The Impact of Haptic Guidance on Musical Motor Learning

by

Graham C. Grindlay

M.S., University of California, Santa Cruz (2005)
B.A., Oberlin College (2000)

Submitted to the Program in Media Arts and Sciences,
Massachusetts Institute of Technology,
in partial fulfillment of the requirements for the degree of

Master of Science in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2007

Author ____
Program in Media Arts and Sciences
August 20, 2006

Certified by ____
Tod Machover
Professor of Media Arts and Sciences
Program in Media Arts and Sciences
Thesis Supervisor

Accepted by ____
Professor Deb Roy
Chairperson, Departmental Committee on Graduate Students
Program in Media Arts and Sciences

The Impact of Haptic Guidance on Musical Motor Learning

by

Graham C. Grindlay

Submitted to the Program in Media Arts and Sciences,
Massachusetts Institute of Technology,
on August 20, 2006, in partial fulfillment of the
requirements for the degree of
Master of Science in Media Arts and Sciences

Abstract

Skilled musical performance provides one of the best demonstrations of the upper limits of the human motor system's capabilities. It is therefore not surprising that learning to play an instrument is a long and difficult process. Teachers and education researchers alike have long since recognized that while learning rate is dependent on the quantity of practice, perhaps even more important is the *quality* of that practice. However, for non-trivial skills such as music performance, just gaining an understanding of what physical movements are required can be challenging since they are often difficult to describe verbally. Music teachers often communicate complex gesture by physically guiding their students' hands through the required motions. However, at best, this gives a rough approximation of the target movement and begs the question of whether technology might be leveraged to provide a more accurate form of physical guidance. The success of such a system could lead to significant advancements in music pedagogy by speeding and easing the learning process and providing a more effective means of home instruction.

This thesis proposes a "learning-by-feel" approach to percussion instruction and presents two different systems to test the effect of guidance on motor learning. The first system, called the *FielDrum*, uses a combination of permanent and electromagnets to guide a player's drumstick tip through the motions involved in the performance of arbitrary rhythmic patterns. The second system, called the *Haptic Guidance System*, uses a servo motor and optical encoder pairing to provide precise measurement and playback of motions approximating those involved in snaredrum performance. This device was used in a pilot study of the effects of physical guidance on percussion learning. Results indicate that physical guidance can significantly benefit recall of both note timing and velocity. When subject performance was compared in terms of note velocity recall, the addition of haptic guidance to audio-based training produced a 17% reduction in final error when compared to audio training alone. When performance was evaluated in terms of timing recall, the combination of audio and haptic guidance led to an 18% reduction in early-stage error.

Thesis Supervisor: Tod Machover
Title: Professor of Media Arts and Sciences, Program in Media Arts and Sciences

The Impact of Haptic Guidance on Musical Motor Learning

by

Graham C. Grindlay

The following people served as readers for this thesis:

Thesis Reader____

Sile O'Modhrain
Lecturer in Haptics and Acoustics
Queens University, Belfast.

Thesis Reader____

Ed Boyden
Assistant Professor of Media Arts and Sciences
Program in Media Arts and Sciences

Acknowledgments

Thanks to:

My thesis committee for their helpful comments and suggestions.

Friends and colleagues in the MIT community.

Everyone in the Hyperinstruments/Opera of the Future Group.

My family for encouraging me throughout my life.

Rachel for her support, friendship, and love.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Humans acquire new motor skills through a multi-stage learning process [14, 15], central to which is of course practice. Through the process of trial and error, we continually refine our motor skills achieving better and more consistent performance. But in order for practice to be productive, a reliable means of evaluating performance is required, making feedback of paramount importance. Feedback can take many different forms in motor learning applications, including verbal communication (i.e. knowledge of results), visual and auditory signals [27], and vibrotactile stimulation [28]. Although different in sensory modality, these examples all share the common trait of being indirect forms of feedback. That is, the information that they provide about performance must be translated into the proprioceptive coordinate system. For simple tasks, this translation may not be a significant difficulty, but for more complex tasks it may be overwhelming, particularly in the early stages of learning. Another, more direct, form of feedback is physical guidance where the learner is physically moved along the trajectory of the target motion. Because the learner experiences the exact proprioceptive feedback that he/she would during a correct execution of the target task, no translation is necessary. Of course this raises the question of whether physical guidance could become a crutch and actually *hinder* learning. This is an important concern that is at least partly addressed by interspersing guided runs through the target motion with unguided runs and is discussed at length in Chapter 2.

The term “Physical Guidance” has come to refer to many different things over the last century of research, including pushing and pulling a learner’s body through a trajectory, simply providing physical constraints to prevent a learner from moving incorrectly, as well as others [39]. Therefore, to avoid confusion, this thesis will follow Feygin et al. [12] and use the term “Haptic Guidance” to refer to the process of physically moving a learner’s body through an example of a motion to be learned.

1.1 Motivation & Applications

Learning to play a musical instrument is one of the most difficult as well as most rewarding human pursuits. Mastering an instrument takes years of practice and hard work, both on the part of the student as well as the teacher. It is therefore imperative that these efforts be carried out in as efficient a manner as possible. For good progress to be made, it is perhaps *more* important that a student practice well than it is that he or she practice often. Although ensuring that students develop good technique and practice habits is largely the job of the teacher, they cannot be present for more than a small fraction of the time that a student spends playing his or her instrument. Moreover, due to the implicit nature of motor skills, teachers cannot directly teach or assess motor programs. The best that they can do is provide correction and guidance in the form of high-level feedback. While some of this feedback is verbal, most aspects of performance instruction necessarily relate to physical skills and are therefore extremely difficult to describe. Therefore, many teachers physically move their students through correct example trajectories to provide more direct instruction. This type of guided instruction is used in the teaching of other physical skills as well, such as sports and dance. But these forms of haptic guidance provide only rough approximations of the target movement; most or all of the finer aspects are absent or even worse, incorrect. This begs the question of whether technology might be leveraged in order to provide a more accurate and more consistent means of haptic guidance. The success of a haptic guidance system could lead to significant advancements in musical pedagogy by speeding/easing the learning process and providing a more effective means of home instruction. Many other areas involving the transfer of physical skill from one person to another could benefit from this research as well. One could imagine guidance systems for teaching sports such as tennis and golf or even dance. There are even potential applications in rehabilitation and physical therapy where patients need to relearn lost motor skills.

1.2 Scope

This thesis proposes a *learning-by-feel* approach to percussion training. The design and construction of two different devices that are able to physically guide a student/player are presented. Experimental evaluation of one of the systems is also presented, including protocol design, experimental results, and analysis of the findings.

The first system, called the *FielDrum*, uses a combination of permanent magnets and electromagnets to guide a player's drumstick tip through the motions involved in the performance of arbitrary rhythmic patterns. The drum itself contains a large electromagnet just under the drum head while the drumsticks contain permanent magnets in their tips. By switching the polarity of the electromagnet from one direction to the other, it will either attract or repell the drumstick tip, producing a "self-playing" or guided effect. The primary advantage of this system is the untethered nature of the device due to the action-at-a-distance property of magnets. This allows a natural interaction with the *FielDrum* as the interface is largely identical to a traditional drum.

The second device, called the Haptic Guidance System (HAGUS), was designed and built with very different goals. It was designed specifically for use as an experimental tool to investigate what effects, if any, haptic guidance has on percussion learning. To do this, it uses a servo motor and optical encoder pairing to provide precise measurement and playback of drumstick motions approximating those involved in snare drum performance. It was used in a pilot study of the effects of haptic guidance on percussion learning, the results of which, along with discussion of their implications for the design of teaching systems, are presented in Chapter 4.

1.3 Contributions

There are several contributions made by this thesis. First, the work examines how percussion learning is affected by haptic guidance. The pilot experiment conducted with the HAGUS device elucidates the different roles played by the auditory and haptic sensory modalities in a non-trivial and perhaps more importantly, *ecologically valid*, motor learning context. This stands in contrast to previous research on haptic guidance which has not considered musical tasks or even tasks involving the auditory system at all.

Second, this thesis presents two novel hardware devices, each of which represents a significant engineering contribution. The software and hardware designed and built as part of HAGUS provide a means for measuring, recording, and synthesizing, to a high degree of spatial and temporal accuracy, the primary motions involved in snare drum performance. The *FielDrum* represents a novel contribution to the musical controller/instrument field as perhaps the first physically-assisted percussion interface.

Finally, this work suggests some goals for the design and construction of musical pedagogy systems based on haptic guidance.

1.4 Thesis Outline

This thesis is organized as follows: Chapter 2 reviews previous research relevant to the current work. Motor behavior literature pertaining to guidance, feedback, and learning is discussed and reviewed as is research on musical controllers that make use of haptic information. Chapter 3 describes the two devices that were built as part of the current work. In addition to detailed descriptions of the hardware and software designs of the *FielDrum* and HAGUS devices, the goals, limitations, and possible extensions are also discussed. Chapter 4 presents the experiment that was conducted using the HAGUS device. Along with a presentation of the experimental design, motivation, and protocol, results are discussed and interpreted. Finally, Chapter 5 summarizes the thesis as a whole and discusses some ideas for future work.

Chapter 2

Background & Related Work

2.1 Motor Learning and Guidance

The past century has seen significant gains in our understanding of how humans learn and perform physical skills. Motor behavior research flourished in the post-war period largely due to the U.S. Air Force's interest in identifying skilled candidates for pilot training [39]. While the next half-century brought a much better understanding of motor control on a basic physiological as well as conceptual level, advancements in motor learning research were much more asymmetric, with higher-level conceptual theories constituting the bulk of the work. Limitations of imaging technologies have made examination and understanding of the underlying physiological processes involved in human motor learning (and indeed many other types of learning as well) difficult, although many recent advancements have been made in this area [9, 8, 7].

Of the many theories of motor learning proposed over the years, Fitts' theory of the stages of motor learning is particularly relevant to our discussion [14, 15]. It provides a framework in which to think about how people learn new motor skills and at what point in the learning process augmented feedback or haptic assistance might be most beneficial. According to Fitts, there are three distinct phases of motor learning: the *cognitive phase*, the *associative phase*, and the *autonomous phase*. The *cognitive phase* involves developing an understanding of what is involved in the task to be performed and how to evaluate its performance. This phase is largely exploratory and the most dramatic gains in performance occur at this stage of learning. As such, it is the best candidate for augmented learning techniques such as haptic guidance. The *associative phase*, which can last for many days or weeks, is largely concerned with refining the baseline skills acquired during the *cognitive phase*. Performance becomes more consistent as the learner settles on a strategy. The *autonomous phase* takes place on a much longer time scale (months or even years) than the previous two phase. At this point, the motor skill has become largely automatic and requires very little attention to perform.

2.1.1 Feedback

Along with practice, feedback is probably the single most important component of the learning process [4]. Without feedback a learner cannot evaluate his or her performance and therefore cannot improve. Perhaps of primary importance is the intrinsic proprioceptive feedback that we all use to evaluate our movements. This information is invaluable for learning new motor skills as well as ensuring correct execution of ones we may already know. However, there are other, *augmented*, forms of feedback that play an extremely important role in the learning process as well. Strictly defined, augmented feedback is information presented about a task that is supplemental to, or that augments, inherent feedback [39]. It can take different sensory forms (i.e. verbal, visual, haptic) and can occur on different delivery timescales (i.e. concurrent or delayed),

Verbal feedback from a teacher is a commonly used form of augmented feedback. Two important subclasses of verbal feedback are *knowledge of results* (KR) and *knowledge of performance* (KP). KR refers to verbal or verbalizable postmovement feedback about the outcome of the movement in terms of the goal [39]. While KR refers specifically to verbal feedback given about goal achievement, KP refers to feedback given about the movement pattern [39]. If, for example, during a percussion lesson a teacher told her student, "You played the rhythm incorrectly", she would be providing KR feedback, while if she were to say, “The second note you played was an eighth note rather than a dotted-sixteenth”, she would be providing KP feedback.

Both of these types of augmented verbal feedback (KR in particular) have been well studied in the motor learning literature. Although research in KP has been somewhat less extensive than research in KR, the research does indicate that KP is most effective when the information provided cannot be derived from other (intrinsic) sources [39].

Feedback Frequency and the Guidance Hypothesis

The past two decades of motor learning research have seen a surge in the study of KP and KR feedback. Much of this work was stimulated by counter-intuitive findings regarding the effects of KR presentation frequency on performance and retention. Hagman [20] found that practicing a task with relatively infrequent (i.e. low percentage of total trials) KR feedback led to worse performance, but better retention than practice with relatively frequent KR feedback. These results, which have been duplicated and expanded upon in a number of other studies, seem to fly in the face of much of what conventional motor learning theory tells us about more feedback being better [40, 44, 41, 44, 27].

The *guidance hypothesis* [37] was proposed to explain these findings. It hypothesizes that while augmented feedback such as KR has useful informational properties that help to correct errors and improve performance on subsequent trials, it may be detrimental to long-term storage by hampering critical between-trial encoding processes [45]. Thus, we seem to be in a losing position no matter what we do: if feedback is presented often, performance is good, but retention is poor, but if feedback is presented infrequently, the opposite occurs.

One possible way out of this seemingly paradoxical situation is to use a “fading” feedback schedule [47, 38, 45]. In this paradigm, high-frequency feedback is presented early in the training schedule and then gradually withdrawn. Studies using this “fade” technique suggest that it can support both short-term performance and long-term retention [44].

The Role of Task Complexity

As mentioned previously, the degree to which augmented feedback provides information that would not otherwise be available to the learner is one of the most important factors in determining how useful the addition of augmented feedback will be. Task complexity is another closely related variable that can change the effect of augmented feedback on the learning process. Schmidt et al. conclude that in terms of KR summary window size (the number of trials that KR gives information about), the optimal length is related to the amount of information provided by the summary and that this is largely determined by task complexity [38]. Guadagnoli et al. provide support for this conclusion as well [19]. They found that for simple tasks the optimal KR summary length was long (15 trials), while for complex tasks the optimal summary length was short (1 trial).

Task complexity appears to be important for KP feedback as well. Physical guidance, which is discussed in detail in Section 2.1.2, can be considered a form of KP where kinetic feedback is delivered concurrent to the execution of the movement. While some researchers have found little or no benefit to physical guidance when used with relatively simple tasks [49], other researchers have found beneficial effects if the task movements are complex [48].

