Physical Models for Music Synthesis, and a Meta-Controller for Real Time Performance

Perry R. Cook

Stanford CCRMA, Stanford, CA, 94305
PRC@CCRMA.STANFORD.EDU

ABSTRACT

The time is approaching when music can be synthesized by direct physical models of the acoustic systems of musical instruments. New understanding of instrument acoustics and digital signal processing, combined with increasing computer power, now allows such instruments to run in real-time under the control of a human player. This paper discusses some techniques for efficiently simulating physical systems in digital computers, and new instruments and controllers based on physical paradigms.

A waveguide instrument has been constructed which simulates the main acoustical mechanisms of the wind instrument family, allowing one instrument to function as flute, recorder, clarinet, saxophone, trumpet, trombone, and other instruments which are hybrids of the various wind instrument sub-families. Multiple voices of the instrument can be run in real-time on a Motorola DSP 56001 signal processing chip, under control from a NeXT scorefile, mouse control with graphical feedback on the NeXT computer screen, or via MIDI. The instrument is controlled by parameters affecting reed stiffness, mass-spring-damper characteristics of the reed/lip oscillator, length of the bore and of secondary delay paths, state of tone/register holes, jet length, and breath pressure. Lower level control allows direct access to control coefficients, allowing hybrid instrument settings to be defined.

A special MIDI controller has been constructed for controlling physical model instruments as well as more conventional MIDI synthesizer instruments. The controller allows a musician to play in the paradigms of most wind instruments, but provides all of these modes and others simultaneously. Controls include linear slide control like that of the trombone, valve control like that of other brass instruments, fingering control like that of the woodwinds, breath pressure, lip tension, bite pressure, and rotational position (similar to rotating the flute or head to change embouchure). A new rotational control is mounted in the mouthpiece which varies playing position from recorder-like (end blown) to flute-like (edge blown). The controller allows a performer to effectively use much more of the real time control bandwidth available in the mouth and hands.

1. INTRODUCTION: PHYSICAL MODELS FOR MUSIC SYNTHESIS

Synthesis of musical sounds by physical modeling is becoming more feasible due to two components: computing equipment is growing increasingly powerful and inexpensive, and more efficient algorithms for computing the physical solutions are emerging. The most popularly used physical model is the Karplus-Strong plucked string algorithm[Karplus87], which came about as more of an accident rather than an intentional act of physical modeling. When doing wavetable synthesis using white noise as the table contents, Karplus and Strong discovered that when the wavetable contents were attenuated by a small amount each time the wavetable was read, a plucked string timbre resulted. Julius Smith helped to explain the physics of this, and Jaffe and Smith [Jaffe89] added refinements to the basic model, allowing simulation of pick position, string stiffness, etc. The technique known as WaveGuide Synthesis[Smith87] has given rise to a family of instrument models, including clarinet [Hirschman91] and trombone [CookICMC91]. Figure 1 shows block diagrams of the original Karplus-Strong string simulation, the same model with a rearrangement of the components to yield a more physical model, and a general physical model instrument. Recently, Sullivan [Sullivan90] has brought the plucked string instrument into the age of rock and roll, with refinements to model tube amplifier distortion, room reverberation, and feedback. Rick VanderKam at CCRMA has demonstrated that this rather complex model is still simple enough to run in real time on a digital signal processor (DSP) chip.

While these models are the most popular physical models used today, they are by no means the first. Working with Max Mathews at Bell Laboratories in 1959, John Kelly and Carol Lochbaum constructed a digital model of the human vocal tract based on transmission line physical models [Kelly62], and synthesized a striking rendition of the song "Bicycle Built for Two" (Daisy). This early model was one starting point for the author’s dissertation work in modeling the singing voice [CookDiss91]. Figure 2 graphically shows the process of turning an acoustic tube like the vocal tract into a digital filter simulation. Figure 3 shows the control screen for SPASM (Singing Physical Articulatory Synthesis Model) constructed for the synthesis of singing on the NeXT computer.

