MODULATION AND DELAY LINE BASED DIGITAL AUDIO EFFECTS

Sascha Disch and Udo Zölzer

Dept. of Signal Processing and Communications
University of the German Federal Armed Forces Hamburg
Holstenhofweg 85 · D-22043 Hamburg · Germany
dsh@iis.fhg.de, udo.zoelzer@unibw-hamburg.de

ABSTRACT

In the field of musicians and recording engineers audio effects are mainly described and indicated by their acoustical effect. Audio effects can also be categorized from a technical point of view. The main criterion is found to be the type of modulation technique used to achieve the effect. After a short introduction to the different modulation types, three more sophisticated audio effect applications are presented, namely single sideband domain vibrato (mechanical vibrato bar simulation), a rotary speaker simulation, and an enhanced pitch transposing scheme.

1. EFFECT CLASSIFICATION

Many audio effects are based on mixing the original signal with delayed and/or amplified copies of it. The delay is implemented in integer multiples of the unit delay. If the delay is greater than one unit delay, the chain of unit delays is referred to as delay line. If the delay time and the corresponding coefficients are constant, the well known FIR (finite impulse response) and IIR (infinite impulse response) filters are the result. In the case of time varying filters the coefficients are time variant, but the delays stay fixed. The resulting filter is a superposition of amplitude modulated (AM) and delayed copies of the original signal. If the coefficients are fixed and the delays are variable, audio effects like vibrato, flanging and chorus are built. The variation of the delay length is a phase modulation (PM) and its dynamic component is perceived as a frequency modulation. Some more elaborated effects use both types of modulation and will be presented in the following sections.

2. MODULATION

2.1. Amplitude Modulation

Performing amplitude modulation (AM) in the digital domain requires just one multiplication, as shown in Fig. 1. The audio input signal $x(n)$ is amplitude modulated by the modulation signal $m(n)$. The classical amplitude modulation is expressed as

$$m(n) = 1 + \alpha \sin(\omega_c nT) \quad 0 < \alpha < 1 \tag{1}$$

where α is the modulation depth. The well known ring modulation is given by

$$m(n) = \alpha \sin(\omega_c nT). \tag{2}$$

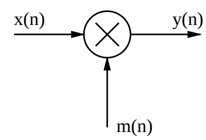

Figure 1: Amplitude modulation.

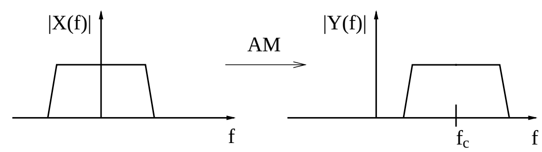

Figure 2: Spectrum of amplitude modulation.

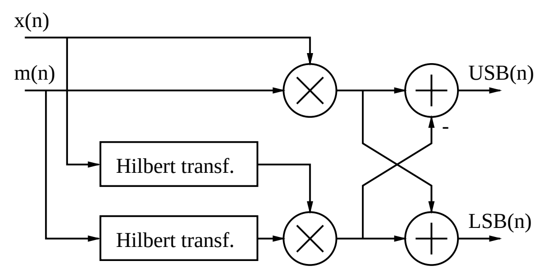

Figure 3: Single sideband modulation.

The spectral properties are depicted in Fig. 2.

Single sideband modulation (SSB mod.), as shown in Fig. 3, shifts the spectrum of the signal a certain amount up or down the frequency axis. The width of the spectrum remains unchanged (see Fig. 4). This is the reason why SSB modulation is not suitable for transposing music signals: the harmonic relations of the signal frequencies are destroyed, e.g. cords will sound 'out of tune'. Nevertheless the SSB modulation can be used as a 'special effect'. Another application is to use SSB modulation as an intermediate transform. First use SSB modulation to shift the input spectrum, then apply filtering or phase modulation and then perform the demodulation of the signal.