2.1.2 Physical Guidance

The term "physical guidance" has been used in the literature to refer to several different forms of augmented feedback. Two of the most common types of physical guidance, *constraint-based guidance* and *haptic guidance*, are review below with an emphasis on the latter.

Constraint-Based

What is referred to here as *constraint-based* guidance is a form of concurrent augmented feedback where the learner has temporal control over movement, but is spatially constrained in some way. Holding and Macrae used a manual positioning task to compare KR, constraint-based guidance, and haptic guidance [21]. They found all training conditions to have considerable advantage over the uncorrected and unguided control condition in terms of short-term recall performance. In studies of positioning tasks where subjects had to learn to move a manipulandum to a specific point in space, researchers found that training with relatively frequent constraint-based guidance produced degraded retention performance [20, 45]. These findings are similar to those of KR-based feedback frequency research, providing further support for the guidance hypothesis. Given these similarities, once might expect the effects of task complexity found with KR to hold for constraint-based guidance. The previously mentioned work of Wulf et al. suggests this to be the case [48]. Using a ski slalom task where subjects either trained with or without ski poles (which provided a form of constraint-based guidance), they found that the use of the pole-based guidance significantly benefited both task performance and retention.

Haptic Guidance

Haptic guidance, sometimes called mechanical guidance, manual guidance, as well as other names in the literature, refers to concurrent augmented feedback where the learner is moved, both temporally and spatially, through an ideal rendition of the task motion.

Early research in this area generally made use of fairly simple reaching or linear positioning tasks [21, 24]. One notable exception is the work of Armstrong who used a complex elbow movement task to compared haptic guidance, KP delivered concurrently using a visual display, and KP delivered at the end of each trial [3]. He found that while the physical guidance and concurrent KP training conditions had superior performance during the trials, they were worse than terminal KP in a retention test. It should be noted, however, that in Armstrong's study, each of the training conditions used 100% relative feedback frequency.

Later research largely focused on comparing haptic guidance to guidance based on other sensory modalities (almost exclusively vision). Yokokohki et al. proposed several different combinations of haptic and visual guidance as part of a record-and-playback system that they called, "What You See Is What You Feel." [49] Although they did not conduct formal experiments, a very preliminary test using a virtual cube manipulation task did not yield any conclusive results. They speculate that this may have been due to the task being overly easy.

Gillespie et al. developed a system called the *Virtual Teacher* to test haptic guidance in a crane-moving task [18]. This device consisted of a free-swinging pendulum attached to a cart which could be slid along a track. The task involved setting the pendulum into motion by moving the cart and then trying to stop the pendulum from swinging as quickly as possible. The optimal movement strategy, which involves first injecting energy into the system and then removing it after a carefully timed interval, was demonstrated to some subjects while others (the control group) simply tried to learn the system dynamics on their own. Although they did not observe any statistically significant advantage of the guidance-trained groups over the control group, guidance did seem to effectively communicate the basic components of the optimal strategy. The authors conjecture that the optimal strategy was probably too difficult to master and that better results might have been had if the *Virtual Teacher* had demonstrated the components of the optimal strategy individually.

Several recent studies have compared the effects of haptic and visual guidance for learning. Feygin et al. looked at these types of guidance using complex sinusoidal task movements [12]. Subjects learned three-dimensional spatial trajectories under several different training conditions (haptic, visual, haptic and visual) and then had to manually reproduce them under two different unassisted recall conditions (with vision, without vision). The experiment contained 15 trials for each combination of training and recall conditions where each trial consisted of two training (presentation) runs followed by a test (recall) run. Performance was measured during each of the recall runs using several different error metrics, including position, shape, timing, and drift. They found that subjects significantly improved their performance in all training conditions under the position and shape metrics, but not under the drift or timing metrics. In terms of performance averaged over the last five trials, haptic training alone was less effective than visual training under the position and shape metrics, but more effective under the timing metric. Recall mode only affected timing and drift (marginally) metrics with the addition of vision benefiting performance. Training and recall mode were found to interact such that performance under haptic training modes decreased when vision was included in the recall condition. The authors suggest that this interaction may be because vision overpowers proprioception, degrading its effect. A separate analysis of haptic guidance and visual training indicated that while position and shape accuracy were predominantly affected by vision, timing accuracy was largely affected by haptic guidance. The finding that haptic guidance benefits timing accuracy irrespective of whether visual information is present, agrees with previous research on observational learning [6, 5].

Recently, Liu et al. re-examined some aspects of the Feygin study, but altering the protocol to make it more similar to a rehabilitation context [30]. One of the more significant changes that they made was to the trial structure. Instead of each trial consisting of two practice runs through the task motion followed by a test run as in the Feygin et al. experiment, each trial in the Liu et al. experiment consisted of seven practice runs followed by seven test runs. This allowed for an examination of learning during repeated unguided practice. Other differences from the Feygin et al. study is that Liu et al. only considered recall with vision and they only looked at position error. Although they found that all training conditions produced a significant improvement between the first and last trials, they did not find a significant difference between training with and without haptic guidance (in fact vision alone was marginally better). Additionally, they found that subject performance degraded over the course of the test runs in each trial with movements gravitating towards an "attractor path". Despite the fact that they did not measure timing error, making a comparison of the positive haptic guidance results found by Feygin et al. impossible, these results largely confirm those of Feygin et al.

Conclusions About Guidance

Given its somewhat mixed experimental history, what can we conclude about guidance? Considering its information-bearing properties and the importance of task complexity, there are a couple of generalizations that might be made about guidance and motor learning.

First, guidance may assist in learning by reducing the "cognitive load" imposed by non-trivial tasks, bringing task difficulty closer to an optimal or "germane load" for learning [42]. This idea is closely related to adaptive-training, where task difficulty is systematically varied according to the learner's abilities [29]. Second, guidance techniques may be particularly helpful during the early stages of learning non-trivial tasks with several degrees of freedom. In this situation, guidance may help the learner to both understand the nature of the task as well as get him/her "into the ballpark" of correct movement.

2.1.3 Haptic and Vibrotactile Feedback

In contrast to haptic guidance where learners are explicitly moved through the task motion, haptic feedback is more general and may refer to any type of information communicated through touch.

Huang et al. used a spring excitation task to investigate the differences between haptic feedback, visual feedback, and a combination of the two on learning to control a dynamical system [22]. They found that haptic feedback produced better performance than visual feedback and that the combination of both feedback forms was more effective than either one alone.

Morris et al. used haptic feedback in a force-learning context [31]. In their experiment, subjects were guided along a trajectory and had to produce a series of one-dimensional forces along the way. Subjects were trained using haptic feedback, visual feedback, and a combination of both forms of feedback. As with Huang et al., they found the combined visual-haptic feedback to be most effective at teaching the force learning task.

Vibrotactile feedback is another means of communicating haptic information. Lieberman used vibrotactile signals to indicate joint angle error to subjects trying to learn arm poses [28]. In an experiment designed to test whether the addition of vibrotactile feedback was beneficial for learning motor skills, subjects wore an arm suit with eight actuators place around the wrist and elbow which provided error feedback about the angle of these joints. Subjects were asked to assume various arm poses as well as perform motions as quickly as possible after they were displayed on a screen. Results of this experiment showed a reduction in steady-state error of approximately 15%.

2.2 Robotic Rehabilitation

Rehabilitative robotics is a relatively recent area of research that deals with how mechanical systems can be used to facilitate motor learning in a rehabilitation context.

Krebs and his collaborators have used robot-assisted therapy techniques as part of a treatment program for stroke patients [26, 25]. They have developed several systems, including the *MIT-Manus* device, which target arm and upper-body movement. Results of several pilot studies indicate that these devices are safe, well-tolerated by patients, and are effective at reducing physical impairments in stroke patients.

Reinkensmeyer et al. also use robotic devices for physical rehabilitation following stroke [36]. They use a device, called the *ARM Guide*, to diagnose and treat arm movement impairment following stroke. Initial results of a small pilot study suggest that the *ARM Guide* can produce quantifiable benefits in the chronic hemiparetic arm.

2.3 Haptics & Musical Controllers

Although haptic feedback may seem to be of relatively little concern in musical instrument design, there is evidence that the sense of touch may play an important role in how musicians understand and interact with their instruments [10]. Therefore haptics is an important area of research not only so that we may better understand how musicians use haptic cues, but also so that new instruments can be designed that make efficient use of this type of information.

Several researchers have explored how the addition of haptic cues can be used to augment existing types of instruments. Perhaps the most definitive body of work in this area is that of O'Modhrain [33]. She examined the role and importance of haptics in a number of musical performance contexts. Using a haptic display device called the *Moose*, she looked at the effect of haptic feedback on theremin performance accuracy. The results indicated that the addition of haptic feedback produced a small improvement in performance of this relatively simple instrument. A second set of experiments was designed to look at a more complex instrument that naturally has haptic cues. These experiments used a virtual bowed string model to test the hypothesis that the presence of friction would affect a performers ability to maintain good Helmholtz motion and tone. Although the results of an experiment with novice players did not find an effect of the presence of friction on performance, results of a second experiment with experienced players actually found a negative transfer effect. The author suggests that this indicates that the experiments had tapped into a relevant component of bowed performance, but that the haptic simulation had not been sufficient to allow performers to take advantage of it. She also notes that, while the presence of friction did not help experienced players in terms of performance, most indicated a preference for it.

Several other researchers have also built virtual instrument models to explore haptics. Gillespie built a one-octave force-feedback keyboard called the *Virtual Piano* [17, 16]. This system is capable of rendering approximations to the forces present in a grand piano (or other type of keyboard) using a rigid-body modeling scheme. Nichols developed the *vBow*, a virtual violin system that simulates the haptic feedback generated by a bow on strings [32].

2.4 Musical Performance Pedagogy

Despite its long history in human culture, there is relatively little formal theory of musical performance pedagogy (at least in Western music). This is not to say that there are no texts or teaching methodologies regarding proper performance techniques, only that there is little theory on how to best communicate those techniques. In addition, many of the documented theories of performance focus on higher level issues such as phrasing, while ignoring the development of basic motor skills. This pedagogical gap seems likely to be the result of inherent difficulties in communicating precise gesture, one of the central issues addressed by this thesis work. There are, however, several modern theories of piano pedagogy that are particularly well-developed and deal with low-level playing technique.

In 1929, Otto Ortmann began conducting experiments on the muscular actions involved in piano performance. He ran numerous studies of professional performers using high-speed photography and other measurement techniques available at the time. Based on the findings, he developed a theory of performance which he published along with his scientific results [34]. The approach outlined in his book is firmly rooted in physiology with an emphasis on rules of performance empirically derived from his studies. Most of these rules are concerned with specific ways that the hands and fingers can most efficiently be used to produce different types of notes (e.g. staccato, legato, etc.).

Not long after Ortmann conducted his research, a music pedagogue named Dorothy Taubman began developing another substantial and influential theory of piano performance. Although she did not arrive at her theory through scientific investigation as Ortmann did, her methods are based on a long career of piano instruction and critical observation of performance. It is difficult to discuss her theory in depth as there is no written documentation of it, however based on descriptions from the Golandsky Institute (the primary center for the Taubman approach) as well as videos released by Taubman, we might summarize the technique as one which emphasizes efficient and ergonomic playing [23]. Also characteristic of the Taubman approach is an emphasis on individual performer differences and the need to tailor lessons and technical corrections to each student.

Somewhat in between the approaches of Ortmann and Taubman, is that of Seymour Fink. Fink's theory is grounded in the idea that players should have an acute understanding of the physiology involved in piano performance [13]. As with Taubman, Fink's methods are not rooted in any scientific examination of performance, but rather are based on his observations as a teacher and a performer. The result is a bottom-up approach to piano technique which begins with what he calls "primary" movements, extends to more complex "integrative" and "joggle" movements, and finally applies these ideas to higher-level musical concepts.

Although these three theories of performance do, for the most part, attempt to break the complex movements required for piano performance down into manageable pieces, they all still suffer from the difficulties inherent in trying to verbally communicate physical motion. It should be noted that there are videos available illustrating both the Taubman and Fink techniques (and Fink does a good job of providing reasonable illustrations in his book). However, despite the arguable superiority of video as a medium for teaching movements, there is still a large amount of information that is difficult, if not impossible, to communicate. At a bare minimum we can say that both verbal and visual descriptions of body movements require translation to the proprioceptive coordinate system and that this translation is likely to be non-trivial. However, if haptic guidance is used for training, the information is communicated in the native (proprioceptive) coordinate system, avoiding the need for translation. This ability to communicate proprioceptive information directly is one of the primary motivations for the work described in this thesis.

2.5 Summary

This chapter presented a review of relevant literature and research. Clearly, much work has been done to further our understanding of how the human motor system learns, how natural feedback impacts that learning, and what forms of additional information might be presented to enhance it. Haptic guidance, along with other forms of augmented feedback, is a promising avenue of research both for what it might tell us about the workings of the motor system as well as what practical applications it may yield. As suggested by the brevity of the musical controllers section, relatively little research has been done to understand the relationship of haptic feedback and musical learning. However, no existing research (that I am aware of) has looked at haptic guidance in a musical training context. In Chapter 3 we shall turn our attention to two devices designed and built for this purpose.

