The GRIP MAESTRO: Idiomatic Mappings of Emotive Gestures for Control of Live Electroacoustic Music

Michael Berger
Stanford University, Center for Computer Research in Music and Acoustics
The Knoll, 660 Lomita Court
Stanford, CA, 94305
1 (650) 723-4971
mpberger@stanford.edu

ABSTRACT
This paper introduces my research in physical interactive design with my "GRIP MAESTRO" electroacoustic performance interface. It then discusses the considerations involved in creating intuitive software mappings of emotive performative gestures such that they are idiomatic not only of the sounds they create but also of the physical nature of the interface itself.

Keywords
sensor-augmented hand-exerciser, emotive gesture and music, musical mapping strategies, human-controller interaction, passive haptic feedback, novel musical instrument, Hall effect sensor, accelerometer, Arduino Nano.

1. INTRODUCTION
In the Spring of 2008 I created the GRIP MAESTRO (hereafter "GM") as part of my ongoing research in physical interactive design (PID) at Stanford University's Center for Computer Research in Music and Acoustics (CCRMA)[1]. The GM is a sensor-augmented hand-exerciser (the "Grip Master" by Prohands™[2]). It is connected via an Arduino Nano[3] microcontroller board to a computer where the data retrieved by the unit's sensors is adjusted, smoothed, extrapolated from, refined, and then finally translated into musical sound.

My approach to PID revolves around a few central precepts:

1. A more meaningful performance is one where the audience can clearly hear *and see* the work being done by the performer, and track the correlation between these two sensory experiences.
2. The physical connection a performer shares with his instrument (vis-à-vis innate physical feedback) is of the utmost importance not only to the establishment of the correlation stipulated in point 1 above, but also to the overall success of the performance; success both in accuracy and in emotivity.
3. The degree of correspondence between the performative work and the resulting sounds has direct bearing not only on the performer's continuation of that same work, but also on the nature of the audience's relationship to the performer, his instrument, and the performed sounds.

The third point above is an interesting one. Work in the neuroscience and psychology of music [1] has shown that when one person sees another performing a task (musical or otherwise), the observer's motor cortex (and in the case of music, the auditory cortex too) is engaged as if he were also performing the task (to a degree). This is increasingly the case when the task is one with which observer is already familiar or skilled. The degree to which an audience member is engaged while viewing a performance is in a large and subtle part due to this phenomenon. In a performance where the performer's work is hidden (by a computer screen, say) and in which the sounds resulting from his actions are unfamiliar, the audience member will have very little to hold on to. Although the science behind these sensory-neurological connections is well beyond the scope of this paper, the existence of such connections pervades my thoughts on matters of musically performative PID, and is extremely pertinent here.

1.1 Definitions
1. "Meaningful" and "Effective" — These are largely synonyms here; both refer to the performance's potential to connect with the performer(s) and the audience, where both are engaged with the performance such that they are not lost or bored. For an audience this can almost evoke a kind of mental participation in the performance itself [1].
2. "Performative" — This refers to those activities that are *intended* as performance.
3. "Emotive" — an event or action that conveys emotion. This is not the case when an observer simply infers an emotion from a performance; it requires that the emotion is present in the conveyor to begin with (i.e. that it is a true conveyance).
4. "Gesture" — a motion or group of motions in time, be they physical, sonic, or other, that accentuate an interpretive gestalt [2]. With regard to physical motion, these could be anything from miming the opening of a soup can to a seemingly random series of jerks of the hand and arm. What matters is that the observed/performed phenomena have perceptual boundaries due to the human scale at which they take place. These may be contextual as in the case of miming, or limited only by the human body of the performer as with the 'random' jerks. Such gestures are very often delineated by the limits of short-term memory (seven +/- two seconds) but may occur at many different cognitive levels (i.e. atoms, groups, groups of groups, and so on) [3].

[1] http://ccrma.stanford.edu/

[2] http://www.prohands.net/

[3] http://www.arduino.cc/

5. "Idiomatic" and "Natural" — No sound that comes from the GM, other than those that it produces acoustically, is truly idiomatic or natural, but my goal when mapping performative physical gestures and interactions to computer-generated sounds is to do so in such a way as to convince the observer that they might as well be. The perceptual distance between the action and the sound is ideally less than that of the just noticeable difference.

1.2 Prior Work

There has been much notable work both in hand-controller electroacoustic performance interfaces, and in gestural control in general. A detailed discussion of this history is beyond the scope of this paper, but what should be noted is how the GM differs from some key past work.