Discrete-time single sideband modulation is given by

$$\left.\begin{matrix}\mathrm{LSB}(n)\\ \mathrm{USB}(n)\end{matrix}\right\} = m(n)x(n) \pm \widehat{m}(n)\widehat{x}(n) \tag{3}$$

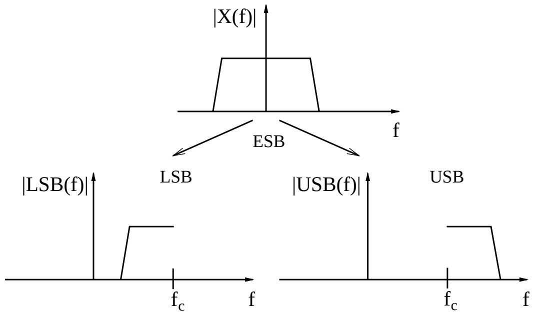

Figure 4: Spectrum of SSB modulation.

where $\widehat{x}$ denotes the Hilbert transform of the signal x. A discrete-time Hilbert transform can be approximated by a FIR filter with the impulse response

$$h(n) = \frac{1-\cos(\pi n)}{\pi n} = \begin{cases} 2/(\pi n) & \text{for} \quad n \quad \text{odd} \\ 0 & \text{for} \quad n \quad \text{even.} \end{cases} \tag{4}$$

These coefficients are multiplied with a suitable window function of length N. Acceptable quality can be achieved with $N \approx 60$. Note that the use of an FIR filter approximation requires the filter delay to be compensated in the direct path. Further applications of modulation techniques for audio effects are presented in [2, 4].

2.2. Phase Modulation

A phase modulation (PM) of the signal $x(n)$ is described by

$$x_{PM}(n) = x(n - D(n)) \tag{5}$$

where $D(n)$ is a continuous variable. Therefore $D(n)$ is decomposed into an integer and a fractional part [1]. The integer part is implemented by a series of M unit delays, the fractional part is approximated by interpolation filters [1, 3, 5], e.g. linear interpolation, spline interpolation or all-pass interpolation (see Fig. 5). The discrete-time Fourier transform of (5) yields

$$X_{PM}(e^{j\Omega}) = X(e^{j\Omega})e^{-j\Omega D(n)}, \tag{6}$$

which shows the phase modulation of the input signal. The phase modulation is performed by the modulation signal $D(n)$.

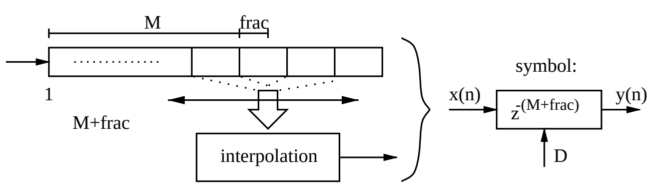

Figure 5: Delay line with interpolation.

For sine type modulation, useful for vibrato effects, the modulation signal can be written as

$$D(n) = M + \text{DEPTH} \cdot \sin(\omega_M nT). \tag{7}$$

For a sinusoid input signal the so-called resampling factor for sine type modulation can be derived as

$$\alpha(n) = \frac{\omega_I}{\omega} = 1 - \text{DEPTH} \cdot \omega_M T \cos(\omega_M nT). \tag{8}$$

The instantaneous radian frequency is denoted by ω_I and the radian frequency of the input sinusoid is denoted by ω. The resampling factor is regarded as the pitch change ratio in [1]. For sine type modulation the mean value of the resampling factor $\alpha(n)$ is one. The consequence is an output signal, which has the same length as the input signal, but has a vibrato centered around the original pitch.

If the modulation signal is a ramp type signal according to

$$D(n) = M \pm \text{SLOPE} \cdot n, \tag{9}$$

the resampling factor $\alpha(n)$ for sinusoid input signal is given by

$$\alpha(n) = \frac{\omega_I}{\omega} = 1 \mp \text{SLOPE}. \tag{10}$$

The output signal is pitch transposed with a factor α and the length of the output data is altered with the factor $1/\alpha$. This behavior is useful for pitch transposing applications and will be exploited in the last section.

