Reproduction and Compression of Violin Tones through Spectral Interpolation Synthesis

Chris Niessl
Advisor: Roger Dannenberg

Introduction:

Synthesizing instrument performances have a variety of useful applications, from providing a computerized 'session musician' for improvisation and composition, to a form of source based compression for audio data. However, synthesizing instruments accurately is a challenging problem; merely splicing recorded samples for pitches ignores variation in timbre and dynamics that are affected by the style of play. In addition, using signal processing techniques such as spectrum analysis can generate large amounts of data.

In 1998, Spectral Interpolation Synthesis was shown to be successful in replicating trumpet notes played in succession by successfully capturing greater features of the sounds produced by the instrument than prior synthesis models. [1] In this paper, SIS is used to attempt synthesis and reproduction of violin tones. However, in attempting to reduce the control rate for reproducing the sound, we notice several limitations that smoothing harmonics has on the original sound. Thus, we introduce splitting synthesis into two parts: A control model using SIS, and a noise model, to recapture noise lost from smoothing, in order to more accurately reproduce violin tones.

Methodology:

Initial testing was conducted on a single violin tone, a middle G pitch (G4) with a fundamental harmonic frequency of approximately 202hz. Phase Vocoder Analysis was used through the SNDAN software package [2] to obtain amplitude and frequency information for all multiple of the fundamental harmonic (202hz, 404hz …) up to the 22050hz, above the Nyquist rate for human audio perception. This resulted in a total of 1186 data points for 109 harmonics over 3 seconds of the sound. This information was then used to re-synthesize the original tone, through SIS, and used as a reference point of comparison against lower data rate signals.

Lower data rate signals were then obtained by encoding all the amplitude points for a harmonic as a new signal, and then re-sampling that signal at approximately 30hz. Each re-sampled harmonics signal was then recombined to form a control signal file. Spectral Interpolation Synthesis was then used to reconstruct the violin sound from the control signal, and compared against the reference point.

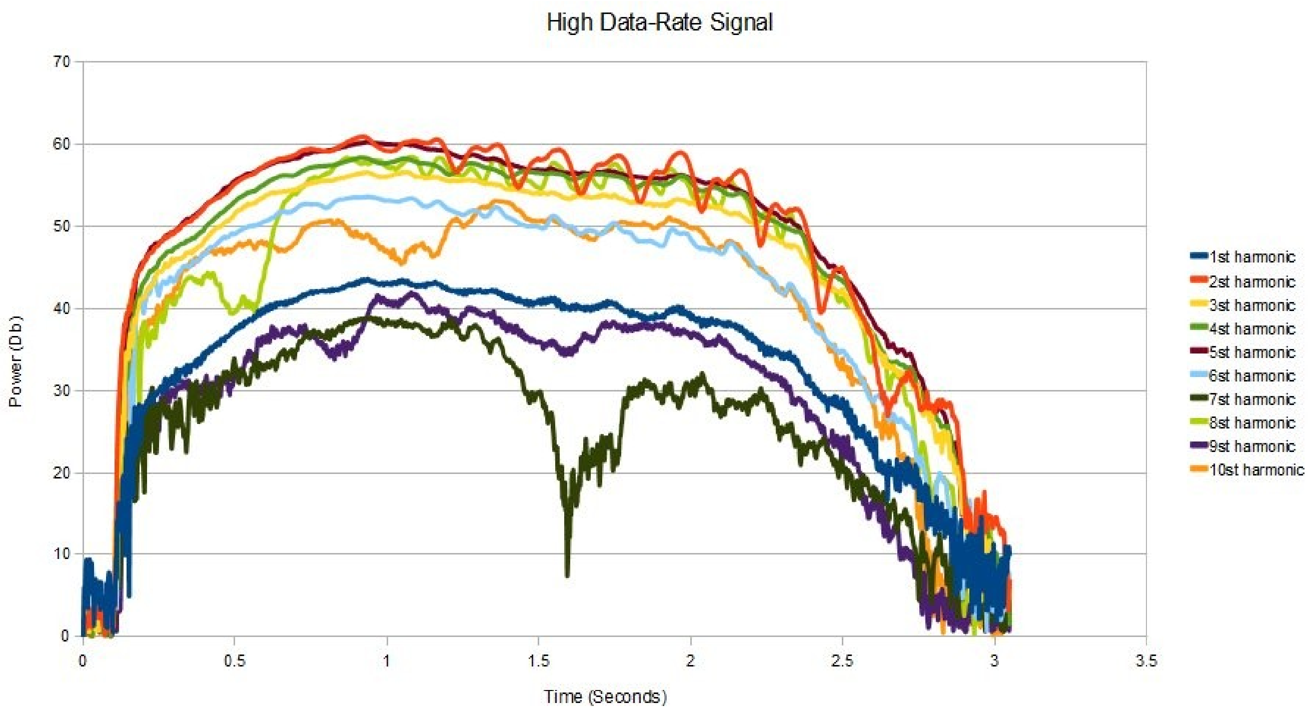
While SIS produced a very good approximation of sound, some finer details from the original tone (Such as scratches due to interactions between bow resin and string), were lacking in the tone due to the smoothing caused by re-sampling the sound at a lower rate. In order to compensate for the lost sound, additive synthesis was used to combine the tone reproduced via the control signal file with another tone produced through modeling noise in the original tone.

