Time Domain Extraction of Vibrato from Monophonic Instruments

Introduction

Vibrato is an essential ingredient in the expressive nature of many musical instruments. Creating slight oscillations in the pitch and/or volume of the musical tone, vibrato allows a long sustained note to become more lively and dynamic. Musical instruments differ in both the technique used to create vibrato as well as the physical characteristics of the vibrato-enhanced sound produced. The goal of this research is to extract information describing the amplitude, frequency, and phase of the vibrato from a section of monophonic music.

Procedure

The steady state of a musical note can be approximated as follows:

$$G(t)=A(t)*\sum_{n=1}^{\infty}[H_n*\cos(n*(f_0+F(t))+\theta_n)]$$

$G(t)$- Musical signal
n- harmonic number (1 (fundamental), 2, 3, etc.)
H_n- amplitude of harmonic
θ_n- phase of harmonic
f_0- fundamental frequency
$A(t)$- amplitude modulation signal
$F(t)$- frequency modulation signal

The amplitude and frequency modulation signals, A(t) and F(t), can be extracted using envelope and pitch extraction techniques respectively. In non-vibrato signals, A(t)=1 and F(t)=0, however with the onset of vibrato they both take on the form $A_v*\cos(f_v*t+\theta_v)$. The amplitude (A_V), frequency (f_V), and phase (θ_V), of each these signals can then be analyzed. These methods only work on monophonic music, and are suitable for real-time implementation.

Envelope extraction

Vibrato, a low frequency signal, sinusoidally varies the envelope of the instrument's waveform (composed of higher frequency components). We can reconstruct the original vibrato signal, by using the standard deviation within a 36.4 ms window as a relative measure of envelope's amplitude. This is the smallest window that will fit the entire wavelength of the musical note A0 (27.5 Hz), the lowest note we are interested in detecting.

The unfiltered vibrato signal is found by sliding the 36.4 ms window one sample at a time and recalculating the standard deviation within the window. We then filter this with a bandpass filter[1] (passing 4-7 Hz)[2], to obtain the original vibrato signal causing the amplitude modulation.

Prior to using the bandpass filter, the signal can be put through an anti-aliasing low pass filter and downsampled. (We downsampled the sampling rate from 44.1 kHz to 44.1 Hz in our experiment). This reduces the order needed for the bandpass filter.

If the onsets and offsets are known, a Hanning window can be used to remove the transients they produce in the signal. Otherwise these transients may be confused with vibrato when passed through the bandpass filter.

Pitch Extraction

[1] The bandpass filter used is a FIR least squares filter Order=100, f_{p1}=4Hz, f_{p2}=7Hz

[2] 4-7 Hz was the observed range of vibrato frequency in our sound files.

Vibrato sinusoidally varies the fundamental pitch (and harmonics) of the instrument's waveform. Because all the frequency components are affected, just observing the fundamental pitch will be enough for determining where the vibrato is present. To monitor the vibrato in the fundamental, the following pitch tracking scheme was used.

We can obtain the fundamental frequency by dividing the sampling rate by the average number of samples within one wavelength of the sound. To find the wavelength, an autocorrelation of a 72.7 ms windowed signal is used. This is the minimum window size that would contain two wavelengths of the musical note A0 (27.5 Hz), the lowest note we are interested in detecting. With the autocorrelation data, the distance (number of samples) between the two largest peaks is calculated, and the most likely wavelength (number of samples) is determined. The windowed autocorrelation is passed over the whole signal, sliding 36.4 ms (half of a window length) each time.

The frequency modulation signal in this form can be put through a bandpass filter (4-7 Hz) as for the amplitude modulation signal, to clean up the non-vibrato noise.

Conclusion

This is an approach to a completely time based analysis of vibrato induced variation of frequency and amplitude in a monophonic musical signal. Amplitude changes are detected by envelope extraction with a sliding windowed standard deviation. Frequency changes are detected by pitch tracking with a sliding windowed autocorrelation of the signal.

These algorithms are suitable for real-time implementation, and could be utilized as a teaching aid for musicians wishing to see what physical parameters affect their vibrato quality. In addition, this information could be fed into a blackboard system as a preprocessing step before music transcription. The parameters of the amplitude, frequency, and phase of the vibrato generated frequency and amplitude modulating signals, can be obtained with this method and used in classification systems and neural network models.

Author Information

Daniel Bendor
Undergraduate School of Electrical Engineering
University of Maryland at College Park
dbendor@glue.umd.edu

Mark Sandler
Department of Electronic Engineering
King's College London
mark.sandler@kcl.ac.uk

Suggested Readings

Herrera, Perfecto; Bonada, Jordi. 1998. *Vibrato Extraction and Parameterization in the Spectral Modeling Synthesis Framework.* Proceedings 98 Digital Audio Effects Workshop.

Rossignol, S., Depalle, P., Soumagne, J. Rodet, X. and Collette J.-L. 1999. *Vibrato Detection, Estimation, Extraction and Modification.* Proceedings 99 Digital Audio Effects Workshop.

Skovenborg, Esben. 1998. *Modelling the Timbre of Musical Sound.* Dissertation submitted to the Department of Electronics and the Music Technology Group of the University of York.