

# WAVETABLE INTERPOLATION OF MULTIPLE INSTRUMENT TONES

*Jonathan Mohr*  
University of Alberta  
Augustana Faculty

*Xiaobo Li*  
University of Alberta  
Department of Computing Science

## ABSTRACT

Previous work on multiple wavetable interpolation synthesis has focused on modeling single instrumental tones, using a different set of basis spectra for each. However, the amount of space used to store the wavetable banks needed to resynthesize many different tones would be reduced if at least some of the wavetables were selected for their general-purpose use in synthesizing multiple tones. This paper presents the results of our research on matching and synthesizing 198 tones played at various pitches by a variety of different instruments using common wavetable banks. We introduce two new techniques: the grouping of tones by pitch so that the highest expected partial frequency will be less than the Nyquist frequency, and the construction of shared wavetable banks by applying a clustering algorithm to all the breakpoint spectra of all the tones in a given group and selecting the spectrum nearest the centroid of each class as a basis spectrum.

## 1. INTRODUCTION

*Multiple wavetable interpolation* [10, 3] is a form of music analysis/synthesis. It begins by converting a digital waveform to the frequency domain by a short-time Fourier transform and reducing it to a set of shared breakpoints [4] by piecewise linear approximation (PLA) of the spectral envelopes of its harmonics. A single weighted-average frequency differential is also computed and stored for each breakpoint, since wavetable interpolation requires that the corresponding harmonics of the wavetables involved in the interpolation be in phase [10].

Next, a number of *basis spectra* are selected to comprise a *wavetable bank*.<sup>1</sup> Typically, these basis spectra are selected from the breakpoint spectra of the analyzed tone, but they could be selected by other means, including spectral principal components analysis (PCA) [9], a genetic algorithm [2, 5], or by hand-selection [10, §2.2].

The spectrum at each breakpoint of the spectral envelope of the tone is then matched by determining weightings for a small number (typically, 2 to 5) of basis spectra selected from the wavetable bank, and the sound is resynthesized using multiple wavetable additive synthesis [2, 5]

<sup>1</sup>The basis spectra are collectively referred to as a wavetable bank, although for the purposes of breakpoint matching they are initially represented in the frequency domain as vectors of harmonic amplitudes. At synthesis time, each vector is converted to an actual wavetable—a table of the time-domain amplitude values of one cycle of the waveform—for use by a table-lookup oscillator.

by interpolating between the weightings for each wavetable at consecutive breakpoints.

A different set of wavetables can be selected for use at each breakpoint, subject to the restriction that a wavetable that is used at one breakpoint but not at the next must be faded out and a wavetable that comes into use at a particular breakpoint must be faded in, since audible clicks and spectral discontinuities would result from the sudden change of wavetables [3].

Previous work on multiple wavetable interpolation synthesis [10, 3] has focused on modeling a single instrumental tone with each set of basis spectra. However, the amount of space used to store the wavetable banks needed to resynthesize many different tones would be reduced if at least some of the wavetables were selected for their general-purpose use in synthesizing multiple tones. This paper presents the results of our research on matching and synthesizing multiple tones played at various pitches by a variety of different instruments using common wavetable banks. We introduce two new techniques: the grouping of tones by pitch so that the highest expected partial frequency will be less than the Nyquist frequency, and the construction of shared wavetable banks by clustering all the breakpoint spectra of all the tones in a given group and choosing a representative spectrum from each cluster.

## 2. MULTIPLE TONE MATCHING

A set of 198 tones played by sixteen different instruments, spanning the range from A1 to B6 by minor thirds, were selected from the *McGill University Master Samples* collection for the purpose of testing the proposed analysis-synthesis method. As indicated in Table 1, all tones of pitch classes<sup>2</sup> A♯, C♯, E and G in the chosen range as played by the bassoon (abbreviated as *bsn* in the table), B♭ clarinet (*cla*), bass clarinet (*clb*), English horn (*eng*), flute (*flt*), glockenspiel (*glk*), French horn (*hrn*), oboe (*obo*), piano (*pno*), the saxophone family<sup>3</sup> (*sax*), C trumpet (*tpt*), trombone (*trb*), viola (*vla*), string bass (*vlb*), 'cello (*vlc*), and violin (*vln*) were selected.

<sup>2</sup>All references to pitches specify sounding pitch, not written pitch.

<sup>3</sup>Because the McGill collection does not include tones spanning the full range of each member of the saxophone family (bass, baritone, tenor, alto, and soprano) but uses about an octave from each instrument such that the recorded tones span the full range of the family, the saxophones were regarded as a single instrument for the purposes of this research.

