

운동기능 학습보조를 위한 촉각 가이드런스:
드럼 리듬 학습에의 응용

**Vibrotactile Guidance for Motor Skill
Learning and Its Application to
Drumming Learning**

Vibrotactile Guidance for Motor Skill Learning and Its Application to Drumming Learning

by

In Lee

Department of Computer Science and Engineering

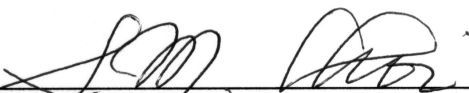
POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

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

December 20, 2014

Approved by


Seungmoon Choi, Academic Advisor

Vibrotactile Guidance for Motor Skill Learning and Its Application to Drumming Learning

In Lee

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

12/20/2014

Committee Chair

최 승 문


(Seal)



Member

이 근 배

(Seal)



Member

이 승 용

(Seal)



Member

한 성 호

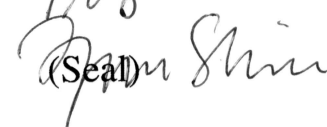
(Seal)



Member

신 혜 수

(Seal)



DCSE 이 인 In Lee, Vibrotactile Guidance for Motor Skill Learning and Its
20065073 Application to Drumming Learning. 운동기능 학습보조를 위한 촉각
가이던스: 드럼 리듬 학습에의 응용,
Department of Computer Science and Engineering, 2015, 105P,
Advisor: Seungmoon Choi. Text in English

Abstract

With recent medical and economical advances, people are spending more and more time on leisure-time physical activities such as jogging, swimming, or playing music. These activities, or motor skills, often require to perform a certain sequence of unit movements at a given performance speed, which can be difficult to learn. One efficient way of skill learning is to observe another's demonstration and to practice the skill based on it. It seems not efficient to use the sight or hearing for demonstration observation because those channels are usually occupied by the skill. Via the sense of touch, it is possible to deliver the guidance information while making the sight and hearing available for the acquisition of other vital information about the skill. Also, it can provide guidance to whom visual demonstration is not applicable, i.e., the blind.

In this regard, we propose a vibrotactile guidance method for learning complex procedural motor skills using vibrotactile cues generated by multiple vibration actuators worn by the learner. Drumming is used as a target skill representing the motor skills requiring fast, patterned, coordinated discrete movements of multiple limbs. A natural egocentric mapping of our system from the body site of vibrotactile stimulation to a target percussion instrument (PI) in a drum set enables intuitive guidance for striking movements. The method also informs the learner of two levels of PI striking strength by varying both the intensity and duration of vibrotactile cues.

To evaluate the performance of our method in delivering guidance information, a series of human-subject experiments were conducted. An initial perceptual assessment of the

system showed 96.18% of accuracy and 0.77 s of time in delivering the information on the target PI and strength level for a single strike, and it was 55.03% and 1.11 s for a pair of concurrent strikes. When provided with a sequence (4 items) of single or paired vibrotactile cues, the participants showed 88.4, 56.3, 23.3% of response accuracy and 7.53, 10.15, and 13.71 s of response time for simple, moderate, and complex sequences, respectively.

The effectiveness of our guidance system was also evaluated with an actual experimental scenario of drum rhythm learning. Three sets of short drum rhythms were learned for three days using different learning methods (practice only, practice with video guidance, and practice with vibrotactile guidance), and the participant's performance was compared among the learning methods. The experimental results indicated that vibrotactile guidance was as helpful as video guidance in learning the temporal pattern of a drum rhythm, which suggests that our vibrotactile guidance method is a viable alternative to video guidance.

Contents

1	Introduction	1
1.1	Motivation and Goal of Research	1
1.2	Target Motor Skill	2
1.3	Approach	3
1.4	Contribution	4
1.5	Organization	5
2	Background	6
2.1	Motor Learning	6
2.2	Haptic Guidance for Motor Learning	7
2.2.1	Force-feedback Guidance	7
2.2.2	Vibrotactile Guidance	8
2.2.3	Cuncurrency in Haptic Guidance	10
3	Initial Study: Multimodal Guidance of Random Drum Sequences	11
3.1	Experiment Design	12
3.1.1	Task and Apparatus	12
3.1.2	Guidance Methods	14
3.1.3	Experimental Conditions and Participants	17
3.1.4	Procedures	17
3.1.5	Performance Measures	18
3.2	Results	19
3.2.1	Familiarization	19

3.2.2	Error Ratio	20
3.2.3	Task Completion Time	21
3.2.4	Subjective Evaluation	22
3.3	Discussion	23
3.3.1	Familiarization	23
3.3.2	Error Ratio	24
3.3.3	Task Completion Time	25
3.3.4	Subjective Evaluation	26
3.4	Conclusions	26
4	Vibrotactile Drumming Guidance System	28
4.1	Hardware	28
4.2	Guidance Method	29
4.3	Vibrotactile Guidance System Improvements	30
4.3.1	System Improvements	32
5	Experiment I: Identification of a Cue	34
5.1	Methods	36
5.1.1	Participants	36
5.1.2	Three Cue-Response Tasks	36
5.1.3	Procedures	38
5.1.4	Performance Measures	39
5.2	Results	41
5.2.1	Effects of Processing Stage	41
5.2.2	Effects of Cue Position and Strength	43
5.3	Discussion	44
5.3.1	Effects of Processing Stage	44
5.3.2	Effects of Cue Position and Strength	45
6	Experiment II: Identification of Paired Cues	47
6.1	Methods	47
6.1.1	Participants	47
6.1.2	Experimental Tasks and Vibrotactile Cues	47
6.1.3	Procedures	48
6.1.4	Performance Measures	49

6.2	Results	50
6.2.1	Comparison between Experiment I-1 and I-2	50
6.2.2	Effects of Concurrent Cue Presentation	52
6.2.3	Effects of Cue Position Pair and Strength Pair	52
6.3	Discussion	54
6.3.1	Effects of Concurrent Cue Presentation	54
6.3.2	Effects of Cue Position Pair and Strength Pair	56
7	Experiment III: Series Identification of Single or Paired Cues	58
7.1	Methods	58
7.1.1	Participants	58
7.1.2	Task and Stimuli	58
7.1.3	Procedures	60
7.1.4	Performance Measures	63
7.2	Results	63
7.2.1	Recognition of Single or Paired Cues	64
7.2.2	Series Recognition of Cues	65
7.3	Discussion	65
7.3.1	Recognition of Single or Paired Cues	65
7.3.2	Series Recognition of Cues	68
8	Experiment IV: Application to Drumming Learning	71
8.1	Methods	71
8.1.1	Participants	71
8.1.2	Task	71
8.1.3	Conditions	73
8.1.4	Procedures	74
8.1.5	Performance Measures	75
8.2	Results	77
8.2.1	Performance Before Learning	78
8.2.2	Guidance and Practice During Learning	78
8.2.3	Performance Gains from Learning	79
8.3	Discussion	81
8.3.1	Guidance and Practice During Learning	81
8.3.2	Targeting	81

8.3.3	Timing	82
8.3.4	Strength Level	83
9	Conclusions and Future Work	85
	한글 요약문	88
	REFERENCES	90

List of Figures

3.1	Instrument-note-cue relationships.	12
3.2	Example task and visual guidance given to the participant.	12
3.3	Participant learning the sight reading skill.	13
3.4	Schematic diagram of rhythm generation using amplitude modulation.	15
3.5	Experimental procedures.	19
3.6	Mean error ratios with standard error bars.	20
3.7	Mean task completion times with standard error bars.	22
3.8	Mean subjective evaluation scores with standard error bars.	23
4.1	Hardware for vibrotactile drumming guidance. (a) Electronic drum set with nine percussion instruments (PIs). (b) and (c) Vibrotactile vest and ankle bands (mirror images), respectively. Relationships between PIs and body sites are denoted by numbers.	29
4.2	(a) Hardware for vibrotactile drumming guidance. (b) Modified layout of vibration motors (mirror images). Relationships between PIs and body sites are denoted by numbers.	31
5.1	Visual scenes provided in each experiment task.	37
5.2	Mean error rates (%) for four processing stages. Error bars represent standard errors.	42
5.3	Mean processing times (s) for four processing stages. Error bars represent standard errors.	42

5.4	Tukey’s HSD test results on the processing time (s) of each processing stage for different body sites. Black dots represent stimulated body sites, and those with the same alphabet are of the same performance group by the test.	44
6.1	Mean error rates (%) for four processing stages. Error bars represent standard errors.	51
6.2	Means and standard errors of the processing times (s) for four processing stages.	51
6.3	Graphical representation of the best (upper row) and the worst (lower row) nine cue position pairs for the cue detection time (τ^D), the cue recognition error rate (e^R) and time (τ^R), and the response execution time (τ^E). Each pair is represented by a line connecting two body sites (circles), with thickness representing its rank (thick: 1–3rd; normal: 4–6th; dashed: 7–9th).	54
6.4	Mean recognition error rates (%) for the pairs of cue strengths (n: normal and A: accented). Type A represents the misses caused by incorrect selection of a target circle, while Type B is for those with correct target selection but wrong strength response. Conditions with the same alphabet above the bar are of the same performance group by the test.	55
7.1	Example of visual scenes provided to the participant and brief summary of two experimental sessions.	61
7.2	Performances of single cues and cue pairs. Higher correctness score and shorter response time indicates better performance.	64
7.3	Performances of three cue combination types of different complexity levels.	66
8.1	Three sets of target rhythms.	72
8.2	Example visual scene.	73
8.3	Performance measures of participants measured at T1–7 for three learning methods.	84

List of Tables

5.1	Two-way ANOVA results on the main effects of cue position and cue strength on two performance measures (e : error rate, τ : processing time) measured in four processing stages (D: detection, R: recognition, C: choice, E: execution).	43
6.1	T-test results that compared two experiments (I-1 and I-2; I-2 and II) using two performance measures (e : error rate, τ : processing time) for four processing stages (D: detection, R: recognition, C: choice, E: execution). .	50
6.2	Two-way ANOVA results of Experiment II for the effects of cue position pair and strength pair on two performance measures (e : error rate, τ : processing time) for four processing stages (D: detection, R: recognition, C: choice, E: execution).	53
7.1	List of choices for a single cue and a cue pair. Each body site is represented by a unique number in the same way with Fig. ??	60
7.2	Comparisons of three cue combinations used in the experiment.	60
8.1	Three-way ANOVA results on the effects of test point (T), learning method (M), and existence of visual feedback (F) for three performance measures. .	79

Chapter 1

Introduction

This study pertains to haptic guidance for learning coordinated discrete movements of multiple limbs using vibrotactile cueing. Drumming is used as a target skill representing those requiring fast, patterned, coordinated discrete movements of the two arms and two feet. Our guidance method makes use of all of the spatial, temporal, and intensity aspects of vibrotactile cues using by multiple factors distributed on the learner's body.

1.1 Motivation and Goal of Research

A steady decline in the average work hours and an increased lifespan as a result of medical advances have allowed individuals to have more leisure time. People are spending more and more time on leisure-time physical activities for better health and lifestyle [72, 56]. Frequent physical activities are ranged from those of moderate intensity (e.g., walking or weight lifting) to more vigorous ones (e.g., running, swimming, or aerobics), including artistic activities (e.g., dancing or playing music). These activities, or motor skills, often require to perform a certain sequence of unit movements at a given performance speed, which can be difficult for beginners and requires explicit external guidance. For the guidance, computerized guidance devices can be a good alternative of relatively expensive personal trainers or instructors.

Motivated by this, we aim to develop an efficient guidance method for the learning of

motor skills. Specifically, our research goal is to construct an efficient guidance system for a complex cognitive-motor skill involving a number of actions of multiple body extremities, which can be mastered after a prolonged practice.

1.2 Target Motor Skill

Playing drum rhythms, in which an individual reads sheet music and simultaneously plays a drum set according to the music piece, was selected for the research as it was considered to be one good example of complex cognitive-motor skills. A drum set is a musical instrument comprised of many percussion instruments (PIs), and it provides the *groove* of music by the repetitive and rhythmic presentation of percussion sound patterns, i.e., drum rhythms. Playing all drum rhythms correctly and fluently is vital to good drumming, so learning of drum rhythms is the main content of drum lessons for novice drummers. Playing a drum rhythm involves a series of fast, single or multiple drum strikes, and this requires a series of fast coordinated discrete movements of hands and feet. Every strike must be executed onto a PI with high accuracy in position, timing, and strength; even a small error in a drum strike can cause a substantial change in the overall perception of the drum rhythm.

For practice, learners read musical notations on a drum music piece, interpret their meanings, and execute the designated drumming sequence. They make various execution errors because they lack a well-established knowledge for reading music and motor program for drumming action. During drum lessons, an instructor helps learners in various ways, e.g., showing a demonstration of the desired play for transferring the notation-to-action model of the instructor to the learners. In case of self-learning, the learner can consult video lectures [27, 60, 53], or tutoring software [55]. Video demonstration requires intensive visual processing for recognizing fast and distant striking movements of multiple limbs, especially due to the limited and fixed viewing angles. In addition, it is highly expensive to prepare video demonstration for every drum rhythm that the learner is practicing. As for tutoring software, their principal means of assistance is correctness feedback on the learner's play (e.g., marking played notes with different colors or symbols). Such feedback is helpful for intermediate learners, but may not be suitable to the novices who cannot or barely perform

drum rhythms.

For the research, we assume beginner-level drum playing, which involves up to nine striking positions (one for each PI) and two striking strengths (normal and accented). Playing speed is relatively slow, but still challenging (30–60 BPM; 0.5–1.0 s between strikes for 8-beat rhythms).

1.3 Approach

The two major means of guiding or promoting motor skill acquisition are demonstration and augmented feedback [3, 43]. Demonstration provides the learner with an idea of an *ideal* performance, which can be used as a reference during practice, while augmented feedback gives information on the *learner's* performance to guide and encourage learning. Demonstration is thought to have the most influence at the initial stage of learning, and augmented feedback has a larger effect in the rest period of learning. Because the two guidance methods have effects on different learning stages, we expect that learning efficiency would be greater when both methods are utilized than when only one of them is used.

Humans have limited capacities of attention and short-term memory [2]. For this reason, guidance should concisely emphasize the salient aspects of a target skill, and cueing¹ or feedback consisting of only a small number of single-modal stimuli is expected to be sufficient, even more beneficial than human coaching [66]. For the same reason, we expect that the two types of guidance are better to be delivered via different modalities to avoid sensory overloads. For the selection of sensory modalities, sight can be one good choice as it has been widely used and shown its efficacy in motor learning [43]. Among the remaining modalities, touch would be good since it is simple to deliver the spatial aspects of a motion (e.g., target body part or direction of motion).

The sense of touch has several advantages that make the sense preferable to other modalities for motor skill guidance. First, associating a tactile stimulation site with the body part to be moved can be natural and intuitive, whereas such relations are not always self-evident

¹Here, cueing refers to a simplified form of demonstration that is comprised of single-modal stimuli given at the same time of practice.

for audio stimuli. Second, the touch does not require the sight and hearing of the learner, making these senses available for the acquisition of additional information about the skill (e.g., the verbal instructions of an instructor) or for concurrent activities (e.g., talking with someone or listening to music). Lastly, it is robust against light or sound interference, suggesting high applicability to outdoor activities.

Attracted by the advantages of touch, intensive research on the use of haptic guidance for motor learning has been proposed recently. For example, vibrotactile cues were provided to the hands and feet to guide drumming paces [36] and walking paces [71], respectively, to the fingers of one hand for memorization of piano music [25], to the wrists and ankles for a multi-limb drumming task [22], and to many locations on the body to signal whole body movements in snowboarding [63] and generalized body movements [40]. Several studies used vibrotactile feedback to teach arm motions [4, 6], a gait pattern [61], and a violin playing posture and an arm stroking motion [68]. There also exist many studies that used force/torque feedback devices for motor learning, and they spread out widely in their applications, tasks, and teaching strategies [45, 52].

1.4 Contribution

The expected contributions of this study is as follows. 1) A vibrotactile guidance method is proposed for complex procedural motor skills by utilizing various properties (location, intensity, and duration) of vibration stimuli. 2) A systematic evaluation procedures for guidance methods are also introduced. The procedure is beneficial to us to understand which sensory-cognitive processing stage is the main bottleneck for guidance delivery, to determine the applicable area of the guidance method, and to decide solutions for further improvement. 3) a drum rhythm learning system is developed. The system helps the learner in various ways, including multi-modal guidance and visual feedback, so that the learner can learn target drum rhythms more easily. 4) The effectiveness of our vibrotactile guidance in guiding drum rhythms is shown by a human-subject experiment under a realistic drumming learning scenario.

1.5 Organization

The remainder of this study is organized as follows. In Chapter 2, the background theories and literature related on motor learning and haptic guidance are introduced in detail. Our initial study to obtain the basic knowledge and system requirements about haptically-guided drumming learning is described in Chapter 3. Based on the results from the initial study, we developed a vibrotactile drumming guidance system with a guidance method for drum rhythms (Chapter 4). We also conducted four human-subject experiments to evaluate the effectiveness our guidance system in a systematic manner, and they are explained in Chapter 5–8. In Chapter 5, we focus on the performance of a single vibrotactile cue in the point of view of information delivery, and do the same for a pair of simultaneous vibrotactile cues in Chapter 6. Chapter 7 describes the performance evaluation of our method in instructing a short sequence of single or multiple drum strikes. Finally in Chapter 8, we apply our haptic guidance method to drumming learning and compare its effectiveness with those of two other learning methods (practice only and video-guided practice) under a realistic learning scenario and two different feedback conditions. The findings and results of this study is summarized in Chapter conclusions, with a short outlook for future work.

Chapter 2

Background

2.1 Motor Learning

The theory of demonstration-based learning or observational learning was proposed by psychologist Albert Bandura [3, 43]. He argued that individuals can learn a skill or behavior by observing someone else's performance and explained that there are four processes involved in this type of learning: attention, retention, production, and motivation. Attention is the process of grasping important information (i.e., motor memory in our case) from the demonstration by paying attention to it. Retention is the process of storing the observed knowledge in the memory for future retrieval. Production is the process of physical practice to become capable of producing the target skill. Motivation is a situation where the obtained skill is needed to be reproduced. Attention and retention account for the acquisition of a skill, while production and motivation determine the reproduction quality. This theory states that, for effective demonstration-based learning, the demonstration should be salient and well organized, with sufficient practice time and motivation. Demonstration-based learning is distinguished from another major principle, augmented feedback, which feeds back the practice performance of the learner during learning [44, 62], in that it does not provide such feedback information.

The demonstration and augmented feedback are not necessarily provided by a human. Drawings, voice recordings, films, or any type of media that describes the target skill can

be a good substitute. Furthermore, the media does not have to be very realistic. A more abstracted guidance that contains only a critical aspect of the skill is known to be sufficient and even result in better performance. For instance, Thompson and Russell recently showed that young children could more easily open a door when they observed the door opening on its own than when they observed a human opening it [65].

Facilitated by the above characteristics of guidance, many computer-aided learning methods have been proposed [57, 76, 12, 66, 77, 64]. Although the majority of them are of visual methods, vision is not always necessary or the best modality. It is known that some skills can be delivered more effectively through non-visual modalities. Doody et al. showed that an auditory or audiovisual demonstration was more beneficial than a visual one in a timing task that required complex hand movements [12]. Similar results were also found for a dancing task where a person synchronized a series of dance steps to two auditory rhythms [76], and in a tapping task where a person sequentially pressed four buttons [28]. For touch, Feygin et al. argued that it was better to use a haptic demonstration to teach a person about the time-related aspect of a 3D path-following task, while a visual demonstration was better for the spatial aspect [14].

2.2 Haptic Guidance for Motor Learning

Motor learning via the sense of touch can be found in the literature on haptic guidance, which has been actively researched recently. Haptic guidance refers to the methods that provide kinesthetic or tactile stimuli for the purpose of assisting in the learning or completion of a task.

2.2.1 Force-feedback Guidance

Early studies on haptic guidance aimed at facilitating motor learning by providing the force feedback that enables the learner to experience the ideal, desired movements during training. For example, it was demonstrated that active force guidance can be beneficial for learning a 3D trajectory-following task [14, 5], a 2D trajectory-following task [7], steering [10], and handwriting [49], particularly in timing-related aspects. Force guidance applied to the

rehabilitation of stroke patients enabled them to regain their arm functions better than those who received a conventional therapy [42]. Grindlay carried out a haptic guidance of a one-handed drumming task using a one-degree-of-freedom force-feedback device. In his study, he argued that force guidance had better learning effectiveness than auditory guidance for a rhythmic drumming task in movement velocity accuracy [18].

However, a considerably greater number of studies reported no positive effects of force guidance on motor learning [41, 62, 73, 39, 74]. It is presumably due to the facts that force guidance results in differences in the task context between practice and actual execution and that the learner's attention level decreases as the learner's dependency on guidance stimuli grows over the course of training, both of which lead to inefficient motor learning. Approaches for improvements include progressive haptic guidance, which adaptively controls the intensity or frequency of guidance stimuli depending on the learner's performance [54, 26, 38], and haptic disturbance, which makes the task more challenging to prompt the learner to pay more attention to training [37, 52, 34, 13]. These approaches resolve some disadvantages of the previous fixed-gain force guidance, but extensive research is still required before understanding the ultimate benefits of force guidance. See [48, 52, 62, 73] for a comprehensive review on this topic. As well as skill acquisition of normal people, force-feedback guidance has also widely been applied for the rehabilitation and assistance of the injured, and [42, 45, 33] provide a detailed review for this.

2.2.2 Vibrotactile Guidance

An alternative of force guidance is vibrotactile guidance, and it has been the subject of recent research. Although vibrotactile guidance is unable to provide direct kinetic feedback unlike force guidance, it does have several distinctive merits. Vibrotactile actuators are much more compact and inexpensive, and they can easily stimulate multiple body sites if embedded in a chair or a wearable interface. Therefore, vibrotactile guidance has the potential for an effective delivery of movement instructions, especially for complex coordinated movements of multiple limbs.

Vibrotactile guidance has considered two classes of motor tasks: continuous and pro-

cedural. For continuous tasks, a popular approach is to present vibrotactile stimuli to the body part to move to specify its movement direction. The direction coding scheme is either attractive or repulsive; e.g., a vibration applied onto the palm means to move the hand to the direction of the palm (attractive) or to the direction of the back of the hand (repulsive) [4]. The strength of the vibration is fixed or proportional to the distance to the target position. A vibrotactile sleeve with eight tactors distributed on the elbow and wrist showed that repulsive guidance was effective in guiding complex arm motion trajectories in terms of position accuracy and learning rate [40]. Using similar hardware [30], Bark et al. [4] guided three arm motions (wiping, eating, and cutting) performed with a non-dominant hand, but they found no statistically significant differences in position error or in subjective measures between attractive and repulsive mode. Repulsive vibrotactile guidance to the upper body helped young children learn the proper body posture and bowing motion of playing violin [68]. It is also effective for gait pattern guidance [61] and virtual environment navigation assistance [29].

Vibrotactile guidance for procedural tasks is often called vibrotactile cueing. In this kind of methods, every movement comprising the target task is represented by a unique vibrotactile cue. Then, during learning, the vibrotactile cues are presented in series to teach the learner the movement order. It is important to design vibrotactile cues and their mapping to the movements of a task in such a way that minimizes the learner's effort to recognize the cues and subsequently determine the corresponding movements. For example, to teach two two-measure-long piano phrases composed of five piano keys each, the five piano keys were one-to-one mapped to the five fingers of a hand, and then a series of short vibrotactile stimuli were presented to the fingers to designate which key should be pressed when [24, 25]. Similarly, short vibrotactile stimuli to the wrists and ankles could guide several drum rhythms [22]. A subjective evaluation on this guidance system reported that subjects often missed the vibrotactile cues due to the impact that occurred at drum strikes. Watanabe and Ando found that a haptic demonstration that alternately stimulated the learner's feet with short vibrations had a positive effect on the learning of walking pace [71]. The learners could easily perceive and recognize the vibration signals even when they were actually

performing the task in a real environment, and their responses to the vibration instructions were quicker than those of audio instructions. Spelmezan et al. [63] searched for an intuitive mapping from vibrotactile patterns generated by many factors distributed over the entire body to snowboarding movements, e.g., leaning forward or turning left, through a series of human-subjects experiments.

2.2.3 Cuncurrency in Haptic Guidance

In many haptic guidance studies, unlike traditional motor learning studies, the practice is often performed concurrently with the observation of demonstration or augmented feedback. Such concurrency can provide more observation and practice within a limited learning time, but there also exists the possibility that they interfere with each other, causing insufficient learning. For instance, in a 2D sequential point-selection task, the subject group who practiced with both visual and haptic demonstrations showed worse performances than those who passively observed a visual demonstration [21]. Only well-designed demonstrations that rarely affect the learner's voluntary motion and require less cognitive effort may benefit from the concurrent process. Despite the wide use of concurrent observation and practice, its effect has not been well analyzed and needs further study.

