MODELING HIGH-FREQUENCY MODES OF COMPLEX RESONATORS USING A WAVEGUIDE MESH

Patty Huang, Stefania Serafin and Julius O. Smith III

CCRMA, Department of Music
Stanford University, Stanford, CA
pph@ccrma.stanford.edu, serafin@ccrma.stanford.edu, jos@ccrma.stanford.edu

ABSTRACT

This paper describes the use of a digital waveguide mesh which provides certain desirable components of the frequency response of the body of an instrument. An application for the violin is illustrated, showing that meshes can be designed to have a modal distribution which is psychoacoustically equivalent to the resonances of the violin body at high frequencies.

1. INTRODUCTION

The body of a musical instrument represents a complex resonator whose filtering contributes strongly to the characteristic timbre of the instrument. In a previous paper [2], a waveguide mesh was used to model the high-frequency resonances of a violin body in order to improve the quality of violin sound synthesis. It was shown that the mesh provides a high-quality alternative for modeling body resonators having nonlinear excitations (which precludes use of the commuted synthesis technique [7]). In this paper, we extend our previous results with a more careful study of the psychoacoustical match in each Bark band. Furthermore, waveguide meshes of different dimensions were synthesized in order to obtain resonances with a distribution which is perceptually equivalent to the original at high frequencies.

2. PSYCHOACOUSTIC CHARACTERIZATION OF HIGH-FREQUENCY MODES

When analyzing the high-frequency content of complex resonators, we notice that the mode density is greatly increased, which forces us to deal with *bands* of high-frequency mode distributions. In fact, it is impractical in most cases to attempt to resolve individual high-frequency modes, because they are so heavily overlapped in frequency, where many modes occur per mode-bandwidth. To conform to the characteristics of human hearing, we form bands by summing the spectral power in each *critical band of hearing*, as defined by the Bark frequency scale.[1] Within each band, statistically similar mode distributions can be expected to sound musically equivalent.

The relevant parameters in each critical band include mode spacing (in frequency), bandwidth, amplitude, and phase. These parameters may be characterized statistically in a variety of ways, such as by a simple average (mean) value in each critical band, or as a per-band distribution having, e.g., a mean and variance. There appears to be very little literature available characterizing how faithfully such distributions should be preserved in order to retain psychoacoustic equivalence.

Preliminary simulations were made to examine how many modes are necessary to reach a psychoacoustically equivalent density in each critical band. Sinusoidal components uniformly spaced on a logarithmic scale and centered in a Bark band were synthesized with random phases and amplitudes and then summed together. The amplitudes were scaled accordingly, in order to have the same average power in the band regardless of the number of modes. An exponentially decaying envelope was applied to the combination for uniform decay at all frequencies.

Experiments show that in a critical band as few as 3 to 5 modes are sufficient for achieving the effect of maximum density in a critical band, depending on the phase values of the sinuosoids.

3. WAVEGUIDE MESH DESIGN WITH APPLICATION TO THE VIOLIN

3.1. Characteristics of complex resonators

Bodies of musical instruments are complex resonating systems which strongly influence the timbre of the instrument itself. For example, the timbre of a musical feature such vibrato in a bowed violin is determined by the effect of its body, which modifies the spectral content of the note played as its harmonics are filtered through the body response. Since the detailed high-frequency response of a complex resonator contributes to timbre, sound synthesis quality should be highest when the entire frequency response of the resonator is simulated in a perceptually accurate way.

3.2. Characteristics of waveguide meshes

In order to implement the results presented in the previous section, we chose a waveguide mesh structure [8], which has the desirable properties of increasing modal density with frequency and at the same time is relatively inexpensive to implement.

As described in [2], the goal of the mesh design is to find a mesh having an impulse response which sounds identical to the high-frequency residual obtained after removing the low-frequency resonators.

Figure 1 shows the two-stage exponential decay in the violin body response. The slower, second decay contains 6-8 main frequency components from 115 Hz to 1600 Hz that can be extracted from the measured response [4]. The residual impulse response, containing the quickly decaying components, then becomes the target of a waveguide mesh design.

[1] A more accurate choice would be the ERB scale [5], and we plan to compare results for the ERB in later work. The Bark scale can be viewed as a scale based on slightly overestimated critical bandwidths.

