Typology of Tactile Sounds and their Synthesis in Gesture-Driven Computer Music Performance

Dr. Joseph Rovan
email: rovan_j@cmr.fsu.edu
InterMedia Performance And Composition Technologies (IMPACT)
School of Music
Florida State University
Tallahassee, Florida 32306-1180
USA

Dr. Vincent Hayward
email: hayward@cim.mcgill.ca
McGill Research Centre for Intelligent Machines
McGill University, Montreal
Canada H3A2A7

Abstract

In this paper we outline the fundamentals for a tactile feedback system to be used in conjunction with open-air computer music performance devices. Some underlying physiological and perceptual mechanisms of haptics are examined, some currently available open-air controllers are reviewed, and previous technologies and experiments regarding haptic/tactile feedback are surveyed. Our VR/TX system is proposed as a solution for adding tactile feedback to open-air controllers; experiments show that the VR/TX vibrotactile stimulators provide invaluable perceptually-significant tactile feedback when used in conjunction with an open-air music controller. A typology of tactile sound events is also described, as well as the notion of a tactile simulation event (TSE).

1. Introduction

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.

Beginning with the Theremin, and continuing today with contemporary devices such as Don Buchla's “Lighting II”, many of these alternative performance interfaces have sought to “unchain” the performer from the physical constraints of holding, touching and manipulating an instrument. By using non-contact sensing technologies such as near-field capacitive measurement, infrared, ultrasound, video, etc. , these instruments have given musical control to “open air” gestures not traditionally associated with music making. Performers now stroke invisible marimbas, dancers generate sound with a twist of their bodies. Without the necessity of manipulating an instrument, these touchless performers magically conjure music from “thin air.”

In spite of the freedom such open-air controllers allow, however, at the same time they have eliminated one of the key channels of information through which performers gauge the response of their instrument and the accuracy of their performance: haptic sensations (defined below). In the traditional closed loop between performer and instrument, intention gives rise to gesture, gesture gives rise to sound, and feedback—including visual, aural and haptic/tactile information—is used to gauge the result. Matched against intention for immediate correction, feedback ultimately stimulates short term as well as long term learning. Indeed, motor control—essential in all musical performance—*cannot* be separated from haptic sensation. Kinesthesia and touch are a part of performance, from the buzz of the reed to the bruise of the chin rest. But in the electronic world of synthesis algorithms and control rate data, open-air controllers are both divorced from the actual sound producing mechanisms, as well as from the body itself. Simply put, the computer music performer loses touch. (See Figure 1)

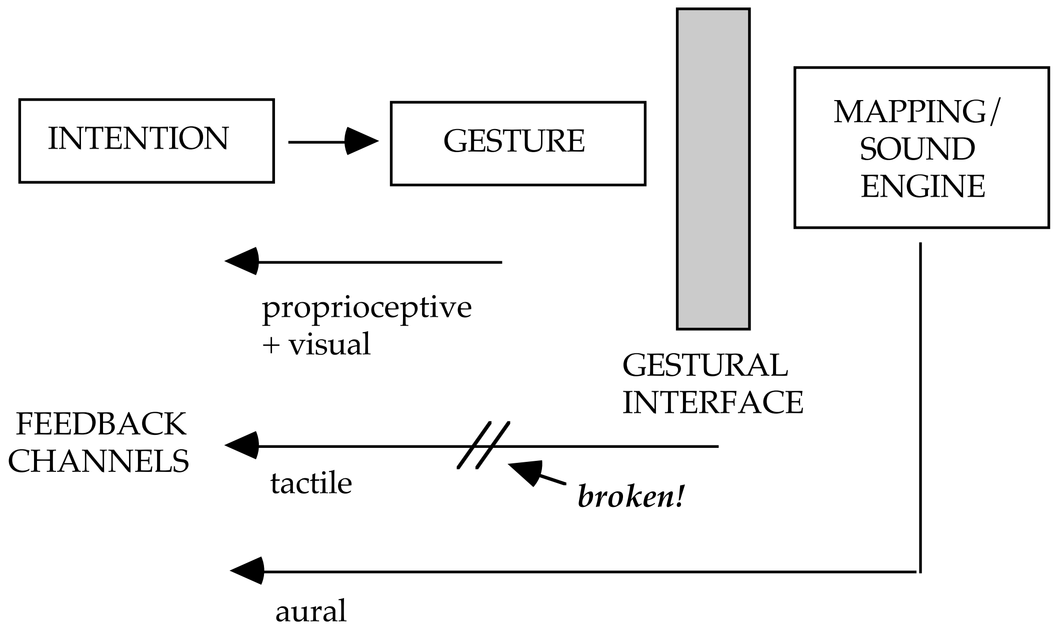

Figure 1: Performing with an open-air controller

In systems which track a performer's gesture through remote sensing technologies—such as infrared, ultrasound, or video—the tactile feedback loop is broken, forcing performers to rely on proprioceptive, visual and aural cues.

2. Haptic sensations

Being "in touch" with an acoustical instrument infers that a performer is receiving mechanical feedback cues via a variety of physiological and perceptual mechanisms. These *haptic* sensations broadly comprise two main channels:

> *Tactile sensation:* associated with discriminative touch as in the perception of surfaces. It comprises sensations that include pressure, local features such as curvature (arising from edges, embossings and divets), orientation, puncture, texture, thermal properties, softness, wetness, friction-induced phenomena such as incipient slip, adhesion or lack

thereof, micro failures, persistent vibration. Tactile sensation facilitates or makes possible the execution of many perceptual tasks; it participates in motor control at almost every level, from spinal reflexes, to motor programs, from involuntary reflexes to gesture plans.