Chapter 3

Initial Study: Multimodal Guidance of Random Drum Sequences

As a initial study, we introduce a learning system for the sight reading skill of simple drum sequences. Sight reading is a cognitive-motor skill that requires reading of music symbols and actions of multiple limbs for playing the music [46, 31]. The system provides knowledge of results (KR) pertaining to the learner’s performance by color-coding music symbols, and guides the learner by indicating the corresponding action for a given music symbol using additional auditory or vibrotactile cues.

To evaluate the effects of KR and guidance cues, three learning methods were experimentally compared: KR only, KR with auditory cues, and KR with vibrotactile cues. The task was to play a random 16-note-long drum sequence displayed on a screen. Thirty university students learned the task using one of the learning methods in a between-subjects design. The experimental results did not show statistically significant differences between the methods in terms of task accuracy and completion time. This suggests that visual KR can be dominant for learning the task and the role of auditory or vibrotactile guidance cues can be subsidiary.



Fig. 3.1 Instrument-note-cue relationships.

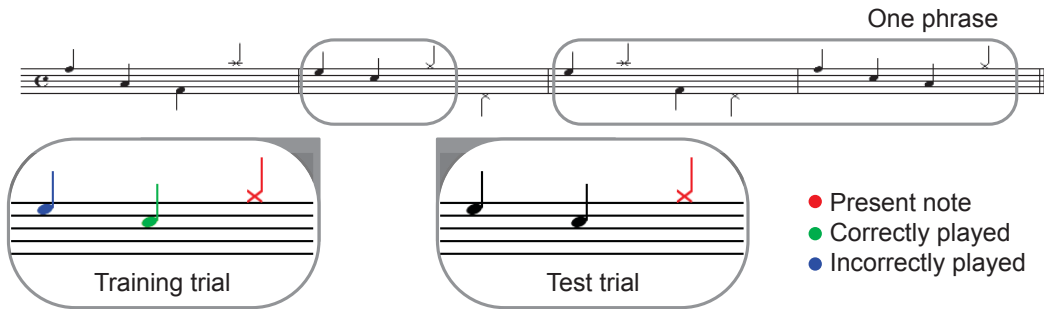


Fig. 3.2 Example task and visual guidance given to the participant.

3.1 Experiment Design

3.1.1 Task and Apparatus

The task was to play random drum music at one’s own pace in order to acquire the ability to perform arbitrary music pieces (i.e., sight reading of music). Considering the initial nature



Fig. 3.3 Participant learning the sight reading skill.

of our research, in the present study, we only focused on the teaching of the sight reading of the tonal patterns of simple music that involved no chords. Thus, the task was modeled as a sequence of single notes, and the required action was striking a certain percussion instrument (PI) specified by the pitch of a note (see Figure 3.1)¹. Specifically, a task was defined as a group of two phrases and, in each phrase, all the PIs of the drum set appeared only once in random order. This was to ensure balanced learning of the pitch-PI relations. A phrase had eight notes since the drum set used in the system composed of eight PIs (two pedals, one snare drum, three tom-toms, and two cymbals). Because our goal was to teach the sight reading skill, not a specific tonal sequence, the exact sequence of a task varied with the trial. An example of the task is shown in Figure 3.2, with two insets describing the visual guidance used in our system (discussed in the next section).

The system consisted of a control computer, a digital drum set (Model DD506; MEDELI

¹The music notation for a drum set has not been standardized yet. Thus, we used one common convention with slight modifications for this study.

Electronics), a 24-inch LCD monitor, headphones, two vibration actuators (Haptuator; Tactile Labs) with a digital-to-analog converter and a current amplifier (Figure 3.3). The control computer executed the experimental software and controlled the other devices. The drum set was used for gathering data from the learner's playing. The LCD monitor displayed the striking sequence, and the headphones provided the auditory cues and the playing sound from the drum set. The two actuators were attached to the learner's upper arms using arm-bands (see the bottom-right inset of Figure 3.3) to provide vibrotactile cues.

3.1.2 Guidance Methods

As the task is sight reading of a drum sequence, the learner should be provided with the target sequence in the form of music score during the performance. Because the learner's sight was already involved in performing the task, visual KR with haptic cueing seemed more reasonable than the opposite configuration (i.e., haptic KR with visual cueing) in terms of balanced distribution of information; KR required 1 additional bit of information capacity to deliver the correctness of performance, whereas cueing used 3 bits to designate one of eight PIs.

During training, visual KR regarding the learner's performance was provided using color coding (Figure 3.2). At first, the present note to be played was highlighted in red to facilitate identification from the others. Then, the note turned green if the learner correctly responded by striking the corresponding PI, whereas it turned blue for an incorrect response. We used this color-coding convention because green has positive meanings while blue is negative in Korean culture, and red, which is more attentive than green or blue, was reserved for the present note. During the tests, only the current note was shown in red, while all the other notes were in black.

Vibrotactile cues are commonly obtained by combining different levels of vibration frequency, strength, rhythm (or amplitude envelope; the shape of strength change over time), or location. Among these design factors, vibration strength seemed inappropriate because weak vibrations would not be perceived by the learner, especially in a situation of high cognitive and motor requirements. Further, it seemed a better idea to reserve the strength

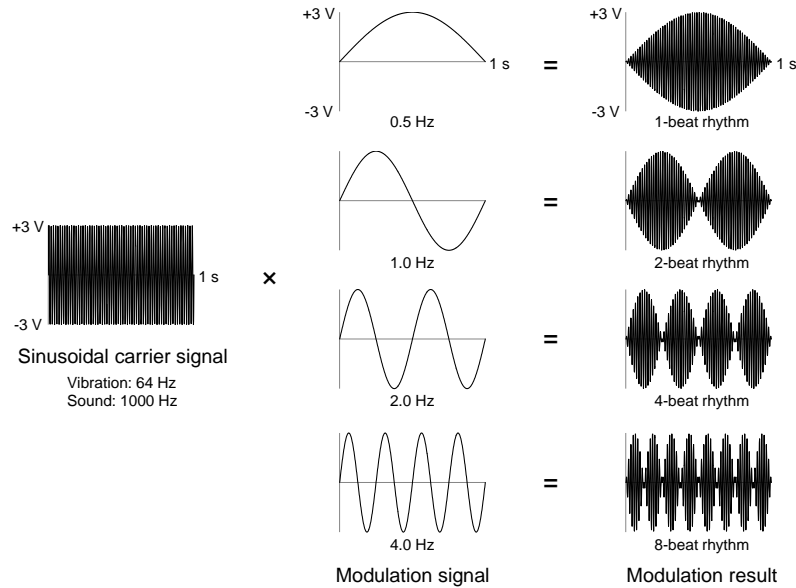


Fig. 3.4 Schematic diagram of rhythm generation using amplitude modulation.

for the guidance of striking strength, which will be added in the future. As regards vibration frequency, though the actuator used (Haptuator; Tactile Labs) was a wideband actuator that can render vibrations of 50–500 Hz, only a very narrow range around its resonance frequency (60 Hz) could generate vibrations strong enough for our purpose, restricting the use of multiple levels of frequency. Hence, the vibration rhythm and location were the only usable factors in designing the cues, and they are known to be much more discriminative than strength and frequency.

For haptic cueing, we used four monotonic vibrational rhythms (1, 2, 4, and 8 beats per second) in combination with two stimulation locations (left and right upper arms), thereby resulting in a total of eight vibrotactile cues. The rhythms were generated by modulating the amplitude of a 64-Hz (near the resonance frequency of Haptuator) sinusoidal vibration signal with the modulation frequencies of 0.5, 1.0, 2.0, and 4.0 Hz, respectively (see Figure 3.4). It should be addressed that the above implementation of rhythms is an indirect display of lower (modulation) frequency signals, which cannot be rendered directly, using a higher frequency signal, thus the rhythms are actually a set of frequency levels. It is

known that 0.0 (no modulation), 1.0, 2.0, and 5.0 Hz signals delivered via a 150 Hz carrier signal are perceptually distant and easily distinguishable [50]. Our modulation frequencies are similar to those frequencies and expected to be easily distinguishable, which was also confirmed in our pilot experiments. We selected the upper (proximal) arms instead of the hands to prevent the cues from being masked by lower arm movements or vibrations produced by striking the PIs [51, 22, 36]. Each cue was assigned to one of the eight PIs of the drum set based on the relative spatial locations of the PIs. The PIs at higher positions (in terms of the distance from the ground) were associated to the rhythms with more beats (i.e., higher modulation frequencies), and the PIs on the left/right side received cues on the left/right upper arm, respectively (see Figure 3.1). Due to the intuitive and straightforward mapping to the PIs, compared to that of drum music notation, the cues were expected to be helpful in performing the task correctly, thereby delivering the idea of correct performance and promoting the learning, especially at the beginning of learning,

An auditory cueing method was also devised to assess the effect of modality difference. This method was essentially the same as the vibrotactile method, except for the carrier frequency (1 kHz) and the stimulation locations (left and right ears). We also tested two other auditory methods in pilot experiments: one using the percussion sounds and the other using the spoken names of the PIs. However, they were excluded from the present experiment since their performances were below that of the rhythm-based method. Participants often faced difficulties in matching the percussion sounds or names with the PIs. It is probably because the participants were unfamiliar with the sounds and names, and there were no easily memorable rules in the relations to the PIs.

The vibrotactile and auditory cues were 1-s long. Their magnitudes were sufficiently strong to be perceived clearly. The cue for the present note was initiated as soon as the note was highlighted, and it was repeated every 3 s until the note was answered. If the note was answered, then the present cue was immediately stopped, and the cue for the next note was initiated.

3.1.3 Experimental Conditions and Participants

The learning method (KR, KR+AC, and KR+TC) was the main factor of the experiment. The KR method was a method which the learner was required to learn the task using the KR feedback only. In addition to the KR feedback, the KR+AC and KR+TC methods provided auditory and vibrotactile cues, respectively. In this experiment, the KR method was a baseline condition because our purpose was to evaluate the effectiveness of cueing-based guidance additionally given with visual KR. We did not include a no-guidance condition because it is impractical. Without guidance, subjects cannot determine the correctness of their performance, and thus they are expected to show no improvement from learning.

We recruited 30 healthy male university students (18–28 years old with an average of 20.9) and assigned 10 participants to each method. The experiment was restricted to male participants to prevent influences from gender differences. The participants were screened by self-report to ensure that they were unable to read drum sheet music or play any musical instrument. It was also confirmed that they had not received additional music lessons at least for the last five years, except formal school education. After the experiment, the participants were compensated for their participation.

3.1.4 Procedures

Upon arrival, the participants received brief explanations on the experiment. The KR+AC and KR+TC groups were also informed of the mapping rules of the respective guidance cues to the PIs. Then, they were asked to strike the corresponding PI to a given guidance cue to become familiar with the cues. This was continued until they correctly responded to all the cues of two consecutive phrases. The KR group freely played two phrases without the guidance cues for familiarization with the system. For all groups, this familiarization session was started without any prior exposures to the cues, and no visual display was given during the stage.

After the familiarization, the participants studied a pitch-PI diagram, which was similar to Figure 3.1, for 30 s to gain initial knowledge on sight reading for drum music. Next, they were asked to play two phrases displayed on the monitor as accurately and quickly as

possible, in order to evaluate their basic ability before the training session (PRE). In the test, the participants were required to perform the task using their own knowledge only; no extrinsic guidance (cueing or KR) was provided. Then, the participants learned sight reading by repeatedly performing the training task seven times (T1–T7; each consisted of two phrases) without a break, using the learning method assigned to them. The complete note sequence for the training session (7 trials \times 2 phrases \times 8 notes) was given at the beginning of the session, and the played notes were color-coded (green if correctly played and blue otherwise) for the feedback of results. Immediately after the learning, the participants underwent a test (POST) that measured their performance improvements.

After training, the participants temporarily left the experimental site and returned 2.5 h later. The rest time of 2.5 h is considered as a very long intermission compared to the time of the training session (about 4 min). Then a test (RET) was conducted to assess how well participants retained the improvements that they had made even after the recess. The task involved was identical to those of PRE and POST. Finally, the participants completed a questionnaire in which they subjectively rated the learning method assigned to them. These experimental procedures are summarized in Figure 3.5.

During the experiment, all of the participants wore earplugs to exclude ambient sound noise and headphones to listen to the sounds from the drum set. The drum sounds and the auditory cues generated in the KR+AC method were sufficiently loud to be heard despite the earplugs. Only the participants of the KR+TC method wore armbands during the main experiment to receive the vibrotactile cues. The main experiment lasted for about 30 min, and the retention experiment and questionnaire session took about 10 min.

3.1.5 Performance Measures

During the experiment, the time (at a resolution of 1 ms), the target PI of each note, and each response (i.e., striking a PI) were monitored and logged by the control computer. From these data, the error ratio and the task completion time were calculated for each test and training trial. The error ratio was defined as the ratio of incorrect responses to the total number of responses (16). The task completion time was defined as the time from the initiation of a

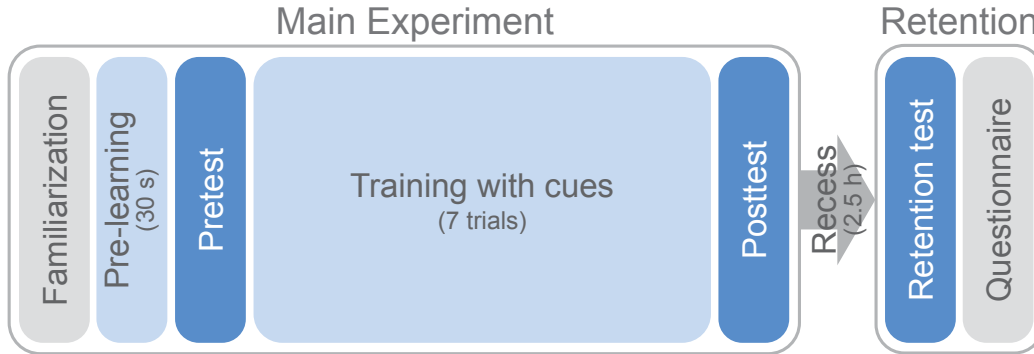


Fig. 3.5 Experimental procedures.

trial to the detected time of the last response. The measurement results of the three tests (PRE, POST, and RET) were compared to evaluate the learning efficacy, and those of the seven training trials (T1–T7) were compared to see the trends during training.

In the questionnaire, the participants subjectively evaluated their learning method by answering the following four questions: Q1. Was the method useful for learning sight reading? Q2. Was it easy to perform the task? Q3. Was the learning method interesting? Q4. Was the method convenient to use? Each question was answered by marking on a horizontal line whose left end represented “*strongly disagree*” and right end represented “*strongly agree*”. The answers were linearly mapped to real numbers from 0.0 (strongly disagree) to 6.0 (strongly agree) for the analysis.

3.2 Results

3.2.1 Familiarization

The KR+AC group required 2–3 phrases (2.3 on average, with a standard deviation of 0.5) to complete the familiarization session, while the KR+TC group required 2–6 phrases (an average of 3.7, with a standard deviation of 1.3). Considering the fact that the last two phrases were actually used for terminating the familiarization session, the KR+AC and KR+TC groups respectively used 0.3 and 1.7 additional phrases, on average, for the familiarization. This result implies that the auditory cues were distinguishable with virtually

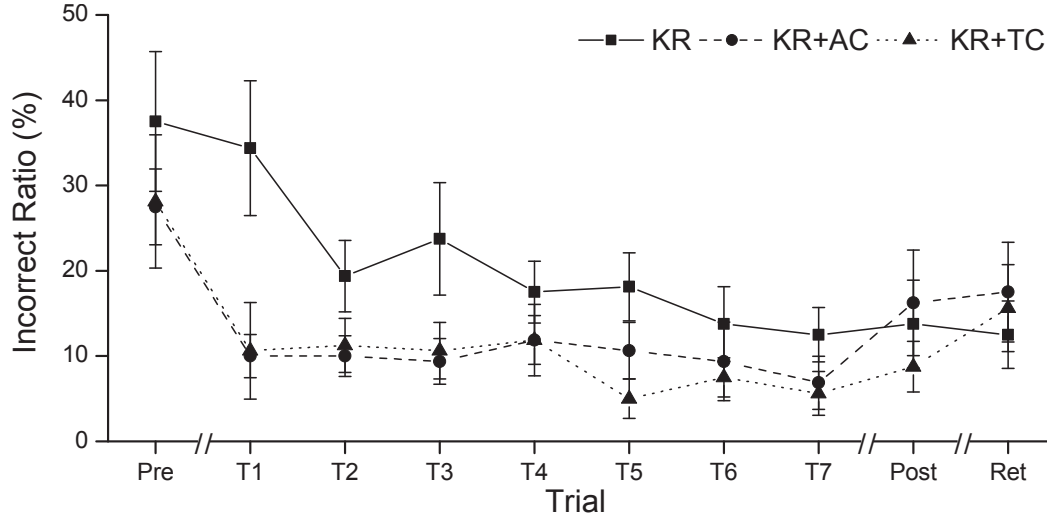


Fig. 3.6 Mean error ratios with standard error bars.

no training, whereas the vibrotactile cues required some training. It is noteworthy that the KR+TC group was also expected to have no difficulty in identifying the vibrotactile cues during the training session since they completely learned the cue-PI pairs during this familiarization session.

3.2.2 Error Ratio

The error ratios of the three participant groups averaged across each test (PRE, POST, and RET) and training trial (T1–T7) are shown in Figure 3.6. Overall, the large initial error ratios (PRE) were considerably reduced after learning (POST), and this improvement was well retained after the recess (RET). The KR+TC group showed the smallest error in the POST test, while the KR group did the same in the RET test. The KR+AC group showed the largest error ratios in both tests. The participants of the KR+AC and KR+TC methods made some errors during training, although they were expected to have virtually no errors. It can be because the process of perceiving a guidance cue and identifying the correct response was hindered by additional cognitive loads for recognizing music symbols and memorizing their relations with the PIs.

A simple main effect analysis (a series of one-way ANOVAs that varies only one factor while the other factors are fixed) was performed for analysis. The differences between the methods were not significant for all the tests and training trials, except significant differences in T1 ($F_{2,27} = 5.15$, $p = 0.0127$) and T5 ($F_{2,27} = 3.63$, $p = 0.0402$). This result suggests that the effects of the guidance cues were strong at the early stages of learning, but the effects were rapidly reduced as the training session continued. Between the trials, statistically significant differences were observed in all participant groups ($F_{9,81} = 5.66$ and $p < 0.0001$ for KR; $F_{9,81} = 3.32$ and $p = 0.0017$ for KR+TC; $F_{9,81} = 3.96$ and $p = 0.0003$ for KR+AC). A Tukey's HSD test was performed for a post-hoc analysis. The errors in the POST and RET tests were significantly smaller than the error in the PRE test with the KR method, and only the error in the POST test was significant with the KR+TC method. The KR+AC method showed no significant difference between the tests; significant differences were only found between the PRE test and all the training trials. This shows that the gains from the learning were well retained with the KR method, but not with the KR+AC method, with an intermediate retention performance with the KR+TC method. In summary, the vibrotactile and auditory cues resulted in no significant improvements in the error ratio, but rather hindered the retention of learning to some degree.

3.2.3 Task Completion Time

The average task completion times for each participant group and trial are presented in Figure 3.7. The task completion time generally showed a similar tendency to that of the error ratio. The task completion time was large before learning (PRE) and gradually decreased during the training session (T1–T7). It was slightly increased immediately after the learning (POST), but the gains from learning were well retained after 2.5 h (RET). In general, the KR+TC group had the largest task completion times. The KR+AC group showed the smallest task completion times during learning, and the KR group did the same in the POST and RET tests.

An ANOVA analysis was performed in the same manner as used for the error ratio. The learning methods had no significance differences for all tests and training trials, whereas the

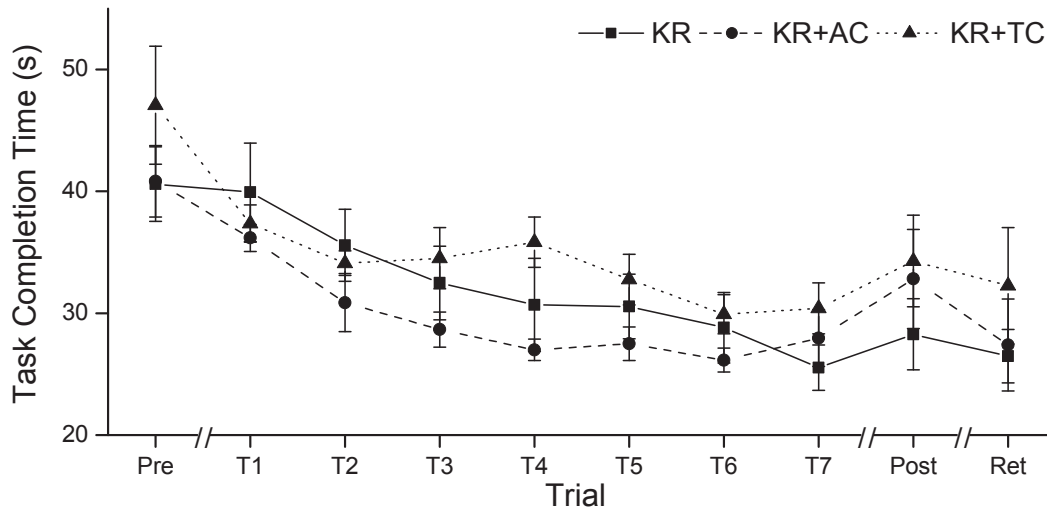


Fig. 3.7 Mean task completion times with standard error bars.

trials showed statistical significance for all groups ($F_{9,81} = 9.76$ and $p < 0.0001$ for KR; $F_{9,81} = 5.36$ and $p = 0.0001$ for KR+TC; $F_{9,81} = 4.08$ and $p = 0.0002$ for KR+AC). In a Tukey's HSD test, it was confirmed that all participant groups had significantly reduced their task completion time after learning (POST and RET) compared to the times before learning (PRE), with an exception of the POST test with the KR+AC method. In summary, the three learning methods resulted in negligible differences in the task completion time, but they all provided considerable learning effects.

3.2.4 Subjective Evaluation

The subjective evaluation results are shown in Figure 3.8. As for the learning methods, the KR method was reported as the most convenient to use, but the least interesting and not easy. The KR+AC method was considered to be the easiest but the least useful. The KR+TC method received the highest score for usefulness, while it had the lowest score for convenience. In one-way ANOVAs that used the learning methods as a main factor, we could not find statistically significant differences for all qualitative measures ($F_{2,27} = 1.70$ and $p = 0.2023$ for usefulness; $F_{2,27} = 1.12$ and $p = 0.3406$ for easiness of the task;

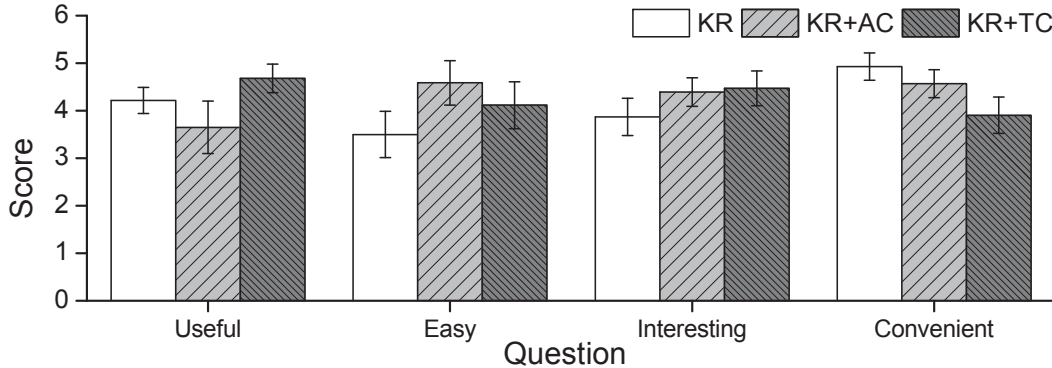


Fig. 3.8 Mean subjective evaluation scores with standard error bars.

$F_{2,27} = 0.85$ and $p = 0.4390$ for interestingness; $F_{2,27} = 2.81$ and $p = 0.0781$ for convenience).

3.3 Discussion

3.3.1 Familiarization

The vibrotactile cues required more effort for the familiarization than the auditory cues despite the fact that both cue sets were essentially the same except for the sensory channel for information transmission. However, considering that people are less familiar with vibration signals than sound signals, especially amplitude-modulated ones, it is an acceptable result that the vibration cues required one or two additional exposures for perfect discrimination.

During the familiarization session, the KR+TC group made 3.1 incorrect responses on an average. Among these errors, there was no stroke in which the participants in the group misunderstood the stimulation location (left or right upper arm). More than half of the errors (1.7 strokes; 0.8 from the left and 0.9 from the right side) were made by incorrectly judging the 8-beat rhythm as the 4-beat rhythm. It appeared that the participants undercounted the number of beats in the 8-beat rhythm, possibly because of the limited capacity for tactile numerosity judgments [32]. Indeed, the participants often reported that they felt only six or seven beats from the 8-beat rhythm. Because they were told that the training system would present 1-, 2-, 4-, and 8-beat rhythms before the familiarization, when they perceived six or

seven beats, they could have judged them as the 4-beat rhythm.

