CENTER FOR COMPUTER RESEARCH IN MUSIC AND ACOUSTICS MAY1988

Department of Music
Report No. STAN-M-50

IMPLEMENTATION OF SINGLE REED INSTRUMENTS WITH ARBITRARY BORE SHAPES USING DIGITAL WAVEGUIDE FILTERS

Perry R. Cook

Research sponsored in part by
The System Development Foundation
and
Dynacord, Germany

CCRMA
DEPARTMENT OF MUSIC
Stanford University
Stanford, California 94305

IMPLEMENTATION OF SINGLE REED INSTRUMENTS WITH ARBITRARY BORE SHAPES USING DIGITAL WAVEGUIDE FILTERS

by Perry R. Cook

I. INTRODUCTION

The accurate physical modeling of musical instruments is desirable from the standpoint of analysis, and the efficient duplication of instrument behavior is desirable from the standpoint of synthesis. Certain musical instruments, including some members of the wind instrument family have been successfully and efficiently modeled in the past using waveguide filters [1][2][3]. The one-dimensional propagation model of most wind instrument bores, coupled with the assumption of a single input / single output system, makes the use of waveguides an exceedingly efficient method of simulating the vibration of many wind instruments. Moreover, the approximation of a bore instrument by differential cylinders (an analog to the scattering section) is the basis for many finite element techniques, and musical instruments have been accurately analyzed using this technique [4]. My starting point for the instrument models presented in this paper was the waveguide clarinet model developed by Julius Smith [2], based on the work of McIntyre, Schumacher, and Woodhouse [5]. Figure 1 shows a basic block diagram for this clarinet and all of the instruments modeled in my investigation. The principal features are the non-linear reed section, the bore, and the transmission/reflection termination section.

Figure 1. Basic Bore Instrument Model

In this project, a synthesis-by-modeling / analysis-by-synthesis process was used on a variety of single reed driven wind instruments. The procedure is described below:

1. Significant features of the instrument in question were analyzed. This included taking physical measurements, consulting the acoustical literature as related to the instrument, and recording reference tones of the actual instrument. The features of each instrument are described in that instrument's section of this paper.

2. A waveguide model of the instrument bore was constructed purely from physical measurements. This process is described in detail in Section III of this paper.

3. Termination characteristics for the bore were determined. A basic acoustics approach was used, along with input/output transfer data from recorded sounds.

4. The single reed non-linear mechanism was adjusted for the particular instrument being modeled. This usually required only a scaling of the reed parameters (which very much maps to overall reed stiffness and mouthpiece aperture size).

5. The model was built and tested using various input breath pressure functions. Fine tuning was done during the testing phase, specifically the breath pressure control envelope, the termination filter characteristics, and the reed parameters. This experimentation phase continued until a set of reasonable and natural short sound examples was generated using the model.

II. THE CLARINET

The basic waveguide clarinet model is well documented in [2], so only a brief explanation will be given here. As with all of the systems investigated in this paper, the only input control variable is breath pressure, and pressure is the variable which is computed throughout the model. Pressure is inserted into the reed simulation section, and is typically on the order of the dynamic range of the system itself. That is, for 16 bit synthesis, the input pressure is ramped from zero to about 10^5.

The reed section is the only non-linear block in the system, and is responsible for much of the realistic behavior of the model. Incoming breath pressure is compared to the current left-traveling wavefront in the bore, and this difference is proportional to the net force on the reed. The reed is assumed massless, and thus the net instantaneous force corresponds to an instantaneous position of the reed (like a massless spring system). The position of the reed corresponds to a slit aperture size, which in turn yields a reflection coefficient for the impinging wave value. The slit size also controls the amount of incoming breath pressure which is allowed through the slit and into the bore. The compute time relationship of a value of pressure difference (breath - wave) to a reflection coefficient is done via a table lookup scheme, greatly simplifying computation.

