Final Technology Review

Pitch Modulated Vibrato Function

Justin Leong 306179806

Digital Audio Systems, DESC9115, Semester 1 2012

Graduate Program in Audio and Acoustics

Faculty of Architecture, Design and Planning, The University of Sydney
The digital synthesis of natural sounding, pitch modulated vibrato is a deceptively challenging task due to all of the subtle variations and fine details that are present in vibrato. Examining a spectrogram of human violin vibrato, such as figure 1, illustrates some of these nuances.

**Figure 1.** Spectrogram of violin vibrato with the note D6 as its fundamental frequency.
Simply modulating the pitch of an input signal in the shape of a sine wave yields a very artificial result. One of the reasons for this is because every cycle of vibrato is identical. Close inspection of the spectrogram shows that human vibrato contains slight variations in the shape of each pitch modulation from one cycle to the next. There is also a period of no pitch fluctuation when the note is initially sounded, followed by a short “warm-up” period as the vibrato gradually reaches its full width. Furthermore, the shape of the pitch modulation is not exactly that of a sine wave. We can see that the maxima of each period are much smoother and rounder compared to the minima, which are sharp and almost triangular. These points expose the quasi-periodic nature of human vibrato, and this is one of the key elements to replicating this effect realistically.

Another point of concern is that most instruments have a fixed resonant response as a result of the fixed shape of their bodies. A graph of an instrument’s resonant response shows that the body of the instrument radiates certain frequencies much more effectively than others. An example of this can be seen in figure 2.

![Resonant response of the body of a violin](image)

**Figure 2. Resonant response of the body of a violin [1].**

This resonant response is different for every instrument, and gives each instrument its own unique timbre [1]. This physical characteristic is an important point of consideration with respect to vibrato synthesis because when the pitch of the signal is altered, the resonant response of the signal must remain unchanged. If the vibrato effect produces a fluctuation in the instrument’s resonant response then it will sound as though the instrument is periodically shrinking and expanding in size.

This is by no means an exhaustive list of qualities that give rise to natural sounding vibrato, but those mentioned are the ones that this vibrato function seeks to address. In addition to being able to alter the rate of pitch oscillation and width of the vibrato, users will have control over:

- Delaying the onset of the vibrato; and
- Choosing from a number of pitch modulation signals. The choices are filtered fractal noise, a sine wave and an inverted $\sin^4$ wave. Examples of these can be seen in figure 3.
Figure 3. Pitch modulation shapes for the vibrato effect. Filtered fractal noise (top), sine wave (middle) and inverted sine^4 wave (bottom).
The function also uses an algorithm that preserves the resonant response of the input signal when applying the vibrato effect.

The function generates pitch variation by putting the input signal through an oscillating time delay system. By varying the delay of the signal, the function is changing the apparent distance between the signal and the listener. This procedure induces a periodic Doppler effect on the signal, and can be likened to the perceived change in pitch of an ambulance’s siren as it drives by a listener [2]. The shape of the modulating delay signal governs how the vibrato’s pitch modulation will sound. For example, if the delay signal was a simple sine wave in the form of:

\[ y[n] = \text{Acos}(\omega n) \]

and we increased its frequency and amplitude, both of which are designated by the user, then this would result in, respectively, a faster vibrato and more pitch variation.

The use of filtered fractal noise as the modulating delay signal is how this function attempts to replicate the quasi-periodic nature of human vibrato. The function generates random fractal noise, which is then sent through a low pass filter to obtain a quasi-periodic sine wave with the approximate frequency of that specified by the user. The extent to which the resulting delay signal adheres to this specified modulation frequency is determined by the quality factor (Q) of the filter used; also designated by the user. The higher the Q value, the more strictly the delay signal’s modulation frequency will adhere to the value assigned by the user. Examples of these signals can be seen in figure 4.

It can be seen that this process is capable of generating quasi-periodic pitch variation. However, the disadvantage of this method is that it generates a different delay signal every time the function is run. This means that there is a degree of uncertainty as to what waveform will be produced by the function. In the interest of producing more predictable delay signals, the function also allows the option of choosing a sine wave or an inverted sine^4 wave as the modulation signal. The significance of the inverted sine^4 wave is that its shape more accurately represents the pitch modulation of real vibrato seen in the spectrogram in figure 1.

![Unfiltered Fractal Noise](image)

**Figure 4.** Random fractal noise signal with no filtering.
Figure 4 (continued). Examples of filtered fractal noise using filters with different quality values to produce quasi-periodic modulating delay signals.

The temporal delaying of the onset of vibrato is achieved through the use of a scaling system applied to the generated modulating delay signal. The user specifies the length of time for there to be no vibrato, and another length of time to determine how long it takes for the vibrato to reach its full width. The function generates a scaling vector of the same length as the modulating delay signal. This consists of an appropriate number of zeros and ones bridged by a half Hanning-Bartlett window of suitable length. Each sample of this scaling vector is then multiplied by its respective sampling index in the delay signal to yield the result seen in figure 5.
Figure 5. The original modulation signal (top) is multiplied by the scaling signal (middle) to produce a modulation signal with a delayed onset of modulation (bottom). This example has a no vibrato length of 0.5 seconds and a transition length of two seconds.
The function preserves the resonant response of the input signal by using a linear predictive coding (LPC) algorithm which filters the input signal into its excitation source and resonances. The chosen modulating delay signal is then applied only to the excitation source, leaving the filtered resonances unaltered. These are then recombined using a LPC synthesis function to produce the final result [3].

An evaluation of this function shows that it addresses the highlighted challenges. It offers a means of synthesising vibrato effects which have subtle differences from one vibrato cycle to the next. As well as offering a variety of different types of pitch modulation signals, it also gives the user control over delaying the onset of the vibrato, which imitates the natural tendency of performers when executing this technique.

A survey of six volunteers revealed that most, but not all, listeners were able to distinguish real violin vibrato from the three different types of synthesised vibrato (applied to a recording of a violin playing a single, held note). There was a general consensus that the real vibrato sounded more resonant or reverberant, whilst the synthesised vibratos sounded duller. This is most likely due to how the pitch fluctuation of the fundamental note and its overtones interact with the fixed resonant response of the instrument. Some overtones will be radiated more effectively when their pitch is increased, while others may be attenuated. This is also the case for when the pitch decreases [4]. Unfortunately, this function does not replicate this complex spectral characteristic of vibrato.

It is worth noting that this survey involved listening to these vibrato effects in isolation on single notes. The volunteers mentioned that the effect would be much more difficult to distinguish if it was applied in the context of a recorded piece of solo, unaccompanied music which originally had no vibrato. This shows that there is merit in this function as a viable method of synthesising natural sounding vibrato.
References


