Physical Modeling Synthesis Update*

Julius O. Smith III

Center for Computer Research in Music and Acoustics (CCRMA)
Department of Music, Stanford University, Stanford, California 94305 USA

Abstract

Recent research in physical modeling of musical instruments for purposes of sound synthesis is reviewed. Recent references, results, and outstanding problems are highlighted for models of strings, winds, brasses, percussion, and acoustic spaces. Emphasis is placed on digital waveguide models and the musical acoustics research on which they are based.

Introduction

A musical instrument should be "alive" in the hands of the performer. A performance is naturally an *interaction* between the player and the instrument. While the main attributes of each note are predictable from a score, for example, many subtle qualities are not, contributing to the delight of the player and audience. Recently, commercial music synthesizers have been progressing toward more interactive, model-based instruments, and there seems to be growing interest in them among performing musicians. A new breed of "virtual acoustic" synthesizer is now on the market, and there has been a resurgence of interest in "virtual analog" synthesis. One can even hear former leaders of the development of wavetable ("sampling") synthesizers claiming that, for the future, "ROM is dead."

The principal source of "life" in most acoustic instruments (aside from the performer) is *resonance* of one kind or another. For example, in a cello, the strings resonate to provide pitched notes, and the whole body resonates to provide subtle variations in the tone. Resonance gives memory and variable character to the sound. The player interacts with body resonances in unpredictable ways, sometimes reinforcing, sometimes partially canceling or beating against the accumulated resonating state.

Physical Modeling Synthesis

Physical models used in music sound synthesis are generally one of two basic types: *lumped* and *distributed*. Lumped models consist, in principle, of masses, springs, dampers, and nonlinear elements, and they can be used to approximate physical systems such as a brass player's lips, a singer's vocal folds, or a piano hammer.

When a mass and spring are connected, an elementary second-order resonator is formed. In digital audio signal processing, a second-order resonator is implemented using a two-pole *digital*

*Original version published in *The Computer Music Journal*, vol. 20, no. 2, , pp. 44–56, Summer 1996.

filter. As a result, lumped models are typically implemented using second-order digital filters as building blocks.

On the other hand, distributed model implementations typically consist of delay lines (often called "digital waveguides" when used for physical modeling), in combination with digital filters and nonlinear elements. They model wave propagation in distributed media such as strings, bores, horns, plates, and acoustic spaces. In digital waveguide models, distributed losses and dispersion are still summarized and lumped at discrete points as digital filters, separating out the pure delay-line which represents ideal propagation delay. Distributed waveguide models can be freely combined with lumped filter models; for example, a brass instrument model typically consists of a lumped model for the "lip reed" and a distributed waveguide model for the horn.

Summary

This paper provides a status report on physical modeling synthesis from the point of view of a researcher working primarily on digital waveguide models [54]. First, an overall status is proffered, followed by status reports by instrument family. In each family, some successes and shortcomings are listed, and some recent highlights in the literature are cited. For brevity, recent activity is emphasized over earlier developments. The last section briefly summarizes a few outstanding problems that seem to warrant further consideration.

Overall Status

Historically, we appear to be approaching *parity* between real and virtual acoustic instruments in the context of recorded music playback. That is, we are approaching the time when many virtual instruments can be considered interchangeable with their real-world counterparts for recording purposes. Already, sampling synthesis gives us full interchangeability for the case of a single played note—the note which was sampled. Model-based techniques, however, are beginning to provide parity over a wider variety of performance expression, and they require far less memory (though more computational power) to achieve this.

The continuum between sampling and modeling is analogous to the extremes of motion photography versus computer-generated animation in film making. Just as computer-generated graphics is finding increasing use in films, model-based musical instruments are likely to grow in importance over time as the quality/cost ratio associated with their use increases. Note that "cost" should include overall ease of use as well computational and hardware costs.

Often students will ask why we bother simulating traditional instruments when the computer is capable of generating any possible sound. Why don't we focus mainly on exciting new instruments that are light-years beyond preexisting instruments? One short answer to this question is that artificially computed sounds tend to sound artificial. In other words, we simply don't know very many ways to generate deeply communicative sounds from scratch. Another answer is that traditional musical instruments are important because they are recognizable. Communication via sound waves is generally symbolic, referring to a shared library of common experience. Traditional musical instruments provide the important vocabulary necessary to articulate musical statements in terms of the prior repertoire. Also, attempting parity with traditional musical instruments provides an excellent test of our ability to construct efficient models, since we can test by direct comparison how well the model performs.

Once the natural behavior of traditional instruments has been conveniently captured in the form of computational models, the evolution of musical instruments can transfer from "corporeal form" in terms of wood and metal, etc., to the virtual world of computational models. At this point, they can begin an accelerated evolution. For example, virtual physical instruments may be distorted in ways that would be impossible in the real world, yet because they are model-based, distortions tend to remain recognizable to the listener as "morphs" of the basic instrument. Anyone who has tweaked "parameter fifteen" of a complex FM patch knows that this "morphability" property is not typical. Often, small changes in an instrument parameter will change the sound so drastically that there is no obvious connection between the "before" and "after" sounds. It is straightforward to provide recognizable morphs using sampled instruments, but in that case it's still difficult to obtain nonlinear phenomena such as saxophone growl, or the overblowing of a flute. Again, it all comes down to shared experience: A physical model captures in concise form a wide *variety* of sounds, all recognizable as coming from a particular kind of instrument by a wide variety of people. It is a valuable resource for composers to have instrument models capable of creating such a rich collection of vivid illusions in the mind of the listener, with intuitive controls.

