

**Perception-based Design,  
Including Haptic Feedback in Expressive Music Interfaces**

by

Ricardo Pedrosa

Eng. “José Antonio Echeverría” Higher Polytechnique Institute, Cuba, 1994

MSc. “José Antonio Echeverría” Higher Polytechnique Institute, Cuba, 1998

A THESIS PROPOSAL SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

The Faculty of Graduate Studies

(Computer Science)

The University of British Columbia

June 2007

© Ricardo Pedrosa 2007

## ABSTRACT

When designing haptic feedback into human-computer interfaces the most common approach is to restrict its deployment to only one of the haptic channels (either force or vibrotactile feedback) and this decision is often biased according to the technological boundaries imposed by the interface design or its purpose.

The research presented here outlines a methodology developed to include haptic feedback in a gesture interface used to implement a computer based music instrument. The primary goal of this research is to achieve an increased understanding of how different flavors of haptic feedback should be combined to support both controllability and comfort in expressive interfaces such as computer based musical instruments.

The work reported here has two parts: 1) a description of an initial user experiment conducted for the purpose of assessing the use of vibrotactile stimuli as rhythmic cues, and 2) a proposed research trajectory which will constitute the theme for a PhD project.

The main hypothesis guiding this research is that the best way to include haptic feedback into any design should start by understanding the role each haptic channel plays in the “real life scenario”. The assumption we will put into test is that the feedback provided for each haptic channel could be developed individually and later on fine tuned when mixed in the final design, and this includes that the addition of any extra cues we might be interested in supplying to the end user to enhance the real life experience should be designed and mixed in each haptic channel before the final blend occurs.

This methodology could be defined as a “*perception-based design*”: a design approach that starts with a perceptual characterization of the interaction that is being addressed and fine tune the results according to the perceptual load put on the user senses.

We expect that the knowledge gathered will be useful not only to a broader spectrum of expressive interfaces but could also be extrapolated to including haptic feedback in other interface designs.

# TABLE OF CONTENTS

<b>ABSTRACT</b>	<b>ii</b>
<b>TABLE OF CONTENTS</b>	<b>iii</b>
<b>LIST OF FIGURES</b>	<b>v</b>
<b>LIST OF TABLES</b>	<b>vi</b>
<b>ACKNOWLEDGEMENTS</b>	<b>vii</b>
<b>1. INTRODUCTION</b>	<b>1</b>
<b>1.1. Problem statement</b>	<b>1</b>
<b>1.2. Music and perception</b>	<b>1</b>
<b>1.3. Haptics as a possible design approach</b>	<b>3</b>
<b>1.4. Preliminary study: Use of vibrotactile feedback as a haptic cue for rhythm</b>	<b>4</b>
<b>1.5. Case Study Approach: Grounded Gesture Controller</b>	<b>5</b>
<b>1.6. Proposed Research Scope</b>	<b>6</b>
<b>1.7. Proposal structure</b>	<b>7</b>
<b>2. RELATED WORK</b>	<b>9</b>
<b>2.1. Haptic perception in music performance</b>	<b>9</b>
2.1.1. Perceptual attributes of music performances	9
2.1.2. Musician-instrument interaction: a feedback control approach	9
2.1.3. Role of haptic channels in music performance	11
<b>2.2. Gesture interfaces and gesture-based communications</b>	<b>13</b>
2.2.1. Affordance in electronic music interfaces	13
2.2.2. Gestures in every day life	14
2.2.3. Gestures in music performance	15
2.2.4. Gesture interfaces for music performance	16
<b>2.3. Haptic feedback in electronic music instruments</b>	<b>17</b>
2.3.1. Ergotic relations in Human Computer interfaces	17
2.3.2. Design issues for computer music interfaces	18
2.3.3. Directions for including haptic feedback in computer music interfaces	19
<b>2.4. Conclusions</b>	<b>21</b>
<b>3. OBJECTIVES</b>	<b>23</b>
<b>3.1. Motivation</b>	<b>23</b>
<b>3.2. Research hypothesis</b>	<b>24</b>
<b>3.3. Goals and contributions</b>	<b>25</b>
3.3.1. Goals	26
3.3.2. Anticipated contributions	26
<b>3.4. Impact</b>	<b>28</b>
3.4.1. Human-computer interaction (HCI)	28

3.4.2. Haptics	28
3.4.3. Psychophysics and Music	29
<b>3.5. Disambiguation of Research Scope (Negative Space)</b>	<b>29</b>
<b>4. EFFECTS OF HAPTIC AND VISUAL CUES ON ABILITY TO FOLLOW RHYTHMS</b>	<b>30</b>
<b>4.1 Approach</b>	<b>30</b>
<b>4.2 Prototype</b>	<b>31</b>
<b>4.3. Experiment Design</b>	<b>33</b>
<b>4.4. Experiment Results</b>	<b>34</b>
<b>4.5. Discussion</b>	<b>35</b>
<b>4.6. Conclusions from the experiment</b>	<b>39</b>
<b>5. MANAGEMENT PLAN</b>	<b>41</b>
<b>5.1 Approach</b>	<b>41</b>
5.1.1. Platform	41
5.1.2. Development of a semantic base for a gesture interface	42
5.1.3. Sound Generation	43
5.1.4. Design and deployment of the haptic feedback	43
5.1.5: Merging of tactile/force feedback	44
<b>5.2. Road map</b>	<b>45</b>
5.2.1. Stage 0: Phantom enhancement with vibrotactile feedback	45
5.2.2. Stage 1: Building a Semantic Base	45
5.2.3. Stage 2: Design of the force feedback	46
5.2.4. Stage 3: Design of vibrotactile feedback	47
5.2.5. Stage 4: First Mixing Stage	48
5.2.6. Stage 5: Second Mixing Stage	49
<b>5.3. Risk Points</b>	<b>49</b>
<b>5.4. Resources</b>	<b>50</b>
<b>5.5. Future Work</b>	<b>52</b>
5.5.1. Haptic feedback and different end applications of the interface	52
5.5.2. Degradation of quality	52
5.5.3. Expanding the scope of the proposed methodology	52
<b>REFERENCES</b>	<b>54</b>
<b>APPENDIX A User Study Consent Form</b>	<b>59</b>
<b>APPENDIX B Ethics Approval Certificate</b>	<b>63</b>

## LIST OF FIGURES

Figure 1. Main research Stages	7
Figure 2. Hardware components connection diagram	32
Figure 3. Software components connection diagram	33
Figure 4. Difference in performance between groups who received audio feedback of their tapping vs. those who did not	36
Figure 5. Differences on observed difficulty level for all subjects	37
Figure 6. Errors observed between users for different difficulty levels and conditions	37
Figure 7. Example of variation between reference and subject performance for a single rhythm pattern	39
Figure 8. Research Stages and Research Time Frame	45

## LIST OF TABLES

Table 1. Sample test design _____	34
Table 2. Statistics results. ANOVA analysis on a 2x4x3 Split Plot design _____	36

## ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Karon MacLean for all her support, encouragement and understanding. Also I would like to thank Dr. Michiel van de Panne for being my second reader and Dr. Jim Little for his understanding and consideration.

I would also appreciate the time and advice provided by Dr Alan Kingstone and Dr Keith Hamel. It will be a pleasure to continue the work presented in this thesis with you.

Also to the people at the SPIN Research Group: Mario Enriquez, Steve Yohanan, Colin Swindells and Dave Ternes: thanks for all the support and advice and for making my life here more pleasant.

And last but not least to my wife Beatriz for being there. This thesis is for you.

## 1. INTRODUCTION

In this section, we provide an overview of the whole proposal and highlight the main topics that will be covered in detail later on.

### 1.1. Problem statement

Today's computer-based music interfaces depart so greatly from traditional acoustic instruments that they can be hard to recognize as instruments at all. A computer-based music instrument has two distinct components: an interface controller and a sound synthesis module. The sound is synthesized using a software module whose parameters can be arbitrarily mapped to any interface controller, and thus designers have complete freedom in terms of user control approaches and methods. This freedom represents an opportunity to improve on usability issues such as ease of learning and ergonomics which are often associated with traditional acoustic music instruments.

Though some of the most common electronic music interfaces take the shape of already existent music instruments (like keyboards, electric guitars or wind controllers, to mention a few) they do not provide significant usability improvement over their acoustic counterparts beyond the ability to control various sound synthesis methods. Some more novel interfaces, such as gesture controllers, provide a higher level of expressivity and are easier to learn than acoustic instruments; but they are limited in terms of controllability and do not capture the characteristic "*feel*" of a music instrument.

The research proposed in this document aims to deliver one path through which the *feeling* of using a musical instrument can be attained in gesture-controlled computer-music interfaces without compromising the expressivity and ease-of-use that have already been achieved by this route. Our goal is to better understand, using a scientific, controlled approach, how to best combine the expressivity of a gesture controller with the feel and control of a contact-type interface, meanwhile taking advantage of the new abilities of digital music synthesis.

### 1.2. Music and perception

There are three perceptual modalities involved in a musical performance, whether solo or by a group:

- **Auditory:** Channel through which the performance's main outcome flows. It is involved in



musician-musician, musician-instrument and musician-audience interactions. It is the principal source of control feedback for musicians.

- **Visual:** Though the main outcome in a music performance (the sound) could be augmented with the use of other visual stimuli to enhance the performer-audience interaction (lights, video or still images), we are more interested in the visual aspects of a music performance that are linked with the way the sound is produced and that have some incidence in the musician-musician and musician-audience interactions. These include both the gestures used to produce the sounds as well as the particular “*body language*” of this artistic expression through which more abstract meanings and emotions are conveyed; both can be visually available to the audience.
- **Haptic:** Primarily involved in the musician-instrument interaction, this channel is a complex combination of (broadly speaking) two distinctive components. From the standpoint of a haptic interface, these are *force feedback* (which relate the human’s proprioceptive and kinesthetic sense of bodily movements and forces) and *vibrotactile feedback* (supplying the human tactile sense with object or system features such as vibrations, texture, shape and temperature).

These three modalities are involved in one way or another in the construction of musical instruments. To fulfill the main objective of achieving some distinctive and enjoyable sound, a *luthier* (a craftsman who makes stringed instruments) must incorporate ergonomic issues into his/her design in such a way that a musician could then control the sound-generating mechanism. These ergonomic precepts will define the movements the musician must make to generate the music and to establish an extra channel of communication with other musicians and the audience. In the construction stage, the *luthier* has to carefully choose the different materials involved in the process: from the type of wood for the different parts, their shape and thickness or the glue used to put everything together, to the paint to provide an adequate finishing. All of this aims to fulfill one objective: the instrument's body should amplify the sound produced by the vibration of the sound-generating mechanism and, most important, should resonate in a way that reinforces certain frequencies of the sound spectrum, thus defining the timbre for the particular instrument [FleRos1998].

To reach a high performance level, the musician must spend years to achieve an understanding of the nuances of the instrument’s mechanical sound-generating interface. Details varying from the amount of force necessary to press a string in a guitar to knowing the moment

in which the hammer is released to hit the string in a piano are felt through the haptic channels and used to close a control loop at reflexive rates, once trained. As will be shown in Section 2.1.3, the performer not only uses the forces that the instrument exerts against her/him, but also feels the instrument vibrations. The latter are part of the feedback the musicians use to fine tune their performance.

### 1.3. Haptics as a possible design approach

Despite all this, the haptic channel has received little attention since new computer interfaces for music began to appear, and this is because designing these interfaces departs drastically from the *luthier's* approach of selecting parts and configuring the instrument to amplify the sound generated.

To control a software synthesis module we need several triggers. From a strictly engineering point of view, buttons, sliders or knobs are sufficient for this purpose. However, because the sound in these systems comes from speakers that are usually placed away from the performer, the performer doesn't directly experience sound-related vibrations. To a classically-trained musician who is used to feel these vibrations, this basic interface feels no different than any other computer peripheral, and is inadequate. A further casualty of losing the intimate relationship with the instrument is the visual aspect of the performance that communicates the artist's effort to the audience: an interaction with knobs, sliders or a mouse with a GUI is not compelling to watch.

The first inclusion of haptic feedback in music came as a result of musicians' requests to incorporate the mechanical feeling of an acoustic piano into electronic keyboards. Designers started to notice that sound is not everything and that the way instruments *feel* in the hands of musicians is also important. But as opposed to the case of acoustic instruments where the haptic feedback is a physical artifact of the sound-generating mechanism (albeit one which has played a co-evolutionary role over a long period of time), in electronic instruments it must be deliberately added. This is especially true in the case of those interfaces that depart more drastically from the traditional concept of a musical instrument but yet have the potential for the highest levels of expressivity and usability. Section 2.2.4 presents some examples of this type of interface.

Besides the feeling of playing a musical instrument there is another issue with the type of interfaces known as *gesture controllers*. These interfaces map the gestures of performers in open air into sound. Despite being considered one of the most expressive interfaces for music, gesture

controllers rely on a level of proprioception (the performer's knowledge of his/her own body position) to a degree of accuracy similar to that required by dancers. Humans are better at controlling forces than position; so one of the ways to improve this interface is to supply resistive haptic feedback to increase the movement accuracy. Section 2.3.3 presents some of the approaches in this direction.

To our knowledge, there is no established method for including this type of feedback or which of the two "flavors" of haptic feedback (tactile and kinesthetic) are most appropriate for a given communication aspect. Even more, the inclusion of haptic feedback into an already designed and/or tuned interface could do more harm than good by overloading the user's perception or creating a distraction from the main purpose of the interface, whether it is music related or not. On top of all this, in expressive interfaces the feeling of comfort is as valuable as the feeling of being in control of the process.

The primary goal of this research is to achieve an increased understanding of how the different channels of haptic feedback should be combined to support both controllability and comfort in expressive interfaces such as computer-based musical instruments.

#### **1.4. Preliminary study: Use of vibrotactile feedback as a haptic cue for rhythm**

Many commercially available musical instruments include visual cues as a means of improving user performance. However, the exact extent of how helpful the presence of additional stimuli is has not been examined to a satisfactory degree. We present a preliminary study of the possible effects of visual and haptic cues on the ability to follow rhythms developed as a group project for a Course in Human Computer Interaction in which the author had a leading role. By exposing users to a variety of different conditions and then gauging their ability to follow rhythms of varying difficulty, we subsequently collected and analyzed their error rate based on the number of notes misplaced.

While we were unable to prove with a significant degree of confidence that the presence of additional stimuli improved a subject's performance in learning/following rhythms, our experiment yielded some unexpected insight regarding the considerations one must take into account in the design of these types of studies, such as possible unintentional cultural bias in rhythm selection. Section 4 presents and discusses the results of this study.

## 1.5. Case Study Approach: Grounded Gesture Controller

A perceptual study of any real-life target context (be it music, sculpting, painting, driving or drilling a tunnel) could help us determine the role each perceptual channel plays there. A computer interface targeting the context could then be designed using a methodology that takes into account these roles. In this way, regardless of the eventual interface's morphology, we could be sure that the user will find some similarity to the real-life case.

Here, our target context is the general class of music instruments. As each music instrument has its own method of interaction, we are interested in identifying broad perceptual issues that hold across the instrument spectrum. We aim to equip the interface with the haptic feeling of a music instrument, as opposed to any other tool or instrument. The perceptual characterization of the musician-instrument interaction should focus on issues related to controllability and comfort in performance, as well as on those related to bring the feeling of playing a music instrument. This perceptual approach could be a solution to the challenge of appropriate and effective inclusion of haptic feedback in music interfaces, considering the broad variation and inherent lack of constraints in computer-music interface design and design approaches.

A methodology based on such perceptual factors could be tested through a case study focused on one particular musical interface type. We see two major advantages of a case-based approach. First, the general effectiveness of incorporating perceptually guided decisions into the interface design procedure could be tested. Secondly, we can evaluate whether the inclusion of haptic feedback in one particular instance helps to mitigate known problems that occur when that interface is used as a music controller.