Pitch	Instrument															Count	
	bsn	cla	clb	eng	flt	glk	hrn	obo	pno	sax	tpt	trb	vla	vlb	vlc		vln
A#1	•								•	•				•			43
C#2			•						•	•				•	•		
E2	•		•				•		•	•		•		•	•		
G2	•		•				•		•	•		•		•	•		
A#2	•		•				•		•	•		•		•	•		
C#3	•		•				•		•	•		•	•	•	•		
E3	•	•	•	•			•		•	•		•	•	•	•		67
G3	•	•	•	•			•		•	•	•	•	•	•	•	•	
A#3	•	•	•	•			•	•	•	•	•	•	•	•	•	•	
C#4	•	•	•	•	•		•	•	•	•	•	•	•	•	•	•	
E4	•	•		•	•		•	•	•	•	•	•	•	•	•	•	
G4		•		•	•		•	•	•	•	•	•	•		•	•	46
A#4		•		•	•		•	•	•	•	•	•		•	•	•	
C#5		•		•	•		•	•	•	•	•	•		•	•	•	
E5		•		•	•		•	•	•	•	•	•		•	•	•	
G5		•		•	•	•		•	•	•	•		•		•	•	34
A#5		•			•	•		•	•	•	•		•			•	
C#6		•			•	•		•	•	•	•		•			•	
E6					•	•		•	•							•	
G6					•	•			•							•	8
A#6					•	•			•							•	
<b>Count</b>	11	12	9	10	12	6	12	11	21	18	11	12	13	11	15	14	198

**Table 1.** Pitches and grouping of the instrument tones selected for testing.

## 2.1. Grouping of Tones by Pitch

Since these tones have fundamental frequencies that span a broad range of frequencies, it was necessary to partition the tones into groups, each spanning a smaller range of fundamental frequencies, and to select a different bank of wavetables for each group, due to the restrictions imposed by the sampling theorem. If a spectrum were selected from a low-pitched tone for inclusion in the wavetable bank, complete with all its harmonics, and then used in synthesizing a tone at a higher pitch, the upper partials would wrap around the Nyquist frequency, creating synthesis artefacts. If all the spectra in the wavetable bank were band-limited to the frequency range between the highest expected fundamental frequency and the Nyquist frequency, then all the energy in the upper harmonics of the lower-frequency tones would be lost on resynthesis, resulting in audibly degraded tone quality. As indicated in Table 1, the sample tones, spanning five octaves, were partitioned into five groups so that more partials could be retained in the wavetable banks for the lower-pitched tones than for the higher-pitched ones. The number of harmonics retained in the basis spectra for each group are indicated in Table 2.

The scheme of selecting tones a minor third apart was used by Horner in his testing of multiple tone matching [2]. Horner tested his multiple wavetable synthesis method<sup>4</sup> on ten English horn tones, twelve trombone tones, fourteen violin tones, and an unspecified number of clarinet, saxophone, viola, and glockenspiel tones; results are given for only the first three instruments. How-

<sup>4</sup> This method does not use interpolation; it uses the same set of wavetables throughout the synthesis of a tone, with no changing wavetables.

Group	Harmonics
1	146
2	61
3	31
4	15
5	11

**Table 2.** Number of harmonics retained in the basis spectra for each group of tones.

ever, Horner does not discuss the problem of avoiding audible artefacts due to wrapping around the Nyquist frequency when using harmonic-rich basis spectra selected from lower tones in the synthesis of higher tones; the fourteen violin tones were divided into two sets of seven tones each, but this was done because “matching this extensive space of tones with just six basis spectra did not work” [2, p. 119].

Stapleton and Bass [11] grouped tones into classes according to the instruments that produced the tones, but did not group the tones by pitch; they did observe, however, that their “basis functions must be band-limited to prevent aliasing, resulting in loss of information at higher frequencies for any tone” [11, p. 318].

Beauchamp and Horner [1] clustered spectral envelopes from 15 trumpet tones according to their spectral centroid, a correlate of the perceptual attribute brightness. Tones were synthesized from the average spectral envelopes of 10 clusters using spectral centroid, RMS amplitude, and frequency differential as control functions. The method worked well for trumpet tones, but not as well for other instruments.

