), as well as both amplitude and frequency. Interactions of direction-weight with amplitude and frequency were also significant ($F(4,92) = 6.63$, $p < 0.001$, and $F(8,184) = 2.72$, $p = 0.0074$, respectively).

For post-hoc analysis, we used pairwise comparisons with the Bonferroni correction between the five direction-weight conditions. Among the three direction conditions (Depth–110 g, Width–110 g, and Height–110 g), the differences of Height–110 g with Depth–110 g and Width–110 g were statistically significant ($t(92) = 4.70$, $p < 0.0001$, and $t(92) = 6.05$, $p < 0.0001$, respectively), while that between Depth–110 g and Width–110 g was not. The differences between the three weight conditions (Depth–90 g, Depth–110 g, and Depth–

130 g) were not statistically significant. This means that the vibrations in the height direction were perceived to be stronger than those in the other two directions, as can be seen in Fig. 3.2, while the device weight did not affect the perceived intensity noticeably. These multiple comparison results were the same in the tests done with the data of each amplitude (0.2 and 0.8 G). The same was true for each of the three frequencies, except that for the 60 Hz vibrations, the effect of vibration direction was significant only between the height and width conditions.

3.2.3 Discussion

As reviewed earlier in Section 2.1, no previous study has clearly demonstrated the effect of vibration direction on the absolute level of perceived intensity. Our study measured perceived intensities against a condition-independent, participant's internal scale and was able to show significant differences in the perceived intensity among the three cardinal vibration directions. We note that the same effect of vibration direction was also observed in Exp. II. The underlying reasons for that, however, are still to be investigated.

For the effect of weight, our result may appear inconsistent with the findings of Yao et al. [84], which demonstrated the substantial effect of mobile mockup weight on the perceived intensity of vibration. This difference can be attributed to the differences in experimental methods. As reviewed in Section 2.1.3, Yao et al. relied on the relative comparisons between stimuli (as opposed to the absolute judgments used in our study) in a much wider weight range (55–200 g; in comparison to the 90–130 g used in our study). Thus, the methods used by Yao et al. [84] could have had much higher discriminability. In contrast, our results suggest that, in real life scenarios, users are unlikely to distinguish the differences in absolute vibration strength caused by the relatively small weight differences of contemporary mobile devices.

3.3 Exp. II: Perceived Intensity Model

Exp. I confirmed the statistical significance of vibration direction on the vibration intensities perceived via a mobile device grasped in the hand. The device weight was shown to

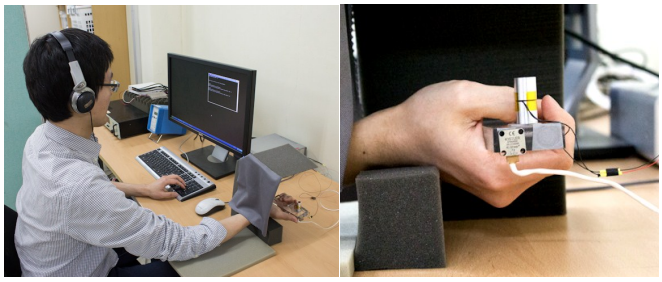

Fig. 3.3 Participant's posture used in Exp. II.

be rather insignificant in the range tested. Hence, in Exp. II, we measured the perceived intensities of various vibrations produced along each of the three cardinal directions of a mobile device. Psychophysical magnitude functions were then constructed for each vibration direction.

3.3.1 Methods

Participants

Twenty-four participants (19 males and 5 females; 18–26 years old, with a mean of 21.6) participated in this experiment. All participants reported that they were users of mobile devices and they had no known sensorimotor impairments. Each participant was paid 15,000 KRW (about 13 USD) after the experiment.

Experimental Conditions

Exp. I demonstrated the strong significance of vibration direction on perceived intensity, while no significant effect was found for small changes in device weight. Thus, Exp. II was conducted with vibration amplitude, frequency, and direction as independent variables, while the device weight was fixed using the 110 g mockup. For each of the three vibration directions, vibrations were tested for six frequencies (60, 80, 120, 180, 250, and 320 Hz) and six amplitudes, resulting in a total of 108 experimental conditions (see Table 3.2). The amplitudes for each of the six frequencies were selected to be within the amplitude range that could be generated at that frequency with the Haptuator.

Table 3.2 Vibration amplitudes (G, peak) used in Exp. II.

Frequency (Hz)	Direction		
	Width	Height	Depth
60	0.2, 0.5, 0.8, 1.1, 1.4, 1.7	0.1, 0.2, 0.3, 0.4, 0.5, 0.6	0.2, 0.4, 0.6, 0.8, 1.0, 1.2
80	0.2, 0.5, 1.0, 1.5, 2.0, 3.0	0.2, 0.5, 0.8, 1.2, 1.8, 2.5	0.2, 0.5, 1.0, 1.5, 2.0, 3.0
120	0.2, 0.5, 1.0, 1.5, 2.0, 3.0	0.2, 0.5, 0.8, 1.1, 1.4, 1.7	0.2, 0.5, 1.0, 1.5, 2.0, 3.0
180	0.2, 0.5, 0.8, 1.2, 1.8, 2.5	0.2, 0.4, 0.6, 0.8, 1.0, 1.2	0.2, 0.5, 0.8, 1.2, 1.8, 2.5
250	0.2, 0.5, 0.8, 1.2, 1.8, 2.5	0.1, 0.2, 0.3, 0.45, 0.65, 0.9	0.2, 0.5, 0.8, 1.2, 1.8, 2.5
320	0.2, 0.5, 0.8, 1.2, 1.8, 2.5	0.1, 0.2, 0.3, 0.45, 0.65, 0.9	0.2, 0.5, 0.8, 1.2, 1.8, 2.5

Procedures

The experiment consisted of three sessions, and each session used vibrations along one of the three cardinal vibration directions. Each session comprised four blocks of trials. In each block, 36 vibrations (6 frequencies $\times$ 6 amplitudes) were presented in random order. The first block was used for training, and its results were discarded. Thus, the perceived intensity of each vibration used for data analysis was measured three times.

Fig. 3.3 shows the posture used by the participants in the experiment. To simulate the real-world use of mobile devices, the participants lightly grasped the mockup in their right hand during the experiment. We also tested in pilot experiments whether this postural change can affect the significance of vibration direction on the perceived intensity observed in Exp. I, but we did not find any noticeable difference.

Each experimental session lasted for approximately 25 min. The participants were given a 3 min rest at the end of each block and were allowed to take additional rests if they felt tired. For each individual, the entire experiment took approximately 1.5 h to complete.

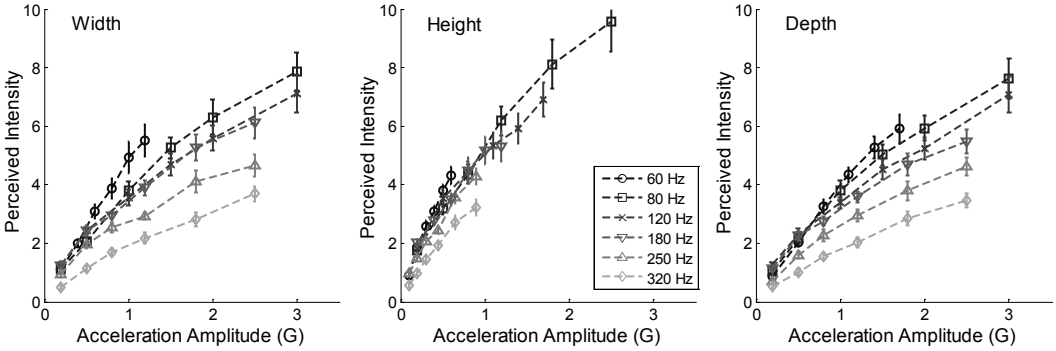

Fig. 3.4 Mean perceived intensities and their standard errors for all the conditions of Exp. II.

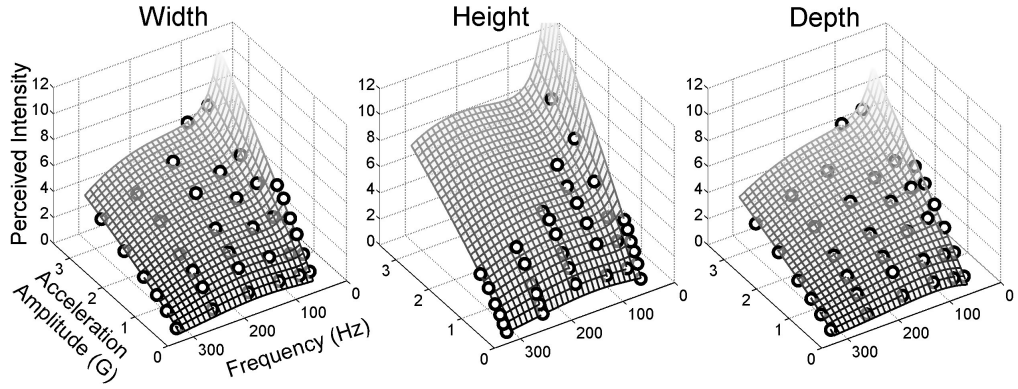

Fig. 3.5 Mean perceived intensities (circles) and the best fitting surfaces for the three directions obtained in Exp. II.

3.3.2 Results

The mean perceived intensities for the three cardinal directions are shown along with their standard errors in Fig. 3.4. In each plot, the perceived intensities of the same frequency increased with the amplitude, while the steepness of the increase decreased with the amplitude. Given the amplitude, the perceived intensities decreased with the frequency, with more salient effects in the low frequency region (< 100 Hz). In addition, the perceived intensities along the height direction were larger than those along both the width and depth direction, and the effect of frequency was much weaker with the height direction.

We built a psychophysical magnitude function that maps vibration frequency and ampli-

Table 3.3 Coefficients of the psychophysical magnitude function.

Direction	i	0	1	2	3
Width	α_i	267	-373	177	-27.8
	β_i	-2.28	8.04	-5.48	1.10
Height	α_i	358	-493	229	-35.5
	β_i	8.14	-6.73	1.36	0.0611
Depth	α_i	90.2	-128	63.9	-10.7
	β_i	23.7	-28.6	11.6	-1.54

tude to perceived intensity for each vibration direction as follows. Our magnitude function is based on Stevens' power law:

$$\psi = k\phi^e, \tag{3.1}$$

where ψ is the perceived intensity for the stimulus of the physical amplitude ϕ (in our case, the vibration amplitude in acceleration). We regressed the slope k and the exponent e using the following equations:

$$k = \sum_{i=0}^{3} \alpha_i (\log_{10} f)^i \quad \text{and} \quad e = \sum_{i=0}^{3} \beta_i (\log_{10} f)^i, \tag{3.2}$$

where f is the vibration frequency in Hz. The third-order polynomial functions describing k and e were chosen empirically to obtain the best fits with sufficiently high R^2 values (> 0.99), as in our previous study [72]. The values of α_i and β_i found by this procedure are shown in Table 3.3. These psychophysical magnitude functions are graphically depicted in Fig. 3.5. The plots show smooth valleys at around 150 Hz, where the lowest detection threshold appears [60].

3.3.3 Discussion

Equal Sensation Contours

In our context, an equal sensation contour represents the frequency and amplitude of vibrations that lead to a given perceived intensity. The equal sensation contours derived from our

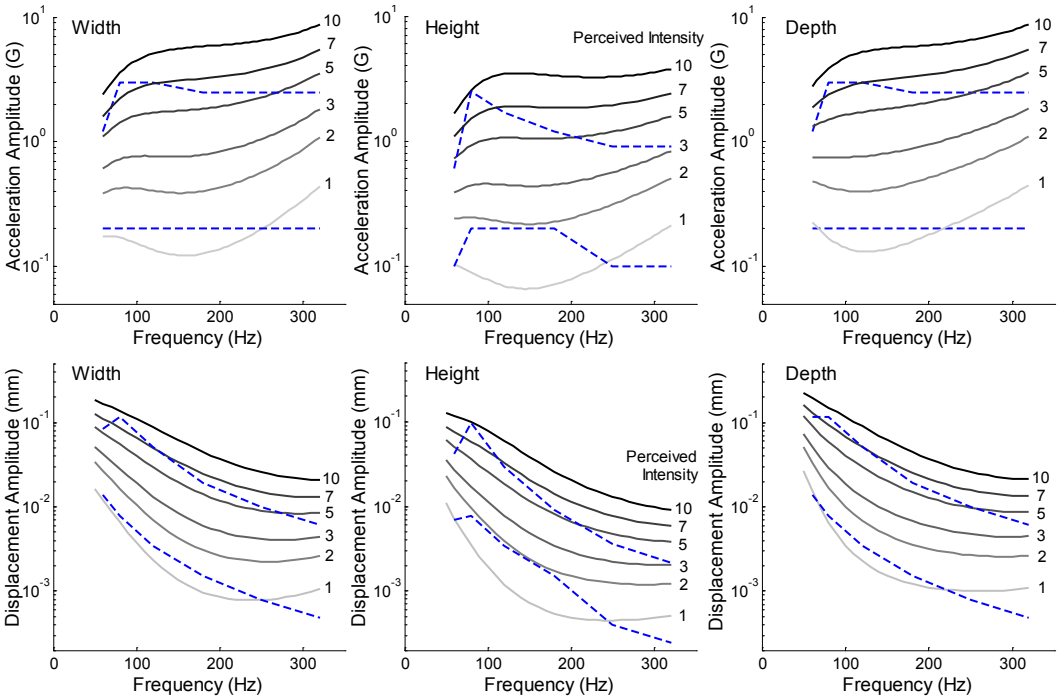

Fig. 3.6 Equal sensation contours for each vibration direction. Vibration amplitudes are represented in acceleration (upper panels) or displacement (lower panels), where acceleration amplitude = displacement amplitude $\times (2\pi f)^2$. Note the use of a logarithmic scale in the ordinates. Dashed lines represent the lower and upper limit of amplitude used in Exp. II. The contours outside these bounds should be considered extrapolated values.

perceived magnitude functions are shown in Fig. 3.6. The vibration amplitudes were represented in terms of acceleration (upper panel) or displacement (lower panel). The dashed lines represent the lower and upper bounds of the parameter range used in the experiment.

The equal sensation contours form a U-shaped curve at low perceived intensity levels less than 2 (approximately 0.4 G in acceleration amplitude). This is similar to the typical detection threshold curve of the PC channel, which has lower thresholds than the other Non-Pacinian (NP) channels in the frequency range tested in our experiment. At higher intensity levels, other NP channels such as the RA and SA1 channels also respond. For example, the detection threshold of the RA channel is approximately 0.1 G in acceleration for 125 Hz vibrations and increases with frequency [58]. In this range, each individual channel con-

Table 3.4 Experimental methods of related studies about the perceived intensity of hand-transmitted vibrations.

Reference Source	Ryu2010[72] Exp. 2	Morioka2006[60] Exp. 2	Verrillo1975[82] Without Surround
Stimulus Medium	Cell Phone	Cylinder	Small (0.29 cm^2) Contactor
Vibration Actuator	Desktop Shaker	Floor Shaker	Desktop Shaker
Test Variables	Amplitude and Frequency	Amplitude, Frequency, and Direction	Amplitude and Frequency
Frequency Range (Hz)	20–320	8–400	60, 250
Amplitude Range (G)	0.02–2	0.02–14	0.01–1.3
Experiment Method	Absolute Magnitude Estimation	Magnitude Estimation	Absolute Magnitude Estimation

tributes to the intensity function [4, 58]. Our equal sensation curves were consistent with this general tendency. They monotonically increased with vibration frequency, sometimes including a flat region, similar to the equivalent comfort contours reported by Morioka and Griffin [60].

The perceived magnitude functions and the equal sensation contours are instrumental in designing vibrotactile actuators or stimuli for mobile devices. To design vibration actuators, a number of physical constraints, such as maximum output amplitude, frequency bandwidth, size, and power consumption, need to be taken into account. The data shown in Figs. 3.5 and 3.6 can effectively visualize the effects of such trade-offs on perceived intensity for optimal actuator design. Furthermore, given a vibration actuator, we can map its frequency and amplitude range, which is likely to be a subset of our tested parameter range, to perceived intensity. This perceptual information can help designers choose vibrotactile stimuli with the frequency and strength appropriate to their goals.

Table 3.5 Exponents of Stevens' power law representing the rate of sensation growth.

Reference	Direction	20 Hz	40 Hz	60 Hz	80 Hz	120 Hz	180 Hz	250 Hz	320 Hz
Our Exp. 2	Width	.	.	0.87	0.76	0.64	0.60	0.66	0.77
	Height	.	.	0.83	0.69	0.59	0.60	0.68	0.80
	Depth	.	.	0.92	0.74	0.62	0.64	0.71	0.78
Ryu2010 [72]	Height	0.72	0.68	$0.61^{\top}$	0.55	$0.54^{\top}$	$0.54^{\top}$	$0.55^{\top}$	0.58
Morioka2006[60]	Width	0.67	0.62	$0.55^{\top}$	0.48	$0.40^{\top}$	$0.37^{\top}$	0.53	$0.23^{\top}$
	Height	0.81	0.52	$0.44^{\top}$	0.46	$0.46^{\top}$	$0.38^{\top}$	0.44	$0.32^{\top}$
	Depth	0.66	0.52	$0.45^{\top}$	0.38	$0.36^{\top}$	$0.31^{\top}$	0.41	$0.32^{\top}$
Verrillo1975[82]	Depth	.	.	0.40	.	.	.	0.35	.

$^{\top}$: These values are interpolated from neighbors.

Comparison with Previous Studies

To compare the major results of our study with previous related studies, the experimental designs of three related studies, [72, 60, 82], are summarized in Table 3.4. They are all concerned with vibrotactile intensity perceived by the hand. The rightmost column of Table 3.4 is taken from [82], where a small-area contactor stimulated the thenar eminence on the hand.

The psychophysical magnitude functions of the present and three previous studies showed similar shapes. In Table 3.5, we compared the exponents of Stevens' power law, which represent the rates of sensation magnitude growth, for eight frequencies between 20 and 320 Hz. The exponents measured in [72, 60], and this study all followed a U-shaped curve, with the minimum found between 120 and 180 Hz. In addition, all the exponents were less than 1.0, indicating that perceived intensity is eventually saturated with the increase in amplitude.

A notable difference is that the exponents of this study are considerably larger than those of the other studies, even compared with our previous study [72] that used experimental methods most similar with those of the present study. This may be due to the use of mechanically ungrounded holding posture in this study. As demonstrated in our recent study [28], the ungrounded condition leads to significantly lower perceived intensity than the

grounded condition, possibly because of the tactile suppression caused by active muscle movements required in the ungrounded condition. The lower device mass (90–130 g) could also have contributed to the lower perceived intensities, as the weight difference (201.7 g in [72]) was quite large [84]. Psychophysical magnitude functions with lower absolute perceived intensities generally show steeper increases with vibration amplitude [83].

3.4 General Discussion

3.4.1 Stimulus Context Effect

The absolute magnitude estimation procedures used in both Exps. I and II were identical, enabling direct comparisons between the results of the two experiments. For example, the perceived intensity of a 120 Hz, 0.2 G vibration along the height direction was 2.82 in Exp. I and 1.82 in Exp. II, with about a 35% difference. At 0.8 G, the difference ratio was similar to the perceived intensities 7.01 and 4.46 (a 36% difference), respectively. This tendency that the perceived intensities of Exp. I were larger than the corresponding intensities of Exp. II was common to all the vibrotactile stimuli used in both experiments.

This scale difference may have resulted from two possible sources. First, the two experiments had different contact conditions. In Exp. I, the participants' hand was fully open, limiting the contact area with the mockup to the palm only. In Exp. II, the mockup was lightly grasped, also allowing contact with the fingers. However, the increased contact area in Exp. II cannot account for the lower perceived intensities in Exp. II compared to those in Exp. I. This is because a larger contact area generally increases perceived intensity due to the spatial summation of the PC channel until the summation is saturated [59]. Hence, the different contact postures cannot be a reason for the scale difference between the two experiments.

A more plausible explanation is the stimulus context effect present in magnitude estimation procedures [15]. Even though the absolute magnitude estimation used in our experiments shows relatively weaker stimulus context effects, the same stimuli evaluated with other lower amplitude stimuli can be scaled to be higher than those evaluated with higher

amplitude stimuli [18]. The amplitude range of Exp. II was larger than that of Exp. I, and this seems to have caused the scale difference described above.

In actual applications, the range of vibration amplitudes that can be produced by miniature actuators in mobile devices is expected to be smaller than the range tested in Exp. II. In this case, for the same stimuli, the lower amplitude context of actual mobile devices is expected to result in subjective intensities distributed in a wider range than those reported in this study. This means that to use the psychophysical magnitude functions shown in Fig. 3.5, the corresponding range of perceived intensities should be stretched to some degree.

Use of the free-modulus scale with a reference stimulus can reduce the effect of stumulus context. For the vibration stimuli, many participants seems to do not have a firm internal scale of intensity perception. Therefore, a reference stimulus would be helpful to hold their reporting scale constantly.

3.4.2 Physical Metrics

Our vibratory magnitude functions were estimated from psychophysical measurements, which generally require a great amount of time and effort. A more convenient and practical approach is predicting the perceived intensity of a vibrotactile stimulus from its physical parameter using a physical metric. For example, it has been demonstrated in [55, 2] that for a sinusoidal vibration with acceleration $a(t) = A\sin(2\pi f_s t)$, its efficacy in simulating the Pacinian system can be represented by the sinusoid's spectral power weighted by the detection threshold $T(f_s)$:

$$P_P = \left(\frac{A}{T(f_s)}\right)^2. \tag{3.3}$$

Fig. 3.7 shows these Pacinian-weighted powers against the perceived intensities measured in Exp. II. The computation used the detection thresholds measured with a mobile device [72]. A good fit was found with a quadratic function for each vibration direction ($r^2 = 0.88$–0.95). Hence, this metric, P_P, can be used as an adequate metric for perceived intensity. However, it should be noted that the validity of the Pacinian-weighted power for low-frequency (e.g., 60 Hz) vibration remains in question, since those vibrations are also

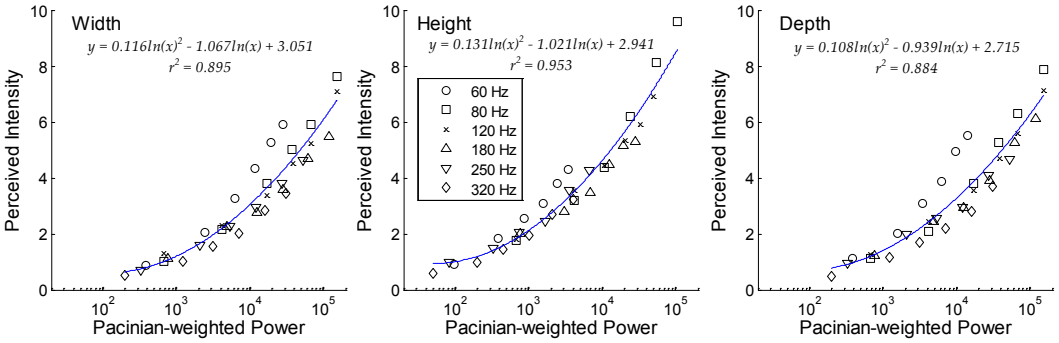

Fig. 3.7 Perceived intensity vs. threshold-weighted vibration power. A log scale is used in the abscissa.

mediated by the RA channel. We indeed obtained better fits when the 60 Hz data were excluded from analysis ($r^2 = 0.95$–0.97).

We also explored another metric that relies on the signal power absorbed by the hand. When we hold a mobile device, the physical energy of a vibrotactile stimulus transmitted to the hand is determined by the bio-mechanical property of the hand-arm system, which is represented by mechanical impedance Z. Z is defined in the frequency domain as $Z = F/v$, for applied force F and skin movement velocity v. Here, the real part $\mathrm{Re}[Z]$ is a damping term related to the vibratory energy absorbed by the hand, while the imaginary part $\mathrm{Im}[Z]$ contains spring and mass terms related to the energies stored in the hand-arm system and transmitted to the movement of the other body parts, respectively [50]. $|\mathrm{Im}[Z]|$ is known to be much smaller than $|\mathrm{Re}[Z]|$ in the frequency range (60–320 Hz) tested in our experiments, and the vibration power absorbed by the hand is likely to be highly correlated with the subjective sensation magnitude or discomfort [10]. Therefore, we compared the absorbed power of the vibrotactile stimuli used in our experiment and their perceived intensities.