Because of the cylindrical shape of the clarinet bore, the bore section of the clarinet model is very simple, requiring only two delay lines to simulate left-going and right-going wave components within the bore. No losses or disturbances (such as tone holes) are included in any of the basic models, although section V briefly discusses tone hole considerations.

The final section in the model is the reflection/transmission section. If the clarinet bell is neglected, a simple model of the junction at the end of the instrument is that of a low-pass filter reflection function, and a complementary high-pass transmission function (our output). The physical justification for this assumption is that the cap of air which loads the end of the bore can be viewed as a piston, and the physics of driving a frictionless piston dictate a low-pass reflection behavior and a complementary high-pass transmission [6]. Wavelengths much longer than the bell diameter are reflected, and using the clarinet bore diameter dictates a cutoff frequency of about 5,000 Hz. This can be approximated very efficiently at a sampling rate of 20kHz by a simple one-zero filter.

The model worked quite well as described, with the only difficult task being the development of good breath pressure control functions. The clarinet model is extremely sensitive to the slope of the onset of breath pressure, displaying characteristic clarinet overblowing (locking on higher order vibration modes). Once in the steady state, a rather small decrease in breath pressure yielded a large change in oscillatory behavior. A very small amount of white noise mixed with the breath pressure function makes the system more robust, yielding more natural sounding attack transients. Figure 2 shows an FFT of a typical steady state clarinet tone, superimposed with the expected clarinet resonance modes from published tables [7].

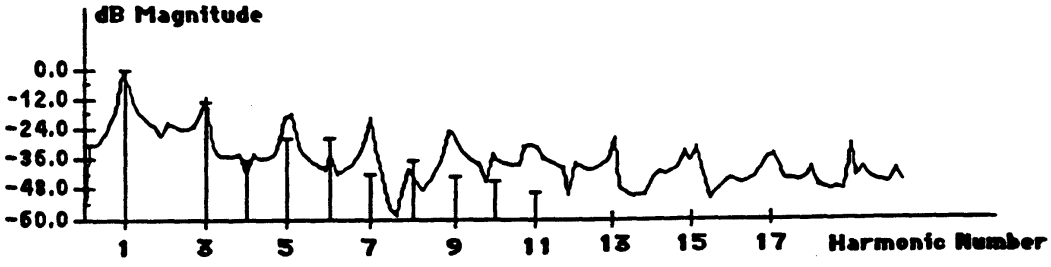

Figure 2. Steady State Waveguide Clarinet Spectrum vs. Predicted Spectrum

III. THE SOPRANO SAXOPHONE

The soprano saxophone is a nearly perfect conical bore, but with length and bell diameter measurements very close to those of the clarinet. This made it a very good choice for the next simulation. The reed table and bell transmission/reflection filters used for the clarinet were left intact, but the simple delay lines used for the clarinet bore were replaced by the characteristic saxophone conical bore. A general purpose wavequide bore was constructed using one scattering junction per spatial sample (one junction per 1.6cm at 20kHz sampling rate). Each junction scattering coefficient is a function of the characteristic impedances of the two sections it joins, and since characteristic impedance is a function of area, the scattering coefficients can be computed from the area function. The conical bore exhibits the characteristic of a cone, in which the radius is a linear function of the position down the bore.

$$\text{radius} = \text{initial radius} + \text{flare} * (\text{bore position})$$

$$r_n = r_0 + f * x_n \qquad (1$$

This corresponds to an area function which is a function of position squared.

$$\text{area of nth section} = a_n = \pi * r_n^2 = \pi * (r_0 + f * x_n)^2 \qquad (2$$

The reflection coefficient is found by forming the ratio of the difference of the areas and the sum of the areas in the two adjacent sections.

$$c_n = [a_n - a_{(n-1)}] / [a_n + a_{(n-1)}] \qquad (3$$

Using the measured data from a soprano saxophone with an initial radius of 3.75mm, a flare of .0359mm/mm, and a length of 66.2cm, a 40 section saxophone waveguide bore was constructed and integrated into the model. The physics of an expanding bore dictate that the pressure shifts from high to low (with an accompanying reciprocal velocity relationship) as the wave travels from the mouthpiece to the bell. This required that high pressures be accomodated in the reed table in order for significant output to be realized. The high mouthpiece pressure corresponds physically very well with the stiffer reeds typically used on saxophones. The saxophone is also quite different from the clarinet in the region where the mouthpiece connects to the horn, in that the sax mouthpiece fits over the horn. There is thus a large discontinuity (abruptly shrinking area) at this point, making the first junction reflection coefficient large and positive, while all others within the bore are negative and decreasing.