Half of the remaining errors (0.8 strokes; 0.3 from the left and 0.5 from the right side) were due to incorrectly judging the 1-beat rhythm as the 8-beat rhythm. We believe that this may be due to the amplitude-modulated nature of vibrotactile rhythms. The 1-beat rhythm had the slowest amplitude change rate (0.5 Hz) among the rhythms used in the experiment, and some participants could not notice such slow change in the amplitude. This probably caused the participants to attend to the 64 Hz carrier frequency and judge the 1-beat rhythm as the fastest (8-beat) one among the rhythms.

After the familiarization session, all participants of the KR+TC method could correctly distinguish the vibration cues, and they did not show any noticeable tendency in their errors during training in comparison to the other participant groups.

3.3.2 Error Ratio

The three participant groups did not show significant differences in the error ratio for the POST and RET tests, despite the fact that the KR+AC and KR+TC groups made significantly smaller error ratios in the early stages of the training session. Even though the three groups showed similar performances for the tests, they differed in the degree of retention; the KR+AC group showed the lowest retention performance, while the KR group led to the highest. This suggests that the contribution of the KR feedback was dominant in the learning of the sight reading skill, and the effects of the auditory and vibrotactile cues were not salient in comparison to the KR feedback, hindering the retention of the task in some degree.

It is also possible that the effect of cueing was underestimated due to the simple task. As explained earlier, cueing expedites learning by providing knowledge about correct performance. Because of the simplicity of our task, the participants of the KR group could have acquired such knowledge easily without the guidance of cueing, canceling the difference between the methods. In addition, the experimental result was probably caused in part by large individual differences. People often show a wide range of performance differences in motor skill learning, and this also applies to sight reading skills [19]. Large individual dif-

ferences result in large within-group differences, and due to this, between-group differences can be statistically not significant.

As to the retention, it is possible that excessive guidance interfered with the active learning efforts of the participants. The KR group required to use more effort during training because they could only see the correctness of their responses, whereas the KR+AC and KR+TC groups could know the correct answer before making a response. Many motor learning studies have observed similar results; the gains of learning disappear rapidly after learning when the learner used demonstration-based learning methods [37, 52]. It is generally considered that differences between the training and test tasks and insufficient learning due to the divided attention of participants are the major reasons for that. It is also known that frequent guidance during learning often degrades the retention of a skill [59]. In our experiment, the KR+AC and KR+TC groups received guidance twice (each for cueing or KR) for each note, whereas the KR group did only once (for KR). Regarding the retention difference between the KR+AC and KR+TC groups, it could be because of the higher distinguishability of the auditory cues. Because the KR+AC group could readily know the correct answer from the auditory cues, they probably performed the training task by relying more on the cues, resulting in insufficient learning.

3.3.3 Task Completion Time

We found no statistical significance in the task completion times of the three participant groups. This result can be explained in a similar manner to the results for the error ratio. Compared with the error ratio, the task completion time showed less variability between the learning methods even in the training sessions. This result indicates that the guidance cues had no statistically significant effects on the task completion time not only for the tests but also for the training trials. This is in contrast with our initial expectation that the guidance cues would reduce the time for learning by allowing the learner readily perform the task. The KR+AC and KR+TC groups favored to use sufficient time to obtain information on the pitch-PI pairs from the cues and to memorize the information.

Though the differences were not statistically significant, the KR+AC group had a ten-

dency to spend less time during the training trials than the other learning methods. It seems that the cue dependency of the KR+AC group, which appears to be caused by the high distinguishability of the cues, led the group to spend less effort and time in memorizing the pitch-PI relations. In contrast, the KR+TC group usually used more time. This seems that the time needed to identify vibrotactile cues was longer than that required for auditory cues. Also, the result was probably influenced in part by the cautiousness of the group, considering its relatively large initial task completion time in PRE.

3.3.4 Subjective Evaluation

We could not find statistically significant differences between the three learning methods from the questionnaire results. It is possible that the subjective perception of a learning method received during learning had weakened after the long recess (about 2.5 h) between the training session and the subjective evaluation session.

The relatively low usefulness and high easiness scores of the KR+AC method seem to reflect the dependency of the participants on the auditory cues and the relatively high discriminability of the cues, respectively. In addition, the KR+AC and KR+TC methods received higher scores for the easiness and interestingness than the KR method. The provision of additional guidance cues may have improved the subjective perception of easiness and interest regardless of the sensory channel used. As to the convenience, the KR+TC method showed the least score, while the KR+AC method showed the second-least score. This is a reasonable result because the two methods required additional effort for the familiarization with the cues, and in particular, the KR+TC method required the participants to wear the arm bands.

3.4 Conclusions

In this study, we introduced a learning system for the sight reading skills of simple drum sequences. The system feedbacks knowledge of results (KR) on the learner's performance by color-coding the music symbols, and it can additionally provide vibrotactile and auditory cues that indicate which action should be performed for a given music symbol. Three learn-

ing methods were experimentally compared to evaluate the effects of the KR feedback and the guidance cues: KR feedback only, KR feedback with auditory cues, and KR feedback with vibrotactile cues. In the experimental results, we found no statistically significant differences between the three learning methods in terms of the task accuracy and completion time. Both auditory and vibrotactile cues were effective in reducing errors during training, but such effect did not result in better retention performances. These results indicate that the KR feedback is more dominant in the learning of sight reading skills, while the role of additional cueing is rather subsidiary. Nevertheless, we do not draw a solid conclusion as we are in the initial stage of our research.

It is likely that our task was much easier than actual sight reading and the efficacy of the guidance cues, which accelerates learning by directly providing the knowledge of how to perform the task, could be underestimated. In actual sight reading, multiple notes in a music piece are processed in parallel in order to meet the temporal requirements (i.e., tempo and rhythm) of music [31], whereas the participants of our study processed the notes one by one. To reflect temporal aspects of music, we need a set of short, but still identifiable, cues. For this, we may need to add additional vibration actuators or make modifications to the cueing patterns. Also, cues will be provided slightly ahead of the desired playing time of a note in order that the learner can process the cues beforehand. Striking strength is another important factor in actual drum music, and multiple levels of cueing intensity will fit to guide this factor. As regards the guidance methods, we may vary the guidance frequency during learning for more active and effective learning [75].

Chapter 4

Vibrotactile Drumming Guidance System

Our vibrotactile guidance system designates striking position by the body site stimulated. The trunk and ankles, which are relatively stationary during drumming, are used to avoid masking between vibrotactile stimuli during active motion [51, 17]. The exact stimulation positions are selected in such a way that they preserve the egocentric orientations from the body sites to the PIs. Striking strength is mapped to the stimulus strength and duration using redundant coding. All vibrotactile stimuli are sufficiently short (<0.2 s) for the beginner-level playing speed.

4.1 Hardware

Our haptic drumming guidance system is shown in Fig. 4.1a. The key component is an electric drum set (Model DD506; Medeli Electronics, Hong Kong). If a player strikes a PI in the drum set, the PI hit and the strength of that stroke are measured and sent to a computer that renders visual and haptic stimuli. Visual scenes are displayed on a 24-inch LCD monitor. Haptic guidance is provided by a custom-made vibrotactile vest and ankle bands.

The vibrotactile vest and ankle bands are made of elastic rubber bands to which bar-type

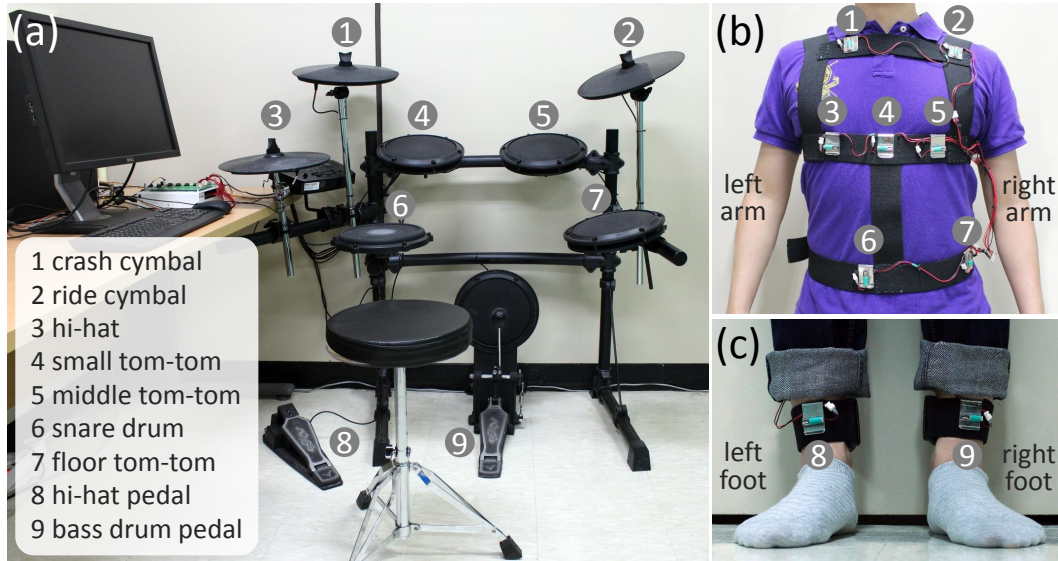


Fig. 4.1 Hardware for vibrotactile drumming guidance. (a) Electronic drum set with nine percussion instruments (PIs). (b) and (c) Vibrotactile vest and ankle bands (mirror images), respectively. Relationships between PIs and body sites are denoted by numbers.

vibration motors ($\phi 7.0 \times L 25.0$ mm, 5 g; Sejoo Electronics, Korea) are fastened. Seven tactors are attached to the vest using metal clips ($W 25 \times H 50$ mm), and one tactor is attached to each ankle band in the same way. The use of clips allows us to adjust tactor positions to individual learners while maintaining stable contacts between the tactors and the learner's body. The placement of the tactors is shown in Fig. 4.1b and 4.1c.

4.2 Guidance Method

Haptic guidance delivers the three main elements (target PI, strength, and timing) of a drum strike by a vibrotactile cue. The target PI of the strike is designated by a stimulated body site. For this, each PI of the drum set is mapped to the body site that is near to the PI and also relatively stationary during drumming. This mapping for the vest is illustrated in Fig. 4.1a and 4.1b. This design preserves the egocentric orientation in the transverse plane from each body site to the corresponding PI, while reflecting correspondence in their

relative heights. The mapping for the two ankle bands is also depicted in Fig. 4.1a and 4.1c. Vibrotactile cues stimulate the distal frontal part of each shin to prevent hindrance during pedaling while matching the egocentric orientations between the stimulation sites and the target PIs. The exact stimulation locations were determined by a series of pilot tests so that absolute identification of individual vibrotactile cues could have nearly perfect accuracy.

We also transmit two levels of PI striking strength (normal and accented) using two vibration strengths. Since vibrotactile magnitude perception depends on body site, input voltage to each tactor has been adjusted so that the vibrations of the same strength level are perceived to be of the same (or at least very similar) intensities at all the nine body sites. Tactors at the epigastrium (no. 4 in Fig. 4.1b) or umbilical region (no. 6) use higher input, those at the upper thorax (no. 1 and 2) or right lumbar region (no. 7) use lower input, and the other tactors (no. 3, 5, 8, and 9) use medium input. Their input ranges are 1.2–1.8 V for normal vibrations and 2.8–3.5 V for accented ones.

Striking timing is represented by the stimulation timing of vibrotactile cues. To guide the timing precisely while preventing overlaps between consecutive cues, short but clearly perceptible vibrotactile stimuli are required. We use 100-ms long vibration signals for normal cues and 150-ms signals for accented cues, which result in actual vibration durations of about 94 ms and 199 ms (threshold 1 G), respectively. Accented cues have a longer duration for better recognition of strength level (redundant coding).

4.3 Vibrotactile Guidance System Improvements

Our guidance design and initial vibrotactile guidance system was presented in [35]. The key component is an electric drum set (Model DD506; Medeli Electronics, Hong Kong) which consisted of nine PIs. The drum set measures the PI, strength, and time of every drum strike made by a learner and provided percussion sounds accordingly. It also sends the measurement data to a computer that renders visual and haptic stimuli. Visual scenes are displayed on A 24-inch LCD monitor. Haptic guidance is provided by nine bar-type vibration motors ($\phi 7.0 \times L 25.0$ mm, 5 g; Sejoo Electronics, Korea) that are attached to the learner using elastic rubber belts and clips.

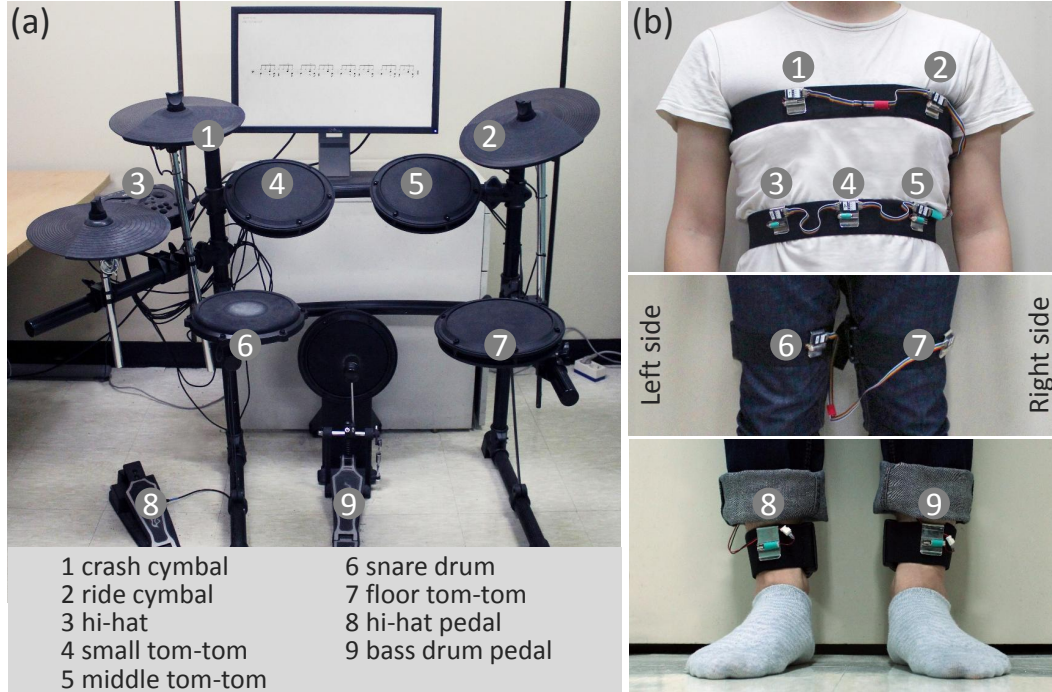


Fig. 4.2 (a) Hardware for vibrotactile drumming guidance. (b) Modified layout of vibration motors (mirror images). Relationships between PIs and body sites are denoted by numbers.

For guidance, the system delivers the target PI, strength, and timing of a drum strike by a vibrotactile cue. Target PI is designated by the body site of stimulation. To this end, for each PI of the drum set, a vibration motor was placed on the body site that is near to the PI and also relatively stationary during drumming. The exact locations of the motors were selected in such a way that the egocentric orientations in the transverse plane are well matched between the motors and the corresponding PIs, while preserving the correspondence in their relative heights (see Fig. 4.2(b)). As to striking strength, our system can guide two levels of striking strength (normal and accented) using two vibration strengths. Since vibrotactile magnitude perception depends on body site, input voltage to each motor was adjusted so that the vibrations of the same strength level were perceived to be of the same (or at least very similar) intensities at all the nine body sites. In addition to higher vibration intensities, accented cues has longer vibration durations (roughly 199 ms) than normal ones (94 ms) for

better recognition of strength level. The striking timing is represented by the stimulation timing. Drumming sometimes requires multiple simultaneous drum strikes, and in this case, we provide all the guidance cues for the strikes at the same time.

4.3.1 System Improvements

A systematic evaluation of our guidance system revealed that the system was less suitable to guide two simultaneous drum strikes, though it was highly successful in guiding a single strike [35]. Spatial masking between vibrotactile cues was thought to be one of the main reasons of this discrepancy. To improve upon this problem, we have rearranged the vibration motors on the trunk to increase the distances among them (Fig. 4.2(b)). First, the motor next to the navel (no. 6 in the figure) was moved to the left vastus medialis muscle (the distal anterior medial part of the thigh), and the motor just above the iliac bone (no. 7) was relocated to the right vastus lateralis muscle (the distal anterior lateral part of the thigh). Then, the motors under the chest (no. 3, 4, and 5) were moved down to the level of the navel so that they can be more apart from the motors on the upper chest (no. 1 and 2). We have also lowered the positions of the motors on the upper chest to some extent for a more stable attachment to the body while drumming.

Limited memory and attention capacities were considered the other reasons of the lower simultaneous guidance performance. People can attend to only a few items at a time, and without a conscious effort to retain the impression of a stimulus, it quickly decays and disappears from the memory [1]. Due to this, when a pair of vibrotactile cues are given simultaneously, one of the cues may fade out while processing another cue, resulting in an incomplete delivery of guidance. Normal cues, which transmit weaker vibration stimuli for a shorter time to instruct normal strength strikes, were thought to be less impressive and consequently easier to be forgotten. One simple solution is to assign longer and more intense vibration stimuli to the cues, so that their impressions can be strong sufficiently. However, it is not applicable to our case because the normal cues were required to be distinguishable from the accented cues by their short duration and low intensities.

When provided with two vibrotactile cues, it is obviously a better choice to process the

more forgettable cue first and then proceed to the other to avoid forgetting. In this regard, we have modified our guidance method to present normal cues slightly before accented cues, with an assumption that the processing order is determined by the order of perception. For this, we first adjusted the cue presentation timing to each vibrotactile cue so that a pair of cues can be perceived at the same (or at least, at a very similar) time. The perception time of a vibrotactile cue depends on the body site of stimulation [35]. Such difference was compensated by introducing a delay to each body site, from 0 ms (ankles) to 30 ms (upper chest). Considering the faster vibration output for higher voltage input, the accented cues were given a delay of 20 ms while the normal cues received no delay. The rendering of vibrotactile cues was processed 50 ms earlier than visual rendering to maintain the synchrony between the modalities. Then, the stimulus-onset asynchrony (SOA; 30 ms) between a pair of vibrotactile cues was determined by a pilot experiment that compared the recognition performances of cue pairs with different SOAs.

The above method is also expected to be beneficial to the case of both normal cues. Because the cue transmitted later can remain longer in the memory (in terms of the time from the initiation of the former cue) as much time as the SOA, there is higher chance of processing the cue before forgetting. For this case, the same SOA was used, and the order of presentation was decided by another pilot experiment. The cue to the lower body site (to the right body site in case of same height) had a higher priority.

Chapter 5

Experiment I: Identification of a Cue

The long-term goal of our research is to develop methods that can help novice drum players learn playing of drumming sequences by providing guidance on the target, strength, and timing of PI strikes using vibrotactile cues. To this end, the first priority is with ensuring that learners correctly recognize the information embedded in each vibrotactile cue. The time required for recognition is also important since it determines the extent of drumming speed to which vibrotactile guidance is effective. Hence, our first evaluation was concerned with the accuracy and speed of information transmission of our guidance design.

In pilot experiments, we found that striking a PI with high positional accuracy while controlling its strength is difficult for novice participants even after some hours of practice. This means that using the actual drum set to collect participants' responses is subject to a large amount of motor errors, thereby preventing us from looking into the true information transmission performance of our design. Therefore, we needed to use the most reliable means for response collection, i.e., the mouse that was the most familiar interface to participants. It was assumed that response interface has a negligible effect to the perceptual accuracy and speed of cue response.

Responding to a guidance cue can be regarded as a choice reaction task (CRT) in which participants need to give a response in accordance with a randomly given stimulus [11].

To perform a CRT, participants go through four processing stages: *detection*, *recognition*, *choice*, and *execution*. When a stimulus is presented, the participants first detect the stimulus, and recognize the stimulus based on its properties such as location and strength. Then the participants make a choice of what to do, and finally execute a response. Among these stages, the first three are sensory-cognitive processes that determine the perceptual performance of cue response. Understanding their respective effects is a cornerstone for optimal vibrotactile guidance design.

The processing stages of CRT often roll back (e.g., recognizing the stimulus again for the assurance of a choice) and overlap each other (e.g., making a choice while moving). Thus, it is almost impossible to measure their effects independently. To estimate the respective effects of the cognitive processes on reaction time, Donders compared three reaction tasks with different cognitive requirements [11]. The three tasks were made to involve the processing stages of CRT one after another, starting from stimulus detection to stimulus recognition, then response choice. The effect of a processing stage was calculated from the difference in reaction time between two adjacent tasks, assuming that the addition of a new stage did not affect the existing stages and the difference was entirely originated from the new stage.

Based on Donders' method, we devised three reaction tasks to evaluate the perceptual performance of our guidance design: (1) detection of vibrotactile cues, (2) recognition of the vibrotactile cues, and (3) selection of PIs in a drum set according to the vibrotactile cues. By comparing the performance in the three tasks, we estimated the effects of the four processing stages in terms of accuracy and time.

For the evaluation of our guidance design, we first examine the simplest scenario in which a single vibrotactile cue is presented and the participant responds to the cue in this chapter, and then proceed to a more complex scenario in which participants perceive and respond to two vibrotactile cues at a time in Chapter 6.

5.1 Methods

5.1.1 Participants

We recruited 12 male university students (aged 19–28 years; mean 21.0) for the experiment. They reported that they had no known sensorimotor disorders, had no experience of playing drum sets, and were naive to this kind of experiments. The participants were paid 10,000 KRW (\simeq 9.35 USD) for their participation.

5.1.2 Three Cue-Response Tasks

The experiment consisted of three reaction tasks named DE, DRE, and DRCE. For the three tasks, the participants were asked to perceive a vibrotactile cue presented by one of the factors in the vest or ankle bands and then enter its perceived location and strength to the computer using a mouse. On each trial, nine targets were displayed on the screen as gray circles (outer diameter 10 mm, inner diameter 5 mm; Fig. 5.1). The targets had a one-to-one correspondence to the body sites for stimulation (and also to the PIs of a drum set). The target positions were consistent with the stimulated body locations. The positions of the target circles were the same in all the three tasks, ensuring the movement required for target selection remained identical. The participants were instructed to indicate the location of each vibrotactile cue by selecting the corresponding circle and its strength by pressing a left button on the mouse for normal cues and a right button for accented cues.

The three experimental tasks differed in the cognitive processing stages involved for a systematic assessment of our vibrotactile guidance design. This was done by providing different levels of visual information as described below.

Task DE (Detection and Execution) was to measure the performance for vibrotactile cue detection and subsequent response execution. The location and strength of the correct answer was provided visually before a vibrotactile cue was presented. For this, the target circle was filled with a red inner circle with different diameters (small for normal cues and large for accented cues; Fig. 5.1a). The participant was asked to wait and stay still and to enter a response immediately after perceiving the cue. Then, a vibrotactile cue was

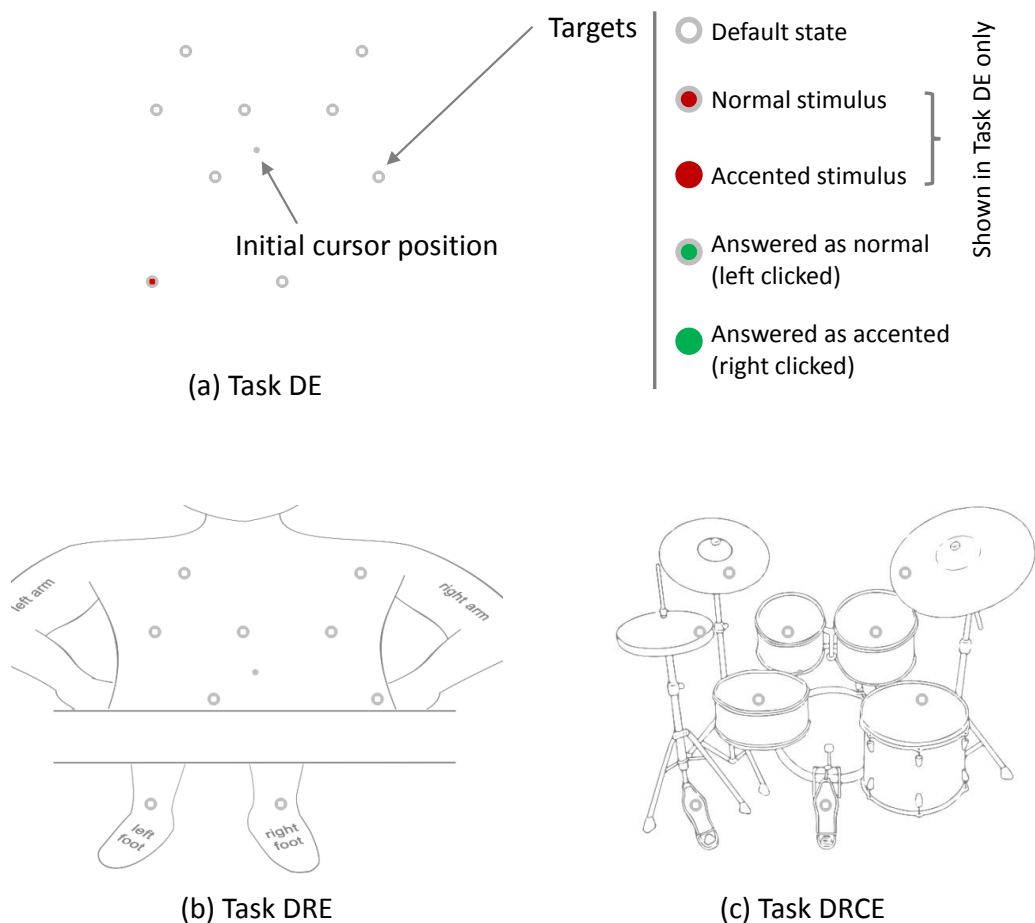


Fig. 5.1: Visual scenes provided in each experiment task.

provided randomly after 1–3 s. The random waiting time was to prevent the participant from initiating response actions without perceiving cues. Any cursor movement (threshold 2.5 mm) in the wait period made the participant wait for another random 1–3 s to receive the cue. The visual guidance lasted until the participant entered the response. In this task, the necessity for cue recognition and response choice is removed or at least minimized.