Chapter 3

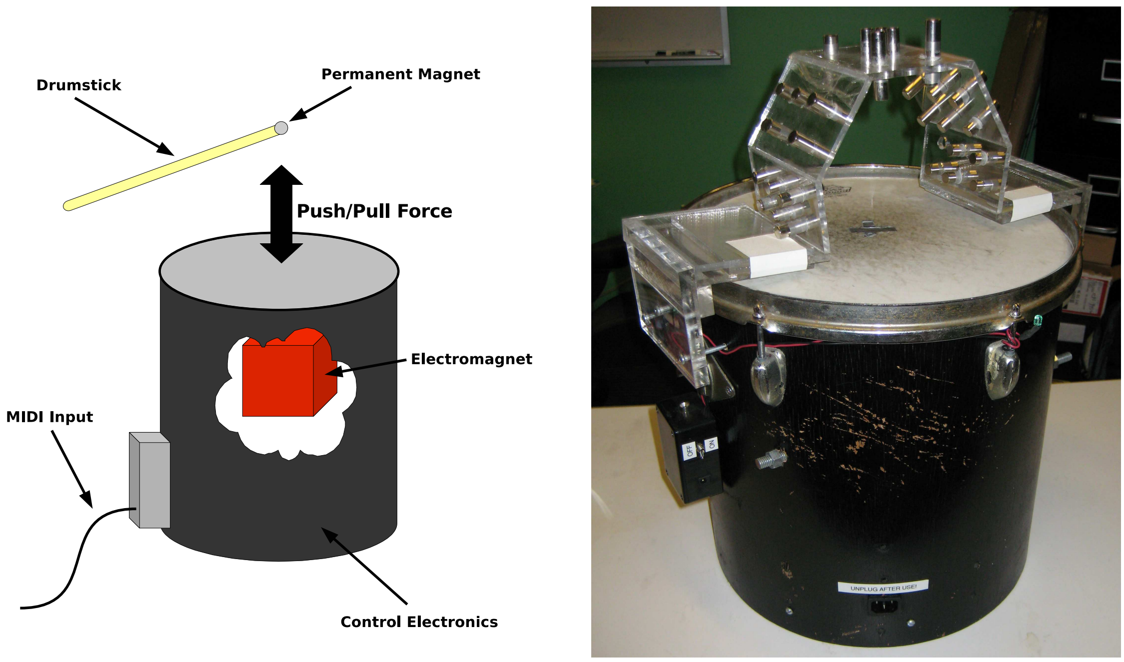
Devices

This chapter describes the two haptic guidance systems that were designed and built to study the effects of haptic guidance in the context of percussion performance. The first system, called the *FielDrum*, was designed as a general tool for providing percussion guidance in a realistic setting. It has several advantageous qualities, perhaps the most important of which is its untethered nature. Because the method of actuation is based on magnetics, its drumstick need not be physically connected to anything, allowing for a natural playing experience. The design does, however, have several drawbacks as well which make it difficult to use in a controlled experimental setting. This lead to the design of the second system, called the *Haptic Guidance System* (HAGUS). This system was primarily intended for use in an experimental setting and as such, its design differs quite substantially from that of the *FielDrum*. The design and construction of both systems is described below.

3.1 The FielDrum

3.1.1 System Design

The *FielDrum* consists of an off-the-shelf 13" tom-tom drum, large electromagnet (181mH), powersupply (24V at 8A), and switching circuitry (see Figure 3-1). The electromagnet sits



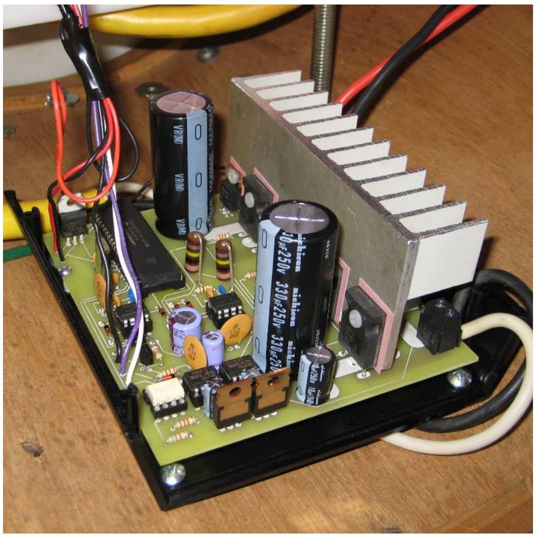
(a) Diagram of the *FielDrum* (b) Photo of the *FielDrum*

Figure 3-1: Components of the *FielDrum*

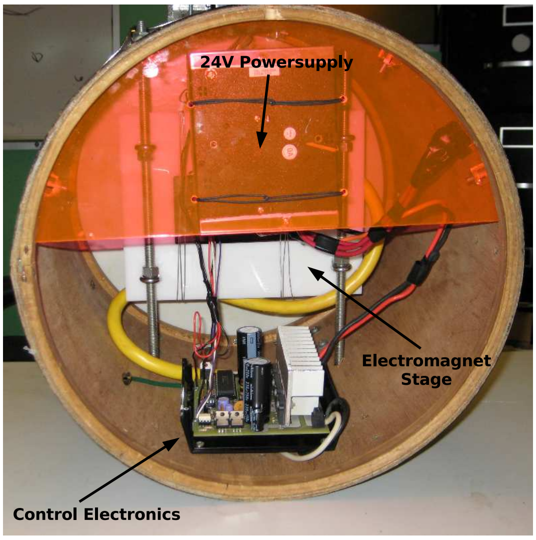
approximately 1" below the drum head on an acrylic stage. The powersupply, switching circuitry, and other support electronics are mounted internally as well. A PIC microcontroller is used to interpret MIDI note-on and note-off messages (pitch, velocity, and control messages are ignored) that arrive on an input line. These messages are then mapped to voltage polarities on the electromagnet using a modified H-bridge switching circuit (see Figure 3-3) such that either a north (note-on) or south (note-off) magnetic pole emanates from the portion of the electromagnet closest to the drumhead.[1]

The drum sticks have 0.5" spherical neodymium permanent magnets attached to their ends. These are arranged such that the south poles are emanating from the tips of the sticks. At rest the sticks are repelled from the electromagnet and therefore the drum head. When a note-on message arrives, the sticks become attracted to the electromagnet and strike the drum head if they are sufficiently close. Then when a note-off message arrives, the sticks once again become repelled from the drumhead. By carefully timing the spacing between

[1]I am indebted to Winfield Hill, Tony Williams, and others for help with the H-bridge design.



(a) The *FielDrum* control electronics



(b) The underside of the *FielDrum*

Figure 3-2: *FielDrum* control electronics

note-on and note-off messages, a "self-playing" effect can be achieved assuming that the drumsticks are in the vicinity of the drumhead.

3.1.2 Limitations

One major limitation of the *FielDrum* is its lack of sensors and therefore ability to measure the position of the drumstick above the drumhead. This has consequences both for the suitability of the device as an experimental tool (we have no record of a player's performance and therefore no means of evaluating that performance) and for the accuracy of its control system. Without a feedback loop the control system cannot guarantee that the drumstick is following the path that the electromagnet is trying to move it along. A capacitive sensing system was developed to solve these problems, but early tests showed that it was extremely difficult to achieve the necessary level of accuracy.

A more fundamental and perhaps serious limitation to the *FielDrum* design, is the use of magnets for actuation. The force exerted between two magnetic dipoles falls off proportionate to the inverse fourth power of the distance between them. This creates a

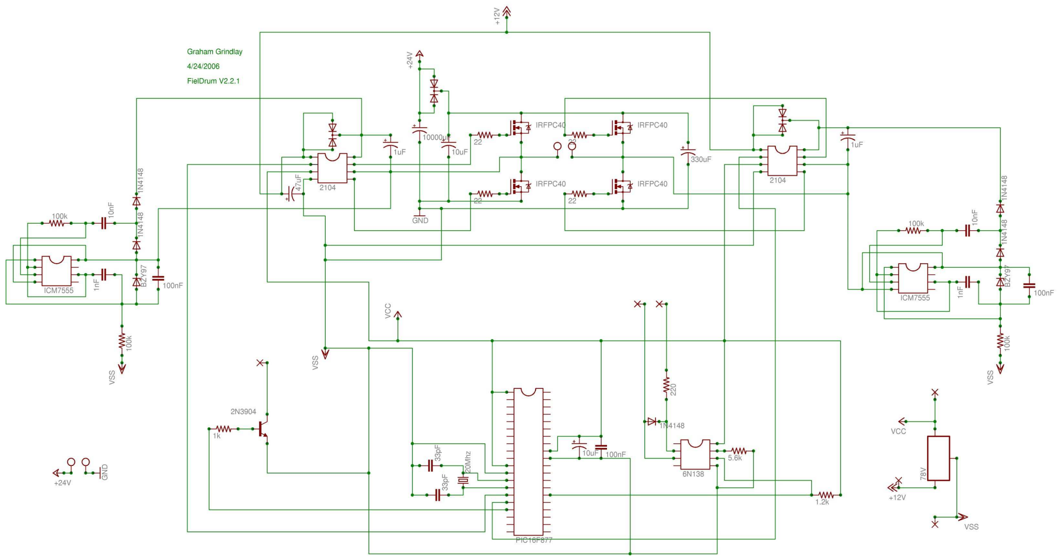


Figure 3-3: Switching circuitry used in the *FielDrum*

number of challenges, both because the force-distance relationship is non-linear and because the power requirements become substantial (the *FielDrum* consumes 250W of energy and really only has a range of influence of a few inches). Furthermore it is extremely difficult, if not impossible, to produce force completely perpendicular to the drumhead. This is because the magnetic field lines of a magnet arc around to connect its poles together. A "magnetic buttress" (a grid of permanent magnets all "pushing" inwards) was used (see Figure 3-1(b)) in the *FielDrum* to counteract this horizontal force. While the buttress did help to reduce the horizontal component of the force vector, it is a somewhat invasive addition as it seriously limits the distance that the drumstick can be raised above the drumhead.

3.1.3 Anecdotal Findings

The *FielDrum* was setup in the *Hyperinstruments* common area at the MIT Media Lab for roughly 8 months. During this time numerous demonstrations of the system were given to people of all ages and levels of musical experience. Although no formal conclusions can be drawn, there are a couple of interesting anecdotal observations that are worth reporting.

A particularly salient observation that seemed to hold true for almost all users was the instability and difficulty experienced when they tried to fight the influence of the magnets. When users tried to play along with the rhythm or otherwise take control, they invariably got out of phase with the push-pull force cycle. However, if users were too passive and did not firmly keep the drumstick in place over the drumhead, then other problems arose (such as the drumstick getting stuck to the buttress magnets). The people who had the most success with the system were those who held the drumstick firmly while still yielding to the influence of the magnets. Although this finding is not surprising given the lack of a feedback loop in the *FielDrum* control system, it provides valuable usability data that would be helpful in designing future revisions of the device.

Another interesting trend in user interactions with the *FielDrum* had to do with the age of the player. Although probably 90% of the demonstrations of the *FielDrum* were given to adults, at point a group of roughly 30 fourth-graders toured the Media Lab and tried out the *FielDrum*. The difference in the difficulty experienced by these children (roughly ten year-olds) as compared to adults was striking. The children appeared to have a significantly more difficult time wielding the drumstick, which often got stuck to the buttress magnets. While some of this difficulty was clearly due to differences in physical strength, it also seemed as though the children were more willing to explore the interface. These observations would be worth exploring in more depth as it is certainly possible that the design of a physical guidance based teaching system may need to be tailored to user age.

3.2 The Haptic Guidance System

The *Haptic Guidance System* (HAGUS) was designed to be used as an experimental tool for exploring the effects of haptic guidance on motor learning. Much of the impetus for its construction came from lessons learned from the *FielDrum*. Of primary concern was the need for precision both in measurement of position and in control of actuation forces. While these issues might have been at least partially addressed through a substantial redesign effort, moving to a servo motor-based design seemed a far easier option. Initially, it was thought that a *PHANToM*[2] would be an ideal experimental tool for testing the effects of haptic guidance. However early experiments proved the *PHANToM* to be underpowered for percussion and so a custom solution was designed and built. In the interest of tractability and to keep the experimental question as simple as possible, HAGUS was designed to target only wrist movement and therefore can only record and playback rotational motions about a single axis. Although percussion performance in general certainly isn't restricted entirely to the wrist, I believe that this simplification provides a reasonable first-order approximation. Additionally, it should be noted that particularly for persons with no prior percussion experience, there are many non-trivial rhythmic tasks possible with a single degree of freedom.

3.2.1 Goals

The goals for HAGUS were the following:

1. Highly accurate positional measurement
2. Actuation sufficiently powerful to ensure accurate guidance
3. The ability to both "playback" and "record" drumstick motions
4. Minimal physical impedance during recording
5. Safe operation!

3.2.2 System Design

Figure 3-4 shows a functional diagram of HAGUS. A PC running software described below is responsible for high-level control. It stores data files containing the raw position information that HAGUS generates when in recording mode and uses during playback mode.

[2] SensAble Technologies, `http://www.sensable.com`

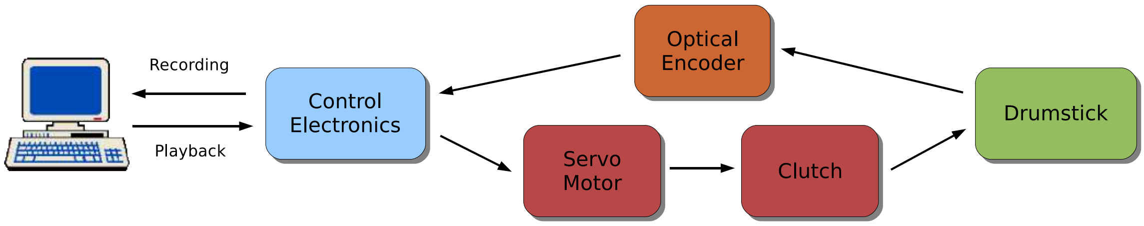


Figure 3-4: Functional diagram of the Haptic Guidance System (HAGUS)

Hardware

HAGUS uses a combination of onboard electronics and PC-based computing power. The actuator (see Figure 3-5 and Figure 3-6(b)) consists of a 40 Watt servo motor[3] connected via an electromagnetic particle clutch[4] to the primary drive shaft. This drive shaft connects to a high resolution optical encoder[5] and is then geared-up (1:3 ratio) before connecting to the drumstick. This gearing brings a number of advantages, such as three-fold increases in effective motor torque and clutch holding torque. However, as with virtually all mechanical couplings, these gears suffer from a small amount of backlash.[6] Empirically, the system backlash in the system was minimal and corresponded to about 0.16 degrees of play in the drumstick.

The servo motor is run by motion control hardware[7] running a proportion-plus-derivative (PD) control filter. This filter is updated at a frequency of 1.953kHz. Representing the proportion gain as K_p, the derivative gain as K_d, and the error at time t as $e(t)$, the standard PD control filter formulation is as follows:

$$Output(t) = K_p e(t) + K_d \frac{de}{dt}$$

[3] Sanyo Denki Super V, model V404-011, 24V, 2.9Amp DC servo motor

[4] Placid Industries model C5-24 electromagnetic particle clutch. This clutch has 80oz-in of holding torque and 1oz-in of drag torque.

[5] RENCO model RCML15 2000 line quadrature encoder. This effectively gives 8000 counts/revolution or 0.045 degree accuracy.

[6] Backlash is the excess of space between the teeth of one gear over the thickness of the teeth of the other gear [46]. Although, one generally strives to minimize backlash, it is generally unavoidable due to manufacturing imperfections, alignment inaccuracies, and thermal expansion.