The specific platform and configuration we will therefore use throughout this research is a *grounded gesture controller*, supplying both force and vibrotactile feedback. By *grounded* we mean that the interface is fixed to a rigid surface and can thus provide reaction forces to the user, a quality that is difficult to achieve in wearable and mobile devices. Our focus is on designing and evaluating the *effects* of haptic feedback in relation to computer music applications, rather than the design of haptic interfaces themselves; therefore, we plan to base our investigation on an existing haptic device (a PHANTOM Premium 1.5 from SenSable Technologies [Sens]). This will also allow testing our hypothesis in a best-case scenario, since this device is considered an industry standard for a force feedback haptic interface. The PHANTOM will be augmented with a handle to supply tactile feedback in addition to the force feedback the device already provides.

## 1.6. Proposed Research Scope

Following this hardware construction, our research roadmap is depicted in Figure 1. These stages (three “design” and two “mixing”) are detailed in Section 5 along with a roadmap and timeframe, and described in overview here.

**Stage 1 – Creation of a Semantic Base:** The semantic base contains the gesture vocabulary that will later be used to control each of the perceptual music attributes of interest. We are considering the evaluation of pitch, loudness, tempo and rhythm as the music parameters to control. However, their final selection will depend on a group of experiments designed in this stage to build a set of meaningful basic gestures customized to match the interface constraints. These basic gestures will be used to construct more complex music commands or phrases in the same way that we construct sentences in oral or written communication. The name “*semantic*” hints that in this stage we are interested in the meaning of each gesture, rather than in the structure of control commands. As in any expressive interface, the syntactic structure is up to the performer. For instance, we can indicate the door with a hand-wave to communicate that someone should leave a room. The difference between “*Please, go this way*” and “*Get out!*” could be conveyed by how quickly the gesture is performed. In this research, we are specifically interested in finding gestures that convey the meaning of raising a pitch or sustaining a note, for example. The level to which a pitch is raised and how fast it is reached (the music phrase syntaxes) will be up to the performer.

**Stages 2-3 – Design of Force and Tactile Feedback:** A unique aspect of the proposed research is our plan to separate the use of force and vibrotactile feedback according to the perceptual purpose these channels play in the target application: using a music instrument. As in a traditional acoustic music instrument, the force feedback designed in Stage 2 will be directed at enforcing instrument controllability while the vibrotactile feedback designed in Stage 3 will provide the feeling of playing a music instrument as well as some enhanced perception over certain music parameters.

**Stages 4-5 Mix Stages:** The integration and fine-tuning that will take place at each mixing stage serve the purpose of obtaining the desired level of *control* (Stage 4) and the *feeling of playing a music instrument* (Stage 5). On top of this, when tuning for their specific purpose in each stage we will also try to maximize *comfort* during performance. As a result of this approach, we could end up with an optimum combination of comfort and control (ie. in Stage 4) that might not match the case where maximum controllability is achieved. With comfort here we

refer to a combination of perceptual/cognitive load and other physical factors including ergonomics. Rather than depending on one interface feature and its correct tuning, the expected level of comfort should be achieved by a careful combination of features during the whole design process.

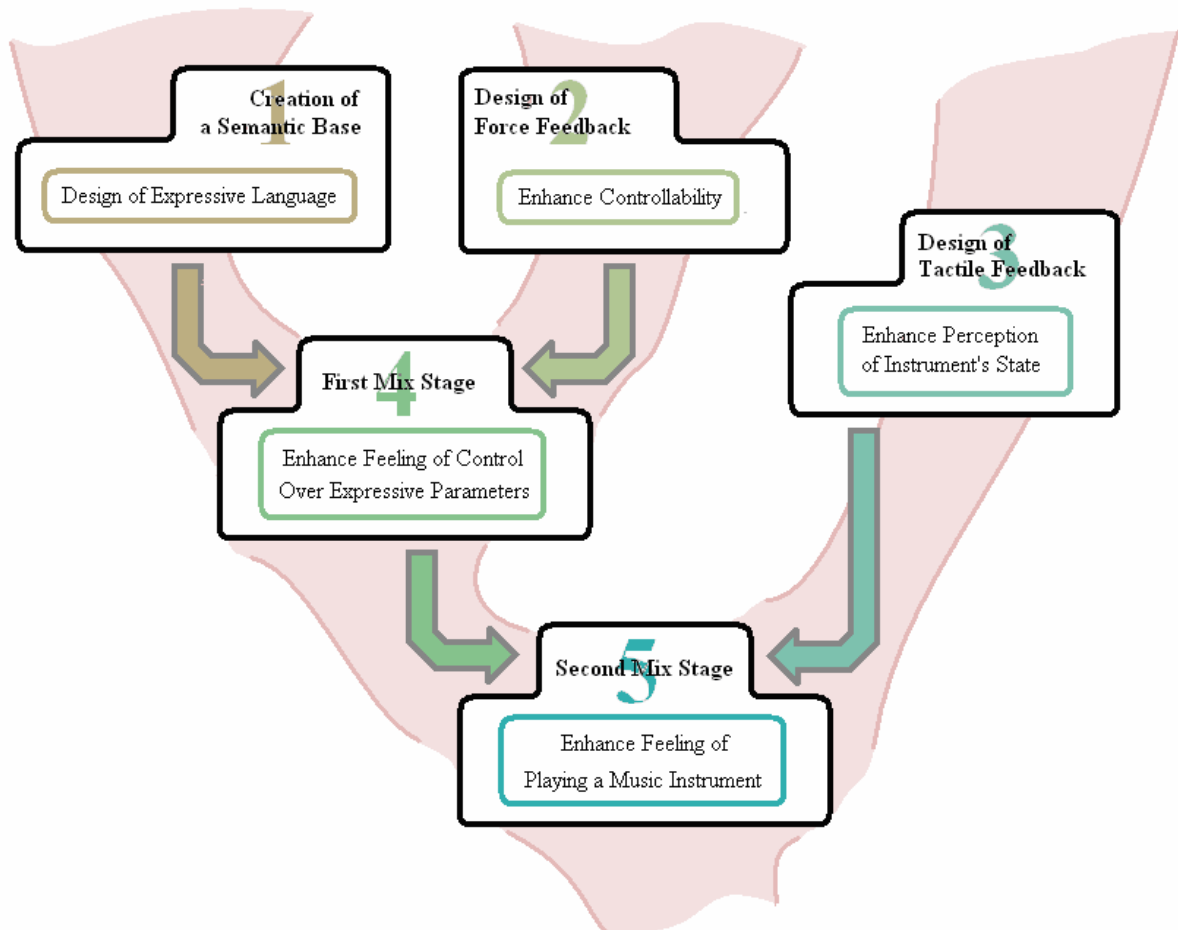


Figure 1. Main research Stages

## 1.7. Proposal structure

The rest of this proposal will be structured as follows. Section 2 provides a literature review of the work of others that is related to this research. Section 3 details the hypotheses, goals, expected contributions and impact of the proposed research. Also we identify in this section the “negative space” of those issues that we are *not* dealing with in this research. Section 4 presents the experimental work already performed to test the effect of visual and haptic cues on the ability to follow rhythms. Section 5 describes the management plan we intend to pursue, as well as a

---

deeper description of the approach, each research stage, resources needed and features we are proposing as future work.

## 2. RELATED WORK

This section presents a literature review of three main areas relevant to the research proposed: the perception of haptic parameters in music performances, gesture-based interfaces and the inclusion of haptic feedback in computer music interfaces.

### 2.1. Haptic perception in music performance

#### 2.1.1. Perceptual attributes of music performances

A performance of music contains the following 7 perceptual attributes: pitch, timbre, loudness, rhythm, tempo, contour, and spatial location [Lev1999]. All of them are psychological constructs that relate to one or more various physical properties of the sound we hear. *Pitch* and *loudness* relate directly to frequency and amplitude. *Timbre* is defines the source of a particular sound and distinguishes it from other sources (e.g. the sound of a violin from the sound of a flute); it is related to the sound's frequency content and the relative magnitudes of the various spectral components. *Rhythm* and *tempo* relate to time: the first is directly connected to patterns of sounds that repeat in a music piece, while the second is related to the overall speed at which the melody or music piece is played. *Spatial location* defines where the sound source is located. It is a more complex perceptual attribute that arises from several factors including the intensity and the frequency content of the sound plus the time differences between the perception of sounds coming from the direct sound source and its reflection in the components of the environment. *Contour* refers to the shape of a melody when musical interval size is ignored.

These musical attributes are independent and any one could be changed without changing the others. In this research we are going to deal only with pitch, loudness, tempo and rhythm. The remaining three parameters, though an important part of the expressive nature of this performance, are either instrument specific (ie. timbre), could be set-up before hand (ie. spatial location) or could be modified by introducing changes in any or various of the already selected parameters (ie. contour).

#### 2.1.2. Musician-instrument interaction: a feedback control approach

Musical instruments are complex mechanical resonators, and this means that in order to produce sounds they must vibrate. These vibrations are also felt by the performer and it is known



that they are used to help him/her in tasks that range from determining if the note is already settled and stable in the instrument [Cook1996] to tune the instrument to a nearby one relying on the way one instrument vibrates in response to the vibrations transmitted through air or floor [Gill1999] [AskJan1992].

Gillespie [Gill1996] provided a broad definition of a musical instrument as: “*a device which transforms mechanical energy (especially that gathered from a human operator) into acoustical energy*”. Thus, the musician-instrument relationship could be represented as a feedback control system [Gill1996][Gill1999][O’Mod2000] where the control loop is closed over two paths via the auditory and the haptic channels. The major feedback from the instrument to the musician occurs through the air in the form of sound waves. The mechanical contact between the musician and the instrument (fingertips, hands or mouth just to mention a few possibilities) acts as an additional bidirectional flow channel through which the musician sends information in the form of mechanical energy, and receives from the instrument the haptic information associated with the sound generated and the status of the sound control mechanism.

Nevertheless, a skilled musician can do (at least briefly) without auditory feedback during the performance of a known piece or while reading a score. After years of training, the musician is able to perform a form of anticipatory control over the instrument where she/he anticipates an instrument response to a given manipulation. Finney [Fin1997] showed that piano performance was only affected when the musicians were presented with a delayed audio feedback and not when they were deprived from it. Also the experiments he conducted showed that changing the pitch in a synchronized auditory feedback cause little or no impairment but changing it in the delayed feedback significantly reduced the amount of delayed auditory feedback impairment. In this case a change of pitch could be understood as if the sound is coming from another source and not from the performer actions, diminishing the negative effects of delaying the feedback due to a weaker action-effect association.

However, the effect of delayed auditory feedback is only a problem when it is not consistent and predictable by the musician. Though most instruments react almost instantly to the actions performed by the musicians, some instruments (e.g. church pipe organs) are known to possess a noticeable delay between the performer action and the production of the sound. This delay is assimilated as part of the instrument internal representation and mastered by skilled organists in a way that virtuoso performances are achievable.

Repp [Rep1999] measured several expression parameters in piano performance (as opposed to

accuracy in the case of Finney's studies) and showed how the effects of auditory deprivation were very small except for some substantial changes in pedaling by some pianists. As Repp said:

*“Although expression seems to be controlled primarily by an internal representation of the music, auditory feedback may be important in fine-tuning a performance and in the control of pedaling. However, it is also possible that the effects of auditory feedback deprivation merely reflect a lack of motivation to play expressively in the absence of sound.”*

### **2.1.3. Role of haptic channels in music performance**

Gillespie in the works leading to his PhD degree [Gil1996] uses impedance theory and bond graph representation to describe a music instrument as a two-port system: two variables, an “effort” and a “flow”, are present at each of the ports (in this case, mechanical on one side of the junction, and acoustic on the other). At each port, if one variable is an input, i.e. constrained by the upstream system, then the other is forced to be an output, i.e. determined by the value of the input and the dynamic system on the other side of the junction. The variables used in this model are force and velocity at the mechanical port, and pressure and flow at the acoustical port. As his work was focused on reproducing the haptic feeling of the piano action mechanism, Gillespie was more concerned with the mapping from mechanical input to mechanical output or in other words, how the interface presents itself mechanically (haptically) to a performer when it is manipulated.

However, for the purposes of the research presented in this proposal, this model has an important limitation: it includes no path between the acoustic output (the pressure or flow rate of the produced sound waves) and the mechanical output (velocity or force, respectively, as experienced by the musician). In other words, the path through which the musician feels the vibrations produced by the instrument's body resonating at the frequency of the produced sound. This path does occur naturally in music instruments and has an important effect in distinguishing a music instrument from a sledge hammer, as it refers to the way the instrument vibrates to produce a sound.

Verillo [Ver1992] presents results which support the theory that the sensory capacities of the skin are such that tactile feedback could be used by some instrumentalists in controlling the tone of their instruments; he offers further evidence that singers utilize vocal cord vibration and the chest cavity's resonance in the same way. His own studies together with relevant literature which

he cites signal the fingers as the most sensitive sites for vibration sensing. Finger (non-glabrous) skin frequency response ranges from near 0 Hz to approximately 1000 Hz, enabling feedback from sounds in that range to be felt if the sound source is in contact with the skin and the sound is loud enough. Even when most instruments have a dynamic range that is considerable higher than this, it is also known that there are sub-harmonics of higher frequencies that appear at this range before a note is completely settled and their amplitudes are high enough as to be felt by the performer [Gill1999a]. However, frequency discrimination is quite poor: only differences in frequency on the order of 20 or 30% are detected at the fingers.

Verillo also presented further evidence supporting the notion that tactile feedback can be used to control timing, in addition to tonal control, and that kinesthetic cues especially coming from the upper extremities are an important factor in musical performance. These kinesthetic cues, as we will see later, are part of the gesture language used in the communication between musicians and between musicians and the audience.

In a thorough study of vibrations in four traditional stringed instruments, Askenfelt and Jansson [AskJan1992] provide enough evidence to assert that mechanical vibration occurring in musical instruments (generated either by proper instrument operation or as a reaction to the sound coming from other instruments) are powerful enough to be felt by the musician during regular performance and that these vibrations are not limited to the parts of the instrument designed to radiate sound (e.g. vibrations at the guitar neck could be felt, though they are weaker than those at the guitar body). In the conclusions the authors say:

*“It would be tempting to conclude that the major feature of the instrument’s vibrations is to convey a feeling of a resonating and responding object in the player’s hands. (...) Regardless of the type of instrument, it could be assumed that the kinesthetic finger forces offer more guidance for timing purposes than the vibrotactile stimuli supplied by instruments’ vibrations.”*

This last statement also has ties with the statement by Repp cited at the end of the previous section where he talks about the musician’s internal representation of music. This internal representation is an asset enforced with training and the development of what came to be known as muscle memory or motor programs [LedKla1997][Mag1989][Gill1999]. The motor programs are believed to reside in the cerebellum and are part of a higher level of control, the ones in charge of triggering major events. A lower level of control is located at what is known as spinal reflexes, a closed loop formed by mechano-receptors and the spinal cord, and responsible for

tasks such as tightening a grip to avoid a sliding object to fall from the hand. After repeating the same movements for years and getting used to the mapping between action-sound, the musician could achieve high degrees of perfection in performance. A motor program is triggered at a higher level, dictating the melody: spinal reflexes take charge of regulating the force and speed of finger or arm movements to perform the expected melody. Signals coming from mechanoreceptors at skin, muscles and bone junctions can provide enough information to fine-tune the movements at low levels.

