Polyrhythmic Texture Derivations of Time-Lapsed Field Recordings

Scott Smallwood
Rensselaer Polytechnic Institute
110 8th St. DCC 135
Troy, NY 12180 USA
Smalls@rpi.edu

challenged us to ‘connect with the acoustic environment and all that inhabits it’ by ‘going below the surface of what is heard and also expanding to the whole field of sound whatever one’s usual focus might be.’ (Oliveros 1998: 114-15) Composers and acoustic ecologists such as Christina Kubisch, R. Murray Schafer, and Hildegard Westerkamp have shown us that the sounded environment can itself be composition, and have, in fact,

II. A Brief Survey of Pitch/Time Shifting Methods

At this point it might be worthwhile to outline the various techniques for time-compressing audio, and the resultant artifacts that accompany these processes. The problem involved in changing the time of a recording without changing the pitch (or vice versa) involves two possible types of processes. The first type is the application of a mathematical process to the soundfile. The second type of process requires applying what we know about the human auditory system to the soundfile.

There are a variety of mathematical processes that can be used to change the pitch or the time of a sound-file independently of one another. The primary methods include using the phase vocoder, the wavelet transform, time-granulation, linear predictive coding, and simple sample-rate conversion. While comprehensive explanations of these methods can be found in Curtis Road's *Computer Music Tutorial* (Roads 1996: 440-48), I will give brief descriptions of each:

The Phase Vocoder: This method employs a frequency-domain analysis of the soundfile (using the fast Fourier Transform, or FFT), and then re-synthesizing the soundfile using additive synthesis. The analysis is done in a series of short, overlapping segments, or frames. Once the pitch information in the soundfile is analyzed and catalogued, it is a simple matter to change the time or pitch of the sound. To alter the time, the frames can simply be read back at different rates by changing the amount of overlap between frames. To alter the pitch, the frequency information in the frame can simply be re-scaled. This method is particularly useful for pitched-oriented sounds, but can be problematic for complex sounds, or sounds with complex attacks. Artifacts include smearing and reverberation effects.

The Wavelet Transform: This method is similar to the phase vocoder. A frequency-domain analysis is made, but unlike the FFT used in the phase vocoder, the segments are not of fixed duration; rather, they are dependent on the frequency information. The higher the frequency content in the analysis, the shorter the duration of the segment, or wavelet. Wavelet transforms tend to be more accurate for sounds with lots of high-frequency content because they are better at determining the timing-resolution of higher frequencies.

Time-Granulation: A time-domain-based process, time-granulation involves a kind of rapid-fire, microscopic sampling of the sound-file into hundreds or even thousands of grains per second. These grains can have their own unique envelope structure, and may overlap each other. To alter the time of the original sound file, grains can be dropped or doubled. To alter the pitch, sample-rate conversion is performed, and grains are doubled to account for the temporal change that would normally result. The obvious problem with this approach is that the transitions between one grain and the next can cause transient artifacts due to changes in level, waveform zero crossings, and other inconsistencies.

Linear Predictive Coding: This method involves analyzing a sound-file and extracting information for modeling the excitation/resonance properties of the instrument that is the subject of the recording. For this reason, linear predictive coding is limited to sounds that have the characteristics of the voice or of musical instruments. The information stored in the analysis, which includes pitch information, filter coefficients, duration, and other kinds of data, can be changed and then re-synthesized in order to change the time or pitch.

determining what the closest length ratio appeared to be, and adjusting their length appropriately to fit that ratio exactly.