A Comparison between Local Search and Genetic Algorithm Methods for Wavetable Matching

Simon Wun, Andrew Horner and Lydia Ayers
Department of Computer Science,
Hong Kong University of Science and Technology

Abstract
Picking spectral snapshots from the original tone is a successful approach to finding basis spectra in wavetable matching of musical instrument tones. Spectral snapshot selection usually requires optimization by a sophisticated technique such as a genetic algorithm (GA). This paper compares GA selection and a straightforward local search method, which first selects an initial set of basis spectra from the original tone by some simple selection strategy, and then modifies and improves them. Matching results for a range of instruments show that local search, when using the best composite results found by various simple selection strategies, is as effective as the GA. Moreover, surprisingly, starting from various initial basis spectra, local search can usually find several different matches, all with errors within 1% of the best solution, indicating that there are many near-optimal solutions in the wavetable search space.

1 INTRODUCTION

Wavetable synthesis reproduces a musical instrument tone of a definite pitch by adding several enveloped static waveforms (see Fig. 1) (Serra, Rubine, and Dannenberg 1990; Horner, Beauchamp, and Haken 1993). Each wavetable stores one period of a waveform which is a sum of harmonic sinusoids with fixed frequencies and amplitudes. The spectrum of a wavetable is called its basis spectrum. Fig. 2 shows a 1024-sample wavetable and the corresponding basis spectrum. During synthesis, the wavetable samples are multiplied by an amplitude envelope, and the signals from different enveloped wavetables are added to produce the matched tone.

First, a spectral analysis of the original tone transforms the sound into a time-varying spectrum. Phase vocoder analysis, short-time Fourier transform for harmonic sounds, can be used (Dolson 1986). It typically divides the original tone into thousands of frames a few milliseconds long, each representing a spectral snapshot at a particular time point. The complete set of spectral snapshots can also be viewed as discrete time-varying amplitudes and frequency deviations.

The target spectrum is typically limited to only some of the original frames to save computation in the later matching phase. Using 10 evenly spaced frames from the attack (that is, before the peak rms amplitude is reached) and 10 from the rest of the original tone keeps most of the perceptually important spectral information.

Next, assuming we have found the basis spectra, the amplitude envelopes can be determined by solving the following system of linear equation:

$$
\begin{bmatrix}
a_{1,1} & \cdots & a_{1,NTAB} \\
a_{2,1} & \cdots & a_{NTAB,NTAB}
\end{bmatrix}
\begin{bmatrix}
w_{1,1} & \cdots & w_{1,NFRM} \\
w_{2,1} & \cdots & w_{NFRM,NTAB}
\end{bmatrix}
$$

Figure 1: Multiple wavetable synthesis block diagram.

Wavetable matching tries to find the basis spectra and amplitude envelopes that can best approximate the original tone.
Local search is an iterative improvement heuristic. It is based on the hillclimbing idea: during the search, we repeatedly transform the current solution into a close and better new solution which then becomes the current solution (for the next iteration). The search starts from an initial solution, which can actually be randomly determined, and ends when improvement is not found with any transformation. In effect, it looks for a local optimum in the neighborhood of the initial solution.

Translating wavetable matching into relative spectral error minimization, we can immediately apply local search to our problem. The relative spectral error is the objective function that depends on the basis spectra. A solution can thus
be represented by a variable vector of harmonic amplitudes for the basis spectra. For example, with 3 wavetables and 30 harmonics, the solution vector is as follows:

\[ \mathbf{a} = [a_{1,1}, a_{1,2}, \ldots, a_{1,30}, a_{2,1}, \ldots, a_{2,30}]^T \]

where \(a_{j,k}\) is the amplitude of the \(k\)th harmonic of the \(j\)th basis spectrum. We used Powell's method (Press et al. 1992), a common multivariable numerical optimization technique, as the transformation. It is a "zeroth-order" direction set method that requires no derivatives but instead a sequence of line minimizations for which we used golden section search.

3 RESULTS

3.1 Simple Strategies versus GA

Fig. 4 shows how the four simple strategies compare to GA selection of the basis spectra. The GA consistently outperforms the simple strategies on a tenor voice tone. With 1–3 wavetables, excluding the end frames does better than including them, though EqHalfEndpts selection performs almost as well as the GA with 4–5 wavetables. The EqHalf strategies usually outperform their EqAll counterparts on the tenor.

![Figure 4: Matching results for a G3 tenor tone using simple and GA selection.](image)

Table 1 lists the 39 total tones we tested, from 16 instruments. Fig. 5 shows the average results over all of them. The simple strategies are not too bad, except EqAllEndpts with 1–2 wavetables. None of the simple strategies stands out as being consistently best, but EqHalf is better than EqAll overall. This is partly due to a bias inherent in the relative spectral error formula which uses an EqHalf target spectrum of 10 frames from the attack and 10 from the rest of the original tone. For fair comparison, we also tested with EqAll target spectra, and found that EqAll was indeed better than EqHalf 80% of the time.

![Table 1: 39 tones tested, from 16 instruments.](image)

Figure 5: Average matching results for all test tones using simple and GA selection with EqHalf target spectra.

3.2 Simple Strategies and GA followed by Local Search

This subsection compares the performance of the simple strategies and the GA when they are followed by local search. Fig. 6 shows the results for the tenor tone. The five methods mostly converge in the sense that nearly all find matches of approximately the same error level. Although the errors are about the same, the GA and simple strategy basis spectra are themselves quite different. Local search significantly improves the results of simple selection. For example, the improvement in the one- and two-wavetable matches is often greater than 8%. Local search only improves on GA selection by about 0.25%. EqAllEndpts gets stuck on a local optimum with two wavetables, though there is still a significant improvement. EqHalfNoEndpts also gets stuck with four
wavetables, so that the error appears to slightly increase compared with three wavetables.

Figure 6: Matching results for a G3 tenor tone using simple and GA selection followed by local search.

Fig. 7 shows the average results for the 39 test tones listed in Table 1. Overall, local search gives the most dramatic improvement when the number of wavetables is small. In particular, with one wavetable local search almost never gets stuck because it works very well when the number of parameters is small. Regardless of whether we start with a simple strategy or GA selection, local search results in approximately the same error level, which is close to that found by the GA before local search. Local search also sometimes gets stuck by chance. In other words, starting from an initial best does not guarantee the best final result. A practical solution to this problem is to use local search after each of the four simple strategies and keep the best solution.

Figure 7: Average matching results for all test tones using simple and GA selection followed by local search.

4 CONCLUSION

This paper has described a local search method for wavetable matching. It starts the search for artificial basis spectra using four simple spectral snapshot selection strategies, and repeatedly looks for local improvement. In effect, this locates small, promising regions of the vast search space and finds local optima.

We have compared the performance of GA selection and four simple strategies (namely EqAllEndpts, EqAllNoEndpts, EqHalfEndpts, and EqHalfNoEndpts) before and after local search. Before local search, GA selection always outperformed the simple strategies, while none of the simple strategies stood out as being consistently best. After local search, the simple selection matches were dramatically improved—by up to 30%, especially when the number of wavetables was small. As a result, all five methods converged, with mean errors within a range of 1%, though occasionally local search got stuck on a bad local optimum. To avoid getting stuck on bad local optima, we took the best solution found by the four simple selection strategies followed by local search. This simpler alternative to GA optimization produced the same quality results. In fact, local search actually improved on GA selection slightly.

Moreover, even starting from various initial basis spectra, local search can usually find different matches of approximately the same error level. These results suggest that the GA matches are near-optimal, and also that there are many near-optimal solutions in the wavetable search space.

References