Taken together, these past works strongly suggest that the haptic channels (force and vibrotactile feedback) have distinct functions and play an important role in both fine-tuning the performance and defining the instrument status. The development of low-level sensori-motor programs through practice and training depends heavily on the musician's apprehension of the relationship between force exerted and sound produced. Vibrations coming from the instrument's sound generation engine close the loop in the same mechanical channel to keep the control loop at the same low level that started the action, while reinforcing the feeling of using a musical instrument, or a *“resonating and responding object”* in the words of Askenfelt and Jansson.

## **2.2. Gesture interfaces and gesture-based communications**

*“For more things affect our eyes than our ears.”*

*Jean-Jacques Rousseau, “Essay on the Origin of Languages”*

### **2.2.1. Affordance in electronic music interfaces**

The common denominator between electronic music instruments and especially those based in computers is that the sound-generating mechanism or synthesis module is physically (mechanically) decoupled from the interface that controls it (the controller) [ParO’Mod2003]. The major consequence of this is that the affordance of an electronic music controller can be very different from that of the acoustic instrument whose sound is being synthesized. In other words, one interface can now be used to control different sound engines; and correspondingly, when the mapping changes, any transfer effect due to the physical knowledge one develops by interacting with acoustic instruments is lost.

The mechanical sense arising from past acoustic-instrument interactions leads to the immediate apprehension or cognition, without reasoning, of an instrument’s basic capabilities; as opposed to detailed knowledge about how to operate it to obtain a particular sound, or to make it

function properly. For example, we “know” what type of action (whack, pluck, blow or bow) should be exerted upon a particular instrument just by looking at it and can approximately predict how this will sound; although unfortunately, this does not mean we will also have the physical know-how to extract the right or desired sound from it.

This lack of “*intuitiveness*” in electronic music instruments comes from the fact that the synthesis module in these instruments is effectively a mathematic algorithm whose parameters are (from a technological point of view) easily mapped to knobs, pushbuttons and levers. Further, extra “functions” like panning, reverb and compression, which generally entail enabling followed by amplitude control, is generally achieved with switches or toggle commands. And this is only one side of the problem. To increase versatility these generic user interfaces are intended for many different synthesis modules. As a result, in one setup a knob might control the oscillator frequency, and a modulation index in another.

One good example of how this need for increased expressivity leads to hidden or arbitrarily mapped functionality is the inclusion of *aftertouch* in digital keyboards. This feature is activated (on those keyboards that are equipped with it) when an already-depressed key is pressed harder. The resulting effect on the sound is up to the performer, for she/he has to map it to the desired parameter. A simple exploration of the interface won’t reveal this capability. As this feature needs to be activated (mapped to a particular sound parameter or effect) the only way of discovering it is by reading the instrument manual.

### **2.2.2. Gestures in every day life**

Humans communicate with each other not only through speech but also with physical expressions conveyed through gestures, ranging from frowning to waving to the settling of body postures. Some authors consider gestures as the opposite of postures: “*gestures are dynamic, and postures are static*” [Mul2000]. Others limit the gestures to “*hand movements that contain information*” [KurHul1990]. We prefer the definition given by [Hum-etal.1998] as “*a movement that conveys meaning to oneself or to a partner in communication*”. Much can be conveyed through gestures, from simple orders to nuanced emotive responses. It makes sense then that gesture interfaces have found a great support both from industry and from the scientific community as a way to improve human-computer interaction. As Camurri and Volpe [CaVo2004] point out:

“*On one hand, recent scientific developments on cognition, on affect/emotion, on*

*multimodal interfaces, and on multimedia have opened new perspectives on the integration of more sophisticated models of gesture in computer systems. On the other hand, the consolidation of new technologies enabling “disappearing” computers and (multimodal) interfaces to be integrated into the natural environments of users are making it realistic to consider tackling the complex meaning and subtleties of human gesture in multimedia systems, enabling a deeper, user centered, enhanced physical participation and experience in the human-machine interaction process”*

### **2.2.3. Gestures in music performance**

One of the most important features in music performances is the communication between the artist and the audience. Not only is the audience able to perceive the effort and ability of the performer through his/her gestures, but also “*the musician is able to communicate to the public the inner meanings and movements of the music being played*” [LemSty2005]. In music performances we found both grounded and un-grounded gestures. By *grounded* we mean those gestures performed in force-exchanging contact with the instrument, and the ones that enforce the musician’s skills. Ungrounded gestures are performed in air, like the ones executed by orchestra conductors or those resulting from musician’s body movements and are the one responsible for conveying the music “*inner meaning*”.

On most computer-based music instruments, the most complex pieces are performed by pressing buttons or modifying on-screen parameters through standard computer interfaces and the audience awareness of the complexity of these pieces and their inner meaning is lost.

Music, as a form of art, could take us into an emotional state by the way different sounds are combined. Those sounds are generated by performers’ gestures acting upon their instruments and it is not a new discovery that many of those gestures *encode* a particular evoked emotion. Gesture interfaces are therefore considered by many as the new paradigm for computer-based music instruments, both because of the inherent intuitiveness they accommodate and because of the capability to form a “common language” that could be used to master control ranging from note-level events to more complex musical phrases. Therefore from now on, we will use the term gesture in the same fashion as Miranda and Wanderley [MiWa2006], to mean “*any human action that could be used to generate sound*”.

As the sound engine could be considered an autonomous and exchangeable section of electronic music instruments (ie. one MIDI synthesis module could be effectively controlled by

different MIDI controllers) , there's no reason the interface design could not take advantage of other research fields and incorporate those elements that foster expression and communication as well as intuitiveness as long as it remains flexible enough to control several sound modules.

Of course the problem is how to build an interface to capture those gestures. Considering our area of application, the biggest challenge is deploying an interface that allows at least the same freedom a musician has when playing a given traditional instrument. This means: not to be an unbearable burden both from the cognitive and from the physical point of view, not be distractive and above all not be annoying.

#### **2.2.4. Gesture interfaces for music performance**

The most logic candidate to solve this problem, given the current state of the art of sensors and transducers, is to use some sort of remote sensing of an artist's gestures, mapping each one of them into an acoustic event or a control method. This group of controllers receives the generic name of "*open air controllers*" or "*immersive controllers*" according to the classification given by Mulder in [Mul2000]. Mulder further divides this group into *Internal*, *External* or *Symbolic Controllers*, taking into account the visualization of the virtual control surface and the mapping strategies rather than the input device used. From these types of interfaces there are several variations in use. Some, like the *Lightning* family of Buchla's devices [Buch], rely on infrared sensing; others like the *Radio Baton* [Mat1991] rely on electric field sensing through capacitive transducers. These two types of controllers restrict the performer movements to the sensor's working space. Other less restrictive controllers falling in the immersive classification are those based on data gloves, like Laetitia Sonami's *Lady's Glove* developed by Bert Bongers [Bon2000] or body suits like Sergi Jorda's *Exoskeleton* [Jord2002] or Bromwich and Wilson's *Bodycoders* [BroWil1998].

Though this type of controllers seems to be the ultimate solution to our problems, there is still one issue they don't address: the intimate physical interaction that exists between a musician and the instrument she/he plays. The most important reason behind why these controllers have not found a warmer welcome inside the music community lies in the extreme departure from the established definition of music instruments and the change in the control paradigm that a gesture controller carries with it. Gesture controllers demand from the performer a high degree of proprioception (the unconscious perception of movement and spatial orientation arising from stimuli within the body itself) and egolocation (awareness of one's overall position within a

defined space). There is a high level of reliance on visual and auditory feedback to achieve accuracy in the planned movements [RoHa2000]. Musicians are forced to develop the same body control as a dancer, and this is something that also requires many years to develop, a totally different set of skills and continual concentration.

### 2.3. Haptic feedback in electronic music instruments

#### 2.3.1. Ergotic relations in Human Computer interfaces

Cadoz and Wanderley [CaWa2000] developed an operational gesture typology that would allow giving a clear answer to the question “*When is haptic feedback needed?*”

By focusing on gestures in the music environment, Cadoz and Wanderley took on the division provided by Delalande [Del88] while studying the playing technique of the pianist Glenn Gould. This division groups the gestures in three categories:

1. *Effective gesture* - necessary to mechanically produce the sound - bow, blow, press a key, etc.
2. *Accompanist gesture* - body movements associated with effective gestures - chest, elbow movements, mimics, breathing for a piano player, etc.
3. *Figurative gesture* - perceived by the audience but without a clear correspondence to a physical movement - a melodic balance, etc.

The typology proposed identifies three different functions associated with the gesture channel:

1. *The ergotic function*: material action, modification and transformation of the environment.
2. *The epistemic function*: perception of the environment.
3. *The semiotic function*: communication of information towards the environment.

If with environment here we identify the sound or music produced, we can see how the majority of computer music controllers (as well as some electronic music instruments) are part of a non-ergotic chain reduced to simple sensors and a sound synthesis process. The link between this controller and the sound synthesis is done through a control process mapping gestures into sound [HuWaPa2003]. There exists no direct interaction/modification of the element generating the sound and therefore there is no ergotic relationship with the environment.

The ergotic action-sound relation, even when it is the most frequent in daily life, was the last



one to be considered in general computer environments [LuFiCa2005], Luciani et al. notes:

*“Ergotic action-sound situations suppose an energetic consistency throughout the chain from the action to the sound. Their implementation in a computer context requires the energetic consistency not to be broken at any stage from the haptic and the acoustical transducers. For example, the rendering of hitting or rubbing a sound object, requires not only force feedback devices but also adequate real-time simulation of adequate physically-based models and adequate links between them. This means that all these elements must be designed in a same, modeling process of the whole.”*

This brings on a set of requirements on bandwidth, latencies and dynamic ranges that become more important when dealing with virtual reality scenarios or when simulating the “feeling” of an existent acoustic instrument into an electronic one.

### **2.3.2. Design issues for computer music interfaces**

Computer-based music instruments tend (as a general rule) to go in a different direction than exact reproduction of existing acoustic music instruments. As Rován and Hayward [RoHa2000] said:

*“In spite of the ubiquitous MIDI keyboard, the performance tool for many computer musicians today does not fit the piano paradigm. Whether it is a bank of MIDI faders, data glove, video tracking device or biological sensor system, the instrument of choice is increasingly a system that allows for a diverse and personalized set of performance gestures. They are often one-of-a-kind instruments, tailored to the needs of a particular performer. Idiosyncratic and possibly temperamental, these instruments are designed by musicians who are part performer, composer, as well as hardware engineer. With a collection of sensors they seek to translate their personal language of gesture into sound.”*

There are two important ideas shining in the background of this statement. The first one is related to the reduced scope of these types of instruments, something that Cook [Cook2001] also pointed out in a paper that is considered by many a classic in this field:

*“Musical interfaces that we construct are influenced greatly by the type of music we like, the music we set out to make, the instruments we already know how to play, and the artists we choose to work with, as well as the available sensors, computers, networks, etc. But the music we create and enable with our new instruments can be even more greatly influenced by our initial design decisions and techniques.”*

Cook also presented in this paper as one of the principles for designing computer music controllers the necessity of making “*a piece, not an instrument or controller*”. While this is a strong argument for not wasting time reinventing the wheel in a way that if a musical piece could be played with an already existent instrument then there's no need for a new one, sadly this could also lead to advocate for designing “*disposable instruments*”. And not that there's something wrong with this approach in certain musical pieces, but the same principle holds for re-using those already created instruments. Of course, this could only be done if they were designed with enough flexibility to allow the inclusion of new functionality.

The other point Rovin and Hayward mention is related to the persons behind the design of these new musical interfaces; in their words: “*(...) musicians who are part performer, composer, as well as hardware engineer*”. And again besides this being the ultimate solution to the always present problem in the field of Human-Computer Interface Design of excluding the final users from the design process, this could also be considered the cause of most of the naive designs we found in this field. By using a computer as a music instrument, a person with a bit of interest could become a *luthier* from dusk to dawn. The availability of software synthesizers waiting for simple signals to trigger their sounds, plus the availability of cheap sensors and data acquisition system combined with easy to use programming environments facilitates the development of “*musical instruments*”.

However, only with a true knowledge of the perceptual nuances occurring in the performer-instrument interaction can these interfaces could pass the border signaling the land of musical instruments. The vast majority of (if not all) the computer-music instruments designed so far remain in the hands of a selected *avant-garde* group of musicians looking for new sounds or ways of expression and who are willing to accept their limitations.

The jump of this type of instruments into a broader community is blocked by a series of factors emerging mostly from the limited or closed perspective of the design (as a direct consequence of the “*one piece, not one controller*” strategy mentioned above) and from the drastic departure taken from the established concept of musical instrument.

### **2.3.3. Directions for including haptic feedback in computer music interfaces**

The inclusion of haptic feedback in computer music instruments is often the approach taken in order to make these new interfaces more attractive to the general music community, by fostering

(among other attributes) the expressivity or controllability of the interface. Sometimes the use of haptic feedback is the only way to foster interaction with certain new sound synthesis algorithms, such as the case of the interfaces developed by O'Modhain and Essl to control granular sound synthesis models and friction-induced sounds [O'ModEss2004][EssO'Mod2005]

Another way to achieve this effect of controllability is by injecting a dose of realism when trying to mimic the feeling and dynamic behavior of a traditional music instrument. In this direction it is worth noticing the work of Gillespie [Gil1992][Gil1994], who designed and built a seven-key touch-programmable keyboard where the dynamical effects of all the elements of the piano action (excluding the wooden key) could be recreated. This system used force feedback to present several renderings ranging from a pianoforte to a grand piano action using a modified virtual wall model controller.

In this same view, Chafe [Chaf1993] showed how the performance of a controller for a real-time physical model of a brass instrument is improved when the audio output of the synthesis is fed back to the controller. Nichols [Nich2002] developed a violin-like haptic interface to control a bowed-string physical model as a way to overcome the difficulties of achieving an accurate or expressive translation from performance gestures to synthesized sounds using a commercial MIDI violin.

Cadoz [Cad-etal.1990] provided another view of the subject. The *Modular Feedback Keyboard* designed at the *Association pour la Creation et la Recherche sur les Outils d'Expressions (ACROE)* could control any sound synthesis engine by mapping the synthesis parameters to the appropriate instrumental gestures that would cause them. The interface was based on a piano keyboard and contains custom designed motors and sensors for each key. The forces fed back to the performer derive from the modeled object (sound or images) as it evolves over time as a result of the actions executed over the synthesis model. The interface used forces to enhance the controllability of the media and took its base in the forces presented by a piano to a performer. The interface was designed in a modular way in that each key has its own sensors and actuators. Therefore many configurations aside from a traditional piano could be constructed, including some other means of interactions like joysticks allowing 2 or 3 degrees of freedom. There was no independent vibrotactile feedback included, though the motors could provide the feeling of vibrations. The major scope was to foster the “*touch-synthesis*” approach through which the performer could modify a sound by feeling it through the responses of a known interface.

A different approach was conducted by O'Modhain [O'Mod2000][O'ModChaf2000] by showing the potential benefits of including haptic feedback in instruments that in the real world provide no haptic cues to the player. The results of her experiments showed how the accuracy of playing a musical piece on a Theremin-like instrument increased when a force proportional to the magnitude of the controlled parameter opposed the performer movements. In other experiments, O'Modhain also showed how the presence of those same forces confused skilled players trying to use a simulated bowed instrument with bow-string friction incorporated in the haptic feedback. She assumed that this effect was due to a severe mismatch between the performer's inner representation of the instrument and how accurately it was modeled in the interface used. Even though no improvement in performance was found on this last case, the participants in the study preferred the model with friction to the one without it, a result that points out to an improved quality of the simulation as perceived by those experienced players.