[7] PIC-SERVO SC, Jeffrey Kerr, LLC., `http://www.jrkerr.com`

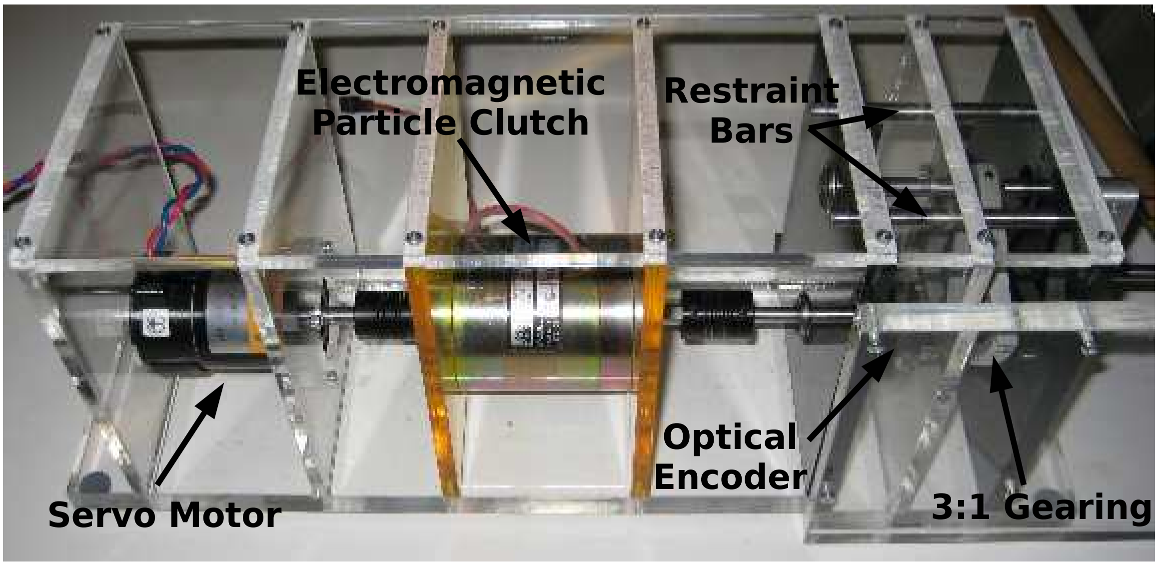


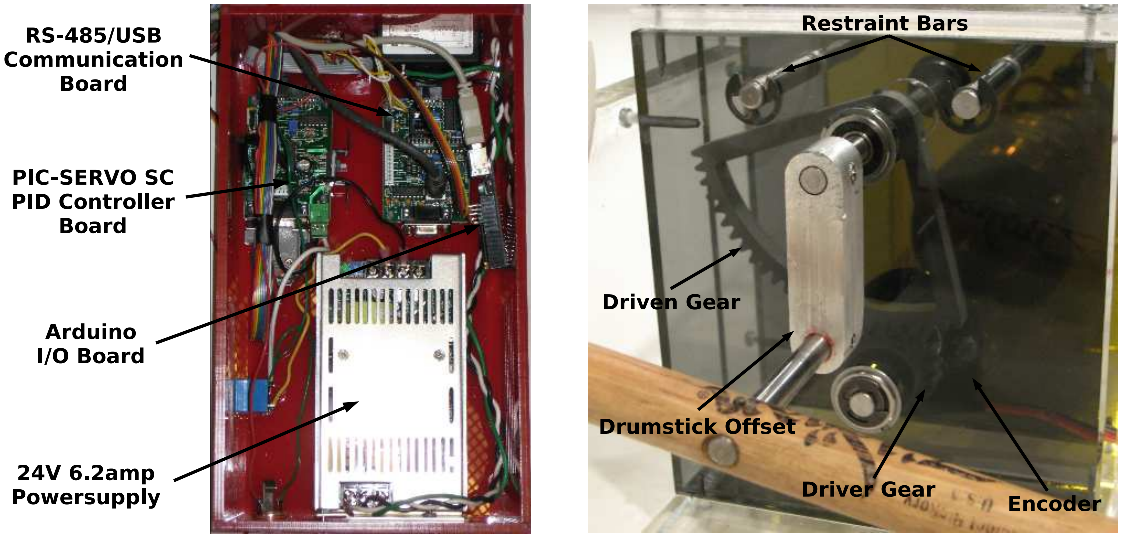
Figure 3-5: The HAGUS actuation hardware

In the experiment described in Chapter 4, gains were typically set such that $K_d = 10K_p$. Additionally, a small amount of deadband compensation was used to reduce jitter.

Position measurements are recorded and played back at a 60Hz sampling rate. To record a motion, the clutch is first disengaged to disconnect the servo motor from the rest of the drive train. This minimizes the amount of physical impedance presented to the user when he/she is freely playing. During recording, HAGUS' motion control electronics stream encoder readings to the host PC over a USB connection which then logs these data to disk. To playback a previously recorded motion, the direction of information is simply reversed; the host PC streams the position data from a file to the control electronics where they are used with the PD filter loop described above to reproduce the motion.

Along with a powersupply and servo control boards, the HAGUS system electronics (see Figure 3-6(a)) include an *Arduino*[8] board for general I/O purposes. The *Arduino* was primarily used to control a set of LED lights which in turn were used during the experiments described in Chapter 4 to cue subjects as to when to expect motion playback to begin, when they should begin playing back a rhythm, etc.

[8] `http://www.arduino.cc`



(a) HAGUS control electronics (b) Parts of the HAGUS actuator

Figure 3-6: Electronic and hardware components of HAGUS

The complete HAGUS setup is shown in Figure 3-7. This figure shows the wrist cradle and strap which are mounted to the end of the actuator and ensure that all subjects are positioned ergonomically and consistently. Several safety features were also included in the HAGUS design, such as an emergency shutoff switch (see Figure 3-7) and over-rotation protection. Restraint bars (see Figure 3-5) were used in the actuator to physically prevent over-rotation. These bars were placed such that if the user (or motor) tried to rotate the drumstick outside of a 40 degree fixed range, the driven gear (the partial gear that is connected to the drumstick) pushes up against them, preventing it from rotating further.

Software

As mentioned previously, high-level control of HAGUS is the responsibility of a host PC. Software was developed in C++ to handle all aspects of streaming position data to and from HAGUS as well as organizing and running experimental sessions. The software's GUI, which is shown in Figure 3-8, uses the *Qt* development framework[9] and runs under the Linux operating system. It provides control over a number of PD filter loop settings as well

[9] `http://www.trolltech.com`

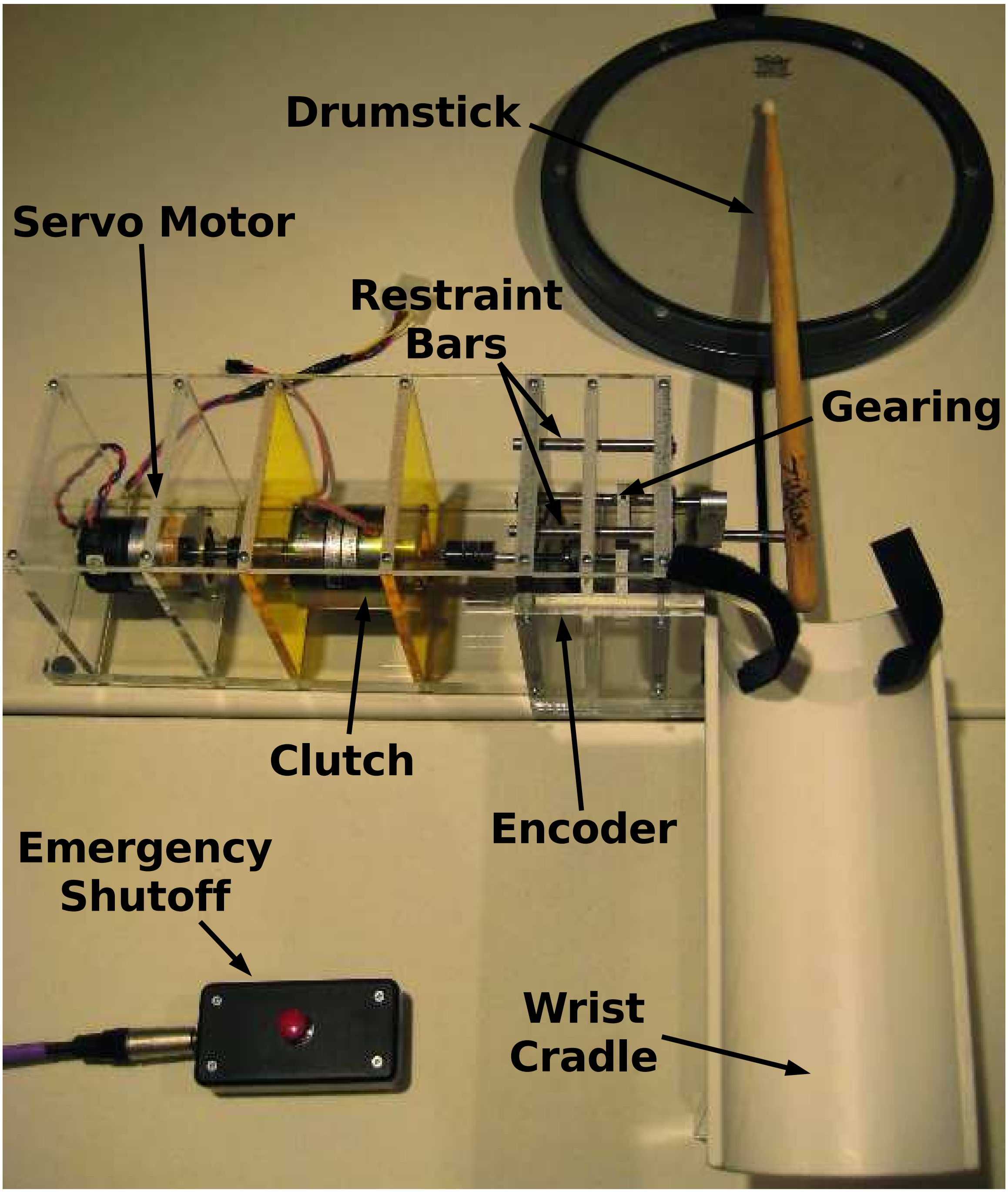


Figure 3-7: The Haptic Guidance System hardware

as low-latency scheduling of position data. The experimental design features were intended to reduce the amount of experimenter intervention required and minimize the chance of human error. Automating control of the presentation and recall runs has the added benefit of making the experimental flow more consistent across subjects. Other features include error logging and the ability to schedule practice runs.

To record a motion sequence, the HAGUS software launches a “recorder” thread which samples the USB line (which gets encoder readings from the control hardware) at 60Hz. These frames are stored in memory until the recording period is over at which point they are logged to disk. The software can either record for a predetermined amount of time or until the user hits a “Stop” button. In the experiment described in Chapter 4, the former was used as the experimental sequence was largely automated.

Motion playback works similarly to recording. Position data are read from disk to memory and then streamed to the control hardware. A major difference from the recording setup is that the control hardware is responsible for timing. It has a small local buffer from which it reads data before using them in the PD control filter. The HAGUS software therefore only needs to worry about keeping the buffer full while not overflowing it.

3.2.3 Limitations

One obvious limitation of HAGUS, which is inherent in its intended design, is the inability to target more than a single degree of freedome. However, depending on the task to be studied, expansion to several axes could be relatively straightforward. The software and control electronics would be particularly easy to extend (the PIC-SERVO platform supports up to 16 axes). Another area in which HAGUS could be extended/improved is the sampling rate at which recording and playback operates. Although the control hardware is capable of handling data points at 120Hz as well as 60Hz, the software cannot reliably operate at this rate. The problem is rooted in the standard Linux kernel which like all non-realtime operating systems, cannot guarantee function calls at frequencies above about 100Hz. To do better, a realtime kernel or hardware-based buffering scheme would be needed.

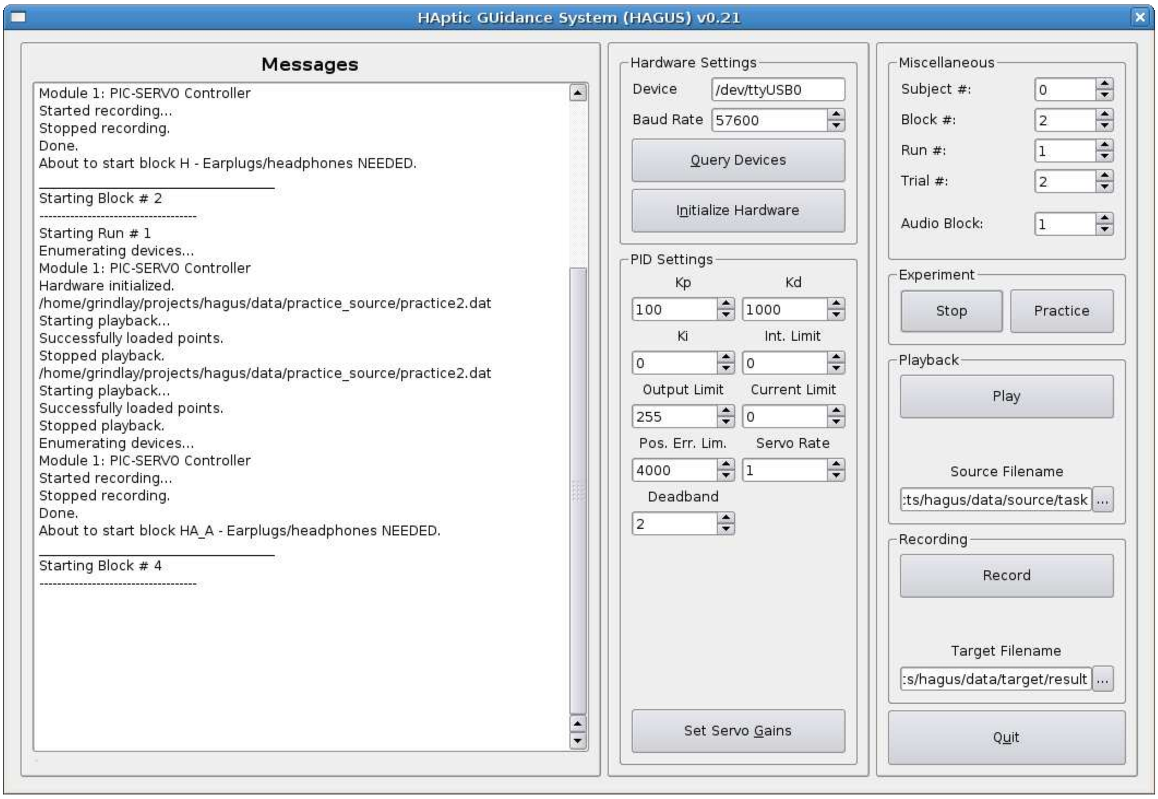


Figure 3-8: The Haptic Guidance System software

Chapter 4

Experimental Evaluation

4.1 Experimental Design

The primary purpose of the experiment was to investigate the differences between haptic guidance and audio-based training on percussion learning. The hypothesis being tested was that haptic guidance combined with auditory feedback would result in participants being able to reproduce rhythms more accurately than either auditory feedback alone or haptic guidance alone. Differences in accuracy were measured in terms of note timing as well as note velocity.