Task DRE (Detection, Recognition, and Execution) did not provide the visual guidance

of Task DE. Instead, a mirrored drawing of a human body was displayed on the background (Fig. 5.1b). The participants had to identify the location and strength of each vibrotactile cue. The mirror image provides a reference as to the associative mapping between the body sites and the target locations, making involvement of the choice stage unnecessary or minimal. No references for vibration strength were given because the participants learned very quickly the cue-response mapping for strength that used the mouse buttons.

Task DRCE (Detection, Recognition, Choice, and Execution) was designed to involve all the four processing stages. This task is the same as Task DRE, except that the background mirror image of a human body was replaced with a drawing of a drum set (Fig. 5.1c). Each PI of the drum set included a target circle associated with the body site of vibrotactile stimulation. This spatial relationship was informed to the participants before the experiment, and they were instructed to select the corresponding PI for a given vibrotactile cue. The latter requires understanding the meanings of the cue and making a decision of which PI to select with which mouse button.

5.1.3 Procedures

The main experiment consisted of three sessions for each of the three experimental tasks. Each session had 180 trials (9 locations \times 2 strengths \times 10 repetitions). The session order was fully balanced across the participants, and the order of trials was randomized for each session and each participant.

Prior to the experiment, the participant was informed of the experimental task and procedures, and then signed on a written consent form. Then the participant wore the vibrotactile vest and ankle bands and went through a short training session to become accustomed to the system. The participant also wore earplugs to mask ambient noise and the sound produced by the factors.

On each trial, the mouse cursor was initially positioned at the center of the screen (a small gray point in Fig. 5.1). After the participant selected a target following the procedure described in Section 5.1.2, the target turned green for 500 ms for confirmation. Then the trial ended, and the mouse cursor was returned to its initial state. For Task DE, visual

guidance for the next trial was given right after the end of the trial.

To avoid fatigue, the participant was required to have a break for 5 min between the experimental sessions and also could take a rest whenever necessary. The experimental procedures took about 1 hr.

5.1.4 Performance Measures

To respond to a vibrotactile cue, the participant went through some or all of the four processing stages depending on the experimental task. Our design of the three experimental tasks enabled to estimate the respective effects of the four stages by comparing the measured data between the tasks, upon an assumption that the operation of each stage was unaffected by the inclusion or omission of other stages.

In each trial, we collected the response time t and the response correctness index c . t was the time difference from vibrotactile cue initiation to the participant's response made using the mouse buttons. c was 1 if the participant's answer was correct in both stimulus location and intensity, or 0 otherwise. t and c for Task DRE and DRCE are denoted by t^{DRE} and t^{DRCE} and by c^{DRE} and c^{DRCE} , respectively.

Special care is needed for Task DE. This task involved two mental processing stages, cue detection and response execution. These stages could be hardly processed in parallel because the random waiting time made it uncertain when to start response unless perceived a cue. This allowed us to estimate the individual performance of cue detection and response execution. For response time, the time t^{D} for cue detection was measured as the time difference from cue initiation to mouse cursor movement detection (threshold 2.5 mm). The time t^{E} for response execution was the time difference from cursor movement detection to response detection. Regarding response correctness, it is noted that the participant could still make the correct response owing to visual guidance even when he missed to perceive a vibrotactile cue. To account for this, we introduced the cue detection correctness index c^{D} : c^{D} was 1 if t^{D} was less than a time threshold of 1.0 s, or 0 otherwise. This is based on an observation that the participant would have waited to perceive a cue for a sufficient time to be sure of missing the cue. We also denote the response correctness index measured in

Task DE by c^E as it represents only the errors in response execution.

The above measurement data were used to compute two performance measures, the response error rate e and the processing time τ , for each processing stage. The error rates were computed as follows:

$$\begin{aligned}
 e^D &= 1 - \bar{c}^D, \\
 e^E &= 1 - \bar{c}^E, \\
 e^R &= 1 - \bar{c}^{\text{DRE}} - e^D - e^E, \\
 e^C &= 1 - \bar{c}^{\text{DRCE}} - e^D - e^R - e^E,
 \end{aligned} \tag{5.1}$$

where e^x is the error rate for a processing stage $x \in \{\text{D: Detection, R: Recognition, C: Choice, E: Execution}\}$ and \bar{c}^x is the mean of c^x .

The mental processing times τ^x were estimated as follows. First, in Task DE, missing to perceive a vibrotactile cue results in an extraneously long t^D . Such measurements were rare owing to strong vibrotactile cues and were excluded (threshold 1.0 s) for computing τ^D by averging:

$$\tau^D = \bar{t}^D \tag{5.2}$$

Second, incorrect responses also have effects on t^E , t^{DRE} and t^{DRCE} due to the difference between the desired and actual responses in the cursor travel distance or the pressed mouse button. In Task DE that provided visual guidance on the correct responses, we observed only few incorrect responses ($c^E = 0$) and so simply removed those measurements of t^E in estimating τ^E :

$$\tau^E = \bar{t}^E \tag{5.3}$$

Third, we were not able to simply exclude the trials with response errors in determining τ^R and τ^E since much more frequent response errors were expected in Task DRE and DRCE. Instead, we compensated t^{DRE} and t^{DRCE} at each trial using the estimated time for response

execution (τ^E):

$$\begin{aligned} t'^{\text{DRE}} &= t^{\text{DRE}} + (\tau_t^E - \tau_a^E), \\ t'^{\text{DRCE}} &= t^{\text{DRCE}} + (\tau_t^E - \tau_a^E), \end{aligned} \quad (5.4)$$

where t'^y is the adjusted response time for a trial of task y , and τ_t^E and τ_a^E are the processing times for the execution stage for the target and the actual responses at the trial, respectively. Then, τ^R and τ^D were determined as follows:

$$\begin{aligned} \tau^R &= \bar{t}'^{\text{DRE}} - \tau^D - \tau^E, \\ \tau^C &= \bar{t}'^{\text{DRCE}} - \tau^D - \tau^R - \tau^E. \end{aligned} \quad (5.5)$$

5.2 Results

The mean error rates e and the processing time τ of the participants for each processing stage are shown in Fig. 5.2 and 5.3, respectively.

5.2.1 Effects of Processing Stage

In Experiment I, the participants missed a total 5.00% of trials. This level of accuracy is sufficient for our purpose, particularly considering the large number (18) of vibrotactile cues and the minimal pre-training given to the participants. The error rate was the lowest in the detection stage ($\tau^D = 0.19\%$), followed by $\tau^C = 0.46\%$ in the choice stage and $\tau^E = 0.65\%$ in the execution stage. The error rate was the highest for the recognition stage with $\tau^R = 3.70\%$, but it was still in an acceptable level. The four error rates were further compared using Tukey's HSD tests, and results are shown in Fig. 5.2 by connecting each pair of the two stages that had a statistically significant difference ($\alpha = 0.05$) with a line. The recognition stage had the significantly larger error rate than the others.

The participants spent a total 0.91 s for detection, recognition, and choice. This implies that, unless a learner can process multiple cues in parallel, at least 0.91 s of time should be provided to the learner to detect and identify a vibrotactile cue and to make a corresponding decision. The choice stage showed the fastest processing time ($\tau^C = 0.15$ s), followed by

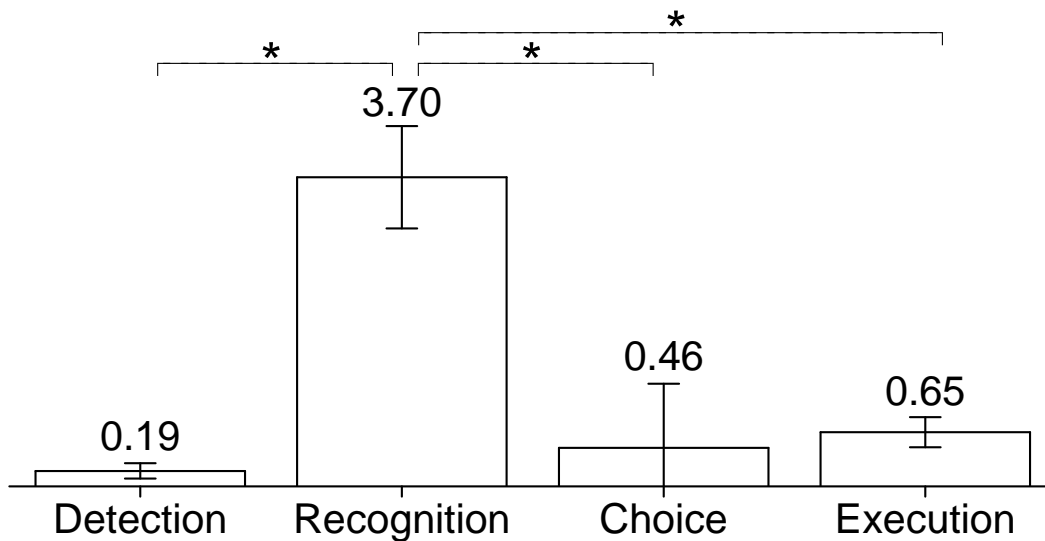


Fig. 5.2 Mean error rates (%) for four processing stages. Error bars represent standard errors.

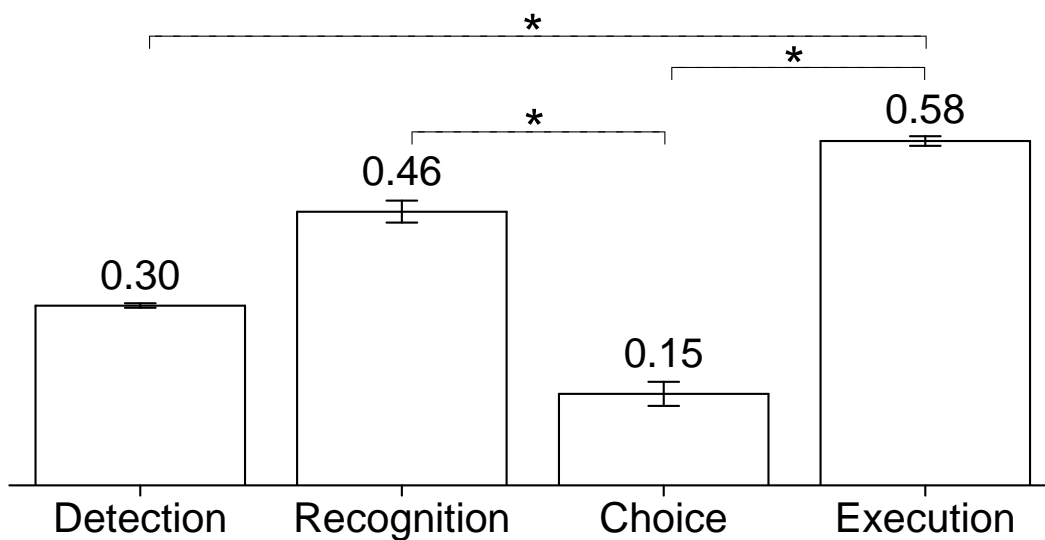


Fig. 5.3 Mean processing times (s) for four processing stages. Error bars represent standard errors.

Table 5.1: Two-way ANOVA results on the main effects of cue position and cue strength on two performance measures (e : error rate, τ : processing time) measured in four processing stages (D: detection, R: recognition, C: choice, E: execution).

Effects of Cue Position								
Statistic	e^D	e^R	e^C	e^E	τ^D	τ^R	τ^C	τ^E
$F_{8,88}$	1.19	1.13	0.87	0.86	8.01	8.64	1.76	34.01
p	0.316	0.355	0.544	0.551	* <0.001	* <0.001	0.096	* <0.001

Effects of Cue Strength								
Statistic	e^D	e^R	e^C	e^E	τ^D	τ^R	τ^C	τ^E
$F_{1,11}$	1.00	0.83	0.35	5.58	0.93	0.42	0.10	0.18
p	0.339	0.383	0.567	*0.038	0.356	0.530	0.757	0.683

the detection stage ($\tau^D = 0.30$ s) and the recognition stage ($\tau^R = 0.46$ s). The execution stage required the longest time ($\tau^E = 0.58$ s). According to Tukey's HSD test, the recognition stage had a significantly longer processing time than the choice stage. The processing time for the detection stage was not statistically different from that for the recognition stage or that for the choice stage.

5.2.2 Effects of Cue Position and Strength

We conducted two-way ANOVAs to assess the effects of cue position and strength on the two performance measures for each processing stage. Results are summarized in Table 5.1. For each statistically significant case, we performed Tukey's HSD tests for multiple comparisons, and results are shown in Fig. 5.4.

Cue position inflicted no significant differences to e for all the processing stages, indicating that the response accuracy was independent of the stimulated body site. As to τ , cue position had a significant effect on τ^D . The longest values of τ^D were measured on the ankles, while the shortest on the upper thorax, with the greatest mean difference of 0.05 s. Cue position was also significant for τ^R . The longest recognition times were measured

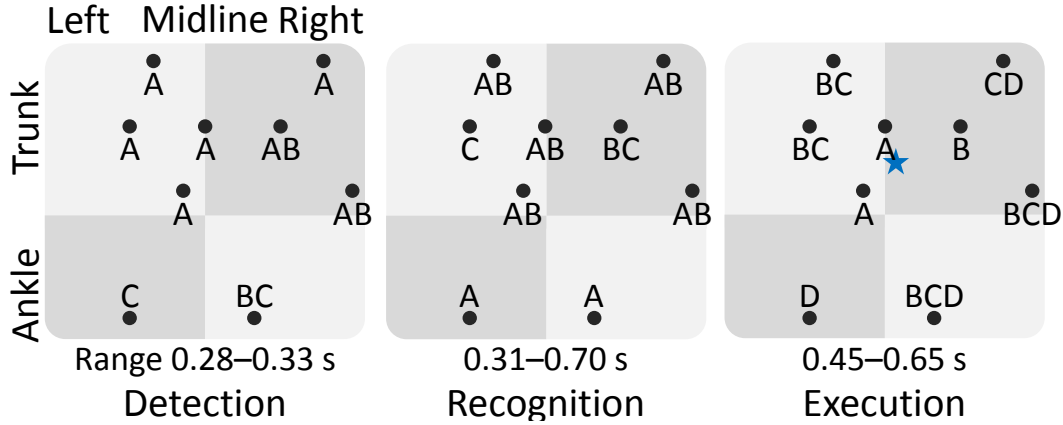


Fig. 5.4 Tukey's HSD test results on the processing time (s) of each processing stage for different body sites. Black dots represent stimulated body sites, and those with the same alphabet are of the same performance group by the test.

on the hypochondriac region, while the shortest were on the ankles, with the largest mean difference of 0.39 s. The effect of cue position was not significant for τ^C , suggesting that the choice process was relatively independent of body site. Lastly, cue position had a significant effect on τ^E as expected, reflecting the different distances from the initial mouse cursor position (denoted by a star in Fig. 5.4) to the target circles.

Cue strength caused significant differences in only e^E , but this result is not robust because of the extremely small number of misses (0.65%) in the execution stage. The effect of cue strength on e was not significant for the other processing stages. Cue strength had no significant influences on τ , though the accented cues resulted in slightly shorter response times in all the processing stages.

5.3 Discussion

5.3.1 Effects of Processing Stage

Most response errors were observed in the recognition stage, albeit the small rate (3.70%). Adjusting the positions and strength levels of the vibrotactile cues may further improve the recognition performance. In the detection and the choice stage, we encountered only

few errors. This indicates respectively that all the vibrotactile cues were strong enough to be perceived and that our mapping from the location and strength of a vibrotactile cue to the target PI and striking strength was highly intuitive. It should be addressed that the response accuracy in the execution stage done with the mouse interface is likely to have been overestimated. In actual drumming guidance, learners would produce more errors due to higher cognitive load and less precise motor control caused by the learners' unfamiliarity to drumming.

Excluding the processing time for the execution stage, which was determined by the response input method, the recognition process required the longest processing time (0.46 s on average). In contrast, the choice stage took the least time with a very small value (0.15 s), confirming the intuitiveness of our guidance design. It is noted that the processing times for the detection stage could have been overestimated to some extent. The measured detection times included a time delay of the tactor to output a perceptible vibration and the time from the motor command to the actual movement for the response (part of response execution rather than cue detection).

5.3.2 Effects of Cue Position and Strength

Since the tactors attached around the waist were more dense, they were expected to have a higher chance for incorrect position perception. However, Experiment I showed similar response accuracies regardless of body site. It seems that the participants spent more time for the cues difficult to identify to improve the response accuracy. The time for cue recognition was the longest around the waist, while it was the shortest on the ankles.

The detection time of vibrotactile cues was generally increased with the distance from the stimulated site to the central nervous system, suggesting that the neural transmission distance was a significant factor for the detection time. The neural transmission speed of tactile sense is about 34 m/s [47]. Therefore, the ankle takes approximately 0.04 s more time to transmit a tactile stimulus if the neural path from the ankle to the brain is assumed to be 1.5 m longer than that from the upper chest to the brain. This value is comparable to our result of 0.05 s difference in the detection time between the ankle and the upper chest.

However, the differences in the detection time were of little practical significance because of their very small values compared to the total processing time ($\tau^{\text{DRCE}} = 1.49$ s).

Providing guidance cues to the ankles resulted in the fastest recognition. The ankles were far from the other body sites, which must have enabled easier identification. More time was required to recognize the cues given on the trunk, particularly on the hypochondriac region, due to the relatively dense positioning of the tactors. This seems to also be associated with the fact that a vibrotactile stimulus is better localized when presented to an anatomical landmark (anchor point) such as the wrist, elbow, or navel [9, 8]. In our guidance design, most tactors were attached on or near to such structures (see Fig. 4.1; Tactor no. 1 on the left collarbone, 2 on the right shoulder, 6 next to the navel, 7 just above the right iliac crest, and 8 and 9 above the ankles). However, the tactors (3 and 5) on the hypochondriac region are relatively far from the body landmarks, resulting in longer localization time. The tactor (4) attached on the epigastrium is also not close from the landmarks, but it showed better performance than those on the hypochondriac region. This result is probably due to the directional sensitivity in vibrotactile stimulus localization. It is known that localization accuracy is higher for the body midline than the other sites around the waist [8, 69]. In our experiment, the participants tended to spend more time for the cues difficult to identify to avoid response errors, resulting in significant differences in recognition time instead of error rate.

Chapter 6

Experiment II: Identification of Paired Cues

Drum rhythms often require the player to strike multiple PIs simultaneously, e.g., stroking a hi-hat while kicking a bass drum. This led us to examine a more complex scenario in which participants perceive and respond to two vibrotactile cues at a time.

6.1 Methods

6.1.1 Participants

Twelve healthy male university students (aged 18–24 years; mean 21.4) were recruited for the experiment. They reported no prior experiences of playing drum sets and participating in this kind of experiments, including Experiment I. Before the experiment, all participants were informed of the experiment, and then signed on a written consent form. They were paid 70,000 KRW (\simeq 65.48 USD) after the experiment.

6.1.2 Experimental Tasks and Vibrotactile Cues

The participants performed three cue-response tasks (Task DE, DRE, DRCE). The tasks were identical to those in Experiment I, except that two vibrotactile cues were presented simultaneously to different body sites and the participant responded to both cues. In drum-

ming learning, two coincident cues must be responded to at once because they represent concurrent multiple drum strikes. In the experiment, however, the participants answered the cues one by one, since the mouse interface did not allow them to select different target circles at the same time.

Vibrotactile cues were generated using the same hardware used in Experiment I, with a minor modification in the vibration profiles. We slightly increased the duration of accented cues (165 ms) for better discrimination between cue strength levels.

6.1.3 Procedures

Our guidance design provided two vibrotactile cues to nine body sites with two strength levels, and thus had a total of 144 ($= {}_9C_2 \times 2^2$) possible combinations. To avoid participants' fatigue that can be caused by the large number of experimental conditions, we divided the experiment into six identical blocks and required the participants to complete them in six consecutive days.

Before the experiment, each participant received verbal instructions about the experimental procedures and tasks, and then wore the vibrotactile vest and ankle bands. The participant also wore earplugs and noise-canceling headphones to remove any sound effect. On each day, the participant completed three experimental sessions, each of which tested all the possible 144 cue pairs under one of the three task conditions (DE, DRE, DRCE). The session order was fully balanced across the participants, and each participant followed the same session order throughout the experiment. The order of trials was randomized for each session, day, and participant.

On each trial, the mouse cursor was located at the center of the screen, and nine circles were displayed on the screen along with the visual guidance assigned to each task (see Section ??). Randomly after 1–3 s, two vibrotactile cues with respective strengths were presented simultaneously to the participant at different body sites, and the participant first answered one of the cues using the mouse. Immediately after this first response, the cursor automatically returned to its initial position. This was to allow participants to have almost the same cursor trajectory regardless of cue response order. Then, the participant proceeded

to answer the remaining cue. After the participant entered the second answer, the current trial was terminated and the next trial was initiated.

It took about 1 hr to complete three sessions on each day, including 5 min breaks between sessions. The same procedure was repeated for six consecutive days, resulting in a total experiment time of 6 hrs. For data analysis, we used the data measured in only the last five days, regarding the first day as training.

The effects of the simultaneous presentation of two vibrotactile cues can be understood by comparing the results of Experiment I and II. However, the two experiments had differences in vibrotactile cue design, training time, and participant group. For more precise comparisons, we conducted Experiment I again on the last day after completing Experiment II, with the same participants and vibrotactile cues as Experiment II. We denote the previous Experiment I by I-1, and this one by I-2.

6.1.4 Performance Measures

In Experiment II, each participant repeated five trials for each pair of vibrotactile cues (excluding one training trial on the first day). In each trial, the participant answered the positions and strengths of two concurrent vibrotactile cues, and the response time (t) and the response correctness index (c) were recorded.

Similar to Experiment I, t was defined as the time from the initiation of cue generation to the detection of the participant's second target selection. For Task DE, t was divided into the cue detection time t^D and the response execution time t^E using the detection time of cursor departure to the first target (threshold 2.5 mm). t^{DRE} and t^{DRCE} were the response times measured in Task DRE and DRCE, respectively. c was 1 if the participant correctly responded to both cues, otherwise it was 0. The correct cue detection index c^D was 1 if $t^D < 1.0$ s, or 0 if not. c^E , c^{DRE} , and c^{DRCE} were recorded in Task DE, DRE, DRCE, respectively. These measurement data were fed into (5.1)–(5.5) to compute the error rate e and the processing time τ for each processing stage.

The performance measures of Experiment I-2 were obtained following the same procedures of Experiment I-1 (Section 5.1.4).

Table 6.1: T-test results that compared two experiments (I-1 and I-2; I-2 and II) using two performance measures (e : error rate, τ : processing time) for four processing stages (D: detection, R: recognition, C: choice, E: execution).

Experiment I-1 vs. Experiment I-2								
Stat.	e^D	e^R	e^C	e^E	τ^D	τ^R	τ^C	τ^E
t	0.82	0.82	0.29	2.90	-6.48	9.84	5.18	2.54
p	0.412	0.410	0.774	*0.004	*<0.001	*<0.001	*<0.001	*0.012

Experiment I-2 vs. Experiment II								
Stat.	e^D	e^R	e^C	e^E	τ^D	τ^R	τ^C	τ^E
t	2.66	43.13	0.55	3.89	11.63	29.80	2.22	92.35
p	*0.008	*<0.001	0.581	*<0.001	*<0.001	*<0.001	*0.027	*<0.001

6.2 Results

The mean error rates e and processing times τ for the four processing stages measured in Experiment II are shown in Fig. 6.1 and 6.2, respectively, together with those of Experiment I-1 and I-2.

6.2.1 Comparison between Experiment I-1 and I-2

Table 6.1 shows the t -test results that compared the results of Experiment I-1 and I-2 for each combination of the performance measures and the processing stages. Overall, Experiment I-2 showed shorter τ compared Experiment I-1, with comparable e (also see Fig. 6.1 and 6.2). Experiment I-2 provided slightly improved guidance cues and prolonged training (through Experiment II) to participants, and this may account for the improvements in processing time. No significant improvement in accuracy is probably due to ceiling effects (the values of e were already very low in Experiment I-1).