With respect to open air gesture controllers, several approaches have been taken to address issues such as lack of consistent control and unrepeatability results. Bongers [Bon1994] used *Muscle Wires* (a type of thin shape memory alloy that shortens in length when electricity is applied) to control the flexibility of the fingers in a glove controller. The resulting controller enhanced with force feedback was used to provide the performer with a feeling of a sound as it traveled through space. Cadoz [Cad1994] instead of shape memory alloys used pneumatic devices to mimic the contraction of muscles. Chu [Chu1996] proposed a haptic MIDI controller to localize sound in space by providing tactile audio signals to both hands. Rovin and Hayward in the above mentioned paper enhanced an open-air controller based on a glove with tactile actuators interfaced with either the hands or feet of the performer to explore the perceptual attributes of a simple vibrotactile vocabulary synthesized in response to a gesture. This approach was aimed to reduce the reliance on visual cues and proprioception by restoring one of the missing haptic channels.

## 2.4. Conclusions

The literature review presented in this section illustrates and reinforces the following points:

- Gesture interfaces possess a high level of expressivity and are fostered as one of the solutions for multimodal communications.
- The two haptic channels (force and tactile perception) are actively involved in the intimate musician-instrument interaction.

- Force feedback could be used to foster the feeling of controllability in computer music interfaces, since this is the role that this haptic channel plays in the real life interaction between musician and music instruments.
- Vibrations generated by music instruments are felt by the musicians and play an important role in perceptually defining a music instrument as such.
- The design of computer-based music interfaces is usually done taking a limited scope that excludes the feeling of playing a musical instrument in favor of an improved controllability.
- Some of the most expressive music interfaces such as gesture controllers lack *both* controllability features as well as the feeling of playing a musical instrument.

The research proposed is aimed at supplying a way to bring computer-music interfaces closer to the standard concept of music instrument by addressing the issues of controllability and “feeling of playing a musical instrument” through an integrating methodology developed over a perceptual basis.

### 3. OBJECTIVES

This section will provide the motivation behind the proposed research. This includes the main hypothesis we are trying to prove, our goals and the contributions we expect to deliver to the involved fields as well as impact on several major research areas. Finally, we will describe those issues we are not dealing with in the proposed research.

#### 3.1. *Motivation*

When designing haptic feedback into human-computer interfaces, the most common approach is to restrict the design to only one of the haptic channels, either force or vibrotactile. The decision of which one should be included is often based on the technology constraints of the interface. For instance, in wearable or mobile devices where displaying forces is cumbersome at best, the preferred way to include haptic feedback is through vibrations.

This leads to a second important point: no matter what channel we use, some sort of coding is necessary in order to convey a meaning to the end user. The intended user needs to understand what's being presented to him. The more abstract the information we intend to present, the higher the cognitive load in the user. For instance, estimating the magnitude of a variable from the length of a buzz is not very demanding if we use a perceptually proportional ratio: a larger buzz indicates a big amount, a short one a small amount. But distinguishing two long buzzes with different frequency characteristics to identify the information as pertaining to forecasted rain levels versus expected wind speed is more demanding. Depending on the intuitiveness of the mapping, the user will have to go through some degree of a learning process to learn this association.

Synthesized haptic feedback can be separated into force and vibration, and the designer can potentially utilize this division to distribute the cognitive load. For instance, let's take the case of a car's GPS interface, an approach that will help us demonstrate this issue at a broad scale and introduce another one directly related to expressive environments .

The original visual interface of a generic car GPS system has recently been enhanced in many available systems with voice commands to inform the driver which way to turn, allowing the driver to maintain visual attention on the road. The already overloaded visual channel was relieved from one task and the user could still feel in total control. Some results obtained in our group also suggest that an even more intuitive effect could be achieved by providing some force

feedback to the steering wheel [ForMacL2006], without the problems of masking in an auditorily noisy car environment. In this case a force might be a better (more natural) way to indicate direction than a vibration. If a force is used the driver doesn't need to “decode” which way to turn (as with either a vibration or a non-spatial voice command), but simply follows the force presented. If a non-directional vibration is applied to the steering wheel the driver must first learn which vibration represents left and which right and later on recognize this in the distracting car environment, potentially leading to increased cognitive load; although this load might decrease with learning as it generally does for interpretation of well-known encoded auditory signals, including spoken language. However, the force approach has another issue. Though the force presented should be strong enough to be felt, it shouldn't limit the driver actions at all. The decision of turning or not must remain under the driver control at all times, as it is when an audio signal or a vibration is used. *Thus, whether visual, auditory or haptic, there can often be a basic tradeoff between the natural mapping of the system's signal, and on the other hand, undesired interference with the user's intention.*

In expressive environments, the criterion that the user “has the last word” is one of the most important ones. Expressive interfaces should inform and even provide “hints” on how to achieve certain solutions, but never interfere with the user's decisions. Every piece of information provided through feedback to improve control must be balanced in a way that the user feels comfortable with the interface and that everything blends into supporting the main task. How to reach this balance is one of the most difficult and sought-after goals of expressive interfaces design. The main motivation behind this research is to find a solution to this problem.

### **3.2. Research hypothesis**

The major hypothesis behind this research is that in any interface where haptic feedback is to be included, each haptic channel (force and/or vibrotactile) could be designed individually upfront and then fine-tuned when mixed in the final design. This includes that the addition of any extra cues we might be interested in supplying to the end user should be designed and mixed in each haptic channel before the final blend occurs. These extra cues might not play a direct role in enforcing controllability, but could reinforce the perception of more complex music parameters like contour. In order to put this into practice there are (at least) two conditions that must be satisfied.

First, it is strictly necessary to have a thorough knowledge of what is happening in the context

approached by the interface (e.g., the context of an expressive musical interface) and the role of each of the haptic channels in that situation. With that knowledge and a design strategy that takes it into consideration, the designer could achieve a higher transfer effect from related non-digital contexts (e.g. acoustic interfaces) and increase the usability criteria for both expert users (for them it is like using the tools they are use to) and beginners (already existent experience developed in other contexts could be applied here).

On the other hand, this “*divide and conquer*” strategy has a drawback that should be considered carefully. As stated in the previous section, in expressive interfaces the feeling of comfort is as valued as the feeling of being in control. There are several ways a feeling of discomfort may arise. We addressed one of them previously (undesired system feedback interfering with the user's own goals when these are not in accord). Most other ways are due to an overloaded perceptual channel or to interactions between information presented through different channels. For instance, when a supplied stimulus is too strong or it is maintained for too long it could numb the receptors to a point that consequent stimuli are not felt by the user. In other cases, one stimulus might mask another if they appear too closely in time, and information might be lost; further, tactile and kinesthetic signals might be more vulnerable to masking than, say, tactile and visual cues. Designers could take advantage of this last issue to avoid overloading, presenting some important information as the masking signal, because only that one will be felt. Again, to achieve this goal a thorough knowledge of perceptual issues is required.

This means that in order to improve an interface we should carefully pick which feedback signal best represents which control feature, such that the actions used to correct an event could be controlled by the user at lower levels while minimizing overhead for the main control unit: the user's brain.

### **3.3. Goals and contributions**

In the broadest sense, this research will examine ways of using current technologies to improve the design of computer interfaces for music applications. The focus will be on enhancing through the use of haptic feedback the feeling of playing a musical instrument, providing both the comfort and controllability necessary for such an expressive application.



### 3.3.1. Goals

Our primary goal is to achieve an increased understanding of how the different channels of haptic feedback should be combined to support both controllability and comfort, in expressive interfaces such as computer-based musical instruments. There are several areas that need to be investigated to achieve this goal.

**1. Design a natural gesture language that affords expressive performance by a novice.**

As stated in Section 1.4, the proposed research focus will be centered on a case study consisting of a gesture interface for an expressive music instrument. Therefore a portion of the research is devoted to creating a gesture language, to be used as the semantic base by which the performer can communicate his/her intentions to the sound synthesis algorithm. A successful gesture language will allow a novice musician to achieve acceptable results in less time than using a traditional music instrument, while maintaining expressiveness and engagement needed for skilled or even virtuoso performances.

**2. Demonstrate duality of haptic channels for musical control, and their effectiveness in their proposed functions.** The role of haptic feedback in traditional acoustic instruments as determined per the literature review will be evaluated in order to probe the main hypothesis of this research, i.e. that feedback for the two haptic channels can be designed independently to provide different roles, then successfully merged. Success can be measured in a number of ways: even though the interface design will depart from the standard form factor of a traditional music instrument, the presence of haptic feedback should still follow the same purposes as in any music instrument. Further, the haptic feedback should reinforce some features not present in acoustic music instruments as a way to foster several usability criteria such as learning rate and intuitiveness.

### 3.3.2. Anticipated contributions

What follows is a list of outcomes expected as a result of this research.

**Methodologies:**

- ***“Perception-based” guidelines for including haptic feedback in computer-based music instruments.*** This could be considered the major contribution of the proposed research. By departing from a typical process for including haptic feedback in human-computer interfaces (through application/technological constraints) towards a design

based on a perceptual evaluation of the most suitable use of haptics in each application, this methodology if proven correct could be used in any type of expressive interface design – not just limited to music.

The remaining contributions outlined below could be considered as being part of this mayor contribution but also are strong enough as to be considered relevant contributions *per se*.

#### **Models:**

- ***A semantic base appropriate for a set of grounded gesture interfaces focused on mapping the hand position in space.*** Though designed with music applications in mind, this semantic base will be formed from gestures that are used in everyday activities and not only those related to music. Therefore it could be usable as the language building blocks for other applications not strictly devoted to music, but using a similar gesture interface with the same constraints of workspace, user's limb used to activate and grasp.

#### **Generalizable Design Criteria:**

- ***Set of perceptual music parameters useful for mapping in an interface designed for real-time performance in order to increase its controllability.*** Music perceptual parameters are defined in the auditory domain (Section 2.1.1). However, not all contribute to enhancing the feeling of control over an instrument during performance. Determining which ones are suitable to be mapped through an auxiliary channel (not audio) will prove useful to enhance music interfaces generally.
- ***Natural associations between perceptual music parameters and haptic channels; and identification of their sensory limits.*** Due to the known bandwidth limitations of human perceptual channels, it is a priority to define which of the useful music perceptual parameters are more salient and give more meaning to the performer when mapped through haptic feedback. To further limit cognitive load it will be useful to know if there is some sort of natural association between one perceptual stimulus and a particular feature of a given haptic channel (e.g. level of force applied against the performer, vibration frequency or amplitude, etc.).
- ***Merging haptic feedback design parameters.*** Following optimization in isolation, the haptic feedback channels need to be combined and further revised. For example, the

user's cognitive load could suffer when several stimuli are presented at once even when each one of them is unimodally optimized in terms of performance. Another issue that could be affected is the feeling of comfort. Various signals that by themselves are pleasant could turn an interface unusable/annoying when combined. This criterion will prove useful to any type of interface that uses haptic feedback by providing representative cases that should be avoided.

### **3.4. Impact**

This section lists some of the fields that may be impacted by the proposed research.

#### **3.4.1. Human-computer interaction (HCI)**

The proposed research falls in the general context of HCI. A user-centered approach will be followed at all times during every stage of development, from designing the necessary upgrades to the chosen platform to the evaluation of the results. In particular, the area of expressive interfaces will be studied. The major impact on this field will come from providing a semantic base (a library of gesture primitives, together with a 'natural' mapping to their meanings) for a gestural language aiming at intuitiveness in representing both general and specific expressive commands. Such a semantic base will benefit not only the type of application envisioned in this research but other areas of HCI, especially those related to multimodal and immersive scenarios where a gesture language could be used to foster usability and inter-person communication.

#### **3.4.2. Haptics**

The proposed research presents an alternative approach to the way haptic feedback is currently designed in expressive interfaces, especially those designed for music applications. Our "*perception-based*" approach follows efforts initiated in our research group aiming to foster design centered in what "feels good", in an emotional sense as opposed to a purely functional perspective [Swin2007]. The major impact in this field will arise from providing a reduction of the user's cognitive load based on a better distribution of the haptic feedback, achieved by recognizing the natural function of each haptic channel in the real-life context and incorporating this into the interface design. The haptic feedback will also be used to reinforce the feeling of playing a musical instrument, enhancing the user experience from a qualitative point of view.

### **3.4.3. Psychophysics and Music**

Perceptual issues in music are studied both in psychology and music. Specifically, the perceptual mechanisms involved in the musician-instrument interaction are still under heavy scrutiny. The proposed research could shed light in some of those issues, as well as serve as a platform to conduct other experiments in this area. For instance, the role of haptic feedback in the perception of music parameters is still unknown, mostly because the deprivation of the haptic channel is only achievable by numbing the areas involved in the activity studied. Such induced numbness in the case of music will carry a noticeable decay in performance by itself not related to the perception of tactile stimuli.

As a result of our research we will have an interface where the haptic feedback could be turned on or off or varied according to some criteria. Further research could use this interface to achieve a better understanding of the role of the haptic channel in music performance.

### **3.5. *Disambiguation of Research Scope (Negative Space)***

This research is not aimed at reproducing the feeling or handling of any particular music instrument. Prior research [O'Mod2000] has demonstrated that haptic feedback intended to improve certain usability features of a modeled music instrument could end up not being an asset if the model aims for realism but does not adequately achieve this. Instead, we plan to grasp the common features that define computer interfaces perceptually as a music instrument.

## 4. EFFECTS OF HAPTIC AND VISUAL CUES ON ABILITY TO FOLLOW RHYTHMS

This section includes the experimental work done to test the effects of including haptic and visual cues in a percussive music interface while learning how to play a rhythmic pattern. The statistical analysis of the results will be presented as well. This experimental work presented here was part of the requirements for the course Physical User Interface Design and Evaluation offered by the Department of Computer Science (CPSC543) [Ped-etal.2006].

Traditional methods of learning music are arduous at best, relying on rote note memorization and a very limited set of individual's senses to guide them (namely, hearing), which not all students might possess. While there are many commercial electronic musical instruments on the market that boast the use of visual cues as a means to improve user performance, there has been no research we could find that gauges the effects of different types of stimuli on user performance.

The study presented here is the first step in an attempt to design a viable computer-based musical instrument with haptic feedback. Our immediate goal is to examine how various stimuli or cues could affect a person's ability to follow (and to a larger extent, learn) rhythms. In order to keep the scope of our goal feasible, we decided to examine only a very limited set of common stimuli, in this case, visual and haptic. To simplify things even further, we based our experiment on a simple drum machine and asked subjects to follow single drum beat percussion patterns, facilitating the user interaction by providing large pads and eliminating the extra level of complexity found while following melodies where note selection is also a factor.

### **4.1 Approach**

The approach followed was simple: to expose users to various stimuli while testing their ability to follow rhythms. We were measuring if and how the user's performance changes in the presence of visual, haptic and a combination of both stimuli delivered on top of a reference drum pattern when compared to their performance when only the reference sound pattern was given. The primary hypothesis was that there would be a significant improvement in user performance when any type of extra cues is given and, particularly, that they would perform better whenever a haptic cue is present.

The experiment setup was designed also to test how the presence of the audio signal generated during user's performance (audio feedback) might influence the results when it is presented to them together with the reference sound pattern. In this case, we hypothesized that subjects will perform better when they do not hear what they are playing.

