Methods for Measurement and Manipulation of Timbral Physical Correlates

James W. Beauchamp

School of Music and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Abstract: Perhaps the most fundamental methodology of timbre research is to explore or test the perceptual effect of varying some physical aspect of the signal. In this regard, the frequency domain has been supreme, as timbre has classically been thought of as most strongly correlated with the amplitude spectrum of the signal. An examination of musical instrument sounds makes it obvious that in the "real world" most physical parameters are time-varying. Thus, time-varying spectral analysis is necessary as the first step toward discovering important physical parameters. The next step is to determine measures on the time-variant spectrum (i.e., time-varying parameters) which have strong and identifiable effects on perception. The spectral data for sounds may then be modified and the sounds resynthesized in order to explore the perceptual saliences of particular physical parameters.

INTRODUCTION

Since John Grey's marvelous thesis [1] a number of studies have been made to correlate perceptual and physical parameters of timbre of musical sounds [e.g., 2, 3]. Some perceptual testing methods used have been: 1) multi-dimensional scaling of perceived dissimilarity data with subsequent correlation to measured physical parameters; 2) analysis/parameter-simplification/resynthesis and testing of same/different perceptual thresholds.

To carry out the physical parameter measurements and parameter simplifications needed for the above tests, certain software is needed. This can be done with packages such as SNDAN [4] for the Unix environment. SNDAN consists of four basic components: 1) time-varying spectral analysis; 2) graphic display of various physical parameters; 3) parameter modifications; and 4) sound resynthesis, using the sine wave additive method. SNDAN is written in simple C and is very modular; thus, experimenters may easily write custom modules to realize their own ideas, assuming their environment allows recompilation. The purpose of this paper is to outline how analysis/synthesis software such as SNDAN can be used to measure physical parameters and prepare stimuli for experiments in timbre research. Primary long term objectives are to determine which parameters are most important and how they can be most fruitfully manipulated for musical applications.

SPECTRAL ANALYSIS

Spectral analysis (time-varying) can be thought of as the "front end" of an analysis/modification/resynthesis system, and many possible algorithms have been designed for this purpose. An analysis method generally must be complementary to a particular synthesis method, and a very desirable property (some would say a mandatory property) is that a sound resynthesized from the unmodified spectral analysis data be indistinguishable from the original sound. Moreover, it is also desirable that the parameters obtained from the analysis appear to be "reasonable", given the nature of the sound analyzed, and that if these parameters are modified before resynthesis, the resynthesized sound is audibly reasonable.

SNDAN contains two methods of analysis: 1) the pitch-synchronous phase vocoder [4] and 2) the frequency tracking method [5]. Both of these methods deliver sets of frequencies ($f_k$) and corresponding amplitudes ($A_k$) when the signal is representable as follows:

$$s(t) = \sum_{k=1}^{K} A_k(t) \cos(2\pi f_k(t) dt + \theta_k(0))$$  \[1\]

where $k$ is the index of the partial frequency, $K$ is the number of partials, and $t$ is time. $\theta_k(0)$ represents the starting phase of each partial. In the case of quasi-harmonic sounds, an approximate harmonic relationship $f_k(t) = k f_1(t)$ is valid. Method 1 may be viewed as a bank of harmonically-spaced equal-bandwidth bandpass filters which have a certain degree of overlap. For this method to work the output of each filter should be either nil or a single sine wave whose amplitude and frequency can then be determined. This puts a severe restriction on the type of signal which can be properly analyzed by this method, but happily it is a restriction satisfied by the vast majority of musical instrument individual tones. Method 2 is complicated by the fact that the number of partials $K$ varies with time and
the integrities of the partials are not generally intact throughout the sound; i.e., they tend to be born and die throughout the sound. Therefore, for simplicity, we usually use method 1. However, when frequencies are grossly inharmonic or when frequencies change by more than a few percent, method 2 may be more appropriate.

PARAMETER MEASUREMENT AND MANIPULATION

Multi-dimensional scaling (MDS) calculations based on listeners's estimates of perceptual distance between musical instrument sounds almost always yield a dimension which is strongly correlated with the spectral centroids of the sound stimuli [e.g., 6]. Spectral centroid can be defined as

\[
f_c = \frac{\sum_{k=1}^{K} f_k A_k}{\sum_{k=1}^{K} A_k}
\]  

[2]

In general, the centroid will vary with time during a sound. This variation can be eliminated by insuring that the ratios of the partial amplitudes do not vary with time. We do this by replacing the partial amplitudes by those which are proportional to the sound's rms amplitude while keeping the ratios of their relative amplitudes the same as those of the average amplitude spectrum of the original sound. We can manipulate the spectral centroid by altering the spectral envelope of the spectral data. Smoothly increasing the relative strengths of the higher frequency partials increases the centroid, and the rms amplitude can be easily retained by appropriate uniform time-variant scaling of the amplitudes.

Other physical parameters which have been proposed as important for timbre perception are attack time, spectral flux, spectral irregularity, spectral tilt, pre-attack noise, general noisiness, amplitude and frequency microvariations, and inharmonicity. While there is insufficient space here to give formulas, all of these parameters may be measured and manipulated. Manipulation may involve reducing or even eliminating parameters (if appropriate) or equalizing like parameters of distinct instrument sounds. In comparing two sounds, such equalization would allow listeners to focus on timbral differences besides the ones that have been equalized.

RESYNTHESIS AND TESTING

Resynthesis can be accomplished by the straightforward additive method, which is essentially an implementation of Equation 1, after the amplitude functions \(A_k(t)\) and frequency functions \(f_k(t)\) have been modified. For sounds with many partials, faster resynthesis may be accomplished by resampling in the frequency domain, inverse FFT, and overlap-add of the results [7]. Some tests that can be run are: 1) discrimination between sounds with modified parameters and corresponding original sounds [3]; we expect discrimination to improve as parameters are modified. 2) study the effect on a perceptual multi-dimensional space due to modified parameters; we would expect instrument locations to move closer together as parameters are simplified or eliminated. 3) compare ability to recognize instruments from a large collection before and after parameter simplification; we expect recognition to deteriorate as parameters are stripped. 4) MDS mapping of instrument sounds which have a particular parameter equalized; this would hopefully better reveal dimensions of important ancillary timbral correlates.

REFERENCES