Several kinds of receptors have been found to mediate tactile sensation in the skin or in the subcutaneous tissues; consequently, it is customary to designate the skin as the seat of this sense (a very large organ, indeed; we carry 5 kg of it). The biophysical attributes of the skin vary tremendously with the parts of the body it covers, from the tongue to the toe. Overall, the tactile system occupies a great part of the afferent pathways of the peripheral nervous system as well as a significant part of the central nervous system, all in a beautiful hierarchical organization culminating at the primary somatic sensory cortex. The study of tactile sensations were at the core of the work of such pioneers as Weber and Katz.

Proprioceptive, or kinesthetic perception: awareness of one's body state, including position, velocity and forces supplied by the muscles through a variety of receptors located in the skin, joints, muscles, and tendons. It is often associated with limbs, but of course applies to all articulated parts of the body subject to voluntary motor control. While proprioception is often described as providing feedback in the execution of motor control, it is also essential to formation of internal models (along with vision and vestibular inputs) which make possible the generation of feedforward signals to the motor system. Proprioception participates in a three-level hierarchical organization mapped anatomically onto the spinal cord, the brain stem and the cortex along with two important subcortical systems (basal ganglia and cerebellum). Together, proprioception and tactile sensation are fundamental to manipulation and locomotion, two most remarkable biological phenomena.

Recently, it has become common to speak of the "haptic channel" as a means to collectively designate these sensory and motor modalities under one unified umbrella, thus escaping the traditional Aristotelian classification of the senses. In this way it is possible to discuss more conveniently such topics as the influence of mechanical experience (in addition to acoustic and optical experience) on memory; it is also possible to think of certain anatomical parts—in particular the hand—as unitary organs where action on the world and perception of it are no longer divorced and discussed separately. In particular, both tactile and kinesthetic channels work together to create haptic cues, providing humans with a sense of orientation within and in contact with an environment.

In the context of a musical performance, one can recognize another important feedback component tightly coupled with the above: *egolocation,* meaning the awareness of one's overall position within a defined space, or with respect to objects in that space. It is also clear that *visual feedback* participates together with tactile and proprioceptive feedback to give a musical performer the ability to orient, critique and calibrate their performance gestures. As mentioned earlier, due to the lack of tactile feedback, egolocation becomes the primary feedback skill necessary to develop in open-air controller performance.

2.1 Performing with open-air controllers

A prototypical computer music performer using an open-air controller might control his or her performance through a combination of feedback loops operating at different rates. Consider the tasks in turn:

1. Think ahead (initial intention)
2. Initiate movement (performance gesture)
3. Gauge performance gesture through vision and proprioception
4. Hear resulting sound
5. Adjust gesture from vision and proprioception
6. Adjust intention from hearing
7. repeat...

In this scenario, the performer receives feedback through hearing, proprioception and visual cues. Nevertheless, this scenario presents a severe obstacle to the evaluation of gestural accuracy. This is due to the fact that—by their very nature—such systems ignore the fastest and most immediate channels of feedback which control movement: muscle tonus (thin air) and tactile sensations (non-contact).

While it is beyond the scope of this paper to analyze in detail how these signals converge in order to correct the execution of an intended gesture (as well as the intention itself), it is known that the cerebellum (10% of the brain) is the seat of such mechanisms and that it receives inputs from the periphery as well as from all levels of the central nervous system, including information about the gesture plan itself (the corollary discharge).

The approach taken here is to restore one missing channel, that of tactile sensations, in an effort to assist gesture execution, while at the same time retaining the initial premise: to relieve the performer of the tedium of manipulation. The hypothesis is that minimally but adequately designed tactile stimulation, produced in response to movement, will provide a valuable aid to the open-air performer.

2.2 Touch—sensory systems/physiological components

The ability to sense objects through touch is actually the combined result of several specialized receptors within the skin, each of which responds to a different type of stimulus. There are four types of receptor, which can be classified according to whether they have punctate sensitivity (finely focused deformation sensitivity) versus diffuse sensitivity, and whether or not they are fast adapting or slow adapting. The receptors are differentiated by their rate of adaptation (Fast: Pacinian and Meissner corpuscles; Slow: Merkel's disks and

Ruffini cells). They are also differentiated by the size of the receptive fields (Punctate: Meissner and Merkel's; Diffuse: Pacinian and Ruffini) and whether these receptive fields are stimulus specific. They are further differentiated by their anatomical structure as well as by their location and structural relationship with respect to the surrounding tissues (Shallow: Merkel's and Meissner; Deep: Pacinians and Ruffini's). It should be noted that Pacinian corpuscles are also found between layers of muscles and in interosseous membranes. In addition, the density of inervation varies greatly within different areas of the body.

It is not surprising that certain areas of the skin are extremely stimulus specific. Throughout the known history of musical instrument making we can see how this phenomenon has been exploited. For example, the lips respond exquisitely to pressure, while they are less proficient than the fingertips at appreciating the textural properties of surfaces. In playing a clarinet, the instrumentalist is acutely aware of the pressure on the lower lip as a function of embouchure tightness and intonation, while sensing the reed tip with the tongue is an intrinsic part of executing different types of articulations. As another example, the concert pianist Richter is said to have carried pieces of sand paper in his pocket to prepare the surfaces of piano keys before performances, modifying their tactile properties[5].