Most notable perhaps among past interfaces, is Michel Waisvisz' "The Hands" developed at the Studio for Electro-Instrumental Music (STEIM)[4] in Amsterdam (1984-2006). "The Hands," were created to "allow [Waisvisz] to 'touch' sound. [He] wanted to operate, navigate, compose, mold and play sound in a sensible, refined and even sensual and groovy way!" [4] The interface is something of a grab-bag of buttons and sensors that allow for the triggering and control over sounds in real-time. The design of this instrument is less about capturing a specific gesture and using it musically, than it is about shortening the 'distance' from Waisvisz' own creative brain to the produced sound [5].

Laetitia Sonami's work on glove interfaces, specifically the "Lady Glove" (1991-94 - present) is also of significant note. It is not so much an instrument (in the same way that a trumpet is an instrument) as it is a system for motion capture. Equipped with "five micro-switches, four Hall effect transducers, pressure pad, resistive strips… two ultrasonic receivers… a mercury switch on the top of the hand and an accelerometer which measures the speed of motion of the hand," [6] this controller is decked out in measuring apparatuses. The instrument in "Lady Glove" is actually the performer's hand (and body); the glove itself is but an ingenious collection of sensors. Both "Lady Glove" and "The Hands" translate their sensor data into MIDI messages before music is mapped to it.

With the GM, not only is sensor data not filtered by MIDI standards, but more importantly, it has been designed to *be* an instrument in a more traditional sense. Its specific physical idiosyncrasies demand a certain type of interaction from the performer. The intended sound of the GM is not of the sounds of the whim or of the motions of the hand and body; rather it is the sound of the GM itself[5]. It is the performer's interaction with the device that is producing the sound, not the motion of his hand or of the unit individually. While this is technically true with the examples above, it is not their focus.

What might further separate the GM from its predecessors is its inherent simplicity. It is designed to measure only two parameters of its interactions with the performer, the resisted gripping of the device, and its motion in space relative to the body of the performer.

[4] http://www.steim.org/steim/

[5] "sound of the GM itself" — I am referring to the sound of the instrument as an abstracted ideal (made up of both hardware and software).

1.3 Origin/Design

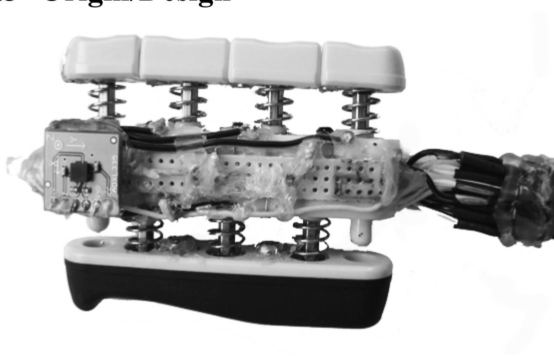

Figure 1 ~ Photograph of GRIP MAESTRO unit

Instead of designing with a specific sound or sonic metaphor in mind, I wanted to take advantage of a specific physical gesture for its emotive potential, with little forethought given to the sounds it would eventually produce. I chose the gesture/act of gripping with the hand because of its general familiarity and under-utilization in musical instrument design hitherto. The extant Prohands™ "Grip Master" is a terrific device to build on because of the variety of gripping motions that it offers in conjunction with passive haptic feedback[6]. The hand exerciser has four spring-resistant finger pads, and a single three-spring palm-pad attached to a centerpiece, so that each finger may squeeze independently and that the palm may squeeze evenly, or more or less towards the top or bottom of the hand. The model used for the development of the GM, the "Extra Light Tension" with 3lbs of resistance per spring, offers the least resistance of any "Grip Master" model, for ease of use in a musical context.

