Remutualizing the Instrument: Co-Design of Synthesis Algorithms and Controllers

Perry R. Cook

Computer Science Dept. (also jointly in Music)
Princeton University, Princeton, NJ, USA
prc@cs.princeton.edu

ABSTRACT

From the advent of electronic music, and even from early organ consoles and other remote manipulated instruments, much of the design and research of new musical interfaces has focused on abstracting the "controller" from the "synthesizer" and then investigating how to best interface those two classes of hardware with each other and the player. Yet, many of the striking lessons from our history of intimate expressive musical instruments lie in the blurred boundaries between player, controller, and sound producing object. Bowed strings, winds, and certainly the human voice all blur these boundaries, both in the design and construction of the "instrument" and in the resulting controls and expressions.

This paper looks at some of the issues involved with creating new expressive electronic musical instruments, and presents a number of recent projects in the co-design of musical controllers and computer sound synthesis algorithms. Specific cases are described where the traditional engineering approach of building a controller (a box), and connecting it to a synthesizer (another box) would never have yielded the final product that resulted from the tightly coupled development of a complete musical system all at the same time. Examples are given of where a discovery from synthesis algorithm development suggested a new control metaphor, and where a control component suggested a new aspect of synthesis.

1. INTRODUCTION

On first reflection, one might assume that the notion of a musical "controller" has a relatively recent history, dating perhaps only since the advent of MIDI and digital controllers and synthesizers. But the controller/generator metaphor has a longer history than might be expected. By necessity of it's size and other factors, the organ console is likely the earliest of common instruments to have separated the controller (the console) from the tone-generator (the pipes). The earliest specimen of a hydralis (water powered) organ dates to before 300 BC [1]. By the late 1700's, organs were growing in size, were finding they way into more churches, and builders were devising new means to physically separate the console and pipes. Tracker mechanisms, pneumatically assisted valves, and eventually electronic switches and valves all increased the distance between player and sound production [2].

In the industrial era of the early 1900's, various player, reproducing, and recording pianos possessed features such as speed and volume controls, and some even had remote controls for those functions (a flexible cable with handle that, when turned, changed a motor speed, or a handle/cable that opened and closed shutter doors for muting the sound). Many Victrolas and cylinder players also possessed remote speed and volume controls.

Of course the notion of transforming music listening into an afferent act, with a human issuing commands to an instrument (playing machine), the machine responding with a "perfect" performance, and with the auditory channel being the only mechanism of feedback, has a distinct bias. The notion that everyone in the world wants to be an "armchair conductor" is a well-held tenet of modern times, but it has little to do with the construction, playing, and personal enjoyment of most musical instruments.

The major flaw in the controller/synthesizer paradigm is the loss of intimacy between human player and instrument. I pose three primary reasons (lacks) for this intimacy loss:

- Lack of haptic feedback from the controller/instrument to the player. Haptic (combined senses of touch, including skin vibration and pressure, and the muscle senses of motion, position, and force) feedback has been increasingly addressed in musical interface research projects. Commercially, the most successful haptic systems are electronic keyboards that copy (passively, through weights and levers) the feel of piano keys.

- Lack of fidelity in the connections from the controller/sensor to the generator, primarily delays and distortions in response to gestures. "Distortion" here refers to any response that doesn't meet some usual, learnable, or repeatable expectation.

- Lack of any sense that sound comes from the instrument (the controller) itself. More generally, this is a subset of a larger feeling that no meaningful physics goes on in the controller. Trends toward larger concert venues, greater amplification, and larger loudspeakers have consistently worked to diminish the importance of the actual acoustical instrument sound. The aesthetic toll of this has been great, most importantly for the player and composer, and shows profoundly in the musical results.

Figure 1 shows a block diagram of a typical electronic music controller system. Obvious potential flaws in a system of this type include the one-way flow (lack of feedback to the player) from controller to synthesizer to speaker, and the potential latencies and other distortions that might result from the sensors, quantization, transmission, etc.