Some instruments have already begun their evolution in the virtual world. For example, in the attempt to imitate the piano, various "standard" electric pianos have been created, such as the "Rhodes" and "Wurlitzer" electric piano sounds. Nowadays, it is relatively rare to see a real Rhodes or Wurlitzer keyboard; instead, they have become a family of presets on various synthesizers.

It does not necessarily follow that human performers will be replaced by computational models. The best way to test a virtual acoustic instrument is with a real player connected via physical controllers to the parameter dimensions of the model. On the other hand, the development of virtual performers is also greatly facilitated by virtual instruments, and such instruments will be upward compatible with arbitrary levels of sophistication in the automated enforcement of accuracy, style, and idiom.

Status by Instrument Family

Strings

In general, strings are in great shape. In most cases, parity with real strings is possible at low cost (e.g., several voices in real time on a single processor). Strings are relatively easy to model efficiently because in the real world they are generally uniform, tightly stretched, and nearly rigidly terminated. As a result, they are highly *linear* under normal playing conditions. The digital waveguide approach to string modeling therefore works very well for strings typically used in musical instruments. In these models, the wave propagation delay along the string is implemented using an ordinary *delay line*, while damping and dispersion characteristics associated with propagation on the string are lumped into low-order *digital filters*.

The first-order effect of *nonlinearity* in strings is normally to sharpen the fundamental frequency slightly at the beginning of a hard pluck or strike, particularly on a low tension string such as a banjo string. Since variable pitch is routinely implemented for purposes of vibrato anyway, it is quite easy to add this main effect of nonlinearity to the extent it is there.

Another complicating factor for strings is the important *coupling* that exists in most instruments, e.g., in the piano [78]. At CCRMA, we have almost never seen real string measurements that do not exhibit beating or otherwise modulated decay rates due to coupling, and these effects add an important quality to the sound. As an extreme example, even a solid-body electric guitar (a 1969 Les Paul Deluxe) shows pronounced beats by the seventh partial. In this case, the coupling is primarily between the vertical and horizontal planes of vibration on a single string. In principle, the two transverse vibration planes are always a little out of tune with each other because the bridge in most instruments moves more easily in the direction normal to the body than in the horizontal direction.

To simulate both vertical and horizontal planes of transverse string vibration, two digital waveguides are needed, slightly out of tune. It is generally sufficient to implement coupling only at the bridge, although in principle they are coupled along the entire length of the string [22, see appendix]. To simulate *longitudinal* compression waves as well, which are quite audible in the low piano strings, a third waveguide is needed which is much shorter than the two transverse waveguides because compression waves generally travel much faster in strings than transverse waves. In bowed strings [40], it is argued that *torsional waves* are also important, thus adding yet a fourth digital waveguide per string.

In the piano, for key ranges in which the hammer strikes three strings simultaneously, *nine* coupled waveguides are required per key for a complete simulation (not including torsional waves); however, in a practical, high-quality, virtual piano, one waveguide per coupled string (modeling only the vertical, transverse plane) suffices quite well. It is difficult to get by with less than the correct number of strings, however, because their detuning determines the entire amplitude envelope as well as beating and aftersound effects [78]. Efficient implementation of N coupled strings is discussed in [55]. The best existing digital waveguide piano implementation appears to be a SynthBuilder patch [43] based on commuted waveguide synthesis [67]. In this technique, applicable to all linear instruments excited over a short duration, the soundboard is commuted with the string and hammer so that the soundboard model—otherwise a giant digital filter or waveguide network—can be replaced by a simple recording of its impulse response; in other words, the soundboard impulse response can be "played into the string" via a small digital filter (representing the hammer) which adjusts the brightness according to striking velocity.

The most cost-effective guitars to date also appear to be based on commuted waveguide synthesis. Matti Karjalainen's excellent flamenco guitar, played at his ICMC-93 talk in Tokyo [35], implemented six virtual strings in real time on the TI TMSC30 signal processing chip, controlled from a Common Lisp/CLOS environment. The Sondius SynthBuilder classical guitar patch has been ported to a 120 MHz Pentium where each real-time voice occupies less than two percent of the processor. The SynthBuilder distortion-guitar patch sounds authentic to most listeners, and it is possible to get six strings running in real time on a single DSP56002 clocked at 40 MHz, with room left over for effects. (A 25 MHz DSP56001 can run five strings and a flanger in real time.) The distortion-feedback model employs a saturating "virtual amplifier" whose output feeds back to the string after a propagation delay [59], and is a good example of how important nonlinear extensions become straightforward when the building blocks have physical interpretations.

The best bowed strings so far seem also to be based on commuted wageuide synthesis [32]. (In this case, the bowed string is modeled as a periodically plucked string.) Due to the great expense of implementing explicit models for resonating structures such as cello bodies or piano soundboards, commuted models, which replace the resonator model by its recorded impulse response, have an enormous cost advantage over non-commuted physical models. As a result, it's probably safe to say that *all* acoustic stringed instruments are best synthesized today using commuted waveguide models, as long as the rich resonating body is deemed important. Electric instruments, on the other hand, such as a Zeta violin or solid-body electric guitar, can be modeled as nothing but a string and a pick-up, so there's little to commute, and direct waveguide models are appropriate.

The commuting of body and string is only an exact model when the string is held steady (i.e., when the overall system is linear and time-invariant). As a result, artifacts are encountered in the simplest implementations of bowed strings during highly expressive legato playing. The solutions of these difficulties lead away from true physical modeling, but the results so far are quite promising.

There are several areas for future development in string modeling. For example, transverse and longitudinal waves are really *nonlinearly coupled* to each other [75, 25, 22]. A related phenomenon is that bridge geometry typically causes nonlinear frequency doubling [42], even though all elements meeting at the bridge, including the bridge suspension itself, may be linear. A notable exception to the general absence of nonlinear string research is the model for the Finnish Kantele [34]. Further remarks on nonlinearity are given below under "Outstanding Problems."

