Modeling Piano Tones

Joseph Markel and John Puckett
Center for Computer Research in Music and Acoustics (CCRMA)
Department of Music, Stanford University
Stanford, California 94305

1. ABSTRACT

The complex spectrum of the piano tone demands that several methods be considered in attempting to recreate the tone with a maximizing degree of perceptual accuracy. In modeling the piano tone, our approach endeavor to incorporate the most advantageous aspects of several synthesis techniques in an effort to maintain both economy and accuracy throughout the range of the piano.

2. INTRODUCTION

In designing an "instrument" to reconstruct the piano tone, the qualities of timbral and flexibility were deemed most important. Both Additive Synthesis and Frequency Modulation Synthesis were chosen as primary means of reconstruction (see sections on Additive Synthesis and Frequency Modulation Synthesis). While modeling piano tones, it is essential that the instrument user has dynamic control over the most influential parameters. Specifically, the following parameters are required for a sufficient degree of control: the number of additive oscillators; the number of FM oscillators; the amplitudes and frequencies envelope controls over each oscillator or oscillator pair; dynamic control over modulation oscillators; dynamic placement of both additive- and FM-based amplitudes; and FM oscillator pairs within the frequency bandwidth; relative amplitudes control over both additive and FM oscillator pairs; stretch factor control over partial spacing as well as to harmonize the stretch factor applied.

Each of these parameters was incorporated into the user interface, giving an enormous amount of flexibility to the instrument. The user can therefore incorporate either or both of the synthesis techniques at any location within the spectral range, thus allowing a "building-block" type of synthesis using both additive and FM synthesis.

2. DATA REDUCTION TECHNIQUES

Two types of data reduction analysis were employed on the piano recordings: the first, a spectral editor developed by John Strawn at CCRMA, was an interactive FFT analysis with fixed frequency bins (EMERGE). The second method, a peak finding method to track the exact stretched frequency and amplitude of each partial (PASHEL), developed by John C. Smith at CCRMA and supported by Digital First, Inc.

Each method of data reduction analysis offered its respective highlights. EMERGE uses an incredible FFT on its own and is an exact copy of the sound it detects at each fixed frequency bin. It is able to capture the initial amplitude spike caused by the hammer strike. This attack is of vital importance in the perception and recognition of the piano tone and can be seen in the spectral excerpt shown in Fig. 1.

Figure 1: First 10 harmonics of C3 (440 Hz) taken from EMERGE. The first 10 partials are sufficiently close to harmonics.

Because EMERGE assumes exact harmonic frequencies, it is not able to locate the positions of the stretched partials.
inherent in the piano tone. PARISIAN, using inharmonic transfer analysis, is designed to make each partial at its actual stretched frequency. This provides not only proper frequency placement, but also the actual amplitude function of each partial as it appears. PARISIAN does, however, have difficulty in capturing the quick attack of the hammer strike.

The advantages of each method were utilized in order to gain an accurate data reduction analysis of the piano. Amplitude and frequency functions were approximated to fit the analysis output. Furthermore, it was noted that one "typical" frequency function could be applied to every partial without loss of perceptual accuracy, reducing by 50 percent the total number of functions required.

4. ADDITIVE SYNTHESIS

An Additive Synthesis approach, employing both frequency and amplitude envelopes derived from the data reduction techniques discussed above, provides results over each and all of the inharmonic components comprising the piano tone. Additive synthesis was used first to control over the lowest harmonics of each piano tone, as the amplitude structure of the lower harmonics does not particularly lend itself to the "spectral-group" type of the higher harmonics, a structure which can be efficiently synthesized using frequency modulation synthesis techniques.

An investigation into the importance of the characteristics of the piano partials necessitated a piano "instrument" in which both the range and degree of partial stretching could be controlled. Experimentation was done on an S1 [19.35 Hz], as it provided a large frequency spectrum which could be realized (see discussion on Frequency Modulation).

Tones synthesized with completely harmonic spectra resulted static and modified, although not quite to the degree expected. These tones lacked an essential "metallic" quality which is necessary, especially in lower piano tones. To approximate the observed frequency stretching of the piano harmonics, the following stretch formulas, derived by Julian Smith at CCRI as using PARISIAN on C2 (262 Hz), was incorporated into the piano model:

\[ f(t) = v(t) + 0.000053(\pi^2) \]

where \( n \) is the harmonic number and \( f(t) \) is the fundamental frequency (accurate for the first forty partials). Using this formula and the frequency range of C3 (approximately 100 partials), and using the full frequency range of equalized tones as it resulted on a "stretched" piano having a "true" metallic quality. However, this approach was helpful in an attempt to push a stretch compressor. A compromise was reached using the stretch formula for the first forty partials and then placing modified spectral groups or specific additive partials at locations indicated by the data reduction from PARISIAN. This compromise provides good sound results, "brightening" the completely harmonic piano and "softening" the stretched metallic piano - making the stretch preprogrammable but not overbearing. Thus, the stretching of the piano harmonics plays a crucial role in realistic piano modeling, creating a more dynamic sound, especially for piano tones at the very lowest frequency range.