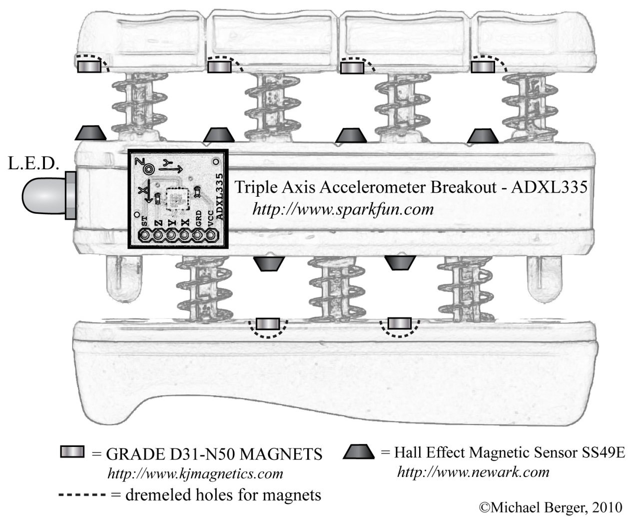

Figure 2 ~ GRIP MAESTRO Sensor Layout Diagram

Attached near each finger-pad spring and between each spring on the palm-pad is a small yet powerful magnet (grade D31-N50 magnets) opposite a Hall effect magnetic sensor. The data that each sensor sends out is directly related to its distance from its corresponding magnet, and thereby indicates the amount that

[6] "passive haptic feedback" — the "Grip Master" incorporates weighted springs into its design, thereby providing an innate resistance to motion. *Active* haptic feedback would be that which is assisted, augmented, or entirely created by mechanical and/or electrical means.

each pad is depressed. The palm-pad's two sensor/magnet pairs are arranged so that its 'tilt' maybe calculated by the computer; that is the angle of the palm-pad relative to the centerpiece. An average of the two palm-pad sensors is taken to approximate the amount that the performer is squeezing the whole unit. An ADXL335 triple-axis accelerometer attached near the top of the front face of the unit detects its orientation to gravity and relative motion in three-dimensional space. A light emitting diode (L.E.D.) is attached at the top of the unit to indicate its operational state (*figure 1 and figure 2 above*).

The circuit design is very simple (*figure 3 below*). Each Hall effect sensor and the accelerometer are powered in parallel by 5 volts directly from the Arduino microcontroller. Their internal mechanisms provide resistance to this voltage potential so that no additional resistors are required. All of the sensors share ground and each has a single connector to convey its readouts to the Arduino.

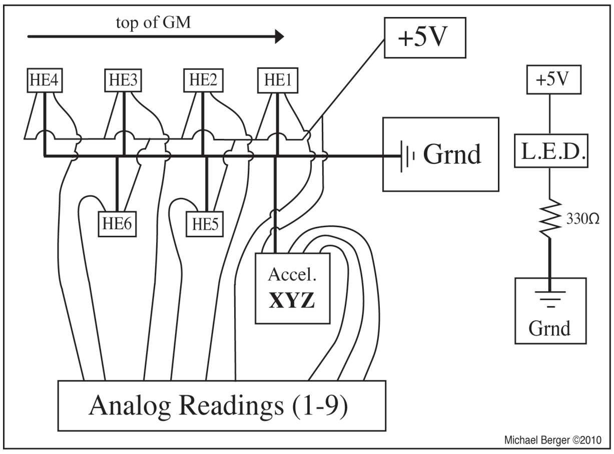

Figure 3 ~ GRIP MAESTRO Circuit Design

2. MAPPING

Since the GM is modeled after a gestural interaction and not after a desired sound (or method of sound production), finding/developing a musical sound that works idiomatically for the instrument is the fundamental challenge it poses. An ideal software mapping will produce sounds that, to an observer and to the performer, seem to come *naturally* from the device and the gestures its performance requires. This is not to say that there is only one ideal mapping for the GM; on the contrary, the multidimensionality of the challenge ensures the possibility of many such ideals. It is this wealth of options that is simultaneously the instrument's best strength and its biggest potential hindrance.

This challenge becomes one of balancing not only the large- and small-scale aspects of the performative gestures, but also of the analogous aspects of the musical sounds the gestures seek to create (all with a mind to the physical idiosyncrasies of the instrument itself). A good mapping of the GM to sound is one that is developed dialectically from the two approaches, i.e. asking oneself the questions: 'what gestures do I want to perform?' versus: 'what sounds do I want to produce?'

2.1 Interpreting the Data

The sensor data coming from the Hall effect sensors (after calibration and scaling — *see section 2.2 and figure 6 below*) gives a clear indication of the depression of the finger-pads on the GM. This data is perceptually very close to the performer's feel of the interaction, and so it is already in very good shape for sound mapping. One may easily consider the data from each finger pad separately or build relationships between them for more complicated interactions that retain this perceptual coherence.

The data given by the accelerometer however, is more perceptually distant from the actions a performer might take with it. It senses the amount of force acted upon it (either due to gravity or physical acceleration) in each of three dimensions (X, Y, and Z). To effectively interpret this incoming data is to shorten its perceptual distance from the performer's experience, i.e. one must extrapolate elements of motion in human terms from these objective data points. References to the human experience in space like, which direction is "up," and which direction is "in," become very important. Limitations of the body itself also play key roles in this extrapolation; a GM in free space might be able to rotate in one direction indefinitely, but a human arm cannot.