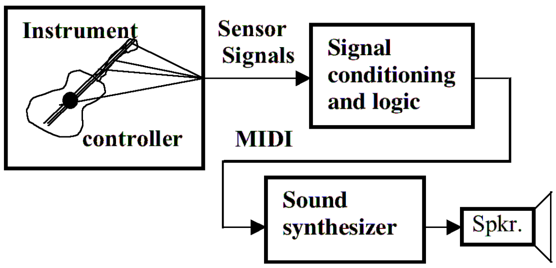

Figure 1: *Standard controller/synthesizer system*

Figure 2 shows a block diagram of a human performer connected to a musical instrument, specifically noting the multiple auditory and haptic feedback paths. Degradation in the playing "feel," (and in any resulting music) can result from distortion of any path in either system of Figure 1 or 2.

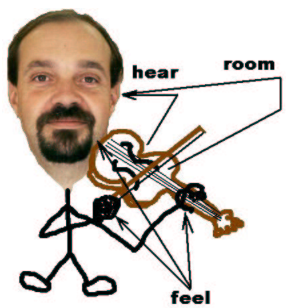

Figure 2: *Traditional acoustic player/instrument system*.

Most work in musical interfaces, protocols, and systems has worked to improve the fidelity and flexibility of the forward control channels (Figure 1), by improving the sensors, connection speeds, signal processing algorithms, computer hardware, and synthesizer quality. Much less work has been done on the feedback (Figure 2) channels. The next sections will describe a number of recent projects that attempt to address the "lacks" listed earlier, by completing and improving both the feedforward and feedback paths.

2. CLOSED-LOOP MULTI-MODAL SYNTHESIS: THE HAPTIC MARACA

Based on the Physically-Informed Stochastic Event synthesis Model (PhISEM) [3], a number of controllers containing accelerometers, force-sensing resistors, and switches were developed to translate the gestures of shaking and scraping into parameters for the particle synthesis model. Early controllers (Figure 3) were selected for their relation to the origins of the computational model. Some controllers were constructed to give a natural "feel" of the sound/instrument being controlled. For example, a ratchet device with multi-position rotary switch can feel like the ratchet sound it controls. This is an example of "passive haptics," where no explicit forces are computed and fed back to motors in the controller, but rather the basic feel of the controller is matched as closely as possible to the sound (or the sound is matched to the feel). This is the type of haptic interface provided in commercial "weighted" electronic pianos.

In constructing the early controllers shown in Figure 3, it was noted that the synthesis algorithm automatically provides useful information about the sound production physics. In PhISEM, each sound producing particle collision is computed. The Haptic Maraca was an early attempt at mutualizing synthesis and control, by feeding the PhISEM particle collision impulses to a small solenoid motor (mass glued to a speaker coil) located in the maraca gourd. Figure 4 shows the PhISEM haptic maraca controller, with the Analog Devices 2181 DSP board used to compute the model. Using the DSP board afforded essentially no delays between control gestures and audio/haptic responses. Even though the system had wires attached, testers reported an uncanny feeling of connection between gesture, sound, and feel.

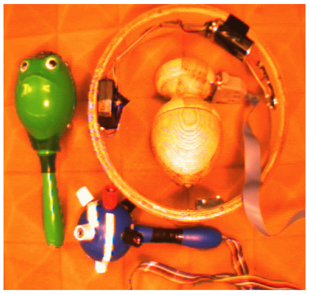

Figure 3: *Early PhISEM synthesis controllers.*

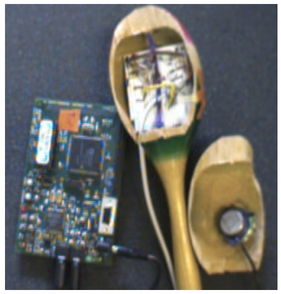

Figure 4: *The PhISEM haptic maraca controller.*

3. A CRAZY IDEA: SOUND FROM THE ACTUAL INSTRUMENT

3.1 The Bowed Sensor Speaker Array (BoSSA)

The Nbody project [4] was inspired by a desire to make new musical instruments that engage the performers, the performance space, and the audience more effectively than the forward-facing stage speakers most often used in computer music concerts. It was also desired to extend the work of Weinreich [5], Causse [6], and others in the directional sound radiation of acoustic instruments, but to collect and freely distribute a large number of radiation transfer functions for a variety of instruments. Multi-directional (12–72) impulse responses were collected from six different stringed instruments. These were studied and used to construct various simulated and virtual performance instruments.