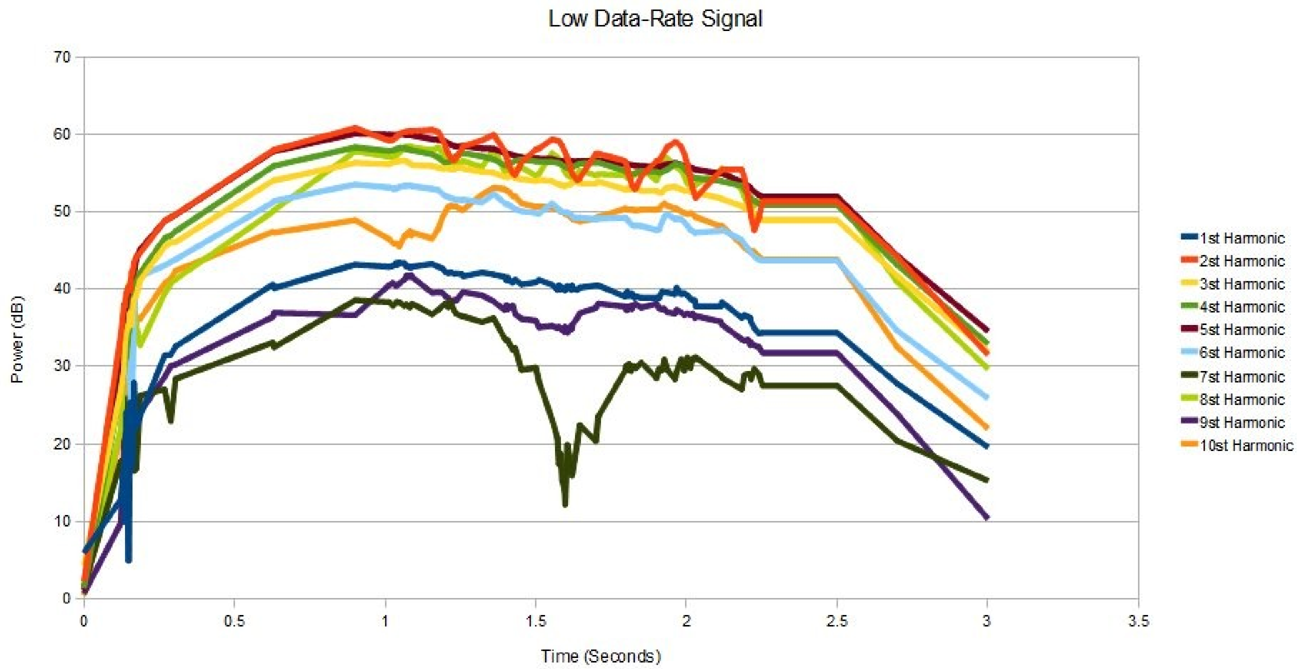
The Noise model was created by analyzing differences between the control signal and original tone for several pitches. Noise for each pitch was analyzed by capturing the differences in amplitude data points for each harmonic between the reference signal and the control signal. Spectral analysis was then taken on each harmonic via FFT in order to determine if the noise signal had frequency dependent properties.