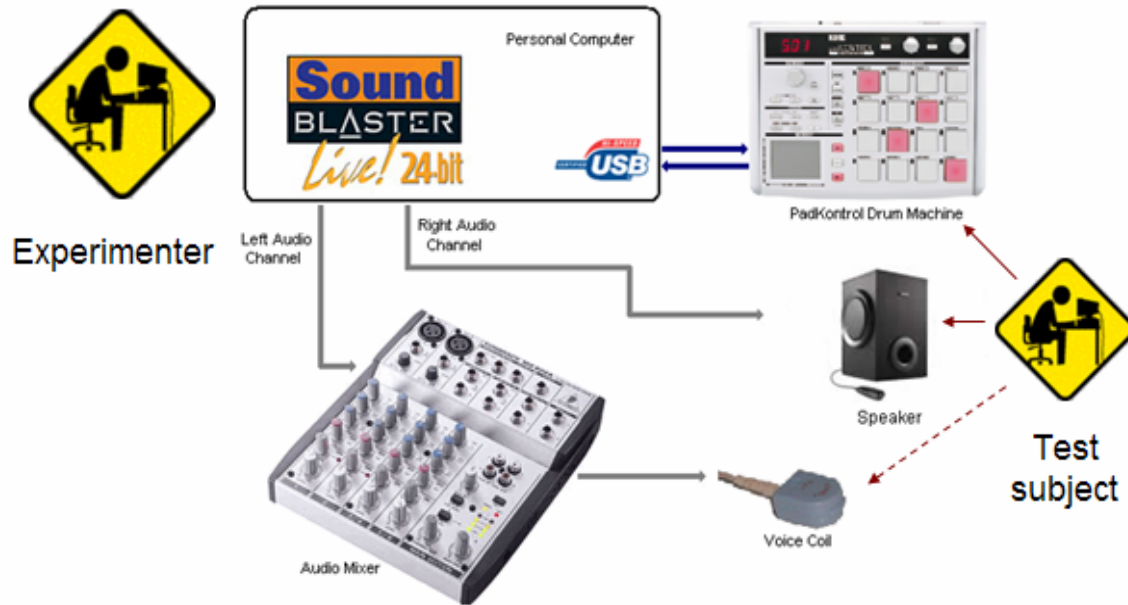
The subjects for the experiments were randomly taken from a population of people mostly without significant music education. By not significant, we mean that they have not received musical education, on a permanent basis or that the music education was received only in their early years. The tests were designed to be short enough to prevent muscle memory to take place, guaranteeing in this way that we are really measuring the process of learning rather than user's ability to play a learned sequence.

We also decided to eliminate the extra complexity of selecting a particular note from a piano-like keyboard and so chose to use the equivalent of an electronic drum machine where users could easily tap in any given pad always producing the same sound. This will limit the cognitive load on the users, helping them focus on the task given. Also, the fact that the pads on the drum only need to be tapped to react (as opposed to a piano keyboard where the key needs to be pressed and moved a certain distance before registering the action) will help in reducing muscular fatigue and simplify the task given to the users.

## **4.2 Prototype**

The prototype ran on a Centrino 1.4 GHz laptop with 512 MB RAM. Figure 2 shows a schematic of the hardware setup.

The prototype used in our experiment was based on a commercially available MIDI studio controller (Korg PadKontrol). This controller was connected to a laptop through USB and served both as a mean to acquire user's input and as a display for the visual cues because it can be programmed in a way that its pads blink synchronously with the reference sound. To generate the haptic stimuli, we used a voice coil directly plugged into the audio output of the computer. In order to compensate for the potential reduction in signal strength, we attached a mixer console to amplify the audio signal, and therefore increase the strength of the haptic signal delivered by the coils. As a side bonus, the inclusion of the mixer console allows for panning and muting of some of the audio signals, simplifying a lot the hardware requirements.



**Figure 2. Hardware components connection diagram**

The prototype was running under DeMuDi [3], a low latency Linux distribution. All the software components used are Open Source Software provided with the distribution of choice. To connect the different software components and route the Drum Machine signals we used Jack [Jack], a low latency audio server that provides a virtual patch bay, through which different audio applications and MIDI devices could be linked to exchange both audio and MIDI information. The connection diagram is shown in Figure 3.

The rhythmic patterns and the synchronized visual stimuli conveyed to the user were created using the software sequencer Rosegarden [Ros]. Rosegarden data was sent over MIDI channel 1 to the software synthesizer ZynAddSubFX [Zyn] which was in charge of transforming those patterns into sound, and through MIDI channel 10 to the external drum machine (PadKontrol) to be displayed as visual cues to the user. Rosegarden also was used to record a MIDI track of the user's performance sent by the drum machine over MIDI channel 10. The haptic stimulus was transmitted to the user via the previously mentioned voice coil activated by the reference sound. The computer's sound card was used for the audio component.

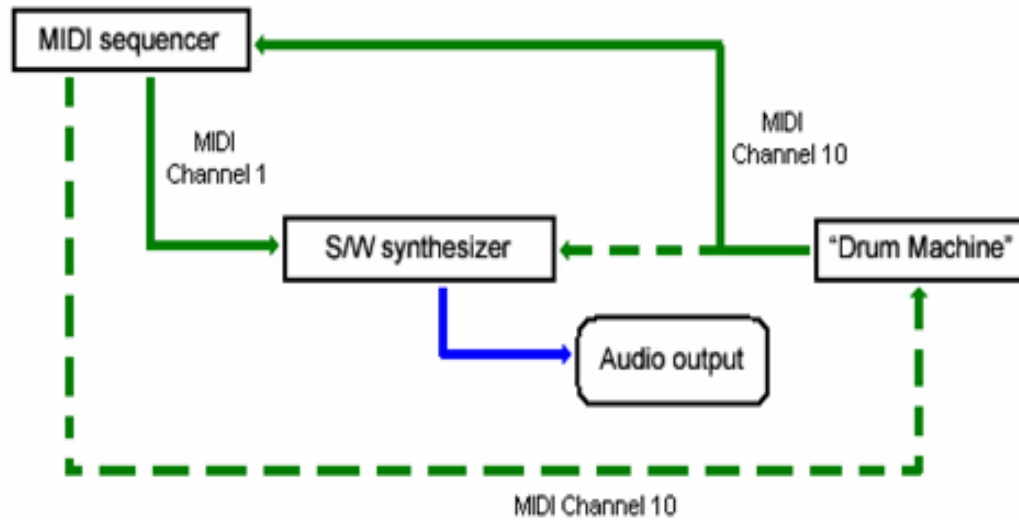


Figure 3. Software components connection diagram

### 4.3. Experiment Design

Our experiment exposed test subjects to four different conditions or combinations of stimuli while they attempted to follow a rhythmic sound. These combinations were: no extra stimuli (only the sound), visual, haptic, and visual + haptic. All rhythmic patterns differed by the number of beats and intervals between them. They were divided into three groups of increased level of difficulty: easy, moderate and hard. For each group we created 16 rhythmic patterns of similar level of complexity. The simplest rhythms were enclosed in one bar, while the moderate and hard ones were enclosed in two. As we were interested only in analyzing how useful an extra stimulus could be in the process of learning, we presented the test subjects to a maximum of 4 repetitions per rhythmic pattern, so as to restrict any processes such as muscle memory from influencing our results.

We tested 16 test subjects divided in two groups of eight. One of the groups received audio feedback from their own performance on top of the audio reference and the other one was given only the audio reference.

Each of the test subjects in each group performed the same set of patterns across complexity level, as shown in Table 1 for two subjects in a same group. Under each of four combinations of



stimuli, test subjects listened to and tried to follow 12 rhythms of various difficulties that steadily increased with time. First, they were exposed to four easy rhythms, then four medium and lastly four hard ones. Rhythms within each difficulty level provided to the subjects during the experiment were random but never repeated for the same subject throughout the experiment. For instance, as we see from the table, the Easy Set #1 was given to user 1 under the Only Audio condition and to User 16 under the +Visual condition. We employed this randomization to reduce the error of assuming that we correctly matched the complexity of each set, making some conditions *a priori* more difficult than others.

**Table 1. Sample test design**

Test subjects	Rhythm Complexity Level	Conditions			
		Only audio	+ Visual	+ Haptic	+ Visual and Haptic
Test subject 1	Easy	Easy Set 1	Easy Set 2	Easy Set 3	Easy Set 4
	Medium	Med. Set 1	Med. Set 2	Med. Set 3	Med. Set 4
	Hard	Hard Set 1	Hard Set 2	Hard Set 3	Hard Set 4
.					
.					
.					
Test subject 16	Easy	Easy Set 3	Easy Set 1	Easy Set 4	Easy Set 2
	Medium	Med. Set 3	Med. Set 1	Med. Set 4	Med. Set 2
	Hard	Hard Set 3	Hard Set 1	Hard Set 4	Hard Set 2

In total, each test subject tried to reproduce 48 rhythmic patterns and each rhythm set was assigned *a priori* to each condition. This way we made eight combinations that were given randomly to each of the eight test subjects in each group.

#### **4.4. Experiment Results**

The recorded performances from the 16 users were processed by applying a quantization to the same time grid we used to create the reference patterns. This allowed us to eliminate small nuances of time differences that could be considered as belonging to the same rhythmic

sequence. Then we counted how many notes (or pad hits in our case) they misplaced when contrasted with the reference pattern. We calculated the percentage of error for each user under each condition for each of the patterns in the four sets.

The qualitative analysis from the questionnaire given to the subjects showed that the trials were somewhat physically exhausting, that the users felt an increased level of complexity from the easy sets to the hard sets and that they on average felt a slight improvement in performance when an external stimuli was given on top of the reference sound. The visual cues were the preferred ones or the ones with which they felt they performed better.

Users approached the experiment in different ways. Most of them held the voice coil between the thumb and the index finger while others grabbed it, completely pressing it against the palm of the non-tapping hand. Some of them waited until they were sure they comprehend the rhythmic pattern before tapping on the drum machine and others just started tapping from the first or the second bar and try to adapt their performance to the reference as they went.

One user reported that she kept looking at the haptic device while it was buzzing because it helped her concentrate on this type of cue, even when she recognized that there was nothing there to see. We also noticed that several subjects closed their eyes even when the visual cue was given, especially when they were trying to follow the hard rhythms.

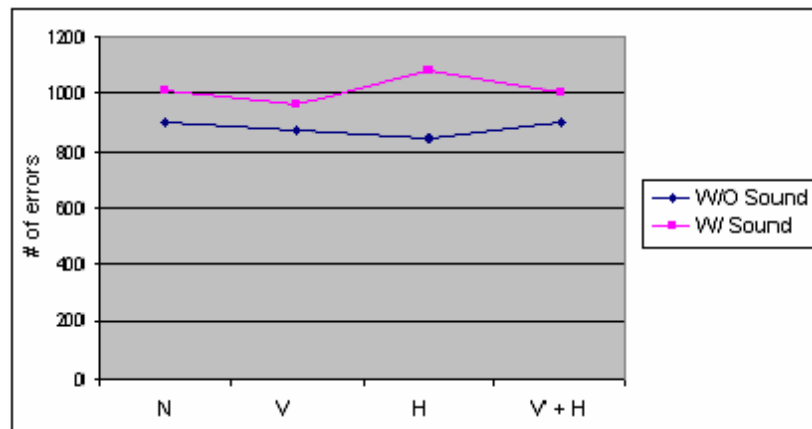
#### **4.5. Discussion**

Since our experiment design matches the case of a 2x4x3 Split Plot with repeated measures in each cell, the data was processed according to this methodology and the results are presented in Table 2.

From the data in Table 2, we can see that we could only prove our second hypothesis: the presence of sound from the user's own performance negatively affects their ability to follow the rhythms ( $F_{(1,168)}=6.4923$ ,  $p < .025$ ). Figure 4 clearly shows that the number of errors was lower when users did not hear the sound they were producing together with the reference, being the most notable difference achieved when the haptic stimuli were present.

**Table 2. Statistics results. ANOVA analysis on a 2x4x3 Split Plot design**

	SS	Df	MS	F	sig
A (Presence or not of sound from user performance)	1377.68	1	1377.68	6.4923	<.025
B (Complexity Level)	27464.9	2	13732.4	64.714	<.01
C (External Factors)	124.54	3	41.5143	0.1956	-
AB	404.41	2	202.20	0.9529	-
AC	299.46	3	99.820	0.4704	-
BC	593.19	6	98.866	0.4659	-
ABC	182.39	6	30.398	0.1432	-
Within cell (Experimental Error)	35649.59	168	212.19		
Total	66096.19				



**Figure 4. Difference in performance between groups who received audio feedback of their tapping vs. those who did not**

From that same set of curves in the graph, it is obvious that the subjects performed slightly better when haptic stimuli were provided and they did not have sound feedback from their performance. Their performance was the worst in case when they did have that reference. However, the statistical analysis also clearly indicates that there is no significant difference between performances under each of the external cues when compared to the case where they were not present.

Figure 5 shows a significant difference in performance among the three difficulty levels. ( $F_{(2,168)}=64.714$ ,  $p < .010$ ). This graph shows that our hard sets were harder than we initially

thought, and in addition, the presence of sound feedback only significantly increased the error rate in the easy, but not the medium or hard subsets.

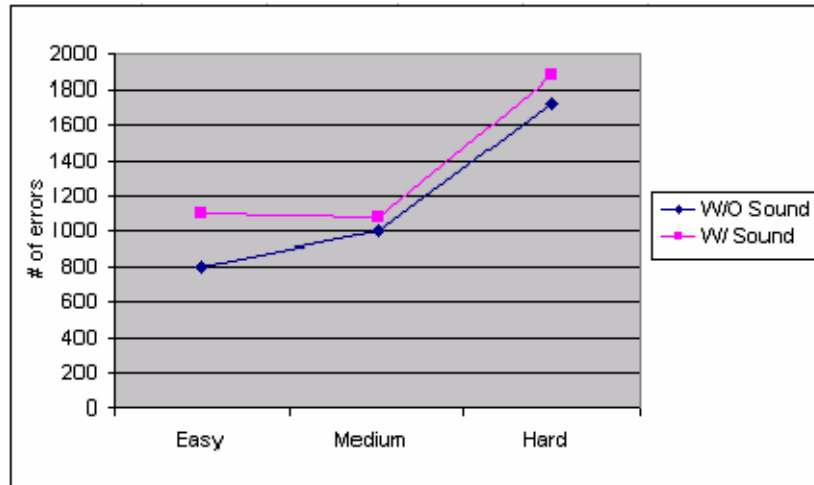


Figure 5. Differences on observed difficulty level for all subjects

Figure 6 shows how subjects performed under each condition on the three difficulty levels. Even though these results were not statistically significant, it is interesting to note how, with the harder rhythms, subjects performed slightly better when no extra stimulus was given together with the reference sound. This reinforced our suspicion that those rhythms had a greater level of difficulty than we expected, too high for the participants of our experiment, and that any other stimuli constituted an extra cognitive load interfering with their performance.

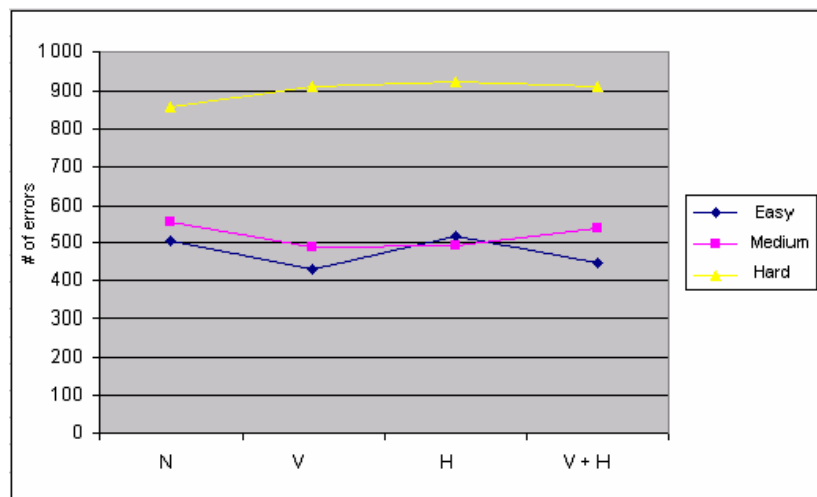


Figure 6. Errors observed between users for different difficulty levels and conditions

While analyzing the data we uncovered some details we never even considered when designing the experiment. The most significant one was related to the cultural background of our subjects. The vast majority of our subjects were voluntary university students from Asia (mostly Chinese or Japanese) and most of them have been in Canada for less than five years. The performances of these subjects were worse than those of the subjects with a western background (North or Latin America). The only hypothesis we could draw from this is that there is a very likely possibility of cultural bias in the selection of our rhythms. As will be detailed in the subsequent section, this fact prompted us to refine our selection of rhythms for future experiments and include a broader variety of them. It seems apparent that the cultural environment where the subjects were raised, contributing to factors such as the types of rhythms that they were more commonly exposed to, can influence their performance in this type of tasks.