Hrn G3		Pno G3		Sax G3	
Time	Class	Time	Class	Time	Class
0.015	0	0.003	72	0.020	5
0.028	9	0.013	72	0.041	57
0.043	0	0.018	72	0.054	57
0.059	20	0.033	17	0.120	57
0.151	20	0.049	17	0.261	57
0.200	63	0.056	17	0.562	57
0.276	63	0.074	17	0.680	57
0.384	20	0.079	17	0.828	57
0.588	20	0.095	17	0.907	57
0.634	20	0.102	17	1.025	57
0.783	20	0.130	17	1.130	57
0.813	20	0.143	17	1.176	57
0.952	20	0.161	17	1.331	57
1.003	20	0.245	17	1.380	57
1.125	20	0.427	4	1.526	57
1.212	20	0.527	4	1.738	57
1.264	20	0.685	4	1.771	57
1.315	20	0.959	4	1.945	57
1.376	20	1.237	4	1.975	57
1.489	20	1.618	1	2.006	57
1.591	9	2.109	1	2.037	27
1.701	9	2.521	26	2.060	4
1.791	15	2.838	26	2.078	27
1.824	15	3.259	26	2.190	18

**Table 3.** The classes to which the breakpoint spectra of three example tones are assigned by clustering. Each spectrum is identified by its time index in seconds.

## 2.2. Selection of Basis Spectra

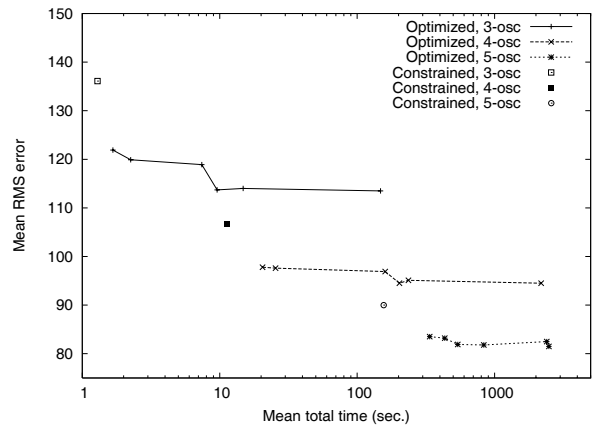
The basis spectra to be used to construct the wavetable bank for each group of tones were selected using the public domain unsupervised Bayesian classification system `AutoClass C`<sup>5</sup> on the breakpoint spectra of all the tones in that group; the spectrum nearest the centroid of each cluster was used as a basis spectrum.

To illustrate the results of clustering with respect to the spectral envelopes of individual tones, Table 3 shows, for three example tones, the class to which each breakpoint spectrum was assigned in the best clustering of the breakpoint spectra in Group 2. Each breakpoint is identified by its time index in seconds relative to the start of the tone.

The horn and sax tones illustrate a commonly occurring pattern in which all or most of the spectra from the sustain portion of a tone are clustered together, while spectra from the attack and release segments are assigned to various other clusters. The piano tone shows spectra passing through a sequence of classes as the tone decays from an extremely quick attack.

The set of basis spectra was then augmented with some hand-picked spectra, selected in order to reduce the approximation error for certain waveforms. Statistics were also gathered on the number of times each basis spectrum

<sup>5</sup> Available at <http://ic.arc.nasa.gov/ic/projects/bayes-group/autoclass/autoclass-c-program.html>.



**Figure 1.** Comparison of Horner's constrained matching with optimized multi-level exhaustive search results.

in the wavetable bank was used in matching the breakpoint spectra of the tones in each group; on the basis of these usage statistics, it was decided to remove the least-used wavetable from two wavetable banks. A summary of the final size of each wavetable bank and the ratio of the size of each bank to the number of tones in its corresponding group is provided in Table 4. The size of each bank in bytes is also indicated, assuming that harmonic amplitudes are represented as 4-byte floating-point values.

The overall ratio of wavetables to tones of 1.2 is approximately twice that of Horner's experiment [2] with multiple-tone matching of 10 English horn, 12 trombone, and 14 violin tones with five, six, and two sets of five wavetables, respectively. If Horner had divided the English horn and trombone tones into at least two groups each and the violin tones into at least three groups (compared to the four used in this research) in order to avoid artefacts due to upper harmonics wrapping around the Nyquist frequency, then the ratio of wavetables to tones would have been approximately equal to the ratio reported here.

## 3. RESULTS

Breakpoint spectra were matched using a new method that optimizes wavetable matching and oscillator assignment, the results of which have been reported elsewhere [6, 7, 8].

Figure 1 compares the average results of the constrained matching method proposed by Horner [3] with those of optimized matching [6] using 3, 4, and 5 oscillators for the tones of Group 1. In the graph, lines connect the data points for a given number of oscillators, where each data point represents a different depth of search for an initial match to the breakpoint spectra. It shows that, while the constrained matching method is faster than any of the types of optimized matching for a given number of oscillators, the error levels produced by constrained matching are significantly higher than those of the optimized matches, and are closer to those achieved by optimization with one fewer oscillators.