In many cases, the quality of contact with these sensitive areas is actually *too intense*, and players must spend hours of practice time developing calluses—as a sort of "filter"—in order to tune the bandwidth of the tactile channel. Thus pain also figures into the family of skin sensations which contribute to the feedback necessary to play an instrument; one could as well include the sensation of thermal properties, and finally, the movement of hair via the hair root plexuses, as a sort of "warning signal" in advance of actual skin contact.

2.3 Sensation of texture—spatial versus temporal coding

The perception of texture is one of the most important functions of tactile perception. Many studies have explored the manner in which texture is neurophysiologically encoded; it now appears that it is perceived as both a temporal phenomenon (as vibration) and as a spatial phenomenon, depending on the textural scale. For surface details measured in microns—*microgeometric textures*—texture seems to be perceived via temporal encoding. With the aid of the fast-adapting Pacinian corpuscles, such surfaces are typically perceived as a vibration (LS90). When surface details are separated by more than .125 mm, however, texture is sensed as being "rough," not vibrating, and seems to be encoded spatially instead of temporally, perceived via slowly-adapting receptors. Depending on the scale of surface detail stimuli, different tactile receptors are operative. The neural machinery eventually translates these large collections of elementary stimuli into stable percepts.

In order to ascertain a surface's finer details, it also appears that lateral motion between fingertip and surface is important; the relative shearing motion may be produced by either the surface or the observer moving. Typical exploratory gestures often involve a rubbing back-and-forth motion, the lateral gesture creating a temporal encoding of the surface in the form of multiple, repeating tactile stimuli (LK87). These periodic sensations can be understood as a *vibrotactile* sensation with a particular frequency and spectrum. In our current experiments we have used temporal encoding to represent texture, while some other experiments have been carried out synthesizing texture via spatial encoding.

3. Gestural performance systems: out of touch

Today there are many commercially-available open-air controller systems, as well as many custom one-of-a-kind devices. Commercial systems include: Big Briar's Theremin, Buchla's Lightning II, Interactive Light's Dimension Beam (also built into several Roland keyboards), the Polhemus device, STEIM's BigEye, David Rokeby's Very Nervous System, Palindrome Inter-media Performance Group's Dynamic Fields, WaveRider, the BodySynth and BioMuse, to name a few. (See Figure 2)

DEVICE	SENSOR TECHNOLOGY
Theremin	near-field capacitive measurement
Lightning II	infrared
Dimension Beam	infrared
BigEye	video tracking
Very Nervous System	video tracking
Palindrome Dynamic Fields	video tracking
WaveRider	biological
Polhemus	electromagnetic

Figure 2: Some commercially-available open-air controllers

Custom-built controllers often include several sensor technologies, and may only partially use open-air sensing. Examples would include the Twin Towers infrared controller designed at CNUCE/C.N.R. (TB97), the ultrasound sensors on Michel Waisvisz' "The Hands" developed at STEIM, the video dancer tracking system described by Antonio Camurri (Cam95), as well as accelerometers used as rotation sensors on the glove designed by one of the authors.

3.1 Features of currently available systems

As shown above, there are many types of sensor technologies used in these systems, including near-field capacitive measurement, infrared, ultrasound, accelerometers, magnetic, radio, etc. Each technology has advantages and disadvantages when it comes to sensitivity, cost, complexity of implementation, portability, etc. Most important, however, is the match of sensor technology to the type of open-air gesture to be used by the performer; different sensors are more suitable for particular performance situations, and each system has typically been tuned to suit a particular vocabulary and scale of gesture.

> *(N.B.—IRCAM has developed an extensive online database of sensor classification, including details on implementation, cost, availability, and suitability for particular gestures. Unfortunately it remains only accessible to those inside IRCAM at present.)*

As a result of this fine tuning, the range of sensitivity differs greatly from system to system, even between those using the same sensor technology. For example, the infrared technology of the "Twin Towers" is calibrated for close proximity gestures resembling piano performance technique, while the infrared system of the Lightning II is designed to allow for much larger gestures, like those of a percussionist (although it does track fine movement detail as well). Likewise, video tracking systems are more suitable for larger-scale gestures, like those of an entire dance company, while accelerometers may be used to measure extremely subtle variations in limb rotation.

Common to all of these open-air controllers, nevertheless, is the paradigm of a virtual performance space which surrounds the performer, created by either the limits of the sensors' sensitivity range, or by arbitrary boundaries set by the designer. This virtual space may have as its locus the entire body, as with a video tracking system, or only a particular body part, such as a hand tracked by infrared or sensed via biological sensors.

Regardless of the size and shape of this virtual performance space, the different systems typically transduce continuous gesture of a performer into an analog signal. From there an analog-to-digital converter samples the analog input and translates it into a digital representation, which is then mapped via various mapping strategies (RWD97) onto control of a sound synthesis algorithm. More often than not this digital representation is in the form of MIDI continuous controller data, which is then used to trigger MIDI sound sources. There is, however, the well-known problem of responsiveness using MIDI due to its low bandwidth and serial transmission protocol; some systems have attempted to bypass the MIDI bottleneck by using direct proprietary triggering mechanisms. Experiments have also been carried out by one of the authors using the IRCAM Signal Processing Workstation (ISPW) that bypass the MIDI problem altogether by treating the gestural data as an audio-rate signal within a real-time DSP environment. These experiments have been carried further by researchers at the Center for New Music and Audio Technology (CNMAT) in Berkeley, California, using similar environments such as MAX/MSP (FW98).