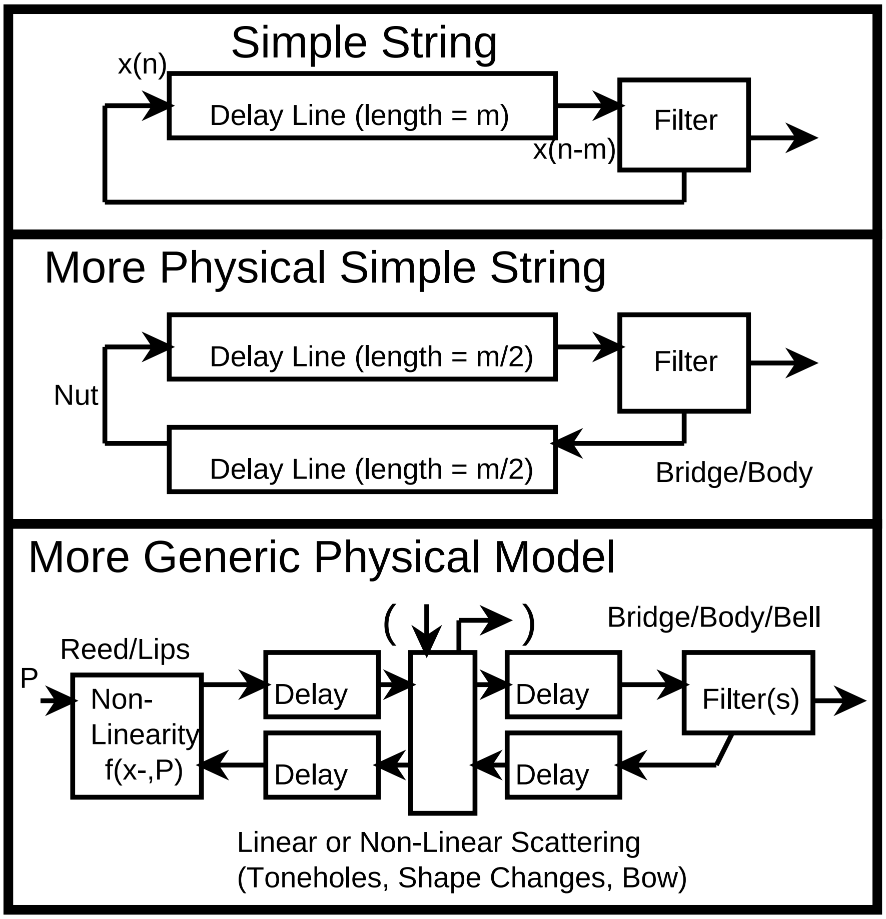


Figure 1. Simple Karplus-Strong string model (top), more physical diagram of the same model (center), and a generic physical model diagram capable of simulating stringed or wind instruments (bottom).

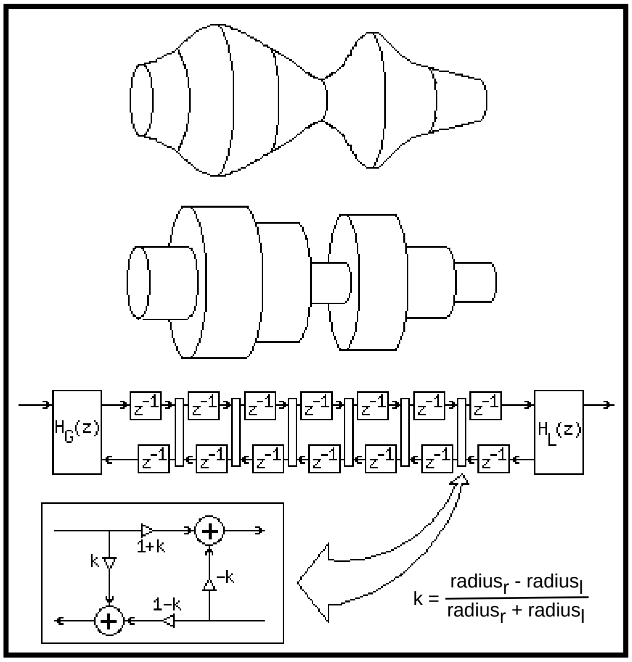
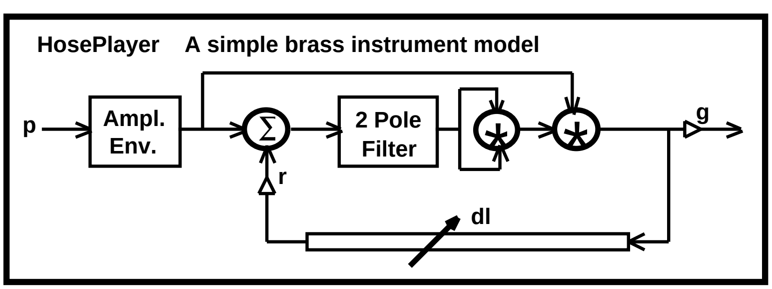


Figure 2. A smooth acoustic tube (top) can be sampled in space, similar to sampling waveforms in time. This sampled tube can be directly converted into a digital filter (bottom).