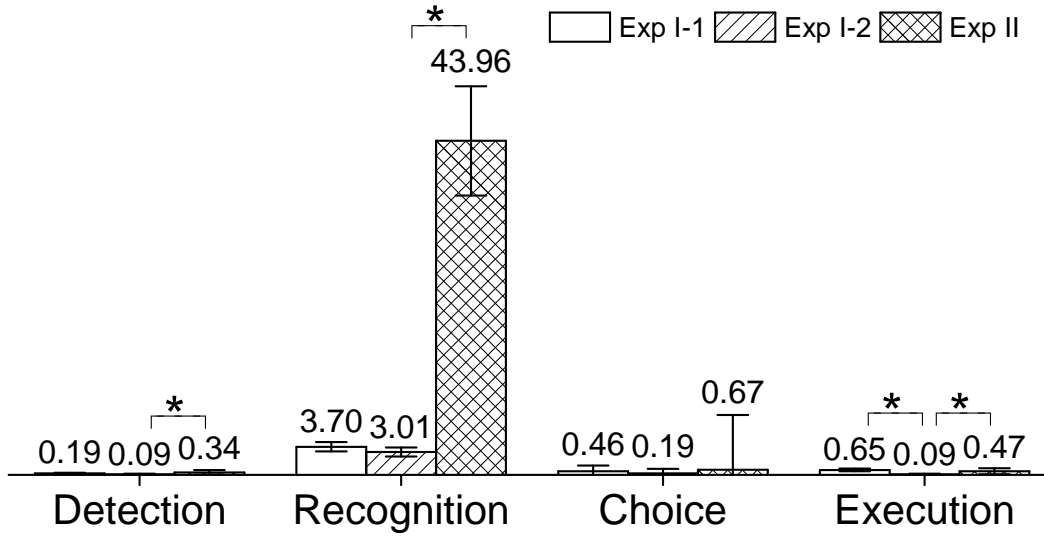


Fig. 6.1 Mean error rates (%) for four processing stages. Error bars represent standard errors.

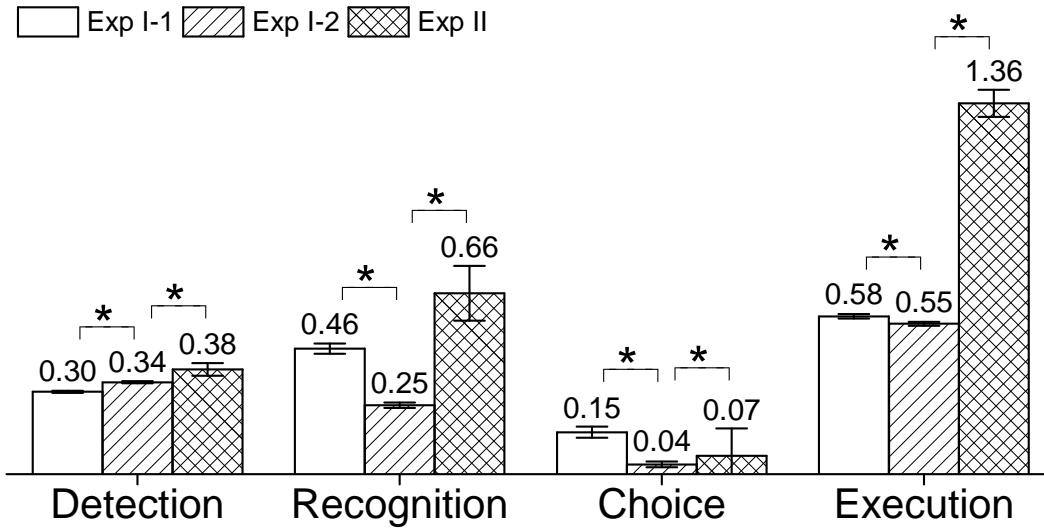


Fig. 6.2 Means and standard errors of the processing times (s) for four processing stages.

6.2.2 Effects of Concurrent Cue Presentation

In Experiment II, the recognition stage caused the largest error rate ($e^R = 43.96\%$; Fig. 6.1). Those of the other stages were very low: $e^D = 0.34\%$, $e^C = 0.67\%$, and $e^E = 0.47\%$. The processing times were $\tau^D = 0.38$ s, $\tau^R = 0.66$ s, $\tau^C = 0.07$ s, and $\tau^E = 1.36$ s (Fig. 6.2). Except the execution stage that depends on the response interface, recognition required the longest processing time.

e^R and τ^R were significantly larger than those of Experiment I-2 (Table 6.1). This result indicates that the two vibrotactile concurrent cues were more difficult for recognition than the single cues. τ^E was also significantly longer than that of Experiment I-2, but this simply reflects the longer cursor movement distance for multiple target selections. Statistically significant differences were also found in most of the other measures. However, they were of little practical importance considering their very small values (mean difference e^D : 0.25%; e^R : 0.48%; e^E : 0.38%; τ^D : 0.04 s; τ^C : 0.03 s). Recall that Experiment I-2 was conducted after Experiment II.

6.2.3 Effects of Cue Position Pair and Strength Pair

We performed two-way ANOVAs for Experiment II to evaluate the effects of cue position pair and cue strength pair on each processing stage. Results are summarized in Table 6.2. The interactions between the position pair and the strength pair were not analyzed because the large number (36) of the position pairs.

Table 6.2 showed a significant effect of cue position pair on e^R . Cue position pair also had significant influences on τ^D , τ^R , and τ^E . The large number of the cue position pairs made ordinary post-hoc methods for multiple comparisons ineffective, so we relied on graphical analysis. Fig. 6.3 shows the top and bottom cue position pairs, nine each, for τ^D , e^R , τ^R , and τ^E .

Comparisons between the good and bad performance groups can reveal the spatial aspects that caused the performance differences. τ^D was shorter when one of the two cues was applied to the right upper chest, but it was longer if no cues were given to the chest. e^R was obviously better when one or both cues were presented to the ankles than when both

Table 6.2: Two-way ANOVA results of Experiment II for the effects of cue position pair and strength pair on two performance measures (e : error rate, τ : processing time) for four processing stages (D: detection, R: recognition, C: choice, E: execution).

Effects of Cue Position Pair								
Statistic	e^D	e^R	e^C	e^E	τ^D	τ^R	τ^C	τ^E
$F_{35,385}$	1.12	10.91	1.24	0.80	2.61	5.48	0.92	35.05
p	0.301	* <0.001	0.172	0.785	* <0.001	* <0.001	0.608	* <0.001

Effects of Cue Strength Pair								
Statistic	e^D	e^R	e^C	e^E	τ^D	τ^R	τ^C	τ^E
$F_{3,33}$	0.45	5.80	0.24	2.09	0.11	0.38	0.01	0.26
p	0.717	*0.003	0.869	0.121	0.951	0.766	0.998	0.855

cues were provided to the trunk. τ^R was similar to e^R with respect to their performance groups, suggesting a high correlation between the two measures in the recognition stage. τ^E was small if one or both target circles were near to the initial cursor position (a star in Fig. 6.3), indicating τ^E was simply determined by the cursor movement distance.

Cue strength pair had a significant effect on only e^R (Table 6.2). The mean e^R for the four cue strength pairs are shown in Fig. 6.4. Tukey's HSD test showed that e^R was the smallest when both of the cues were of normal strength (mean 31.71%; n-n in the figure), and the largest when the two cues had different strengths (n-A: 51.15%; A-n: 46.39%)¹. e^R was in-between for the pairs of both accented cues (A-A: 44.58%). In this case, the response errors were made by missing one or both cue positions (Type A) more than by misunderstanding their strengths (Type B). Type B errors were more frequent in the other conditions.

In summary, the participants recognized the two simultaneous vibrotactile cues more quickly and accurately if one or both cues were presented to the ankles. The recognition accuracy was also higher if the two cues had the same strength level. The participants spent

¹For labeling, a cue given to a higher (left in the case of the same height) body site is denoted on the left side.

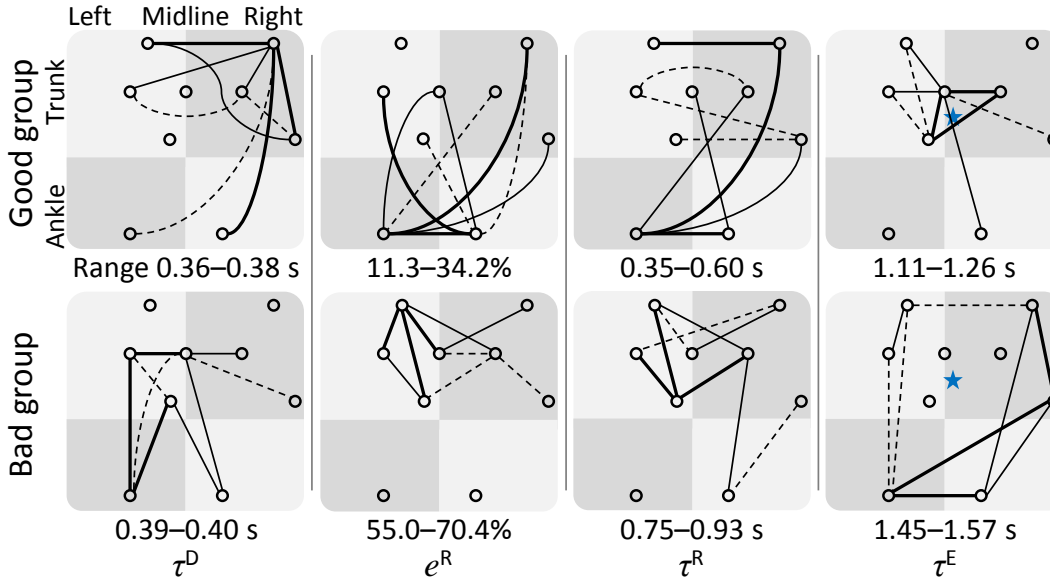


Fig. 6.3 Graphical representation of the best (upper row) and the worst (lower row) nine cue position pairs for the cue detection time (τ^D), the cue recognition error rate (e^R) and time (τ^R), and the response execution time (τ^E). Each pair is represented by a line connecting two body sites (circles), with thickness representing its rank (thick: 1–3rd; normal: 4–6th; dashed: 7–9th).

less time for detection if one or both cues were given to the right chest.

6.3 Discussion

6.3.1 Effects of Concurrent Cue Presentation

The detection performance measures were comparable between Experiments I-2 and II, albeit statistically significant differences (Fig. 6.1 and 6.2), indicating that the detection stage was mostly free from the effect of multiple cue presentation. At the time of detection, participants are not able to discern the number of vibrotactile cues, which is only possible after recognizing the cues. Therefore, the detection of two cues is essentially the same task as that for a single cue. It is reasonable to understand the small performance differences were resulted from the practice effect (Experiment I-2 was carried out after Experiment II).

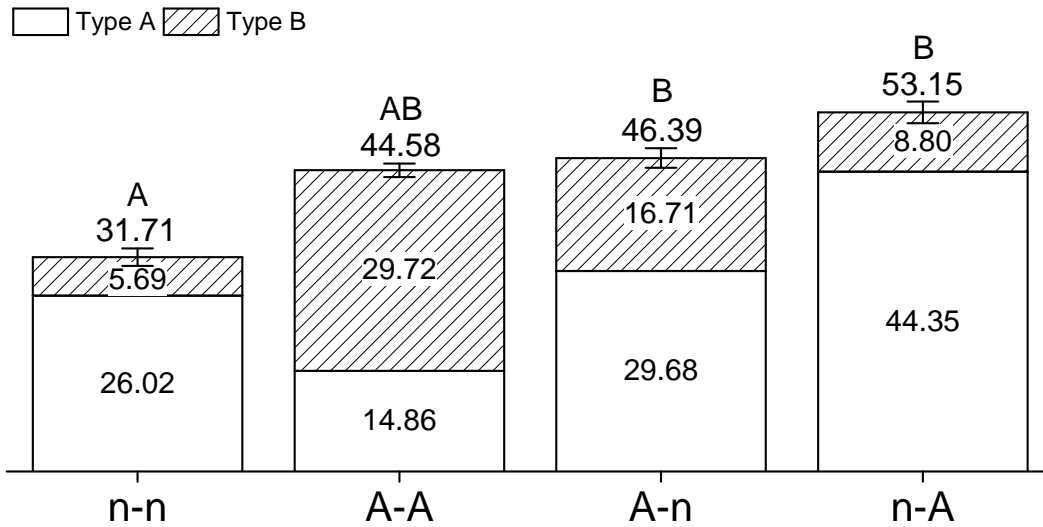


Fig. 6.4 Mean recognition error rates (%) for the pairs of cue strengths (n: normal and A: accented). Type A represents the misses caused by incorrect selection of a target circle, while Type B is for those with correct target selection but wrong strength response. Conditions with the same alphabet above the bar are of the same performance group by the test.

The recognition stage, however, showed apparent performance differences. In Experiment I-2, the participants showed 3.01% of error rate for single-cue recognition. If the recognition of different vibrotactile cues were independent from each other, the error rate for double-cue recognition would be 5.92% (expected probability of having error in recognizing two single-cues). The actual error rate (43.96%) measured in Experiment II was much greater, suggesting strong interference.

This result can be accounted for by two main reasons. First, one or both cues can be perceived incorrectly due to sensory interference between the cues. If two vibrotactile stimuli are given concurrently onto close body sites, one of the stimuli can be perceived to be weaker than its actual strength (masking), or the two stimuli can be perceived as one stimulus (funneling effect). These effects of sensory illusion are discussed in detail in the next section. Second, the sensory impression of a vibrotactile cue may fade out while processing another cue [15]. According to human memory models [1], any sensory information

is first stored in a temporary space called sensory memory for a very short time. If suitable attention is paid to the information within the time limit, it is transferred to working memory. Otherwise, the sensory information is dropped from sensory memory. Working memory can also lose information without conscious effort to retain the information. Due to these two reasons, cue recognition can be hindered greatly by other cognitive processes that interrupt attention to the cue.

The participants spent 0.66 s on average to recognize two vibrotactile cues in Experiment II, which is roughly two times longer than that for a single cue measured in Experiment I-2 (0.25 s), considering the gain by practice. This result suggests that multiple cue recognition did not affect the recognition time as significantly as the recognition accuracy.

In the choice stage, the performance differences between Experiment I-1 and II were negligible with the very small error rates and response times. Unlike cue recognition, response decision is a simple cognitive process that can be done instantly. Thus, it was almost independent of sensory illusions or memory limits.

For execution, the error rates were very low in both Experiment I-2 and II. The time used to enter two answers was expected to be twice longer than that for one answer, but the actually measured execution time was even longer. We noticed that the participants tended to have a short pause after entering the first answer to move to the next target precisely.

6.3.2 Effects of Cue Position Pair and Strength Pair

The participants detected the cues more quickly when one or both cues were given to the upper chest. Since the upper chest showed faster single-cue detection times (Section 5.2.2), it indicates that the detection time of a two-cue pair was determined by the cue that was perceived earlier. However, the detection time differences, although statistically significant, have little practical importance because of their very small values with respect to the total processing time.

The recognition of two vibrotactile cues was less accurate when both were presented to the trunk (Fig. 6.1). This is likely due to spatial vibrotactile masking, wherein a vibrotactile stimulus is perceived weaker than its actual strength when other vibrotactile stimuli are also

presented in proximity [17]. In our vest, the distances between the adjacent tactors on the trunk were not large (roughly 10–15 cm) for natural directional mapping to the target PIs. This design was sufficient for single-cue recognition, but seems to need improvements for double-cue recognition.

Vibrotactile masking also degraded the recognition accuracy among the cue pairs with different strength levels. Since the degree of masking is proportional to the amplitude of a masker stimulus [70], masking should have more effect on the cue pairs including one or more accented cues in Experiment II. In the case of both accented cues (A-A in Fig. 6.4), one or both of the cues could have misperceived as normal ones (Type B error in the figure). In the cases of two cues with different strength levels (A-n and n-A), it was possible that the weaker cue was not perceived as a result of masking, causing errors in the position response (Type A error). Our experiment results are in agreement with the expected effects of vibrotactile masking, suggesting that spatial masking was the main source of recognition error.

Type A error was also dominant in the case of both normal cues (n-n). This implies that, in addition to vibrotactile masking, there was another source of position recognition error for normal cues. One plausible explanation for this result is that the sensory memory of a cue has decayed and removed before it was recognized correctly while processing another cue. Because normal cues had shorter duration than that of accented cues, there was higher chance of such occurrence for normal cues. A similar explanation can be found in [16], which experimentally assessed human precision in the recognition of the number of vibrotactile stimuli given simultaneously over the body.

Overall, the recognition time was proportional to the recognition error rate. It is notable that some cue pairs in which the two cues are horizontally distant each other and vertically close had shorter recognition times, while having moderate recognition error rates.

Chapter 7

Experiment III: Series Identification of Single or Paired Cues

For the evaluation of our guidance design, we examined the scenario in which a series of single or multiple vibrotactile cues are presented and the participant responds to the cues.

7.1 Methods

7.1.1 Participants

We recruited 24 healthy male students (aged 19–24 years; mean 21.2) for the experiment. All the participants reported that they had no known sensorimotor disorders and had no experience of playing drum sets. They were paid 10,000 KRW (\simeq 9.37 USD) for their participation.

7.1.2 Task and Stimuli

The task was to perceive a series of vibrotactile cues, and then to strike the corresponding PIs of a drum set in the same order of the cues. In each trial, a participant was provided with vibrotactile cues four times, and one or two simultaneous vibrotactile cues were presented

at each time. For brevity, we call the case of one vibrotactile cue as a *single cue*, while that of two simultaneous cues was a *cue pair*. A single cue and cue pair require a single- and multiple-drum striking motion for the response, respectively.

The number of cue presentations (4), which result in 4–8 vibrotactile cues for a trial, was decided by considering the human working memory capacity (4–7; [2]). The interval between two consecutive presentations was set to 1 s, by taking into account the cognitive processing times of a single cue and a cue pair (0.77 and 1.11 s, respectively; [35]).

A single cue was given to one of the nine body sites (see Fig. 4.1(c)) except the left ankle, with one of two strength levels (normal and accented). The left ankle, which corresponds to the hi-hat pedal, was not used in this experiment, reflecting the fact that the pedal is rarely used in beginner-level drumming. In the case of a cue pair, their positions were chosen from six predefined position pairs, and their strength levels were both normal or both accented. The case of two cues with different strength levels was not included in the experiment since beginner-level drumming hardly involves simultaneous strikes with different strengths. We prepared the six pairs by combining each of the three cymbals (no. 1, 2, and 3 in Fig. 4.1(a)) with the snare drum (no. 6) and with the bass drum pedal (no. 9). These combinations are the most frequent ones in drumming. To sum up, total 28 choices (8 positions \times 2 strengths for single cues and 6 position pairs \times 2 strengths for cue pairs) of presentation were possible for each time.

Due to the extremely many number (28^4) of possible cue presentation combinations in a trial, it was almost impossible to test all the combinations. This led us to test only three types (simple, moderate, and complex) of combinations. The simple combination was defined as a sequence of four single cues that are given to the same, randomly chosen body site. Because only one body site was involved and stimulated repeatedly, it was very simple to recognize and memorize the position, which allowed the participant to use more attention for the cue strength levels and their order. The moderate combination was also composed of four single cues, but in this case, the target body sites were all different. Consequently, the participant also needed to recognize all the stimulated body sites and memorize their order. As to the complex combination, it was a random sequence of two cue pairs and

Table 7.1 List of choices for a single cue and a cue pair. Each body site is represented by a unique number in the same way with Fig. 4.1.

Type	Target body site(s)	Strength level
Single cue	1, 2, 3, 4, 5, 6, 7, or 9	normal or accented
Cue pair	1-6, 2-6, 3-6, 1-9, 2-9, or 3-9	both normal, or both accented

Table 7.2 Comparisons of three cue combinations used in the experiment.

Type	Composition	# body sites involved
Simple	4 single cues	1
Moderate	4 single cues	4
Complex	2 paired and 2 single cues	6

two single cues. Because it involved many (6) vibrotactile cues to process and complex multiple-limb movements, it was highly challenging to respond to this type of combination correctly. A summary of the single cues and cue pairs is shown in Table 7.1, and that of the three combination types is given in Table 7.2.

7.1.3 Procedures

Upon arrival, the participant was informed of the experimental task and procedures. Then, the participant signed on a written consent form and sat in front of the drum set. After that, the participant wore five vibrotactile belts, to each of which 1–3 vibration motors were fastened by clips, and adjusted them to place the motors on their corresponding body sites. The participant also wore earplugs and noise-canceling headphones to exclude any sound noise from the motors and that from the environment.

Before the experiment, the participant was well told with our guidance design and how to make drum strikes using drum sticks and a pedal, and then went through a familiarization session. The purpose of the familiarization session was to accustom the participant to the vibrotactile cues and the drum striking motions required for the task. In this session,

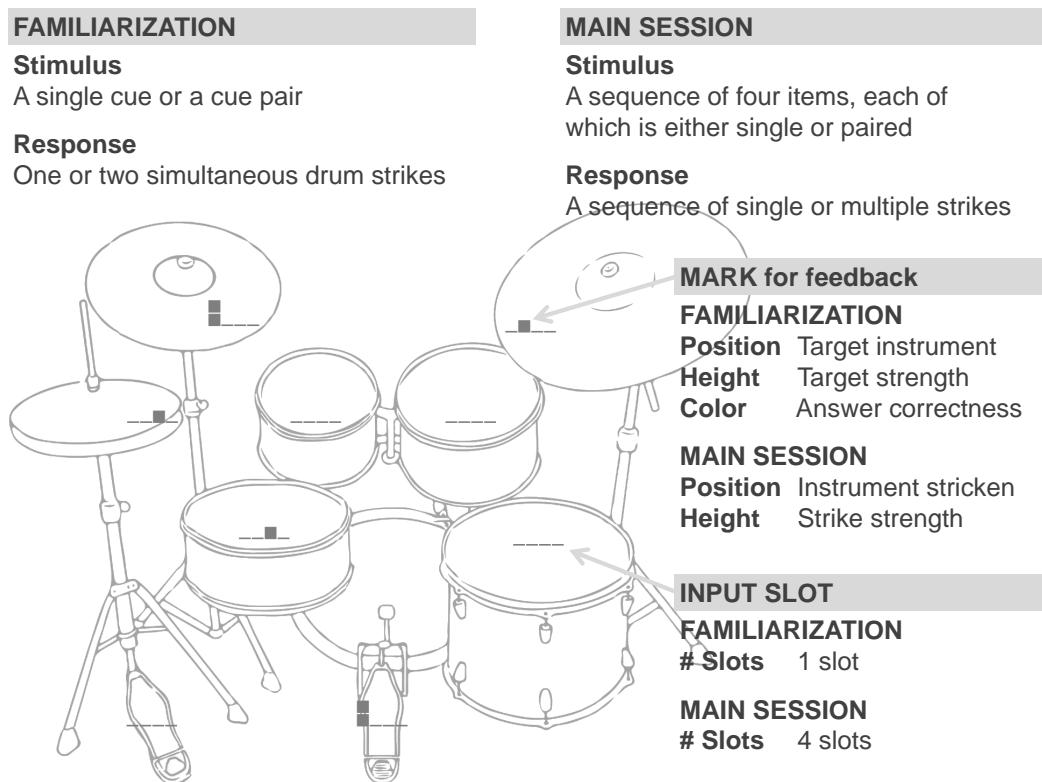


Fig. 7.1 Example of visual scenes provided to the participant and brief summary of two experimental sessions.

the participant performed 140 trials (8 positions and 6 position pairs \times 2 strengths \times 5 repetitions). The order of the trials was randomized by participant. In each trial, one single cue or cue pair was presented, and the participant answered to the cue(s) by performing the corresponding single- or multiple-drum striking motion. When making an answer, the participant was required to input all the drum strikes within a time period of 150 ms from the detection of the first drum strike by the drum set. If only one drum strike was detected within the threshold, the participant's answer was understood as a single-drum striking motion, otherwise it was a multiple-drum striking motion. Immediately after the answer, to promote the familiarization, the participant was provided with the information about the correct answer and the response correctness. This was done by displaying an image of a

drum set with a small rectangle to each target PI as shown in Fig. 7.1. The target strength level was represented by the height of the rectangle, and the accented strength level had a two times taller rectangle than that for the normal level. For the correctness feedback, the rectangle was filled green if the target PI was stricken with the proper strength level, or it was filled red. The feedback was given for 1 s, and then the current trial was terminated. The next trial started after another 1 s.

After completing the familiarization session, the participant took a short rest and then proceeded to the main experiment session. The main session was composed of 45 trials (3 combinations \times 15 repetitions), with a random trial order for each participant. In each trial, one of the three cue combinations explained in Section 7.1.2 was presented, and the participant answered to the combination by performing a sequence of drum striking motions, which corresponds to the given cue combination. No visual feedback about the correct answer or response correctness was provided to minimize the learning effect. Instead, the participant's actual input was displayed to help the participant to be aware of own input and how many drum striking motions are remaining to complete the task. For this, an image of a drum set was displayed (see Fig. 7.1). There were four input slots for each PI, with an input cursor on the leftmost slot. The participant's input to a PI was represented as a small gray rectangle on the current cursor position of the PI. The height of the rectangle represented the input strength level, and it was twice high for the accented level than that for the normal level. After received input to any PI, all the PIs moved their cursors to the next slot and waited for the next input. Multiple drum strikes to different PIs made within a short time (threshold 150 ms) resulted in only one cursor shift as they were regarded as one multiple-drum striking input. The trial ended 2 s after the participant's input to the fourth slot, and the next trial started after another 1 s.