The tones produced by the soprano saxophone model were characteristically horn-like, and as expected, quite distinct from the sound of the clarinet model. The sensitivity to the breath pressure envelope was much less severe than in the clarinet, with a wide range of acceptable values for onset slope and steady state value. It was found that by adding a slow periodic deviation to the breath pressure function in steady state, a shift in pitch as well as amplitude was achieved. By experimenting with the attack envelope length and slope, and the breath tremolo/vibrato amplitude, several convincing saxophone tones were realized. Figure 3 shows the amplitude envelope and corresponding fundamental pitch trajectory from a typical saxophone tone, demonstrating the pitch deviation which results from only breath pressure modulation.

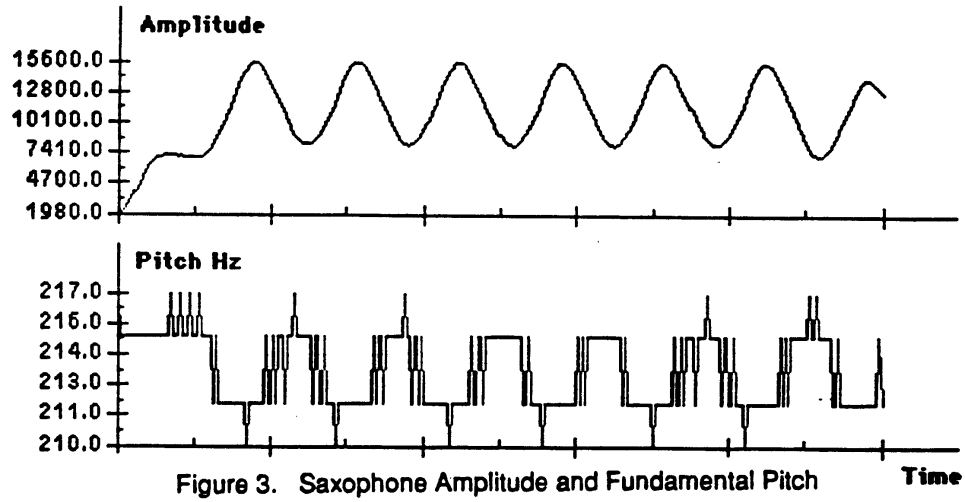

Figure 3. Saxophone Amplitude and Fundamental Pitch

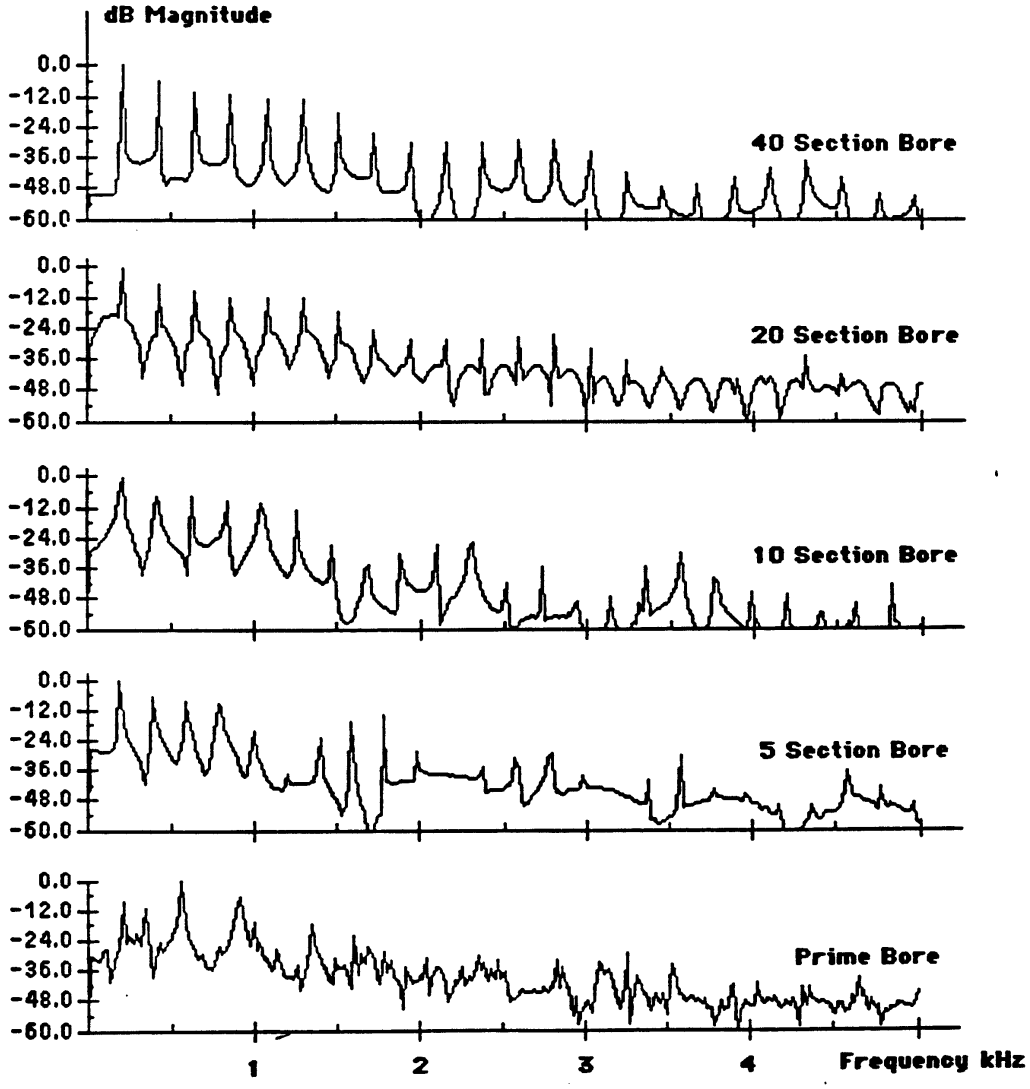

Figure 4. Saxophone Steady State Spectra

The most accurate model of the saxophone bore, using one junction per spatial sample, approximates the bore as a series of cylindrical segments whose lengths are determined by the sampling rate. This has been shown to be an exact simulation at the sample points of one dimensional vibration in a linear lossless system [8]. However, the computational overhead for such a model is quite large. The next step in the saxophone experiments was to simplify the model, by approximating the bore with fewer and longer cylindrical segments. Of course this 'model decimation' causes a compromise in accuracy, but a tempting decrease in computational complexity makes the simpler models worthy of investigation. The model was decimated by factors of 2, 4, and 8, yielding bores of 20, 10, and 5 sections. Noting that particular harmonics were being increased by the approximation process, a decimation scheme was done in which the bore was approximated with 5 sections of uneven lengths. The sections were of relatively prime sample lengths, and it was hoped that this scheme would spread out the effects of the approximation. The 3 even decimation schemes still exhibited characteristic horn waveforms, although progressively coarser for the more severe approximations. Rather than the expected result, the prime decimation scheme yielded an interesting multiphonic oscillation, with many unfused tones present. Figure 4 shows steady state FFT plots of the original model, the three evenly decimated models, and the prime decimated model.