The model also provides the ability to isolate any range of harmonics. This feature allows any direction of the piano tone into simple quantities which can then be accurately recreated using either additive synthesis or frequency modulation synthesis, depending on which seems more appropriate.

While additive synthesis produces very effective results, due to its cost in both computational elements and data storage, its use was restricted to the lower 10-12 partials and for space "patch-work" among spectral peaks, which could not be completely modeled using frequency modulation synthesis. The piano range above 53 (880 Hz), however, can be modeled economically using strictly additive synthesis, due to the fact that there is an average of only 7 partials throughout this range at a sampling rate of 22.05 kHz.

5. FREQUENCY MODULATION SYNTHESIS

Due to the "cost" disadvantages of additive synthesis, a "shaper" synthesis technique is desirable. Frequency Modulation (FM) Synthesis provides such an environment in which complex spectra can be synthesized using relatively few computational elements.

Following the fundamental rules of FM synthesis, a model was made of each "spectral group" using a simple FM patch, consisting of an oscillator pair with the output of the second oscillator (see modulation) fed into the input of the first oscillator (the carrier). The oscillator parameters were determined using the formula for the placement of side-band
components in a simple FM oscillator pair. This formula gives integer multiples of the modulating frequency above and below the carrier in potential placements for sideband components [Bowering, (1972)].

\[ C = f \cdot \beta \cdot j \cdot \lambda \]

where \( C \) = carrier frequency, \( f \beta \) = modulator frequency and \( j \cdot \lambda = \ldots \).

The actual power of the sideband components will then be determined by the "index" of the modulator, a measure of the modulator's strength. The index in a convenient form of expressing the deviation of the modulator; the actual deviation may be obtained by multiplying the index by the modulating frequency. The piano instrument was then considered to allow the unendormingly well-formed frequency peaks by specifying a carrier frequency, an amplitude envelope, and a modulating index. Again, as in the additivity system, a value of 4 is needed to ensure that a spectral peak's carrier frequency was placed in the proper location.

FM synthesis also has its disadvantages, preventing components which had to be made. First, where FM is used to create a "spectral peak", one must carefully control envelope width over each individual harmonic (partial) of the spectral peak. Specifically, the amplitude envelope of the carrier rectangle is the primary basis for the amplitude envelopes of the sideband components. In addition, the index of the modulator also contributes to the relative amplitudes of the sideband components. While this remains a theoretical disadvantage, in practice it was found that spectral composition within each spectral group played a much greater role than did actual amplitude control over individual partials. In fact, the majority of the partials that needed to be synthesized using FM were partial existing primarily in the attack portion of the piano tone - partials with quick exponential decays. Thus, for those groups of "attack" partials, amplitude control was more important only to the extent that a proper "spectral mix" had to be obtained at the onset, allowing the partial groups to decay in a house-grown manner governed by the envelope of the carrier generator.

A second disadvantage in using FM synthesis is that it creates a musical spectrum around the carrier. Unfortunately, not all spectral groups are "natural" symmetrical, possessing certain asymmetry, shown in Fig. 2.

Figure 2. Spectral Group taken from F2.

These asymmetries, however, could not be readily perceived, especially when considered in the context of the entire piano tone.

In an effort to model the partial spacing of a spectral group using FM, two types of spacing lead to be considered. First, it had to be realized that the carrier frequency of the carrier was placed at a proper extended partial location. A second and less obvious concern was the partial spacing within a spectral group. Perhaps the most significant disadvantage of FM synthesis occurs due to the equal spacing of the sideband components which FM produces. This spacing (determined by the modulating frequency) is not able to accommodate the observed "stretching" of the piano harmonics. A compromise was obtained by using the frequency distance between the carrier frequency of a potential FM group and its adjacent lower partial, then using this distance as the modulating frequency for that FM group. By doing this, the partials of an FM model (where carrier frequency is placed at the proper extended harmonic location) will be able to approximate the increasing harmonic stretch, seen in Fig. 3.