The target body site(s) and strength level of each single cue or cue pair in a combination and the order of the cues were randomized by trial. When selecting the positions, it was required to involve as many body sites as possible (1, 4, and 6 sites for simple, moderate, and complex combinations, respectively). These were to maximize and maintain the respective difficulties of the three combination types throughout the experiment.

The whole experimental procedures completed in about 30 min.

7.1.4 Performance Measures

For each drum strike, we measured the PI number, striking strength, and response index. The PI number was the MIDI number assigned to the target PI. The striking strength was originally a value that ranges from 1 to 127 and shows how strong the strike was. For the analysis, we converted the value to a binary digit (0 for the normal strength level and 1 for the accented) by simple thresholding. The strength threshold for each PI was prepared by a pilot test that required a professional drum player to strike all the PIs with two different strength levels. The response index was the cursor position number (1–4) at the time of measurement, which indicates to which combination component (i.e., a single cue or a cue pair) the measurement data is related.

Using the measurement data, the response correctness c was calculated for each combination component in a trial. For a single cue, c was 1 if only one drum strike was made for the cue and its target PI and strength were the same as those of the cue, otherwise it was 0. c for a cue pair was 1 if there were two drum strikes and they well matched to the pair, otherwise 0.

For each trial, the correctness score C was obtained by adding all the c values in a trial together. We also measured the response time t , which was defined as the time from the initiation of a trial to the detection of the participant's input to the input slot. The measured C s and t s were averaged across the trials in the same experimental condition and participant, and then fed to the statistical analysis.

7.2 Results

The performance measures for a single and a cue pair are shown in Fig. 7.2, and those for the three combination types are given in Fig. 7.3, respectively.

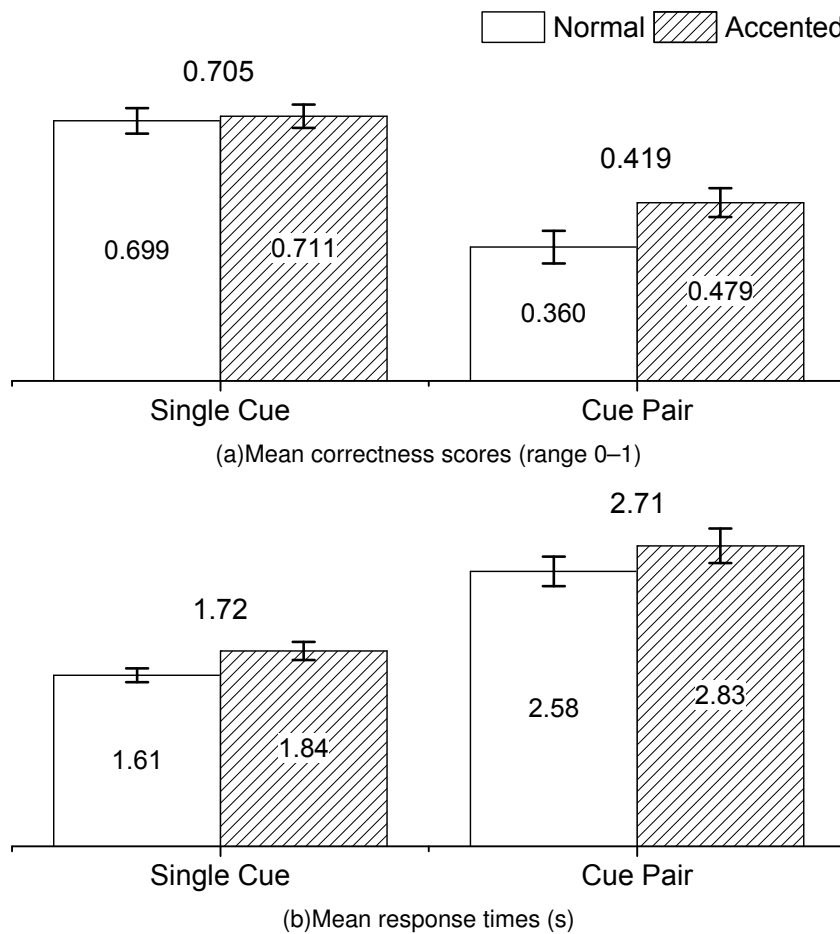


Fig. 7.2: Performances of single cues and cue pairs. Higher correctness score and shorter response time indicates better performance.

7.2.1 Recognition of Single or Paired Cues

In the familiarization session, the participant answered to a single cue or a cue pair at each trial. The task performed was actually a subtask of the main experiment task. In this regard, the understanding of its performance would help us understand the main experiment results.

The participants showed C of 0.705 for a single cue on average, and it was 0.419 for a cue pair. A two-way ANOVA was performed on the effects of cue type and cue strength, and the difference between the two cue types was statistically significant ($F_{1,23} = 167.93$ and

$p < 0.001$). There was a significant interaction effect between the main factors, and the simple effect tests indicated that cue strength was a significant factor for cue pairs ($F_{1,23} = 5.10$ and $p = 0.034$; mean difference 0.119), but it was not for single cues ($F_{1,23} = 0.06$ and $p = 0.815$; mean difference 0.013).

As to t , the participants required 1.72 s of time to respond to a single cue and 2.71 s to a cue pair on average, with a significant difference between the cue types ($F_{1,23} = 104.30$ and $p < 0.001$). Normal-strength cues were responded 0.24 s faster than accented ones, and this difference was statistically significant ($F_{1,23} = 35.02$ and $p < 0.001$).

7.2.2 Series Recognition of Cues

C was the highest for the simple combination (mean 3.536), the second highest for the moderate combination ($=2.250$), and the lowest for the complex combination ($=0.928$). For t , the simple combination had the shortest (mean 7.54 s), while the complex combination had the longest (13.71 s), with the moderate combination in between ($=10.15$ s).

For both measures, it is obvious that the performance had decreased as the combination complexity increased. This was confirmed by one-way ANOVAs that evaluated the effects of combination complexity on the two measures ($F_{2,46} = 414.64$ and $p < 0.001$ for C , $F_{2,46} = 55.34$ and $p < 0.001$ for t). The following Tukey-Kramer tests showed that none of the combinations belonged to the same performance group for both of the measures.

7.3 Discussion

7.3.1 Recognition of Single or Paired Cues

In our previous study [35], the response accuracy was 96.2% for a single cue and 55.0% for a cue pair, which correspond to C s of 0.962 and 0.550, respectively. Compared to the previously reported data, the accuracy values of the present study (0.705 and 0.419, respectively) are rather small. This difference is mainly due to the different means of response collection. In the previous work, we assessed how correctly our system can deliver the guidance information, and a mouse interface, which is one of the most familiar input devices to most

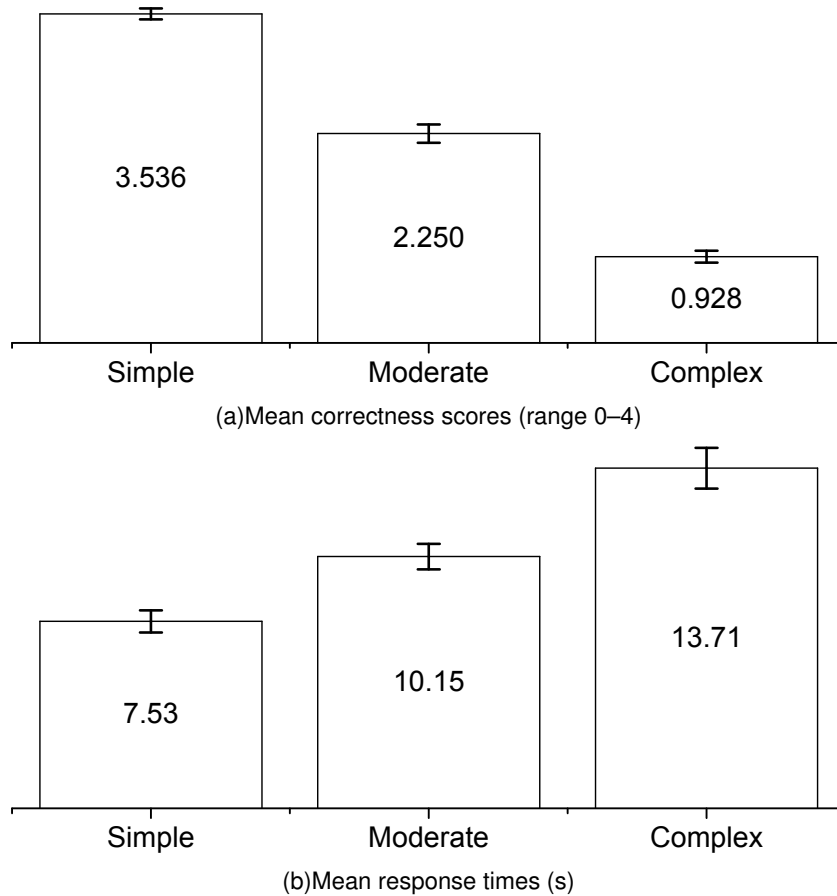


Fig. 7.3: Performances of three cue combination types of different complexity levels.

participants, was used for minimizing the error in entering a response. For the present study, instead of a mouse interface, we used a drum set for an evaluation of our guidance system under a more practical use scenario. Because all the participants had no prior experience of drumming, they made a lot of response errors when striking a PI. The error was more obvious in the strength level. This is because entering a strength level required a precise control of both striking strength and position¹, while entering a target PI did not require

¹For a given strength, a drum strike to the boundary results in a much small strength reading compared to a strike to the center.

position control that much. If we ignore the strength error, the mean C of the participants was 0.880 for a single cue and 0.677 for a cue pair.

An experiment with expert drummers, who can control the striking strength at their will, would result in a more accurate evaluation of guidance performance of our system. However, hiring a sufficient number of expert drummers was costly and hardly achievable. Training novice participants before the experiment was also not a solution because the ability to control striking strength requires prolonged training to be mastered.

The accuracy for a normal cue pair (mean C 0.360 out of 1.000) was much lower than that for an accented one (0.479). A similar result was also found in our previous study, and we conjectured that this problem was resulted from the rather weak and short vibration stimuli of a normal cue pair. To improve the problem while keeping normal cue pairs distinguishable from accented ones, we introduced stimulus-onset asynchrony (SOA) between the two vibration stimuli of a normal cue pair instead of increasing their vibration intensities or durations. The experiment result suggests that our solution was not effective in increasing the response accuracy of a normal cue pair. It is likely that the length (30 ms) of SOA was not sufficient. We think that the temporal synchrony of cues is important to guide a multiple simultaneous drum striking motions, and thus a short, not noticeable SOA was used in the experiment so that the two stimuli in a cue pair were perceived to be synchronous despite the SOA. However, such a short SOA seems not adequate for the second cue to stay in the memory until the participant becomes available for the cue. A longer SOA would improve the response accuracy, but it is also expected to deteriorate the perceived simultaneity of cues. It needs more study to determine the optimum value of SOA.

The participants used more time (0.99 s) to answer a cue pair than a single cue. Since the two cues of a cue pair were presented almost in parallel, the time to transmit the cues was comparable to the time for a single cue. The time to enter an answer was also similar as multiple PIs were entered simultaneously. In this regard, it seems that the time difference is mostly due to the time to recognize an additional cue and determine its target PI. The result is in contrast with our previous result that a cue pair required only 0.28 s of additional time for the cognitive processing compared to a single cue. This is because to some extent no

training was provided prior to the familiarization session and the learning gain during the session was not strong due to the relatively small number of trials. Also, it is likely that the search for the target PI of a real drum set was more time-consuming than searching a drawing of a drum set, due to the larger size and non-planar structure of the real drum set.

The response time difference (0.24 s) between the normal and accented cues is mainly due to the larger striking motion of accented strikes. For an accented strike, the participant first raised the drumstick (or released the pedal) and then initiated a strike from a farther location to the target PI than that for a normal strike, and the longer range of striking motion contributed to the longer response time.

7.3.2 Series Recognition of Cues

The participants were accurate (mean C 3.536 out of 4.0) in responding to a simple combination, which consisted of four single cues to the same body site with random strength levels. Considering the minimal training for the task and the error in entering the answer, it seems that the participants could recognize the cues and determine and memorize the answer almost perfectly. Two reasons may contributed to this result. First, except the first cue, the participants did not need to move their attention and recognize the stimulated body site, and thus they could recognize the cue strength levels more quickly and accurately. Second, the memory for an answer (one target PI and a sequence of four striking strengths) was retained easily owing to the relatively small amount of information.

The moderate combination presented a sequence of four different single cues, and about half (mean C 2.250) of the cues were correctly answered by the participants. If the cues were processed independently of each other and the participants had no difficulty in memorizing their answers, the correctness score C of the moderate combination should have been around 2.820 ($=4 \times C_{single}$). The measured value of C is rather lower than the expected value, suggesting that the task was hindered to some extent for a certain reason. It is probably due to the insufficient time interval (SOA 1.00 s) between the cues. We decided the interval based on the previous result that 0.77 s of time were used on average to detect and recognize a single cue, and to decide its answer. However, as discussed in Section 7.3.1,

the participants of the present study required much longer times for the recognition and decision. This suggests that the processing of a cue could not be completed within the time interval, and the processing of the subsequent cue might be delayed for a while to complete the current one. Because the sensory impression quickly decays and disappears from the memory [67], the processing delay could have led to the loss of information about a cue, and consequently to an incorrect processing of the cue. Regarding the information (four target PIs with their respective striking strengths and the order) needed to be remembered to respond to a moderate combination, it is within an acceptable level considering the working memory capacity. Therefore, it is not regarded as a main reason of the lower correctness score.

The complex combination was composed of two cue pairs and two single cues, and its expected value of C was 2.248 ($=2 \times C_{single} + 2 \times C_{pair}$), assuming independent processing of the cues. The actual value (mean C 0.928) was much lower than the expected value, which indicates strong hindrance to the task. One such hindrance is the cue processing delay, as discussed earlier. Because a cue pair requires longer processing time than that of a single cue, the processing of subsequent cues may be delayed for a longer time, which results in a higher possibility of missing the sensory impression of the cues. The large amount of information for the answer also accounts for the low accuracy of complex combination. For a correct answer to a complex combination, the participants required to retain the information on four target PI groups, their respective striking strength, and their order, which is rather difficult to achieve especially in a situation of high cognitive and physical workload.

The low accuracies of the moderate and complex combinations do not necessarily mean that our guidance system is not applicable to the combinations of these complexity levels. Even for a human tutor, it is difficult to instruct learners in complex movements at once, and repeated demonstrations of the target task is necessary to deliver the target task. It is also a common method to demonstrate the task more slowly than the desired speed. In this regard, with an adequate number of repetitions and a longer time interval between adjacent cues, our system may also be successful in delivering more complex combinations.

As the response time was 1.72 s for a single cue and 2.71 s for a cue pair, the moderate

combination was expected to be answered in 6.88 s and the complex combination was in 8.86 s. For the simple combination, its response time was expected to be shorter than that of a moderate combination because the cognitive processing was much simple and there was no transition movement from a target PI to the next target. In the experiment, the response times were much longer than the expected time for all combinations. The longer response times were partly attributed to the fact that the participants often took a short pause after making a drum striking motion to balance their body and to assure that the answer was inputted as intended. The difference between the actual and expected response time is proportional to the task complexity (3.27 and 4.85 s for the moderate and complex combinations, respectively). This suggests that the process of detecting and recognizing multiple cues and deciding and retaining their answers was slowed down as the complexity increases. However, it is also possible that the result was resulted from the longer pause time after a more complex movement, and more study is required for the verification.

Experiment IV: Application to Drumming Learning

In Experiment IV, we evaluate our vibrotactile guidance method under a scenario of actual drumming learning and compare its efficacy with those of visual-based methods.

8.1 Methods

8.1.1 Participants

The same participants with Experiment III participated in Experiment IV. This was to minimize our resources for recruiting and training. For each participant, Experiment II was performed for 4 consecutive days from the day after Experiment III. The participants were paid 40,000 KRW (\simeq 37.48 USD) for their participation.

8.1.2 Task

The task was to learn a 1-measure long drum rhythm for 3 min. Each participant learned three sets of rhythms (S1, S2, and S3; see Fig. 8.1) under the learning condition assigned to each rhythm set and participant. Each rhythm set was composed of three rhythms (R1, R2, and R3) to assess the effectiveness of guidance for a variety of rhythms with different difficulty levels. R1 was the simplest rhythm that consisted of eight 8th notes with accents

	Rhythm 1 (2 PIs, metric)	Rhythm 2 (3 PIs, metric)	Rhythm 3 (7 PIs, partially non-metric)
SET 1			
SET 2			
SET 3			

Fig. 8.1 Three sets of target rhythms.

on the first and fifth ones (i.e., on the first and third beats). The notes were evenly distributed to two different PIs, the ride cymbal and the snare drum, that were played by different hands. R2 was such that the participant intermittently kicks the bass drum 3 times while stroking the hi-hat at every half beat and the snare drum at the second and fourth beats. This rhythm involved three different limbs, two hands and one foot, and included many simultaneous striking movements of two limbs. As to R3, it resembled to R2 for the first half, and its second half consisted of six rhythmic strokes that strike twice for each of the three tomtoms. In this rhythm, many PIs were involved, and the target PI of each hand was frequently changed to perform the rhythm.

The three rhythm sets were prepared in consultation with a local drum tutor (17 years of drumming and 6 years of tutoring experiences). The rhythms were designed in such a way that the rhythms of the same difficulty level are different in body movement sequence (i.e., motor memory) but very similar in sensory-motor difficulty. For all rhythms, the target tempo was set to 40 BPM.

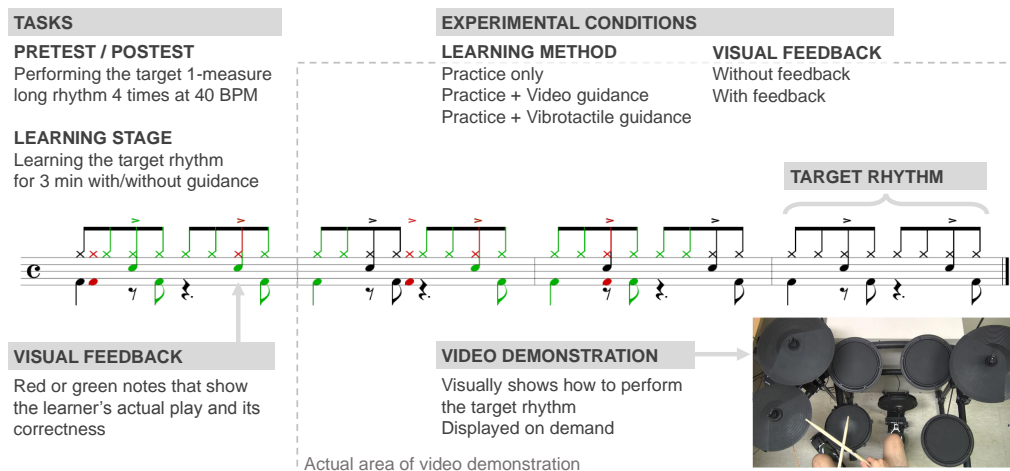


Fig. 8.2: Example visual scene.

8.1.3 Conditions

Three learning methods (P0, PV, and PT) were prepared for the experiment, and each participant experienced all the conditions by learning three rhythm sets with different learning conditions. Method P0 was a baseline condition. In this method, the participant learned the target rhythm by practice only. A practice was defined as performing the target rhythm four times in succession. For this, four measures of the same target rhythm were displayed on the screen, as shown in Fig. 8.2. To begin a practice, the participant pressed the hi-hat pedal. Then, metronome-like sound guidance, which played a ticking sound periodically, was provided, and the participant performed the target rhythm four times while matching the play speed to the guidance. A metronome is a common means of timing guidance in music learning, and we provided metronome-like auditory guidance for all learning methods for a more realistic evaluation of our system. After that, the hi-hat pedal was pressed again to terminate the practice. The participant repeated the practice for a given learning time (3 min). Method PV was a representative of the guidance methods that have been used in the situations of tutoring or self-teaching. In this method, the participant was provided with, as well as practice, an additional option of watching a short video recording (see the

inset in Fig. 8.2) in which an expert drum player performs the target rhythm once. By watching the video, the participant could easily obtain the idea of how to perform the target rhythm correctly. The video was provided at the beginning of learning and when the participant expressed the need for the guidance by stroking the crash cymbal. Method PT was the same as Method PV, except it provided a vibrotactile guidance instead of a video guidance. The vibrotactile guidance delivered the idea of correct performance by transmitting short vibrotactile cues, each of which instructs the target PI, striking strength, and time of a drum strike for the target rhythm. The time for learning was kept constant for all methods, and thus the participant could have less practice in Method PV or PT than in Method P0. This was more clear if the participant spent more time to observe the guidance.

It was expected that the effectiveness of guidance is influenced by the existence of augmented feedback that provides information about the performance of practice. To evaluate the influence from the feedback, the participants were randomly distributed into two groups (12 participants each). Then, one participant group was provided with visual feedback (Condition F1) that shows how accurate the participant's practice is, while the other group was not given such feedback (Group F0). For the participants of Group F1, their actual performance was displayed on the screen during the practice along with the target rhythm. A drum strike was correct and represented as a green note, if it was made within a time threshold (185 ms = 32th note duration at 40 BPM) from the desired time of a note in the target rhythm and with the correct PI and strength level of the target note. Otherwise, the strike was incorrect and represented as a red note.

In summary, total six experimental combinations (3 learning methods \times 2 visual feedback conditions) were evaluated in the experiment.

8.1.4 Procedures

For brevity, the experimental procedures that were also used in Experiment I are not repeated here. Before the experiment, the participant was briefly introduced to how to read a piece of drum music. Also, a diagram that shows the mapping between the PIs and note pitches was provided throughout the experiment.

The experiment was performed for four consecutive experimental days, and for each day, each participant carried out three experimental sessions. In each session, the participant learned a rhythm set, from S1 to S3, using one of the three learning methods. The order of learning methods was balanced across the participants and kept constant throughout the experiment for each participant. A rhythm set was comprised of three rhythms, and the participant learned the rhythms in the order of R1, R2, then R3. We used the same rhythm order for all sessions, days, and participants, because there was no reasonable method to balance the order effect on the rhythms with different difficulties, and also because the performance differences among the rhythms were not of our main interest.

For each rhythm, the participant first took a pretest to measure the participant's ability to perform the target rhythm before learning. The test was initiated by the participant by pressing the hi-hat pedal. Then, the participant performed the target rhythm displayed on the screen four times without any guidance or visual feedback, and then pressed the hi-hat pedal again to complete the test. After the test, the participant learned the target rhythm for 3 min, using the given learning method and visual feedback, following the procedures described in Section 8.2. Immediately after the learning, the participant carried out a posttest, whose procedure is the same as the pretest, to measure the immediate gains from learning. After completing the test, the participant took a short break, and then proceeded to the next rhythm.

A session was completed with the posttest of R3, and the experimental procedures for a day were completed after completing all the three sessions, which took about 50 min. The same procedures were repeated for the first three experimental days, which included total 6 test points (T1, T3, and T5 for the pretest, and T2, T4, T6 for the posttest at Day 1, 2, and 3, respectively). The experiment was completed with the fourth experimental day that tested (T7) the participant's final ability to perform each rhythm.

8.1.5 Performance Measures

During the experiment, for each rhythm and learning method, each participant performed 7 tests (T1–7). For each test, the experimental program measured the target p , strength level

s , and time t for each drum strike \vec{x} . p was a one-digit number unique to each PI of the drum set. s was a binary number, whose value was 1 if the strike strength exceeded a predefined threshold for the target PI, or 0. t was the time at which the drum strike was sensed by the drum set.

To evaluate the accuracy of a test performance, we compare the participant's actual performance data with that of desired one. For this, we first define each actual and desired performance data as a sequence of drum strikes,

$$\vec{X} = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n), \quad (8.1)$$

$$\vec{Y} = (\vec{y}_1, \vec{y}_2, \dots, \vec{y}_m), \quad (8.2)$$

where, \vec{x}_i and \vec{y}_i are the i th drum strike of actual performance \vec{X} and desired performance \vec{Y} , respectively, and n and m are the total number of items in \vec{X} and \vec{Y} , respectively. Then, using dynamic time warping [58], \vec{X} and \vec{Y} were aligned in such a way that the number of matching items between the sequences is maximized¹. Because the dynamic time warping is only applicable to time series data but some of our target rhythms included simultaneous drum strikes, we used the item index i instead of the exact striking time t for the alignment. Also, in actual performance, simultaneous strikes to different PIs were sensed as successive single strikes, which can lead an incorrect alignment result depending on the order of sensing. We prevented such occurrences by sorting simultaneous (time window 125 ms) drum strikes in the order of p when constructing \vec{X} . As a result of the alignment, we obtain a list of matching strike pairs between \vec{X} and \vec{Y} ,

$$\vec{M} = (\vec{m}_1, \vec{m}_2, \dots, \vec{m}_l), \quad (8.3)$$

$$\vec{m}_i = (\vec{x}_j, \vec{y}_k), \quad (8.4)$$

where, \vec{m}_i is the i th pair of a match list \vec{M} , l is the number of items in \vec{M} .