2. SIMPLE WAVEGUIDE INSTRUMENTS FOR REAL-TIME SYNTHESIS

Once the basic physics of an instrument or instrument family are understood, a computational model can be derived, and that model can be simplified for more efficient computation. Optimization often takes advantage of the idiosyncrasies of specialized DSP hardware. Figure 4 shows four simplified physical models, based loosely on the physics of brass instruments, jet-reed instruments, and cane-reed instruments. Non-linearities are modeled as polynomials for efficient DSP calculation. In the brass instrument simulation, called HosePlayer, a second order resonator (two pole digital filter) is used to model the lip oscillator. In the flute simulation, called SlideFlute, an extra delay line is added to model the jet delay. In the clarinet simulator,called ClariNot, an extra delay line is used to coarsely model the effects of a register hole. All three instruments are bundled together into one meta-wind instrument called WhirlWind. These models are described in detail in [Cook92].

Figure 3. Singing Physical Articulatory Synthesis Model (SPASM) main control screen.



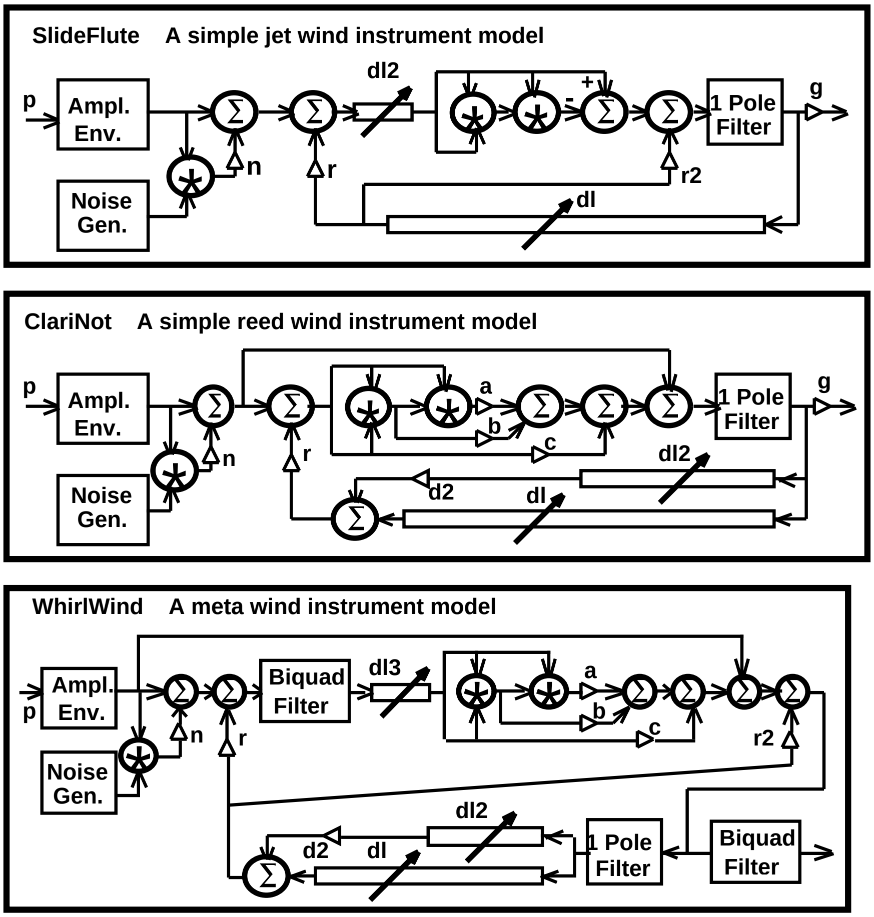


Figure 4. Simple physical instrument models for DSP implementation.

3. HIRN: A NEW REAL-TIME SYNTHESIZER CONTROLLER

Based on the premise that wind players have much unused control bandwidth available while playing their instruments, a Meta-Wind Instrument Controller was designed and constructed. Called HIRN, this controller exploits the common control modes found in the wind instrument family, and adds many more degrees of control freedom. Figure 5 shows the HIRN Meta-Wind Instrument Controller. Signals detected in the instrument mouthpiece include breath pressure, bite tension, lip tension as estimated by measuring the myoelectric activity of the upper lip [Knapp90], and a pitch detector so the player can sing or buzz the lips directly into the instrument mouthpiece. Fingering control is provided via 8 buttons, controlled by four fingers on the right hand, and three fingers and thumb on the left hand. The right hand can be slid linearly along the axis of the instrument, as well as radially rotated. The left hand can be rotated radially. Finally, the head of the instrument can be rotated to vary the playing style from end blown as in a clarinet or soprano saxophone, to edge blown as in a flute. Optional continuous and switch foot controllers can be added to the HIRN controls. Suggested mappings of HIRN control features to standard MIDI and WhirlWind physical model control are shown in Table 1.