The vibratory power absorbed by the hand at frequency f is

$$P(f) = \mathrm{Re}[F(f) \cdot v(f)] = \mathrm{Re}[Z(f)] \cdot |v(f)|^2. \tag{3.4}$$

Then, for a sinusoidal vibration with acceleration $a(t) = A\sin(2\pi f_s t)$, its skin-absorbed

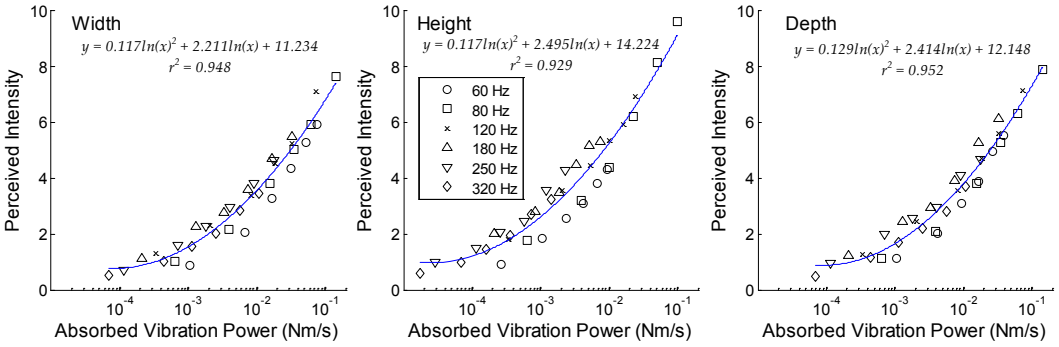

Fig. 3.8 Perceived intensity vs. skin-absorbed vibration power. A log scale is used in the abscissa.

power integrated over all frequencies is

$$P_A = P(f_s) = \mathrm{Re}[Z(f_s)] \cdot \left| \frac{A}{2\pi f_s} \right|^2 . \tag{3.5}$$

For Z, we took the mechanical impedance contour measured at a 20 N grip force from [38], as this study used a grip condition similar to Exp. II, albeit a large difference in grip force.

Fig. 3.8 depicts the skin-absorbed vibration powers of the vibrotactile stimuli used in Exp. II with respect to their perceived intensities for each vibration direction. A strong quadratic relationship ($r^2 > 0.92$) was found between the log-scaled P_A and the perceived intensities for each vibration direction. These fitted lines are represented in Fig. 3.8 by solid lines with their equations. Therefore, the skin-absorbed vibration power computed from the mechanical impedance of the hand-arm system and the physical parameters of vibrotactile stimuli can also be a robust metric for perceived intensity.

Compared to the Pacinian-weighted power, the skin-absorbed power can be simplified further since the real part of the mechanical impedance above a 60 Hz frequency is almost constant [38]. Thus, for simplicity, the impedance term in (3.5) can be omitted and this may suffice for practical purposes. In addition, the skin-absorbed power does not distinguish between the responding mechanoreceptor types, unlike the Pacinian-weighted power, which may justify its use also for low frequency vibrations.

Chapter 4

Perceptual Space of Sinusoidal Vibration

In this chapter, we report the perceptual space of sinusoidal vibrotactile signals perceived through a mobile device held in the hand. In Exp. I, we measured perceived dissimilarities among sinusoidal vibrations with seven frequencies in 40–250 Hz and two amplitudes of 30 and 40 dB SL. Multi-dimensional scaling was then applied to the perceptual distances, and led to a two-dimensional perceptual space. In Exp. II, we evaluated the subjective qualities of simple sinusoidal with different frequencies and via adjective rating. Thirteen adjective pairs were carefully selected, and rated for the vibrations played through the mobile device. In Exp. III, we had same procedure to Exp. II for superimposed bi-frequency vibrations with different intensity mixture ratios.

4.1 Exp. I: Perceptual Space Estiamtion

In this experiment, we measured the perceptual distances between the pairs of sinusoidal vibrations with various frequencies (40–250 Hz) and two amplitudes (30 and 40 dB SL). The vibrations were presented through a mobile device mock-up grasped in the hand. From the perceptual distances, we computed a dissimilarity matrix, and performed MDS on the matrix in order to find a Euclidian perceptual space of sinusoidal vibrations.

4.1.1 Methods

Participants

Ten male participants participated in this experiment. Their ages ranged in 20–28 years with an average of 23.0. Six participants were right-handed, and the other four were left-handed or ambidextrous by self-report. All participants were everyday users of a mobile phone. No participants reported any sensorimotor abnormalities regarding their upper extremity. The participants were compensated after the experiment.

Apparatus

We used a mini-shaker (Brüel & Kjær; model 4810) to generate sinusoidal vibrations in the experiment. The mini-shaker can produce linear vibrations in a wide range of frequencies (nominal bandwidth = 18 kHz). Miniature vibration actuators that can be embedded into a mobile device were inappropriate in this experiment due to their limited output ranges in terms of vibration amplitude and frequency, as reviewed earlier in Section 2.4.

A mobile device mock-up made of acrylic resin ($10.5 \times 4.5 \times 1.5$ cm) was attached on top of the mini-shaker using a screw-type aluminum bracket as shown in Fig. 4.1. Input signals to the shaker were computed by a control PC, and were transmitted to the mini-shaker via a data acquisition board (National Instruments; model PCI-6229) and a power amplifier (Brüel & Kjær; model 2718). Vibrations produced by the mini-shaker were measured using a single axis accelerometer (Kistler; model 8630C) attached on the lower part of the mock-up. The sampling rate for signal I/O was 10 kHz. This value was 40 times larger than the highest vibration frequency used in the experiment (250 Hz), and guaranteed faithful signal sampling and reconstruction. We also carefully calibrated the relationship between input voltage amplitude and output vibration amplitude for each frequency, following the calibration procedures in [31].

Stimuli

The parameters of sinusoidal vibrations used in this experiment were determined by factorially combining 7 frequencies (40, 80, 100, 120, 150, 200, and 250 Hz) and 2 amplitudes (30

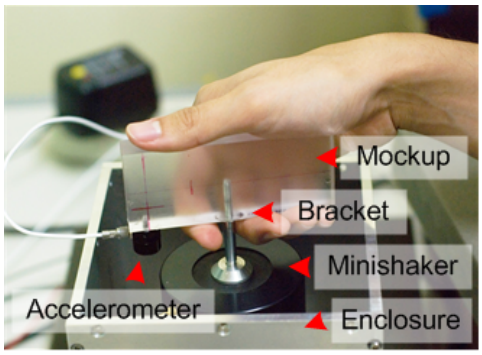

Fig. 4.1 Exp.al hardware that simulates vibration generation and perception in a mobile device.

and 40 dB SL), resulting in the 14 sinusoidal vibrations. The amplitudes were specified in sensation levels with respect to the absolute detection thresholds measured previously using a mobile phone [71]. Since two vibrations were presented in a pair to evaluate their perceptual dissimilarity, the experiment had the 91 pairs ($= 14 \times 13/2$) of different vibrations and the 14 pairs of same vibrations, thus a total of 105 pairs.

Procedure

In each trial, two vibrations of 1-s duration were presented to the participant one by one with an inter-stimulus interval of 1 s. After perceiving the vibration pair, the participant was asked to report absolute perceptual dissimilarity between the two vibrations in a free modulus scale (without a standard stimulus) using a keyboard. The participant was instructed to represent the perceptual distance as a positive number and the distance between vibrations perceived identical as 0. In a magnitude estimation experiment, participants have a tendency to answer the perceived magnitude of a stimulus in their own scale. The paradigm of absolute magnitude estimation is immune to the problem that a standard stimulus may bias the internal scale of a participant [15].

In the experiment, each vibration pair was presented in four trials. In two trials, one vibration was rendered first, and in the other two trials, the other vibration was rendered first. The pairs of the same vibrations were also included to provide the perceptual anchor of

zero difference. As a result, the experiment had a total of 420 ($= (14 \times 13/2 + 14) \times 4$) trials. The order of presenting the vibration pairs was randomized for each participant. Prior to the main experiment, all participants went through a training session where the 105 vibration pairs were presented once each.

During the experiment, the participant was seated in front of a computer, and held the mobile device mockup comfortably with the right hand similarly as s/he held a mobile phone. A computer monitor displayed texts necessary for the progress of the experiment. The participant also wore headphones that played white noise to block faint auditory noise emanating from the mini-shaker.

Data Analysis

The first step of data analysis was the normalization of perceptual distances collected in the experiment. The responses of the participants were in different subjective scales, and these were normalized into a 0–100 scale. A scaling factor $R(k)$ for a participant k was calculated as follows. Let the response for a pair (i, j) in the n-th repetition be $r_{ij}(n)(1 \leq n \leq 4)$. The responses $r_{ij}(n)$ were averaged over n, and a maximum value was picked from the averages for normalization, such that

$$R(k) = \frac{100}{\max\limits_{i,j} \frac{\sum_{n=1}^{4} r_{ij}(n)}{4}}. \tag{4.1}$$

The normalized perceptual differences were averaged again across the participants to obtain a dissimilarity matrix of the 14 sinusoidal vibrations.

The next step was the estimation of a Euclidean perceptual space by applying MDS to the normalized dissimilarity matrix. For a perceptual space of dimension N, a distance between two elements $\mathbf{x}_i^N$ and $\mathbf{x}_j^N$ was defined as the Euclidean distance:

$$d_{ij} = ||\mathbf{x}_i^N - \mathbf{x}_j^N||. \tag{4.2}$$

From the dissimilarity matrix $\{\delta_{ij} \,|1 \leq i, j \leq 14\}$, we estimated the positions of the 14 sinusoidal vibrations, $\{\mathbf{x}_i^N | 1 \leq i \leq 14\}$ so that the error between $\{\delta_{ij}\}$ and $\{d_{ij}\}$ was

minimized using MDS [86]. The goodness of fit was evaluated using Kruskal's stress [42], defined as

$$S = \left[\frac{\sum_{i=1}^{14} \sum_{j=i}^{14} \left(\delta_{ij} - d_{ij}\right)^2}{\sum_{i=1}^{14} \sum_{j=i}^{14} d_{ij}^2} \right]^{\frac{1}{2}} . \tag{4.3}$$

The stress S varies between 0 and 1, and approaches to 0 as agreement between the measured and estimated distances improves.

In general, a perceptual space in a higher dimension has lower S than a lower dimensional representation, but S is usually saturated from a certain dimension. This dimension is selected as an optimal dimension for the perceptual space. Note that in MDS representations, only relative positions between elements are meaningful; a rotation or translation of entire elements does not affect S or change the analysis of results.

4.1.2 Results

Table 4.1 shows the dissimilarity matrix of the 14 sinusoidal vibrations averaged across the participants. The standard deviations of the elements ranged in 5.21–24.79 with an average of 12.71.

To find a perceptual space that best preserves the perceptual distances in Table 4.1, we performed MDS while increasing the dimension of the perceptual space. The stress value representing the goodness of fit was sufficiently small in 2D MDS (Kruskal's stress $S = 0.096$). Therefore, we showed the relative positions of the 14 vibrations in a two-dimensional plane, as in Fig. 4.2. In the figure, the squares represent the positions of vibrations with 30 dB SL amplitude, and the circles represent those with 40 dB SL amplitude.

In the MDS plot, the 14 vibration points formed two distinct groups depending on their amplitudes. Seven vibration points of different frequencies in group exhibited similar distributions. In both groups, as frequency increased from 40 Hz, the vibration points moved to the northwestern direction until the frequency reached 100 Hz. Further frequency increase from 100 Hz changed the direction of point movements to the northeastern direction,

Table 4.1 Dissimilarity matrix of the 14 sinusoidal vibrations measured in Exp. I. The numbers in the first row and column indicate the parameters of vibrations in frequency (Hz) - amplitude (dB SL).

,Hz-dB	40-40	80-30	80-40	100-30	100-40	120-30	120-40	150-30	150-40	200-30	200-40	250-30	250-40
40-30	46.3	31.0	69.4	38.9	87.3	33.1	84.1	34.7	77.0	29.0	68.7	31.0	77.4
40-40		32.8	39.0	35.2	61.2	41.0	63.4	44.6	57.0	43.4	56.0	42.5	59.2
80-30			55.0	21.4	63.3	16.6	60.0	18.3	57.6	20.1	47.9	27.6	54.4
80-40				46.7	17.4	46.4	22.8	56.2	37.7	61.8	39.4	58.9	50.7
100-30					57.0	14.7	49.9	21.8	49.0	29.1	47.5	27.9	47.8
100-40						53.6	23.0	71.2	26.1	70.0	41.2	73.8	46.0
120-30							52.3	14.8	53.6	17.9	45.9	20.8	50.9
120-40								64.8	22.7	68.5	30.1	66.3	34.9
150-30									56.9	18.0	51.4	16.1	54.0
150-40										62.7	20.3	58.9	30.5
200-30											51.7	13.1	51.8
200-40												53.3	21.7
250-30													47.2

demonstrating a clear elbow point at 100 Hz. Consequently, the overall shape of the point distribution was similar to a '$<$' shape in both amplitude levels.

We further assessed the effect of vibration amplitude as follows. In each group of vibration points with the same amplitude, we regressed a line to the vibration points with frequencies of 40–100 Hz, and another line to those with frequencies of 100 – 250 Hz. The two fitted lines were then connected, making a triangle. In Fig. 4.2, the fitted lines are shown in solid lines, and the connecting line between them is in a dashed line, for each amplitude group. The shapes of the two triangles were fairly similar, but had different sizes. The areas of the triangles were 251.2 in the 30 dB SL group and 1444.1 in the 40 dB SL group. This suggests that vibrations with a higher amplitude were perceived to be more different as their frequency varied.

4.1.3 Discussion

The perceptual space of sinusoidal vibrations perceived via the mobile device mockup indicated that the vibrations in the frequency bands of 40–100 Hz and 100–250 Hz have fundamentally different perceptual properties. This observation is consistent with the categories

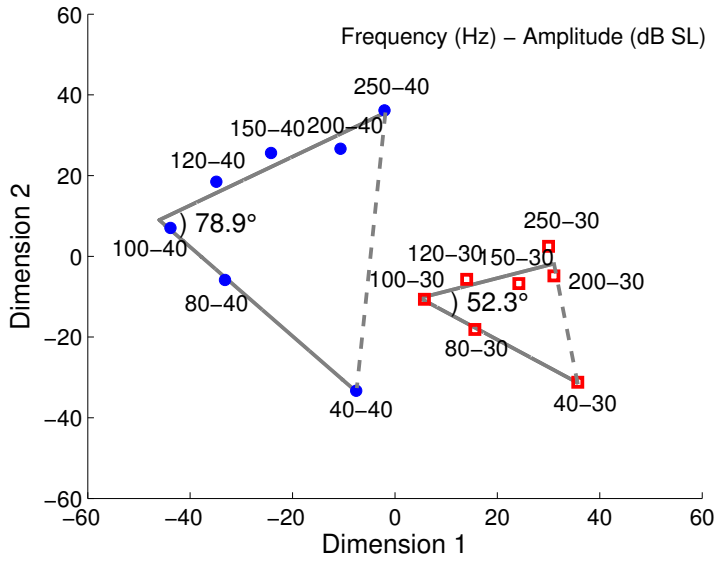

Fig. 4.2 Two-dimensional perceptual space of the 14 sinusoidal vibrations.

of sinusoidal vibrations made by Tan [79]. Based on subjective descriptions, Tan classified sinusoidal vibrations into three groups along their frequency. Vibrations in a 1–3 Hz band were described as a slow kinesthetic motion, those in 10–70 Hz as rough motion or fluttering, and those in 100–300 Hz as smooth vibration. Vibrations with in-between frequencies were reported to share the qualities of both adjacent bands. This implies that the perceptual space of sinusoidal vibrations spanned by one physical variable, frequency, may consist of two perceptual dimensions in the frequency range tested in Exp. I. Indeed, the dimension of our perceptual space was two, and the '<'-shaped point distributions are likely to be resulted from the differences in the subjective impressions of sinusoidal vibrations depending on their frequencies. If we had included a very low frequency range below 10 Hz in the experiment, a resulting perceptual space might have had three perceptual dimensions.

Another important issue regarding a perceptual space is the orthogonality between perceptual dimensions that span the perceptual space. In our case, the two fitted lines in Fig. 4.2 for each amplitude level can be considered as two perceptual dimensions. The inner angles of the 100-Hz elbow points at which the two fitted lines crossed were 52.3° in the 30 dB SL

group and 78.9° in the 40 dB SL group. It appears that increasing the vibration amplitude makes the two perceptual dimensions more orthogonal.

The above results can also be explained based on the characteristics of the neural channels responsible for tactile perception. As is well known, time-varying tactile stimuli with frequencies over 100 Hz are mostly mediated by the PC (Pacinian) channel, and feel like smooth vibrations. In contrast, tactile stimuli with lower frequencies, e.g., 40 and 80 Hz in this experiment, evoke the responses of both PC and NP (Non-Pacinian) I channels if their amplitudes are larger than the absolute thresholds of the both channels [4][70]. The stimuli mediated by the NP I channel induce the sensations of fluttering [19]. Thus, sensations resulted from the 40 and 80 Hz vibrations with a large amplitude possess the perceptual qualities of both neural channels. As signal frequency decreases from 100 Hz, the fluttering sensation of the NP I channel becomes more evident, since the absolute thresholds of the NP I channel remain relatively invariant but those of the PC channel rapidly increase [4]. This tendency was well observed at the 40 and 80 Hz points in the perceptual space; the 40 Hz vibration was far more distant from the high-frequency vibrations than the 80 Hz vibration. Furthermore, since the absolute thresholds of the NP I channel are significantly higher than those of the PC channel, tactile stimuli with a higher amplitude are expected to exhibit the characteristics of the NP I channel more clearly. This tendency is also well reflected in the orthogonality of the two perceptual dimensions, which increased with the vibration amplitude.

4.2 Exp. II: Adjective Rating of Simple Sinusoids

Exp. II aimed at finding the qualitative attributes of sinusoidal vibrations perceived through a mobile device using adjective rating. The experiment consisted of two steps. In the first step, we collected adjectives suitable to describe the subjective feelings of the sinusoidal vibrations. In the second step, participants rated the similarities between the adjectives and the sinusoidal vibrations. The results were projected to the perceptual space estimated in Exp. I to determine the adjective pairs that can account for the perceptual space in the largest extent.

4.2.1 Methods

Participants and Apparatus

We recruited two groups of participants for this experiment. Group 1 was devoted to the adjective collection, and comprised nine male university students. Group 2 consisted of 10 participants (7 males and 3 females), and participated in the similarity rating experiment. Their ages ranged in 19–24 years with an average of 21.3 years. All participants of both groups were native speakers of Korean, and reported no known sensorimotor abnormalities. The participants were compensated for their efforts after the experiment.

The participants were presented with mobile device vibrations under the same hardware configuration used in Exp. I.

Adjective Collection

As a first step, we collected Korean adjectives that can represent the subjective impressions of sinusoidal vibrations perceived through a mobile device. The participant group 1 took part in this step. The vibrations experienced by the participants had 1-s duration, four frequencies (40, 80, 150, and 250 Hz) and one amplitude. The vibration amplitudes were adjusted to be level 11 in the perceived magnitude regardless of their frequency using a psychophysical magnitude function measured in our previous study [72]. The perceived magnitude approximately corresponded to 40 dB SL (see Section 4.2.3 for further details).

The participants were asked to write down adjectives that could be associated with the feelings of the vibrations. After that, objective questions were given to the participants in a list of 56 adjectives gathered from related studies [24, 40], web pages, and Korean dictionary. The participants marked the adjectives that they thought were appropriate to describe the sensations of the sample vibrations. After the collection, we sorted the most frequently appeared adjectives, and prepared 13 adjective pairs such that each pair had opposite meanings. The adjective pairs translated to English are provided in Table 4.2. They are denoted in a form of 'adjective 1-adjective 2' in the rest of this study.

Table 4.2 List of the 13 adjective pairs used for adjective rating in Exp. II (translated from Korean to English).

Pair No.	Adjective 1	Adjective 2
1	slow	fast
2	sparse	dense
3	blunt	sharp
4	bumpy	smooth
5	hard	soft
6	jagged	aligned
7	thick	thin
8	vague	distinct
9	heavy	light
10	deep	shallow
11	dark	bright
12	gentle	brisk
13	dull	clear

Adjective Rating

In this step, the participants rated the similarities of 1-s sinusoidal vibrations of seven frequencies (40, 80, 100, 120, 150, 200, and 250 Hz; the same to Exp. I) with the 13 adjective pairs shown in Table 4.2. Thus, the experiment consisted of seven trials. The vibration amplitudes were adjusted to be level 11 in the perceived magnitude as were in the adjective collection.

We made a GUI-based program for this experiment (see Fig. 4.3 for its appearance). In each trial, the participant could perceive a vibration stimulus by pressing a play button on the program window as many times as s/he wanted. The participant moved slide bars to rate similarities between the vibration and the adjective pairs. All slide bars were centered in the beginning of the trial. Moving a slide bar to a certain adjective indicated that the vibration was perceived close to the feeling of that adjective. The horizontal length of each slide was 127 mm on the screen, following the recommendation of Shiffman et al. [75]. The order of presenting the vibrations in the seven trials was randomized for each participant. On average, the experiment took an hour per participant. The participants could take a rest

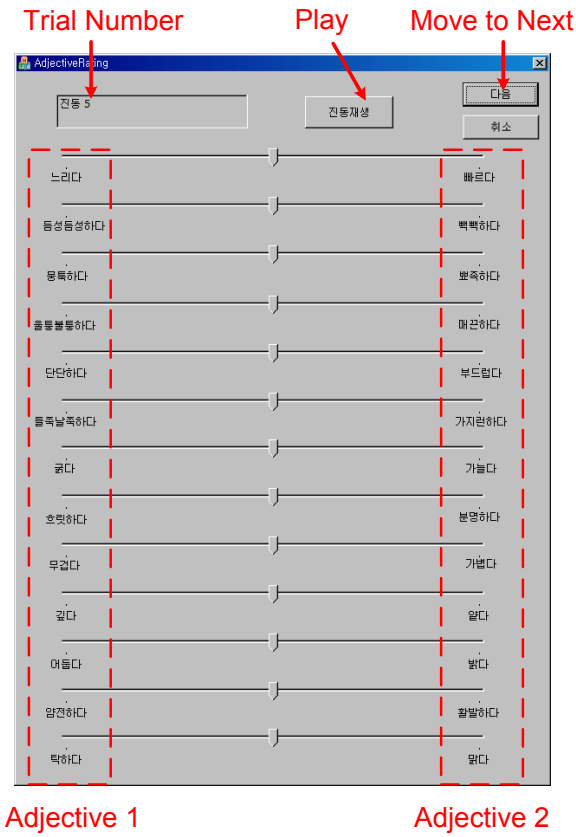

Fig. 4.3 Screen shot of the experiment program used for adjective rating. The order of the adjective pairs shown on the window in Korean are identical to that in English listed in Table 4.2.

whenever necessary.

Data Analysis

At the end of each trial, the position of a slide bar for 'adjective 1-adjective 2' was linearly mapped to a similarity score between 0 and 100. The slide bar positions closest to adjective 1 and adjective 2 corresponded to similarity scores of 0 and 100, respectively. The individual similarity scores were averaged across the participants for each vibration.

Since the vibrations used in this experiment only had one amplitude close to 40 dB SL,

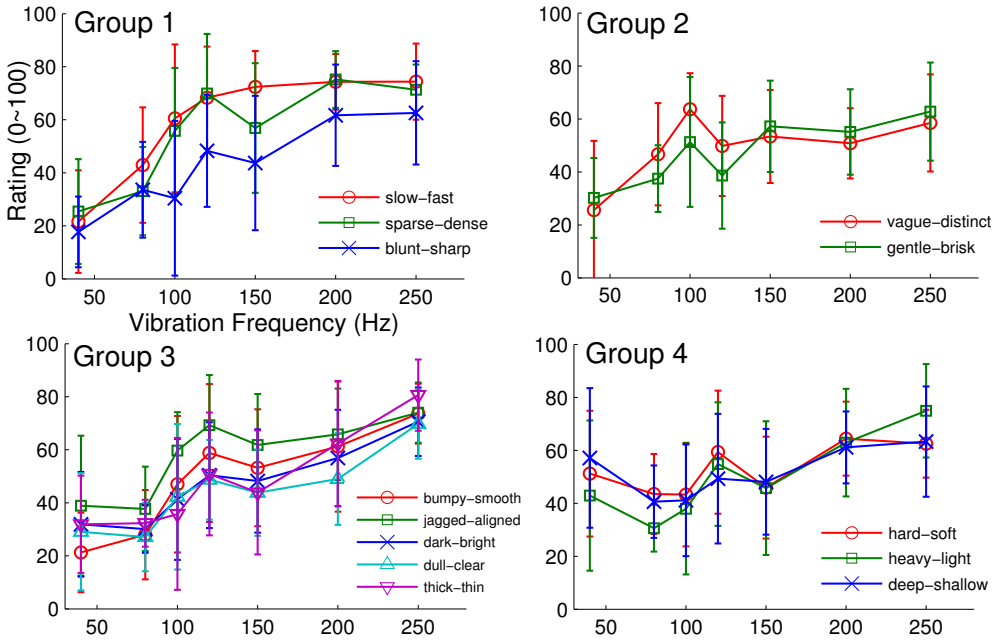

Fig. 4.4 Results of adjective rating in Exp. II. The error bars represent standard deviations. In the data of 'adjective 1-adjective 2', a score close to 0 indicates the corresponding vibration felt more similar to adjective 1, and a score close to 100 indicates it felt more similar to adjective 2.

we applied MDS again to the elements with 40 dB SL amplitude in the dissimilarity matrix shown in Table 4.1, and obtained a 2D perceptual space for the 40 dB SL sinusoidal vibrations. The adjective pairs were then projected to the perceptual space using the multiple linear regression. The seven vibration coordinates in the perceptual space were input variables, and the similarity ratings for the 13 adjective pairs were response variables. The standardized regression coefficients were used to represent the slopes of the adjective pairs in the 2D perceptual space.