Group	Tones	Breakpoints	Wavetables in Bank	Wavetables per Tone	Breakpoints per Wavetable	Bank Size (bytes)
1	43	1347	48	1.1	28.1	28032
2	67	3059	74	1.1	41.3	18056
3	46	2632	64	1.4	41.1	7936
4	34	2125	48	1.4	44.3	2880
5	8	416	12	1.5	34.7	528
<b>Total</b>	<b>198</b>	<b>9579</b>	<b>246</b>	<b>1.2</b>	<b>38.9</b>	<b>57432</b>

**Table 4.** Number of basis spectra selected for each wavetable bank, and the average number of wavetables per tone, breakpoints per wavetable, and bank size for each group.

Significant data reduction can be achieved through multiple wavetable interpolation analysis using shared wavetable banks. As shown in Table 4, all five wavetable banks occupy only 57.4 kilobytes. The oscillator assignment control stream for each analyzed and matched tone consists of a single-precision floating-point value (4 bytes) for the time index of each breakpoint, another for the pitch differential at that breakpoint, and, for each oscillator, an unsigned integer (4 bytes) for the index of a wavetable in the bank and a floating-point weighting (amplitude coefficient) for that wavetable. For a 5-oscillator control stream, this totals only 48 bytes per breakpoint. The total space occupied by the 198 5-oscillator control streams plus the space required for the five wavetable banks is 513 kilobytes, which is only 0.88% of the 58.2 megabytes used by the 198 original CD-quality WAV files.

#### 4. CONCLUSION

The generalization of multiple wavetable matching can be thought of as the use of a horizontal rather than a vertical grouping of tones. In previous studies that used a common set of basis spectra to match multiple tones, the tones were those of different pitches being played by the same instrument [2] or by instruments that had been determined to have similar timbres [11]. This might be characterized as a vertical grouping of tones, since the tones ranged from low to high in pitch, but did not range across different instruments or groups of instruments. In the current study, a horizontal grouping of tones was used instead—grouping together all the tones within a narrow pitch range from across all the instruments being considered—in order to address directly the Nyquist limit while generalizing the multiple wavetable technique across many different instruments.

#### 5. REFERENCES

[1] J. W. Beauchamp and A. Horner. Wavetable interpolation synthesis based on time-variant spectral analysis of musical sounds. In *Ninety-eighth Convention of the Audio Engineering Society*, Paris, 1995. Audio Engineering Society, New York. Preprint 3960.

[2] A. Horner. *Spectral Matching of Musical Instrument*

*Tones*. PhD thesis, University of Illinois at Urbana-Champaign, 1993.

- [3] A. Horner. Computation and memory tradeoffs with multiple wavetable interpolation. *Journal of the Audio Engineering Society*, 44(6):481–496, June 1996.
- [4] A. Horner and J. Beauchamp. Piecewise-linear approximation of additive synthesis envelopes: A comparison of various methods. *Computer Music Journal*, 20(2):72–95, Summer 1996.
- [5] A. Horner, J. Beauchamp, and L. Haken. Methods for multiple wavetable synthesis of musical instrument tones. *Journal of the Audio Engineering Society*, 41(5):336–355, May 1993.
- [6] J. Mohr. *Music Analysis/Synthesis by Optimized Multiple Wavetable Interpolation*. PhD thesis, University of Alberta, Edmonton, Alberta, Canada, 2002.
- [7] J. Mohr and X. Li. Computational challenges in multiple wavetable interpolation synthesis. In P. M. A. Sloot et al., editors, *Computational Science—ICCS 2003, International Conference, Melbourne, Australia and St. Petersburg, Russia, June 2–4, 2003*, number 2657 in Lecture Notes in Computer Science, pages 447–456. Springer-Verlag, 2003.
- [8] J. Mohr and X. Li. Optimized multiple wavetable interpolation. *WSEAS Transactions on Information Science and Applications*, 2(2):265–273, Feb. 2005.
- [9] G. Sandell and W. Martens. Prototyping and interpolation of multiple musical timbres using principal components-based analysis. In A. Strange, editor, *Proceedings of the 1992 International Computer Music Conference*, pages 34–37. International Computer Music Association, 1992.
- [10] M.-H. Serra, D. Rubine, and R. Dannenberg. Analysis and synthesis of tones by spectral interpolation. *Journal of the Audio Engineering Society*, 38(3):111–128, Mar. 1990.
- [11] J. C. Stapleton and S. C. Bass. Synthesis of musical tones based on the Karhunen-Loève transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-36(3):305–319, Mar. 1988.