In the simplified "commuted synthesis" models, difficulties must be overcome to provide correct behavior during note-to-note transitions, a situation that arises not only in legato performance, but also in any melodic context other than isolated notes separated by rests. Even the most commonly used plucked string model [37, 31], arguably the easiest modeling task of all, has problems on legato transitions. Legato problems arise when a new note begins on a string that is already sounding, or when the string length is changed suddenly while sounding. The reason is that really the model itself should be changed during an excitation from that of an isolated string to that of a string with a excitation or "finger" attached. While the "pick," "finger," etc., is in contact with the string, the string is divided into two sections joined by a time-varying, damped scattering junction [51, 52]. There is also new signal energy injected in both directions on the string in superposition with the scattering (partial wave reflection and transmission). This physically accurate model of string excitation/partial-termination has been used for bowed strings [50], and it applies equally well to plucked or struck strings. It is very difficult to get high quality legato performance using only a single delay line.

A reduced-cost, approximate solution for obtaining good sounding note transitions in a basic string model was proposed in [32]. In this technique, the string delay line is "branched" during the transition, i.e., a second feedback loop is formed at the new loop delay, thus forming two delay lines sharing the same memory, one corresponding to the old pitch and the other corresponding to the new pitch. A cross-fade from the old-pitch delay to the new-pitch delay sounds good if the cross-fade time and duration are carefully chosen. Another way to look at this algorithm is in terms of "read pointers" and "write pointers." A normal delay line consists of a single write pointer followed by a single read pointer, delayed by one-period. During a legato transition, we simply cross-fade from a read-pointer at the old-pitch delay to a read-pointer at the new-pitch delay. In this type of implementation, the write-pointer always traverses the full delay memory corresponding to the minimum supported pitch in order that read-pointers may be instantiated at any pitch-period delay at any time. Conceptually, this simplified model of note transitions can be derived from the more rigorous model by replacing the scattering junction at the excitation or finger by a single reflection coefficient.

Comparatively little work has appeared on calibrating string model parameters to recorded measurements [49, 35, 65, 67]. Working from the opposite direction, it is somewhat difficult to find information on the fundamental physical properties of the strings used in real musical instruments [11, 41]. In general, as more physical parameters are pinned down by *a priori* knowledge, techniques for automatic calibration become more successful. Finally, there seems to be no end in sight to future research in the area of distortion algorithms for electric guitar simulations (although strictly speaking, the string is not normally where the nonlinearity lies, but rather somewhere else in the processing, such as in the virtual preamp or speaker).

While digital models for vibrating strings are "ready for prime time," the various ways of exciting them are still under active development. The simplest cases are plucked and struck excitations, and sufficiently good models exist for these [55, 35, 57, 67, 69]. Bowed strings, however, are not yet quite considered to be "parity ready" due to the problems mentioned earlier.

Winds

Woodwinds are sounding quite good in some cases, particularly on the Yamaha VL1 synthesizer which uses digital waveguide techniques [50, 26, 15]. Single-reed theory appears to be pretty well understood, although there are some outstanding questions regarding flow separation in the mouthpiece [27]. Double-reeds, on the other hand, such as the oboe and bassoon, are not yet fully understood theoretically, as evidenced by the fact that new theoretical models are still being proposed [79]. Nevertheless, it is generally acknowledged that the oboe and bassoon presets in the VL1 are excellent as far as they go.

Implementations of woodwind finger-hole models have recently been developed using fractional-delay filters [63] with coefficients based on both theoretical and experimental values from musical acoustics [38].

While the theory of air-jet driven instruments is perhaps the most slippery in all of musical acoustics, existing flute models nevertheless sound good and are highly expressive [15, 33], including register shifting ("overblowing") in response to modulation of the virtual "jet delay." The VL1 shakuhachi patch is a good example of the present state of the art along these lines. Convincing flute models have also been constructed using purely lumped elements to model the bore [77, 76]. In principle, discrete-time simulations of distributed systems can always be approximated to arbitrary precision using lumped modeling elements. (An example is the standard LC model of a transmission line, which corresponds to a mass-spring model of an acoustic waveguide.) Therefore, simple resonators are fully general building blocks for linear systems. However, the use of digital waveguides greatly reduces computational complexity and enables accurate nonlinear extensions [54].

Recorder-like instruments have recently been given a strong boost lately [71, 72], including new applications of the theory of *fluid dynamics*. Marc-Pierre Verge has recently implemented a practical synthesis model at IRCAM based on his thesis work at Eindhoven [74].

Identification of acoustic tube parameters is a fairly classical subject in acoustics, but only recently have papers begun to appear on practical techniques for estimating the filters needed for digital waveguide models of wind instruments [48, 36, 58].

Brasses

Current waveguide brass instruments appear to capture the essential features of brass tones [14, 15, 20, 21, 60]. The horn itself is well understood up to the bell [6, 8, 3, 7, 21]. However, there is apparently no complete theory which describes what happens at the bell exit aperture [8]. For example, nonlinear vortex shedding has been measured downstream from the bell output [28].

Models for the brass-player's lips are rapidly developing [**?**, 18, 1, 14, 46], and there appears to be a trend in the direction of the one- and two-mass models which have been used for years to model the vocal folds in speech [29]. It is clear that the lip model should vary with pitch since the motion of the lips qualitatively changes as the pitch increases [**?**, see Figure 3].