4.2.2 Results

The average ratings and standard deviations measured in Exp. II are shown in Fig. 4.4 for each adjective pair. The adjective pairs were grouped based on the similarities of the

Table 4.3 Correlation matrix of the 13 adjective pairs. Values for highly correlated adjective pairs are marked in boldface.

Adjective Pair No.	2	3	4	5	6	7	8	9	10	11	12	13
1 (slow-fast)	0.684	0.579	0.584	0.331	0.587	0.551	0.448	0.363	0.220	0.516	0.558	0.572
2 (sparse-dense)		0.662	**0.815**	0.501	0.684	0.658	0.410	0.561	0.369	0.626	0.423	0.651
3 (blunt-sharp)			0.660	0.521	0.568	0.746	0.309	0.596	0.487	0.620	0.421	0.636
4 (bumpy-smooth)				0.599	**0.760**	**0.806**	0.308	0.707	0.449	0.716	0.397	0.676
5 (hard-soft)					0.558	0.728	0.028	**0.789**	0.492	0.578	0.187	0.577
6 (jagged-aligned)						0.654	0.291	0.583	0.305	0.631	0.331	0.682
7 (thick-thin)							0.154	**0.868**	0.586	0.735	0.360	0.734
8 (vague-distinct)								0.042	-0.099	0.417	0.537	0.400
9 (heavy-light)									0.578	**0.759**	0.229	0.701
10 (deep-shallow)										0.498	0.069	0.375
11 (dark-bright)											0.498	**0.870**
12 (gentle-brisk)												0.572
13 (dull-clear)												

results, and were shown in separate plots for visibility. In adjective group 1 including 'slow-fast,' 'sparse-dense,' and 'blunt-sharp,' the feeling of the vibrations became closer to 'fast,' 'dense,' and 'sharp,' almost monotonically as the frequency increased. In adjective group 2, the changing patterns were dependent on the frequency. With increasing frequency, the ratings of 'vague-distinct' and 'gentle-brisk' increased to 'distinct' and 'brisk' in 40–100 Hz, dropped to 'vague' and 'gentle' at 120 Hz, and then saturated in 150–250 Hz. In adjective group 3 for 'bumpy-smooth,' 'jagged-aligned,' 'dark-bright,' 'dull-clear,' and 'thick-thin,' the sensations for 'smooth,' 'aligned,' 'bright,' 'clear,' and 'thin' increased in 40–120 Hz, slightly decreased to 'bumpy,' 'jagged,' 'dark,' 'dull,' and 'thick' at 150 Hz, and then increased again in 150–250 Hz. The adjective group 4 consisted of 'hard-soft,' 'heavy-light,' and 'deep-shallow,' and showed similar patterns to group 3, except that they contained a U-shaped curve in 40–120 Hz. Overall, as vibration frequency increased, the subjective impressions changed from the negative (yin in Chinese) adjective of an adjective pair to the positive (yang in Chinese) adjective. The varying patterns in the low frequency band (40–100 Hz) and the high frequency band (100–250 Hz) were noticeably different in the ten adjective pairs of groups 2, 3, and 4. Correlations between the adjective pairs are also

provided in Table 4.3. High correlations over 0.750 are marked in the bold face.

We then projected the results of adjective rating to the 2D perceptual space of 40 dB SL vibrations using the multiple linear regression, and presented the results in Fig. 4.5. Each adjective pair is represented as a line that intercepts the origin with a slope proportional to the ratio of its standardized regression coefficients. The sum of the squares (SS) of the standard coefficients is approximately proportional to R^2 of an adjective pair [24]. To visualize the fitness of an adjective pair, the length of a line from the origin is set to 75 times of the SS of its standard coefficients, and is also mirrored to the opposite quadrant. Thus, a long adjective axis indicates that the adjective pair is highly correlated to the vibration positions in the 2D perceptual space.

In addition, we projected the vibration points to the axis of each adjective pair to visualize the regressed positions of the vibration points, and examined whether the order of vibration frequencies was preserved in the projected positions and the projected positions were also well separated. Such adjective pairs for all vibration frequencies were 'dark-bright' and 'dull-clear,' as shown in the top panel of Fig. 4.6. The two adjective pairs showed a monotonically increasing pattern for increasing frequency except at 150 Hz in Fig. 4.4, and also had a high correlation of 0.870 in Table 4.3. Thus, both 'dark-bright' and 'dull-clear' are most appropriate perceptual dimensions to describe the subjective sensations of mobile device vibrations in a frequency range of 40–250 Hz with one dimension. For the low frequency range of 40–100 Hz, the adjective pairs that monotonically span the vibrations with reasonable discriminability were 'slow-fast,' 'vague-distinct,' 'sparse-dense,' 'jagged-aligned,' and 'bumpy-smooth,' as is in the middle panel of Fig. 4.6. Among them, 'slow-fast' and 'vague-distinct' spanned the largest distances, which can be seen in Figures 4.4 and 4.5, thus are recommended perceptual dimensions for the vibrations in 40–100 Hz. Their correlation was relatively low (0.448). For the high frequency range of 100–250 Hz, 'thick-thin,' 'heavy-light,' 'deep-shallow,' and 'hard-soft' were the candidates as in the bottom panel of Fig. 4.6. The adjective pairs that distributed the vibrations in larger distances were 'thick-thin' and 'heavy-light', which had a high correlation of 0.868. In summary, we recommend 'dark-bright' or 'dull-clear' to represent the subjective impressions of

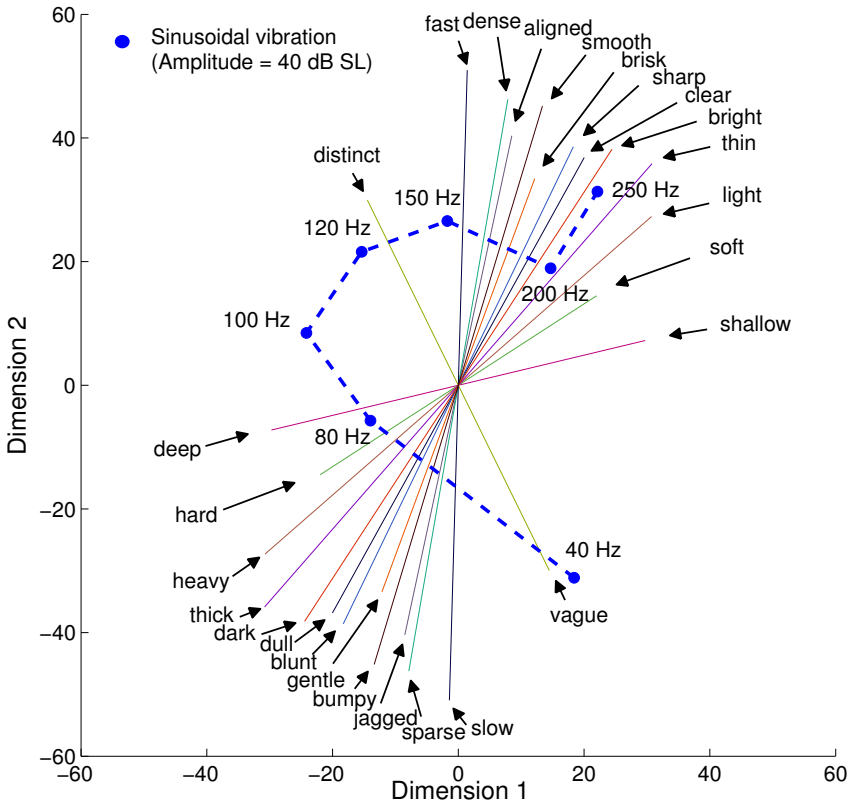

Fig. 4.5 Adjective pairs regressed to a 2D perceptual space of the sinusoidal vibrations of 40 dB SL amplitude. The length of each axis is proportional to the correlation magnitude of the corresponding adjective pair to the vibration points.

mobile device vibrations in one perceptual dimension. To use two perceptual dimensions, 'slow-fast' or 'vague-distinct' is appropriate for the low frequency range (40–100 Hz), and 'thick-thin' or 'heavy-light' is for the high frequency range (100–250 Hz).

4.2.3 Discussion

In Exp. II, the vibration amplitudes of different frequencies were regulated to be level 11 in their perceived intensities using a psychophysical magnitude function measured in [71]. This was necessary since we further needed to compare the results of Exp. II to those

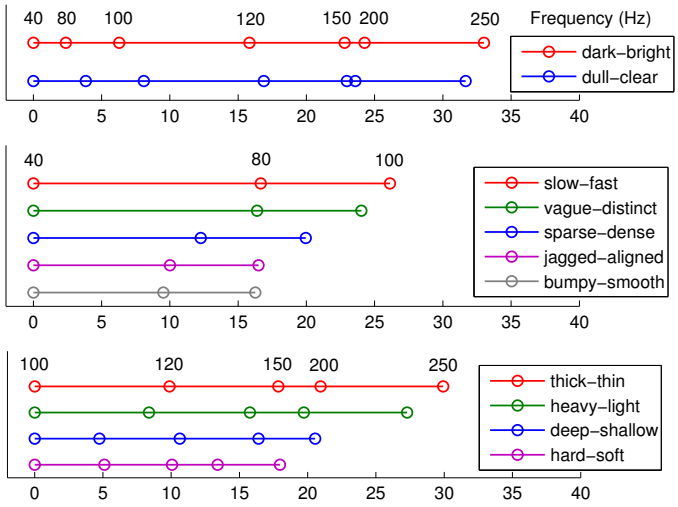

Fig. 4.6 The positions of vibrations projected to the axes of adjective pairs reproduced from Fig. 4.5. Results of the adjective pairs where the order of vibration frequencies is preserved and the projected positions are well distributed were only selected.

of another experiment (not reported in this study due to a non-disclosure agreement with a funding sponsor). Since this experiment used a different vibration actuator that makes more complex vibration outputs, representing the perceived intensity of its output in sensation level was not straightforward. The perceptual space obtained in Exp. I, however, controlled the vibration amplitudes in sensation level. Despite this mismatch between Exps. I and II, its effect is not so significant to undermine the conclusions drawn in Exp. II.

Fig. 4.7 shows the perceived magnitudes of mobile device vibrations reproduced from [71]. For an vibration amplitude of 40 dB SL, 40 Hz vibration resulted in the highest perceived intensity of 14.1, and 200 Hz vibration resulted in the lowest of 7.6. Thus, the perceived intensity difference from 11 was between -3.4 and 3.1. Also note that in all experiment we did not strictly control the hand grip forces of the participants holding the mobile device in order to collect data under the natural uses. This may cause relatively large individual variations between the participants, thus increasing the overall variances in the perceived intensity measurements. Due to these reasons, we ensured in Section 4.2.2

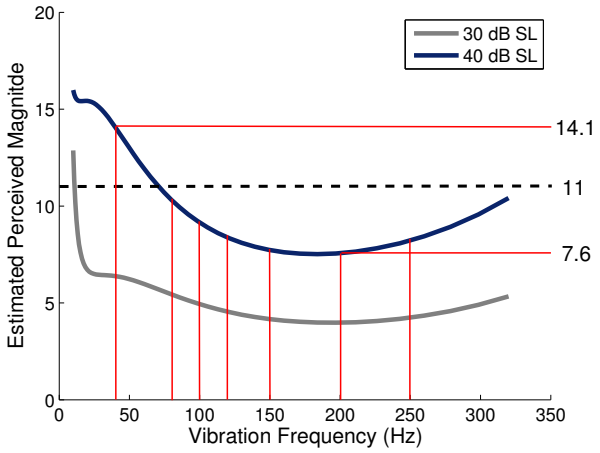

Fig. 4.7 Perceived magnitudes of mobile device vibrations as a function of vibration frequency (reproduced from [71]).

that all conclusions about the adjective ratings made using the perceptual space of 40 dB SL vibrations in Fig. 4.5 were also confirmed in the raw data in Fig. 4.4 and Table 4.3.

In Kim et al. [40], an adjective rating experiment similar to our study was conducted using a 2D tactile pin array. The vibrotactile sensations generated by a pin array are largely different from those of mobile device vibrations. Nonetheless, since this study used 10 single adjectives some of which were common to the 13 adjective pairs of our study, we compare the results in what follows. A common observation was that higher frequency stimuli were perceived as denser (less sparse) in both studies. Other adjectives had different ratings between the two studies. In the study of Kim et al., negative adjectives such as 'prickly' and 'lumpy' had increasing ratings as vibration frequency increased, whereas positive adjectives such as 'smooth' and 'tender' had decreasing ratings. In contrast, our study showed that the high frequency vibrations had closer feeling to positive adjectives including 'smooth' and 'soft.'

A primary difference between the two studies lies on the type of a contactor and the contact area. In [40], the array consisted of 8×6 pins with a pin spacing larger than 1 mm, and each pin with a diameter of 0.5 mm stimulated a very small area on the index

finger. Since the movements of all the pins were synchronized in their experiment, the total contact area was slightly over 1 cm^2. In such stimulation, the sensation of each pin may feel 'prickly' if its amplitude and frequency are large. In addition, the amplitudes of vibrations were fixed to one physical value over different frequencies [40], without taking into account that the vibrotactile detection thresholds that rapidly decrease with frequency until around 250 Hz. Thus, it appears that higher frequency vibrations had larger perceived intensities than lower frequency vibrations, which can enhance the impressions such as 'prickly' and 'lumpy.' On the other hand, a mobile phone in contact with the palm and all five fingers provides much smoother and softer vibration sensations. We also controlled the perceived intensities of the vibrations to be on the same level regardless of vibration frequency. This comparison suggests that the subjective impressions of vibrotactile stimuli can depend on a device and various contact conditions.

4.3 Exp. III: Adjective Rating of Bi-frequency Vibrations

Goal of Exp. III is to compare the qualitative characteristics of simple and superimposed bi-frequency vibrations. Adjective rating was conducted on the bi-frequency vibrations with various intensity mixture ratio. Most of the procedures were identical to those of Exp. II. Experiment results were projected on the perceptual space estimated in Exp. II.

4.3.1 Methods

Participants and Apparatus

Six of the participants in Exp. II agreed to participate on Exp. III. The participants were compensated after the experiment.

In this experiment, a prototype of DMA was used to generate superimposed bi-frequency vibration stimuli. Detailed explanation of DMA is described in Section 2.4.1. Two resonance frequencies of the DMA used were 150 and 250 Hz. We attached a DMA on the front face of mock-up in the hardware system of Exp. II. The rest part of the apparatus was identical to that of Exps. I and II.

Bi-frequency Vibration Stimuli

Bi-frequency sinusoidal vibrations were used as stimuli in this experiment. We assumed that perceived intensity of superimposed sinusoidal vibration is additive [2]. In Bensmaia's model, perceived intensity of superimposed sinusoidal stimulus S, I_S is given by,

$$P_S(f) = \frac{A_f^2 f^2}{T_f^2 f^2}, \tag{4.4}$$

$$I_s = \sum_f P_S(f)^{a_f} \tag{4.5}$$

where A_f and T_f are amplitudes of the spectral component and detection threshold at frequency f, respectively. a_f is the exponent of Stevens' power law that vary from frequency to frequency. Since we measured and modeled perceived intensity of simple vibration $P_S(f)^{a_f}$, we can control them in acceleration magnitude. Thus bi-frequency sinusoidal vibration can be composed additively as a superimposition of two simple vibrations to have the desired perceived intensity. Since the phase diference between two components affects little on perception in Pacinian channel [1], it was not considered in this study.

Adjective Rating

Five bi-frequency sinusoidal vibrations were used as stimuli in this experiment. Two element frequencies of the stimuli, f_1 and f_2 were fixed as 150 and 250 Hz, which are resonant frequencies of the actuator. Intensity mixture ratio of the two frequency components, I_1:I_2 were varied in five levels: 0.1:0.9, 0.3:0.7, 0.5:0.5, 0.7:0.3, and 0.9:0.1. The intensity was controlled in perceived magnitude using Ryu's perceived magnitude model [72]. Level 11 in perceived magnitude were linearly divided into the two frequency components along with their intensity ratio. In this way, all bi-frequency sinusoidal stimuli in this experiment are assumed to have the identical perceived intensity with the stimuli in Exp. II. The 13 adjective pairs in Exp. II were used. Rest of the setup was identical to the procedure of Exp. II.

Data Analysis

At the end of each trial, the position of a slide bar for 'adjective 1-adjective 2' was linearly converted into a similarity score between 0 and 100. The positions closest to adjective 1 and adjective 2 corresponded to similarity scores of 0 and 100, respectively. The individual similarity scores were averaged across the participants for each vibration. Each stimulus was projected as a point on the perceptual space of Exps. I and II, regarding its adjective similarity scores. We can estimate adjective rating scores of a point on the perceptual space using the adjective axes in Exp. II. For each bi-frequency stimulus, we iteratively found a point with minimum sum of errors between the measured scores and estimated scores of a position.

4.3.2 Results

The means and standard deviations of the measured similarity scores are shown in Fig. 4.8 for each adjective pair. We used the adjective groups in Exp. II, and showed them in separate plots for visibility. Most adjective pairs show U-shape on their score plot, with low scores on the equal-intensity mixture of 150 Hz and 250 Hz components. When $I_1 : I_2 = 0.5{:}0.5$, the participants' ratings are mostly close to adjective 1, such as; slow, sparse, blunt, bumpy, jagged, dark, dull, thick, deep, and heavy. Meanwhile, scores of 'vague-distinct', 'gentle-brisk', and 'hard-soft' pairs did not vary much with the intensity mixture ratio. As shown in the results of Exp. II, ratings close to adjective 1 (negative adjective) is a characteristic of low frequency sinusoidal vibrations.

From the result, we can suppose the perceived feeling of the bi-frequency vibrations in this experiment are similar to low frequency vibrations ($>$100 Hz). Their perceptual similarity with the low frequency vibration is maximized when the perceived intensity of two frequency components are even.

To analyze and show this effect visually, coordinates of the bi-frequency stimuli were estimated on the perceptual space. The result is shown in Fig. 4.9. Each bi-frequency vibration is represented as a square on the perceptual space in Exp. II. When a component of bi-frequency vibration is dominant ($I_1 : I_2 = 0.1{:}0.9$ or $0.9{:}0.1$), coordinates of the stimulus

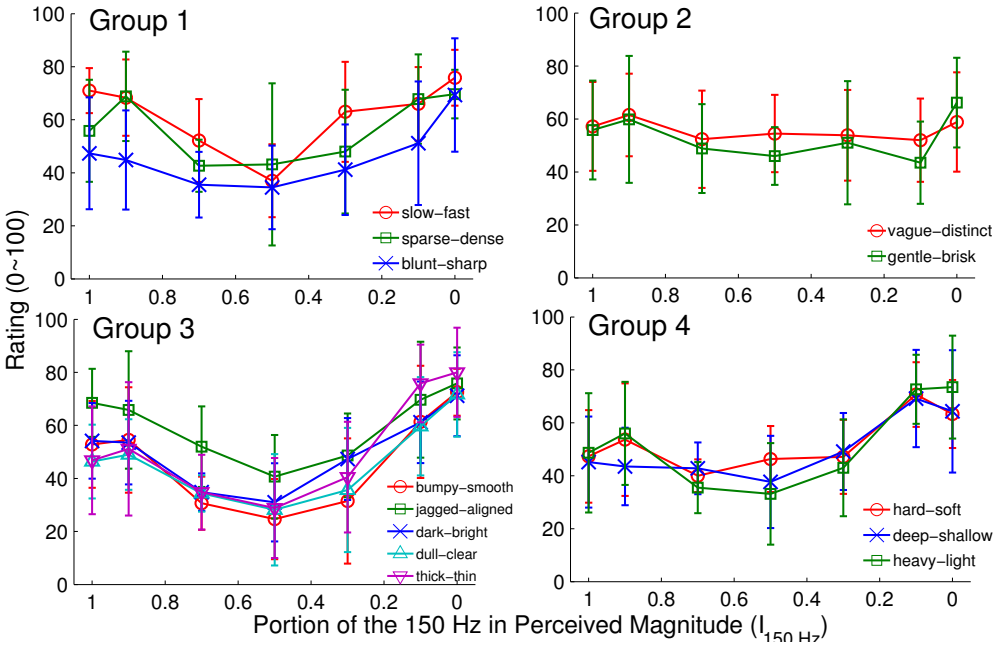

Fig. 4.8 Results of adjective rating in Exp. III. The error bars represent standard deviations.

are close to coordinates of the dominant component. While the mixture ratio closes to even, the coordinates closes to the region of low frequency simple vibrations. In the even mixture level, the bi-freuency stimulus is located near 80 Hz vibration, on a way to 40 Hz vibration. Consequently, we can suppose that the even mixture of 150 Hz and 250 Hz components is felt similarly to a 80 Hz simple vibration rather than its component vibrations, 150 Hz or 250 Hz.

4.3.3 Discussion

In Exp. III, superposition of 150 Hz and 250 Hz sinusoidal vibrations showed the different adjective rating to their components. Superposition of two different frequency components can be thought as amplitude modulation (AM) of a signal. The carrier and modulation frequency are mean and a half difference of two component frequencies, respectively. Therefore, superimposition of 150 Hz and 250 Hz corresponds to 50 Hz modulated 200 Hz carrier signal. In the study of Park and Choi, perceptual space was configured for several modu-

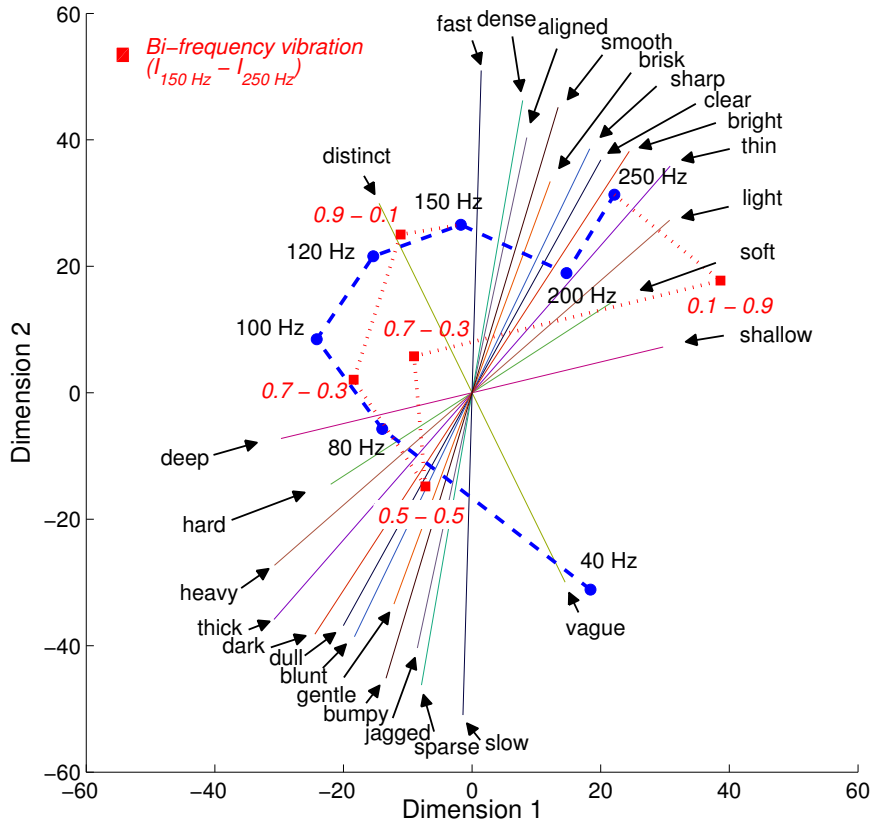

Fig. 4.9 Superimposed bi-frequency vibrations of 150 Hz and 250 Hz projected on the perceptual space in Exp. II.

lation frequencies [63]. Perceptual difference was maximized in 5 Hz modulation in their experiment, with an even intensity in acceleration level. A similar tendency was shown in our previous study [85]. Degree of consonance was lower in smaller frequency differences between two components, which is related to the negative adjectives. We can suppose more distinct effects of superimposition will be shown in smaller differences of two component frequencies.

In this study, the bi-frequency vibrations were not directly compared with simple vibrations. We located the bi-frequency vibrations on the perceptual space with simple vibrations from their adjective ratings. This indirect method imposes possibilities of ignoring hidden

perceptual dimensions, which can classify the simple and superimposed vibrations. In addition, our adjective rating method forced participants to evaluate every scores. All participants were novice to vibrotactile displays and might not have firm perceptual dimensions for the vibrotactile stimuli. Sometimes they might evaluate scores relying on a vague feeling without confidence, which degrades accuracy of the resulted perceptual space. Rating of subjectvie confidence in evaluation could be useful to resolve the effect of forced rating. In this research, we conducted a following study comparing simple and superimposed vibrations on a broad range of frequency to find clear answers for the discussion points.