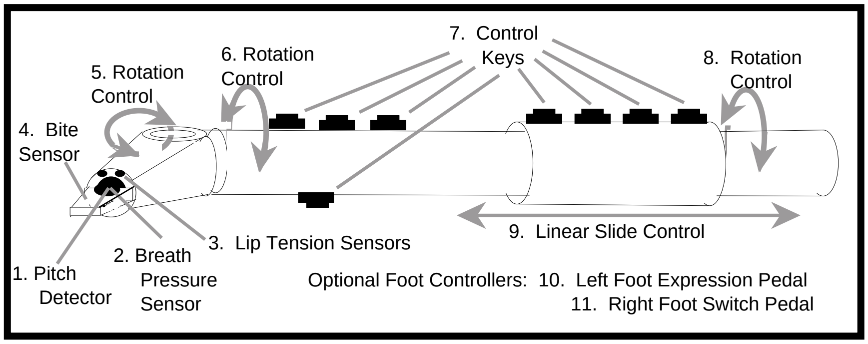


Figure 5. The HIRN meta-wind instrument controller.

Table 1. Suggested Mappings of HIRN Controls for Standard MIDI and WhirlWind Control

Control type   Standard MIDI Control Function   WhirlWind Control Function

1. Mouthpiece Pitch Detector: Note # with pitch bend   N/A
2. Breath Pressure Sensor: Aftertouch / Breath Controller Breath Pressure
3. Lip Tension Sensor: Note # (with 7.) (Brass Mode) Lip Tension
4. Bite Pressure Sensor: Note # (with 7.) (Wood Mode) Reed/Lip polynomial
5. Head Rotation Control: Cross Fading MIDI Volume   Feed-Forward Delay Line
   Chs.1<-->2 and 3<-->4   (Clar. to Flute)
6. Left Hand Rotation Control: Cross Fading MIDI Volume
   Flute Embouchure Control
   Chs.1&2<-->3&4
7. Control Keys: Note # Select: Two implemented
   Penny Whistle Fingering (Wood)   as Register Holes
   Three Keys as Valves (Brass)
8. Linear Slide Control: Pitch Bend   Delay Line Length
9. Right Hand Rotation Control: Modulation Wheel   Noise Volume
10. Left Foot Expression Pedal: Global Volume Chs. 1, 2, 3, & 4
   Output Volume
11. Right Foot Switch Pedal: Sustain   Sustain (Breath Pressure)

4. REFERENCES

Cook,P. R. 1991 ‘‘TBone: An Interactive WaveGuide Brass Instrument Synthesis Workbench for the NeXT Machine,’’ Proceedings of the ICMC, Montreal, 297 -299.

Cook, P. R. 1991 "Identification of Control Parameters in an Articulatory Vocal Tract Model, With Applications to the Synthesis of Singing," Elec. Engr. PhD Dissertation, Stanford.

Cook,P. R. 1992 ‘‘A Meta-Wind-Instrument Physical Model, and a Meta-Controller for Real Time Performance Control,’’ Proceedings of the ICMC, San Jose, CA.

Hirschman, S. E., P. R. Cook and J. O. Smith 1991 "Digital Waveguide Modelling of Reed Woodwinds: An Interactive Development,’’ Proceedings of the ICMC, Montreal, 300 -303.

Jaffe, D. A. and J. O. Smith 1983 "Extensions of the Karplus-Strong Plucked String Algorithm," Comp. Music Journal, 7:2, 56-69.

Karplus, K. and A. Strong 1983 "Digital Synthesis of Plucked String and Drum Timbres," Comp. Music Journal, 7: 2, 43-55.

Kelly, J. L. and C. C. Lochbaum 1962 "Speech Synthesis," Proc. Fourth International Congr. on Acoust. Paper G42: 1-4.

Knapp, R.B. and H. S. Lusted 1990, "Bioelectric Controller for Computer Music Applications," Comp. Music Journal, 14: 1, 42-47.

Smith, J. O. 1987, ‘‘Musical Applications of Digital Waveguides." Stanford University Center For Computer Research in Music and Acoustics, Report No. STAN-M-39.

Sullivan, C. R. 1991 "Extending the Karplus-Strong Algorithm to Synthesize Electric Guitar Timbres With Distortion and Feedback," Comp. Music Journal, 14:3, 26-37.