Subjects were trained to perform four different rhythms under four different training conditions. During each trial, the subject was run through two training presentations of the task rhythm which was then followed by a recall run where the subject tried to reproduce the task rhythm with no assistance. Each training condition consisted of 15 consecutive trials followed by a five minute break.

The experiment compared four different training techniques, three primary and one ancillary. The first primary training technique was an audio-only (A) condition. During this training condition, subjects did not move their hands, but only listened to a recording of the task rhythm being played by the HAGUS device. This condition was designed to mimic a typical at-home self-instruction situation where a student may have an instructional book and CD with audio examples. The second primary training technique was a haptic guidance only (H) condition. During this condition, subjects were physically moved through the motions required to perform the task rhythm, but were unable to hear. Subjects wore -32dB earplugs[1] as well as headphones which played white noise masking sound. A pilot experiment testing this setup confirmed its efficacy at preventing subjects from hearing drumpad sounds. The third primary training condition (A+H) was a combination of the first two where subjects were physically guided through the ideal task motion and were also able to hear its results.

Ideally, all conditions would have identical recall run setups. However, this was not possible as it would have meant that the subjects would need to remove the earplugs and headphones between each of the H condition trials. Because the masking noise was necessary to effectively prevent subjects from hearing, I opted instead to allow subjects to leave the earplugs in and headphones on during the haptic guidance only (H) recall runs. Even though subjects reported being able to hear the drumpad fairly well with this setup, a fourth, ancillary training condition was included to test for the effects of attenuated hearing with the presence of the earplugs and headphones (when not playing the masking noise). This condition (A+H(atten)) was similar in all ways to the A+H condition with the exception that subjects wore earplugs and headphones (without masking noise) throughout the condition.

	Condition Order			
Group 1	H	A+H	A	A+H(atten)
Group 2	A+H	A+H(atten)	A	H
Group 3	A+H(atten)	A	A+H	H
Group 4	A	H	A+H(atten)	A+H

Table 4.1: Balanced Latin Square design used to order training conditions.

Each of these four training conditions was considered a within-subjects factor in a repeated measures experiment. Subjects were randomly assigned to one of four groups and

[1] Aearo Co., `http://www.aosafety.com`

a balanced Latin square design (see Table 4.1) was used to order the training conditions differently for each group. Four different rhythmic tasks (see Section 4.1.2) were used to prevent learning transfer between training conditions. The assignment of rhythmic task to training condition was varied across groups using a balanced Latin square design (see Table 4.2).

	Condition			
Group 1	Task 4	Task 3	Task 2	Task 1
Group 2	Task 3	Task 1	Task 4	Task 2
Group 3	Task 1	Task 2	Task 3	Task 4
Group 4	Task 2	Task 4	Task 1	Task 3

Table 4.2: Balanced Latin Square design used to assign rhythmic tasks to training conditions.

4.1.1 Subjects

Thirty-two right-handed subjects (20 females and 12 males), all of whom were between the ages of 18 and 50 (the median age was 27), were recruited for the study. None of the subjects had any percussion training or significant playing experience, although some subjects did have training and /or experience with other instruments (see Table 4.8). Each subject was paid $10 for their participation.

4.1.2 Tasks

Four different rhythmic tasks were used in this experiment. These were designed to be non-trivial while still being learnable within the 15 trial period of each condition in the experiment. A set of four rhythms containing eight notes each (one quarter note, three eighth notes, three sixteenth notes, and one dotted-eighth note) was devised (see Figure 4-1). A small pilot study with three subjects suggested that these rhythms were of an appropriate level of difficulty (each subject's data showed reasonable learning curves).

(a) Task rhythm 1

(b) Task rhythm 2

(c) Task rhythm 3

(d) Task rhythm 4

Figure 4-1: The four task rhythms used in the experiment

Template audio and haptic performances of each of these rhythms (the target standard that the subjects were trained on and therefore compared against) were generated by the author. A tempo of 80 beats-per-minute was used, which meant that each rhythm was exactly three seconds in duration. The haptic guidance templates were produced by playing each rhythm on the HAGUS device while listening to an audio rendition of that rhythm (to ensure accurate timing).[2] Several takes of each rhythm were recorded and the best one (to my ears) was retained as the template for that rhythm. An audio template for each rhythm was also produced by playing the haptic guidance template back on HAGUS and recording the sound that it produced. This ensured that the audio used in the A training condition closely matched the audio that was produced during the A+H and A+H(atten) conditions.

4.2 Procedure

Subjects were first familiarized with the purpose of the study and equipment. Verbal instructions were given and informed consent was obtained. Each subject then practiced one trial of the A training condition and one trial of the H training condition (the A+H and A+H(atten) were judged to be similar enough that practice was unnecessary). Each trial consisted of two presentation (training) runs immediately followed by a recall (test) run.

[2] Although the use of artificially constructed motion sequences (i.e. sequences of single stroke motions stitched together) would be advantageous in terms of rhythmic precision, it is not clear *prima facie* how this could be done while ensuring that the sequences are ergonomically sensible. Therefore, it was decided that it was preferable to use non-quantized human performances rather than risk potentially awkward and unnatural task motions.

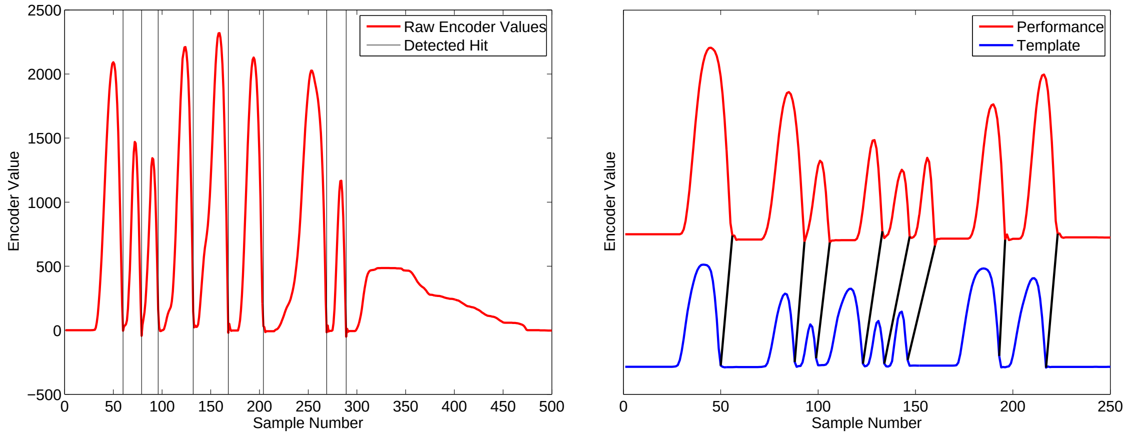
Subjects were instructed to "play along" with the HAGUS device during the training runs that included haptic guidance (H, A+H, and A+H(atten)) and to just listen during the A training condition. Each subject was also instructed to try to reproduce the task rhythm as accurately as possible in all respects during the recall runs. There were 4 training conditions, each of which consisted of 15 trials. A set of LED lights was used to cue subjects as to when each training and testing run was about to begin. This light sequence took 3 seconds to complete which, when added to the 3 second task duration and a 1 second delay after each task, produced a total run length of 7 seconds. Given 3 runs (2 training and 1 test) per trial, each trial lasted 21 seconds and each training condition lasted 5 minutes and 15 seconds (15 trials per condition). Subjects completed each training condition (all 15 trials) and then were given a short (5 minute) break. After all 4 training conditions had been completed, subjects filled out a brief questionnaire.

4.3 Measurement & Preprocessing

4.3.1 Onset Detection

The data from each recall run during the experiment was recorded by HAGUS and logged by a host PC. However this data consists of raw encoder readings which provide a very accurate history of the drumstick movements but are too low-level to be useful in assessing percussion performance. Therefore, the first step in analyzing a performance is to translate it into symbolic form by finding the times at which each drum stroke occurred.

For the most part the method used is a fairly straightforward trough-picking algorithm, although in practice data pathologies necessitated several ad hoc modifications. Perhaps the most difficult has to do with deciding when a trough in encoder readings actually corresponds to a percussive hit. Visual inspection of the raw data indicated that the last note of a rhythm was often played with stick rebound such that the subject then had to slowly lower the stick to return it to its starting position for the next trial (see Figure 4-2(a)). This "lowering" could potentially confuse the onset detector and so thresholding on



(a) Example of onsets detected in raw encoder data.

(b) Dynamic time warping alignment of performance and template data.

Figure 4-2: Onset detection and alignment of subject data.

the first-derivative of the encoder readings was used. Additionally, other thresholds were added to prevent false onset detection of encoder noise and drumstick movements small-enough that they were unlikely to be intended as hits. Finally, all sequences of onset times were normalized to begin at time 0 (i.e. each element of each sequence had the first element of that sequence subtracted from it).

4.3.2 Stroke Velocity

In addition to evaluating subject performance in terms of timing, I also assessed their accuracy in reproducing note velocity. This data was relatively easy to procure given the onset times extracted from the techniques described in the previous section. Pseudocode for this algorithm is given in Algorithm Listing 1.

4.3.3 Scoring

Once we have obtained a symbolic representation of a rhythmic performance we need to find a way to compare it to the template rhythm. However, assessing the total similarity or difference between two performances of a rhythm is a fairly difficult task. There are

Algorithm 1 Stroke velocity detection algorithm

Require: $onsetTimes[1..T], encoderData[1..N]$
  $velocities[1..T] \leftarrow []$
  for $t \leftarrow 1 \ldots T$ **do**
    if $t == 1$ **then**
      $lastIndex = 1$
    else
      $lastIndex = onsetTimes[t-1]$
    end if
    $maxIndex = indexOfMaxValue(encoderData[lastIndex..onsetTimes[t]])$
    $velocities[t] = \frac{encoderData[lastIndex+maxIndex-1]-encoderData[onsetTimes[t]]}{onsetTimes[t]-(lastIndex+maxIndex-1)}$
  end for
  return $velocities$

several dimensions in which the rhythms may differ and it is unclear how these differences should be weighted in a complete metric. For example, suppose one rhythm differs from another by only the timing of a single note while another differs in that an extra note has been added (or deleted). Or suppose that the global tempo of a rhythm has been sped up or slowed down during its performance. How should these differences be compared and combined in a global similarity metric? The approach taken here was to use five different evaluation metrics, one to assess velocity accuracy and four to examine different aspects of timing precision.

Timing Metrics

The first and most complete timing metric is referred to as the *unnormalized distance* (UD). Pseudocode for the algorithm is given in Algorithm Listing 2. The UD metric does not require that the two sequences to be compared be of the same total length or have the same number of elements. The metric uses a variant of the well-known *dynamic time warping* algorithm to align and compare two sequences of data and produce a single scalar number representing their distance/similarity [43]. Dynamic time warping is a algorithmic technique for comparing sequences that may vary non-linearly in time or speed. It is based on dynamic programming and is guaranteed to produce the optimal (in terms of the cost function used) alignment. Figure 4-2(b) shows an example alignment of onset data superimposed onto the corresponding raw encoder values.

The cost of matching any two elements (onset times) between sequences is taken to be the absolute value of their difference and an additional cost to insert or delete elements is also included to reflect the severity of this type of error. An extra cost of 50% of the total length (in seconds) of the longer of the two sequences was used for both insertions and deletions.

Algorithm 2 Unnormalized rhythmic distance algorithm

Require: $t[1..N], p[1..M]$
  $DTW[1 \dots N, 1 \dots M] \leftarrow 0$
  $DTW[1, 2 \dots M] \leftarrow \infty$
  $DTW[2 \dots N, 1] \leftarrow \infty$
  $cost_{ins} \leftarrow \frac{1}{2} max(t[N], p[M])$
  $cost_{del} \leftarrow \frac{1}{2} max(t[N], p[M])$
  for $i \leftarrow 2 \dots N$ **do**
    for $j \leftarrow 2 \dots M$ **do**
      $cost_{dist} = |t[i] - p[j]|$
      $DTW[i, j] \leftarrow min(DTW[i-1, j] + cost_{dist} + cost_{ins},$
        $DTW[i, j-1] + cost_{dist} + cost_{del},$
        $DTW[i-1, j-1] + cost_{dist})$
    end for
  end for
  return $DTW[N, M]$

The second distance metric is referred to as the *normalized distance* (ND). Pseudocode for it can be found in Algorithm Listing 3. The UD metric assigns cost based on absolute onset time and will therefore punish sequences that have different global timescales/tempos, but are otherwise identical. In contrast, the ND metric first normalizes each sequence (by dividing each element in each sequence by the last element in that sequence) such that differences due to global tempo are removed. It should be noted, however, that this normalization step potentially introduces an added penalty when the two sequences to be compared have different numbers of elements (i.e. insertions or deletions need to take place). Consider a performance sequence that matches the template perfectly except for an extra note at its end. After normalization, the notes in the performance sequence that had been correct, will have been shifted in time and this will incur error during the alignment process.