The Bowed Sensor Speaker Array (BoSSA) [7] was the first project to fully integrate a spherical radiating speaker body and a complete electronic performance interface into one musical artifact (Figure 5). Inspired by the violin, BoSSA consists of a set of four bowed sponges (the Bonge) to emulate the strings, a bow with force and acceleration sensors (the Rbow), a violin fingerboard with linear force and tilt sensors (the Fangerbored), all mounted to a spherical speaker with 12 drivers (the Critter). This instrument allows a new kind of intimacy for computer music performance, with the sound coming from the actual controller itself.

Figure 5: *Dan Trueman's BoSSA.*

3.2 SqueezeVoxen: "Singing" Accordions

As might be expected, attempts to create controllers for computer voice models consistently point up the same problems. From a technical standpoint, the sheer number of parameters that need to be controlled in an expressive voice model present daunting issues of sensors, bandwidth, systems, and mappings. From a musical, linguistic, and perceptual standpoint, all humans posses a voice, have years of experience "playing" it (not necessarily musically, but expressively) and attending to the voices of others. Thus humans are extremely critical of synthesized voices.

Even given the problems, a further attempt at creating a meaningful and expressive (or at least fun) interface for controlling computer voices was undertaken with Colby Leider in 2000. The SqueezeVox project [8] recognized that to control a vocal model, independent controls are needed for pitch, breathing, and phoneme articulation (spectral features). The accordion is an instrument with components that map somewhat naturally to these requirements. Melody pitch is controlled with the right hand keyboard, "breathing" is provided naturally in the bellows mechanism (though the accordion "sings" when breathing both in and out), and the left hand provides an array of buttons (10 or so in a concertina or 12-bass instrument, 100+ in a full sized instrument).

The SqueezeVox project exploited these features of the accordion to control a variety of voice models (formant models, acoustic tube models, FOF synthesizers, orchestras of chanting monks, and more). Air pressure sensors, a linear FSR located next to the keyboard, left-hand buttons and four nearby bend sensors (addressing/controlling phonemes, articulator positions, or formant positions), internal speakers, tilt sensors, and other sensors were incorporated into a variety of instruments. Bart, Lisa (shown in Figure 6), Maggie (a concertina) and Santa's Little Helper (a child's toy accordion) make up the complete fleet of SqueezeVoxen.

Figure 6: *SqueezeVox (Lisa) vocal synthesis controller.*

4. FULLY INTEGRATED PHYSICAL MODEL SYNTHESIS & CONTROL: THE NUKELELE

Research on physical models, and work at Interval Research on control projects such as the Virtual Haptic Maraca and "The Stick" [9] led to a question of just how tightly one could bind a virtual instrument to it's form, algorithm, and playing technique. The result was the Digital Koto, or the "Nukelele" (dubbed this name by Michael Brook). The main idea of the Nukelele was to attempt to create a true virtual stringed instrument, with responsiveness to all the subtleties of rubs, damps, and plucks of many types. Further, as with most projects in virtual instrument design, it was desired to extend the capabilities beyond what an actual stringed instrument can do.

The Nukelele uses a simple Karplus-Strong plucked string model, controlled by linear force sensors, to simulate the interactions of a virtual string. The left hand FSR behaves as a fretless fingerboard, and can also sense hammer-on and hammer-off gestures. The right hand "string" is pressure sensitive (to sense damping), position sensitive (for pluck position), and also sends an audio signal into the plucked string model in response to striking and rubbing. This audio signal is the key to the intimacy of the interface, where even the slightest scratching or rubbing causes the virtual string to oscillate. Figure 7 shows a block diagram of the Nukelele synthesis system, and Figure 8 shows the original Nukelele.