Chapter 5

Perceptual Characteristics of Bi-frequency Vibration

From the results of Exp. III in Ch. 4, we found some clues about the perceptual characteristics of superimposed bi-frequency vibration. In this chapter, two psychophysical experiments were carried out to explore the perceptual characteristics of superimposed bi-frequency vibrotactile stimuli transmitted to the user's hand via a mobile device. Since we focus on the use of vibrotactile feedback in mobile devices, the frequency range of spectral components and their intensity were determined considering the physical performance of miniature actuators can be equipped in a mobile device. From the experimental results, we analyzed effects of the structural factors that can affect on the percept of the superimposed bi-frequency vibrations. However, some of the structural factors: component frequency, frequency ratio, and absolute frequency difference are dependent and confounded each other. Hence, we carefully selected the tested component frequencies to observe the individual effect of each factor on the perceptual space.

In Exp. I, we examined the effect of intensity mixture ratio between the two frequency components on perceived dissimilarity. A dissimilarity was evaluated from every pair of seven mixture ratios from 0.0:1.0 to 1.0:0.0, for each of three different frequency pairs. In results, the superposed vibration of two equal-intensity components showed the largest perceptual difference from the single frequency vibrations. Next, in Exp. II, we measured the

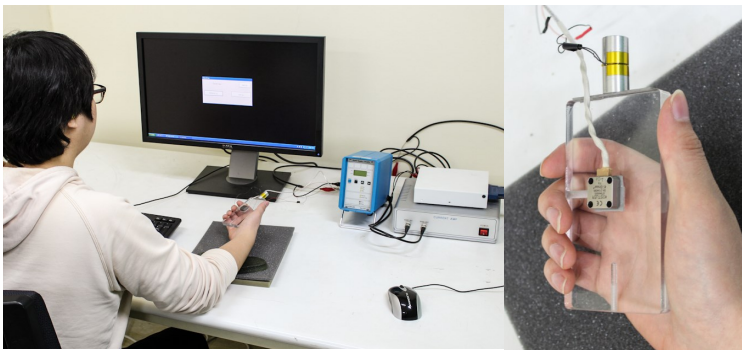

Fig. 5.1 Experimental setup and participant's posture.

dissimilarities among ten equally mixed bi-frequency vibrations and five single frequency vibrations. Then we estimated a two-dimensional perceptual space of the single and bi-frequency vibrations.

The remainder of this chapter is organized as follows. In Section 5.1, general methods common to both experiments are presented. The experimental methods used and the results obtained from Exps. I and II are reported in Sections 5.2 and 5.3, respectively, followed by a general discussion in Section 5.4. The results of experiments were utilized to determine bi-frequency rendering conditions in the bass mode of Haptic Music Player in Ch. 7.

5.1 General Methods

In this section, we describe the experimental methods used commonly in Exps. I and II.

5.1.1 Apparatus

Hardware setup for the experiments are shown in Fig. 5.1. To produce vibrations in a wide range, a miniature linear vibration actuator (Tactile Labs; Haptuator TL002-14-A; weight 12.5 g) was used in both experiments. This actuator is adequate for our experiments in its stronger output amplitude over a broad bandwidth (50–500 Hz, with a weak resonance around 60 Hz) than the other miniature actuators in commercial mobile devices. A PC controlled the actuator via a 16-bit data acquisition (DAQ) board (National Instruments;

model PCI-6251) at a 20-kHz sampling rate. We also used a custom amplifier to supply sufficient current for the operation of the actuator. A Haptuator was attached on the center of the top side of a mobile device mockup by means of an adhesive rubber tape. To measure the generated vibration amplitude, an accelerometer (Kistler; model 7894A500; weight 7.5 g) was attached on middle of the wide face of the mockup. The mockup was made of acrylic resin and has a similar size to commercial mobile phone (110$\times$ 60$\times$ 10 mm). The total moving mass of the mockup was 104.5 g.

5.1.2 Stimuli

In both experiments, a vibration stimulus was continued for 1.5 s. A bi-frequency vibration was generated by addition of two different single frequency components (f_1 and f_2, $f_1 < f_2$). For the single frequency components, our vibratory perceived intensity model for height direction was used to set their amplitudes [30]. Amplitude mixture ratio of a bi-frequency vibration was controlled in a ratio of the two perceived intensity levels from the two frequency components. We used a closed-loop PD control to control the vibration amplitude accurately, avoiding influences of the individual-dependent hand-arm mechanical impedance and other time-varying error sources such as grip force changes.

5.1.3 Procedures

In both experiments, participants sat on a chair and grasped the mockup comfortably with their dominant hand while resting their wrist on a silicon support pad on a table, as shown in Fig. 5.1. They can give their response to the trial via a keyboard using the other hand. Participants wore earplugs to block the operating noise of the actuator which can be an auditory cue.

Each experiment was consisted by two stages: intensity matching and pairwise dissimilarity rating. The intensity matching stage was introduced to avoid the effect of different perceived intensities among the vibration stimuli on dissimilarity rating. In the intensity matching stage, the participants were asked to adjust amplitudes of the tested vibration stimuli to an equal perceived intensity of a reference stimulus. The reference stimulus

was fixed as 140 Hz simple sinusoidal vibration with 3.5 perceived intensity level (about 0.57 G). In a trial, amplitude of a test stimulus can be adjusted linearly in the perceived intensity level by pressing up/down keys on keyboard.

Participants could feel the reference and test stimuli repeatedly without any fixed order. When they felt the two stimuli have a same intensity, they pressed 'NEXT' button on GUI of the experimental program and the final amplitude of the test stimulus was stored. A trial took about 30–40 s and after each trial, they had a 20-s rest. An additional 3-min break was given in every 15 trials to reduce the effect of temporal adaptation. Each test stimulus appeared four times in a session without notifying the repetition to the participants. The initial amplitude of the test stimulus was chosen randomly to be much lower or higher than the expected PSE, evenly two times each, to minimize the effects of the participants' expectation and habituation. There were instructions of the procedure and a 3-min training session prior to the main session.

The next stage was the pairwise dissimilarity rating of the stimuli. Amplitudes of the test stimuli were decided from the participant's mean of the intensity matching results. In a trial, two vibration stimuli were generated in a series with a 2-s interval. After feeling the two stimuli, the participant reported the degree of perceived difference between the two stimuli in 0–100 scale without a reference. A zero rating means the two stimuli were identical and 100 was used to evaluate very different stimuli pairs. The next trial started after a 2-second break. All participants evaluated each vibration pair four times evenly in two presenting orders. They had a short training session of 20 trials to be familiar with the experimental procedure and scale. In every 30–35 trials, they took a rest for 3 mins.

5.1.4 Data Analysis

The measured pairwise dissimilarities were averaged for the four repetitions in a participant. Since there was no reference rating, the participants' scales for dissimilarities are different each other. Hence, we used geometric mean to get a global mean of the dissimilarity between a vibration pair. From the global means, we configured a dissimilarity matrix of the tested vibration stimuli. Using the non-parametric MDS, we estimated a perceptual

space that representing the perceptual distance relationship among the tested stimuli. We used MATLAB 7.14 throughout these analysis processes.

5.2 Exp. I: Effects of Intensity Mixture Ratio

We investigated the significance of intensity mixture ratio on the perceptual space of bi-frequency vibration in mobile device. The results obtained were subsequently taken into account for the design of Exp. II.

5.2.1 Methods

Participants

Twenty participants (10 males and 10 females; 19–27 years old with a mean of 21.9) participated in this experiment. They were everyday users of mobile devices and had no known sensorimotor impairments, both by self-report. Each participant was paid 40,000 KRW (about 36 USD) after the experiment.

Experimental Conditions

We used seven levels of intensity mixture ratio of two frequency components to generate the superimposed vibratory stimuli ($I_{f_1} : I_{f_2}$ = (1.0:0.0), (0.9:0.1), (0.3:0.7), (0.5:0.5), (0.3:0.7), (0.1:0.9), (0.0:1.0). The amplitude of each vibration component was controlled in perceived intensity level estimated from our vibration intensity function for height direction in Section 3.3 [30]. Both ends of the seven conditions, (1.0:0.0) and (0.0:1.0) are identical to the simple sinusoidal vibrations of f_1 and f_2, respectively. The experiment was consisted of three sessions differed by composition of frequencies (f_1 + f_2), which are 50 Hz + 140 Hz, 140 Hz + 230 Hz, and 50 Hz + 230 Hz. In each of the three sessions, a dissimilarity was evaluated at every pair of two stimuli with different mixture ratio conditions. A session had $7 \times 6 \times 2$ = 84 trials including four repetitions for each pair without considering orders of presentation. A session took about 70–90 mins. including the intensity matching of stimuli. Each participant had a session in a day and finished the three sessions in three days.

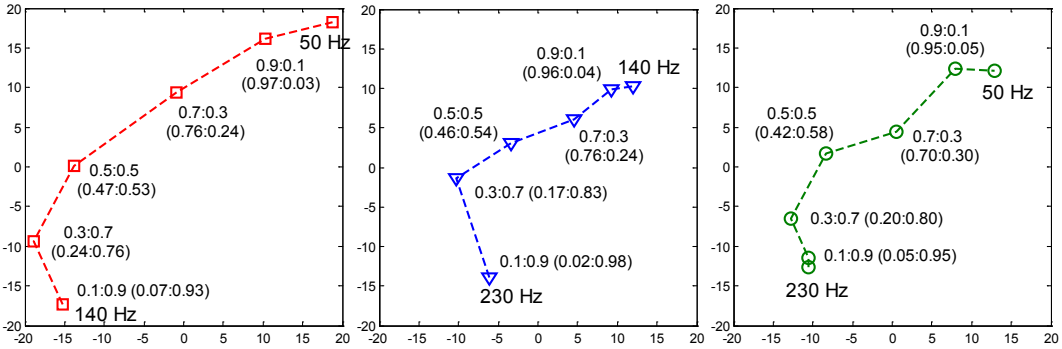

Fig. 5.2 Perceptual spaces obtained from dissimilarities in Exp. I. Amplitude mixture ratios in acceleration were represented in parentheses.

5.2.2 Results

Fig. 5.2 shows the estimated perceptual spaces of three frequency conditions (Kruskal's stress1 < 0.002). Each point on the plot represents a stimulus condition. The plots were rotated to same direction for better comparisons among them. When the percepts of bi-frequency vibrations can be represented by the linear combination of the percepts of two consisting frequency components, all seven points for the stimuli will lie in a line on the perceptual space. However, the traces of points are curved and skewed to the higher end of simple vibrations commonly in the three plots, despite the perceived intensities of all stimuli were calibrated equally. The sum of measured dissimilarities from the two simple vibrations, (0.0:1.0) and (1.0:0.0), was largest at (0.5:0.5) intensity mixture level in 50 Hz+140 Hz and 50 Hz+230 Hz conditions (69.5 and 48.8, respectively). In 50 Hz+230 Hz condition, the sums of dissimilarities were similar among bi-frequency mixture levels. The largest sum was about 16% greater than the distance between two simple vibrations, which is quite smaller than about 22% difference in 50 Hz+140 Hz condition. On contrary, the largest sum of dissimilarities is observed at (0.3:0.7) in 140 Hz+230 Hz condition and the largest sum of distance from simple vibrations was 35% greater than the distance between two simple vibrations. We can suspect that the difference increased with the ratio of two component frequencies. We will discuss about this later with the results of Exp. II.

We also represented the mixture ratio in acceleration unit in parentheses on Fig. 5.2.

Since the perceived intensity of vibration mostly decreases with vibration frequency, when the amplitude of (0.5:0.5) condition in acceleration has a larger portion of a higher frequency component as shown in plots. From the results of Exp. I, we can suspect that the mixture of two equal-intensity frequency components is perceived differently to the simple sinusoids in vibration.

5.3 Exp. II: Perceived Intensity Model

The results of Exp. I showed the perceptual differences of bi-frequency vibrations from simple sinusoidal vibrations. The difference was maximized at (0.5:0.5) intensity mixture level. Hence, in Exp. II, we measured the pairwise dissimilarities of various bi-frequency vibrations of equal-intensity frequency components. Then, a perceptual space was estimated for the tested set of frequency composition. From the estimated perceptual space, we analyzed effects of three structural factors of bi-frequency vibration: component frequencies, within frequency ratio, and absolute within frequency difference.

5.3.1 Methods

Participants

Twenty participants (10 males and 10 females; 19–26 years old, with a mean of 20.9) participated in this experiment. All participants reported that they were users of mobile devices and they had no known sensorimotor impairments. Each participant was paid 40,000 KRW (about 36 USD) after the experiment.

Experimental Conditions

In Exp. II, we used sinusoidal vibrations with five frequencies of 50, 90, 140, 230, and 320 Hz. Ten bi-frequency vibrations were composed by pairwise equal-intensity mixtures of the five frequency components. The frequencies were selected to consist a sequence like Fibonacci series, considering the frequency bandwidth of the haptuator. Use of this sequence allows better observation for the effect of within/between spectral differences in bi-frequency vibrations. Pairwise dissimilarity ratings were conducted for the five sinusoidal

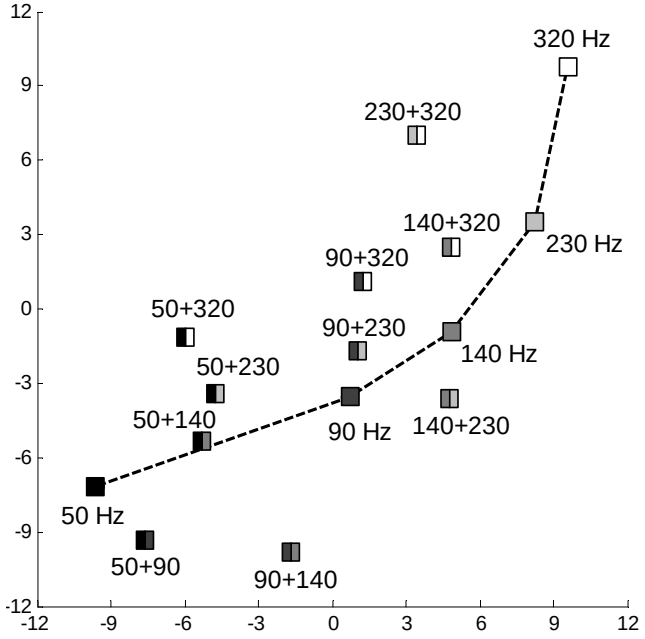

Fig. 5.3 Estimated perceptual space of 15 vibration stimuli in Exp. II.

vibrations and ten bi-frequency vibrations. An experimental session for dissimilarity rating had $15 \times 14 \times 2 = 420$ trials including four repetitions for each pair without considering orders of presentation. The Exp. II took about 3.5 hours including the intensity matching of stimuli. Each participant had a session for intensity matching in the first day and a session for dissimilarity rating in the second day of the experiment.

5.3.2 Results and Discussion

Fig. 5.3 shows the estimated perceptual spaces of the five simple vibrations and ten bi-frequency vibrations of equal-intensity components (Kruskal's stress1 $= 0.077$). The trace of simple vibrations, represented as dotted lines on the plot has an elbow point between 90 Hz and 140 Hz, as shown in our previous study [27].

We analyzed the effects of structural factors of bi-frequency vibrations on the measured perceptual dissimilities. Three structural factors of component frequencies (f_1 and f_2), within frequency ratio (f_2/f_1), and within frequency difference ($f_2 - f_1$) can be arised from

the two component frequencies. Since an independent effect of within frequency difference was not found significantly in our analysis, we only presented analysis for the other two factors below.

First, we can find the dominant effect of the lower component frequency (f_1) than the higher component frequency (f_2) in the measured perceptual dissimilarities. For an example, perceptual distance from 50 Hz+230 Hz condition to 50 Hz is closer than the distance to 230 Hz (8.2 vs. 16.6). In average, distance from a bi-frequency vibration of f_1+f_2 to the simple vibration of f_1 was only 57.5% of the distance to the simple vibration of f_2. On the estimated perceptual space, we can find another effect of component frequency. When one of the two frequency components is fixed, the variation of the coordinates with the change of other frequency component have a distinguished direction to that of simple vibrations.

Effect of within frequency ratio (f_2/f_1) can be seen in comparisons of distances among a bi-frequency condition and its two components. Sum of distances from a bi-frequency condition to its lower component, and to its higher component, $D_1 + D_2$ is equal or longer than the distance between the two components, D_{12}. The ratio between two distances, $(D_1 + D_2)/D_{12}$ can represent how much the bi-frequency vibration is perceptually distinguished from linear sum of simple vibrations. When $f_2/f_1 < 2.0$, the average $(D_1 + D_2)/D_{12}$ was 2.28, which is much larger than the average of 1.07, when $f_2/f_1 \geq 2.0$. Our previous study found a trend of increasing consonances with the within frequency ratio in bi-frequency vibrations [85]. The consonances were saturated when the $f_2/f_1 \geq 2.0$. Since the simple vibrations showed high consonance, the reduced dissimilarity of bi-frequency vibrations can be related to the larger consonance in this frequency ratio range. From these results, we can insist that the lower frequency ratio in a bi-frequency vibration results larger perceptual difference to simple vibration components.

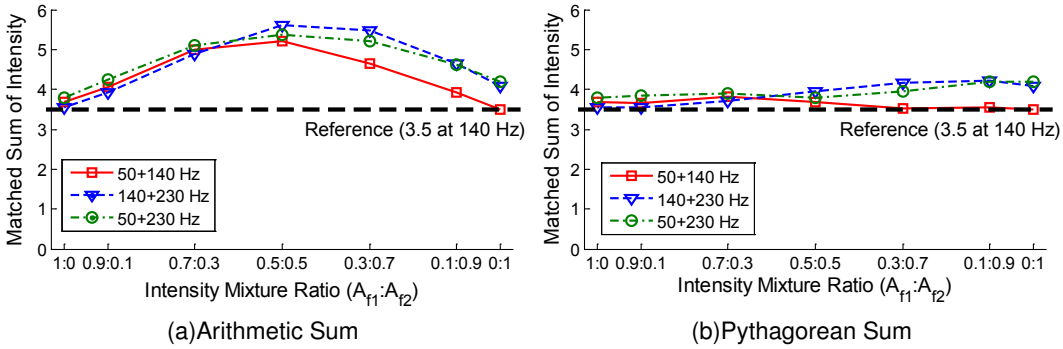

Fig. 5.4 Averaged results of intensity matching in Exp. I. Arithmetic and Pythagorean sums of the perceived intensities for the superposed components. Dotted horizontal line represents reference perceived intensity.

5.4 General Discussion

5.4.1 Perceived Intensity of Bi-frequency Vibration

Spectral Summation of Intensity

In Exp. I and Exp. II, perceived intensities of the tested vibration stimuli were equalized for each participant through the intensity matching process. Fig. 5.4a presents means obtained from the intensity matching in Exp. I. The y-axis means arithmetic sums of the intensity levels for the frequency components of stimuli. The plot shows mostly similar results among the three mixture conditions of component frequencies. When the two frequency components are close to have a same intensity (0.5:0.5), we can see that the sum of intensities should be larger than the reference, to be perceived as an equal intensity. At the (0.5:0.5) condition, the matched arithmetic sum of component intensities were 5.2–5.6. In other words, when the two equal-intensity vibration components are superposed, their overall perceived intensity was 62.5–67.3% of the arithmetic intensity sum. The lifted right sides of 140 Hz + 230 Hz and 50 Hz + 230 Hz can be explained by the effect of deviation error induced from the perceived intensity model for 230 Hz simple vibration. In 50 Hz + 140 Hz condition, both ends of the trace are very closely stuck on the reference level.

To find a better fit for the perceived intensity of bi-frequency vibration stimuli, we tried

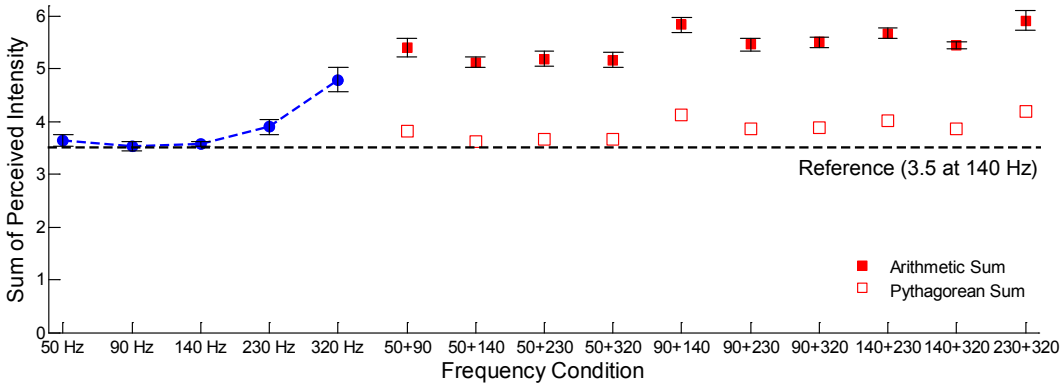

Fig. 5.5 Arithmetic sum and Pythagorean sum of component intensities calculated from the results of intensity matching in Exp. II. Dotted horizontal line represents reference perceived intensity.

to adjust Pythagorean summation and represented the results in Fig. 5.4b. The Pythagorean summation model fits well with the results of intensity matching except some errors on 230 Hz component. From these results, we can insist that the perceived intensity of bi-frequency vibration can be represented by the Pythagorean summation model rather than the arithmetic summation model.

We also compared the two summation models with the intensity matching results of Exp. II. The results are represented in Fig. 5.5. As the results of Exp. I, the Pythagorean summation offers better estimations for the perceived intensities of superposed vibrations. The estimated perceived intensity by Pythagorean summation ranged in 3.6–4.2, slightly higher than the reference 3.5. On the other hand, the intensity by arithmetic summation ranged in 5.1–5.9.

In Fig. 5.5, matched intensity levels for high frequency simple vibrations are showing error induced from the perceived intensity model. In particularly, the error is quite large at 320 Hz, about 40% of the reference intensity. This tendency appeared also in the results at 230 Hz of Exp. I, as shown in Figs. 5.4a and 5.4b. To eliminate the effect of this error, the matched intensity level was linearly scaled using the matched intensity of simple vibrations for each frequency component. We recalculated the arithmetic and Pythagorean sums from

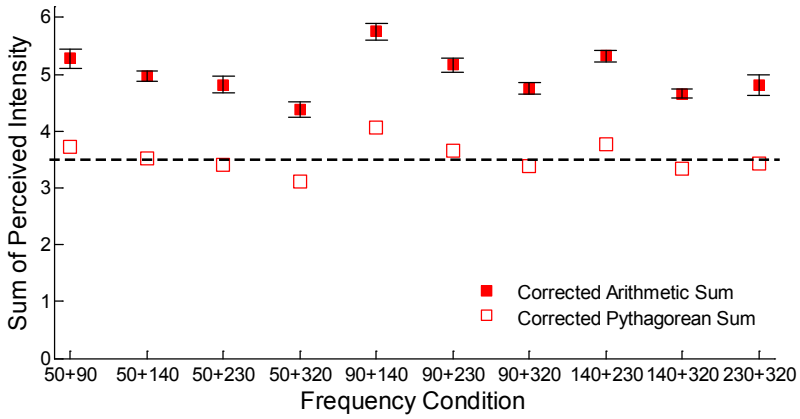

Fig. 5.6 Corrected Arithmetic and Pythagorean sums calculated from the results of intensity matching in Exp. II.

the corrected intensity levels and represented in Fig. 5.6. The corrected Pythagorean sums range in 3.1–4.1, which is much closer to the reference (3.5), than the arithmetic sums range in 4.3–5.8.

In Fig. 5.6, we can see a trend of decreasing sum of intensities with increasing frequency of the higher frequency component. We suspected the relationship between the sum of perceived intensities and frequency ratio between two components, shown in Fig. 5.7. Solid line on the plot is the linear fit of plot and showing the negative correlation (Pearson coefficient $r = -0.70$). This result means that spectral summation of bi-frequency vibration increases with spectral difference of components.

Consequently, the perceived intensity of bi-frequency vibration can be estimated by Pythagorean sum of perceived intensities of components and seems to increase with spectral difference, like the two-tone loudness summation explained by the critical band theory in auditory perception [89]. However, more concrete evidence from further studies is needed to get a fundamental understanding on the characteristic of vibratory spectral summation.