IV. THE TIME VARYING, SINGLE REED DRIVEN, HUMAN VOCAL TRACT

In order to investigate truly arbitrary bore shapes, various bores were constructed from data on the human vocal tract [9]. The shape functions were copied and translated into reflection coefficients for waveguides. Taking a vocal tract length of 8 inches for an adult male (this is somewhat longer than average), only about 13 scattering sections are required for a waveguide bore model at 20kHz sampling rate. Two bores were constructed, one fashioned after the neutral Uh sound (as in *rubber*), and the other shaped like the Oo sound (as in *food*). The clarinet reed calibration proved suitable for the models, as the excitation end of the vocal tract is not significantly smaller than the exit opening. In fact, the mouth opening area is often much smaller than the excitation end, and thus problems can occur with excessive output pressure at the exit point.

Since the network was driven with a single reed, which does not correspond very closely with the highly damped vocal folds, the fundamental pitch of oscillation is much higher than that of the male voice. The massless reed does not display the sluggishness of the vocal folds, and thus the oscillation is likely to lock on to one of the many formant peaks. The fundamental which most commonly occurs is the first resonance formant, but many modes are possible. The principal difficulty with 'playing' this instrument was the mode locking phenomenon, and this prompted the final selection of the more neutral vowels. Very small changes in breath pressure, bore shape, or vibrato caused the vibrational behaviors to rapidly shift. All of this is consistent with the massless reed non-linear oscillator, in that the reed is extremely sensitive to the loading of the tube to which it is connected.

A higher fundamental pitch means that there are less harmonics in the final waveform, so the formant envelope was much more difficult to trace. The two steady state vibrations did show differences (most noticeably a difference in fundamental pitch), and the next step in modeling was to make a transition gradually from one shape to the other. The shape was modulated between the two vowel shapes sinusoidally, that is, each coefficient varied from its value in the Oo bore to its value in the Uh bore in a sinusoidal fasion. The transition period from one bore to the other and back was selected to be one second. This experiment produced an effect which was more interesting musically than it was speechlike, which again is consistent with the high fundamental produced by the reed, and the resulting lack of ease in hearing formants. The actual sound was like a musical instrument which changed from a clarinet-like characteristic, to a more sax-like sound. Figure 5 shows steady state FFTs of the oscillations of the two bores, along with the normal formant curves of the two vowels.

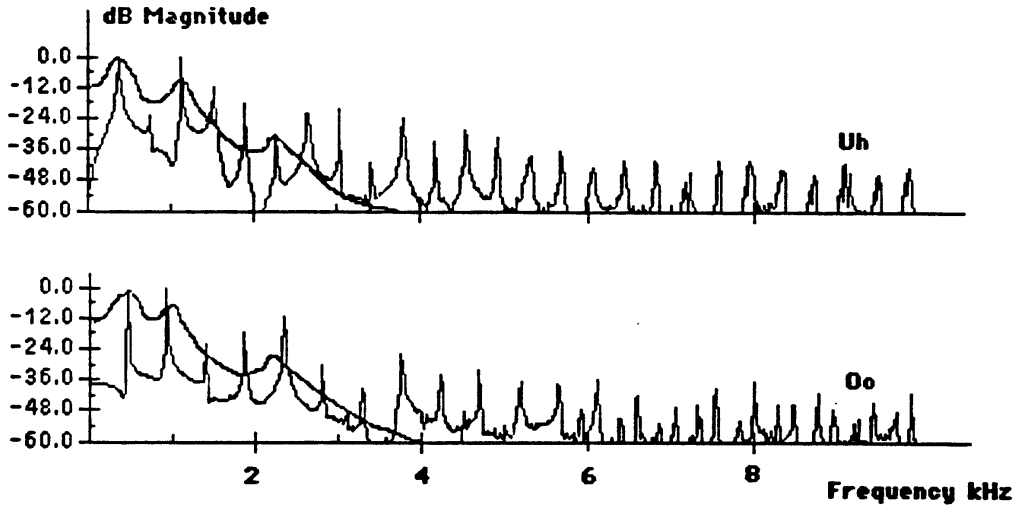

Figure 5. Single Reed Vocal Tract Spectra and Expected Formants

V. TONE HOLES

The behavior of these bores is quite interesting, even in their sterile, lossless form. However the addition of tone holes is required in order to bring the models into a more musical form. A tone hole performs many important functions in a wind instrument. One is the obvious function of changing the effective length of the bore, thus allowing pitch changing. Another is to filter out resonance modes, as in the clarinet register key, which allows the player to overblow more easily, and thus this function is also related to pitch control. Another important function is that of radiation, allowing some of the sound energy to escape and add to the overall percieved sound of the instrument.