We also realized that several factors influenced the complexity of the patterns that we had not considered. For instance, faster rhythms (patterns where notes were closer to each other) were more difficult to some subjects than to others. In these cases, subjects seemed to comprehend the rhythm and indicated that it was not difficult but they just could not tap with such speed. The exact opposite of this, long periods of silence between notes, was also found to be highly difficult for the subjects, in contrast to our predicted case where the only main contribution to difficulty would be different and uneven time intervals between notes in the same pattern.

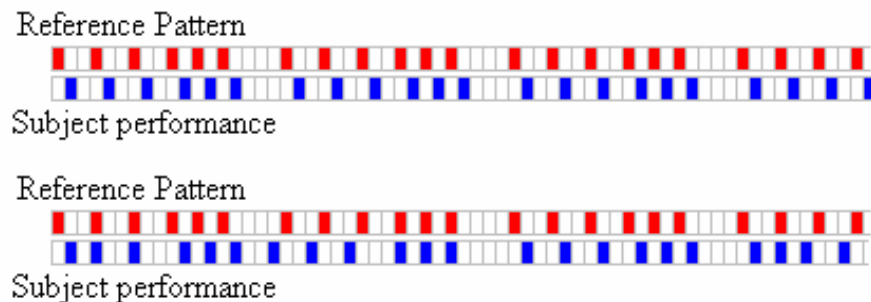
The last factor we noticed that has added to the complexity of the hardest rhythms was the lack of a marked or clearly indicated ending of the pattern. While the two simpler cases had a marked ending of the pattern defined with a relatively long (compared to the average note separation in the rhythm) period of silence, in the hard ones this period of silence was too similar to some of the time intervals used in the pattern, which made it harder for the subjects to realize when/where each rhythm ended.

We also found that the approach of delivering the haptic stimuli we took was not optimal. Firstly, there was no uniformity in the placement of the haptic device, as subjects held it in different ways as explained above. We derived the signal sent to the voice coil from the sound given as a reference in order to foster synchronization. The haptic cue conveyed this way, despite being a well-defined burst, lacked a clear, sharp starting edge. We suspect this could be one of the reasons why we could not prove our hypothesis regarding an increase in performance with the addition of haptic cues.

Also, we realized that we were dealing with an extremely qualitative point of view or

component that is hard to judge by cold quantitative analysis. For instance, when counting misplaced notes we could end up with a 100% error rate for a subject even if he/she managed to follow the rhythm exactly but delayed with respect to the original one, as shown in the upper two patterns in Figure 7.

The bottom two patterns of Figure 7 show the problem as we found it when analyzing the raw results. In this case, we can derive completely different conclusions depending on how we execute a quantitative analysis. If we focus merely on misplaced notes, the percentage of errors is 16/22. If we analyze the time difference between the notes it would be 8/21. In fact, if we listen to the pattern performed by the subject we could evaluate the performance in a totally different manner, perhaps giving it a 50% correctness rating.



**Figure 7. Example of variation between reference and subject performance for a single rhythm pattern**

We eventually decided to employ error rating analysis via the ‘notes misplaced’ approach because overall it seemed to give the most drastic results, and if we managed to prove our hypothesis in this stronger case we would find no difficulty in proving it on the weaker ones. Nevertheless, as we have shown, there are other more subjective issues than those that could come up from a pure quantitative analysis, and finding ways to analyze the data in short time and without a bias of an expert evaluation is another challenge.

#### **4.6. Conclusions from the experiment**

Even though we were unable to prove our primary hypothesis that the presence of external stimuli, and especially haptic stimuli, improves user performance, we found the results we obtained encouraging in that, at the very least, subject performance was not significantly worse either. Furthermore, we were able to prove that there was a significant difference in user performance between the presence and absence of audio feedback from their own tapping, proving our second hypothesis. The data and results we gathered from this experiment are

essential for future work in this area, especially those concerned with possible improvements to our experiment's design.

The issue of the cultural background, which we had never considered, seems now the most relevant factor when planning future studies, taking into account the extremely subjective nature of these types of studies and how previous exposure to music influences the subject performance. In order to eliminate this factor, we will take the approach of expanding and altering our rhythm selection to include a broader range of non-cultural-specific rhythms, as opposed to coming up with different sets for each 'culture' due to the obvious inherent problems with such solution.

As mentioned above, the pattern complexity will also need to be taken into account. In particular, we need to factor in the effects of the change in rhythm speed on the subjects' performance, since there were noticeable differences in performance between, for example, the first 'fast' medium rhythm and the second 'fast' medium rhythm a subject is exposed to. A possible way is to group all rhythms of a similar speed in each difficulty level together, so that the subject will go through all the 'slow' rhythms of that difficulty level before the 'fast' ones, hopefully eliminating the effects of adapting to speed change on subject performance.

Furthermore, a better (sharper) haptic stimuli presentation and specified co-location of the haptic device's placement relative to the hand or finger directly involved in the musical activity might improve subject performance. It will also help eliminate the variability incurred when subjects were allowed to place the haptic device anywhere they wanted. Also, we are interested in seeing whether different types of haptic stimuli (i.e different waveforms) will have any effect on subject performance. Regarding the delivery of cues, it will also be of interest to examine how (if at all) the displacement of the extra stimuli in time relative to the reference sound (such as delivering the haptic cue in advance, for example) affects subject performance, and whether the results we obtained with this experiment will hold with future improved setups.

Since previous work in this area showed that what the skill level is an important factor to consider in evaluating music interfaces, another viable and necessary avenue for a future study will be for us to examine the effectiveness of extra stimuli (specifically haptic cues) on expert users to contrast our pool of subjects in this experiment.

## 5. MANAGEMENT PLAN

This section includes the approach we are proposing to guide the research, and the road map devised to obtain the expected results. A detailed description of the proposed research stages will be provided framed in time.

### 5.1 Approach

To achieve the thesis goals as stated in Section 3.3.1, the proposed research should address three major points: developing a platform, development of a semantic base for a gesture interface and finally, the design, implementation and merging of the two channels of haptic feedback. In this section we will lay out the basic approach for each of these tasks.

#### 5.1.1. Platform

Based on our case-study approach, we require a specific platform to act as the music instrument interface. Throughout this proposal, we have developed the premise that gesture interfaces (whether grounded or not) are one of the most intuitive and promising interfaces for expressive applications and therefore our approach will be built upon the idea of expressive gesture.

However, we further believe that in the name of comfort and control, grounded force feedback is required to support these gestures; this grounding will certainly affect the type of natural gestures that appear in our semantic base. Gesture interfaces are notoriously challenging in terms of incorporating force feedback, in that the grounding device generally imposes a restriction on workspace. Our case study will be that of a grounded gesture interface capable of mapping in 3D the movement of a performer's hand at the same time that it provides both force and vibrotactile feedback. This should be enough to control a set of expressive parameters in an arbitrarily selected music performance, and will provide the means of evaluating the main design criteria of comfort and controllability.

As our interest is concentrated on designing and evaluating the effects of haptic feedback in relation to computer music applications and not on the design of the gestural / haptic interface itself, our approach will take as a base a commercially available device that provides both a 3D mapping mechanism and haptic feedback through the force channel. Such device (a PHANTOM Premium 1.5 from SensAble Technologies) is already available through ICICS and is considered



an industry standard for deploying high-quality force feedback. This device has a workspace of 381x267x191 mm (WxHxD) throughout which a continuous force of 1.4N (8.5N max) can be exerted in 3 directions (degrees of freedom or DF). We will design and build an add-on to the original interface to include the desired vibrotactile feedback that the original device is unable to provide.

The selection of this interface will also allow us to test this proposal's primary hypothesis (the separable design of haptic feedback channels based on their perceptual roles in a given task context) in a best-case scenario, in that the platform, while not able to reproduce real-world fidelity, does deliver relatively high-quality and high-DF haptic feedback. This approach is preferred to an approach of using commodity-level hardware, because we will be able to (a) eliminate feedback combinations that do not work by focusing on perceptual issues rather than technological ones; and (b) demonstrate that those combinations that work can be deployed using available technology. Furthermore, (c) satisfactory results obtained through this research could be narrowed down to the optimum point of performance/cost through a degradation study (not included in this proposal) where the top-of-the-line characteristics of this interface are deliberately reduced to determine the values necessary to produce required performance ratings, as in [Bro-etal.2004 ].

### **5.1.2. Development of a semantic base for a gesture interface**

The *semantic base* that we aim to gather at this stage will consist of meaningful gestures that could suit our purpose of controlling a music interface. The movements' *syntax* (the structure by which a performer combines each of the control actions, or gestures, to achieve certain aesthetic results) will be left to the performer working within the interface constraints. As what we are building is an *expressive* interface, we are most interested in the meaning of a gesture and how it should be mapped to control a certain music parameter.

The semantic base will be constrained by the PHANTOM's kinematics and workspace. As this stage does not include force feedback, the gestures gathered here will determine only how the music control parameter of interest should be mapped in the interface to match the gesture meaning. It is known that an ungrounded gesture can not be repeated in the same way when some impedance is added by using a grounded interface. By focusing on the meaning of the free gesture, we anticipate that this semantic base will be applicable to a broad spectrum of actual gesture interfaces, with customization happening in the ensuing haptic feedback design step and

with an eventual syntax depending on the particular interface's constraints and resources. Also, by providing a semantic base built upon gestures taken both from music and from everyday life, we expect to provide the necessary freedom that any expressive interface should have.

A deeper explanation of the procedures to achieve this point will be provided in Section 4.2.2.

### 5.1.3. Sound Generation

To generate the sounds we will rely on the MIDI protocol; once the semantic base has been defined, it will be easy to map the sensed PHANTOM endpoint position in space to the appropriate MIDI parameter, and this will place this phase at the end of Stage 1 or could be considered as part of this Stage. Due to its simplicity, this design step is not included in Figures 1 or 2.

### 5.1.4. Design and deployment of the haptic feedback

The bulk of the research is concentrated in this stage. Our hypothesis asserts that the design of the haptic feedback should be governed by a perceptual framework, separating each haptic channel according to the use and meaning it possesses in the real-life context. Research will focus on assigning an appropriate haptic response (tactile or force environment) to any useful perceptual (auditory) music parameters. Only when these responses and environments are tuned and tested will they be matched to the gestures defined with the semantic base as the ones appropriate to control those parameters. Thus, the haptic environment will tend to be associated with the *identity and value* of the parameter, and the gesture with the *action* to be performed upon it.

For instance, if increasing the pitch is better controlled by presenting a force that opposes the user's movements (providing a resistance that enables more precise control over pitch), then first the designer would tune the force feedback parameters to a point that they feel comfortable. Next, he/she would match this feedback with the gesture for controlling the pitch, and further adjust it until this feels right. This incorporation of haptic feedback might also induce some change in the (ungrounded) gesture originally defined during the semantic base stage, but this adjustment is executed in the mixing stages described below. The haptic feedback design deals only with the best way to *perceive* a particular music parameter through a haptic channel, while carrying out the semblance of a particular gesture.

We base this approach on the possibility of separating the force feedback from the tactile

feedback channels to address different interface features: the former to foster controllability and the latter to foster the feeling of using a musical instrument.

A set of user studies will help the final fine-tuning of the haptic feedback parameters (still in isolation from one another, but in conjunction with the gestures associated with those parameters / actions) and will serve to evaluate ensuing performance in parameter control, as well as to identify relevant perceptual limits. With these user studies we plan to gather a list of some representative restrictions/suggestions pertaining to an appropriate division of control/perception functionality between the two haptic channels that will serve to generalize our research to other interfaces in the same general range as ours. A detailed explanation of each research stage included in this point is provided in Section 4.2.3 and 4.2.4.

### **5.1.5: Merging of tactile/force feedback**

There are two mixing stages in the proposed research, and each one fulfills one particular objective. The particulars of each will be explained in Sections 4.2.5 and 4.2.6.

The first mixing stage has the objective of enhancing the controllability of a gesture controller by providing some impedance (force feedback) to the gestures controlling the music parameters. This stage accepts two inputs: a general set of gestures designed to control the music parameters defining the music/expression space, and a set of forces perceptually fine-tuned to achieve the best apprehension (identity and setting) of the same set of music parameters. In this first mixing stage, the main objective is to create a force space that matches the music/expression space in a way that each music parameter could be controlled effectively. As the gestures included in the semantic base are not grounded, we anticipate some changes but not enough to require us to iterate on the basic gestures themselves.

The second mixing stage is aimed at providing the feeling of playing a music instrument in an interface already tuned for control. This is the stage that presents a higher risk, since some perceptual overloading might occur. However, there may be ways to reduce overloading by using masking and or neglecting a haptic cue that provides less information. The goal in this stage is not to optimize the design for everything that could be controlled, but to achieve the best ratio of controllability and perception of a music instrument. As some music instruments are more expressive/versatile than others, the same situation occurs here. We are aiming to obtain the best performance features with the resources at hand.

## 5.2. Road map

The proposed research is structured in almost a linear way. Figure 1 (Section 1.4) represents the main objectives for each stage and how they relate to each other. Figure 8, below, places these stages as well as a Stage 0 (incorporating the vibrotactile feedback into the original Phantom) in a time frame. Each of the Stages (except Stage 0) contains at least one phase involving user studies. Each will be discussed in detail in separate subsections.

Given these clarifications the different stages of the research could be explained as follows.

### 5.2.1. Stage 0: Phantom enhancement with vibrotactile feedback

The interface will be built taking a Phantom as a base. This device will capture the gestures from the performer and is able to provide force feedback in three dimensions. We will augment the PHANTOM with a handle equipped with vibrotactile actuators. The design and implementation of this handle will take into account both ergonomic issues and the mechanical requirements that will allow it to be attached to the existent Phantom handle without damaging it. The author possesses enough mechatronic experience to guarantee that this process is done properly. This stage is a pre-requisite only for Stage 3 (design of vibrotactile feedback) and could be executed at any time in parallel with Stages 1, 2 and 4.

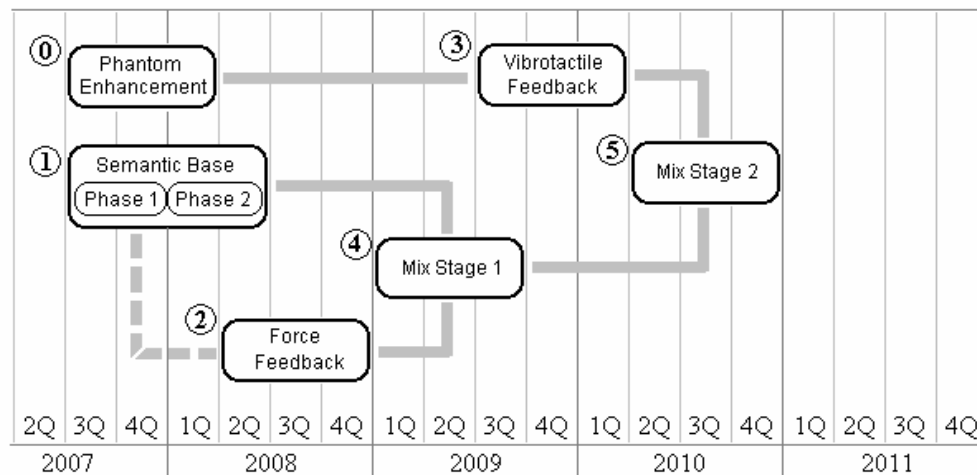


Figure 8. Research Stages and Research Time Frame

### 5.2.2. Stage 1: Building a Semantic Base

Stage 1 will consist of a set of two experiments used to build the semantic base for the gesture language. The procedure to create this language will be based on the one proposed by Nielsen et

al. in [Niel-et al.2004]. This method was designed to guarantee that the gesture vocabulary obtained contains gestures that are easy to perform and remember, and are intuitive, metaphorically and iconically logical, and ergonomic. This stage can be summarized as follows:

An initial user study (Phase 1 inside Stage 1 in Figure 2) will be devoted to building a general database of gestures taken from a balanced mix of experimental users selected with and without a music background (e.g. music students or music professionals). The general procedure will be to present a set of sound clips to the users and ask them to simulate a performance of that sound using the Phantom without force feedback implemented. The movements and the explanation of their choice will be videotaped.