Figure 3. Partial Approximation.

265 ICME 8E Proceedings
Since the harmonic stretch is always positive as frequency is increased, using this difference will always produce upper side-band components which will choose approximate the partial's location. This method is preferred to an "averaging method", which would use as a modulating frequency the average of the upper and lower partials around the carrier frequency. The former method ensures that the lower side-band components will not deviate too far away from the actual lower side-band components' spacing.

5. INTERPOLATION

Another problem exists in obtaining a homogeneous sound quality throughout the entire range of the keyboard. When the amplitude functions of an analyzed tone are applied to a range of frequencies other than that of the original tone, the effect is much like that heard in digital sampling when re-sampling is used to plot a tone at a frequency other than that of the originally sampled frequency. This problem arises due to the fact that the piano exhibits a resonantly changing system across the range of the keyboard. To solve this problem in sampling, one must sample every third semitone along the keyboard in order to maintain a smooth spectral transition without any "liminal breakpoints". This method, however, is quite costly in terms of the data space required. In this synthesis, functions were derived every fourth semitone instead of every third, thus cutting the analysis down to one-quarter of the time originally thought to be required.

To obtain a smooth spectral and timbral transition throughout the chromatic scale from one octave to the next, a linear interpolation scheme was applied which eliminated all breakpoints on the keyboard. The amplitude functions for fundamental frequencies lying within successive octaves were derived using a weighted average of the adjacent octave amplitude envelopes. Successive octaves beginning with $\frac{3}{2}$ (1125 Hz) through $\frac{1}{2}$ (2000 Hz) were chosen as the breakpoints for analysis. The following equation was used to determine the relative weights of the surrounding functions to be interpolated:

$$x = \frac{\text{freq}_1 - \text{freq}_2}{\text{freq}_1 - \text{freq}_0}$$

where $\text{freq}_1$ is the frequency of the $\text{AE}$ above the desired pitch, $\text{freq}_2$ is the desired pitch, and $x$ is the relative weight.

The formula for the interpolated function, $F(x)$, becomes:

$$F(x) = F(x1)(1 - x) + x F(x2)$$

where $F(x1)$ is the amplitude function of $\text{AE}$ above the desired pitch, $F(x2)$ is the amplitude function of $\text{AE}$ above the desired pitch, and $x$ is the relative weight calculated above.

The most difficult problem in interpolating between octaves lies in the attack of the piano. The amplitude functions for the piano exhibit a large "spike" (see Fig. 4) lasting on the average about 100 milliseconds.

Figure 4. Interpolation without curve checking.

To overcome this, an algorithm was employed to check the probability of a spike and eliminate the double spike by gliding the first spike towards the second spike at the frequency increase between successive octaves. See Fig. 5.

Figure 4. Interpolation with amplitude function curve taken from 51 and approximated using EAKMMS.

This distinctive spike of the hammer hitting the string does not coincide in every case precisely with the spike of its corresponding octave. A direct linear interpolation of the two skewed spikes yields the resulting "double epiled" function depicted in Fig. 5.
Figure 6. Direct interpolation with simple blending.

The selection for eliminating double spikes also eliminates the creation of unnecessary points in the resulting interpolated function, thereby increasing computational economy and efficiency.

7. HAMMER NOISES

Filtering techniques tend themselves particularly to the attack portion of the piano tone, which contains a highly complex frequency spectrum that is sometimes difficult to reproduce accurately among other synthesis techniques. First, the sound of a hammer striking tuned piano strings was recorded. A filter was created in 99.9% of the hammer strike. The hammer strike is then recreated by driving the filter with a noise source, passing only those frequencies modulated at the hammer strike through the filter, creating a replica of the frequency spectrum found in the recording.

8. CONCLUSION

While attempts at piano synthesis have been numerous, this approach does not single out any particular synthesis method; rather, attempts to obtain both realistic and "unrealistic" sounds through the combination of various synthesis techniques, utilizing the optimal qualities which each technique has to offer.

9. ACKNOWLEDGEMENTS

We would like to thank several people who have all contributed to this ongoing research effort including Xavier Serra, Johann Wolfmann, Dong Kende, and John Chowning. We would also like to thank those who helped in the editing process (especially Julian, Janine, and Donna). We would particularly like to thank those personas who have provided the various analysis tools, which are in the research effort would not have been possible. These personas include: Julian Smith (Pascal), John Strawn (George), and Bill Schottstaedt for his infinite programming help. We extend our greatest appreciation to Julius for his genuine concern and guidance.

References


267

ICMC 86 Proceedings