Vibratory Power Model

In our previous study, relationships of the vibration power transmitted to a hand and its perceived intensity were derived [30]. The relationship describes the transmitted vibra-

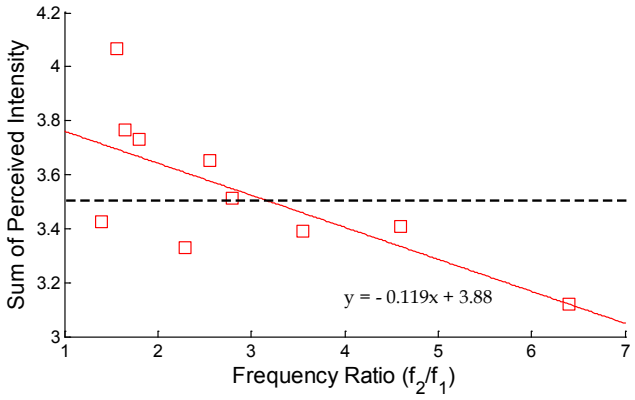

Fig. 5.7 Corrected Pythagorean sum vs. frequency ratio of two components in bi-frequency vibrations.

tion power as a proportional term to A^2/f^2, where A is the acceleration amplitude. A second-order fitting function was derived in the logarithm of transmitted vibration power and perceived intensity. This power-based model can be used without a psychometric data to estimate the perceived intensity of vibration.

We tested feasibility of this model to the bi-frequency vibrations. For a bi-frequency vibration, the power is equal to the sum of two component powers. Using the fitting function for height direction in [30], perceived intensities were estimated for the stimuli in Exp. II, as represented in Fig. 5.8. In the plot, we can see the power-based model tends to overestimate at low frequency region and underestimate at high frequency region, in a range of 2.6–4.8.

We also computed the Pythagorean sum of the estimated component intensity from the power-based model and compared with the intensity of bi-frequency vibration directly estimated from the power-based model. The comparison showed a proportional relationship as shown in Fig. 5.9. Despite more discussions are needed on this topic yet, the current results imply the feasibility of our perceived intensity estimation based on the transmitted vibration power in bi-frequency vibration.

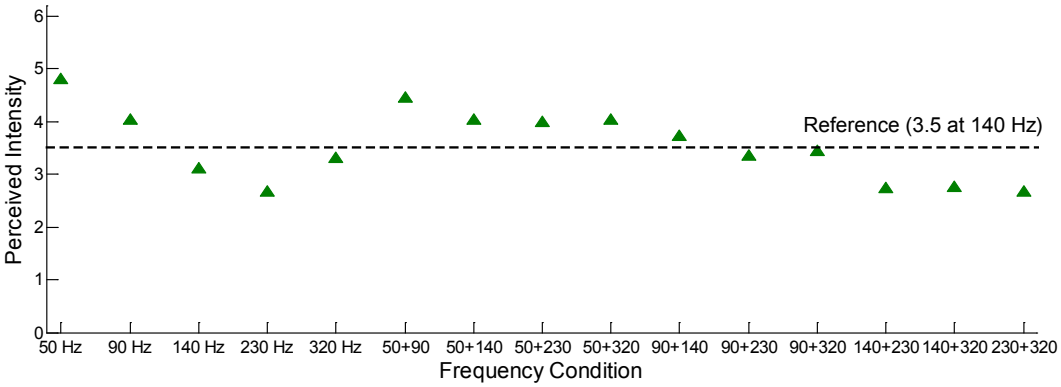

Fig. 5.8 Perceived intensity of the vibration stimuli in Exp. II, estimated from the transmitted vibratory power.

Error in Perceived Intensity

In the results of intensity matching, mismatch of the perceived intensity model and matched intensity was found in high-frequency simple vibrations (230 and 320 Hz).

As a primary error source, we may suspect error from fitting of the measured perceived intensity in our previous study. However, the fit showed a high R^2 value ($>$0.99), and the fitting errors near the reference intensity level (3.5) were about 10% at 250 Hz and 1.4% at 320 Hz. This error range is quite smaller than the error occurred in this study, particularly at 320 Hz.

Another plausible sources of error can be different participants group in our previous and current studies. The standard errors between participants were less than 10% of the measured perceived intensities for high frequency conditions ($>$ 230 Hz) in our previous and current studies.

Difference in the experimental methods of our two studies also may contributed to the error. Our previous study used the absolute magnitude estimation method. Since there was no reference or modulus, the results are subject to be influenced by the stimulus context effect [15]. Meanwhile, we used the method of adjustment with a 140 Hz reference stimulus in the current study. The participants felt a vibration repeatedly to compare the perceived intensities of the reference stimulus and a test stimulus. Despite a 20-s rest was forced be-

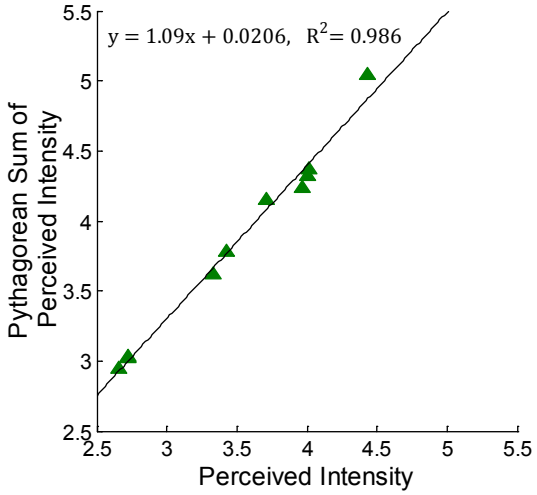

Fig. 5.9 Pythagorean sum of component intensities vs. bi-frequency perceived intensity, both estimated from power-based model in [30].

tween every two conditions, sensory adaptation of the Pacinian receptors can be suspected. Reduced sensitivity of the Pacinian channel can be a plausible reason for the large error of intensity matching in high frequency range. However, a firm evidence for the effects of these error sources could not be found in this study.

5.4.2 Scale of Perceptual Space

In Exp. I, we estimated a perceptual spaces for each of the three frequency conditions. The dissimilarity ratings were conducted by same participants in different days. We compared the spanned length of each perceptual space to test the participants' consistency in dissimilarity rating. When we consider the dissimilarity relationship presented in Fig. 5.3, 230 Hz vibration should be more distanced than 140 Hz vibration from 50 Hz vibration. Surprisingly, the spanned length at 50 Hz+140 Hz condition was the longest (48.9), and 50 Hz+230 Hz was the second (34.0) and 140 Hz+230 Hz was the shortest (30.1).

This inconsistency in dissimilarity rating can be resulted to the stimulus context effect [15]. In a session, participants might establish their 0–100 scale for the given stimulus condition. Also the contrast between stimuli would affected participants' rating criteria.

Hence, the perceptual distances measured in Exp. I should not be directly compared across the frequency conditions and only have their meaning in a same frequency condition. For the comparison among frequency conditions, a reference for dissimilarity will be needed in further studies.

Chapter 6

Haptic Music Player

In this chapter, the author presents the initial version of haptic music player for mobile devices developed to enrich music listening experience. The haptic music player has the following four major features.

First, we use a new miniature actuator called *Dual-Mode Actuator (DMA)*. The DMA can produce vibrations composed of two principal frequencies, which can lead to greater diversity in vibrotactile perception [2, 63, 49]. This is in contrast to the vast majority of commercial mobile devices that use a simple actuator, e.g., an Eccentric Rotating Mass (ERM) or Linear Resonance Actuator (LRA). The dynamic performance of these actuators is insufficient for creating expressive vibrotactile effects for haptic music.

Second, the haptic music player enables *real-time*, on-the-fly playback of vibrotactile effects. Since thousands of new musical pieces are published every year, producing vibrotactile music directly from musical sources without any preprocessing is a highly desirable benefit. Our vibration generation algorithms satisfy this requirement using digital signal processing techniques with very low computational complexity.

Third, our haptic music player supports *dual-band vibration playback*. As reviewed earlier, previous attempts for haptic music playback relied on the rhythms or beats extracted from the bass band of sound signals. This was attributable partly to the performance limits of the vibration actuators used. Other salient aspects of music, such as a singer's voice or

guitar solo, were ignored. In our haptic music player, the rhythmic variations in music are encoded in a low-pitch vibration signal (bass band), whereas high-frequency salient sounds are transmitted in a high-frequency signal (treble band), both produced by one DMA. To deal with high-frequency music saliency that varies greatly among music genres, we introduce the concept of a haptic equalizer. The haptic equalizer mixes the signal energies from different frequency bands using genre-dependent weights, analogously to an audio equalizer.

Lastly, in our haptic music player, all the conversion and scaling processes between sound and vibrotactile signals are based on *perceptual data* taken from the relevant literature. Human perception of vibrotactile stimuli is complex and depends on various factors such as signal frequency, contact site, and contact area. However, the previous methods of vibrotactile music tended to control the physical amplitude of vibrotactile stimulus, without explicit consideration of their perceptual consequences. We also compensate for the actuator input/output relationships to minimize the perceptual distortions that might occur otherwise because of the actuator dynamics and the human perception process.

In addition, a user study was conducted to evaluate the perceptual merits of our dual-band vibrotactile music playback in comparison with the conventional single-band playback. Two types of actuators (LRA and DMA) were used, and 16 musical pieces were selected to represent four music genres (rock, dance, classical, and vocal; four pieces each). The experimental results elucidated the benefits of dual-band playback and their dependence on music genre, also providing insights that can facilitate further improvements of simultaneous audio-haptic music playback.

6.1 Software

The software structure and algorithms of the haptic music player are described in this section.

6.1.1 Structure

Fig. 6.1 shows the overall structure and computational processes of our haptic music player. The first step is to open a music source file and store it as a timed sequence. Each element in the sequence represents a sound amplitude. The sequence is then partitioned to a number of segments with the same length. Fast Fourier Transform (FFT) is applied to each signal segment to extract the spectral densities between 0 Hz and the Nyquist frequency (half of the music sampling frequency). The segment length can be determined based on the processing power of the computing platform.

The spectral densities are divided into two bands, which are denoted as "bass" and "treble" in Fig. 6.1. The bass band is used to extract the beat information. This information is played through the superimposed, perceptually low-pitch vibrations of DMA. The treble band, which is unique to our haptic music player, tracks high-frequency salient features of music. These features are delivered by the high-frequency signal of DMA. To handle the saliency dependence on music genres, the treble band is further partitioned into many sub-bands. The signal energies of the sub-bands are then merged using a haptic equalizer with mixing weights dependent on music genre.

The final step is a nonlinear scaling procedure that converts the extracted audio energies into the voltage commands of DMA in (2.1). The overall conversion process is: sound signal energy $\rightarrow$ auditory perceived magnitude $\rightarrow$ vibratory perceived magnitude $\rightarrow$ physical vibration amplitude $\rightarrow$ voltage command amplitude to the DMA. This transformation uses appropriate perceptual data acquired from the literature.

The entire procedures are repeated as a loop for each music signal segment until the end of music playback. For real-time performance, our computational algorithms are designed to be as efficient as possible while maintaining perceptual plausibility, as described further in the remainder of this section. The current implementation is tailored to DMA, but it can be extended to other wideband vibration actuators.

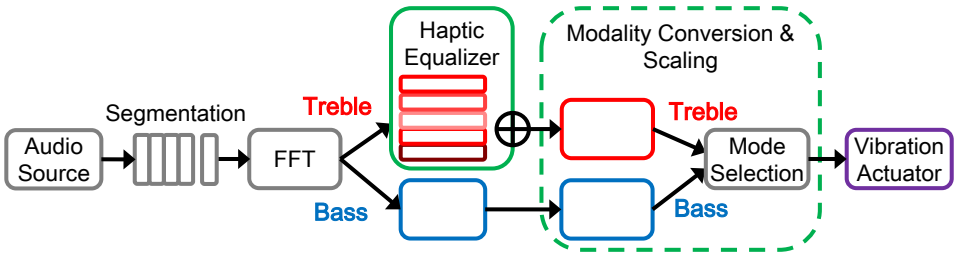

Fig. 6.1 Process loop of the haptic music player.

6.1.2 Haptic Equalizer

Musical instruments have their own frequency band. For example, the general frequency band of human voice is 80–1,300 Hz, while that of a drum is 50–1,000 Hz. We also examined the frequency ranges of various pop songs and found that the bass accompaniment was in 50–200 Hz, vocal sound was in 200–1,500 Hz, while the treble percussive sound was around 4,000 Hz. Thus, in order to fully exploit the dual-vibration playback mode, feature extraction algorithms need to handle this wide frequency range. Our approach for this requirement involves the use of a haptic equalizer similar to the audio equalizer found in audio systems and traditional tactile vocoders [7].

We partition a music signal of frequencies ranging from 0 to 6,400 Hz into the bass and treble bands at a 200 Hz boundary. This is the most common setting for bass-treble separation in audio systems. The treble band is further divided into five sub-bands, linearly in a log scale: 200–400, 400–800, 800–1,600, 1,600–3,200, and 3,200–6,400 Hz. Separate mixing weights are assigned to each of the five frequency bands. Each weight determines the amplification gain of the corresponding frequency band.

Table 6.1 shows the preset weights used in our haptic music player for four representative music genres: rock, dance, classical, and vocal. Their initial values were taken from the equalizer gains of a popular music player program (jetAudio; Cowon Systems Inc.). These weights were then tuned manually to better express the genre characteristics via touch. The preset weights for rock music emphasize the low-frequency bass guitar sounds and the high-frequency percussive drum sounds. The weights used for dance music were

Table 6.1 Preset weights of haptic equalizer for music genres.

Frequency (Hz)	Rock	Dance	Classical	Vocal
200–400	0.25	0.25	0.25	0.05
400–800	0.15	0.15	0.20	0.15
800–1,600	0.12	0.15	0.15	0.70
1,600–3,200	0.18	0.20	0.15	0.05
3,200–6,400	0.30	0.25	0.25	0.05

adjusted slightly from the rock preset for milder expression. The classical preset has evenly-distributed weights with some emphasis on 400–800 Hz, which is the main frequency band of many classical music instruments. The focus of the vocal preset is on the high tone voice of a singer. A user can freely adjust the preset weights according to their preference, thereby providing maximum control to the user.

6.1.3 Modality Conversion and Intensity Scaling

The signal energies that are divided and weighted by the haptic equalizer are converted into voltage commands to DMA, $V_1(n)$ and $V_2(n)$ in (2.1), following several computational steps that rely on relevant perceptual data. Compared with previous approaches, this step is unique to our haptic music player.

Computation of Sound Signal Energies

Let $E_{bass}(n)$ be the sound signal energy of the bass band and $E_i(n)$ be that of the i-th sub-band in the treble band. The first step is to compute the signal energies by summating the absolute spectral magnitudes over all frequencies in the corresponding band.

Conversion to Auditory Perceived Magnitudes

We use a two-step procedure to convert the sound signal energies to two auditory perceived magnitudes, $L_{bass}(n)$ and $L_{treble}(n)$, for the bass and treble bands, respectively. First, we

compute intermediate auditory perceived magnitudes, $A_{bass}(n)$ and $A_{treble}(n)$, as

$$A_{bass}(n) = \left\{ \frac{E_{bass}(n)}{l} \right\}^{e}, \tag{6.1}$$

$$A_{treble}(n) = \left\{ \sum_{i=1}^{5} w_i \frac{E_i(n)}{l} \right\}^{e}, \tag{6.2}$$

where l is the length of a signal segment and w_i is the weight of the i-th sub-band in the haptic equalizer. Since $E(n)$ represents the integrated energy in a wide frequency band, its exact transformation to auditory loudness can be very complex, and it is more so with perceptual weights. (6.1) and (6.2) are plausible approximations based on Stevens' power law for fast computation. Stevens' power law provides a well-defined relation between physical and perceptual magnitudes for simple auditory stimuli [77]. In our implementation, the exponent $e = 0.67$ was derived from the experimental data for 3 kHz tone loudness [77].

Next, we calculate an additional gain $g(n)$ for the treble band and then determine $L_{bass}(n)$ and $L_{treble}(n)$, such that

$$L_{bass}(n) = A_{bass}(n), \tag{6.3}$$

$$L_{treble}(n) = g(n)A_{treble}(n). \tag{6.4}$$

During our initial implementation, we realized that sub-band sound is very salient in terms of perception when the energy of the sub-band is significantly greater than those of the other sub-bands, e.g., during the solo performance of an instrument. However, the total signal energy of the treble band reflected in $A_{treble}(n)$ may not be strong enough to deliver this dominant band effect. To compensate for this, we introduce $g(n)$ ($1 \leq g(n) \leq 2$). If $A_{treble}(n) \leq A_t$,

$$g(n) = 1 + \frac{A_t - A_{treble}(n)}{A_t} \frac{w_{max}E_{max}(n)}{\sum_{i=1}^{5} w_i E_i(n)}. \tag{6.5}$$

If $A_{treble}(n) > A_t$, $g(n) = 1$. Here, $E_{max}(n)$ represents the energy of the sub-band with the maximum energy, and w_{max} is the weight of that sub-band. The rightmost term converges to 1 if $E_{max}(n)$ approaches the total energy of the treble band. The middle term that includes $A_{treble}(n)$ reduces the effect of dominant band amplification if the total

treble band energy increases, preventing overcompensation. Dominant band amplification is activated only if $A_{treble}(n) \leq A_t$, also to avoid overcompensation. We set A_t to the 90% level of the range of auditory perceived magnitudes.

Conversion to Vibratory Perceived Magnitudes

Matching the scales of perceived magnitudes between sound and vibration requires great care. The absolute sound levels in music files vary significantly from file to file, and vibration actuators also have different ranges of producible vibration strength. Our approach relies on our previous study [29], which presented a psychophysical magnitude function of vibration frequency and amplitude to the resulting perceived magnitude in a wide parameter range for a mobile device held in the hand. One can measure the range of vibration amplitudes generated by an actuator at a given frequency, input this range into the perceived magnitude function, and obtain a range of vibratory perceived magnitudes. The range of auditory perceived magnitudes can then be scaled to match the range of vibratory perceived magnitudes. Nonetheless, the physical signal level of music files is still beyond our control, which demands a "haptic volume" control.

After the scales are determined for the auditory and vibratory perceived magnitudes, we can compute the desired perceived magnitudes of vibration, $I_{bass}(n)$ and $I_{treble}(n)$, from $L_{bass}(n)$ and $L_{treble}(n)$ as follows:

$$I_{bass}(n) = w_{bass} c L_{bass}(n), \tag{6.6}$$

$$I_{treble}(n) = w_{treble} c L_{treble}(n), \tag{6.7}$$

where c is the cross-modal scaling constant, and w_{bass} and w_{treble} are the amplification gains of the bass and treble bands in charge of haptic volume control.

Rendering Mode Selection

To drive DMA, a subtle adjustment is required on the perceived intensity because of its superposition mode. Activating DMA with positive $V_1(n)$ and $V_2(n)$ using (2.1) generates superimposed vibration of the frequency f_1 and f_2. The sensation of such vibration is

much rougher and feels like a lower frequency than f_1 or f_2 [63, 49]. Therefore, we use this superposition mode to render bass signals. When the treble intensity is dominant, we only use vibration frequency f_2, i.e., $V_1(n) = 0$. Furthermore, we observed via a magnitude matching experiment that the superimposed vibration of two equally-intensive vibrations with frequencies f_1 and f_2 has a perceived magnitude that is about 1.25 times higher than the perceived magnitude of each individual vibration. To compensate for this, we scale down the individual perceived magnitudes of the two vibrations by 0.8 times in the superposition mode, so that the resulting superimposed vibration would have the same perceived magnitude as the bass band auditory perceived magnitude $I_{bass}(n)$. In summary, the desired perceived magnitudes for two vibration frequencies f_1 and f_2, $P_1(n)$ and $P_2(n)$, are as follows: If $I_{bass}(n) > I_{treble}(n)$, then

$$P_1(n) = 0.8I_{bass}(n) \text{ and } P_2(n) = 0.8I_{bass}(n). \tag{6.8}$$

If $I_{bass}(n) \leq I_{treble}(n)$, then

$$P_1(n) = 0 \text{ and } P_2(n) = I_{treble}(n). \tag{6.9}$$

An example demonstrating this rule is shown in Fig. 6.2. The high frequency unit (for f_2) is always activated, whereas the low frequency unit (for f_1) is activated only when the bass component is dominant.

Conversion to Physical Vibration Amplitudes

The desired perceived magnitudes of the two vibrations, $P_1(n)$ and $P_2(n)$, can be readily converted to the desired vibration amplitudes at frequencies f_1 and f_2 using the inverse of the perceived magnitude function of vibratory stimuli [29].

Conversion to Voltage Command Amplitudes

The final step is to convert the desired vibration amplitudes to voltage command amplitudes for f_1 and f_2. To do this, we need the I/O mappings of DMA for each frequency, which are derived from input voltage amplitude to output vibration amplitude measured when the

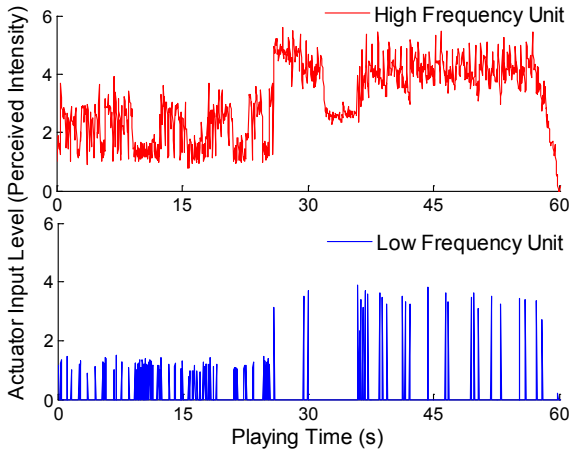

Fig. 6.2 Example of input signal to a DMA for dual-band playback.

DMA is attached to a mobile device. This I/O calibration can be done easily (e.g., see [31]), especially for DMA which has fairly linear responses at both resonance frequencies. The input voltage amplitudes $V_1(n)$ and $V_2(n)$ are determined using these I/O relations.

6.1.4 Implementation and Processing Speed

We implemented the algorithms described above on an MS Windows platform, using MS Visual C++ 2008 with external libraries for music file I/O (Audiere 1.9.4) and FFT calculation (FFTw 3.2.2). The haptic music player ran on a desktop PC (3 GHz Intel Core 2 Duo) because of the difficulty of custom signal I/O for DMA on commercial mobile platforms.

A parameter critical for the performance of our haptic music player is the length of a music segment processed in each loop. Increasing the segment length improves the frequency resolution in spectral density estimation. However, it degrades the smoothness of vibration updates and also leads to a longer processing time because of the increased computational load for FFT and subsequent operations.

After extensive tests, we set the length of a music segment to 50 ms as the best trade-off. This value allows a 20 Hz update rate for vibration playback using each 2,205 samples from a music source sampled at 44.1 kHz. This update rate is sufficient for smooth transitions between music segments without causing any perceptible discontinuity and also for fast

responses synchronized with audio playback.

In our desktop system, a single loop took 0.3 ms on average for vibration command extraction. We also ported these Windows codes to Android to assess viability on mobile platforms. When tested with a smartphone (Samsung Electronics; Galaxy S2; without generating vibrations), the Android version took 1.0 ms on average for processing a single loop, including the most time-consuming audio file decoding. Compared to the 50 ms processing interval for 20-Hz updates, this performance clearly allows for real-time rendering with a very low computational burden. Note that the current mobile devices are equipped with a faster multi-core CPU than one included in the smartphone used for our test.

6.2 User Study

We evaluated the subjective performance of our dual-band vibration extraction algorithm compared with the bass-band only algorithm for four music genres via a user study. Details are presented in this section.

6.2.1 Methods

Participants

Twenty four university students (12 males and 12 females) participated in this experiment. They were 18–30 years old with a mean 22.3 (SD 3.5). Young participants were preferred as they are generally more enthusiastic in listening to music and accepting new technology and interfaces. All participants were daily users of a mobile phone with no known sensorimotor impairment. They were paid KRW 20,000 ($\simeq$ USD 17) for the experiment.

Apparatus

We used LRA (LG Innotek; model MVMU-A360G) and DMA (LG Electronics; a prototype model) as a vibration actuator. The resonance frequency of LRA was 178 Hz, while those of DMA were 150 and 223 Hz. Each actuator was attached to one wide face of a handheld mockup made from acrylic resin ($105 \times 45 \times 15$ mm), as shown in Fig. 6.3. The actuators were not in direct contact with the participants' hand. Their input-output relations were

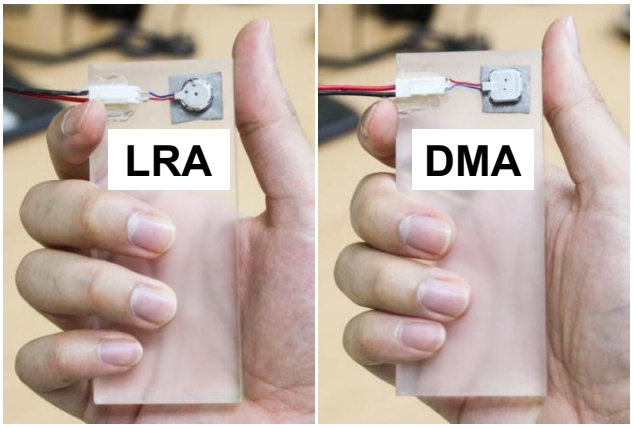

Fig. 6.3 Handheld mockups with two vibration actuators (LRA and DMA) used in the user study.

calibrated using a miniature accelerometer (Kister; model 7894A500; 7.5 g) attached to the center of the mockup. The actuators were controlled by a PC via a data acquisition board (National Instruments; model USB-6251) with a custom-made power amplifier. The sampling rate for signal I/O was 10 kHz for faithful signal sampling and reconstruction.