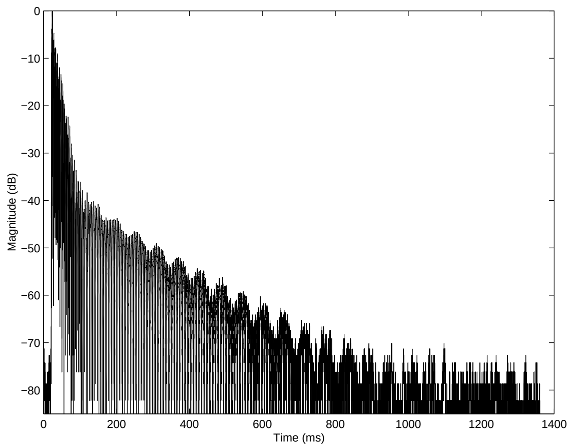

Figure 1: *Log magnitude time response showing the two-stage decay of the violin body impulse response. The noise floor is reached after 820 ms or so.*

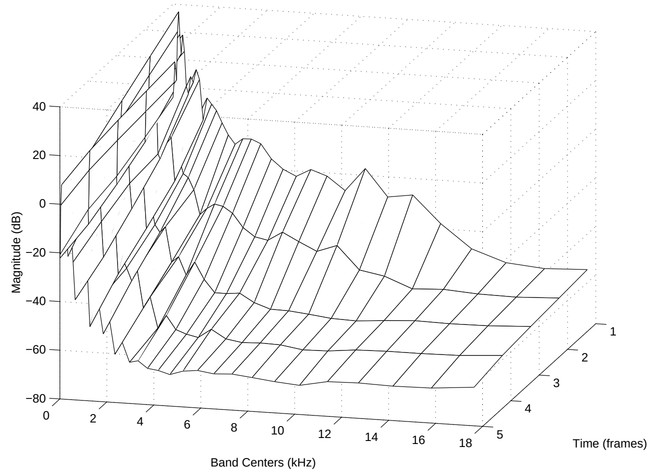

Figure 2: *Bark-Summed Energy Decay Relief of a violin body impulse response, using 30 ms analysis windows with 50% overlap.*

Within each band, we wish to match certain statistics of the mesh response to those of the instrument body response. By starting with mesh dimensions comparable to those of a real violin body, we may expect the mesh to spontaneously have a similar mode spacing at high frequencies where the air modes dominate [1]. Additionally, the mode phase being sufficiently randomized in any mesh, it is not necessary to explicitly worry about setting it. The within-band amplitude distribution is taken to be the natural amplitude fluctuation obtained when summing a set of identical modes at center frequencies chosen randomly according to the appropriate distribution. We therefore expect the choice of proper mesh geometry to take care of this as well at high frequencies.

3.3. Estimation of average decay in a critical band

To match the mode bandwidths, we analyze the violin body impulse response over a Bark frequency axis to determine the average decay time for each "band of modes" in the high-frequency response.

The Energy Decay Relief (EDR) at time t and frequency ω is defined as the sum of all remaining energy at that frequency from time t to infinity. It is a frequency-dependent generalization of Schroeder's Energy Decay Curve [3]. We prefer it over the more usual short-time Fourier power spectrum because it deemphasizes beating decay envelopes due to closely tuned coupled modes (which occur often in acoustic measurements of resonating bodies). This facilitates estimating decay times for ensembles of resonators which are being characterized statistically.

The results of summing power in each critical band can be seen in Fig. 2. A line is fitted to the successive values for each band to estimate the average decay and initial amplitude levels for the modes in the band.

3.4. Bowed violin synthesis model

We applied our mesh model to the waveguide bowed string model described in [6], as proposed in [2]. In this bowed violin model, second-order resonant filters model the low-frequency resonances, and a waveguide mesh is used to approximate the dense modes of the violin body at higher frequencies.

3.5. Mesh geometry

We chose the simple rectilinear mesh on the grounds that we only need a psychoacoustically equivalent distribution of high-frequency modes, so that the greater accuracy of the triangular and warped triangular meshes did not justify the additional programming overhead. Since our approach is a statistical characterization of the mesh properties, the dispersion error is not taken into account and presumably does not significantly influence the perceptual results.