This phase will be followed by an evaluation to extract the optimum gestures from the recorded material and creating in this way a primary semantic base. Nielsen et al. in their methodology insist that this stage should not be limited only to the recordings but should serve as an inspiration from where other gestures could be formed. Some redundancy in the semantic base is preferred in order to allow a fine tuning in the following phase.

In Phase 2 we will test the primary semantic base. Again, following the methodology of Nielsen et al, a balanced mix of experimental users selected with and without a music background will evaluate sets of gestures selected from the ones collected in the previous phase. The experiments will consist of a video presentation featuring the gestures simulating a control over the relevant music parameter. User will later be evaluated for memorability and cognitive overhead. The intuitiveness of each gesture will be tested by evaluating the user ability to identify the represented function.

At the end of this phase a relationship of a defined gesture to a defined music perceptual parameter will be achieved.

### **5.2.3. Stage 2: Design of the force feedback**

We posited in Section 2.3. that by designing appropriate force feedback, the controllability of a gesture interface could be improved. In this stage, we start by doing this design for each of the perceptual parameters chosen to be controlled. Several force renderings will be deployed and matched to different perceptual parameters to determine the best coupling. For example, pitch being a parameter that should be controlled: how should we deploy a force to improve the sense of control over pitch? Should the force increase as the pitch increases? In a vibrato, how should the forces changes should be mapped?

At this point, it is still left to test whether these acoustic parameter/force environment combinations (and/or the net effect of all of them together) will improve the controllability of the interface or reduce the time needed to master a musical piece. Though force feedback design could be performed without taking in consideration the gestures chosen for the interface, it seems productive to wait until at least the first Phase of Stage 1 has completed to have a more solid ground and focus the force feedback design according with the already obtained users' preferences in terms of general meaningful gestures. As this link is not a must, it is illustrated in Figure 2 with a dashed line between Stage 1 and Stage 2.

Thus, the end of Stage 2 consists of a set of user studies designed to fine tune the coupling between force response and perceptual music parameters. Cognitive load and comfort are the most important factors to measure in those studies in order to set the proper force feedback parameters

#### **5.2.4. Stage 3: Design of vibrotactile feedback**

Music instruments are basically resonators (or arrays of resonators) whose output is amplified by the instrument's structure. The literature review presented in Section 2.1 shows how those vibrations could be felt by performers. It is not unreasonable to surmise then that these sound-derived vibrations are the ones responsible for defining the sensation of playing a musical instrument. If this is so we could feedback through the vibrotactile channel a signal derived from the sound produced to simulate the vibration levels achieved in music instruments.

Independently, the literature presented in Section 2.1 suggests that time and rhythm cues could be provided by vibrotactile stimuli. The user study already conducted and described in section 4 suggested that vibrations could be utilized as rhythmic cues. These results together provide some support for including some cues on rhythm, tempo and performance timing in the vibrotactile feedback to enhance the interface controllability.

In this way, the vibrotactile feedback may also play an indirect role in controllability, although we regard the force feedback as being primarily responsible for how a parameter is controlled whereas the tactile feedback is associated with establishing the general feel of playing a music instrument. However, there is an inherent link between the force and the vibrotactile feedback established through the sound generated. Since the force feedback is used to enforce a feeling of control over the sound and we propose to synthesize the vibrotactile feedback from the sound produced, it is expected that the vibrotactile feedback should reinforce the achievement of

a desired sound, and thus it also contributes to the instrument's controllability.

A set of user studies will be designed to fine-tune the experiment on rhythm already conducted and to test for the viability of including sound cues on top of the basic vibratory response expected from a musical instrument. The rhythmic cues in our implementation should come from the same produced sound, e.g. by enhancing certain spectral content to make it more salient to the sense of touch. As in Stage 2, cognitive load and comfort are the most important factors to measure and tune.

#### **5.2.5. Stage 4: First Mixing Stage**

This stage is devoted to tuning the instrument's controllability. The various force feedback response designed on Stage 2 for each music parameter will be implemented together on the Phantom and combined to create a force space that works well with the gestures designed to control each parameter. No vibrotactile feedback will be included in this stage. For instance: let us assume that to increase pitch the performer should move the PHANTOM end effector from left to right; and to increase amplitude should move it upwards. Let us also assume that the preferred force rendering to increase the controllability of pitch and amplitude is an increasing force proportional to the parameter magnitude. In this scenario the resulting force space for pitch and amplitude will be such that the performer will feel an increasing force when raising his/her hand *and* when moving it from left to right.

A set of user studies or pilots will be designed to determine the appropriate tuning of the gestures included in the semantic base to transition them to a grounded environment; and additionally whether, in this integrated (in terms of multiple force feedback 'layers' for the various parameters) environment, the user is being overloaded. In this event, we will determine the added value of forces for each acoustic parameter, and potentially reduce the set to a cognitively manageable (i.e. comfortable) size. Some fine tuning in the force feedback response is also expected. Another set of experiments involving users with music background will be run to measure the accuracy while following a simple music score. Subjective measures of comfort will also be taken. The output of this stage will define those perceptual parameters that are suitable and usable to map through the haptic force channel to achieve an equilibrium between controllability and comfort.

### **5.2.6. Stage 5: Second Mixing Stage**

This is the last stage of the proposed research and will test the effects of including vibrotactile feedback in the instrument which has already been tuned in terms of force feedback for controllability.

The vibrotactile feedback designed without any force component in Stage 3 will be incorporated into the interface. A set of experiments will be designed to determine the level of comfort and cognitive load of the final interface. It is very likely that some iteration will be needed to achieve the proper results. At this stage we are aiming only to tweak the tactile feedback parameters in order to enhance the feeling of playing a music instrument, and comfort is the ultimate goal.

Comfort is a highly qualitative parameter and its measurement could be hard to quantify unless some physiological measures are performed. However, as we are going to be evaluating an expressive interface we propose to leave this evaluation in a qualitative level. Our final set of experiments should concentrate in evaluating the use of the interface while performing some demanding musical piece and obtaining user feedback about using the interface under different configurations where elements of the haptic feedback could be turned on or off.

We expect that the force feedback should remain as it was tuned in Stage 4. However this might not be the case. This constitutes one of the major risks points in the proposed research. In Section 4.3 we will address the implications of reaching these points.

## **5.3. Risk Points**

The major risk points we foresee in the proposed research arise as consequences of separating the design of force and vibrotactile feedback.

As we described in Section 4.1.5, our approach does not attempt to maximize controllability or the feeling of playing a music instrument as separate issues, but to reach the maximum equilibrium where both aspects complement each other given a particular interface. The methodology proposed is open in the sense that in each stage we could aim for maximizing one feature, be it tuning control gestures or perceptually optimizing a feedback signal. The mixing stages are ordered in a way that for a given time frame in the research path, a combination of two or more features could be tuned. This means that not all of the individually ‘optimized’ mappings (some being redundant) could be materialized at once, but at each one of those stages we should



have enough combinations to choose from so as to achieve an acceptable integrated result. To illustrate this more specifically, we note that at the end of the independent design stages (2 and 3) and the first mix stage (4), we hope to have the following elements to work with:

- 1) At least one (but preferably multiple) effective force renderings for each music parameter (which do not necessarily work together when integrated).
- 2) One combination of force renderings (a force space linked to a gesture space) which do achieve good control when integrated (i.e. a positive result from the first mixing stage 4). These would most likely be various combinations of those renderings that work (ie. not all combinations would include a force rendering for a certain music parameter).
- 3) Tactile feedback with and without enhanced rhythmic cues.

In the final mix stage, we can then play with all these variables; if the work is done carefully there should be enough combinations to choose from to find an integrated result that works, and thus redundancy in earlier stages is key.

Our main objective is to try to achieve the point of ‘feedback saturation’, where for a certain interface (defined by specific capabilities and constraints) any further addition of feedback cues will start degrading the user’s overall performance while using the system. This point will be different for different hardware platforms. From this point of view, even if the last mixing stage fails to achieve a usable combination of force and tactile feedback we could present it as a valid result for this type of interfaces with this set of resources; based on our hypothesis, this combination might be successful in a more capable hardware platform. However, we should have enough combinations to tweak in order to achieve some satisfactory results.

The time frame presented in Figure 2 has been constructed in a way to guarantee a certain degree of flexibility in time to accommodate for unplanned iterations in each stage.

#### **5.4. Resources**

The following is a list of the expected resources for the proposed research. It should be noted that references to resources acquired through ICICS are not directly reflected in the cost estimates. Though these items are not free of charge they are available as part of a general fund provided by the supervisor.

- ***Platform: \$300 - \$800***

The proposed research is centered around a PHANTOM V1.5 from SensAble

Technologies. This device was acquired through ICICS and is shared among researchers in ICICS Departments. Although we anticipate a need to time-share this device with other users, it shouldn't be of a magnitude that could put in danger the development of the proposed research. The costs included here will be for adding vibrotactile feedback to the PHANTOM interface. For this task we will require access to the AMPEL/ICICS Student Prototyping Workshop where the mechanical manufacturing will take place.

- ***Platform Software: 0\$***

The main development platform for the software will be Linux. The PHANTOM is fully supported in this platform and the rest of the software tools that will be used will be open-standards.

- ***Computers/hardware upgrades: \$1000 - \$1500***

Our previous studies requiring real-time sound generation showed the necessity of upgrading the hardware of the computer that should be used for this research (lassen). At least a high quality sound card with low latency drivers compatible with the operating system of choice to develop this project (Linux) and enough audio output to satisfy the requirements of this research should be acquired.

- ***Supplemental Developers: \$3000 - \$6000***

It is likely that one or two supplemental developers (skilled mechatronics or CS undergraduates or recent graduates) may be hired to offset certain development responsibilities. Desired skills would be background in low-level hardware/software integration and device testing.

- ***User studies: \$1500 - \$2400***

The majority of costs for the user studies will be for participant compensation. To fulfill the objectives of the proposed research a minimum of 6 user studies averaging 20 users each is necessary. The common user remuneration is set to 10\$/hour and we don't expect to go over that time limit in any of the studies planned. The high-end estimate takes into consideration that 2 other extra user studies could be needed as part of unexpected iterations in the mixing stages. Several pilot studies should incur no charges as they are expected to run on volunteers for no compensation. In the event that compensation will be required, \$500 is included in the high-end estimate. Digital video equipment for recording the user studies is available from ICICS. However, recording

media and related miscellanea estimated at \$300 will need to be purchased.

### **5.5. Future Work**

Though the proposed research presents a project intended to be sufficient in scope for a PhD thesis, there are several elements that could be further studied or extended to round up the results obtained with this one. A discussion of these will follow.

#### **5.5.1. Haptic feedback and different end applications of the interface**

A music interface based on computers has the possibility of being flexible. Perceptual mechanisms involved in different activities could be different, and this means that different stimuli may be necessary to match different applications. The approach presented in this proposal is devoted to use the interface to perform a musical piece and the haptic feedback will be tuned for this scenario. Other music related activities might need a different tuning: for instance the cases of learning how to use the interface or learning how to perform a musical piece. A comparative study of the interface in each of those cases would help to bring some clarity on the nuances inside the interaction Musician-Instrument.

#### **5.5.2. Degradation of quality**

The proposed research is centered in a best case scenario. The device used as the main component for the hardware platform is known to be one of the most precise haptic interfaces available. It was chosen like this to avoid the uncertainties coming from untested hardware. However a result obtained in this platform needs to be brought into a more realistic budget for an interface of this type. A degradation of quality study will serve to determine up to what level the infrastructure could be relaxed (in other words, how much the budget could be cut) in order to achieve relevant results.

#### **5.5.3. Expanding the scope of the proposed methodology**

Though primarily designed for the case of computer music applications, this methodology (if proven) could be extended to other applications. Fine tuning feedback taking into consideration the perceptual parameters involved and specifically designing the feedback taking into account what is the best channel to represent the information could lead to less cognitive overhead. Several areas like supervisory control and multimodal interactions could take advantage of this

---

approach. Other haptic modalities such as tangible interfaces (mostly linked to the shape of the manipulators) could have a higher level of incidence in other applications and might be of greater importance than force or vibrotactile feedback. A perceptual evaluation or classification of several interfaces of relevance in everyday or skilled applications will be an interesting expansion of this works that would have an important incidence in clarifying the way haptic feedback should be incorporated into new designs.

## REFERENCES

- [AskJan1992] Askenfelt, A., Jansson, E.V. “*On vibration sensation and finger touch in stringed instrument playing.*” *Music Perception*. Spring 1992, Vol. 9, No. 3. pp. 311-350.
- [Bon1994] Bongers, B. “*The use of active tactile and force feedback in timbre controlling electronic instruments.*” In *Proceedings of the 1994 International Computer Music Conference*, pp. 171-174, 1994.
- [Bon2000] Bongers, B. “*Physical interfaces in the electronic arts. Interaction, theory and interfacing techniques for Real-Time performance.*” In M. Wanderley and M. Battier eds. *Trends in gestural control of music*. Paris: IRCAM-Centre Pompidou, 2000. pp. 41-70.
- [Bro-etal.2004 ] Brouwer, I., MacLean, K. E., Hodgson, A. J., “*Simulating Cheap Hardware: A platform for evaluating cost-performance trade-offs in haptic hardware design.*” In *Proceedings of IEEE Robotics and Automation (ICRA 04)*, New Orleans, LA, April 22-28 2004. pp. 770-775
- [BroWil1998] Bromwich, M.A., Wilson, J. “*Bodycoder: A sensor suit and vocal performance mechanism for Real-Time performance.*” In *Proceedings of the 1998 International Computer Music Conference*, pp. 292-295.
- [Buch] Buchla Lightning II. Infrared midi controller. Description and specifications available at: <http://www.buchla.com/lightning/index.html>. Last accessed: June, 2007.
- [Cad1994] Cadoz, C., “*Le geste canal de communication homme/machine. La communication “instrumentale”*”, *Technique et science informatique*, vol. 13, no. 1, pp. 31-61, 1994.
- [Cad-etal.1990] Cadoz, C. Lisowski, L. Florens, JL. “*A Modular Feedback Keyboard Design.*” *Computer Music Journal*, Vol. 14, No. 2, *New Performance Interfaces 2*. (Summer, 1990), pp.47-51.
- [CaWa2000]Cadoz C., Wanderley M. (2000). “*Gesture-Music*”. In M. Wanderley and M. Battier, eds. *Trends in Gestural Control of Music*, pp. 71-94 2000.
- [CaVo2004] Camurri, A., Volpe, G. (eds.). “*Preface.*” In *Gesture-based communication in human-computer interaction : 5th International Gesture Workshop, GW 2003 : Genova, Italy, April 2003 : selected revised papers*. Berlin; New York. Springer, 2004.
- [Chaf1993] Chafe, C. “*Tactile Audio Feedback*” *Proc. Intl. Computer Music Conf.* Tokyo, Sept. 1993
- [Chu1996] Chu, L. “*Haptic Feedback in Computer Music Performance*”. *Proceedings of*