Voice

The human voice is the most communicative musical instrument of all. It is interesting to note that while modern voice synthesis techniques go back much farther than those for other musical instruments [24], it is one of the instruments that is farthest from reaching parity between the synthetic and natural versions. This seems to be primarily due to the complex fluidity of control involved and to the incompleteness of the physio-acoustic models used. Nevertheless, excellent singing voice quality based on a digital waveguide model has been achieved for isolated phrases by Perry Cook [12, 13]. Since voice is in a class by itself relative to other musical instruments, and since it is addressed elsewhere in this special issue [17], it will not be discussed further here.

Membranes, Plates, Solids, and Acoustic Spaces

The most prevalent method for simulating distributed media in more than one dimension, such as membranes and plates, is by means of a *modal expansion*. That is, the resonances of the object are explicitly simulated using second-order resonators, typically arranged in parallel. More recently, explicit physical models in higher dimensions have been developed using meshes of coupled digital waveguides [66, 47, 68]. A particularly convincing example is gong synthesis developed by Scott Van Duyne using nonlinearities and lowpass losses around the rim of a lossless waveguide mesh [69].

Since "plate reverbs" are considered better than "spring reverbs" (which are essentially one dimensional), and since three-dimensional acoustic reverberation (such as in a concert hall) is considered superior to plate reverberation, it is logical to ask whether waveguide meshes in dimensions higher than three will provide yet better reverberation. This remains largely a subject of future research.

Virtual Analog

"Virtual Analog" synthesis is defined as simulating classic analog synthesizers (e.g., Moog, Arp, etc.) using digital methods. At the 1994 NAMM show in Anaheim, the Nord Lead virtual analog synth appeared to be a pretty big hit. It turns out it's not that easy to simulate analog synthesizers well in digital form. For example, it is not obvious how to "digitize" the classic four-pole Moog VCF (also known the "Moog ladder" after the appearance of its schematic). Even something as basic as a sawtooth oscillator is difficult to do "right" in digital form since a simple resetting-ramp approach generates an aliased sawtooth waveform. Using bandlimited interpolation techniques described in [56], it is straightforward, but relatively expensive, to generate properly bandlimited sampled analog waveforms.

Further Outstanding Problems

There are several interesting theoretical and practical problems remaining to be solved in addition to those mentioned above.

Delay Line Interpolation

A deceptively simple problem that applies to nearly all digital waveguide models is that of delay-line interpolation. Integer delay lengths are not sufficient for musical tuning of digital waveguide models at commonly used sampling rates [31]. The simplest scheme which is typically tried first is *linear interpolation*. However, poor results are obtained in some cases (such electric guitars) because the pitch-dependent damping caused by interpolation can be much larger than the desired damping in the string model. In these cases, the interpolation filter becomes the dominant source of damping, so that when the pitch happens to fall on an integer delay-line length, the damping suddenly decreases, making the note stand out as "buzzy."

Allpass interpolation is a nice choice for the nearly lossless feedback loops commonly used in digital waveguide models, because it does not suffer *any* frequency-dependent damping [31]. However, allpass interpolation instead has the problem that instantly switching from one delay to another (as in a hammer-on or pull-off simulation in a string model) gives rise to a *transient artifact* due to the recursive nature of the allpass filter. Recently, Vesa Välimäki has developed a general transient elimination scheme for recursive digital filters of arbitrary order [64].

Another popular choice is *Lagrange interpolation* [33] which is a special case of FIR filter interpolation; while the switching problem does not arise since the interpolating filter is nonrecursive, there is still a time-varying amplitude distortion at high frequencies. In fact, first-order Lagrange interpolation is just linear interpolation, and higher orders can be shown to give a maximally smooth frequency response at DC (zero frequency), while the gain generally rolls off at high frequencies. Allpass interpolation can be seen as trading off this frequency-dependent amplitude distortion for additional frequency-dependent delay distortion [16]. A comprehensive review of Lagrange interpolation appears in [61].

Both allpass and FIR interpolation suffer from some delay distortion at high frequencies due to having a nonlinear phase response at non-integer desired delays. This distortion is normally inaudible, even in the first-order case, causing mistuning or phase modulation only in the highest partial overtones of a resonating string or tube.

Optimal interpolation can be approached via general-purpose bandlimited interpolation techniques [56]. However, the expense is generally considered too high for widespread usage at present. Both amplitude and delay distortions can be eliminated over the entire band of human hearing using higher order allpass or FIR interpolation filters in conjunction with some amount of oversampling. A comprehensive review of delay-line interpolation techniques is due to appear in [39].

Time Varying Filters

Time varying recursive filter structures with convenient controls are hard to find in general. For example, given an analog voltage-controlled filter (VCF) that behaves in a valued way as a function of the control voltage, how does one find a similar digital counterpart? The standard technique for digitizing an analog filter is the bilinear transform, and frequency scaling can be done in the digital domain using an allpass substitution in the digital filter transfer function. To obtain a digital VCF, one might think of implementing real-time frequency scaling by replacing each delay element of the unscaled digital filter with a first-order allpass filter; however, when this is applied to a recursive digital filter, such as the Moog Ladder mentioned earlier, a nonrealizable structure is obtained because a delay-free loop is introduced. The general way to eliminate the delay-free loop is to multiply out the filter denominator and renormalize it, but this destroys the nice control structure which led to the choice of the analog prototype filter in the first place. An ad hoc solution which preserves the control structure is to insert a unit-sample delay in the loop to make it implementable. However, this generally degrades the frequency response at high frequencies.