In order to provide the most accurate asymptotic mode density, a 3D waveguide mesh is initially chosen to correspond to a physical box with dimensions $35.5 \times 21.0 \times 3.0$ cm. For a sampling rate of 44.1 kHz, assuming a soundspeed $c = 344$ m/s, these dimensions translate to a rectangular mesh that is approximately $26 \times 16 \times 2$ samples along each edge. (Using the results of the von Neumann stability analysis [8], a physical length d is converted to samples using the formula $N_d = df_s/(c\sqrt{3})$ for the 3D rectilinear mesh.)

3.6. Simulation results

3.6.1. Determination of crossover frequency

Simulations were carried out to determine the crossover frequency at which the synthesis model becomes perceptually equivalent to the original body impulse response. We define the crossover frequency as the upper limit of resonant modes modeled using biquads and the lower limit at which the mesh output models the high frequency modes.

To construct the most accurate mesh impulse response with regard to the features extracted from the violin impulse response data, we first analyzed the average decay rate of the original impulse response as described in subsection 3.3. We then fit this data to our waveguide mesh impulse response.

To determine the lowest crossover frequency we mixed a lowpassed violin impulse response and a highpassed synthesized mesh impulse response and compared the resulting sound to the original violin impulse response.

Tests performed show that the lowest crossover frequency at which the highpassed mesh/lowpassed violin impulse mixture is indistinguishable from the violin body impulse reference is around 1400 Hz. At this level, the number of biquad resonators needed to capture the low frequency modes is about 9. This is more computationally efficient than the crossover frequency used in the [2], which used 13 biquad resonators which went up to a crossover frequency of 3200 Hz. At a crossover frequency that was lower than the threshold, there was an audible difference in the pitch level and timbre of the test signal, which could not adequately match the violin impulse response in a perceptually equivalent way.

These tests show that it is perceptually possible to use a reduced number of low-frequency resonators and still provide high quality results. Connecting the mesh to these resonators produces a high quality synthetic violin impulse response whose high frequency content is shown in Fig. 4. The corresponding time domain signal is shown along with that of the original violin impulse response in Fig. 3. The more pronounced beating of the highpassed waveguide mesh output is effectively masked by the strong, long-ringing body modes present in the original violin body impulse below 1400 Hz.

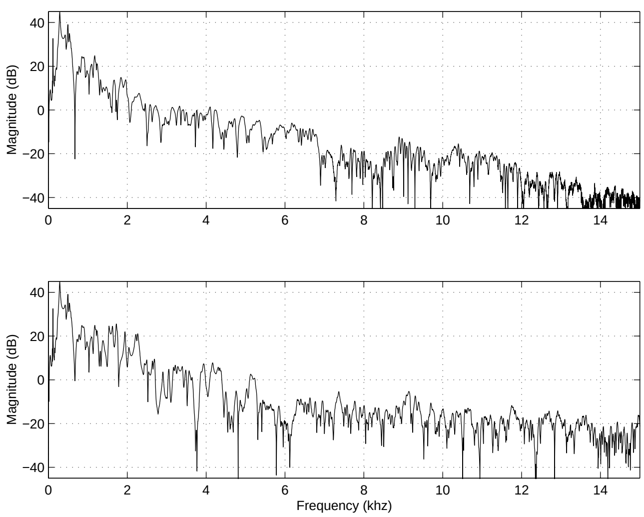

Figure 3: *Frequency responses of the original violin body impulse response (top) and a highpassed synthesized waveguide mesh impulse response plus a lowpassed violin body impulse response (bottom), with the crossover frequency at 1400 Hz.*

3.6.2. Examining smaller mesh dimensions

Since perceptual mode saturation in a critical band can be realistically achieved by a small number of resonators, it is theoretically possible that even smaller dimensions of waveguide meshes would have a distribution of mode phases and amplitudes which can satisfy the perceptual criteria for our synthesis model described above.

Dimensional reductions of the 3D mesh at 44.1 kHz were also tested in the same way to determine their lowest crossover frequency, if any. At a lower sampling rate of 22.05 kHz, the desired physical dimensions translate to the 3D-mesh dimensions $13 \times 8 \times 1$ samples ($35.9 \times 20.3 \times 3.1$ cm). With only 1 sample in the z direction, we expect that a 2D mesh at 22.1 kHz can behave similarly to its 3D counterpart, at 16×9 samples (35.3×19.9 cm).