Experimental Conditions

This study consisted of 16 experimental conditions (2 rendering modes $\times$ 2 actuators $\times$ 4 music genres) in a within-subjects design. For vibration rendering, we used two methods: single- and dual-band modes (SINGLE vs. DUAL). The single-band mode represents the current standard and presents only a bass-band signal. The dual-band mode, the main function of our work, provides both bass- and treble-band vibrations, which is unique to our haptic music player.

As an actuator, LRA or DMA was used. When LRA was used, the single-band mode produced 178-Hz resonance vibrations using only bass signals, i. e., $P_1(n) = I_{bass}(n)$ and $P_2(n) = 0$. In the dual-band mode, both bass and treble signals were encoded in the 178-Hz vibrations using the same algorithm as DMA. The special magnitude adjustments required for the superposition mode of DMA are not necessary for LRA, thus (7.5) became $P_1(n) =$

$I_{bass}(n)$ and $P_2(n) = 0$ and (7.6) remained intact. When DMA was used, the single-band mode used superimposed vibrations (mix of 150- and 223-Hz vibrations) to render bass signals. In the dual-band mode, bass signals were expressed by the superposition method, while treble signals were in 223-Hz vibrations. The maximum vibration amplitude was set to level 6 of the perceived intensity model in [29] (about 0.5, 0.6, and 0.8 G at 150, 178, and 223 Hz, respectively).

In pilot experiments, we realized that participants' preference of vibration playback depends on music genre to a great extent. Thus, we included music genre as an independent factor in the experiment. We selected four genres of rock, dance, classical, and vocal, and chose four music pieces per genre, as listed in Table 6.2. They are familiar music pieces to our Korean participants containing the styles representative of the corresponding genres. For playback, we trimmed one minute of each musical piece and concatenated them for each music genre. Each music clip was played with the preset weights of the corresponding genre shown in Table 6.1. This set of equalizer gains was found by the experimenter to best express the genre characteristics via vibrotactile stimulation. For all music, haptic volume was set to be identical by the experimenter: $c = 0.001$ and $(w_{bass} : w_{treble}) = (6 : 7)$.

Both equalizer gains and haptic volume can affect the subjective preference of vibration playback, but we were unable to include them as independent factors in the experiment because of the large number of continuous variables involved (20 for equalizer gains and 3 for haptic volume). Instead, we used the fixed values found by the experimenter to be the best, concentrating more on our major interests (the effects of rendering method and actuator).

Subjective Performance Measures

We collected four subjective measures using a questionnaire in a 0–100 continuous scale. They were: Precision–"Did the vibration express the music precisely?" (0: very imprecise, 100: very precise); Harmony–"Was the vibration harmonious with the music?" (0: very inharmonious, 100: very harmonious); Fun–"Was the vibration fun?" (0: very boring, 100: very fun); and Preference–"Did you like the vibration?" (0: dislike very much, 100:

Table 6.2 The 16 genre-representative musical pieces used for evaluation.

Genre	Music title
Rock	Don't Look Back In Anger - *Oasis*
	Time Is Running Out - *Muse*
	It's My Life - *Bon Jovi*
	Basket Case - *Green Day*
Dance	It's Gonna Be Me - *'N Sync*
	Let's Get It Started - *The Black Eyed Peas*
	Livin' La Vida Loca - *Ricky Martin*
	Billy Jean - *Michael Jackson*
Classical	Ouverture Solennelle '1812' - *P. Tchaikovsky*
	Violin Concerto in E major, BWV 1042 - *J. S. Bach*
	Cellokonzert, C-dur, Hob.VIIb:1 - *F. Haydn*
	Pomp and Circumstance Marches - *E. Elgar*
Vocal	If I Ain't Got You - *Alicia Keys*
	Because Of You - *Kelly Clarkson*
	Falling Slowly - *Glen Hansard* and *Marketa Irglova*
	You Raise Me Up - *Westlife*

like very much). The participants also described the subjective impressions of vibrotactile feedback in a free form.

Procedure

Prior to the experiment, each participant was given instructions about the experimental procedures and explanations of the meanings of the questions in the questionnaire. A training session was then followed, where two songs that were not used in the main sessions were played using each of the 4 vibration rendering conditions (2 modes $\times$ 2 actuators) for 2 min.

The main experiment consisted of four sessions. Each session used one of the 4-min genre-representative music clips. At the beginning, the music clip was played without vibrotactile playback using over-ear headphones to provide perceptual reference. The participant could adjust audio volume to a comfortable level. Then, the music clip was played with vibration using one of the four rendering conditions. After the playback, the partic-

ipant answered the questions on the rendering conditions using the questionnaire sheets. For each performance metric, the participant gave a score by marking a position on a line labeled on both ends with their meanings. The participant had a rest for a few minutes to prevent tactile adaptation before proceeding to the next rendering condition. This procedure was repeated four times with different rendering conditions.

To remove any possible order effects, we randomized the orders of the rendering conditions in each session and those of the music genres. The entire experiment took about 2.5 hours, and each participant finished it in two days (two main sessions per day).

6.2.2 Results

A within-subjects three-way analysis of variance (ANOVA) was conducted, where rendering mode, actuator, and music genre were fixed-effect factors and subject was treated as a random effect factor. To see the influence of rendering mode (DUAL vs. SINGLE) and actuator (DMA vs. LRA) on the results more clearly, we also provide the analysis results classified by music genre. In the analysis, we excluded data of one male participant, treating him as an outlier. Most of his data lied outside of the confidence intervals, and he reported that he was a hip-hop dancer trained to respond to bass-beat sounds.

ANOVA Results

We presented the results of the three-way ANOVA with the effect sizes (η^2) of the main effects in Table 6.3. In all measures, the effects of rendering mode (R) and music genre (M) were strongly significant ($p < 0.01$), while the effect of actuator (A) was not. Interactions of A$\times$M and R$\times$A$\times$M were significant ($p < 0.05$) in fun and marginally significant ($p < 0.1$) in preference. R$\times$A$\times$M also had marginal significance in harmony.

Since our main interest was finding the effects of A and R, we then conducted a two-way ANOVA for each of the four music genres. The results are summarized in Table 6.4. In most cases, R had very significant effects, consistent with the results of the three-way ANOVA (DUAL$>$SINGLE). The effect of A was significant in fun and preference of dance music and in preference of classical music. Marginal significance of A was also seen in harmony

Table 6.3 Three-way ANOVA results (F-ratios) with effect size (in parentheses)

Source	Precision	Harmony	Fun	Preference
Rendering (R)	***39.39 (0.641)	***28.02 (0.560)	***35.88 (0.620)	***15.68 (0.088)
Actuator (A)	0.45 (0.020)	1.57 (0.067)	1.09 (0.047)	2.11 (0.088)
Music (M)	***12.12 (0.355)	***9.92 (0.311)	***11.47 (0.343)	***6.27 (0.222)
R×A	1.68	0.92	1.35	0.20
R×M	0.82	0.76	2.12	1.73
A×M	2.13	1.75	**3.14	*2.44
R×A×M	2.13	*2.47	**3.70	*2.31

$^{*}: p < 0.10$, $^{**}: p < 0.05$, $^{***}: p < 0.01$

of classical music and precision of vocal music. The interaction effect R×A was significant in fun and marginally significant in the other three measures of dance music.

Overall, the results indicated that the dual-band rendering, the main function of our haptic music player, improved the users' evaluations of music listening compared with the current standard of single-band, bass-only rendering.

Effects of Rendering Conditions

Score differences were analyzed among the four rendering conditions. The average evaluation scores in Fig. 6.4 show much higher scores for DUAL than for SINGLE in all measures, while the scores of DMA and LRA were mostly comparable. We separated the scores by music genre in Fig. 6.5 and conducted the Student-Newman-Keuls (SNK) multiple comparison test. Table 6.5 shows the grouping results. We summarized noteworthy results below, emphasizing comparisons between DUAL-DMA (our main contribution) and SINGLE-LRA (current standard).

For rock music, in precision and harmony, significant differences ($p < 0.05$) were found between DUAL-DMA and the two SINGLE conditions, and marginally significant differences ($p < 0.1$) were found between DUAL-DMA and DUAL-LRA, with the much higher

Table 6.4 Two-way ANOVA results for the four music genres.

Genre	Source	Precision	Harmony	Fun	Preference
Rock	R	$^{***}24.37$	$^{***}10.00$	$^{*}3.36$	1.72
	A	2.61	1.82	1.17	0.50
	R×A	1.07	1.38	1.27	0.05
Dance	R	$^{***}15.19$	$^{***}18.62$	$^{***}13.92$	$^{***}13.57$
	A	0.99	1.47	$^{**}5.97$	$^{**}4.61$
	R×A	$^{*}3.30$	$^{*}3.50$	$^{***}8.54$	$^{*}3.78$
Classical	R	$^{***}30.47$	$^{***}14.31$	$^{***}32.16$	$^{***}12.70$
	A	0.18	$^{*}3.48$	2.71	$^{**}5.01$
	R×A	1.76	2.07	2.14	1.38
Vocal	R	$^{***}8.25$	$^{**}6.20$	$^{***}8.94$	$^{*}4.20$
	A	$^{*}3.45$	0.39	< 0.01	0.08
	R×A	2.18	0.25	0.09	0.14

scores of DUAL-DMA. In particular, DUAL-DMA and SINGLE-LRA showed significant differences in precision and harmony. Hence, it can be said that DUAL-DMA, which acquired the highest scores in all the four measures, was the best rendering mode for rock music.

For dance music, a significant difference was seen between SINGLE-DMA and the two DUAL conditions in precision. In the other three measures, significant differences were between SINGLE-DMA and the other three conditions. In all of these cases, SINGLE-DMA exhibited the lowest scores. Between DUAL-DMA and SINGLE-LRA, no significant difference was present in any of the four measures, although the scores of DUAL-DMA were higher. In summary, SINGLE-DMA was the lowest-rated rendering condition for dance music, and the two DUAL conditions were comparable without evident effects of actuator.

For classical music, a significant difference was found in precision between the DUAL and SINGLE conditions. In harmony, fun, and preference, DUAL-LRA and the other three conditions showed significance differences. In all of these cases, DUAL-LRA resulted in the highest scores. DUAL-DMA and SINGLE-LRA showed significance differences in

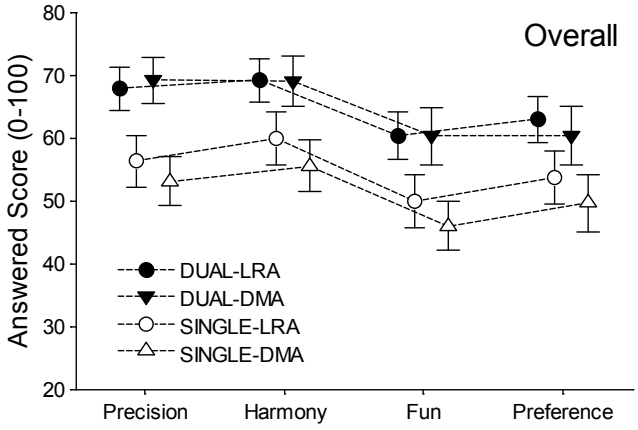

Fig. 6.4 Average evaluation results of the four rendering conditions. Error bars represent standard errors.

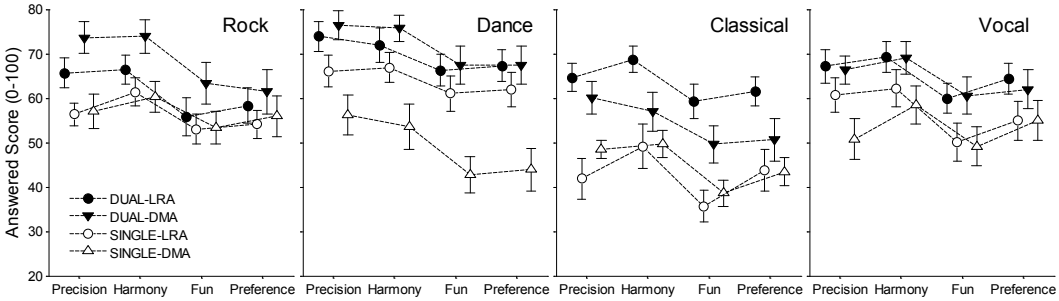

Fig. 6.5 Evaluation results of the four rendering conditions by music genre.

precision and fun, with the higher scores of DUAL-DMA.

For the vocal-oriented songs, a significant difference occurred in precision between SINGLE-DMA and the other conditions, with the lowest score of SINGLE-DMA. In fun, there was a marginally significant difference between DUAL and SINGLE. No significant difference was seen between DUAL-DMA and SINGLE-LRA in any of the four measures, despite the higher scores of DUAL-DMA. Hence, the DUAL conditions were generally better than the SINGLE conditions for vocal music.

Overall, our new rendering method, DUAL-DMA, scored as the highest for rock, dance, and vocal music, except classical music where DUAL-LRA obtained the best scores. DUAL-

Table 6.5 Grouping of rendering methods by the SNK test (α=0.05; α=0.1 in parentheses). The rendering methods represented by the same alphabet belonged to the same group.

Genre	Method	Precision	Harmony	Fun	Preference
Rock	DUAL-LRA	A, B (B)	A, B (B)	A	A
	DUAL-DMA	A	A	A	A
	SINGLE-LRA	B (C)	B	A	A
	SINGLE-DMA	B (C)	B	A	A
Dance	DUAL-LRA	A	A	A	A
	DUAL-DMA	A	A	A	A
	SINGLE-LRA	A, B (A)	A	A	A
	SINGLE-DMA	B	B	B	B
Classical	DUAL-LRA	A	A	A	A
	DUAL-DMA	A	B	B	B
	SINGLE-LRA	B	B	C	B
	SINGLE-DMA	B	B	C	B
Vocal	DUAL-LRA	A	A	A	A
	DUAL-DMA	A	A	A	A
	SINGLE-LRA	A	A	A (B)	A
	SINGLE-DMA	B	A	A (B)	A

DMA also received higher scores than the current standard method, SINGLE-LRA in all the measures for all the genres, with statistical significance in 4 (out of 16) cases. The benefit of DUAL-DMA was the most evident for rock music.

6.2.3 Discussion

The user study elucidated the benefit of our dual-band haptic music player for improving the music listening experience. The dual-band rendering with DMA acquired high subjective scores, especially for rock music among the four music genres. The evaluation results also indicated adequate uses of single-mode vs. dual-mode rendering for each music genre.

The comments of the participants collected after the experiment were quite diverse. Common ones are reported below along with the major experimental results. First of all, 48% participants (11 of 23) said that a vibration rendering method should be tailored to a music genre. This included a choice of single- or dual-mode and individually customizable

weights of the haptic equalizer.

Most participants evaluated vibration playback of the dual-band mode higher than that of the single-band mode. The bass components of music mainly contain regular beat sounds played by drums and bass instruments. The participants reported frequently that regularly-repeated vibration beats in the single band conditions were somewhat flat and boring. This was some exacerbated with the classical music that had low bass-band energy, where the single-band rendering conditions were not able to provide clear vibrotactile sensations for beats. In contrast, the dual-mode rendering attempts to add the playback of the main aspects of music, such as theme and melody. This behavior seems to have resulted in the large score differences between DUAL and SINGLE for classical music.

Between the two DUAL conditions, the effects of the two actuators greatly depended on music genre. DMA outperformed LRA for rock music, while LRA was better for classical music. For rock music, many participants reported that they liked short and highly contrasted vibrations instead of continuous vibrations. DMA could express the intensive drum beats of rock music adequately with superimposed vibrations giving rough sensations with high contrast to high-frequency vibrations. On the other hand, many participants recommended fine and delicate vibrations for classical music. The rough feeling of superimposed vibration does not seem to match with classical music, while the smooth sensations of single sinusoidal vibration by LRA appear to be more preferred.

In addition, most participants expressed strong preference on vibration strength. During the user study, the experimenter set the vibration volume, and the participants were not allowed to change it. Some participants complained of fatigue because of long strong vibration, whereas others complained of vibration strength being too weak.

Lastly, some notable descriptions of the participants for the four rendering conditions are provided: 1) DUAL-LRA: “The vibration is smooth and matches well with the music, but it is somewhat flat and boring.” 2) SINGLE-LRA: “The vibration is good at beat expression, but it is too sparse, regular, and weak.” 3) DUAL-DMA: “The expression of bass beats is good. It is also more fun and feels like a well-tailored vibration to music.” 4) SINGLE-DMA: “The expression of strong bass beats is good. However, it focuses on the bass sound

too much and is a little boring." These subjective reports suggested that bass beat expression using the superimposed vibration of DMA was distinct from the simple sinusoidal vibration of LRA, making good impressions on many participants.

6.3 Limitations

Our current vibration generation algorithm is designed for real-time processing on a mobile platform with relatively low processing power. As such, it has several important limitations. In this section, we discuss these limitations, as well as potential research directions to resolve the issues.

First of all, we noted during the user study that the users' expectations of good vibration playback are quite diverse. For example, some participants preferred vibration playback that faithfully followed the main melody of a song, whereas some others wanted vibration playback to track a particular instrument. Our current vibration extraction algorithm, which relies on sound energy integration, is not capable of such explicit feature tracking; selective feature tracking requires much more sophisticated algorithms. In computer music research, there has been active research on automatic transcription of polyphonic music. In particular, some algorithms can automatically extract musical scores from music sources using the sampled sounds of an instrument, with about 70% accuracy [41, 65]. Such musical scores can be transformed into scores for vibration playback based on signal-level conversion between sound and vibration, as demonstrated in a score-based vibration authoring tool we developed earlier [45, 43]. However, such approaches are likely to require a significant amount of preprocessing and/or off-line authoring. Devising a real-time algorithm with a reasonable trade-off between tracking accuracy and processing speed will be an intriguing research topic.

Matching perceptual variables between sound and vibration while considering actuator limitations and aesthetic quality also remains largely unexplored. Our approach uses perceived magnitude as a medium, but other time-related factors, e.g., signal duration, may also be crucial. Our evaluation results also suggested that the best relationships may depend on music genre. These issues also need significantly more attention, especially with

wideband actuators becoming popular in mobile devices.

Another important practical problem is power consumption. Music listening with vibration playback is inevitably a prolonged task, thus vibration patterns that can save power, e.g., short-duration patterns, are more desirable if their perceptual value is comparable to those of other patterns. The incorporation of these requirements into automatic vibration generation demands further research before this new function can be actively employed in mobile devices.

The current haptic music player is our initial approach to vibrotactile music rendering. The results of this study are not specific to DMA; they can be adapted to other mobile actuators with wide frequency bands. Alternatively, we can use multiple LRAs that have different resonance frequencies to implement superimposed vibration. Apart from these, the vibration superposition approach is expected to have a long-lasting merit because its perception is better understood [63, 85] and considerably simpler than the largely unexplored perception of wideband vibrotactile stimuli. Our haptic music player can also be extended to other applications that contain audio signals, such as movies, games, and music videos.

Chapter 7

Improvement of Haptic Music Player

This chapter describes the improved version of haptic music player with auditory saliency detection algorithm and the use of a wideband actuator.

The intensity extraction algorithm in Chapter 6 is very simple with moderately high subjective performance. However, the algorithm often generates continuing vibration that makes the user bored and tired without recognizing the user's attention in a music clip. Thus, an algorithm that can detect and emphasize users' interest in music is needed to increase users' preference on vibrotactile music. We developed a novel algorithm for haptic music player with several auditory features for estimating saliency in music perception. In this algorithm, the intensity level of vibration is determined by the estimated saliency of audio signal. The users can feel more contrasted vibrations by the saliency in music with a less fatigue.

In addition, we replaced the vibration actuator from DMA to a voice-coil type wideband actuator. The new actuator has much faster response ($\sim$1 ms) and stronger intensity, with a wideband frequency response (90–1,000 Hz). With the use of new actuator, we can freely select the component frequencies for vibrotactile rendering and the number of rendering channels. Hence, vibration rendering in haptic music player may become more transparent and variety in expression.

In this study, a user evaluation was conducted to compare the perceptual performance of our new saliency-based algorithm and the initial algorithm. A new actuator was used rather than LRA or DMA, and the 16 musical pieces were used again to represent four music genres (rock, dance, classical, and vocal; four pieces each). Perceptual benefits of the saliency-based two-channel rendering mode was found in evaluation with dependence on music genre.

7.1 Software

Changes of the software structure and algorithms of the haptic music player from the initial version are described in this section.

7.1.1 Structure

Fig. 7.1 demonstrates the overall structure and computational processes of the saliency-based haptic music player. The current saliency-based haptic music player shares much of processes with the initial version. First, both algorithm are identical from the open of a music file, to the sub-band separation of FFT results for each audio signal segment. In the saliency-based haptic music player, four auditory features were calculated for saliency estimation in each sub-band using the spectral intensities calculated by FFT. Then, the auditory saliency can be computed through the haptic equalizer for bass and treble bands (two-channel) or six sub-bands (six-channel) by the selected number of rendering channels. The computed auditory saliencies were transformed into voltage commands for the wide-band actuator via psychophysical scaling procedures. The overall conversion process is: sound signal energy $\rightarrow$ auditory perceived magnitude $\rightarrow$ vibratory perceived magnitude $\rightarrow$ physical vibration amplitude $\rightarrow$ voltage command amplitude to the actuator. We improved the psychophysical scaling procedures of the initial version, with the knowledge derived from psychophysical studies on simple and complex vibrations. Despite the introduction of additional features and following computations, the algorithm was designed to maintain appropriate real-time performance.

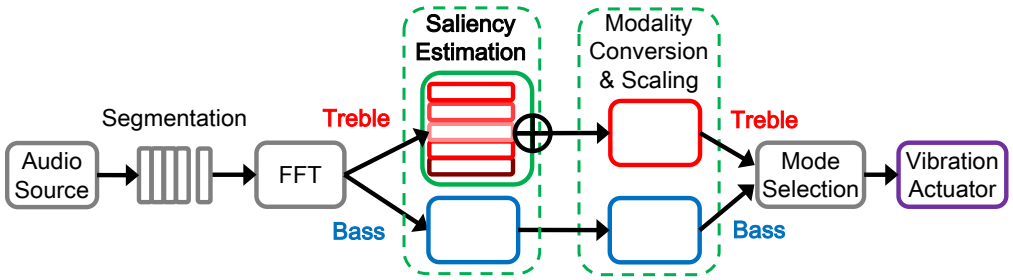

Fig. 7.1 Process loop of the saliency-based haptic music player.

7.1.2 Saliency Estimation

Four features were selected to estimate the auditory saliency of each sub-band: loudness of the peak (L), energy of the peak (E), pitch of the peak (P), and sum of amplitude (A). A frequency component with the largest amplitude was selected as the peak in a sub-band. Two physical features, E and A were introduced into our saliency model from Evangelopoulos' study [12]. The Evangelopoulos' saliency model was designed for detection of salient event in speech, which is quite different to our goal, estimation of a weight variable for the vibration intensity in vibrotactile music rendering. Since we needed continuous and perceptually linear model for music, we added two perceptual features, L and P, on the saliency model.

Loudness of an audio signal can be estimated in dB scale by A-weighting. A 16-bit audio file contains intensity information of music source in a range of 90 dB (0–32267). Considering the sound output level in the real-use environment, we adjusted -30 dB offset on the loudness computed from the signal intensity. After that, the loudness standardized into 0–1.0 scale for the next computation step.

Energy of the peak was estimated by Teager-Kaiser energy, as $E = a^2 sin^2 \omega$, where a is spectral amplitude and ω is angular frequency of the component in a sub-band [35]. Since the temporal variation of the energy so drastical that it can occur undesired discontinuity, we used logarithm of E. The $\log E$ was also linearly scaled into 0–1.0.

For the estimation of pitch which is the perceptual degree of frequency, mel scale was adjusted in the current implementation [78]. Relative pitch of the peak at nth loop, $P(n)$ was calculated by following equation:

$$P(n) = \frac{mel(f_{peak}(n))}{mel(f_{max}(n))} = \frac{2595\log 10(1 + f_{peak}(n)/700)}{2595\log 10(1 + 6400/700)}. \quad (7.1)$$

As the last of the four factors, sum of amplitude was computed by summing up all component amplitudes in a sub-band identically to the initial version. From the four factors, auditory saliency for ith sub-band (0: lowest band, and 5: highest band) at nth loop, $S_i(n)$ were estimated as follows,

$$S_i(n) = \alpha \cdot A_i(n) \cdot L_i(n) \cdot \log E_i(n) \cdot P_i(n)^{\beta}. \quad (7.2)$$

The β determines the weight of attenuation by pitch of the peak. In a music listening experience in real world, vibrations are more distinct from the bass components than the treble components. Thus, we tested the value of β in a range of -0.4–0.0 to attenuate the saliency with the increase of the pitch.