Conical Bores

To a first approximation, a truncated cone can be regarded as a cylinder open on both ends [4]. To make a more precise model, the phase shift between traveling pressure and velocity needs to be taken into account, or, equivalently, the imaginary part of the wave impedance needs to be modeled [5]. Digital waveguide models have been derived for conical-bore instruments [53, 62, 70], and limited simulations have been successful. However, there is a surprising result in the theory: When a conical tube suddenly decreases in taper angle, such as when crossing from a diverging conical segment into a converging one, the impulse response of the junction actually contains *growing* exponentials [2]. This means that, at such a junction, the straightforward waveguide model must use unstable reflection and transmission filters! It seems highly inappropriate to use unstable filters to model a passive physical bore, and numerically it is highly inadvisable without elaborate schemes to reset the growing round-off noise inside the filters. However, so far, no efficient solution has been found. One expensive solution is to replace the unstable IIR junction filters by large FIR filters which explicitly implement truncated growing exponentials; this solution is based on the observation that in realistic (finite-length) bore geometries, the growing exponentials are always ultimately canceled by reflections from the terminations.

Nonlinearities

Nonlinearities are extremely important in many musical instruments for generating a variable bandwidth over time. Examples include woodwinds, bowed strings, sitars, gongs, cymbals, and distorting electric guitars. In many more cases, sparing use of nonlinearity can serve to spice up the spectrum of any harmonic signal, as is used in so-called "aural exciters."

A general problem in the digital domain is that nonlinearities tend to cause *aliasing*. The simplest (weakest) nonlinearity is the squaring operation, and each time a signal is squared its bandwidth doubles. When a nonlinearity is used in a feedback loop, this bandwidth expansion happens over and over again until aliasing occurs. Even outside of feedback loops, large oversampling factors may be needed to avoid aliasing. Additionally, lowpass filters are often needed to push down the expanding bandwidth when it gets above a certain point. In general, there is very little practical theory for working with nonlinear elements in digital audio systems.

Another problem with nonlinearities in a physical modeling context is that they can effectively "create" or "destroy" signal energy. In a digital waveguide, the energy associated with a single signal sample is proportional to the square of that sample. Applying a nonlinear gain will change the signal energy, in general, and so some higher level framework must be introduced to ensure energy conservation in the presence of nonlinearities. Some recent work has been pursued on "passive nonlinearities" [69] which are developed based on analogous passive nonlinearities in continuous-time system (e.g., a nonlinear spring becomes a switching allpass filter in the digital world). However, in the discrete-time case, these analogies are not exact, and there remains the problem of how to achieve exactly lossless nonlinearities. A related problem is how to "feed back" round-off errors in otherwise lossless computations such that energy is exactly preserved; the solution is elementary, but applications do not yet seem to exist. Recent analytical work [45] has helped to characterize under what conditions feedback loops containing nonlinearities will at least be stable; for example, by restricting the class of nonlinearities to certain ratios of polynomials, stability can be guaranteed. A general treatment of the problem of stability of a waveguide network in the presence of nonlinearities may be found in [19].

Automated Turbulence

In wind instrument models, turbulence is generally modeled using ad hoc filtered noise injection. It would be considerably more convenient, and presumably more accurate, to generate the turbulence automatically using a model for turbulence generation. For example, given air velocity and mouthpiece geometry, the noise spectrum can sometimes be predicted fairly well from theory. Early ideas along these lines were pursued in articulatory speech synthesis [23]. A recent turbulence model for flue pipes based on the Lighthill theory is described in [73], and more extensively in [72, Chapter 5]. Another based on sound produced by vortex shedding was proposed in [10]. An advantage of automatically generated noise due to turbulence is that pulsed modulation at the pitch rate falls out automatically [9].

Control

The big advantage of making a virtual instrument based on a physical model is to obtain the entire range of expressive variations in the instrument in response to intuitive controls. Unfortunately, controlling physical models gracefully in real time can be quite difficult, especially with sustained instruments such as bowed strings, woodwinds, and most particularly the human voice. In general, we need to find an "orthogonalizing" software layer to place between the performer and model so that "simple things are simple," yet everything is still possible. The VL1 does a surprisingly good job of fencing in the parameter space so that each voice almost always sounds, and is reasonably in tune. Research is proceeding in this direction, but at present, there seem to be few broadly applicable techniques in the open literature.

Conclusions

An overview of recent results in music synthesis based on physical models has been presented. Basic "distributed" building blocks such as vibrating strings and resonating bores are modeled very effectively using the digital waveguide approach. Additionally, "lumped" systems such as Helmholtz resonators and mass-spring-damper combinations are easily modeled using second-order digital filter sections. However, many special details of various instruments such as mouthpieces, reeds, radiation load, and nonlinearities require further research. In general, we hope that all such components can be modeled effectively using low-order digital filters, nonlinear polynomials/table-lookups, and in some cases filtered noise injection. Promising preliminary results have been obtained along these lines. Finally, there remains much research to be done into estimating model parameters such as filter and polynomial coefficients from measured data. As these methods evolve, virtual musical instruments based on physical models will become easier to devise, like "wavetable voices" based on sampled sound. In fact, the physical models can be seen as a form of "structured sampling synthesis" in which deeper physical parameters are sampled in place of the simple air pressure fluctuations recorded in traditional sampling synthesis.

References

[1] S. Adachi and M. Sato, "On the transition of lip-vibration states in the brass instrument," in *Proceedings of the International Symposium on Musical Acoustics (ISMA-95), Dourdan, France*, (France), pp. 17–22, Société Français d'Acoustique, July 1995.

[2] J. Agullo, A. Barjau, and J. Martinez, "Alternatives to the impulse response h(t) to describe the acoustical behavior of conical ducts," *Journal of the Acoustical Society of America*, vol. 84, pp. 1606–1612, Nov. 1988.

[3] N. Amir, G. Rosenhouse, and U. Shimony, "Reconstructing the bore of brass instruments: Theory and experiment," in *Proceedings of the Stockholm Musical Acoustics Conference (SMAC-93)*, (Stockholm), pp. 470–475, Royal Swedish Academy of Music, July 1993.