- ICMC, 1996, pp. 57-58.
- [Cook1996] Cook, P. "*Hearing, feeling and playing: masking studies with trombone players.*" In B. Pennycook and E. Costa-Giomi, eds., Proceedings of the 4<sup>th</sup> International Conference on Music Perception and Cognition, pp. 513-518. McGill University. 1996.
- [Cook2001] Cook, P. "*Principles for designing computer music controllers*" Proceedings of the 2001 conference on New Interfaces for Musical Expression, April 01-02, 2001, Seattle, Washington.
- [Del1988] Delalande, F. "*La gestique de Gould: éléments pour une sémiologie du geste musical.*" In G. Guertin, ed. Glenn Gould pluriel. Verdun, Québec : L. Courteau, 1988. pp.114-121.
- [EsslO'Mod2005] Essl, G., O'Modhrain, S. "*Scrubber: an interface for friction-induced sounds.*" Proceedings of the 2005 conference on New Interfaces for Musical Expression. Vancouver, Canada. pp. 70 – 75. 2005.
- [Fin1997] Finney, S.A. "*Auditory feedback and musical keyboard performance*". Music Perception. Winter 1997, Vol. 15, No. 2. pp. 153-174.
- [FleRos1998] Fletcher, N.H., Rossing, T.D. "*The physics of musical instruments.*" New York: Springer, 1998.
- [ForMacL2006] Forsyth, B., MacLean, K. E. "*Predictive Haptic Guidance: Intelligent User Assistance for the Control of Dynamic Tasks*", in IEEE Transactions on Visualization and Computer Graphics, vol. 12, no. 1, pages 103-113, January/February 2006.
- [Gil1992] Gillespie, B. "*The Touchback Keyboard*" Proceedings of the International Computer Music Conference, San Jose, CA, Oct 14-18, 1992. pp. 447--448.
- [Gil1994] Gillespie, B. "*The Virtual Piano Action: Design and Implementation*" Proceedings of the International Computer Music Conference, Aarhus, Denmark, Sept 12--17, 1994. pp. 167--170.
- [Gil1996] Gillespie, B. "*Haptic display of systems with changing kinematic constraints: The Virtual Piano Action.*" Ph.D. dissertation, Stanford University.
- [Gil1999a] Gillespie, B. "*Haptics*". In Perry R. Cook, ed., Music, cognition and computerized sound: an introduction to psychoacoustics. The MIT Press. 1999. pp. 229-245.
- [Gil1999] Gillespie, B. "*Haptics in manipulation*". In Perry R. Cook, ed., Music, cognition and computerized sound: an introduction to psychoacoustics. The MIT Press. 1999. pp. 247-260.
- [Hum-etal.1998] Hummels, C., Smets, G., Overbeeke, K. "*An intuitive two handed gestural*

- interface for computer supported product design.*” In I. Wachsmuth and M. Fröhlich, eds. *Gesture and sign language in human-computer interaction: Proceedings of the II Gesture Workshop.* Heidelberg: Springer-Verlag, 1998. pp. 197-208.
- [HuWaPa2003] Hunt A.D., Wanderley M., Paradis M. “*The importance of parameter mapping in electronic instrument design*”. *Journal of New Music Research*, vol. 32, no. 4, pp. 429-440, December 2003.
- [Jack] Jack:: Low Latency Audio Server. <http://jackaudio.org>
- [Jord2002] Jordà, S. “*Afasia: the ultimate Homeric one-man-multimedia-band.*” *Proceedings of the 2002 Conference on New Interfaces for Musical Expression (NIME-02).* Dublin, Ireland, 2002. pp. 132-137.
- [KurHul1990] Kurtenbach, G. Hulteen, E. “*Gestures in human communication.*” In B. Laurel, ed. *The art of Human-Computer interface design.* Reading, PA: Addison-Wesley, 1990. pp. 309-317.
- [LedKla1997] Lederman, S.J., Klatzky, R.L. “*Haptic aspects of motor control*”. In F. Boller and J. Grafman, eds. *Handbook of neuropsychology*, Vol.11, 131-148. New York. Elsevier. 1997
- [LemSty2005] Leman, M., Styns, F. “*Sound, sense, and music mediation: A historical/philosophical perspective.*” In Marc Leman, Damien Cirotteau, eds. *Sound to Sense, Sense to Sound: A State-of-the-Art. S2S<sup>2</sup> Project Deliverable. Version 0.10 (CVS: November 9, 2005).* Available at:  
[http://www.s2s2.org/docman/task,doc\\_download/gid,70/Itemid,65/](http://www.s2s2.org/docman/task,doc_download/gid,70/Itemid,65/)
- [Lev1999] Levitin, D.J. “*Memory for musical attributes*”. In Perry R. Cook, ed., *Music, cognition and computerized sound: an introduction to psychoacoustics.* The MIT Press. 1999. pp. 209-228.
- [LuFICa2005] Luciani, A., Florens, J.L., Castagne, N. “*From Action to Sound: A Challenging Perspective for Haptics*”. *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'05).* March 2005. pp. 592-595.
- [Mag1989] Magill, R.A. “*Motor Learning: concepts and applications.*” Dubuque, Iowa: Wm. C. Brown. 1989
- [Mat1991] Mathews, M.V. “*The Radio Baton and the Conductor Program, or: Pitch, the most important and least expressive part of music.*” *Computer Music Journal*, 15(4), 1991. pp. 37-

- 46.
- [MiWa2006] Miranda, E.R., Wanderley, M.M. “*New digital musical instruments: control and interaction beyond the keyboard*”. Middleton, Wis. : A-R Editions, 2006.
- [Mul2000] Mulder, A. “*Towards a choice of gestural constraints for instrumental performers.*” In M. Wanderley and M. Battier eds. Trends in gestural control of music. Paris: IRCAM-Centre Pompidou, 2000. pp. 315-335.
- [Nich2002] Nichols, C. “*The vBow: Development of a Virtual Violin Bow Haptic Human-Computer Interface*”. Proceedings of the 2002 Conference on New Interfaces for Musical Expression (NIME-02), pages 29-32, Dublin, Ireland, May 24-26 2002.
- [Niel-et al. 2004] Nielsen, M., Störring, M., Moeslund, T.B., Granum, E. “*A procedure for developing intuitive and ergonomic Gesture Interfaces for HCI*”. In Camurri, A., Volpe, G. (eds.). Gesture-based communication in human-computer interaction : 5th International Gesture Workshop, GW 2003 : Genova, Italy, April 2003 : selected revised papers. Berlin; New York. Springer, 2004. pp. 409-420.
- [O’Mod2000] O’Modhrain, S. “*Playing by Feel: Incorporating Haptic Feedback into Computer-Based Musical Instruments*”. Unpublished PhD Dissertation, Stanford University (2000).
- [O’ModChaf 2000] O’Modhrain, M.S., Chafe, C. “*Incorporating Haptic Feedback into Interfaces for Music Applications*”. 8th International Symposium on Robotics with Applications, ISORA 2000, World Automation Congress WAC 2000.
- [O’ModEss2004] O’Modhrain, S., Essl, G. “*PebbleBox and CrumbleBag: tactile interfaces for granular synthesis*”, Proceedings of the 2004 conference on New interfaces for musical expression, p.74-79, June 03-05, 2004, Hamamatsu, Shizuoka, Japan.
- [ParO’Mod2003] Paradiso, J., O’Modhrain, S. “*Current Trends in Electronic Music Interfaces*”. Guest Editors’ Introduction. Journal of New Music Research 2003, Vol. 32, No. 4, pp. 345-349.
- [Ped et al. 2006] Pedrosa, R., Cheung, B., Maksakov, E. “*The effect of visual and haptic cues on ability to follow rhythms*”. Final report for a team project in CPSC543. Winter Session, 2006. Available at: <https://bugs.cs.ubc.ca/twiki/pub/SPIN/ProjectsIndex/Arrhythmia.doc>
- [Rep1999] Repp, B.H. “*Effects of auditory feedback deprivation on expressive piano performance*”. Music Perception. Summer 1999, Vol. 16, No. 4. pp. 409-438.
- [RoHa2000] Rován, J. Hayward, V. “*Typology of Tactile Sounds and their Synthesis in*



- Gesture-Driven Computer Music Performance*". In Trends in Gestural Control of Music. Wanderley, M., Battier, M. (eds). Editions IRCAM, Paris, 2000.
- [Ros] Rosegarden. Software Sequencer. <http://www.rosegardenmusic.com>
- [Sens] PHANTOM Premium 1.5 Haptic Device. Description and datasheets available at: <http://www.sensable.com/phantom-premium-1-5.htm>. Last accessed: June, 2007.
- [Swin2007] Swindells, C. "*Incorporating Affect into the Design of 1-D Rotary Physical Controls*". Ph.D. Thesis, Univ. of British Columbia, 2007
- [Ver1992] Verillo, R.T. "*Vibration sensation in humans.*" Music Perception. Spring 1992, Vol. 9, No. 3. pp. 281-302.
- [Zyn] ZynAddSubFX. Polyphonic Software Synthetizer: <http://zynaddsubfx.sourceforge.net>.

## **APPENDIX A User Study Consent Form**

An example of the consent form given to experimental users is shown on the following pages.



THE UNIVERSITY OF BRITISH COLUMBIA

## PARTICIPANT'S COPY CONSENT FORM

Department of Computer Science  
2366 Main Mall  
Vancouver, B.C. Canada V6T 1Z4  
tel: (604) 822-9289  
fax: (604) 822-5485

**Project Title:** Physical and Multimodal User Interfaces – Usability and Psychophysics

**Principal Investigator:** K. MacLean, tel. 604-822-8169

**Co-Investigators:** Ricardo Pedrosa, MSc Student, Dept. of Comp. Sc., 604-710-1442

This project is designed to investigate the effect of visual and haptic stimuli as cues while following a musical rhythm, in this case a drum pattern. Our main purpose is to gather information that will serve as a basis to the design of haptic feedback in non traditional computer based musical instruments.

You will be asked to follow a rhythmic drum pattern using a drum machine with and without extra cues on top of the reference sound. The cues will be presented randomly and will consist of blinking lights (as visual stimuli), buzzer vibrations (as haptic stimuli) or a combination of both. We will observe you doing this task and will record your performance in a MIDI track from where we can extract the information we need in order to determine how accurate you were while following the rhythm. No other type of video or audio recordings will be made. You will be asked to complete a questionnaire at the end of the experiment. The experiment will consist of one session that will last no more than 1 hour.

Although only a course project in its current form, this project may, at a later date, be extended by one or more of the student investigators to form the basis of his thesis research.

**TIME COMMITMENT:** 60 minute session  
**HONORARIUM:** You will receive \$10 for participating in this experiment.  
**CONFIDENTIALITY:** *Your identity will be confidential: you will not be identified by name in any study reports. Experimental results will be stored in a secure Computer Science account accessible only to the experimenters. The anonymous data from the experiment will be used in scholarly publication(s).*

You understand that the experimenter will ANSWER ANY QUESTIONS you have about the instructions or the procedures of this study. After participating, the experimenter will answer any questions you have about this study.

You understand that you have the RIGHT TO REFUSE to participate or to withdraw from the study at any time without penalty of any form.

You hereby CONSENT to participate in this study and acknowledge RECEIPT of a copy of the consent form.

If you have any concerns regarding your treatment as a research subject you may contact the Research Subject Information Line in the UBC Office of Research Services at 604-822-8598.



THE UNIVERSITY OF BRITISH COLUMBIA

## RESEARCHER'S COPY CONSENT FORM

Department of Computer Science  
2366 Main Mall  
Vancouver, B.C. Canada V6T 1Z4  
tel: (604) 822-9289  
tax: (604) 822-5485

**Project Title:** Physical and Multimodal User Interfaces – Usability and Psychophysics

**Principal Investigator:** K. MacLean, tel. 604-822-8169

**Co-Investigators:** Ricardo Pedrosa, PhD Student, Dept. of Comp. Sc., 604-710-1442

This project is designed to investigate the effect of visual and haptic stimuli as cues while following a musical rhythm, in this case a drum pattern. Our main purpose is to gather information that will serve as a basis to the design of haptic feedback in non traditional computer based musical instruments.

You will be asked to follow a rhythmic drum pattern using a drum machine with and without extra cues on top of the reference sound. The cues will be presented randomly and will consist of blinking lights (as visual stimuli), buzzer vibrations (as haptic stimuli) or a combination of both. We will observe you doing this task and will record your performance in a MIDI track from where we can extract the information we need in order to determine how accurate you were while following the rhythm. No other type of video or audio recordings will be made. You will be asked to complete a questionnaire at the end of the experiment. The experiment will consist of one session that will last no more than 1 hour.

Although only a course project in its current form, this project may, at a later date, be extended by one or more of the student investigators to form the basis of his thesis research.

**TIME COMMITMENT:** 60 minute session

**HONORARIUM:** You will receive \$10 for participating in this experiment.

**CONFIDENTIALITY:** *Your identity will be confidential: you will not be identified by name in any study reports. Experimental results will be stored in a secure Computer Science account accessible only to the experimenters. The anonymous data from the experiment will be used in scholarly publication(s).*

You understand that the experimenter will ANSWER ANY QUESTIONS you have about the instructions or the procedures of this study. After participating, the experimenter will answer any questions you have about this study.

You understand that you have the RIGHT TO REFUSE to participate or to withdraw from the study at any time without penalty of any form.

You hereby CONSENT to participate in this study and acknowledge RECEIPT of a copy of the consent form.

If you have any concerns regarding your treatment as a research subject you may contact the Research Subject Information Line in the UBC Office of Research Services at 604-822-8598.



THE UNIVERSITY OF BRITISH COLUMBIA

You hereby CONSENT to participate in this study and acknowledge RECEIPT of a copy of the consent form:

NAME \_\_\_\_\_  
(please print)

SIGNATURE \_\_\_\_\_ DATE \_\_\_\_\_

If you have any concerns regarding your treatment as a research subject you may contact the Research Subject Information Line in the UBC Office of Research Services at 604-822-8598.

## **APPENDIX B Ethics Approval Certificate**

A copy of the approval certificate is included on the following page.



The University of British Columbia  
Office of Research Services  
**Behavioural Research Ethics Board**  
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

## CERTIFICATE OF APPROVAL- MINIMAL RISK RENEWAL

<b>PRINCIPAL INVESTIGATOR:</b> Karon E. MacLean	<b>DEPARTMENT:</b> UBC/Science/Computer Science	<b>UBC BREB NUMBER:</b> H01-80470
<b>INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:</b>		
<b>Institution</b>	<b>Site</b>	
UBC	Point Grey Site	
<b>Other locations where the research will be conducted:</b> N/A		
<b>CO-INVESTIGATOR(S):</b> Ricardo Pedrosa Colin Swindells Susan Gerofsky Noorin Fazal David Ternes Matt Savage-LeBeau Mario Enriquez Steve Yohanan		
<b>SPONSORING AGENCIES:</b> Innovation and Science Council of British Columbia - "Physical and multimodal user interfaces - usability & psychophysics" Natural Sciences and Engineering Research Council of Canada (NSERC) - "The design of multi-modal symbolic information displays" - "Orsil title - Physical user interfaces: Communication of information and affect" Various Sources		
<b>PROJECT TITLE:</b> Orsil title - Physical user interfaces: Communication of information and affect		
<b>EXPIRY DATE OF THIS APPROVAL: June 14, 2008</b>		
<b>APPROVAL DATE: June 14, 2007</b>		
The Annual Renewal for Study have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.		
Approval is issued on behalf of the Behavioural Research Ethics Board		