The equation for saliency estimation has a multiplication form of the four factors. A high saliency level can be appeared when all factors except P have large values. We can also expect that the multiplications will induce abrupt temporal changes of the computed saliency. Other types of the saliency estimation function in [88], such as weighted sum of the factors can be applied to achieve the smoother temporal variation, but the current function showed the best perceptual performance in our pilot test.

7.1.3 Modality Conversion and Intensity Scaling

The concept of modality conversion in the previous implementation is still remained in the current version of haptic music player. In the current version, the conversion procedures were differred by number of rendering channels: two or six. Since we assumed that the estimated audio saliencies are already in psychophysical scale, the additional conversion to auditory perceived magnitude is omitted from the procedure.

Conversion to Vibratory Perceived Magnitudes

The next step is conversion to vibratory perceived magnitudes. From the saliency of the six sub-bands, S_i, we computed the desired perceived magnitudes of vibration for two-channel

rendering, $I_{bass}(n)$ and $I_{treble}(n)$, as follows:

$$I_{bass}(n) = w_{bass} c S_1(n), \tag{7.3}$$

$$I_{treble}(n) = w_{treble} c \sum_{i=2}^{6} S_i(n). \tag{7.4}$$

For the six-channel rendering, I_i was computed instead of I_{treble}, by $I_i(n) = w_{treble} c S_i(n)$, where c is the cross-modal scaling constant, and w_{bass} and w_{treble} are the amplification gains of the bass and treble bands in charge of haptic volume control, same as those in the previous implementation.

Rendering Mode Selection

When the algorithm is running in two-channel rendering mode, the operation in this step is analogous to that of our previous version. However, Pythagorean summation model in Chapter 5 is adjusted in calculating the perceived intensity of superimposed vibration for bass expression. The superimposed vibration has two frequency components of f_1 and f_2 and their corresponding perceived intensities, P_1 and P_2 are as follows:

If $I_{bass}(n) > I_{treble}(n)$, then

$$P_1(n) = P_2(n) = \frac{I_{bass}(n)}{\sqrt{2}}. \tag{7.5}$$

If $I_{bass}(n) \leq I_{treble}(n)$, then

$$P_1(n) = 0, \text{ and } P_2(n) = I_{treble}(n). \tag{7.6}$$

On the other hand, the six-channel rendering mode uses six simple sinusoidal vibrations from f_1 to f_6 for six sub-bands, simultaneously. Thus, the perceived intensity for ith channel, P_i is equal to the I_i.

Conversion to Physical Amplitudes

The desired perceived magnitudes of the vibrations, $P_i(n)$, can be readily converted to the desired vibration amplitudes at frequencies f_i using the inverse of the perceived magnitude

function of vibratory stimuli [30]. Then the desired vibration amplitudes were converted to the voltage command amplitudes for f_i. We derived the piecewise linear I/O mappings of the vibration actuator for each frequency, as in our previous haptic music player. The input voltage amplitudes $V_i(n)$ are determined using these I/O relations.

7.1.4 Implementation

We implemented the algorithms described above on an MS Windows platform, using MS Visual C++ 2008 with external libraries for music file I/O (Audiere 1.9.4) and FFT calculation (FFTw 3.2.2). The haptic music player ran on a desktop PC (3 GHz Intel Core 2 Duo) because of the difficulty of custom signal I/O for DMA on commercial mobile platforms. The length of a music segment processed in each loop was ramained as 50 ms as the previous implementation.

7.2 User Study

Subjectived performance of two-channel and six-channel rendering modes with saliency-based algorithm were compared to the our previous dual-mode algorithm. The most of experimental conditions are consistent to the user study in Section 6.2.

7.2.1 Methods

Participants

Thirty university students (15 males and 15 females) participated in this experiment. They were 19–25 years old with a mean 21.1 (SD 1.7). All participants were daily users of a mobile phone with no known sensorimotor impairment. They were paid KRW 15,000 ($\simeq$ USD 13) for the experiment.

Apparatus

We used Haptuator Mark II (Tactile Labs Inc.) as a vibration actuator. The rated bandwidth of the actuator is 90–1,000 Hz. Since the Haptuator has much faster time response ($\sim$ 1 ms)

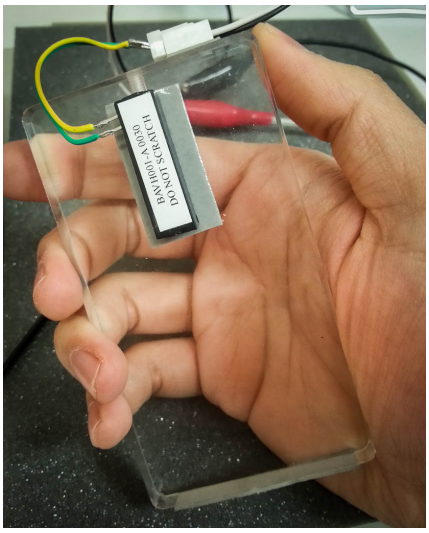

Fig. 7.2 Handheld mockup with a vibration actuator used in the user study.

than the conventional LRA or DMA (50–100 ms), it allows more precise expressions than expressions in the previous haptic music player system. The actuator was attached vertically to the wide frontal face of handheld mockup made from acrylic resin ($110 \times 60 \times 10$ mm), as shown in Fig. 7.2. The vibration generated from actuator was transferred to the participants' dominant hand via the mockup in height direction. The input-output relations of the actuator were calibrated using a miniature accelerometer (Kister; model 7894A500; 7.5 g) attached to the center of the mockup. A PC controlled the actuator via a data acquisition board (National Instruments; model USB-6251) with a custom-made power amplifier. The sampling rate for signal I/O was 10 kHz for faithful signal sampling and reconstruction. In the experiment, overear headphones were used for the auditory music playback.

Experimental Conditions

This study consisted of 12 experimental conditions (3 rendering modes $\times$ 4 music genres) in a within-subjects design. We used three methods for vibration rendering: previous two-channel, saliency-based two-channel, and saliency-based six-channel modes. In the two-channel rendering modes, superposition of 150 Hz and 200 Hz sinuoidal vibrations were

utilized to express the intensity of bass components. We decreased the frequency ratio of f_1 and f_2 than the superposition in the previous implementation (150 Hz and 223 Hz). It may result more distinct feeling of bass expression from the high frequency sinusoidal vibration, as shown in Chapter 5. A lower frequency ratio is available, but the superposed vibration would be perceived as weird due to the decreased consonance [85]. For the six-channel rendering mode, six frequencies were used 80 Hz to 244 Hz with a multiplier of 1.25 in each frequency interval. The interval was determined considering the difference threshold of vibration frequency (about 20%).

The selected music genres and four music pieces per genre were identical to those of our previous study ([26], See Table 6.2). Each music piece was trimmed 45 secs and concatenated them for each music genre to consist a 3-min music clip. The preset weights in Table 6.1 were adjusted when the play of music. The perceptual intensities of vibration were controlled similarly among the three rendering conditions.

Subjective Performance Measures

We collected four subjective measures using a questionnaire in a 0–100 continuous scale. They were: Precision–"Did the vibration express the music precisely?" (0: very imprecise, 100: very precise); Harmony–"Was the vibration harmonious with the music?" (0: very inharmonious, 100: very harmonious); Fun–"Was the vibration fun?" (0: very boring, 100: very fun); and Preference–"Did you like the vibration?" (0: dislike very much, 100: like very much). The participants also reported the subjective preference of vibrotactile feedback for each music genre.

Procedure

Prior to the experiment, each participant was given instructions about the experimental procedures and explanations of the meanings of the questions in the questionnaire. A training session was then followed, where two songs that were not used in the main sessions were played using each of the 3 vibration rendering modes for 2 min.

The main experiment consisted of four sessions. Each session used one of the 3-min

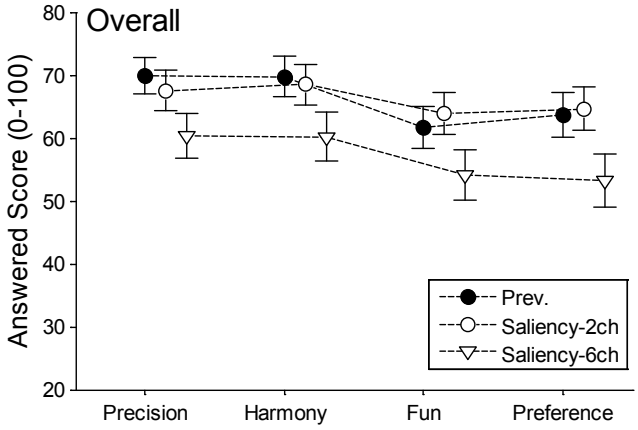

Fig. 7.3 Average evaluation results of the four rendering conditions. Error bars represent standard errors.

genre-representative music clips. First, the participant listened the music clip without vibrotactile playback to be familiar with the music and establish a perceptual reference. The participant could adjust audio volume to a comfortable level. Then, the music clip was played with vibration using one of the four rendering modes. After the playback, the participant answered the questions on the rendering modes using the questionnaire sheets. For each performance metric, the participant gave a score by marking a position on a line labeled on both ends with their meanings. The participant had a rest for a few minutes to prevent tactile adaptation before proceeding to the next rendering condition. This procedure was repeated three times with different rendering conditions.

To remove any possible order effects, we randomized the orders of the rendering conditions in session and those of the music genres. The experiment took about 1.5 hours.

7.2.2 Results

The averaged evaluation results over the four music genres are represented in Fig. 7.3. The two-channel saliency-based rendering mode (Saliency-2ch) shows similar rating with the previous dual-mode rendering mode (Prev.) in all measures. The six-channel saliency-based rendering (Saliency-6ch) follows them with a quite large gap.

The averaged results classified by four music genres are represented in Fig. 7.4. An ANOVA test was conducted for each music genre with SNK test for post-hoc multiple comparison.

In rock music, the three rendering modes were evaluated similarly in terms of precision. However, in the other three measures, Saliency-6ch shows significantly lower scores than the others, and Saliency-2ch precedes Prev. mode. The SNK test revealed marginally significant difference between Saliency-2ch and Prev. in fun ($p < 0.1$). Participants' preference for the vibrotactile playback for the rock music was 73.8, which is a quite high score.

Perceptual merits of the Saliency-2ch were also shown in the results for dance music. The Prev. mode scored lower than the two Saliency modes in all measures. While, the statistically significant difference between Saliency-2ch and Prev. was not shown except the marginal significance in preference ($p < 0.1$). Participants preferred vibrotactile playback with the dance music (74.6).

On the contrary, classical music showed merits of Prev. mode compared to the two Saliency modes. For all measures, statistical significances were found in comparisions of Prev. and Saliency-2ch, and Saliency-2ch and Saliency-6ch ($p < 0.05$). However, participants' preference for the vibrotactile playback was the lowest in classical music among the four music genres (48.7).

In vocal music, scores for the three modes are similar in the four measures. Only a marginal significance was shown in precision between Prev. and Saliency-6ch ($p < 0.1$). Participants' preference for the vocal music with vibrotactile playback was 58.9.

7.2.3 Discussion

The evaluation results showed the superior subjective performance of Saliency-2ch mode than the other two rendering modes in rock and dance music. These two music genres were the most preffered genres by the participants to be well matched with vibrotactile playback function. Hence, the Saliency-2ch mode can enhance users' music experience more than the previous dual-mode rendering, despite the low scores in classical music which is the least preferred genre for vibrotactile music.

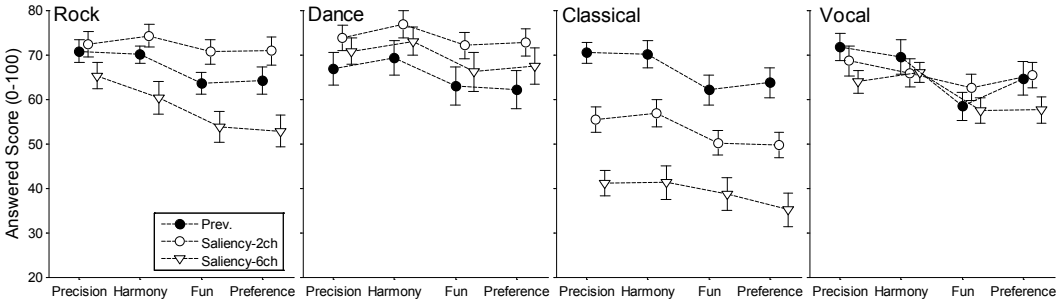

Fig. 7.4 Evaluation results of the four rendering conditions by music genre.

Participants' common comments on the feelings of three rendering modes are as follows: 1) Prev.: "Good at expressing detail of music. Plain and a little boring than the other conditions. Continuing vibration makes the hand tired." 2) Saliency-2ch: "Emphasis of temporal contrast is good and makes me fun. The feeling of vibration is clear." 3) Saliency-6ch: "Temporal intensity change was not expressed properly. The vibration gives rough feeling."

The saliency-based haptic music player aimed to enhance the feeling of music by presenting vibrations generated considering the user's saliency in music. From the participants' reports, we could confirm that the design concept of the saliency-based algorithm was demonstrated well in the generated vibration. With the reduced operation of the vibration actuator, the users can feel more distinct expression of music with decreased fatigue on their hands.

Meanwhile, Saliency-6ch mode showed worse scores than the Prev. mode in the evaluation. The rough feeling in Saliency-6ch mode might be resulted from the superimposition of six vibrations having different frequencies. In our previous study, we showed that vibrotactile consonance increases with the ratio of two component frequencies, analogously to the trend in auditory consonance [85]. Since the frequency bandwidth in sound (25–6,400 Hz) was converted into the six vibration frequencies in a very compressed bandwidth (80–244 Hz), harmonious chords in music is expressed into multiple sinusoidal vibrations with a narrow gap in their frequencies. In Saliency-2ch mode, the rough feeling was generated only when the bass component is stronger than the treble component, usually to ex-

press drum beats or sounds from a bass guitar. However, the superposition in Saliency-6ch mode was occurred when two or more sub-bands have strong saliency, without considering roughness of the sound. Thus the participants could feel mismatch of timbres in sound and vibration. Consequently, perceptual effects of the superimposed vibrations should be considered for simultaneous multi-channel rendering of vibrotactile music.

For the development of the haptic music player, we tried to express the auditory feeling of music into vibration stimuli. Our vibrotactile music is aiming multimodal enhancement of music listening experience by additional vibrotactile stimuli, which is not a total substitution of auditory music experience. Though the current version of haptic music player only concerns about the faithful and synchronized expression of auditory music into vibration stimuli, consideration of auditory-tactile synesthesia may improve the multimodal exprience largely. The contents of vibration stimuli are not need to be similar to the auditory contents and can contain additional information or different contents which can be matched with music. Further studies will be needed to achieve this synesthesia based vibrotactile music rendering.

Chapter 8

Conclusion

The goal of this study is revealing psychophysical characteristics of vibration in mobile device. The perceptual characteristics were utilized on developing vibrotactile rendering methods which can enhance the music experience in mobile device. In the measurements of the perceived intensity of vibration on a mobile device, amplitude, frequency and direction were effective among the four tested factors. Perceived magnitude estimation model for each vibration direction was built from the measured data and Stevens' power law. The power relationship found between stimulus power and the perceived intensity can be practically used for estimation of perceived intensity.

Based on these essential data, qualitative characteristics of vibrations were investigated. Through the three experiments we could configure the perceptual space of simple vibrations with their adjective ratings. The two dimensional perceptual space orthogonally spanned a low frequency range (40–100 Hz) and a high frequency range (100–250 Hz). From the adjective ratings for bi-frequency vibrations, an evidence was found about their similar percepts with low frequency simple vibrations which were close to negative adjectives.

Followed experiments revealed the perceptual space of bi-frequency vibrations. Perceptual differences between the bi-frequency vibrations and simple vibrations were analyzed with the effects of spectral factors. In addition, Pythagorean summation model was suggested to estimate perceived intensities of bi-frequency vibrations.

The characteristics of bi-frequency vibration was used on dual-channel rendering of our haptic music player. Initial version of our haptic music player was developed with several distinguishing features such as, dual-channel playback, haptic equalizer, perception based modality conversion and scaling and real-time rendering. We rendered bi-frequency vibration to express bass signal in music and high frequency vibration for treble signal. An improved version was developed with auditory saliency estimation and use of a wide-band actuator. User study results of the haptic music player showed perceptual merits of bi-frequency vibrations and feasibility of this application on mobile device.

This study is targeting practical use of the results in both of industries and academics. Throughout the study the author found many interesting research issues from the fundamental haptic perception to the application of multimodal interaction. Since all derived results are about the vibration perception in mobile environment, we can check the consistences of perceptual characteristics on other body sites in different platforms. By investigating these research issues, further researches on vibrotactile perception and rendering can be fertilized.

Appendix

Closed-loop Control of Shaker

The open-loop control of a shaker cannot guarantee accurate stimulus delivery since its mechanical load, the hand-arm dynamics of a participant, varies by individual and over time. To improve its control accuracy, we adopted a simple proportional closed-loop control (P-control). In our system, the shaker output was sampled at 10 KHz using the accelerometer. In every 5 ms, we estimated the output amplitude by transforming the acceleration data measured in the previous 50 ms (500 samples) via FFT and taking its amplitude at the input frequency. Then, the input voltage to the shaker was determined by

$$V(n) = V(n-1) + p(A^* - A(n)), \tag{8.1}$$

where n is the time index, p is the proportional gain, and A^* and $A(n)$ are the desired and measured amplitudes, respectively.

The proportional gain was chosen for each mock-up and each vibration direction. During this gain tuning, the experimenter grasped a mock-up with a regular grip force (0.5–2.0 N; an average grip force measured in [72]). The gain was linearly increased from 0 until the steady-state error between A^* and $A(n)$ was reduced to be less than 1.5% of A^*, which is significantly smaller than the difference threshold of vibration magnitude (about 8% [17]). Figure 8.1 shows sample data for a 40 Hz signal. Here, the open-loop shaker gains were calibrated under the *unloaded* condition, and they resulted in a large error in the state-

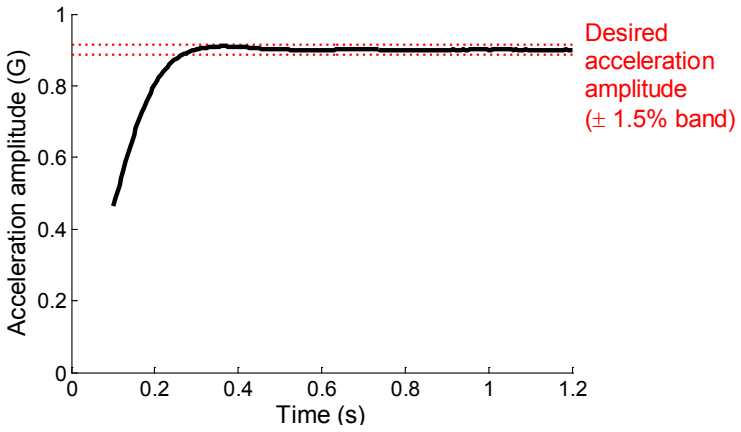

Fig. 8.1 Effect of P-control (40 Hz).

state amplitude. These errors can be mitigated by using the shaker gains calibrated for *each* participant under the *loaded* condition, as was in [72], but this is a cumbersome process and can still suffer from time-varying error sources such as grip force. In contrast, the closed-loop data showed adequate accuracy, converging to the steady state within 0.5 s. This short transient period is acceptable since our stimuli were 3-s long and temporal summation of the Pacinican channel saturates in around 1 s [16].

요 약 문

진동 자극의 인지적 분석과 이를 응용한 음악의 진동촉감 표현기법

다중감각을 이용한 디스플레이는 사용자 경험과 작업 성능을 향상시키는 효과가 있음이 알려져 있다. 특히 촉감 디스플레이는 최근 들어 휴대 단말용 사용자 인터페이스 (UI), 각종 오락기기의 특수효과, 자동차에서의 정보 전달 시스템 등 다양한 분야로 그 활용을 넓혀가고 있으나 지금까지는 단순한 진동신호만이 그 인지적 효과에 대한 체계적인 이해 없이 활용되어 왔다. 본 연구에서는 모바일 기기에서 사용자 중심의 인지적으로 최적화된 진동 신호의 생성을 위해 손으로 전달되는 진동의 인지적 특성과 이를 기반으로 사용자의 음악 감상 경험을 향상시키는 촉감 음악 재생기의 개발을 제안한다. 인지적으로 효과적인 촉감 음악 재생기의 개발을 위해 우선적으로 단순 정현파 진동의 손에서의 인지 특성에 대한 연구가 진행되었다. 다양한 단진동에 대해서 주요한 인지 특성인 인지강도와 인지적 상이성을 정신물리학적 실험을 통해 측정하였다. 정량적 특성인 인지 강도에 대해서는 진동의 진폭, 주파수, 방향, 무게 등 네 가지 요소의 영향을 분석한 결과 진폭, 주파수, 방향에 대해서 Stevens의 지수 법칙에 기반한 진동 인지강도 추정 모델을 도출하였으며 이 과정에서 진동의 인지강도와 진동력 사이에는 지수 관계가 존재함을 밝혀내었다.

단진동의 정성적인 인지 특성을 관찰하기 위해서는 인지적 상이성의 측정과 형용사 평가를 수행하였다. 세 종류의 실험을 거친 결과 다양한 진동 주파수와 진폭에 대해 단진동의 상대적 위치 관계가 2차원 인지공간 상에서 13쌍의 형용사 변화축과 함께 추정되었다. 2차원 인지공간에서 낮은 주파수 (40–100 Hz)와 높은 주파수 (100–

250 Hz) 진동은 서로 수직에 가까운 변화축을 갖는 것이 밝혀졌으며 낮은 주파수의 진동은 느리다, 희미하다, 뭉툭하다 등 부정적인 의미의 형용사에 가깝게 평가되었다.

이러한 단진동의 인지적 특성에 기반하여 두 개의 주파수 성분이 중첩된 진동에 대해서도 인지강도와 인지 공간을 측정하는 실험이 이루어졌다. 다양한 중첩 조건의 인지강도 실험을 통해서 피타고리안 합산 모델이 이중 주파수 중첩 진동의 인지강도를 잘 설명할 수 있음을 보였으며, 이중주파수 진동은 특히 중첩진동의 주파수 성분이 같은 크기에 가깝게 합산될수록, 인지공간 상에서 단진동과 구분되는 인지적 상이성을 가지는 것이 관찰되었다. 다양한 주파수 조합에 대해서 이중주파수 진동을 구성하는 세가지 요소인 성분주파수, 성분주파수 차이, 성분주파수 비가 인지공간에서 미치는 영향 또한 분석되었다.

이와 같은 모바일 기기에서의 진동에 대한 인지적 연구 결과의 응용으로 촉감 음악 재생기의 개발이 이루어졌다. 초기 버전은 중첩진동을 이용한 이중 채널 재생, 햅틱 이퀄라이저, 인지기반 감각 변환, 실시간 재생 등의 특징을 가지고 개발되었으며 사용자 평가에서 기존 방식 대비 우수한 평가 결과를 보였다. 이후 촉감 음악 재생기는 넓은 주파수 대역을 갖는 진동자와 음악의 주목도 추출 알고리즘을 포함하는 개선이 이루어졌으며 사용자 평가 결과 초기 버전에 비해서 인지적 효과의 향상이 이루어졌다.

본 연구는 산업체 및 학계에서의 실용적 활용을 목표로 이루어졌으며, 연구의 결과는 휴대 단말기기에서의 진동촉감의 생성 및 인지 관련 연구의 활성화에 공헌할 수 있을 것으로 기대한다.

REFERENCES

[1] S. J. Bensmaïa and M. Hollins. Complex tactile waveform discrimination. *Acoustical Society of America Journal*, 108:1236–1245, 2000.

[2] S. J. Bensmaïa, M. Hollins, and J. Yau. Vibrotactile intensity and frequency information in the pacinian system: A psychophysical model. *Perception & psychophysics*, 67(5):828–841, 2005.

[3] L. E. Bernstein, S. P. Eberhardt, and M. E. Demorest. Single-channel vibrotactile supplements to visual perception of intonation and stress. *The Journal of the Acoustical Society of America*, 85(1):397–405, 1989.

[4] S. J. Bolanowski Jr, G. A. Gescheider, R. T. Verrillo, and C. M. Checkosky. Four channels mediate the mechanical aspects of touch. *The Journal of the Acoustical society of America*, 84(5):1680–1694, 1988.

[5] S. A. Brewster and L. M. Brown. Tactons: Structured tactile messages for non-visual information display. In *Proceedings of the Australasian User Interface Conference*, pages 15–23, 2004.

[6] A. J. Brisben, S. S. Hsiao, and K. O. Johnson. Detection of vibration transmitted through an object grasped in the hand detection of vibration transmitted through an object grasped in the hand. *Journal of Neurophysiology*, 81(4):1548–1558, 1999.

[7] P. L. Brooks and B. J. Frost. The development and evaluation of a tactile vocoder for the profoundly deaf. *Canadian Journal of Public Health.*, 77:108–113, 1986.