Movement may be tracked within a virtual performance space as absolute position, velocity, acceleration, or a combination of any of the above. The simplest and most straightforward of these is absolute position; in this case the space is often divided into different regions, or zones, with different control data mappings assigned to each. With some systems, simply being within a zone may trigger an event, while other systems may only trigger an event if activity within a zone rises above a certain threshold, as when the Palindrome video system tracks a dancer. Still others, such as the Lightning II, use zone boundary crossings as a triggering parameter. In all of these cases, however, the problem associated with using absolute position is its one-to-one correspondence with body movement, which can easily lead to false triggerings. Using velocity and acceleration obviates this problem by differentiating gestural change over time, decoupling the "hard wiring" of gesture and absolute position tracking. (It should be noted, however, that there may be a corresponding tradeoff in lag time.) Most of the systems mentioned are completely programmable in their response, such that many types of gesture tracking modes are possible. In fact, if collective patterns of position, velocity, and/or acceleration are tracked simultaneously, and then analyzed for global features, higher order events may be extracted and particular gestures may be recognized (SOH95).

3.2 Limitations of current systems

It can be seen that all of the above systems require a performer to attain a high level of proficiency and physical control, as body position is measured/sensed without reference to any physical contact with an external landmark. Due to their design, the haptic/tactile feedback channel is not an option for gauging the accuracy of gesture. As a result, all of these open-air devices require an intense amount of visual feedback as well as exquisite postural control by the performer (and, of course, aural feedback from the resulting sound synthesis). Any non-contact controller that measures position, velocity or acceleration will suffer from these same problems.

Overall, one might make the following observations of most non-contact musical controller systems:

- Reliance on **visual** feedback
- Reliance on **proprioception**
- Reliance on **egolocation**

In such systems the visual feedback derived from the performer's own body position may also be augmented by computer-driven displays, such as watching an "in-zone" or "out-of-zone" indicator light or a real-time graph of sensor activity. Nevertheless, the reliance on visual feedback is easy to critique. Some reasons to avoid a reliance on visual feedback might include:

- Visual feedback is typically only needed by beginning performers. (Regardless of instrument, experienced performers seldom rely on visual cues.)
- Visual displays are inadequate, often impractical in performance.

• There are much more important visual functions during a performance. Important visual functions during a performance could include interacting with the audience, interacting with other performers, reading music as well as dramatic aspects.

• Mechanical channels of feedback are more tightly looped psychologically and physiologically—in time and space—than vision (or audio).

On the other hand, proprioception and egolocation are both vital components of the wide bandwidth feedback channel necessary for performance with open-air controllers. Nevertheless, they both fall short if relied upon as the sole means of orientation. Reasons to avoid a heavy reliance on proprioception and egolocation might include:

• Difficult—if not impossible—to achieve consistent control and repeatable results.

• Requires long training and special skills.

• Never accurate in absolute coordinates. Instead of aiming for an absolute position in space, one typically nudges the body (or a body part) precisely in the direction of certain landmarks, in response to a feedback stimulus.

3.3 Adding Feedback

None of the commercial open-air controller systems available today incorporate haptic/tactile feedback in their design. Perhaps for various aesthetic reasons—or more likely due simply to questions of cost and complexity—the feedback channel has not been widely addressed. Nevertheless, there have been several encouraging experiments in designing custom haptic/tactile feedback musical performance systems, using a variety of methods.

A distinction must first be made, however, between tactile simulators and haptic devices. Tactile stimulators use some mechanism of controlled skin deformation (matrix of pins, typically) or vibrotactile stimulators (devices vibrating at a given frequency, in contact with the skin at one or several locations). Haptic devices, on the other hand, use a mechanical system that links the user to a wide-bandwidth source of mechanical power (HA96). As the operator moves the haptic device across space, acting as a virtual probe, mechanical feedback is transmitted to simulate the passing of the probe over, against or through a virtual object, which itself may be static or dynamic. These systems are also known as "force feedback" devices, due to their ability to exert large-scale mechanical force back onto the operator. The important difference between tactile simulators and haptic devices, besides the difference in scale, is that a tactile simulator simulates the effect of skin touching a surface, while a haptic device serves as a mediating device through which one explores or manipulates a virtual mechanical system.

Perhaps one of the most documented examples of a musical force-feedback implementation is the *Clavier Retroactif Modulaire (CRM)*, or Modular Feedback Keyboard, designed at *Association pour la Création et la Recherche sur les Outils d'Expressions* (A.C.R.O.E. in Grenoble, France). The project has been headed by Claude Cadoz (CLF90), and consists of a piano-like keyboard with computer-driven mechanical (motor) feedback devices for each of its sixteen keys. The system has been integrated with an elaborate gestural recording/editing/playback system and a physical modeling sound synthesis engine called the CORDIS-ANIMA system. Other experiments at A.C.R.O.E. include developing a system for composition using these tools, based not on the manipulation of synthesis parameters, but on the gestures that would cause them (Cad88) (GF88).

Some experiments in designing force feedback musical controllers have relied on "shape memory alloy" actuators that change shape in response to temperature modulation (Bon94). Bongers attached "Muscle Wire"[1] (one commercially available form) to a glove controller, using the wire to control the flexibility of the fingers; he also suggests the use of small plates made of the same substance to stimulate the finger tip. In place of shape memory alloys, other projects have used pneumatic devices to mimic the contraction of muscles (Cad94).