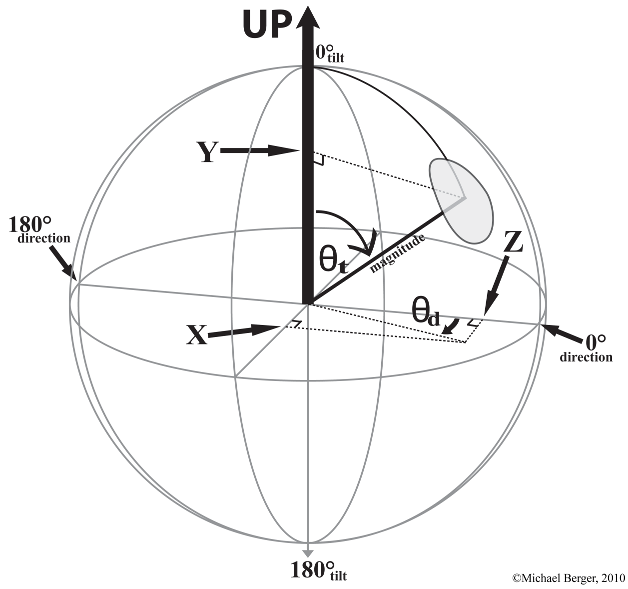

Figure 4 ~ Accelerometer Extrapolation

Figure 4 (*above*) is a diagram of three-dimensional space from the point of view of the GM. The grey circle on the surface of the sphere represents the information coming from the accelerometer (a single point). The figure shows how this data is extrapolated into three main human-scaled elements of motion. These are the magnitude of the displacement (i.e. how much force is being applied to the GM — the sphere's radius), the angle of tilt (θ_t — relative to "up"), and the direction of this tilt (θ_d — starting with directly "inwards" marked as 0° direction). Tilting the GM backwards gives a negative direction angle and tilting it forwards gives a positive direction angle, both from 0° to 180° (θ_d). Tilting the GM downwards will register an angle of tilt (θ_t) from 0° to 180° and back up to 0° if the motion is continued.

By projecting the extrapolated angles from figure 4 onto the X- and Z-axes (relative to the hand of the performer), one can measure the amount of tilt "inwards" vs. "outwards," and "backwards" vs. "forwards" of the instrument (*see figure 5 below*). The data from the accelerometer is now more closely aligned with concepts of motion from a human perspective (i.e. gestalts of motion/space). Gestures of the hand that are intuitive for the performer due to their perceptual boundaries are now immediately evident to the computer.

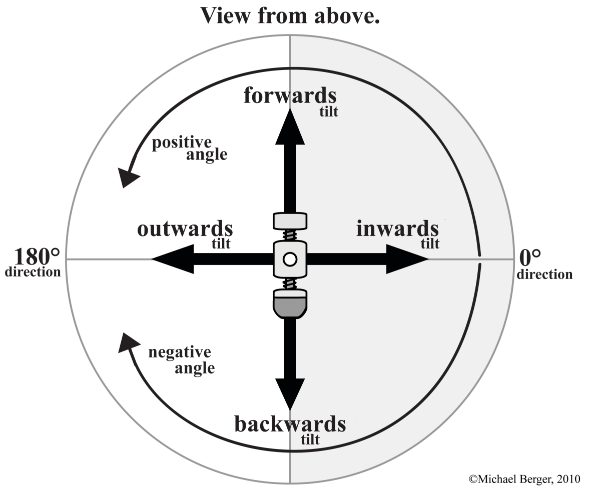

Figure 5 ~ Projections on to "Forwards/Backwards" and "Inwards/Outwards"

2.2 Large- VS. Small-scale

Some motions or gestures, be they with the GM in space (via the accelerometer) or with the fingers and palm on the pads of the GM, are more evident than others to an observer. These differences of scale have a different feeling for the performer as well and carry different emotional contexts. Paying careful attention to balance of scales in both the controlling gestures and in the mapped musical sounds is crucial to a successful mapping.

All gestures have a quality of performative energy about them, and it is the perceived quantity of this energy that one speaks of when one speaks of the *size* of gestures. Often times the perceived energy level is highly correlated with the visibility of the performed gesture or the audibility of the produced sound.