The third comparison metric, referred to as the *global tempo distance* (GT), is essentially the complement of the ND metric. It provides a measure of the global tempo similarity

Algorithm 3 Normalized rhythmic distance algorithm

Require: $t[1..N], p[1..M]$
  $DTW[1 \ldots N, 1 \ldots M] \leftarrow 0$
  $DTW[1, 2 \ldots M] \leftarrow \infty$
  $DTW[2 \ldots N, 1] \leftarrow \infty$
  $t[1..N] \leftarrow \frac{t[1..N]}{t[N]}$
  $p[1..M] \leftarrow \frac{p[1..M]}{p[M]}$
  $cost_{ins} \leftarrow \frac{1}{2}$
  $cost_{del} \leftarrow \frac{1}{2}$
  for $i \leftarrow 2 \ldots N$ **do**
    for $j \leftarrow 2 \ldots M$ **do**
      $cost_{dist} = |t[i] - p[j]|$
      $DTW[i, j] \leftarrow min(DTW[i-1, j] + cost_{dist} + cost_{ins},$
        $DTW[i, j-1] + cost_{dist} + cost_{del},$
        $DTW[i-1, j-1] + cost_{dist})$
    end for
  end for
  return $DTW[N, M]$

between two performances of a rhythmic sequence. This is done by comparing the final onset times in both sequences using the following symmetric function:

$$GT(t[1..N], p[1..M]) = |log_2 \frac{p[M]}{t[N]}| \tag{4.1}$$

The fourth metric compares the number of notes in one rhythmic sequence to the number in another. This provides a simple (and somewhat crude) approximation to the number of insertions and deletions required to match the two sequences. The metric is slightly different from the rest in that it can be positive (indicating insertions) or negative (indicating deletions). There is a potential problem with this formulation as insertions and deletions could potentially wash each other out across subjects, however in practice, the metric appeared robust to this issue. This metric is referred to as the *insertion/deletion distance* and is given by Equation 4.2.

$$ID(t[1..N], p[1..M]) = \frac{M - N}{N} \tag{4.2}$$

Velocity Metric

Because the velocity measurements do not represent temporal measurements like the onset data do, the distance metrics described in the previous section are not appropriate. I chose to use the simplest possible measure of error, the sum of the absolute values of the differences between performance velocities and template velocities. However, we still face the difficulty of sometimes having to deal with sequences of differing lengths. This suggests a dynamic programming approach, but the nature of the velocity data prevents us from using something like dynamic time warping directly on the data. The solution that was chosen was to save the performance-to-template mapping obtained during the alignment of timing data under the UD metric. This information can then be used to determine which elements of the performed timing data (and therefore the velocity data as well) were missing or are extra. Now we can perform the straightforward sum of absolute differences on the velocity data that has been matched using the timing data and we can also add in extra cost for insertions or deletions. In practice, additional insertion/deletion penalties were not included as they did not appear to significantly affect the results. The pseudocode for this distance metric, which is referred to as the *velocity distance* (VD), is given in Algorithm Listing 4.

Algorithm 4 Velocity distance function

Require: $t[1..N], p[1..M], tMatch[1..max(N,M)], pMatch[1..max(N,M)]$
  $distance = 0$
  for $i \leftarrow 1 \ldots max(N,M)$ **do**
    $distance = distance + |t[sMatch[i]] - s[pMatch[i]]|$
  end for
  return $distance$

Ensuring Normality

Normality of data is one of several underlying assumptions made by the ANOVA model which is used extensively in the analyses presented in Section 4.4. While still an active area of research, recent work seems to indicate that ANOVA is fairly robust to modest violations of this assumption [11]. Still, in the interest of using data that was as close to normal as possible, the data produced by the scoring algorithms were transformed before running the ANOVAs. Inspection revealed significant skew in the data generated with the UD, ND, GT, and VD metrics. Therefore, a fourth-root transform was used for these metrics, while a second (square) root transform was used for the GT metric. After transformation, data from all of these metrics appeared to approximate the normal distribution quite well.

4.3.4 Handling Anomalous Data

One issue that occurs in almost all experimental settings is missing data. In the current setup, a (very) intermittent bug in the servo control board firmware caused encoder readings to periodically get stuck on a single value. When this problem occurred, it only corrupted the current trial's data since the control hardware was reset after each recording run. In the end, the problem proved to be relatively minor as only 12 out of the 1920 recorded recall runs, or about 0.6% of the data, were corrupted. Those 12 runs were treated as missing data and simple linear interpolation/extrapolation was used to fill them in once the rest of the data had been preprocessed and converted to sequences of onset times. Eight of the missing runs occurred in separate subjects, while two subjects were each missing two runs. Neither of these sets of two occurred in succession. For each subject missing the first run of a condition (one subject), the value was imputed by linearly extrapolating it from runs two and three. For each subject missing the final run of a condition (two subjects), the value was imputed by linearly extrapolating it from runs 13 and 14. For each subject missing an interior run, the value was linearly interpolated using the two adjacent runs' values.

4.4 Results

This section presents the analysis results of the data gathered during the experiment described above. The results are broken up into two subsections, one for the analyses that pertained to the timing performance data and one for the analyses of velocity performance data.

Differences in training conditions were assessed using repeated-measures analysis of variance (ANOVA) tests with training condition (A, H, A+H, A+H(atten)) as the within-subjects factor. Separate ANOVAs were run for each type of distance metrics and then Bonferroni-corrected pair-wise t-tests were used to compare A+H and A+H(atten) as well as each of the possible pairings of primary training conditions ({A,H}, {A,A+H} and{H,A+H}) [1].

The Bonferroni method is a type of statistical adjustment for paired t-tests that reduces the potential for type I error when multiple comparisons are being made. The method simply divides the statistical significance threshold level by the number of tests being performed, in essence "raising the bar" of significance. It is worth noting that this technique has been criticized as being too conservative in some cases, decreasing type I error at the expense of higher type II error [35]. In the following sections whenever raw p-values are presented, we will indicate whether the value is significant or not when using the Bonferroni correction.

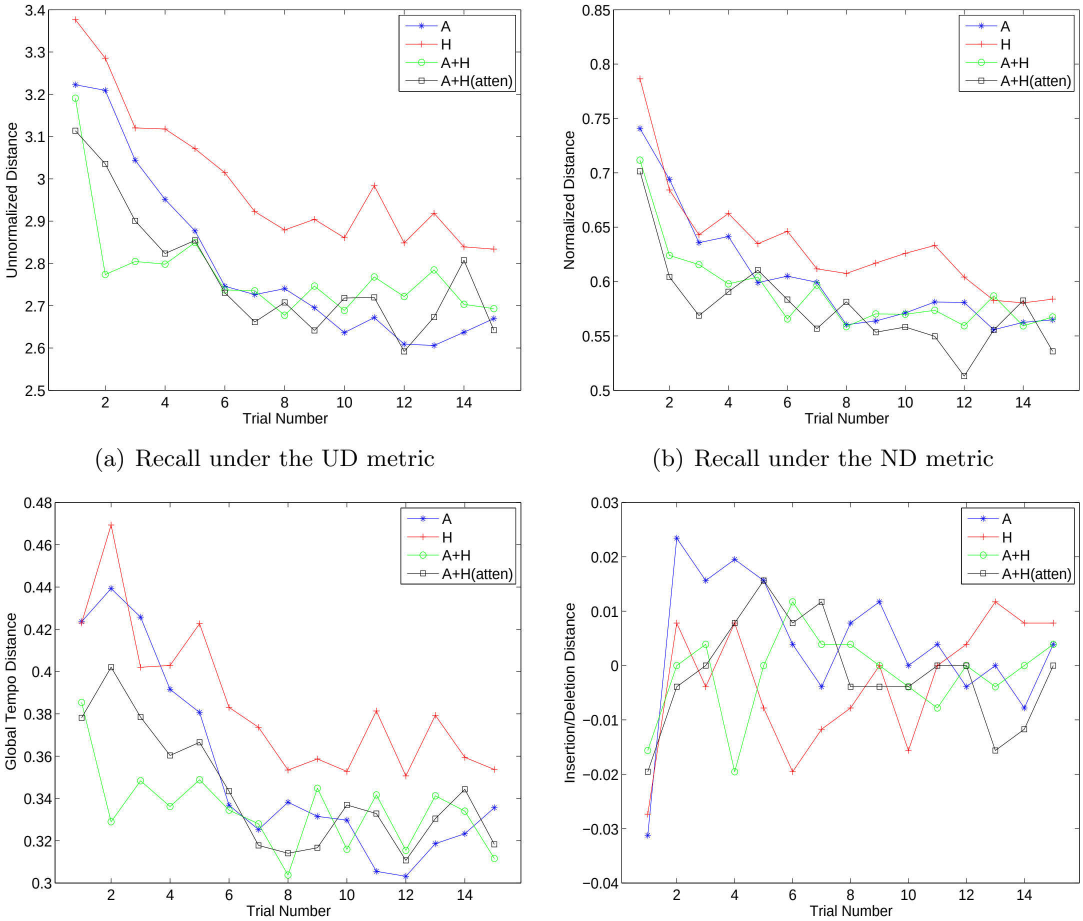
4.4.1 Timing

Learning Across Trials

	A	H	A+H	A+H(atten)
Unnormalized	$p < 0.0008$ (Y)	$p < 0.0028$ (Y)	$p < 0.0282$ (Y)	$p < 0.0064$ (Y)
Normalized	$p < 0.0001$ (Y)	$p < 0.0001$ (Y)	$p < 0.0033$ (Y)	$p < 0.0002$ (Y)
Global Tempo	$p < 0.0699$ (M)	$p < 0.1598$ (N)	$p < 0.1596$ (N)	$p < 0.1582$ (N)
Insertion/Deletion	$p < 0.0256$ (Y)	$p < 0.0835$ (M)	$p < 0.4052$ (N)	$p < 0.2264$ (N)

Table 4.3: Pair-wise t-test results for comparisons between the first and last recall runs. p-values for each combination of training condition and distance metric are given along with whether the value is significant (Y), not significant (N), or marginally significant (M).

Figure 4-3 shows the average (arithmetic mean across subjects) recall curves for the four distance metrics. These curves given an overall sense of differences between training conditions and distance metrics (although again, only UD and ND were statistically significant). We can also see the general trends in error reduction across trials, providing evidence that,



(a) Recall under the UD metric

(b) Recall under the ND metric

(c) Recall under the GT metric

(d) Recall under the ID metric

Figure 4-3: Comparison of recall performance under the four training conditions as measured under the four different distance metrics. Each plot shows the recall curve for its training condition averaged across subjects.

on average, subjects learned to improve their performance. Paired *t*-tests were used to examine performance improvement between the first and last trial of each training condition and under each distance metric. These tests, which are summarized in Table 4.3, show significant improvement for all training conditions under the UD and ND distance metrics. Significant and marginally significant improvements were found for the A and H conditions, respectively, under the ID metric, and only a marginally significant improvement was found for the A training condition under the GT distance metric.

Early Trials

Although one might expect that subject performance on the first trial would not vary significantly between training conditions, the performance curves in Figure 4-3 suggest a possible difference in early performance. They also show that there is some instability in the earlier sections of some of the recall curves. Therefore, to give a representation of the "early" section of training performance (as opposed to the potentially noisy and misleading first trial), the mean of the first three trials worth of data was used instead. Next, a repeated measures ANOVA (again with training condition as the within-subjects factor) was performed using this data and a separate ANOVA was run for each of the four distance metrics. The results are given in Table 4.4.

Distance Metric	F	p	Significant?
Unnormalized	3.462	< 0.0195	Y
Normalized	3.389	< 0.0213	Y
Global Tempo	2.739	< 0.0479	Y
Insertion/Deletion	1.105	< 0.3512	N

Table 4.4: Summary of results of a repeated-measures ANOVA with training condition as the independent variable and early trial recall performance as the dependent variable.

There was a significant main effect of training condition for the UD ($p < 0.019$), ND ($p < 0.021$), and GT ($p < 0.047$) distance metrics. Bonferroni-corrected t-tests were used to both confirm that the A+H and A+H(atten) were statistically similar and allow for further analysis of the ANOVA results.

Under the UD metric, there was a statistically significant ($p < 0.003$) difference between the H and A+H conditions which corresponded to a 10% error reduction for the A+H versus H training condition.

Marginally significant differences were found between the A and A+H conditions ($p < 0.0198$, 18% error reduction for the A+H condition) and between the H and A+H conditions ($p < 0.0170$, 18% error reduction for the A+H condition) under the GT metric.

No significant differences were found for either the ND nor the ID metrics.

Final Trials

The next logical question to ask is whether or not the different training conditions led to different levels of performance by the end of each training block. As with the first trials, the recall curves were somewhat noisy and so the mean of the final five trials worth of data was used as learning had largely stabilized by that point. Repeated-measures ANOVAs with training condition as the within-subjects factor were run (one per distance metric) to check whether there were statistically significant differences between the final sections of trials. Table 4.5 summarizes the results.

Distance Metric	F	p	Significant?
Unnormalized	2.937	< 0.0373	Y
Normalized	1.802	< 0.1523	N
Global Tempo	1.294	< 0.2813	N
Insertion/Deletion	0.942	< 0.4239	N

Table 4.5: Summary of results of a repeated-measures ANOVA with training condition as the independent variable and the final trials' recall performance as the dependent variable.

While the UD distance metric showed a statistically significant ($p < 0.037$) effect of training condition, no significant main effect was found for the ND, GT, or ID distance metrics. Pair-wise Bonferroni t-tests again confirmed the statistical similarity between A+H and A+H(atten) conditions and determined there to be a significant difference ($p < 0.006$) between the A and H conditions with the A condition leading to a 13% improvement in error over the H condition.