The synthesis of tactile feedback by the use of vibrotactile devices has been investigated in several experiments as well. These include experiments by Chafe (Cha93) and Chu (Chu96), as well as the "Dexterous Hand Master" device used at MIT (Mac92). Chafe's experiment involved coupling the audio output from a physically-modeled brass instrument to a device that would allow the performer to sense the modes of vibration through the lip. Chu suggests that tactile localization, by applying tactile audio signals to both hands, could be correlated to audio spatialization. The current authors' experiments focus on this use of audio as a tactile simulator, exploring ways to encode the tactile representation to make it perceptually significant.

4. The VR/TX Vibrotactile Stimulator Project

The VR/TX (an acronym for "virtual texture", pronounced "vortex") experiment initiated out of an attempt to incorporate feedback into a custom glove controller designed by one of the authors while at IRCAM. Initial prototypes were made in collaboration with Mark Goldstein[2] using earphones taken from scavenged head sets. Subsequently, the authors set out to design a vibrotactile stimulator system that would explore in detail the perceptual attributes of a simple vibrotactile vocabulary synthesized in response to a gesture: in short, a basic collection of *tactile sounds.* Overall, the project goal was to augment performance using open-air controllers by creating artificial tactile feedback through the use of custom-designed transducers. Detailed goals of the project included:

Short Term:

- Augment the author's custom glove controller with tactile feedback.

- Enable the performance of existing spatially-controlled pieces without the use of visual displays.

Long Term:

- Make open-air MIDI controllers easier to use.

- Add functionality to existing controllers.

- Develop new controllers.

4.1 VR/TX System—overview

The VR/TX system consists of both hardware and software components. Overall the feedback system presently comprises a Macintosh PowerBook running an application written in MAX/MSP[3], a specially-designed power amplifier, and one or more vibrotactile stimulators (interfaced with either the hands or feet). For open-air controller input an Interactive Light "Dimension Beam" infrared controller was initially used[4]. (See Figure 3)

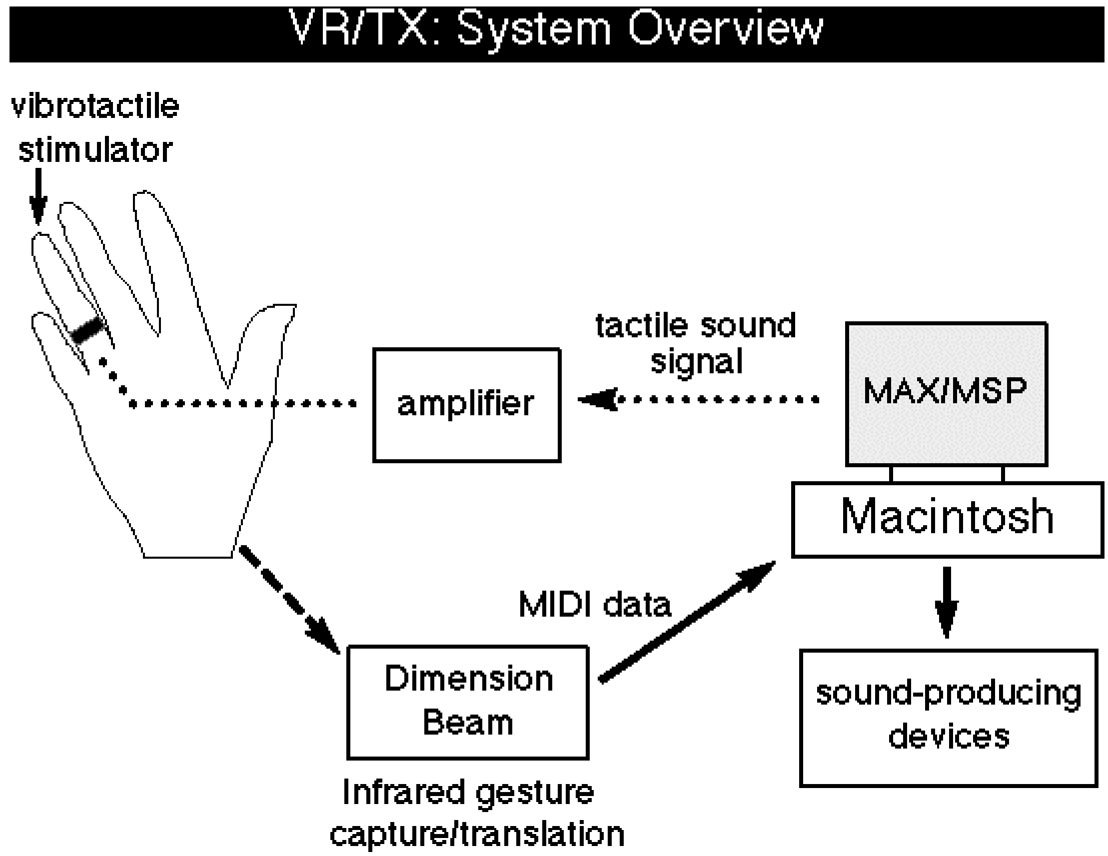

FIGURE 3: VR/TX System

4.2 VR/TX hardware

The core component of the VR/TX system is a specially-designed transducer driven by an amplifier and engineered to transduce texture-simulating signals to the wearer via direct contact. The stimulators have been designed in a variety of shapes and sizes, including: a ring-mounted version, unobtrusive as possible to allow performance on existing instruments; a version for mounting on the custom glove controller; and a larger version embedded in a surface on which a performer can stand, sensing tactile feedback through the feet. Regardless of form factor, however, each of these variants share a common design (See Figure 4). In this design, the coil and the magnet are sized to have similar masses and the suspension is selected so that the principal resonance is well below the frequency operating range. Under these conditions, the transduction is optimized and the acceleration response flat. In one variant of the design, only the housing is in contact with the skin (via a ring on a finger or via a surface on which the foot rests). In the other both the housing and a magnet extension are in contact with the skin; this second case is more appropriate to being embedded in a glove controller. The impedance of the coil was made compatible with ordinary audio amplifiers.