[8] L. M. Brown and T. Kaaresoja. Feel who's talking: Using tactons for mobile phone alerts. In *Proceeding of the CHI: Extended Abstracts*, pages 604–609, 2006.

[9] A. Chang and C. O'Sullivan. Audio-haptic feedback in mobile phones. In *Proceeding of the CHI*, pages 1264–1267, New York, NY, USA, 2005. ACM.

[10] R. G. Dong, D. E. Welcome, T. W. McDowell, J. Z. Wu, and A. W. Schopper. Frequency weighting derived from power absorption of fingers–hand–arm system under z_h-axis vibration. *Journal of Biomechanics*, 39(12):2311–2324, 2006.

[11] M. Enriquez. A study of haptic icons. Master's thesis, Dept. of Computer Science, University of British Columbia, 2002.

[12] G. Evangelopoulos, K. Rapantzikos, P. Maragos, Y. Avrithis, and A. Potamianos. Audiovisual attention modeling and salient event detection. *Multimodal Processing and Interaction*, 33(2):1–21, 2008.

[13] A. Gallace, H. Z. Tan, and C. Spence. The body surface as a communication system: The state of the art after 50 years. *Presence: Teleoperators and Virtual Environments*, 16(6):655–676, 2007.

[14] R. H. Gault. Progress in experiments on tactual interpretation of oral speech. *Journal of Abnormal Psychology and Social Psychology*, 19(2):155–159, 1924.

[15] G. A. Gescheider. *Psychophysics: The Fundamentals*. Lawrence Erlbaum Associate, Mahwah, NJ, USA, 3rd edition, 1997.

[16] G. A. Gescheider, M. E. Berryhill, R. T. Verrillo, and S. J. Bolanowski. Vibrotactile temporal summation: Probability summation or neural integration? *Somatosensory & Motor Research*, 16(3):229–242, 1999.

[17] G. A. Gescheider, S. J. Bolanowski Jr, R. T. Verrillo, D. J. Arpajian, and T. F. Ryan. Vibrotactile intensity discrimination measured by three methods. *The Journal of the Acoustical Society of America*, 87:330–338, 1990.

[18] G. A. Gescheider and B. A. Hughson. Stimulus context and absolute magnitude estimation: A study of individual differences. *Attention, Perception, & Psychophysics*, 50(1):45–57, 1991.

[19] E. B. Goldstein. *Sensation and Perception*. Wadsworth-Thomson Learning, Pacific Grove, CA, USA, 6th edition, 2002.

[20] E. Gunther. Skinscape: A tool for composition in the tactile modality. Master's thesis, Massachusetts Institute of Technology, 2001.

[21] M. Hall, E. Hoggan, and S. Brewster A. T-bars: Towards tactile user interfaces for mobile touchscreens. In *Proceedings of the MobileHCI*, pages 411–414. ACM, 2008.

[22] S. H. Han, M. Song, and J. Kwahk. A systematic method for analyzing magnitude estimation data. *International Journal of Industrial Ergonomics*, 23:513–524, 1999.

[23] E. Hoggan and S. A. Brewster. Crossmodal icons for information display. In *Proceeding of the CHI*, pages 857–862. ACM, 2006.

[24] M. Hollins, S. J. Bensmaia, K. Karlof, and F. Young. Individual differences in perceptual space for tactile textures: Evidence from multidimensional scaling. *Perception and Psychophysics*, 62(8):1534–1544, 2000.

[25] M. Hollins, R. Faldowski, R. Rao, and F. Young. Perceptual dimensions of tactile surfaced texture: A multidimensional scaling analysis. *Perception & Psychophysics*, 54:697–705, 1993.

[26] H. Hwang, I. Lee and S. Choi. Real-time dual-band haptic music player for mobile devices. *To appear in the IEEE Transactions on Haptics*, 2013.

[27] I. Hwang and S. Choi. Perceptual space and adjective rating of sinusoidal vibrations perceived via mobile device. In *Proceedings of the Haptics Symposium*, pages 1–8. IEEE, 2010.

[28] I. Hwang and S. Choi. Effect of mechanical ground on the vibrotactile perceived intensity of a handheld object. *Lecture Notes in Computer Science (Eurohaptics 2012)*, 7283:61–66, 2012.

[29] I. Hwang, J. Seo, M. Kim, and S. Choi. Perceived intensity of tool-transmitted vibration: Effects of amplitude and frequency. In *Proceedings of the IEEE International Symposium on Haptic Visual-Audio Environments and Games (HAVE)*, pages 1–6, 2012.

[30] I. Hwang, J. Seo, M. Kim, and S. Choi. Vibrotactile perceived intensity for mobile devices as a function of direction, amplitude, and frequency. *To appear in the IEEE Transactions on Haptics*, 2013.

[31] A. Israr, S. Choi, and H. Z. Tan. Detection threshold and mechanical impedance of the hand in a pen-hold posture. In *Proceedings of the IROS*, pages 472–477, 2006.

[32] L. Itti and C. Koch. Computational modeling of visual attention. *Nature reviews neuroscience*, 2(3):194–203, 2001.

[33] D. N. Jiang, L. Lu, H. J. Zhang, J. H. Tao, and L. H. Cai. Music type classification by spectral contrast feature. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, volume 1, pages 113–116, 2002.

[34] L. A. Jones and S. J. Lederman. *Human hand function*. Oxford University Press, USA, 2006.

[35] J. F. Kaiser. On a simple algorithm to calculate theenergy'of a signal. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 381–384. IEEE, 1990.

[36] M. Karam, F. A. Russo, and D. I. Fels. Designing the model human cochlea: An ambient crossmodal audio-tactile display. *IEEE Transactions on Haptics*, 2(3):160–169, 2009.

[37] C. Kayser, C. I. Petkov, M. Lippert, and N. K. Logothetis. Mechanisms for allocating auditory attention: An auditory saliency map. *Current Biology*, 15(21):1943–1947, 2005.

[38] S. Kihlberg. Biodynamic response of the hand-arm system to vibration from an impact hammer and a grinder. *International Journal of Industrial Ergonomics*, 16(1):1–8, 1995.

[39] S. Kim, G. Park, S. Yim, G. Han, S. Jeon, S. Choi, and S. Choi. Gesture-recognizing hand-held interface with vibrotactile feedback for 3D interaction. *IEEE Transactions on Consumer Electronics*, 55(3):1169–1177, 2009.

[40] S.-C. Kim, K.-U. Kyung, J.-H. Sohn, and D.-S. Kwon. An evaluation of human sensibility on perceived texture under variation of vibrotactile stimuli using a tactile display system. In *Proceedings of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 429–436, 2006.

[41] A. Klapuri. Automatic music transcription as we know it today. *Journal of New Music Research*, 33(3):269–282, 2004.

[42] J. B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.

[43] J. Lee and S. Choi. Evaluation of vibrotactile pattern design using vibrotactile score. In *Proceedings of the IEEE Haptics Symposium*, pages 231–238, 2012.

[44] J. Lee and S. Choi. Real-time perception-level translation from audio signals to vibrotactile effects. In *Proceeding of the CHI: Extended Abstracts*, pages 2567–2576. ACM, 2013.

[45] J. Lee, J. Ryu, and S. Choi. Vibrotactile score: A score metaphor for designing vibrotactile patterns. In *Proceedings of World Haptics Conference*, pages 302–307. IEEE Computer Society, 2009.

[46] LG Electronics. Apparatus for generating vibration and mobile device employing the same. Korean Patent Application, No. 10-2009-0003264, Jan. 15, 2009.

[47] LG Electronics. Apparatus and method for generating vibration pattern. US Patent Application, No. 12/840,988, Jul. 21, 2010.

[48] K. A. Li, T. Y. Sohn, S. Huang, and W. G. Griswold. Peopletones: a system for the detection and notification of buddy proximity on mobile phones. In *Proceeding of the MobiSys*, pages 160–173. ACM, 2008.

[49] S. C. Lim, K. U. Kyung, and D. S. Kwon. Effect of frequency difference on sensitivity of beats perception. *Experimental Brain Research*, 211:11–19, 2012.

[50] R. Lundström. Local vibrations–mechanical impedance of the human hand's glabrous skin. *Journal of Biomechanics*, 17(2):137–144, 1984.

[51] Y. F. Ma, X. S. Hua, L. Lu, and H. J. Zhang. A generic framework of user attention model and its application in video summarization. *IEEE Transactions on Multimedia*, 7(5):907–919, 2005.

[52] K. E. MacLean. Foundations of transparency in tactile information design. *IEEE Transactions on Haptics*, 1(2):84–95, 2008.

[53] K. E. MacLean. Putting haptics into the ambience. *IEEE Transactions on Haptics*, 2(3):123–135, 2009.

[54] K. E. MacLean and M. Enriquez. Perceptual design of haptic icons. In *Proceedings of Eurohaptics*, pages 351–363, 2003.

[55] J. C. Makous, R. M. Friedman, and C. J. Vierck Jr. A critical band filter in touch. *The Journal of Neuroscience*, 15(4):2808–2818, 1995.

[56] O. Mayor. An adaptative real-time beat tracking system for polyphonic pieces of audio using multiple hypotheses. In *Proceedings of MOSART Workshop on Current Research Directions in Computer Music*, 2001.

[57] T. Miwa. Evaluation methods for vibration effect. Part 3. Measurements of threshold and equal sensation contours on hand for vertical and horizontal sinusoidal vibrations. *Industrial Health*, 5:213–220, 1967.

[58] M. Morioka and M. J. Griffin. Independent responses of pacinian and non-pacinian systems with hand-transmitted vibration detected from masked thresholds. *Somatosensory & Motor Research*, 22(1-2):69–84, 2005.

[59] M. Morioka and M. J. Griffin. Thresholds for the perception of hand-transmitted vibration: Dependence on contact area and contact location. *Somatosensory & Motor Research*, 22(4):281–297, 2005.

[60] M. Morioka and M. J. Griffin. Magnitude-dependence of equivalent comfort contours for fore-and-aft, lateral and vertical hand-transmitted vibration. *Journal of Sound and Vibration*, 295(3-5):633–648, 2006.

[61] M. A. Muniak, S. Ray, S. S. Hsiao, J. F. Dammann, and S. J. Bensmaia. The neural coding of stimulus intensity: linking the population response of mechanoreceptive afferents with psychophysical behavior. *The Journal of Neuroscience*, 27(43):11687–11699, 2007.

[62] S. Okamoto, H. Nagano, and Y. Yamada. Psychophysical dimensions of tactile perception of textures. *IEEE Transactions on Haptics*, 6(1):81–93, 2013.

[63] G. Park and S. Choi. Perceptual space of amplitude-modulated vibrotactile stimuli. In *Proceedings of the World Haptics Conference (WHC)*, pages 59–64. IEEE, 2011.

[64] J. Pasquero, J. Luk, S. Little, and K. E. MacLean. Perceptual analysis of haptic icons: an investigation into the validity of cluster sorted mds. In *Proceedings of the Sympo-*

sium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pages 437–444, 2006.

[65] G. E. Poliner, D. P. W. Ellis, A. F. Ehmann, E. Gómez, S. Streich, and B. Ong. Melody Transcription From Music-Audio: Approaches and Evaluation. *IEEE Transactions on Audio Speech and Language Processing*, 15(4):1247–1256, 2007.

[66] L. J. Post, I. C. Zompa, and C. E. Chapman. Perception of vibrotactile stimuli during motor activity in human subjects. *Experimental Brain Research*, 100(1):107–120, 1994.

[67] I. Poupyrev and S. Maruyama. Tactile interfaces for small touch screens. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, pages 217–220, 2003.

[68] I. Poupyrev, S. Maruyama, and J. Rekimoto. Ambient touch: Designing tactile interfaces for handheld devices. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, pages 51–60, 2002.

[69] D. D. Reynolds, K. G. Standlee, and E. N. Angevine. Hand-arm vibration, Part III: Subjective response characteristics of individuals to hand-induced vibration. *Journal of Sound and Vibration*, 51(2):267–282, 1977.

[70] E. A. Roy and M. Hollins. A ratio code for vibrotactile pitch. *Somatosensory & Motor Research*, 15(2):134–145, 1998.

[71] J. Ryu, J. Jung, and S. Choi. Perceived magnitudes of vibrations transmitted through mobile device. In *Proceedings of the Haptics Symposium*, pages 139–140, 2008.

[72] J. Ryu, J. Jung, G. Park, and S. Choi. Psychophysical model for vibrotactile rendering in mobile devices. *Presence: Teleoperators and Virtual Environments*, 19(4):364–387, 2010.

[73] J. Ryu, C.-W. Lee, and S. Choi. Improving vibrotactile pattern identification for mobile devices using perceptually transparent rendering. In *Proceedings of the Mobile-HCI*, pages 257–260. ACM, 2010.

[74] E. D. Scheirer. Tempo and beat analysis of acoustic musical signals. *The Journal of the Acoustical Society of America*, 103(1):588–601, 1998.

[75] S. S. Shiffman, M. L. Reynolds, and F. W. Young. *Handbook of multidimensional scaling*. New York, NY, USA: Academic Press, 1981.

[76] S. S. Stevens. Tactile vibration: Dynamics of sensory intensity. *Journal of Experimental Psychology*, 57(4):210–218, 1959.

[77] S. S. Stevens, G. Stevens, and L. E. Marks. *Psychophysics: Introduction to its perceptual, neural, and social prospects*. Transaction Publishers, 1975.

[78] S. S. Stevens, J. Volkmann, and E. Newman. A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America*, 8:185–190, 1937.

[79] H. Z. Tan. *Information transmission with a multi-finger tactual Display*. PhD thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 1996.

[80] D. Ternes and K. E. Maclean. Designing large sets of haptic icons with rhythm. *Lecture Notes on Computer Science (EuroHaptics 2008)*, 5024:199–208, 2008.

[81] R. T. Verrillo. *Measurement of vibrotactile sensation magnitude*. Mahwah, NJ: Lawrence Erlbaum, 1991.

[82] R. T. Verrillo and A. J. Capraro. Effect of stimulus frequency on subjective vibrotactile magnitude functions. *Attention, Perception, & Psychophysics*, 17(1):91–96, 1975.

[83] R. T. Verrillo and G. A. Gescheider. *Perception via the sense of touch*. London: Whurr Publishers, 1992.

[84] H. Yao, D. Grant, and M. Cruz. Perceived vibration strength in mobile devices: The effect of weight and frequency. *IEEE Transactions on Haptics*, 3(1):56–62, 2010.

[85] Y. Yoo, I. Hwang, and S. Choi. Consonance perception of vibrotactile chords: a feasibility study. In *Proceedings of the 6th international conference on Haptic and audio interaction design*, pages 42–51. Springer-Verlag, 2011.

[86] G. Young and A. S. Householder. A note on multidimensional psychophysical analysis. *Psychometrika*, 6(5):331–333, 1941.

[87] A. Zils, F. Pachet, O. Delerue, and F. Gouyon. Automatic extraction of drum tracks from polyphonic music signals. In *Proceedings of International Symposium on Cyber Worlds*, pages 179–183. IEEE Computer Society, 2002.

[88] A. Zlatintsi, P. Maragos, A. Potamianos, and G. Evangelopoulos. A saliency-based approach to audio event detection and summarization. In *Proceedings of European Signal Processing Conference (EUSIPCO)*, pages 1294–1298. IEEE, 2012.

[89] E. Zwicker, G. Flottorp, and S. S. Stevens. Critical band width in loudness summation. *The Journal of the Acoustical Society of America*, 29(5):548–557, 1957.

Acknowledgements

감 사 의 글

먼저 햅틱스에 대해 아무것도 모른 채로 연구실에 들어온 저를 연구자로 성장시켜 주신 최승문 교수님께 감사의 말씀을 드립니다. 제 스스로 어리석고 부끄러웠던 순간들이 수없이 많았지만 교수님의 열정적이면서도 세심한 지도 덕분에 제가 일어설 수 있었습니다. 수업과 논문심사를 통해 저에게 많은 가르침과 조언을 주신 김정현 교수님, 이승용 교수님, 한성호 교수님께도 이 지면을 빌어 감사드립니다. 연구 도중에 교수님들께서 가르쳐주신 다양한 분야의 지식들이 제가 새로운 것을 찾아내고, 문제를 해결하여 연구를 완성시켜 나가는 밑거름이 되었습니다. 연구과제의 책임자로서, 또 햅틱스 분야의 선배 연구자로서 격려와 조언을 아끼시지 않으셨던 경기욱 박사님께도 감사의 말씀을 드립니다. HVR 연구실의 선배님들, 동기들, 후배들의 도움도 제게는 잊을 수 없는 고마움으로 남습니다. 성길이형, 종현이형, 석희형, 재훈이형, 성훈이형, 재영이형, 인이, 채현이, 갑종이, 건혁이, 재봉이형, 종만이, 경표, 명찬이, 호진이, 성환이, 용재, Reza, Phoung, 좁은 지면에 그 순간들을 모두 표현할 수는 없지만 연구실에서 더 좋은 후배로, 든든한 동기로, 멋진 선배로 잘해주지 못해 미안한 마음과 함께 생생히 기억속에 간직하며 살아가려 합니다. 저와 함께 포항에서 20대를 보낸 이름보다 별명이 익숙한 윤기, 재현이, 인태, 동혁이, 그리고 먼저 포항을 떠나 사회속에서 각자의 길을 걷고 있는 11분반 친구들에게도 함께한 시간 만큼이나 고마움을 느낍니다. 포항에서의 시간 동안 자주 연락드리지도 찾아뵙지도 못한 부모님께는 아들의 긴 학업을 묵묵히 기다려주시고 응원해주셔서 감사하다는 말씀을 드립니다. 마지막으로 제가 지치고 힘들어할 때 위로해주고 격려해준 여자친구 지경이와 하나님께 깊은 고마움과 사랑을 전하며 이 논문의 끝을 맺고 새로운 연구자로서의 삶을 시작하려 합니다. 감사합니다.

Curriculum Vitae

Name : Inwook Hwang

Education

2002–2006 : Computer Science and Engineering, POSTECH (B.S.)

2006–2013 : Department of Computer Science and Engineering, POSTECH (Ph.D.)
Thesis Title :
진동 자극의 인지적 분석과 이를 응용한 음악의 진동촉감 표현기법**(Perceptual Analysis of Vibrotactile Stimuli and Its Application to Vibrotactile Rendering of Music)**
Advisor: Prof. Seungmoon Choi

Affiliation

HVR Lab., Dept. of Computer Science and Engineering, POSTECH

Publications

International Journals

1. Inwook Hwang, Hyeseon Lee, and Seungmoon Choi, “Real-time Dual-band Haptic Music Player for Mobile Devices,” To appear in the *IEEE Transactions on Haptics*, 2013 (In Press).

2. Inwook Hwang, Jongman Seo, Myongchan Kim, and Seungmoon Choi, “Vibrotactile Perceived Intensity for Mobile Devices as a Function of Direction, Amplitude, and Frequency,” To appear in the *IEEE Transactions on Haptics*, 2013 (In Press).

3. Jongwon Lee, Inwook Hwang, Keehoon Kim, Seungmoon Choi, Wan Kyun Chung, and Young Soo Kim, “Cooperative Robotic Assistant with Drill-By-Wire End-Effector for Spinal Fusion Surgery,” *Industrial Robot: An International Journal*, vol. 36, no. 1, pp. 60-72, 2009.

4. Inwook Hwang, Sunghoon Yim, and Seungmoon Choi, “Haptic Discrimination of Virtual Surface Slope,” *Virtual Reality*, 2013 (In Revision).

International Conferences

1. Inwook Hwang, Jongman Seo, Myongchan Kim, and Seungmoon Choi, “Perceived Intensity of Tool-Transmitted Vibration: Effects of Amplitude and Frequency,” In *Proceedings of the IEEE International Symposium on Haptic Visual-Audio Environments and Games (HAVE)*, pp. 1-6, 2012.

2. Inwook Hwang, and Seungmoon Choi, “Effect of mechanical ground on the vibrotactile perceived intensity of a handheld object,” *Lecture Notes on Computer Science (Eurohaptics 2012, Part II)*, vol. LNCS 7283, pp. 61-66, 2012.

3. Yongjae Yoo, Inwook Hwang, and Seungmoon Choi, “Consonance Perception of Vibrotactile Chords: A Feasibility Study,” *Lecture Notes on Computer Science (HAID 2011)*, vol. LNCS 6851, pp. 42-51, 2011.

4. Inwook Hwang, Karon E. MacLean, Matthew Brehmer, Jeff Hendy, Andreas Sotirakopoulos, and Seungmoon Choi, “The Haptic Crayola Effect: Exploring the Role of Naming in Learning Haptic Stimuli,” In *Proceedings of the the IEEE World Haptics Conference (WHC)*, pp. 385-390, 2011.

5. Ki-Uk Kyung, Jeong Mook Lim, Yo-An Lim, Suntak Park, Seung Koo Park, Inwook Hwang, Seungmoon Choi, Jongman Seo, Sang-Youn Kim, Tae-Heon Yang, and Dong-Soo Kwon, “TAXEL: Initial Progress Toward Self-Morphing Visio-Haptic Interface,” In *Proceedings of the the IEEE World Haptics Conference (WHC)*, pp. 37-42, 2011 (Oral presentation; acceptance rate = 16.6%).

6. Inwook Hwang and Seungmoon Choi, “Perceptual Space and Adjective Rating of Sinusoidal Vibration Perceived via Mobile Device,” In *In Proceedings of the Haptics Symposium (HS)*, pp. 1-8, 2010 (Nominee for Best Student Paper Award; Oral presentation; Acceptance rate = 29.5%).

7. In Lee, Inwook Hwang, Kyung-Lyong Han, Oh Kyu Choi, Seungmoon Choi, and Jin S. Lee, “System Improvements in Mobile Haptic Interface,” In *Proceedings of World Haptics Conference (WHC)*, pp. 109-114, 2009 (Winner of the best student paper award).

8. Jaeyoung Cheon, Inwook Hwang, Gabjong Han, and Seungmoon Choi, “Haptizing Surface Topography with Varying Stiffness Based on Force Constancy: Extended Algorithm,” In *Proceedings of the Haptics Symposium (HS)*, pp. 193-200, 2008.

9. Kyung-Lyong Han, Oh Kyu Choi, In Lee, Inwook Hwang, Jin S. Lee, and Seungmoon Choi,“Design and Control of Omni-Directional Mobile Robot for Mobile Haptic Interface,” In *Proceedings of the International Conference on Control, Automation, and Systems (ICCAS)*, pp. 1290-1295, 2008.

Demonstrations

1. Inwook Hwang, Moonchae Joung, Sunwook Kim, Kyunghun Hwang, Jaecheon Sa, and Seungmoon Choi, “Real-time Dual-band Haptic Music Player for Mobile Devices,” In *Eurohaptics*, 2010.

2. Jonghyun Ryu, Inwook Hwang and Seungmoon Choi, “Graphical Authoring Tools for Vibrotactile Patterns: posVibEditor,” In *World Haptics Conference (WHC)*, 2009.

Domestic Papers

1. Yongjae Yoo, Inwook Hwang, Jongman Seo and Seungmoon Choi, “Multiple Vibration Signal Feedback for Mobile Devices,” *Smart Media Journal*, vol. 1, no. 4, pp. 8-17, 2012.

2. Yongjae Yoo, Inwook Hwang, and Seungmoon Choi, “Consonance Perception of Vibrotactile Chords: A Feasibility Study,” *Journal of Korean Institute of Next*

Generation Computing, vol. 7, no. 5, pp. 24-34, 2011.

3. Yongjae Yoo, Inwook Hwang and Seungmoon Choi, “Consonance Perception of Vibrotactile Chords: A Feasibility Study,” In *Proceedings of the HCI Korea*, pp. 282-284, 2012.

4. Inwook Hwang, Seungmoon Choi, Moonchae Joung, Sunwook Kim, Kyunghun Hwang and Jaecheon Sa, “Dual-band Vibrotactile Music Player for Real-time Playback in Mobile Devices,” In *Proceedings of the HCI Korea*, pp. 251-253, 2011.

5. Inwook Hwang and Seungmoon Choi, “Perceptual Distance Measurement of Vibration Stimulus with Various Frequencies and Amplitudes,” In *Proceedings of the Korean Haptics Community Workshop*, 2009.

6. In Lee, Inwook Hwang, Kyung-Lyoung Han, Oh Kyu Choi, Jin S. Lee, and Seungmoon Choi, “Practical Issues of Mobile Haptic Interface and Their Improvements,” In *Proceedings of the HCI Korea*, pp. 390-395, 2009.

7. Donghoon Lee, Sung H. Han, Gunhyuk Park, and Inwook Hwang, “Usability Evaluation of Expansion Methods on Grouped Icons,” In *Proceedings of Eromonomics Society of Korea*, pp. 276-281, 2008.

8. Jaehoon Jung, Inwook Hwang, In Lee, Chaehyun Lee, Gunhyuk Park, Jane Hwang, Seungmoon Choi and Gerard J. Kim, “Remote Control for Motion-Based Interactions,” In *Proceedings of the HCI Korea*, pp. 115-122, 2007.