Neglecting turbulance at the tone hole, tone holes can be grouped into two categories [10][11]. The two distinct types of tone holes are differentiated by their chimney height. That is, tone holes which have negligible length as tubes, and those which enclose so much air volume that the mass of this air must be included. Tone holes with negligible height can be modeled as simple three way scattering junctions. Further, shallow tone holes which are located between horn sections of the same cross-sectional area can be reduced to a simple single multiply structure. The analysis of a deep chimney tone hole is like the bell reflection/transmission analysis of Section II, with a low pass reflection component, and a high pass transmission component.

Even if the tone hole characteristics are simple, as in the shallow chimney case, the notion of placing all of the tone holes of an actual instrument into a bore model points to a large increase in computation. In order to accurately locate the holes fractionally between spatial samples as required, all pass filters could be employed to yield the necessary fractional sample delays. But this points to even more computation. Add the filters required for deep chimney tone holes, and the computation increases yet further.

The time varying bore aspect of waveguide filters is one possible solution to the problem. In the experiments on both the saxophone and the vocal tract, it was clear that large pitch shifts are possible in bores of fixed length, because of the reaction of the reed with the bore shape. The acoustic lumped circuit model of a chain of tone holes is a ladder filter, with each tone hole being replaced with two series impedances and one shunt impedance. Some testing has been done toward simulating the effects of tone holes with the same time varying bore used for the vocal tract instruments discussed in this paper. One benefit of this type of implementation is that the computational overhead is constant, with each junction requiring some fixed amount of computations per sample.

VI. REFERENCES

[1] K. Karplus and A. Strong, "Digital Synthesis of Plucked String and Drum Timbres," *Computer Music Journal*, vol. 7, no. 2, pp. -55, 1983.

[2] J. Smith, "Efficient Simulation of the Reed-Bore and Bow-String Mechanisms," Proc. 1986 Int. Conf. Computer Music, The Hague, Netherlands.

[3] G. Garnett, "Modeling Piano Sound Using Waveguide Digital Filtering Techniques,", Proc. 1986 Int. Conf. Computer Music, Urbana, Illinois.

[4] G. Plitnik and W. Strong, "Numerical Method for Calculation Input Impedances of the Oboe," *J. Acoustical Soc. Amer.*, vol. 63, no. 3, pp. 816-825, 1979.

[5] M. McIntyre, R. Schumacher, and J. Woodhouse, "On the Oscillations of Musical Instruments," *J. Acoustical Soc. Amer.*, vol. 74, no. 5, pp. 1325-1345, 1983.

[6] P. Morse, *Vibration and Sound*, published by the American Institute of Physics for the Acoustical Society of America.

[7] J. Backus, *The Acoustical Foundations of Music*, 2nd ed., W. W. Norton, New York, 1977.

[8] P. Ruiz, "A Technique for Simulating the Vibrations of Strings with a Digital Computer," Master of Music Dissertation, Univ. Ill., Urbana, 1969.

[9] J. Flanagen, Speech Analysis, Synthesis, and Perception, 2nd ed., Springer Verlang, New York, 1972.

[10] D. Keefe, "Theory of the Single Woodwind Tone Hole," *J. Acoustical Soc. Amer.*, vol. 72, no. 3, pp. 676 - 687, 1982.

[11] D. Keefe, "Experiments on the Single Woodwind Tone Hole," *J. Acoustical Soc. Amer.*, vol. 72, no. 3, pp. 688 - 699, 1982.

The author gratefully acknowledges that this work was supported in part by a grant from Dynacord.