One way to exaggerate the variability of gesture scale is to artificially "bend" the incoming data before it is translated into sound (similar to 'compression' in audio). Figure 6 (*below*) depicts four arbitrary scaling curves. The X-axis represents the energy input (or sensor reading input), and the Y-axis the output (or effect on the sound by the input control).

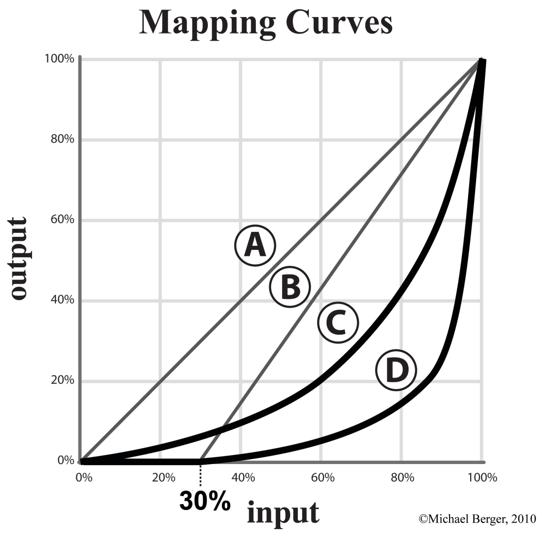

Figure 6 ~ "Bent" Mapping Curves

Curve "A" depicts a linear one-to-one ratio. Curve "B," while also linear, has a region (the first 30% of input) for which no output is generated. This can be understood as a buffer or safe zone. If this mapping were applied to the tilt angle from figure 5, the performer could safely hold his hand 'up' +/- 30% without producing any sounding changes. Paradoxically this allows him *more* control over the sound than curve "A" because it is easier to *not affect* the output. Curves "C" and "D" show exponential scalings that produce very little change at lower inputs and very great change at higher inputs. Curve "D" has the added bonus of a 30% buffer.

2.2.1 Examples of Musical Mappings

Music too, has a continuum of aspects from the micro- to the macro-scopic (eg. from amplitude, pitch, and colour, to phrase, meter, and tempo, respectively). These aspects of musical sound map more naturally onto the analogous scales of gesture.

With the mapping, "AUDIO DESTROY!" (2010), I created a "DJ"-like instrument for live re-mixing of an arbitrary audio track; I tend towards anything by James Brown. The tilt of the performer's hand forwards increases the playback rate of the track, while tilting backwards reduces it. Tilting inwards gives a positive playback direction, outwards negative — one can "scratch" the sound by quickly alternating these two. The amount of tilt downwards lowers the cut-off frequency of a low-pass filter. The average depression of the palm-pad controls the overall gain of the playback, and squeezing it upwards or downwards affects a pitch-shifting algorithm correspondingly. Finally, the finger pad whose depression is the greatest affects the rate and depth of a chorus effect. Assigning only a single function to all four of the finger pads emphasizes and allows for greater flexibility with the palm pad controls. These controls enable a wide range of musical transformations that are also highly repeatable. This facilitates a dynamic richness in both pre-composed and live improvisatory music together with a high level of emotivity and individuality from the performer.

3. CONCLUSIONS

Developing musical mappings using the dialectic approach posited here provides a firm foundation on which very complicated yet intuitive relationships between physical performative actions and musical sounds may be built. The next stage of my research with GM is to conduct a user study from the pool of electroacoustic performers available at CCRMA, with a mind to the questions raised herein.

4. REFERENCES

[1] Peretz, I. and R. J. Zatorre. "Brain Organization for Music Processing." *Annual Review of Psychology,* Annual Reviews, Palo Alto, CA: Vol. 56 (February 2005), 89-114.

[2] Tenney, James and Larry Polansky. "Gestalt Perception in Music." *Journal of Music Theory*, Duke University Press, New Haven, CT: Vol. 24, No. 2 (Autumn, 1980), 205-241.

[3] Snyder, Bob. *Music and Memory: An Introduction.* Cambridge, MA: The MIT Press, 2000.

[4] http://www.crackle.org/TheHands.htm, accessed on February 5, 2010

[5] Krefeld, Volker and Michel Waisvisz. "The Hand in the Web: An Interview with Michel Waisvisz." *Computer Music Journal*, The MIT Press, Cambridge, MA: Vol. 14, No. 2 (Summer, 1990), 28-33.

[6] http://www.sonami.net/lady_glove2.htm, accessed on February 5, 2010