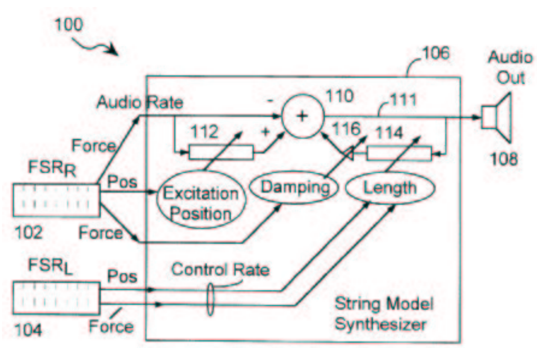

Figure 7: *The Nukelele integrates the components of a physical stringed instrument. The key is in the audio-rate excitation force signal fed into the string, allowing plucking, striking, and rubbing, in addition to position and damping.*

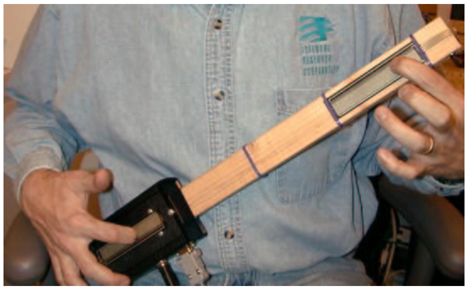

Figure 8: *The original Nukelele instrument.*

A 2nd generation Nukelele, called the Nukelele'elua (Figure 9) was constructed, adding a longer fingerboard FSR, and two speakers in the body of a 1/2 size guitar. The right hand FSR is situated in a "strummer" orientation, allowing multiple virtual strings to be plucked. One speaker faces outward, and the other back toward the player. Different filter functions are used to drive these two speakers, based on impulse responses collected in the NBody Project [4].

Figure 9: *The Nukelele'elua (Nukelele-2).*

5. CONCLUSIONS

With some effort, new electronic instruments can be constructed that extend our notions of acoustical instruments, but keep a sense of intimacy, connectedness, and embodiment for the player and audience. Other experiments with self-contained sensor/speaker systems [10] have yielded provocative and rewarding artistic results, and the author looks forward to building more of such systems in the future.

6. ACKNOWLEDGEMENTS

Thanks to Bob Adams and the Interval Research Expressions group for many inspiring and creative "work-jams." Thanks to Interval Research, specifically Bill Verplank and Geoff Smith, for making the artifacts of my work there available to me after Interval's demise. Thanks to Dan Trueman for great collaborations on Nbody, Rbow, BoSSA, and for good discussions on "that most perfect of musical instruments." Thanks to Colby Leider for cooking up and carrying out the SqueezeVox project with me.

3. REFERENCES

[1] www.organisten.de/oberlinger/agonline/200207.htm, The Hydralis of Dion, D. Panternalis, AGO Online, July 2002.

[2] Snyder, K. ed. "The Organ as a Mirror of Its Time: North European Reflections, 1610-2000," Oxford University Press, 2002.

[3] Cook, P. "Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds," Computer Music Journal, 21:3, 1997.

[4] Cook, P. and Trueman, D. "Spherical Radiation from Stringed Instruments: Measured, Modeled, and Reproduced," Journal of the Catgut Acoustical Society, November 1999.

[5] Weinreich, G. "Sound Radiation from the Violin – As We Know It Today," Journal of the Catgut Acoustical Society, May, 2002.

[6] Caussé, R., Bresciani, J.F., and Warusfel, O. "Radiation of Musical Instruments and Control of Reproduction with Loudspeakers," Proceedings of the International Symposium of Music Acoustics, August 1992.

[7] Trueman, D. and Cook, P. "BoSSA: The Deconstructed Violin Reconstructed," Journal of New Music Research, Fall, 2000.

[8] Cook, P. and Leider, C. "SqueezeVox: A New Controller for Vocal Synthesis Models," Proceedings of the International Computer Music Conference, 2000.

[9] Levitin, D., McAdams, S. and Adams, R. "Control Parameters for Musical Instruments: A Foundation for New Mappings of Gesture to Sound," Organised Sound, 7(2), pp. 171-189, 2002.

[10] Trueman, D., Bahn, C. and Cook, P. "Alternative Voices for Electronic Sound: Spherical Speakers and Sensor-Speaker Arrays (SenSAs)," Proceedings of the International Computer Music Conference, August 2000.