Using the alignment result, we calculate three performance measures for statistical analysis. Targeting error ratio p_{err} is the ratio of the strikes that are incorrect in their target (i.e., PI mismatch, extra strikes, or missing strikes) to the total strike items in \vec{X} and \vec{Y} . For PI

¹The `samplealign` function of MATLAB was used for this data processing stage.

mismatch, \vec{m}_i was counted as one mismatch if \vec{x}_j and \vec{y}_k are different in their PI. Missing strikes are the strikes that were required to perform but the participant missed (=unpaired strikes of \vec{Y}), and extra strikes are those unnecessarily made by the participant (=unpaired strikes of \vec{X}). By definition, the number of missing strikes is $m - l$, and that of extra strikes is $n - l$. p_{err} was calculated by the equation below,

$$p_{err} = \frac{c_{PI} + m + n - 2l}{m + n - l}, \quad (8.5)$$

where, c_{PI} is the number of PI mismatches in \vec{M} .

Strength level error ratio s_{err} was calculated by counting the number of strength level mismatches in \vec{M} and dividing it with the total number of items in \vec{M} ,

$$s_{err} = c_s / m, \quad (8.6)$$

where, c_s is the number of strength level mismatches in \vec{M} .

Similarly to s_{err} , timing error ratio t_{err} was calculated by comparing (threshold 125 ms) the target and actual strike times of each match pair in \vec{M} ,

$$t_{err} = c_t / m, \quad (8.7)$$

where, c_t is the number of timing mismatches in \vec{M} . A special care was required when computing t_{err} . For a test performance, we did not provide a reference signal for performance initiation or guidance on the target speed (i.e., tempo). Due to this, the participant's performance could be delayed or performed at a different tempo, and consequently the direct comparison between the desired and actual strike time was not effective in measuring the accuracy of relative timing of drum strikes (i.e., rhythm). To solve this problem, we removed the time delay and tempo difference between the target and actual performances based on linear regression result, and then used the modified strike times for the computation of t_{err} .

8.2 Results

Figure 8.3 shows three performance measures of the participants at each learning method and at test point T1–7.

8.2.1 Performance Before Learning

At test point T1, two groups of participant measured their initial performance for each learning method. On average, the participants showed similar targeting error ratios among the methods (9.4, 9.1, and 8.8% for P0, PV, and PT, respectively). The participants had the lowest timing error ratio (44.8%) at Method P0 and the highest (52.4%) at Method PT, with an intermediate (46.9%) timing error ratio at Method PV. Method PV had the highest (28.5%) strength error ratio, and Method P0 had the lowest (24.6%), with Method PT in the middle (26.8%). The much smaller values of targeting error ratio than the other measures indicates that the participants generally faced less difficulty in striking the correct PI to perform a rhythm than matching the time and strength level of the strike.

For all performance measures, no statistically significant difference was found in a two-way ANOVA among the learning methods (p -values of 0.927, 0.289, and 0.230 for p_{err} , t_{err} , and s_{err} , respectively), between the participant groups (0.183, 0.210, and 0.179, respectively), and their interaction effects (0.659, 0.894, and 0.256, respectively). This suggest that the initial performances of the six experimental conditions were relatively similar to each other.

8.2.2 Guidance and Practice During Learning

For the first three experimental days, two groups of participants learned target rhythms using three learning methods. During learning, the participants performed 6.3 practice trials on average to learn a target rhythm at Method P0. They performed 5.3 times of practice trials and received 3.2 times of video guidance at Method PV, while having 5.2 practice trials and 3.6 vibrotactile guidance times at Method PT.

A three-way ANOVA on the number of practice trials and following Tukey-Kramer's multiple comparison tests showed that the participants performed significantly more number of practices at Method P0 ($F_{2,44} = 63.62$ and $p < 0.001$). Also, significantly less number of practices were performed on Day 1 ($F_{2,44} = 14.91$ and $p < 0.001$). No statistical significance was found between the participant groups ($F_{1,22} = 1.24$ and $p = 0.277$). For the number of guidance, the participants required significantly more guidance on Day 1 than

Table 8.1 Three-way ANOVA results on the effects of test point (T), learning method (M), and existence of visual feedback (F) for three performance measures.

Effect	p_{err}		t_{err}		s_{err}	
	F	p	F	p	F	p
T _{5,110}	19.25	*<0.001	30.30	*<0.001	16.30	*<0.001
M _{2,44}	0.96	0.393	3.48	*0.039	0.55	0.579
F _{1,22}	0.00	0.965	1.29	0.268	7.71	*0.011
T×M _{10,220}	0.99	0.456	0.29	0.983	0.45	0.921
T×F _{5,110}	1.89	0.103	2.58	*0.030	0.49	0.785
M×F _{2,44}	0.06	0.945	2.89	0.066	0.08	0.928
T×M×F _{10,220}	1.13	0.342	0.56	0.845	0.38	0.956

Day3 ($F_{2,44} = 5.13$ and $p = 0.010$), with no significant difference between Method PV and PT ($F_{1,22} = 1.20$ and $p = 0.285$), and between the groups ($F_{1,22} = 1.17$ and $p = 0.290$).

8.2.3 Performance Gains from Learning

To assess the effectiveness of three learning methods, the performance of two participant groups with different visual feedback conditions were evaluated at six test points (T2–7) by three error measures. In general, the errors were reduced with time, while converging toward certain levels. They were relatively smaller at the tests taken immediately after learning (T2, T4, T6) than at those taken after one day of recess (T3, T5, and T7). Three-way ANOVA results on the effects of test point, learning method, and visual feedback on the three performance measures are summarized in Table 8.1.

On average, targeting error ratio p_{err} , which was initially 9.1%, was improved to 2.4% at T7. The mean p_{err} averaged across T2–7 was the lowest (3.44%) at Method PT and the highest (3.94%) at Method P0, with an intermediate p_{err} (3.78%) at Method PV. A slight difference between the Method P0 and PT was visible at the early stages of learning (T2–3), but it disappeared in the later test points. A three-way ANOVA indicated that statistical significance of test point on p_{err} , but no significance of learning method nor visual feedback.

For all learning methods, timing error ratio t_{err} had decreased greatly (48.0% to 13.4% on

average) by learning. For T2–7, Method PT generally showed the lowest t_{err} (17.3%), despite the relatively high initial performance rate at T1. It was closely followed by Method PV (17.8%), and Method P0 had the highest t_{err} all the time (21.0%). For t_{err} , learning method had a significant effect, as well as test point. A Tukey-Kramer's multiple comparison test indicated that Method PT was significantly different from Method P0 ($p = 0.050$), while Method PV was marginally different ($p = 0.099$). We found a significant interaction effect between test point and visual feedback. A simple effect test showed that visual feedback had a significance only at T3 ($p = 0.031$), with marginal significance difference at T5 ($p = 0.086$). We also found a marginal interaction effect between learning method and visual feedback. Visual feedback had a marginally significant effect ($p = 0.083$) for Method P0, while having no significant effect for Method PV ($p = 0.896$) and PT ($p = 0.228$). Due to this effect, the difference among the methods was not significant ($F_{2,44} = 1.05$ and $p = 0.357$) for the participant group F1 who received visual feedback while practice. For the participant group F0, the effect of learning method was significant ($F_{2,44} = 5.32$ and $p = 0.009$), and Method PV was significantly different from Method P0 ($p = 0.008$) while Method PT was marginally different ($p = 0.070$). The mean performance difference between Method PT and P0 for T2–7 was 4.8% for Group F0 and 2.5% for Group F1, resulting in a 3.6% of difference in overall. It was 6.7% for Group F0 and -0.4% for Group F1 between Method PV and P0, and the total mean difference was 3.2%.

Regarding strength level error ratio s_{err} , it was initially 26.6% on average and reduced to 11.8% at the time of T7. For T2–7, the participants showed mean s_{err} of 14.3, 13.8, 14.8% for Method P0, PV, and PT, respectively. For this measure, Group F1 outperformed Group F0, showing significant effect from visual feedback. No significant effect was observed from learning method.

8.3 Discussion

8.3.1 Guidance and Practice During Learning

In the experiment, the participants learned each rhythm for 3 min for a day. At Method PV and PT, they had two options of practicing the rhythm and receiving guidance, and they used some of the learning time for the guidance and the remaining time for the practice. In contrast, no guidance was provided at Method P0, and the participants could use the learning time solely to the practice, resulting in a more practice at this method. This is important for t_{err} in that the video or vibrotactile guidance of guided methods (PV and PT) was significantly effective despite the more practice of non-guided method (P0).

As proceeded the experiment, the participants had grown in their ability to perform the target rhythms, and the need for guidance had been decreased. This caused gradual decrease in the number of guidance request and gradual increase in the number of practices.

8.3.2 Targeting

During the experiment, the participants showed relatively small amount of targeting error, and neither learning method nor visual feedback had significant effects in reducing the error. This result suggests that striking the correct PI was relatively easy, and no explicit guidance or feedback was necessary to improve the performance.

The effectiveness of guidance can be underestimated because the participants were provided during the experiment with a diagram that explained the relationship between the note pitch and target PI. The diagram was provided as a minimum guidance on reading drum music in that, without such information, the participants were unable to perform the target rhythm at test point T1, and the participants without visual feedback (Group F0) had no means to improve their accuracy at Method P0.

Also, because the relationship between the note pitch and PI was unchanged throughout the experiment, there was a high possibility of skill transfer among the learning methods. That is, the knowledge obtained from Method PT or PV was also useful at Method P0, and consequently the differences among the methods were quickly removed with time.

One solution for preventing the transfer effect is to apply the learning methods to different participant groups (i.e., between-subjects design). However, in this solution, there is another problem that the effectiveness estimation of a learning method can be inaccurate due to the intrinsic difference among the participant groups. A practical problem in recruiting participants also prevented us to utilize between-subjects design for learning method.

8.3.3 Timing

In contrast to targeting error, the participants made a lot of errors in strike timing. This indicates making a strike timely was relatively difficult, and guidance was desirable. The participants showed a significantly better timing performance during learning when they were provided with additional video (Method PV) or vibrotactile (Method PT) guidance than when they performed more practice instead of such guidance (Method P0). The performance difference between Method PT and PV was not significant. Although different sensory modalities were used in the video guidance (vision and hearing) and vibrotactile guidance (touch), both guidance methods delivered the temporal pattern of a target rhythm. In contrast, metronome-like sound display, which provided for all learning methods as a minimum guidance on timing, transmitted a short periodic sound that can be used merely as a reference. In this regard, the experiment result can be understood that the direct presentation of a target rhythm is more effective than a periodic signaling to teach the temporal pattern of the target rhythm.

At a glance, the effect of guidance can be seen as less meaningful compared to the effect of learning time as the participants showed much larger differences among the test points. However, this is partly because of our experimental design that allocated the same participants to all the three learning methods. Since the participants experienced all the learning methods, they could have more time and chance to grow their drumming skill for an experimental day. In contrast, the internal sense of timing obtained from Method PV and PT would also be beneficial at Method P0, which lead an underestimation of guidance effect.

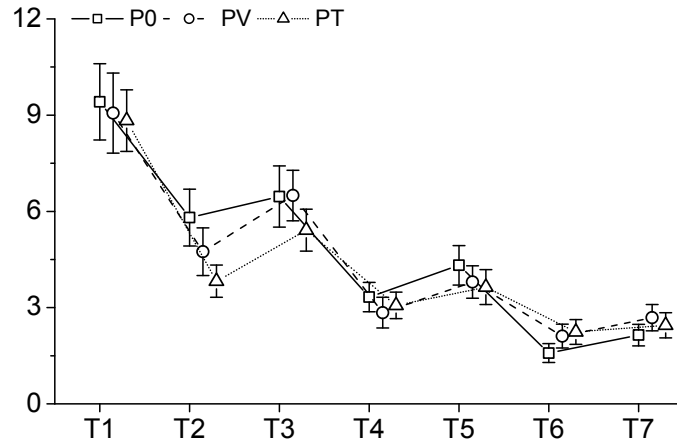
To timing error, visual feedback on the participant's actual performance also had effect to

some extent. The effect was more salient at Method P0, at which no guidance was provided, and reduced the difference between the learning methods. It seems that the participant group F1 made the best use of visual feedback at Method P0 since it was virtually the only means of improvement. Whereas, at Method PV and PT, the group had two different means of improvement, guidance and feedback, and both of the means would not be utilized as much as when they were given alone. This can be supported by the experimental results that significantly less numbers of practices were made by Group F1 at Method PV (5.5) and PT (5.4) than at Method P0 (6.2), and relatively small numbers of guidance requests were made by Group F1 (2.7 at PV and 3.1 at PT) than Group F0 (3.7 and 3.9, respectively).

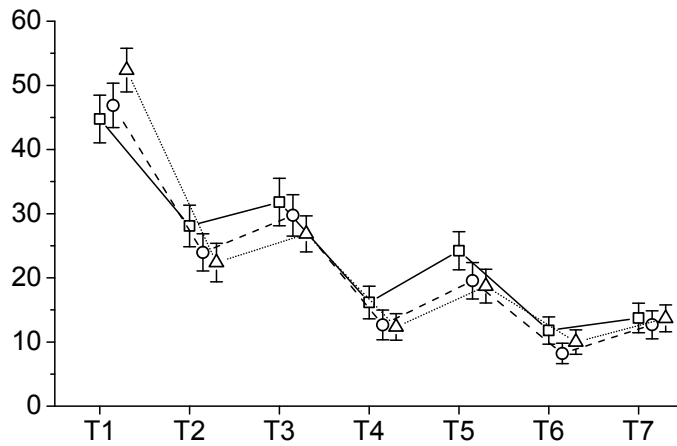
8.3.4 Strength Level

For strength level accuracy, visual feedback on the participant's performance was effective, while video and vibrotactile guidance was not. This is probably because, only with guidance, the participants could discern which note should be played with an accent or not, but not how strong or weak the strike is for an accented or normal note. With visual feedback, the participants could know whether their strike strength was sufficient to each note, and they could adjust their strength in the subsequent practice trial. In the present study, visual feedback was a between-subjects factor, and thus no transfer of its effect to the control condition (Group F0) occurred.

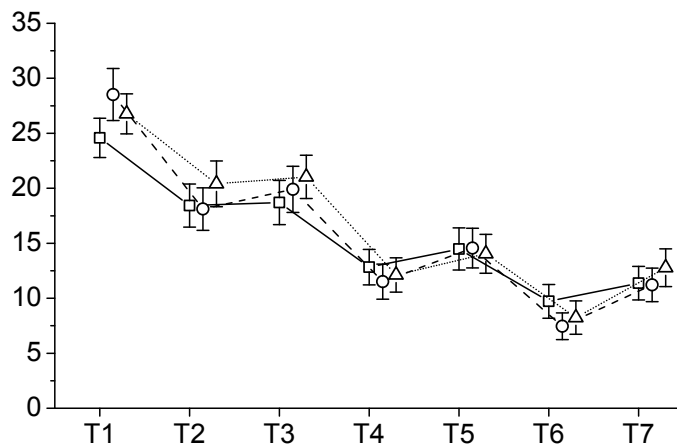
In music sheet, accented notes were easily discriminable from normal notes by an accent mark placed on top of each accented note. This seems another reason of no significant guidance effect in that guidance is usually helpful when the target task is not easily comprehensible. It is expected that the video and vibrotactile guidance would have more effects for the case of no music sheet or for the skills that require different strength levels without explicit notification.



(a) Targeting error ratio (%)



(b) Timing error ratio (%)



(c) Strength error ratio (%)

Fig. 8.3 Performance measures of participants measured at T1–7 for three learning methods.

Chapter 9

Conclusions and Future Work

In this study, a vibrotactile guidance method for complex procedural motor skills was introduced. Each movement for the skill can be specified by a target (or related body part), timing, and strength (or speed) of the movement, and these items were delivered by the location, time, and intensity of a short vibration cue, respectively. For intuitive and correct guidance, our method utilized a natural egocentric mapping from the body site of vibrotactile stimulation to a movement target and the redundant coding of movement strength with the strength and duration of vibrotactile cues. As an application of the guidance method, we developed a vibrotactile guidance system for drumming learning. The system can instruct the learner how to play a drum set using vibrotactile cues generated by nine vibration actuators embedded in six vibrotactile belts worn by the learner. The intensity and duration of the vibration cue was carefully adjusted for the precise delivery of two striking strength levels.

To evaluate and improve our guidance method and system, we conducted a series of human-subject experiments. In the first and second experiments, we tested the accuracy and time of the participant to understand a single vibrotactile cue, and the participant could understand our guidance cue easily and accurately (96.18% accuracy and 0.77 s time). We also tested a situation in which two vibrotactile cues were presented at the same time, and it was not simple to comprehend the cues (55.03% accuracy and 1.11 s time), presumably

due to the sensory illusion and memory limitation. These results suggested that our design can be effective in guiding a successive or repetitive drumming skill that is accomplished by a series of single-limb movements (e.g., linear drum beats or paradiddles) but it is less suitable for more complex skills that involve concurrent movements of multiple limbs (i.e., multi-limb coordination). To increase the recognition accuracy of two simultaneous cues, we changed the layout of vibration actuators so that they can be more apart from each other, and introduced a short stimulus-onset asynchrony when presenting two concurrent vibrotactile cues. Then, we continued to test our system.

To guide a drum rhythm, a sequence of single or multiple vibrotactile cues that consists a drum rhythm needs to be presented and responded at a time. This requires intensive processing and memory efforts of the participant, and the guidance of a rhythm can be unsuccessful even if individual vibrotactile cues are easily recognizable. In this regard, in the third experiment, we asked the participants to respond to a short sequence of single or multiple vibrotactile cues, and measured the response accuracy and task completion time. The accuracy and time greatly depended on the complexity of a cue sequence, and the participants showed 88.4, 56.3, 23.3% of accuracy and 7.53, 10.15, and 13.71 s of task completion time for simple, moderate, and complex cue sequences, respectively. The performance for the complex sequence was rather low, but still acceptable since a large number of response errors were occurred from the imprecise motor control of the participant, not incorrect delivery of guidance information.

Finally, we evaluated our guidance system with a realistic test scenario of drumming learning. Three sets of short drum rhythms were devised, and two groups of participants with different visual feedback conditions learned each rhythm set using one of three learning methods (practice only, practice with video guidance, and practice with vibrotactile guidance). The experimental results indicated that our vibrotactile guidance system is as helpful as video guidance in learning the temporal pattern of a drum rhythm.

Our guidance system showed similar effectiveness to video guidance. However, unlike video guidance, vibrotactile guidance requires the learner to wear on vibrotactile belts, which requires time for preparation and causes inconvenience during learning. In this re-

gard, vibrotactile guidance has lower usability than video guidance, and usability improvement is vital for the practical use of vibrotactile guidance. The use of non-contact tactile display, which provides pressure or tapping sensations from a distance using ultrasound [23] or air vortex [20], instead of a vibration actuator may be a good solution for this.

For the blind, vision-based guidance methods are not applicable, while vibrotactile guidance is not limited to these people as it delivers information through the sense of touch. Moreover, it is known that blind people are much sensitive in touch than normal people. In this regard, vibrotactile guidance is a viable means of helping the blind learn drum rhythms, and its effectiveness is expected to be higher than that for normal people. To examine this, we are considering an experiment that measures the effectiveness of vibrotactile guidance for the blind.

Lastly, in our guidance method, we provided a guidance cue to the body site that is in line with the egocentric direction to the movement target, and this egocentric mapping is shown to be intuitive and effective in guiding the target. However, this does not guarantee that the egocentric mapping is optimal for target guidance, and further study is required that evaluates the egocentric mapping in comparison with other guidance approaches or mappings.

요 약 문

운동기능 학습보조를 위한 촉각 가이드نس: 드럼 리듬 학습에의 응용

최근 경제적, 의학적 발전에 힘입어, 많은 사람들이 조깅, 수영, 악기연주 등 다양한 육체적 활동으로 여가를 즐기고 있다. 육체적 활동, 즉 운동 기능은 일련의 단위 동작을 정해진 순서와 속도로 수행함으로써 이루어지는 경우가 많으며, 이러한 동작의 순서와 흐름을 체득하기 위해서는 많은 노력이 필요하다. 운동 기능을 보다 쉽게 배우기 위한 방법 중 하나는 강사나 교사의 시범을 관찰하고, 이를 따라하는 것이다. 시각이나 청각은 주로 목표 운동 기능을 수행하는데 중요한 역할을 하며, 따라서 시범을 관찰함에 있어 이들 감각을 사용하는 것은 운동 기능의 수행 측면이나 감각 정보의 분산 측면에서 다소 비효율적이다. 촉각을 통한 시범은 시각과 청각을 목표 운동 기능에 온전히 사용하도록 하면서도 시범을 제공할 수 있는 이점이 있으며, 시각장애인처럼 시각적인 시범을 사용할 수 없는 경우에도 활용이 가능한 장점이 있다.

이러한 관점에서, 본 연구에서는 다수의 진동자에서 생성된 진동 큐를 이용해 복잡한 절차적 운동기능 학습을 보조하는 방법을 제시한다. 또한, 제시한 학습 보조 방법의 응용연구로서, 사지의 빠르고 패턴화된 움직임을 요구하는 대표적인 운동기능인 드럼 리듬 연주 기능의 학습을 촉각을 통해 보조하는 시스템을 개발하고 그 효용성을 평가한다. 개발한 시스템은 1인칭 시점에서의 위치적 연관을 바탕으로, 진동 큐가 제공된 신체부위를 달리함으로써 타격해야하는 목표 타악기를 자연스럽게 직관적으로 표현하며, 진동 큐의 강도와 길이를 달리함으로써 목표 타격 강도(2 단계)를

표현한다.

실제 드럼 리듬 학습에의 적용에 앞서, 개발한 시스템의 목표 타악기와 타격강도에 대한 정보 제시 성능을 일련의 사용자 실험을 통해 단계적으로 평가하였다. 사용자 평가 실험에서, 실험참가자들은 단독으로 제시된 진동 큐에 대해 진동 큐의 의미(대응되는 타악기와 타격강도)를 평균 96.18%의 정확도로 0.77 초만에 이해하였으며, 동시에 두 개의 서로 다른 진동 큐가 제공된 경우 각각의 의미를 이해하는 데는 평균 55.03%의 정확도와 1.11 초의 응답시간을 보였다. 한 발 나아가, 한 번에 1-2 개씩 임의의 진동 큐를 순차적으로 4 회 제공하고 응답하는 태스크에 대해서, 참가자들은 단순(88.4%, 7.53 초)하거나 보통 수준(56.3%, 10.15 초)의 시퀀스에 대해서는 참가자들의 진동 큐와 드리밍에 대한 숙련도를 고려할 때 비교적 적절한 수준의 정확도와 이해 속도를 보였다. 복잡한 시퀀스(23.3%, 13.71 초)의 경우에는 다소 성능이 떨어졌으나, 시각 시범에서도 복잡한 운동 기능은 여러 차례의 시범을 필요로 한다는 것을 고려할 때, 동일한 방법을 통해 문제를 해결할 수 있을 것으로 판단되었다.

최종적으로, 개발한 촉각 학습 보조 시스템을 실제 드럼 리듬 학습에 적용하여 그 교육적 효용성을 평가하였다. 참가자들은 3 일간 세 개의 짧은 드럼 리듬 세트를 서로 다른 세 가지 학습 방법(연습만, 연습과 시청각 시범, 연습과 촉각 시범)을 사용해 학습하였으며, 각 방법 하에서의 참가자들의 드럼 리듬 연주 성능을 비교함으로써 각 학습 방법의 성능을 비교하였다. 평가 실험에서, 촉각 시범은 드럼 리듬 연주의 시각적 측면에서 시청각 시범과 매우 유사한 수준의 학습 효용성을 보였으며, 따라서 시청각 시범의 활용이 여의치 않은 경우에 사용할 수 있는 적절한 대안임을 확인할 수 있었다.

REFERENCES

- [1] R. C. Atkinson and R. M. Shiffrin. Human memory: A proposed system and its control processes. *The Psychology of learning and motivation*, 2:89–105, 1968.
- [2] A. D. Baddeley. *Human Memory: Theory and Practice*. Psychology Press, 1997.
- [3] A. Bandura. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice Hall, 1986.
- [4] K. Bark, P. Khanna, R. Irwin, P. Kapur, S. A. Jax, L. J. Buxbaum, and K. J. Kuchenbecker. Lessons in Using Vibrotactile Feedback to Guide Fast Arm Motions. In *Proc. of the IEEE World Haptics Conference*, pages 355–360, 2011.
- [5] B. Bayart, A. Pocheville, and A. Kheddar. An adaptive haptic guidance software module for I-TOUCH: example through a handwriting teaching simulation and a 3D maze. In *Proc. of the IEEE International Workshop on Haptic Audio Visual Environments and Their Applications*, pages 51–56, 2005.
- [6] A. Bloomfield and N. I. Badler. Virtual Training via Vibrotactile Arrays. *Presence: Teleoperators and Virtual Environments*, 17(2):103–120, 2008.