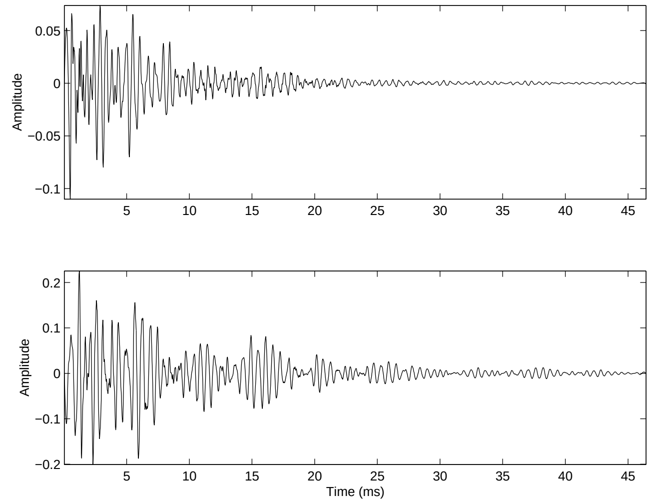

Figure 4: *Time responses of the violin body impulse response (top) and synthesized waveguide mesh impulse response (bottom) above 1400 Hz.*

(For the 2D rectilinear mesh, the correction factor for the speed of sound is $\sqrt{2}$, as opposed to $\sqrt{3}$ in the 3D case.) Simulation results show that these two meshes have about the same crossover frequency as the 3D, 44.1 kHz mesh, at around 1400 Hz.

Using geometries which do not closely follow the physical proportions of the violin edges may result in mode distributions which are less optimal for simulating high-frequency violin body resonances. For example, a $18 \times 13 \times 2$ sample 3D mesh at 44.1 kHz has its lowest crossover frequency around 1600 Hz.

4. CONCLUSIONS

In this paper, we expanded on the use of the digital waveguide mesh as a complex resonator at high frequencies. Results show that the mesh can behave in a perceptually equivalent way as the high-frequency portion (above 1400 Hz) of the body of a complex resonator. Simulations were carried out which indicate the feasibility of using smaller-than-physical waveguide meshes as a statistical spectral modeling tool for simulating high-frequency modes in resonating bodies.

5. REFERENCES

[1] L. Cremer. *The Physics of the Violin*. MIT Press, Cambridge, MA, 1984.

[2] Patty Huang, Stefania Serafin, and Julius O. Smith. A waveguide mesh model of high-frequency violin mode resonances. In *International Computer Music Conference*. Computer Music Association, 2000.

[3] Jean-Marc Jot. An analysis/synthesis approach to real-time artificial reverberation. In *Proc. Int. Conf. Acoustics, Speech, and Signal Processing, San Francisco*, pages II.221–II.224, New York, 1992. IEEE Press.

[4] Matti Karjalainen and Julius O. Smith. Body modeling techniques for string instrument synthesis. In *Proc. 1996 Int.*

Computer Music Conf., Hong Kong, pages 232–239. Computer Music Association, Aug. 1996.

[5] Brian C. J. Moore. *An Introduction to the Psychology of Hearing*. Academic Press, New York, 1997.

[6] Stefania Serafin, Julius O. Smith, III, and Jim Woodhouse. An investigation of the impact of torsion waves and friction characteristics on the playability of virtual bowed strings. In *Proc. IEEE Workshop on Appl. Signal Processing to Audio and Acoustics, New Paltz, NY*, New York, Oct. 1999. IEEE Press.

[7] Julius O. Smith. Efficient synthesis of stringed musical instruments. In *Proc. 1993 Int. Computer Music Conf., Tokyo*, pages 64–71. Computer Music Association, 1993. available online at http://www-ccrma.stanford.edu/˜jos/cs/.

[8] S. A. Van Duyne and Julius O. Smith. Physical modeling with the 2-D digital waveguide mesh. In *Proc. 1993 Int. Computer Music Conf., Tokyo*, pages 40–47. Computer Music Association, 1993.