[4] R. D. Ayers, L. J. Eliason, and D. Mahgerefteh, "The conical bore in musical acoustics," *American Journal of Physics*, vol. 53, pp. 528–537, June 1985.

[5] A. H. Benade, "Equivalent circuits for conical waveguides," *Journal of the Acoustical Society of America*, vol. 83, pp. 1764–1769, May 1988.

[6] A. H. Benade and E. V. Jansson, "On plane and spherical waves in horns with nonuniform flare. I. theory of radiation, resonance frequencies, and mode conversion," *Acustica*, vol. 31, no. 2, pp. 80–98, 1974.

[7] D. P. Berners and J. O. Smith, "Super-spherical wave simulation in flaring horns," in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 112–113, Computer Music Association, 1995.

[8] R. Caussé, J. Kergomard, and X. Lurton, "Input impedance of brass musical instruments—comparison between experiment and numerical models," *Journal of the Acoustical Society of America*, vol. 75, pp. 241–254, Jan. 1984.

[9] C. Chafe, "Pulsed noise in self-sustained oscillations of musical instruments," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Albuquerque*, (New York), IEEE Press, 1990, available as CCRMA Technical Report STAN-M-65, Music Dept., Stanford University.

[10] C. Chafe, "Adding vortex noise to wind instrument physical models," in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 57–60, Computer Music Association, 1995.

[11] A. Chaigne, "Viscoelastic properties of nylon guitar strings," *Catgut Acoustical Society Journal, Series II*, vol. 1, pp. 21–27, May 1991.

[12] P. R. Cook, *Identification of Control Parameters in an Articulatory Vocal Tract Model, with Applications to the Synthesis of Singing*, PhD thesis, Elec. Engineering Dept., Stanford University (CCRMA), Dec. 1990, `http: //www.cs.princeton.edu/~prc/`.

[13] P. R. Cook, "Spasm: A real-time vocal tract physical model editor, controller, and singer: The companion software synthesis system," in *Proceedings of the Colloquium on Physical Modeling*, (Grenoble, France), ACROE, 1990.

[14] P. R. Cook, "Tbone: An interactive waveguide brass instrument synthesis workbench for the NeXT machine," in *Proceedings of the 1991 International Computer Music Conference, Montreal*, pp. 297–299, Computer Music Association, 1991.

[15] P. R. Cook, "A meta-wind-instrument physical model, and a meta-controller for real time performance control," in *Proceedings of the 1992 International Computer Music Conference, San Jose*, pp. 273–276, Computer Music Association, 1992.

[16] P. R. Cook, "Integration of physical modeling for synthesis and animation," in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 525–528, Computer Music Association, 1995.

[17] P. R. Cook, "Singing voice synthesis: History, current work, and future directions," *Computer Music Journal*, vol. 20, pp. 38–46, Fall 1996.

[18] D. C. Copley and W. J. Strong, "Stroboscopic sequences of lip vibrations in a trombone," in *Proceedings of the International Symposium on Musical Acoustics (ISMA-95), Dourdan, France*, (France), pp. 42–46, Société Français d'Acoustique, July 1995.

[19] R. J. de Figueiredo and C. Y. Ho, "Absolute stability of a system of nonlinear networks by lossless transmission lines," *IEEE Transactions on Circuit Theory*, vol. 17, pp. 575–584, Nov 1970.

[20] P. H. Dietz, "Simulation of trumpet tones via physical modeling," Master's thesis, Elec. and Comp. Engineering Dept., Carnegie Mellon University, Dec 1988, available as CMU Report no. LASIP-89-07, Aug. 1989.

[21] P. H. Dietz and N. Amir, "Synthesis of trumpet tones by physical modeling," in *Proceedings of the International Symposium on Musical Acoustics (ISMA-95), Dourdan, France*, (France), pp. 471–477, Société Français d'Acoustique, July 1995.

[22] W. C. Elmore and M. A. Heald, *Physics of Waves*, New York: Dover, 1969.

[23] J. L. Flanagan and K. Ishizaka, "Automatic generation of voiceless excitation in a vocal cord-vocal tract speech synthesizer," *IEEE Transactions on Acoustics, Speech, Signal Processing*, vol. 24, pp. 163–170, April 1976.

[24] J. L. Flanagan and L. R. Rabiner, eds., *Speech Synthesis*, Stroudsburg, Penn.: Dowden, Hutchinson, and Ross, Inc., 1973.

[25] R. Hanson, "Measurements of nonlinear effects in a driven vibrating wire," *Journal of the Acoustical Society of America*, vol. 96, pp. 1549–1556, Sept. 1994.

[26] S. Hirschman, *Digital Waveguide Modelling and Simulation of Reed Woodwind Instruments*, Engineer's thesis, Elec. Engineering Dept., Stanford University (CCRMA), May 1991, available as CCRMA Technical Report Stan-M-72, Music Dept., Stanford University, July 1991.

[27] A. Hirschberg, "A quasi-stationary model of air flow in the reed channel of single-reed woodwind instruments," *Acustica*, vol. 70, pp. 146–154, 1990.

[28] A. Hirschberg, "private communication," 1992.

[29] K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords," *Bell System Technical Journal*, vol. 51, pp. 1233–1268, 1972.

[30] D. A. Jaffe, "Ten criteria for evaluating synthesis and processing techniques," *Computer Music Journal*, vol. 19, no. 1, pp. 76–87, 1995, Also: Proceedings of the First Brazilian Symposium of Computation and Music, Brazilian Society of Computation, Caxambu, MG. 53-61.