-
- [7] J. Bluteau, S. Coquillart, Y. Payan, and E. Gentaz. Haptic Guidance Improves the Visuo-Manual Tracking of Trajectories. *Plos One*, 3(3), 2008.
- [8] R. W. Cholewiak, J. C. Brill, and A. Schwab. Vibrotactile Localization on the Abdomen: Effects of Place and Space. *Perception & Psychophysics*, 66(6):970–987, 2004.
- [9] R. W. Cholewiak and A. A. Collins. Vibrotactile Localization on the Arm: Effects of Place, Space, and Age. *Perception & Psychophysics*, 65(7):1058–1077, 2003.
- [10] L. M. Crespo and D. J. Reinkensmeyer. Haptic Guidance Can Enhance Motor Learning of a Steering Task. *Journal of Motor Behavior*, 40(6):545–556, 2008.
- [11] F. C. Donders. On the Speed of Mental Processes. *Acta Psychologica*, 30(1):412–431, 1969.
- [12] S. G. Doody, A. M. Bird, and D. Ross. The Effect of Auditory and Visual Models on Acquisition of a Timing Task. *Human Movement Science*, 4(4):271–281, 1985.
- [13] J. L. Emken and D. J. Reinkensmeyer. Robot-Enhanced Motor Learning: Accelerating Internal Model Formation During Locomotion by Transient Dynamic Amplification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13(1):33–39, 2005.
- [14] D. Feygin, M. Keehner, and F. Tendick. Haptic Guidance: Experimental Evaluation of a Haptic Training Method for a Perceptual Motor Skill. In *Proc. of the IEEE Haptics Symposium*, pages 40–47, 2002.
- [15] A. Gallace, H. Z. Tan, P. Haggard, and C. Spence. Short Term Memory for Tactile Stimuli. *Brain Research*, 1190:132–142, 2008.

-
- [16] A. Gallace, H. Z. Tan, and C. Spence. Numerosity Judgments for Tactile Stimuli Distributed Over the Body Surface. *Perception*, 35(2):247–266, 2006.
- [17] R. D. Gilson. Vibrotactile Masking: Some Spatial and Temporal Aspects. *Perception & Psychophysics*, 5(3):176–180, 1969.
- [18] G. Grindlay. Haptic Guidance Benefits Musical Motor Learning. In *Proc. of the IEEE Haptics Symposium*, pages 397–404, 2008.
- [19] H. R. Gudmundsdottir. Advances in Music-reading Research. *Music Education Research*, 12(2):331–338, 2010.
- [20] S. Gupta, D. Morris, S. N. Patel, and D. Tan. AirWave: Non-Contact Haptic Feedback Using Air Vortex Rings. In *Proc. of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 419–428, 2013.
- [21] G. Han, J. Lee, I. Lee, S. Jeon, and S. Choi. Effects of Kinesthetic Information on Working Memory for 2D Sequential Selection Task. In *Proc. of the IEEE Haptics Symposium*, pages 43–46, 2010.
- [22] S. Holland, A. J. Bouwer, M. Dalglish, and T. M. Hurtig. Feeling the Beat Where it Counts: Fostering Multi-Limb Rhythm Skills with the Haptic Drum Kit. In *Proc. of the ACM International Conference on Tangible, Embedded and Embodied Interaction*, pages 21–28, 2010.
- [23] T. Hoshi, M. Takahashi, T. Iwamoto, and H. Shinoda. Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound. *IEEE Transactions on Haptics*, 3(3):155–165, 2010.

- [24] K. Huang, E. Y.-L. Do, and T. Starner. PianoTouch: A Wearable Haptic Piano Instruction System For Passive Learning of Piano Skills. In *Proc. of the IEEE International Symposium on Wearable Computers*, pages 41–44, 2008.
- [25] K. Huang, T. Starner, E. Do, G. Weiberg, D. Kohlsdorf, C. Ahlrichs, and L. Ruediger. Mobile Music Touch: Mobile Tactile Stimulation for Passive Learning. In *Proc. of the ACM Conference on Human Factors in Computing Systems*, pages 791–800, 2010.
- [26] J. C. Huegel and M. K. O’Malley. Progressive Haptic and Visual Guidance for Training in a Virtual Dynamic Task. In *Proc. of the IEEE Haptics Symposium*, pages 343–349, 2010.
- [27] T. Igoe. Getting started on drums featuring tommy igoe, 2001. DVD lecture for drumming learning.
- [28] M. Jalilvand, P. Mokhtari, and M. K. V. Mousavi. The Effect of Visual and Auditory Models in Self Regulated and External Controlled Environment on Learning a Timing Task. *World Applied Sciences Journal*, 16(6):776–780, 2012.
- [29] L. Jiang, R. Girotra, M. R. Cutkosky, and C. Ullrich. Reducing Error Rates with Low-Cost Haptic Feedback in Virtual Reality-Based Training Applications. In *Proc. of the World Haptics Conference*, pages 420–425, 2005.
- [30] P. Kapur, M. Jensen, L. J. Buxbaum, S. A. Jax, and K. J. Kuchenbecker. Spatially Distributed Tactile Feedback for Kinesthetic Motion Guidance. In *Proc. of the IEEE Haptics Symposium*, pages 519–526, 2010.
- [31] R. Kopiez and J. I. Lee. Towards a General Model of Skills Involved in Sight Reading Music. *Music Education Research*, 10(1):41–62, 2008.

-
- [32] E. C. Lechelt. Temporal Numerosity Discrimination: Intermodal Comparisons Revisited. *British Journal of Psychology*, 66(1):101–108, 1975.
- [33] B.-C. Lee, J. Kim, S. Chen, and K. H. Sienko. Cell Phone Based Balance Trainer. *Journal of NeuroEngineering and Rehabilitation*, 9(10), 2012.
- [34] H. Lee, G. Han, I. Lee, S. Yim, K. Hong, H. Lee, and S. Choi. Haptic Assistance for Memorization of 2D Selection Sequences. *IEEE Transactions on Human-Machine Systems*, 43(6):643–649, 2013.
- [35] I. Lee and S. Choi. Vibrotactile Guidance for Drumming: Design and Assessment of Single and Concurrent Cue-Response Tasks. *IEEE Transactions on Haptics*, 2015. under review.
- [36] I. Lee, K. Hong, and S. Choi. Guidance Methods for Bimanual Timing Tasks. In *Proc. of the IEEE Haptics Symposium*, pages 297–300, 2012.
- [37] J. Lee and S. Choi. Effects of Haptic Guidance and Disturbance on Motor Learning: Potential Advantage of Haptic Disturbance. In *Proc. of the IEEE Haptics Symposium*, pages 335–342, 2010.
- [38] Y. Li, J. Huegel, V. Patoglu, and M. K. O’Malley. Progressive Shared Control for Training in Virtual Environments. In *Proc. of the IEEE World Haptics Conference*, pages 332–337, 2009.
- [39] Y. Li, V. Patoglu, and M. K. O’Malley. Negative Efficacy of Fixed Gain Error Reducing Shared Control for Training in Virtual Environments. *ACM Transactions on Applied Perception*, 6(1):3:1–3:21, 2009.

- [40] J. Lieberman and C. Breazeal. TIKL: Development of a Wearable Vibrotactile Feedback Suit for Improved Human Motor Learning. *IEEE Transactions on Robotics*, 23(5):919–926, 2007.
- [41] J. Liu, S. C. Cramer, and D. J. Reinkensmeyer. Learning to perform a new movement with robotic assistance: comparison of haptic guidance and visual demonstration. *Journal of NeuroEngineering and Rehabilitation*, 3(20), 2006.
- [42] P. S. Lum, C. G. Burgar, P. C. Shor, M. Majmundar, and M. V. der Loos. Robot-Assisted Movement Training Compared With Conventional Therapy Techniques for the Rehabilitation of Upper-Limb Motor Function After Stroke. *Archives of Physical Medicine and Rehabilitation*, 83(7):952–959, 2002.
- [43] R. A. Magill. *Motor Learning: Concepts and Applications*. McGraw-Hill, 6th edition, 2001.
- [44] R. A. Magill and B. Schoenfelder-Zohdi. A Visual Model and Knowledge of Performance as Sources of Information for Learning a Rhythmic Gymnastics Skill. *International Journal of Sport Psychology*, 27(1):7–22, 1996.
- [45] L. Marchal-Crespo and D. J. Reinkensmeyer. Review of Control Strategies for Robotic Movement Training after Neurologic Injury. *Journal of NeuroEngineering and Rehabilitation*, 6(20), 2009.
- [46] F. Maydwell. *Sight Reading Skills: A Pianist's Guide for Learning to Read Music Accurately and Expressively*. The New Arts Press of Perth, 2003.
- [47] S. Nicolas. On the Speed of Different Senses and Nerve Transmission by Hirsch (1862). *Psychological Research*, 59(4):261–268, 1997.

- [48] M. K. O'Malley, A. Gupta, M. Gen, and Y. Li. Shared Control in Haptic Systems for Performance Enhancement and Training. *Journal of Dynamic Systems, Measurement, and Control*, 128(1):75–85, 2006.
- [49] R. Palluel-Germain, F. Bara, A. H. de Boisferon, B. Hennion, P. Gouagout, and E. Gentaz. A Visuo-Haptic Device - Telemaque - Increases Kindergarten Children's Handwriting Acquisition. In *IEEE World Haptics Conference*, pages 147–152, 2007.
- [50] G. Park and S. Choi. Perceptual Space of Amplitude-Modulated Vibrotactile Stimuli. In *Proc. of the IEEE World Haptics Conference*, pages 59–64, 2011.
- [51] 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.
- [52] D. Powell and M. K. O'Mally. The Task-Dependent Efficacy of Shared-Control Haptic Guidance Paradigms. *IEEE Transactions on Haptics*, 5(3):208–219, 2012.
- [53] Railroad Media, Inc. Freedrumlessons.com. Online drum tutoring service. <http://www.freedrumlessons.com/>.
- [54] D. J. Reinkensmeyer and J. L. Patton. Can Robots Help the Learning of Skilled Actions? *Exercise and Sport Sciences Reviews*, 37(1):43–51, 2009.
- [55] Roland Corporation U.S. DT-1 V-Drums Tutor, 2012. A drumming tutoring software. <http://www.rolandus.com/products/details/1213>.
- [56] B. Román-Viñas, L. Serra-Majem, L. Ribas-Barba, E. Roure-Cuspinera, C. Cabezas, C. Vallbona, and A. Plasencia. Trends in Physical Activity Status in Catalonia, Spain (1992–2003). *Public Health Nutrition*, 10(11):1389–1395, 2007.

-
- [57] M. A. Rosen, E. Salas, D. Pavlas, R. Jensen, D. Fu, and D. Lampton. Demonstration-Based Training: A Review of Instructional Features. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 52(5):596–609, 2010.
- [58] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1):43–49, 1978.
- [59] R. A. Schmidt. Frequent Augmented Feedback Can Degrade Learning: Evidence and Interpretations. *Tutorials in Motor Neuroscience*, pages 59–75, 1991.
- [60] D. Sherrill. Learn & master drums with dan sherrill, 2010. DVD lecture for drumming learning.
- [61] P. Shull, K. Lurie, M. Shin, T. Besier, and M. Cutkosky. Haptic Gait Retraining for Knee Osteoarthritis Treatment. In *Proc. of the IEEE Haptics Symposium*, pages 409–416, 2010.
- [62] R. Sigrist, G. Rauter, R. Riener, and P. Wolf. Augmented Visual, auditory, haptic, and multimodal feedback in motor learning: A review. *Psychonomic Bulletin & Review*, 2012 (online preprint).
- [63] D. Spelmezan, M. Jacobs, A. Hilgers, and J. Borchers. Tactile Motion Instructions for Physical Activities. In *Proc. of the ACM Conference on Human Factors in Computing Systems*, pages 2243–2252, 2009.
- [64] C. Sung and M. K. O’Malley. Effect of Progressive Visual Error Amplification on Human Motor Adaptation. In *Proc. of the IEEE International Conference on Rehabilitation Robotics*, 2011.

- [65] D. E. Thompson and J. Russell. The Ghost Conditions: Imitation Versus Emulation in Young Children's Observational Learning. *Developmental Psychology*, 40(5):882–889, 2004.
- [66] E. Todorov, R. Shadmehr, and E. Bizzi. Augmented Feedback Presented in a Virtual Environment Accelerates Learning of a Difficult Motor Task. *Journal of Motor Behavior*, 29(2):147–158, 1997.
- [67] E. Tulving. How Many Memory Systems Are There? *American Psychologist*, 40(4):385–398, 1985.
- [68] J. van der Linden, R. Johnson, J. Bird, Y. Rogers, and E. Schoonderwaldt. Buzzing to Play: Lessons Learned From an In the Wild Study of Real-time Vibrotactile Feedback. In *Proc. of the ACM Conference on Human Factors in Computing Systems*, pages 533–542, 2011.
- [69] J. B. F. van Erp. Vibrotactile Spatial Acuity on the Torso: Effects of Location and Timing Parameters. In *Proc. of the IEEE World Haptics Conference*, pages 80–85, 2005.
- [70] R. T. Verrillo, G. A. Gescheider, B. G. Calman, and C. L. V. Doren. Vibrotactile Masking: Effects of One- and Two-site Stimulation. *Perception & Psychophysics*, 33(4):379–387, 1983.
- [71] J. Watanabe and H. Ando. Pace-sync Shoes: Intuitive Walking-pace Guidance Based on Cyclic Vibro-tactile Stimulation for the Foot. *Virtual Reality*, 14(3):213–219, 2010.
- [72] R. O. Weagley and E. Huh. The Impact of Retirement on Household Leisure Expenditures. *Journal of Consumer Affairs*, 38(2):262–281, 2004.

-
- [73] C. K. Williams and H. Carnahan. Motor Learning Perspectives on Haptic Training for the Upper Extremities. *IEEE Transactions on Haptics*, 7(2):240–250, 2013.
- [74] C. J. Winstein, P. S. Pohl, and R. Lewthwaite. Effects of Physical Guidance and Knowledge of Results on Motor Learning: Support for the Guidance Hypothesis. *Research Quarterly for Exercise and Sport*, 65(4):316–323, 1994.
- [75] C. J. Winstein and R. A. Schmidt. Reduced Frequency of Knowledge of Results Enhances Motor Skill Learning. *Journal of Experimental Psychology*, 16(4):677–691, 1990.
- [76] I. J. Wuyts and M. Buekers. The effects of visual and auditory models on the learning of a rhythmical synchronization dance skill. *Research Quarterly for Exercise and Sport*, 66(2):105–115, 1995.
- [77] T. Yamabe, H. Asuma, S. Kiyono, and T. Nakajima. Feedback Design in Augmented Musical Instruments: A Case Study with An AR Drum Kit. In *Proc. of the IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*, pages 126–129, 2011.

Acknowledgements

감사의 글

부푼 꿈을 안고 대학원에 온 것이 바로 어제인 것만 같은데 벌써 9년이 지나 어느덧 졸업의 때를 맞이하게 되었습니다. 오늘의 기쁨은 저 혼자만의 힘으로 이룬 것이 아니라 모두의 사랑과 격려가 있었기에 가능했다고 생각합니다.

방향을 잡지 못하고 힘들어할 때마다 이끌어주시고 조언을 아끼지 않으신 지도교수님, 뒤에서 묵묵히 받쳐주시고 감싸주신 어머니와 아버지, 장모님과 장인어른, 그리고 무엇보다도 항상 믿어주고 따뜻한 마음의 안식처가 되어준 사랑하는 아내에게 감사의 말을 전합니다. 입학하는지 얼마 안 되어 모든 것이 낯설 때 도와주고 편안한 말상대가 되어주었던 인욱이, 오랜 재학기간 동안 많은 일을 함께 나눈 성훈이 형과 건혁이, 인생의 가치관과 삶에 대해 많은 것을 가르쳐 준 재봉이 형, 힘든 학업의 와중에 웃음을 잃지 않게 해준 종만이, 호진이, 그리고 성환이, 연구실의 대소사를 책임져준 용재와 송이 씨, 작은 먹거리들로 기쁨을 나누어 준 Reza와 그의 아들 Parsa, 함께 졸업 준비를 하면서 서로 도움이 될 수 있었던 Phuong, 그리고 짧은 기간이지만 대학원 생활의 마지막을 즐겁게 보낼 수 있게 해준 막내 호준이에게도 고마운 마음을 표합니다. 어려울 때 가르쳐주고 함께해 준 성길이 형, 종현이 형, 석희 형, 재영이 형, 재훈이 형, 채현이, 경표, 명찬이를 비롯해 먼저 사회에 나가 왕성하게 활동하고 있는 여러 연구실 선후배들에게도 감사를 드립니다. 마지막으로, 김미자 선생님, 장혜자 선생님, 조동현 선생님을 비롯한 여러 학교 및 학과 사무실 여러분들께도 감사를 전합니다.

이분들의 도움이 있었기에 힘든 학업의 과정을 헤쳐나갈 수 있었습니다. 다시 한번 감사의 인사를 전하며, 모든 분의 앞날이 밝게 빛나기를 기원합니다.

Publications

International Journals

1. **In Lee** and Seungmoon Choi, "Vibrotactile Guidance for Drumming: Design and Assessment of Single and Concurrent Cue-Response Tasks," *IEEE Transactions on Haptics*, 2015 (in review).
2. **In Lee** and Seungmoon Choi, "Discrimination of Visual and Haptic Rendering Delays in Networked Environments," *Int. J. Control, Automation, and Systems*, Vol. 7, No. 1, pp. 25–31, 2009.
3. Hojin Lee, Gabjong Han, **In Lee**, Sunghoon Yim, Kyungpyo Hong, Hyeseon Lee, and Seungmoon Choi, "Haptic Assistance for Memorization of 2D Selection Sequences," *IEEE Transactions on Human-Machine Systems*, Vol. 43, No. 6, pp. 643–649, 2013.
4. Jaemin Chun, **In Lee**, Gunhyuk Park, Jongman Seo, Seungmoon Choi, and Sung H. Han, "Efficacy of Haptic Blind Spot Warnings Applied through a Steering Wheel or a Seatbelt," *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 21, pp. 231–241, 2013.
5. Jaemin Chun, Sung H. Han, Gunhyuk Park, Jongman Seo, **In Lee**, and Seungmoon Choi, "Evaluation of Vibrotactile Feedback for Forward Collision Warning on the Steering Wheel and Seatbelt," *International Journal of Industrial Ergonomics*, Vol. 42, No. 5, pp. 443–448, 2012.

International Conferences

1. **In Lee** and Seungmoon Choi, "Vibrotactile Guidance for Drumming Learning: Method and Perceptual Assessment," in *Proc. IEEE Haptics Symposium*, pp. 147–152, 2014 (Long oral presentation; acceptance rate 7.6%,; Candidates for the best paper award).
2. **In Lee** and Seungmoon Choi, "Effects of Multi-modal Guidance for the Acquisition of Sight Reading Skills: A Case Study with Simple Drum Sequences," in *Proc. of the IEEE World Haptics Conference (WHC)*, pp. 571–576, 2013.
3. **In Lee**, Kyungpyo Hong, and Seungmoon Choi, "Guidance Methods for Bi-manual Timing Tasks," in *Proc. IEEE Haptics Symposium*, pp. 297–300, 2012.
4. **In Lee**, Inwook Hwang, Kyung-Lyong Han, Oh Kyu Choi, Seungmoon Choi, and Jin S. Lee, "System Improvements in Mobile Haptic Interface," in *Proc. Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (World Haptics)*, pp. 109–114, 2009 (Winner of the best student paper).
5. **In Lee** and Seungmoon Choi, "Discrimination of Virtual Environments Under Visual and Haptic Rendering Delays," in *Proc. Int. Conf. Frontiers in the Convergence of Bioscience and Information Technologies (FBIT; A Special Session for Haptics and its Application to Bioscience)*, pp. 554–559, 2007.
6. Jaemin Chun, Gunhyuk Park, Seungwhan Oh, Jongman Seo, **In Lee**, Seungmoon Choi, Sung H. Han, and Wongyu Park, "Evaluating the Effectiveness

- of Haptic Feedback on a Steering Wheel for BSW,” in *Proc. 4th Int. Conf. Applied Human Factors and Ergonomics*, pp. 2047–2054, 2012.
7. Sunghwan Shin, **In Lee**, Hojin Lee, Gabjong Han, Kyungpyo Hong, Sunghoon Yim, Jongwon Lee, YoungJin Park, Byeong Ki Kang, Dae Ho Ryoo, Dae Whan Kim, Seungmoon Choi, and Wan Kyun Chung, “Haptic Simulation of Refrigerator Door,” in *Proc. IEEE Haptics Symposium*, pp. 147–154, 2012 (Best paper nominated).
 8. Hojin Lee, Gabjong Han, **In Lee**, Sunghoon Yim, Kyungpyo Hong, and Seungmoon Choi, “Effect of Active and Passive Haptic Sensory Information on Memory for 2D Sequential Selection Task,” in *Proc. Int Symp. Ubiquitous Virtual Reality (ISUVR)*, pp. 52–54, 2011.
 9. Jaemin Chun, Gunhyuk Park, Seunghwan Oh, Jongman Seo, **In Lee**, Seungmoon Choi, Sung H. Han, Woochul Park, “Development of Human Factors Design Guidelines for Haptic Collision Warning Systems,” in *Proc. IIE Asian Conference*, pp. 249–254, 2011.
 10. Gabjong Han, Jaebong Lee, **In Lee**, Seokhee Jeon, and Seungmoon Choi, “Effects of Kinesthetic Information on Working Memory for 2D Sequential Selection Task,” in *Proc. IEEE Haptics Symposium*, pp. 43–46, 2010 (Oral presentation; Extended abstract; Acceptance rate = 18.7
 11. Seungmoon Choi and **In Lee**, “Research Issues on Mobile Haptic Interface for Large Virtual Environments,” in *Proc. Int. Conf. Advanced Mechatronics (ICAM)*, pp. 626–630, 2010.
 12. 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 *Proc. Int. Conf. Control, Automation, and Systems (ICCAS)*, pp. 1290–1295, 2008.

13. Chaehyun Lee, Min Sik Hong, **In Lee**, Oh Kyu Choi, Kyung-Lyong Han, Yoo Yeon Kim, Seungmoon Choi, and Jin S. Lee, "Mobile Haptic Interface for Large Immersive Virtual Environments: PoMHI v0.5," in *Proc. Int. Conf. Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 106–111, 2007.

Domestic Papers

1. 이인, 최승문, "멀티모달 가이던스가 독보 기능 습득에 미치는 영향: 드럼 타격 시퀀스에서의 사례 연구," *한국로봇학회 논문지*, Vol. 8, No. 2, pp. 217–227, 2013 (초청논문).
2. 이인, 최승문, "이동형 햅틱 장치의 사용성 개선," *한국 HCI 학술대회 논문집*, pp. 199–202, 2010.
3. 이인, 황인욱, 한경룡, 최오규, 이진수, 최승문, "이동형 햅틱 장치의 실제적 문제점과 그 향상 방안," *한국 HCI 학술대회 논문집*, pp. 390–395, 2009.
4. 이호진, Reza Haghighi Osgouei, 이인, 신성환, 최승문, "가상 운전 환경에서의 힘 피드백 제공을 위한 햅틱 가속 페달," *한국로봇종합학술대회 논문집*, pp. 257–258, 2013.
5. 홍경표, 이인, 한갑종, 최승문, "다중 감각 도움을 이용한 드럼 연주 기능 교육 시스템," *한국 HCI 학술대회 논문집*, pp. 161–163, 2011.
6. 이호진, 한갑종, 이인, 임성훈, 홍경표, 최승문, "이차원 순차 선택에서의 능동

- 및 수동 햅틱 정보 제공을 통한 기억력 향상,” 한국 *HCI* 학술대회 논문집, pp. 248–250, 2011.
7. 이재봉, 이인, 한갑종, 전석희, 최승문, “이차원 순차 선택에서의 햅틱 정보 제공을 통한 단기기억능력 향상,” 한국 지능로봇 종합 학술대회 논문집, pp. 455–457, 2009.
 8. 이채현, 홍민식, 이인, 최오규, 한경룡, 김유연, 최승문, 이진수, “대형 가상환경을 위한 이동형 햅틱 인터페이스: PoMHI v0.5,” *로봇공학회 논문지*, Vol. 3, No. 2, pp. 137–145, 2008 (초청논문).
 9. 정재훈, 황인욱, 이인, 이채현, 박건혁, 황재인, 최승문, 김정현, “동작기반의 체험형 리모트 컨트롤,” 한국 *HCI* 학술대회 논문집, pp. 115–122, 2007.
 10. 이채현, 이인, 최승문, “이동형 햅틱 디스플레이를 위한 동작 계획,” 한국 *HCI* 학술대회 논문집, pp. 578–584, 2007.