Summary

In summary, all training conditions showed significant differences between first and last trials under the UD and ID metrics. No significant differences were found under the GT metric and only the A condition showed a significant difference under the ID metric. An ANOVA run on early trial data showed a significant main effect of training condition under the UD, ND, and GT metrics, but not under the ID metric. Under the UD metric, a 10% error reduction for the A+H versus H training condition was found in early trials while error

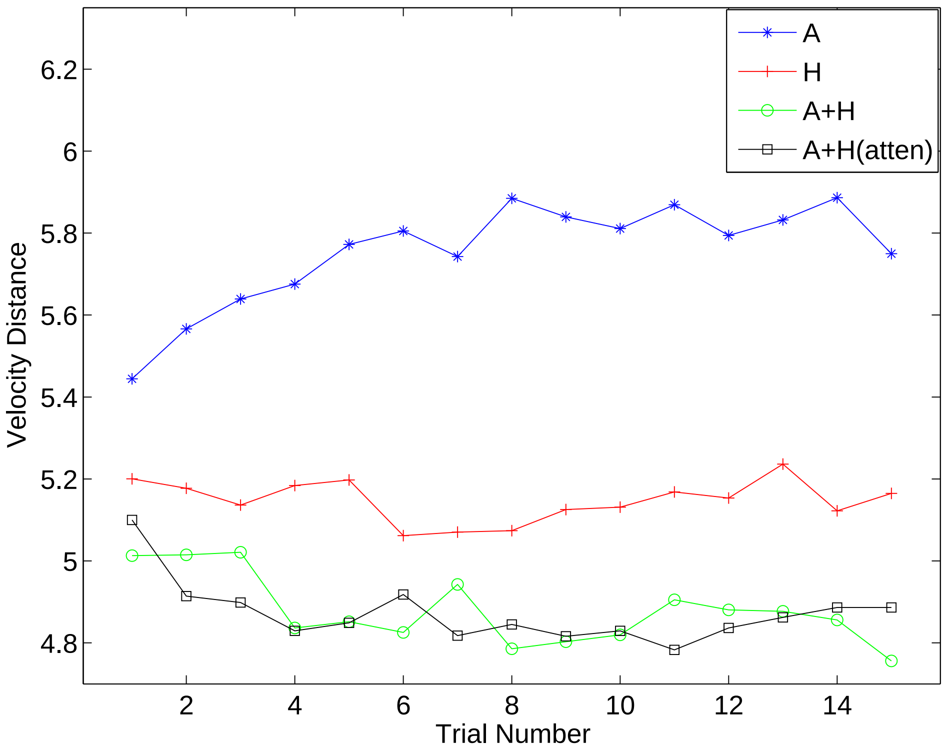


Figure 4-4: Average recall (mean across subjects) under the VD metric.

reductions of 18% were found between the A and A+H conditions and between the H and A+H conditions under the GT metric. ANOVA results for final trials showed a significant effect of training condition under the UD metric with the A training condition leading to a 13% improvement in error over the H condition.

4.4.2 Velocity

We now turn to note velocity and evaluation of subject performance under the VD metric. As with the evaluation of timing performance, a series of repeated-measures ANOVAs with training condition as the within-subjects factor, were used to compare VD scores across training conditions. Bonferroni-corrected pair-wise t-tests were then used to examine the individual effects of training conditions.

	F	p	Significant?
Early Trials	5.768	< 0.0012	Y
Final Trials	13.676	< 0.0001	Y

Table 4.6: Summary of results of three repeated-measures ANOVAs with training condition as the independent variable. The dependent variables were mean recall performance across all trials (global), mean recall across early trials, and mean recall performance across the final trials.

Early Trials

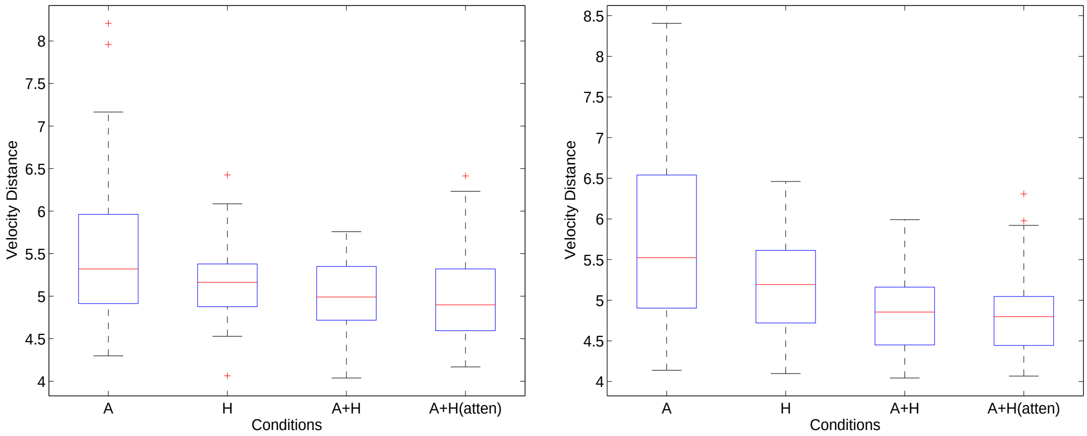
Looking at velocity recall during the early stages of learning as suggested in Figure 4-4 and Figure 4-5(a) we can see a possible difference in training condition. As before, this was tested by running a within-subjects ANOVA with the mean data from the first three trials as the dependent variable. A significant ($p < 0.0012$) main effect of training condition was found and t-tests (see Table 4.7) confirmed that the A and A+H conditions differed significantly which corresponded to a 10% reduction in error under the A+H training condition. The difference between the A and H conditions was of marginal significance with 7% less error under the H training condition than the A training condition.

	{A,H}	{A,A+H}	{H,A+H}	{A+H,A+H(atten)}
Early Trials	$p < 0.0164$ (M)	$p < 0.0009$ (Y)	$p < 0.3194$ (N)	$p < 0.7696$ (N)
Final Trials	$p < 0.0003$ (Y)	$p < 0.0001$ (Y)	$p < 0.0769$ (N)	$p < 0.9813$ (N)

Table 4.7: Pair-wise t-test results for comparisons between training conditions under the VD metric. p-values for each combination of training condition and distance metric are given along with whether the value is significant (Y), not significant (N), or marginally significant (M).

Final Trials

An ANOVA was also run on the mean of the last three trials to test for differences between training conditions in terms of the later portion of the learning process (final five trials as in the timing experiments). A significant main effect was found here as well ($p < 0.0001$) with paired t-tests showing significant differences between the A and H conditions ($p < 0.0003$) as



(a) Early trials summary

(b) Final trials summary

Figure 4-5: Summary of data distributions under the VD metric.

well as between the A and A+H conditions ($p < 0.0003$). The data summaries in Figure 4-5(b) again reveal a fairly clear trend towards lower error when haptic guidance and lowest error when both audio and haptic guidance were used. In fact, the H training condition showed an 11% reduction in error over the A condition, while the A+H condition showed a 17% reduction in error over the A condition.

Learning Across Trials

A somewhat puzzling result suggested by the recall curves in Figure 4-4 is the apparent trend towards *worse* performance across trials when subjects were trained using audio-only (A) guidance. A set of paired t-tests were run to check whether the per-training condition distributions for the first and last trials were significantly different. Surprisingly, the only training condition that differed significantly between the first and last trials was the A+H condition ($p < 0.0464$).

Summary

Of the four training conditions, only the A+H training condition differed significantly between the first and last trials. The results of the velocity data analysis showed significant main effects of training condition for both early and final trial data. Only A and A+H conditions differed significantly during early trials, with A+H leading to a 10% reduction in error. For the final trials, there were significant differences between the A and A+H (A+H lead to a 17% error reduction) conditions as well as H and A (H lead to a 11% error reduction) conditions.

4.4.3 Questionnaire

All subjects filled out a brief questionnaire on their musical experience and background. This questionnaire asked the following:

1. Age
2. Gender
3. Do you or did you ever, play a musical instrument?
4. Do you have any formal training in music theory and if so, how many years?
5. How many hours of music do you listen to per day?

Of the 32 subjects recruited for the study, 24 had played a musical instrument (other than percussion) before. A summary of the remaining questionnaire results is given in Table 4.8.

	μ	σ
Age	27.8	7.95
Years Playing	6.98	5.44
Years Trained	1.29	2.42
Hours Listened	2.97	4.44

Table 4.8: Summary of the questionnaire results

Subject Subpopulations

Given the availability of this additional subject data, one might wonder whether if and how the questionnaire results related to the rhythmic performance results. Perhaps surprisingly, no interesting differences were found when subjects were grouped by age, gender, listening habits, training, or by whether they played an instrument or not. One possible issue with this type of subdivision is that sample size begins to become a problem as groups of subjects are excluded based on certain criteria. Another potential difficulty that could be encountered, particularly when using within-subjects experimental designs, is rebalancing. Recall that in our experiment, subjects were randomly assigned to a group where each group used a particularly order of training conditions and mappings from rhythmic task to training condition in the experiment. If only some of the subjects are used, then the groups may not be represented evenly which could potentially introduces biases to the experimental results.

4.5 Discussion

4.5.1 Effects of Attenuated Hearing

Recall that the reason for including the A+H(atten) condition was to test whether the presence of earplugs and headphones (not playing any masking sound) affected recall results since subjects wore these during the H test condition but not the A or A+H test conditions. If it did not affect the recall results, then there should not be a significant difference between the A+H and A+H(atten) training conditions. The results presented in Section 4.4 confirm this hypothesis as none of the t-tests run for any of the analyses found a statistically significant difference between the A+H and A+H(atten) conditions. This, therefore, validates the experimental design difference between the H and A/A+H test conditions.

4.5.2 Timing

Learning Across Trials

As we might expect, for the most part performance increased between the first and last trials with all conditions showed significant improvement under the UD and ND metrics. Results were more varied under the GT and ID metrics, however. No conditions showed significant improvement under the GT metric, although audio improved marginally. This is surprising given the curve shapes in Figure 4-3(c) which, particularly for the H and A+H conditions, suggest downward trends. However, there was significantly more variance in the GT data of all training conditions, which helps explain this finding and implies that learning of global tempo was somewhat inconsistent as compared to the other metrics.

The non-significant results for the A+H condition under the ID metric is also interesting as it suggests that the combined information present in the A+H training condition may have hindered improvement given that results were significant for the A condition and marginally significant for the H condition. However, when we look at the significantly better performance of the A+H condition after the first trial, it becomes clear that the lack of difference between first and last trial has more to do with data variance than lack of improvement.

Early Trials

As can be seen from the results presented in Section 4.4.1, the combination of haptic and audio guidance appears to provide some advantage over audio or haptic guidance alone, particularly in the early stages of learning. The UD, ND, and GT distance metrics all show some kind of trend in the first few trials towards lower error rates with the A+H training condition than the A or H training conditions alone. From looking at the subject means of the first few trials in Figure 4-3(a) and Figure 4-3(b), it appears that under the UD and ND metrics, the hybrid training conditions (A+H and A+H(atten)), incur less error than the A and H conditions. Indeed this was the case as there was a 10% reduction in error when subjects trained with the A+H condition versus the H training condition.

Under the GT metric, there is some amount of statistical support for this type of difference as well (see Section 4.4.1).[3] In this case, the A+H training condition led to 18% less error than both the A and H training conditions. This is significant, as it suggests that subjects were able to take advantage of both sources of information and learn more effectively

[3]In fact, given the conservative nature of the Bonferroni correction, it is not unreasonable to pay some attention to effects with marginal significance.

than when either one alone was present. The fact that this difference occurred for early learning and for the global tempo metric is also interesting. Recall that this distance metric compares the overall length (in seconds) of two note sequences. Learning overall length or tempo is a more abstract and subconscious type of learning than what was measured under the other metrics (which largely test how well subjects remembered discrete items in time). Therefore, from these results, it appears that combined audio-haptic guidance is particularly effective at quickly giving learners a sense of tempo.

Although it may seem surprising at first that early trial results should show significant differences between training conditions, it is important to keep in mind that subjects had practiced the task rhythm twice before even the first test run (each trial contained two practice runs followed by a test run). Also, recall that early trial analyses used the mean of the first three trials' data as discussed in Section 4.4.1. It is therefore not unreasonable that performance could differ between training conditions during the early stages of learning.

Final Trials

The results of analysis of the final trials show that training with haptic guidance alone led to greater timing error at the end of the trial sequence than did the other training conditions. The results also show no statistical difference between the final error values when subjects were trained using audio alone versus when they were trained using a combination of audio and haptic guidance. Together, these results suggest that the presence of audio information, either alone or in conjunction with haptic guidance, is responsible for the lower final error.

This finding is actually in line with some of the most closely related prior research. Two studies on haptic guidance which were presented in Chapter 2, had results similar to this. Feygin et al. [12] and Liu et al. [30] both looked at haptic guidance in terms of how it compared with vision, both in isolation and in combination. Both of these studies found haptic guidance only based training to be inferior to vision alone and vision with haptic guidance in terms of recall performance. This confirmation of earlier results is interesting given the significantly different nature of the experiments. While the present study compared audition to haptic guidance, these two earlier studies looked at vision. Feygin et al. found that, for position and shape accuracy, haptic guidance did not appear to benefit learning, particularly when vision is available during recall. They relate their results to prior work on the perceptual dominance of vision and suggest that this dominance may degrade proprioceptive influence and interfere with the learning process [2]. In some respects, audition and vision are very similar as they both provide exteroceptive feedback and therefore a similar conjecture may be appropriate. However, there are also clearly significant differences between the two modalities (for instance in terms of information bandwidth) making a direct comparison difficult. Additionally, to date there has been relatively little research on the role of audition in motor learning and behavior [39].

4.5.3 Velocity

Early Trials

The effect of training condition on early trial performance was significant, with particularly pronounced differences between the A and other training conditions (see Figure 4-4). Since the H and A+H conditions did not differ significantly while the A and A+H conditions did (and the A and H pairing differed marginally), it appears that the presence of haptic guidance information was primarily responsible for the better performance levels of the H and A+H conditions. The improvement in performance when training included haptic guidance was substantial. Compared to the early trial error levels of the audio training condition (A), the haptic guidance only (H) condition showed 7% less error while the combined haptic guidance with audio (A+H) condition showed 10% less error. As in the case of timing performance, these results suggest not only that the presence of haptic guidance was particularly effective at reducing error, but also that subjects were able to make effective use of both proprioceptive and auditory information when they were present during training.

The results also show a smaller amount of variance (see Figure 4-5(a)) for the conditions which included haptic guidance than for the audio (A) condition. This suggests that haptic guidance based training can produce more consistent performance as well as more accurate performance in early trials.

Final Trials

The velocity recall results for the final trials show perhaps the most striking differences of all of the experimental analyses. While there was no difference between H and A+H conditions, there were significant differences between the A training condition and the H and A+H training conditions. When subjects trained with haptic guidance only, the average error was reduced by 11% as compared to when they trained with audio only. The difference was even more pronounced for combined haptic and audio training where error was reduced by 17% when compared to audio training alone. The conclusions made above about the primary role of haptic guidance in determining recall performance apply here as well, as does the observation that subjects were able to make more effective use of both sources of sensory information when they were available concurrently.

Looking at Figure 4-5(b) we see a sizable difference in variance between training conditions that use haptic guidance and those that don't (i.e. the audio only condition) which again suggests that haptic guidance leads to more consistent performance than audio only training. Interestingly, the variance for the H condition is larger than the A+H condition, which suggests that not only does the combination of audio and haptic guidance lead to superior performance in terms of measured error, but it also leads to more consistent performance as well.