[31] D. A. Jaffe and J. O. Smith, "Extensions of the Karplus-Strong plucked string algorithm," *Computer Music Journal*, vol. 7, no. 2, pp. 56–69, 1983, Reprinted in [44, pp. 481–494].

[32] D. A. Jaffe and J. O. Smith, "Performance expression in commuted waveguide synthesis of bowed strings," in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 343–346, Computer Music Association, 1995.

[33] M. Karjalainen, U. K. Laine, T. I. Laakso, and V. Välimäki, "Transmission-line modeling and real-time synthesis of string and wind instruments," in *Proceedings of the 1991 International Computer Music Conference, Montreal*, pp. 293–296, Computer Music Association, 1991.

[34] M. Karjalainen, J. Backman, and J. Pölkki, "Analysis, modeling, and real-time sound synthesis of the kantele, a traditional Finnish string instrument," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Minneapolis*, (New York), pp. 229–232, IEEE Press, 1993.

[35] M. Karjalainen, V. Välimäki, and Z. Jánosy, "Towards high-quality sound synthesis of the guitar and string instruments," in *Proceedings of the 1993 International Computer Music Conference, Tokyo*, pp. 56–63, Computer Music Association, Sept. 10–15 1993, `http: //-www.acoustics.hut.fi/~vpv/publications/icmc93-guitar.htm`.

[36] M. Karjalainen, V. Välimäki, B. Hernoux, and J. Huopaniemi, "Exploration of wind instruments using digital signal processing and physical modeling techniques," in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 509–516, Computer Music Association, 1995, Revised version published in the *Journal of New Music Research*, 24(4), December 1995.

[37] K. Karplus and A. Strong, "Digital synthesis of plucked string and drum timbres," *Computer Music Journal*, vol. 7, no. 2, pp. 43–55, 1983, Reprinted in [44, pp. 467–479].

[38] D. H. Keefe, "Theory of the single woodwind tone hole. experiments on the single woodwind tone hole," *Journal of the Acoustical Society of America*, vol. 72, pp. 676–699, Sept. 1982.

[39] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the Unit Delay—Tools for Fractional Delay Filter Design," *IEEE Signal Processing Magazine*, vol. 13, pp. 30–60, January 1996.

[40] M. E. McIntyre, R. T. Schumacher, and J. Woodhouse, "On the oscillations of musical instruments," *Journal of the Acoustical Society of America*, vol. 74, pp. 1325–1345, Nov. 1983.

[41] N. C. Pickering, *The Bowed String*, Mattituck NY: Amereon, Ltd., 1991, Also available from Bowed Instruments, 23 Culver Hill, Southampton, NY 11968.

[42] J. R. Pierce, "private communication," 1994.

[43] N. Porcaro, P. Scandalis, J. O. Smith, D. A. Jaffe, and T. Stilson, "SynthBuilder—a graphical real-time synthesis, processing and performance system," in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 61–62, Computer Music Association, 1995.

[44] C. Roads, ed., *The Music Machine*, Cambridge, MA: MIT Press, 1989.

[45] X. Rodet, "Flexible yet controllable physical models: A nonlinear dynamics approach," in *Proceedings of the 1993 International Computer Music Conference, Tokyo*, pp. 10–15, Computer Music Association, 1993.

[46] X. Rodet, "One and two mass model oscillations for voice and instruments," in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 207–214, Computer Music Association, 1995.

[47] L. Savioja, J. Backman, A. Järvinen, and T. Takala, "Waveguide mesh method for low-frequency simulation of room acoustics," *Proceedings of the 15th International Conference on Acoustics (ICA-95), Trondheim, Norway*, pp. 637–640, June 1995.

[48] G. Scavone and P. R. Cook, "Combined linear and non-linear prediction in calibrating models of musical instruments to recordings," in *Proceedings of the 1994 International Computer Music Conference, Århus*, pp. 433–434, Computer Music Association, 1994, Send email to gary@ccrma.stanford.edu to request a copy of the full paper.

[49] J. O. Smith, *Techniques for Digital Filter Design and System Identification with Application to the Violin*, PhD thesis, Elec. Engineering Dept., Stanford University (CCRMA), June 1983, CCRMA Technical Report STAN-M-14, `http: //-ccrma.stanford.edu/STANM/stanms/stanm14/`.

[50] J. O. Smith, "Efficient simulation of the reed-bore and bow-string mechanisms," in *Proceedings of the 1986 International Computer Music Conference, The Hague*, pp. 275–280, Computer Music Association, 1986, also available in [51].

[51] J. O. Smith, "Music applications of digital waveguides," Tech. Rep. STAN-M-39, CCRMA, Music Department, Stanford University, 1987, CCRMA Technical Report STAN-M-39, `http: //ccrma.stanford.edu/STANM/stanms/stanm39/`.

[52] J. O. Smith, "Waveguide filter tutorial," in *Proceedings of the 1987 International Computer Music Conference, Champaign-Urbana*, pp. 9–16, Computer Music Association, 1987.

[53] J. O. Smith, "Waveguide simulation of non-cylindrical acoustic tubes," in *Proceedings of the 1991 International Computer Music Conference, Montreal*, pp. 304–307, Computer Music Association, 1991.

[54] J. O. Smith, "Physical modeling using digital waveguides," *Computer Music Journal*, vol. 16, pp. 74–91, Winter 1992, special issue: Physical Modeling of Musical Instruments, Part I. `http: //ccrma.stanford.edu/~jos/pmudw/`.

[55] J. O. Smith, “Efficient synthesis of stringed musical instruments,” in *Proceedings of the 1993 International Computer Music Conference, Tokyo*, pp. 64–71, Computer Music Association, 1993, incorporated into [**?**].