Analyzing Reproduced Tones:

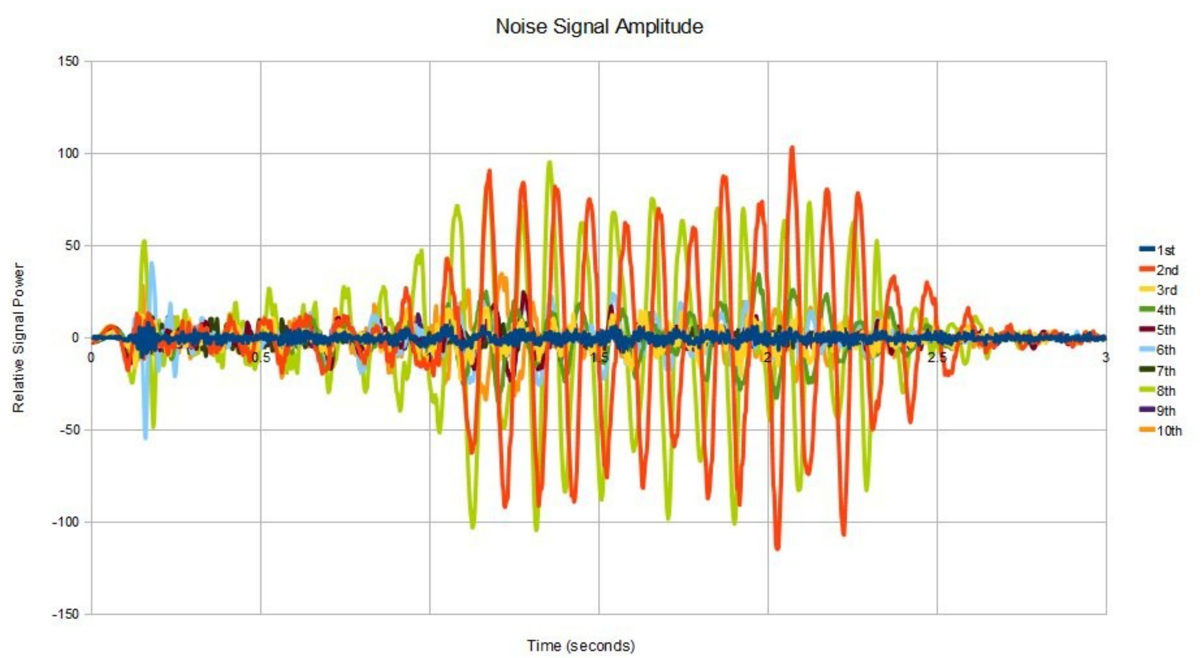
After listening to the original sound, a high data-rate synthesized tone, and low data-rate synthesized tone to listen for differences in the sound, amplitudes for each harmonic in the synthesized versions were plotted over time in order to determine if, and what kind of control envelopes could be used to reduce the amount of information needed for constructing pitch signal tables, and to identify if any features were missing from the low data-rate signal against the reference high data-rate signal.

Below are two plots of the first 10 harmonics of each signal, with fundamental frequency of 202hz. The top graph consists of plotting all data points from the reference signal, captured at 404 data points per second (404hz) while the bottom graph consists of the low data-rate control signal plotted at 30hz. Each data plot represents the volume or power level in decibels (dB) for each individual harmonic.





When comparing the reference synthesized tone against the original tone, the two sound nearly indistinguishable, suggesting that SIS can reproduce the sound, given a sufficiently high data-rate (around 400 samples per second). While listening to tone synthesized through the low data-rate control signal without noise, the tone resembles a violin, however differences become a bit more noticeable when compared against the other two tones. The following shows the difference in signal power between the high and low data-rate signals. A relative signal power of 100 corresponds to approximately a 20dB difference in volume level.



As you can see, several of the harmonics (the 2nd and 8th most prominently) display periodic variation from the control signal. When the noise is re-synthesized through SIS by modifying the original low-rate control signal with modifications from the signal above, a richer tone is achieved, though still not an indistinguishable replication of the original violin sound.

Conclusions and Future Work:

By successfully replicating the original sound with the reference signal, and achieving a significantly accurate approximation with combined noise and control signals, we've demonstrated that SIS is viable for modeling and synthesizing violin tones. Further work can now be done analyzing variations in pitch and gesturing of the violin in order to provide complete synthesis for performances and score.

Acknowledgments / Sources:

In addition to my thesis advisor, I'd like to thank Yuxiang Liu for providing assistance in setting up SNDAN for analyzing the violin tones.

[1] Dannenberg, Pellerin, and Derenyi. ``A Study of Trumpet Envelopes,'' in *Proceedings of the International Computer Music Conference*. San Francisco: International Computer Music Association (1998) pp 57-61.

[2] Rob Maher, James Beauchamp “SNDAN, programs for sound analysis, resynthesis, display and transformation” University of Illinois at Urbana-Champaign (UIUC).
http://dream.cs.bath.ac.uk/software/sndan/sndan.html