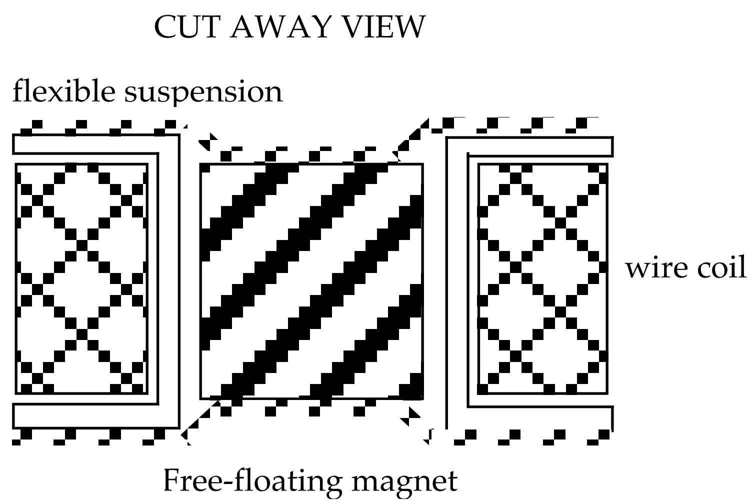

FIGURE 4: Tactile stimulator construction—detail

4.3 VR/TX software:

The input to the stimulator is actually an audio signal, the characteristics of which form the most important determining factor in the performance and perceptibility of the tactile sensation. This audio signal is generated in real-time in the MAX/MSP environment, running on a Macintosh G3 PowerBook. In order to codify the different types of tactile signals and to facilitate the communication between the synthesis and feedback components of the system, a concise protocol was created to relate certain gestural events to tactile stimulation events (TSE). This protocol qualified the following parameters of each TSE:

1. Frequency
2. Waveform
3. Envelope
4. Duration
5. Delay (between repetitions)
6. Amplitude
7. Number of repetitions

An application was created in the MAX/MSP environment that classified incoming controller data (from the Dimension Beam, in this case), and generated a particular TSE according to the above parameters. (See Figure 5)

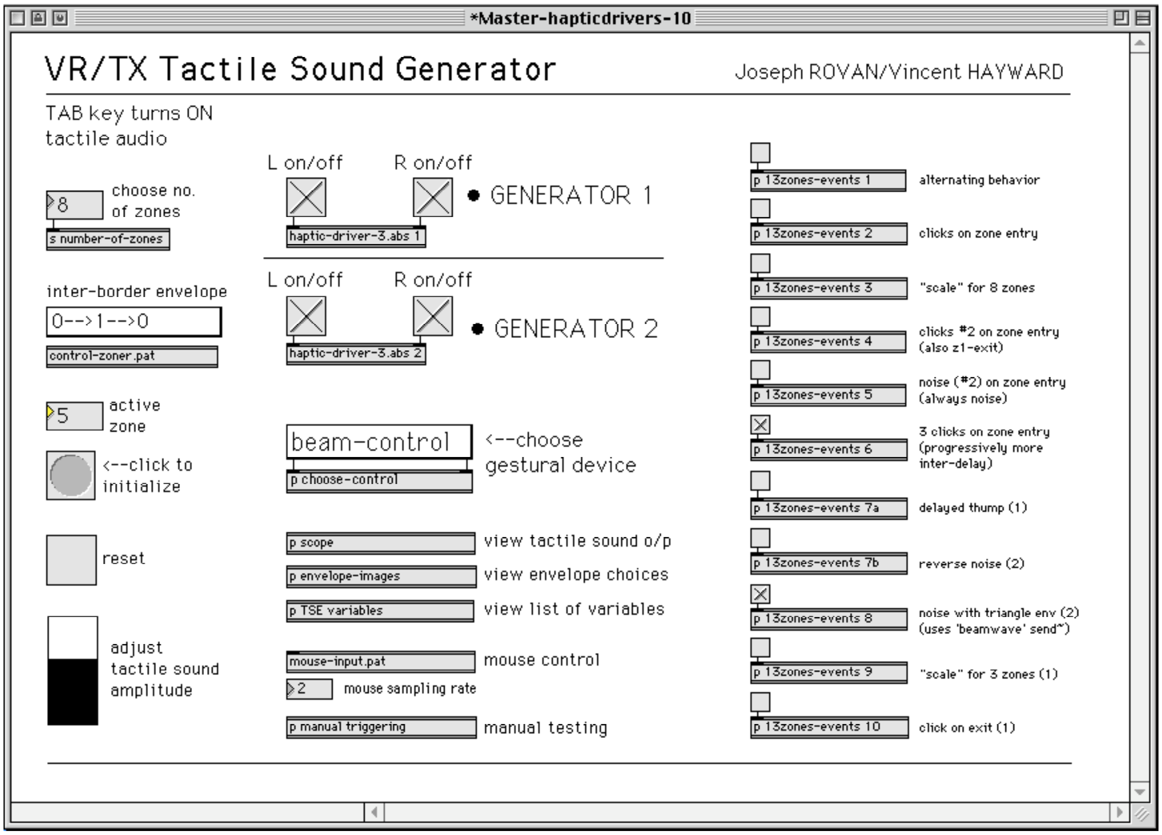

FIGURE 5: MAX/MSP Environment

4.4. Modalities of "Tactile Sounds"

In exploring the myriad ways that audio could be used to simulate tactile feedback, we began by differentiating them according to the perceptual mechanisms of texture (see section 2.3 above). Ultimately we arrived at the following typology of "tactile sound", along with some possible problems associated with their respective implementations:

Type A—Time dependent:

- Examples: periodic simple tones (sine), frequency-rich sounds.
- Variables: frequency, envelope, spectral content, duration, repetition.
- Problem: to describe the set of perceptually meaningful signals

Type B—Space dependent:

• Examples: textures and discontinuities, wavetables, noise.