Learning Across Trials

One interesting aspect of the velocity results is the lack of change in performance over time under the A and H training conditions. Although Figure 4-4 suggests a trend of *worse* performance across trials for the audio training condition, the difference between first and last trial was not statistically significant (providing a good example of why one should never blindly trust plots). The lack of a significant difference between first and last trials under the A condition is also evident from the boxplot summaries in Figure 4-5. They show a larger amount of variance for the A condition (versus the other training conditions) and that the early and final trial distributions for the A condition overlap substantially.

Chapter 5

Conclusions

5.1 Summary

This thesis sought to explore the idea of using haptic guidance in a music pedagogy context. The design and construction of two mechanically-assisted percussion interfaces was presented along with experimental evaluation of one of the devices. When haptic and audio guidance based training was compared to audio only based training, an average error reduction of 17% was observed for recall of note timings during the early stages of training. When the same training conditions were compared for the recall of note velocities during late training, an 18% reduction in error was found. These results indicate that haptic guidance can be effective at teaching non-trivial motor skills. Additionally, the fact that this technique can yield substantially lower error rates than more traditional training schemes, suggests it is an extremely fertile area for further research and development.

5.2 Future Work

System Improvements

There are a number of ways in which the HAGUS device could be improved, although the most useful change would be to increase the number of degrees of freedom addressable by the device. The simplest version of this expansion would be to simply build a second device so that two-hand rhythmic tasks could be taught. This could then be expanded to included foot-actuators so that the system could be used to teach the types of multi-limb coordination required to play a drumset. Increasing the drumstroke realism of the device is also an area that could be greatly improved. The current single-axis approximation to drumming is clearly not realistic and a more complete multi-axis model capable of targeting elbow and shoulder movements is needed.

Another area where the hardware could be improved is the position data sampling rate. Although 60Hz provides an adequate amount of spatio-temporal resolution for the relatively simple and slow rhythms that were used in the experiments described in Section 4, we would like to be able to use the system for more advanced material in the future. This would necessitate a higher sampling rate so that fast rhythms could be used. At present, this problem appears relatively easy to solve by using a real-time or even just low-latency Linux kernel.

Finally, there are minor mechanical issues with the current state of HAGUS that could be worked out. There is a small amount of backlash in the gearing used to drive the drumstick which could probably be improved somewhat. This would likely just require re-fabricating the gear set. Decreasing the amount of static friction in the drivetrain would also be desirable. As most of this drag seemed to come from the electromagnetic particle clutch, it may be worth looking at other ways to engage and disengage the drivetrain.

Future Experiments

The encouraging results presented and discussed in Chapter 4 lead to a number of ideas for future experiments. One major question that was not addressed by the experiments presented here is long-term retention. There is evidence that short-term (within a few minutes) and long-term (24 hours or more) retention can vary greatly, particularly when forms of augmented feedback have been used [20, 20, 45, 3]. The use of dynamic training schedules where guidance or augmented feedback is provided less and less frequently with time is one technique that has been used to counter this effect. Therefore one obvious follow-up experiment would be to test subject timing and velocity recall after 24 hours, comparing subjects who had used a static training schedule versus those who had used a dynamic schedule.

It would also be interesting to test how much, if any, of the observed advantage of the A+H condition is due to arousal effects. One could argue that the confluence of auditory and proprioceptive sensory information could have a stimulating effect on subjects which in turn could boost performance during multimodal training conditions. Because this possibility cannot be ruled out with the current data, a new set of experiments would need to be conducted. One possibility is to train subjects using the A+H condition and once their performance has stabilized, change to the H (or A) training condition. If the error level returns to the stable value reached under the A+H condition, it would suggest that multimodal arousal played a role in performance.

Another natural extension of the experiment presented in this thesis is to look at the effect of haptic guidance on coordinated motor tasks (such as those involved in drumset performance). Coordinated movements are possibly even more resistant to verbal explanation than simple movements and therefore may be particularly well-suited to haptic guidance based training.

There are also numerous ways that assistive technology could be used to explore pedagogical tactics that would not be possible in a traditional setting. For example, using negative dampening, task difficulty could be made artificially difficult to boost the error signal that subjects presumably use during the natural learning process. Another interesting possibility is to exaggerate the task motions so that subtle percussion gestures that might otherwise go unnoticed would be easier to perceive.

Finally, there is the possibility of using the existing hardware in an adaptive training mode [29]. Research suggests that there is a task difficulty “sweet spot” where learners are challenged but not frustrated, and that learning is most effective when the material presented to the learner is kept within this range of difficulty [19, 42]. Given that HAGUS provides real-time data about subject performance, it would be relatively straightforward to extend the system so that it adjusted the task difficulty depending on subject performance. This adjustment could simply involve tempo change or something more complex such as simplifying the rhythm. There are a number of aspects to this setup that could be interesting experimentally, such as how often to adjust the task difficulty and how severe the change should be. One could even imagine varying *these* parameters based on subject performance.

5.3 The Future of Music Pedagogy

Throughout history, music and music pedagogy have undergone constant change as new theories, techniques, and devices have been developed. This, of course, continues to be the case with computers, the internet, and new instruments being developed to advance musical pursuits. Given technology’s rapid adoption into other areas of education, it seems all but certain that it will play a larger and more central role in music pedagogy as well.

It is my belief that the primary way in which technology will influence musical performance education is through devices for home instruction. However, this is not to say that such systems could or should replace human teachers. On the contrary, I would expect that these devices would be used *by* teachers to ensure that students practice properly between lessons. This ability to monitor and correct technique during home practice seems particularly important as most students spend the vast majority of their playing time in the absence of a teacher.

One could also imagine significant implications of such systems for people living in rural communities or other places without physical access to music teachers. With the widespread adoption of the internet, remote instruction has recently started to take place. While the internet is certainly richer than text or video as a medium, it suffers from the same fundamental difficulties when used to communicate gesture. However, imagine if students had in-home haptic guidance systems that teachers in remote locations could control over the internet. This would allow for a rich student-teacher interaction that could rival (and possibly even exceed) standard pedagogical methods.

Bibliography

[1] H. Abdi. *Encyclopedia of Measurement and Statistics*, chapter Bonferroni and Sidak Corrections for Multiple Comparisons. Sage, Thousand Oaks, CA, 2007.

[2] J. Adams, D. Goper, and G. Lintern. Effects of visual and proprioceptive feedback on motor learning. *Journal of Motor Behavior*, 9(1):11–22, 1977.

[3] T. Armstrong. Feedback and perceptual-motor skill learning: A review of information feedback and manual guidance. Technical Report 25, University of Michigan, Department of Psychology, Ann Arbor, MI, 1970.

[4] I. Bilodeau. *Acquisition of Skill*, chapter Information Feedback. New York: Academic Press, 1966.

[5] C. Black and D. Wright. Can observational practice facilitate error recognition and movement production? *Research Quarterly for Exercise and Sport*, 71(4):331–339, 2000.

[6] Y. Blandin, L. Lhuisset, and L. Proteau. Cognitive processes underlying observational learning of motor skills. *Quarterly Journal of Experimental Psychology Section A - Human Experimental Psychology*, 52(4):957–979, 1999.

[7] E. S. Boyden, A. Katoh, J. L. Pyle, T. Chatila, R. W. Tsien, and J. L. Raymond. Selective engagement of plasticity mechanisms for motor memory storage. *Neuron*, 51(6):823–834, 2006.

[8] E. S. Boyden, A. Katoh, and J. Raymond. Cerebellum-dependent learning: The role of multiple plasticity mechanisms. *Annual Review of Neuroscience*, pages 581–609, 2004.

[9] E. S. Boyden and J. L. Raymond. Active reversal of motor memories reveals rules governing memory encoding. *Neuron*, 39(6):1031–1042, 2003.

[10] C. Chaffe and M.S. O'Modhrain. Musical muscle memory and the haptic display of musical nuance. In *Proceedings. Intl. Computer Music Conference*, 1996.

[11] G. Ferguson and T. Yoshio. *Statistical Analysis in Psychology and Education*. Montreal, Quebec: McGraw-Hill Ryerson Limited, 6th edition, 2005.

[12] D. Feygin, M. Keehner, and R. Tendick. Haptic guidance: Experimental evaluation of a haptic training method for a perceptual motor skill. In *Proceedings of the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 40–47, 2002.

[13] S. Fink. *Mastering Piano Technique: A Guide for Students, Teachers, and Performers*. Amadeus Press, 1992.

[14] P. Fitts. *Categories of Human Learning*, chapter Perceptual-Motor Skill Learning. New York: Academic Press, 1964.

[15] P. Fitts and M. Posner. *Human Performance*. Greenwood Press, 1967.

[16] R. Gillespie. *Haptic Displays of Systems with Changing Kinematic Constrants: The Virtual Piano Action*. PhD thesis, Stanford University, Department of Mechanical Engineering, 1996.

[17] R.B. Gillespie. The touchback keyboard. In *Proceedings of the International Computer Music Conference*, pages 447–448, San Jose, CA, October 1992.

[18] R.B. Gillespie, M.S. O'Modhrain, P. Tang, C. Pham, and D. Zaretsky. The virtual teacher. In *ASME International Mechanical Engineering Conference and Exposition*, Anaheim, CA, November 1998.

[19] M. Guadagnoli, L. Dornier, and R. Tandy. Optimal length for summary knowledge of results: the influence of task-related experience and complexity. *Research Quarterly for Exercise and Sport*, 67(2):239–248, 1996.

[20] J.D. Hagman. Presentation and test-trial effects on acquisition and retention of distance and location. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9:334–345, 1983.

[21] D. Holding and A. Macrae. Guidance, restriction, and knowledge of results. *Ergonomics*, pages 289–296, 1964.

[22] F. Huang, R. Gillespie, and A. Kuo. Haptic feedback improves manual excitation of a sprung mass. In *Proceedings of the 12th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 200–207, 2004.

[23] Golandsky Institute. `http://www.golandskyinstitute.org`, 2007.

[24] J. Kelso. Planning and efferent components in the coding of movement. *Journal of Motor Behavior*, 9:33–47, 1977.

[25] H. Krebs, M. Ferraro, S. Buerger, M. Newbery, A. Makiyama, M. Sandmann, D. Lynch, B. Volpe, and N. Hogan. Rehabilitation robotics: Pilot trial of a spatial extension for mit-manus. *Journal of NeuroEngineering and Rehabilitation*, 1(5), 2004.

[26] H. Krebs, N. Hogan, M. Aisen, and B. Volpe. Robot-aided neurorehabilitation. *IEEE Transactions on Rehabilitation Engineering*, 6(1):75–87, 1998.

[27] T. Lee, M. White, and H. Carnahan. On the role of knowledge of results in motor learning: Exporing the guidance hypothesis. *Journal of Motor Behavior*, 22(2):191–208, 1990.

[28] Jeff Lieberman. Accelerated and improved motor learning and rehabilitation using kinesthetic feedback. Master's thesis, Massachusetts Institute of Technology, Auguest 2006.

[29] G. Lintern and D Gopher. Adaptive training of perceptual motor skills: Issues, results, and future directions. *Journal of Man-Machine Studies*, 10:521–551, 1978.

[30] 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 Neurorehabilitation*, 3(20), August 2006.

[31] D. Morris, H. Tan, F. Barbagli, T. Chang, and K. Salisbury. Haptic feedback enhances force skill learning. In *WHC '07: Proceedings of the Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 21–26, Washington, DC, USA, 2007. IEEE Computer Society.

[32] C. Nichols. The vbow: Development of a virtual violin bow controller haptic-human interface. In *Proceedings of New Interfaces for Musical Expression*, pages 29–32, 2002.

[33] Maura Sile O'Modhrain. *Playing by Feel: Incorporating Haptic Feedback Into Computer-Based Musical Instruments*. PhD thesis, Stanford University, 2000.

[34] O. Ortmann. *The Physiological Mechanics of Piano Technique*. E.P. Dutton, Inc., 1962.

[35] T. Perneger. What is wrong with the bonferroni adjustments. *British Medical Journal*, 136:1236–1238, 1998.

[36] D. Reinkensmeyer, L. Kahn, M. Averbuch, A. McKenna-Cole, B. Schmidt, and Z. Rymer. Understanding and treating arm movement impairment after chronic brain injury: Progress with the arm guide. *Journal of Rehabilitation Research and Development*, 37(6):653–662, 2000.

[37] A. Salmoni, R. Schmidt, and C. Walter. Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin*, 95:355–386, 1984.

[38] R. Schmidt, C. Lange, and D. Young. Optimizing summary knowledge of results for skill learning. *Human Movement Science*, 9:325–348, 1990.

[39] R. Schmidt and T. Lee. *Motor Control and Learning: A Behavior Emphasis.* Human Kinetics Publishers, Inc., Champaign, Illinois, 3rd edition, 1999.

[40] R. Schmidt, D. Young, S. Swinnen, and D. Shapiro. Summary knowledge of results for skill acquistion: Support for the guidance hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2):352–359, 1989.

[41] S. Swinnen, R. Schmidt, D. Nicholson, and D. Shapiro. Information feedback for skill acquisition: Instantaneous knowledge of results degrades learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4):706–716, 1990.

[42] J. van Merrienboer, L. Kester, and F. Paas. Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognition Psychology*, 20:343–352, 2006.

[43] K. Wang and G. Gasser. Alignment of curves by dynamic time warping. *Annals of Statistics*, 25(3):1251–1276, 1997.

[44] C. Winstein and R. Schmidt. Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4):677–691, 1990.

[45] 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*, 65:316–323, 1994.

[46] A. Woods and A. Stahl. *Elementary Mechanism: A Text-book for Students of Mechanical Engineering.* D. Van Nostrand Co., 1909.

[47] G. Wulf and R. Schmidt. The learning of generalized motor programs: Reducing the relative frequency of knowledge of results enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15:748–756, 1989.

[48] G. Wulf, C. Shea, and C. Whitacre. Physical-guidance benefits in learning a complex motor skill. *Journal of Motor Behavior*, 30(4):367–380, 1998.

[49] Y. Yokokohji, R. L. Hollis, T. Kanade, K. Henmi, and T. Yoshikawa. Toward machine mediated training of motor skills - skill transfer from human to human via virtual environment. In *5th IEEE International Workshop on Robot and Human Communication*, pages 32–37, 1996.