[56] J. O. Smith and P. Gossett, “A flexible sampling-rate conversion method,” in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, San Diego*, vol. 2, (New York), pp. 19.4.1–19.4.2, IEEE Press, March 1984, expanded tutorial and associated free software available at the Digital Audio Resampling Home Page: `http://-ccrma.stanford.edu/~jos/resample/`.

[57] J. O. Smith and S. A. Van Duyne, “Commuted piano synthesis,” in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 319–326, Computer Music Association, 1995, `http://ccrma.stanford.edu/~jos/pdf/svd95.pdf`.

[58] T. Stilson, “Forward-going wave extraction in acoustic tubes,” in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 517–520, Computer Music Association, 1995.

[59] C. R. Sullivan, “Extending the Karplus-Strong algorithm to synthesize electric guitar timbres with distortion and feedback,” *Computer Music Journal*, vol. 14, pp. 26–37, Fall 1990.

[60] T. K. M. S. R. Team, “OASYS demonstration for the san francisco bay area chapter of the audio engineering society,” tech. rep., Korg R&D, Milpitas, CA, May 18 1995.

[61] V. Välimäki, *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters*, PhD thesis, Report no. 37, Helsinki University of Technology, Faculty of Electrical Engineering, Laboratory of Acoustic and Audio Signal Processing, Espoo, Finland, Dec. 1995, `http://-www.acoustics.hut.fi/~vpv/publications/vesa_phd.html`.

[62] V. Välimäki and M. Karjalainen, “Digital waveguide modeling of wind instrument bores constructed of truncated cones,” in *Proceedings of the 1994 International Computer Music Conference, Århus*, pp. 423–430, Computer Music Association, 1994.

[63] V. Välimäki, M. Karjalainen, and T. I. Laakso, “Modeling of woodwind bores with finger holes,” in *Proceedings of the 1993 International Computer Music Conference, Tokyo*, pp. 32–39, Computer Music Association, 1993.

[64] V. Välimäki, T. I. Laakso, and J. Mackenzie, “Elimination of transients in time-varying allpass fractional delay filters with application to digital waveguide modeling,” in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 327–334, Computer Music Association, 1995.

[65] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy, “Physical modeling of plucked string instruments with application to real-time sound synthesis,” *Journal of the Audio Engineering Society*, vol. 44, pp. 331–353, May 1996.

[66] S. A. Van Duyne and J. O. Smith, “Physical modeling with the 2-D digital waveguide mesh,” in *Proceedings of the 1993 International Computer Music Conference, Tokyo*, pp. 40–47, Computer Music Association, 1993, `http://ccrma.stanford.edu/~jos/pdf/mesh.pdf`.

[67] S. A. Van Duyne and J. O. Smith, "Developments for the commuted piano," in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 335–343, Computer Music Association, 1995, `http: //ccrma.stanford.edu/ jos/pdf/svd95.pdf`.

[68] S. A. Van Duyne and J. O. Smith, "The tetrahedral waveguide mesh: Multiply-free computation of wave propagation in free space," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY*, (New York), pp. 9a.6.1–4, IEEE Press, Oct. 1995.

[69] S. A. Van Duyne, J. R. Pierce, and J. O. Smith, "Traveling-wave implementation of a lossless mode-coupling filter and the wave digital hammer," in *Proceedings of the 1994 International Computer Music Conference, Århus*, pp. 411–418, Computer Music Association, 1994, Also presented at the conference of the Acoustical Society of America, Nov., 1994.

[70] M. van Walstijn and G. de Bruin, "Conical waveguide filters," in *Proceedings of the International Conference on Acoustics and Musical Research*, (Ferrara, Italy), pp. 47–54, CIARM, May 1995.

[71] M. P. Verge, "Jet oscillations and jet drive in recorder-like instruments," *Acta Acustica 2*, pp. 403–419, 1994.

[72] M. P. Verge, *Aeroacoustics of Confined Jets with Applications to the Physical Modeling of Recorder-Like Instruments*, PhD thesis, Eindhoven University, 1995.

[73] M. P. Verge and A. Hirschberg, "Turbulence noise in flue instruments," in *Proceedings of the International Symposium on Musical Acoustics (ISMA-95), Dourdan, France*, (France), pp. 93–99, Société Français d'Acoustique, July 1995.

[74] M. P. Verge, R. Caussé, and A. Hirschberg, "A physical model of recorder-like instruments," in *Proceedings of the 1995 International Computer Music Conference, Banff*, pp. 37–44, Computer Music Association, 1995.

[75] A. Watzky, "On the transverse-axial mode coupling in the non-linear vibration of strings," in *Proceedings of the International Symposium on Musical Acoustics (ISMA-95), Dourdan, France*, (France), pp. 190–195, Société Français d'Acoustique, July 1995.

[76] J. Wawrzynek, "VLSI models for sound synthesis," in *Current Directions in Computer Music Research* (M. V. Mathews and J. R. Pierce, eds.), pp. 113–148, Cambridge, MA: MIT Press, 1989.

[77] J. Wawrzynek, C. Mead, T. Lin, H. Liu, and L. Dyer, "A VLSI approach to sound synthesis," in *Proceedings of the 1984 International Computer Music Conference, Paris*, pp. 53–64, Computer Music Association, 1984.

[78] G. Weinreich, "Coupled piano strings," *Journal of the Acoustical Society of America*, vol. 62, pp. 1474–1484, Dec 1977, see also [**?**] and *Scientific American*, vol. 240, p. 94, 1979.

[79] A. P. Wijnands and A. Hirschberg, "Effect of a pipe neck downstream of a double reed," in *Proceedings of the International Symposium on Musical Acoustics (ISMA-95), Dourdan, France*, (France), pp. 149–152, Société Français d'Acoustique, July 1995.