• Variables: modulate as a function of displacement, velocity and acceleration.

• Problem: how to generate, and how to relate meaningfully to one's movements.

Using this classification, we set up the following correspondences between gesture and typology of tactile sound:

Type A—Time dependent correspondence:

• Use to signify discrete events, either caused by the performer or machine.

• Example: switching program states, zone borders.

Type B—Space dependent correspondence:

• Use to guide continuous movement.

• Example: relation to absolute position, speed of entering a zone, sharpness of a "whip" or hitting motion, virtual friction of "bowing" movements.

4.5 Experiments

We conducted two different experiments, differentiated by the hardware configurations for hand versus foot stimulus. In each experiment, we explored variations in the above-mentioned TSE parameters of frequency, waveform, envelope, duration, delay between repetitions, amplitude, and number of repetitions.

4.5.1 Hand-sensed design

The stimulators were tried on one hand as well as on two hands. For the two-handed scenario there were two variants: the same signal sent to both hands, or different signals sent to each hand. Ultimately, the most perceptually powerful was the two-handed, both same signal, configuration. Several different placements on the hand were tried, including fingertip, back of the hand, wrist, and the phalange of the finger. Coupling the vibrotactile signal to the bone structure, as it is accomplished by the ring design (in effect, only a small tissue layer separates the skin from the bone in that region), greatly improved perception; consequently, the ring-mounted design was found very efficient due to this close coupling with the finger. (See Figure 6)

FIGURE 6: ring-mounted vibrotactile stimulator

With regards to frequency, we first sought to ascertain what was perceptually significant. With discrete frequency, we found that a range from 70 to 800 hertz was perceptible, broadly divided into 8 to 10 discrete steps. This limited range could be called the “scale” perceivable via skin contact; obviously the compass of the vibrotactile connection is not very high. More important, however, was the use of particular audio gestures, such as a rapidly rising pitch curve versus a rapidly descending pitch curve, or simply correlating continuous pitch change (over the perceivable range) to absolute position within a virtual space. Larger-scale audio gestures such as these proved to be more perceptible—and easier to remember—than trying to recognize a discrete pitch step.

Waveform, or spectral content, was very significant. The gamut from pure sine tone to frequency rich spectrum to noise was characterized as a continuous transition from smoothness to roughness. Perhaps most significant of all, however, was the event envelope, combined with duration. We had good results using a fast attack/release envelope to signify a zone-boundary crossing event with a short tone burst, while using noise with a cosine envelope that peaked at each zone boundary created the sensation of a continuous rough surface with ridges at each zone crossing.

The delay parameter was used, together with the number of repetitions parameter, to set up "meta" events, such as a series of pitch bursts that would progressively get closer together as the virtual space was crossed. The use of different sets of repetitions was very perceptually significant, as it proved to be much easier to distinguish between differing numbers of short events—such as tone or noise bursts—versus discrete frequencies. Overall amplitude was simply a global parameter that ended up not being manipulated, as the event amplitudes were controlled by the various envelopes. Also, we noticed that the threshold of tactile "hearing" is tied to duration of the event; with fast, complicated events, there was not much room for fine gradations of amplitude.

4.5.2 Foot-sensed design

Experiments with the foot-sensed version of the vibrotactile stimulators followed a similar design. The major obstacle with this version, however, was the more limited bandwidth of the foot's tactile sense, coupled with the impedance of various types of shoe which did not match that of the transducer itself. Also, it was difficult to combat the resonance of the plate that housed the stimulator; thus, at high amplitudes, the plate would resonate audibly in response to the tactile sound. We foresee that these two problems could be easily addressed in a future version. (See Figure 7)

FIGURE 7: foot-sensed vibrotactile stimulator
(foot raised to show stimulator)

Nevertheless, there were promising results. The benefits include the lack of hand impediment, and the easy adaptation to a variety of scenarios. Noise seemed particularly suited as a tactile signal for the feet, as well as simple high-amplitude bursts for zone orientation. Synthesizing the virtual location of a tactile sound located *between* the two feet (as a generalized version of saltation) by using a slight delay between left and right signals also seems promising; some preliminary experiments were carried out. This localization ability with regards to tactile sound has been documented in previous research (Ges70); it would be natural to correlate this effect to sound spatialization, for example.

5. Tactile Stimulators—Conclusions/Future Work

In this paper we outlined the fundamentals for a tactile feedback system based on audio signals. The underlying physiological and perceptual mechanisms were examined, and experiments were carried out using the custom designed VR/TX system. Experiments showed that the VR/TX vibrotactile stimulators provide invaluable perceptually-significant tactile feedback when used in conjunction with an open-air music controller. A set of parameters were used to describe tactile simulation events (TSE); each of these was explored as a means of encoding and varying the tactile feedback. Two different typologies of tactile sound were identified—time dependent and space dependent—linked to the perceptual mechanisms of texture. In the end, we found that it was perceptually significant to combine the two modes into one TSE; for example, using tone bursts for zone boundaries, a variable noise envelope to signify position within a zone, and continuously variable pitch as a rough correlate to absolute position.

The immediate future goals of the project include:

- Move beyond the current prototype stage.
- Better and more effective hardware design.
- More responsive software implementation.

Longer term goals include further development of the localization effect mentioned above, and linking the system with another current project involving a haptic real-time gestural analyzer/recorder. Also proposed is a tactile interface for sound editing and browsing.

6. Acknowledgments

The first author would like to thank the Florida State University First-Year Professor Grant Program and the Real-Time Systems team at IRCAM for supporting this research. Thanks also go to Mark Goldstein, Marcelo Wanderley, Norbert Schnell and Marie-Hélène Serra. The second author would like to thank McGill University for granting a sabbatical leave which made this research possible, and IRCAM, Paris, for generously hosting him. Funding for this research was also provided by the project "Core Issues in Haptic Interfaces for Virtual Environments and Communications" supported by IRIS (Phase 3), the Institute for Robotics and Intelligent Systems part of Canada's National Centers of Excellence program (NCE), and an operating grant "High Performance Robotic Devices" from NSERC, the National Science and Engineering Council of Canada.

Notes:

[1] Muscle Wire is available from Mondo-tronics, San Anselmo, California, USA.

[2] Mark Goldstein is a percussionist, programmer and researcher based in Menlo Park, California.

[3] MAX and MSP are both available from Cycling'74, at www.cycling74.com

[4] The Dimension Beam (no longer manufactured) uses optics to control the sensitivity zone for an infrared proximity device, giving consistent and predictable results; it produces MIDI controller data as output.

[5] CBC Broadcast: Great Pianists of the Twentieth Century, May 18 ’99.

References:

[Bon94] Bert Bongers. The use of active tactile and force feedback in timbre controlling electronic instruments. In *Proceedings of the International Computer Music Conference*, pp. 171-174, 1994.

[Cad88] Claude Cadoz. Instrumental gesture and musical composition. In *Proceedings of the International Computer Music Conference*, pp. 1-12, 1988.

[Cad94] Claude Cadoz. *Le geste canal de communication homme/machine.* In Technique et science informatiques, volume 13, no. 1, 1994, pp. 31-61.

[Cam95] Antonio Camurri. Interactive Dance Systems. In *Proceedings of the International Computer Music Conference*, pp. 245-252, 1995.

[Cha93] Chris Chafe. Tactile audio feedback. In *Proceedings of the International Computer Music Conference*, pp. 76-79, 1993.

[Chu96] L. Chu. Haptic feedback in computer music performance. In *Proceedings of the International Computer Music Conference*, pp. 57-58, 1996.

[CLF90] C. Cadoz, L. Lisowski, and J-L Florens. Modular Feedback Keyboard. In *Proceedings of the International Computer Music Conference*, pp. 379-382, 1990.

[CR90] C. Cadoz and C. Ramstein. Capture, representation, and "composition" of the instrumental gesture. In *Proceedings of the International Computer Music Conference*, pp. 53-56, 1990.

[FW98] A. Freed and D. Wessel. Communication of musical gesture using the AES/EBU digital audio standard. In *Proceedings of the International Computer Music Conference*, pp. 220-223, 1998.

[Ges70] G. Gescheider. Some comparisons between touch and hearing. *IEEE Transactions on Man-Machine Systems*, pp. 28-35, March 1970.

[GF88] S. Gibet and J-L Florens. Instrumental gesture modeling by identification with time-varying mechanical models. In *Proceedings of the International Computer Music Conference*, pp. 28-40, 1988.

[HA96] Hayward, V. Astley, O.R. 1996. Performance Measures for Haptic Interfaces. In *Robotics Research: The 7th International Symposium*. Giralt, G., Hirzinger, G., (Eds.), Springer Verlag. pp. 195-207.

[LK87] Lederman, S.J., Klatzky, R.L., Collins, A., & Wardell, J. (1987). Exploring environments by hand or foot: Time-based heuristics for encoding distance in movement space. Journal of Experimental Psychology: Learning, Memory, & Cognition, 13(4), pp. 606-614.

[LS91] LaMotte RH and Srinivasan MA. Surface microgeometry: Neural encoding and perception. In *Information Processing in the Somatosensory System,* Eds: O. Franzen and J. Westman, Wenner-Gren Intl. Symposium Series, Macmillan Press, 1991.

[Mac92] Tod Machover. *Hyperinstruments: A Progress Report.* MIT Media Laboratory, Massachusetts Institute of Technology, January 1992.

[RWD97] J. Rovan, M. Wanderley, S. Dubnov, and P. Depalle. Instrumental gestural mapping strategies as expressivity determinants in computer music performance. In *Proceedings of the KANSEI—The Technology of Emotion AIMI International Workshop*, pp. 68-73. 1997.

[SOH95] H. Sawada, S. Ohkura, and S. Hashimoto. Gesture analysis using 3D acceleration sensor for music control. In *Proceedings of the International Computer Music Conference*, pp. 257-260, 1995.

[TB97] L. Tarabella and G. Bertini. Original gesture interfaces for live interactive multimedia performances. In *Quatrièmes Journées d'Informatique Musicale/JIM'97,* pp. 41-45, 1990.