3. MECHANICAL VIBRATO BAR SIMULATION

We now consider the system in Fig. 6 where SSB modulation and demodulation is combined with phase modulation. The system consists of an anti-aliasing filter, a SSB modulator, a vibrato (delay line with sine type modulation) and SSB demodulator. The modulation and demodulation is controlled by a low frequency sinusoid.

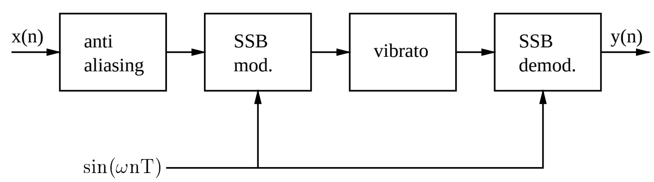

Figure 6: Vibrato with SSB modulation/demodulation.

The overall resampling factor is a function of frequency. Figure 7 shows the distorted overall resampling factor vs. the inner resampling factor. The diagrams demonstrate two cases, namely LSB modulation with 50 Hz and USB modulation with 50 Hz. One can notice, that the resulting vibrato is not harmonic anymore. A range of two octaves is shown (50%-200%). The formula for the overall resampling factor is

$$\widetilde{\alpha}(f) = (\alpha\,(f \pm f_c)) \mp f_c \qquad \begin{cases} \text{for} & \text{USB mod} \\ \text{for} & \text{LSB mod} \end{cases} \tag{11}$$

with α the inner resampling factor of the vibrato, $\widetilde{\alpha}$ the overall resampling factor of the system and f_c the modulation frequency. This phenomenon can be used to simulate the mechanical vibrato bar of an electric guitar. Such a vibrato bar alters the pitch of the lower strings of the guitar in larger amounts than the higher strings and thus a non-harmonic vibrato results. A similar effect can be achieved by using USB modulation, vibrato and the adequate demodulation. Another application could be the construction of modified flangers.

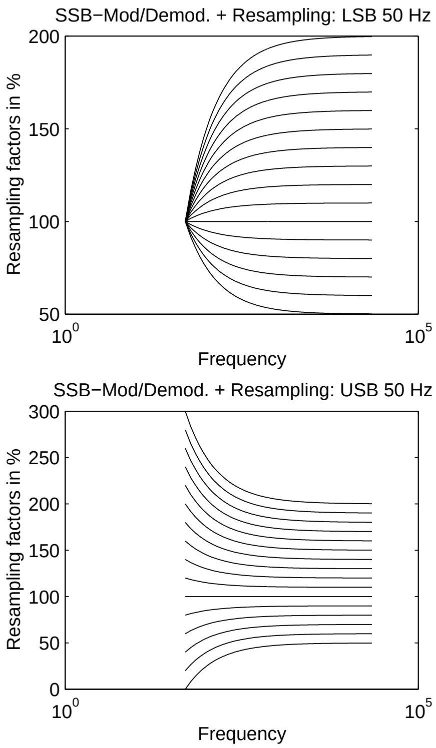

Figure 7: *Resampling* factors in SSB-mod./vibrato/SSB-demod. setup.

4. STEREO ROTARY SPEAKER EFFECT

The rotary speaker effect simulates the sound effect achieved by rotating horn speakers and a bass cylinder, as first produced for organs. The sound is altered by the Doppler effect, the directional characteristic of the speakers, phase effects due to air turbulences, etc. Figure 8 is a coarse physical model, which is used to derive the topology for a digital implementation of the effect.

Doppler effect: There are 2 rotation frequencies (‘choral’and ‘tremolo’) with 15-120 rpm and 300-498 rpm. One can switch between these frequencies during the performance. The speed of the moving horn aperture is 0.35-2.8 and 7-11.7 m/second. The Doppler effect of moving sound sources is described by

$$\frac{f_{\text{Doppler}}}{f} = \alpha = \frac{1}{1 \pm \frac{v_{\text{source}}}{c}} \tag{12}$$

with f_{Doppler} as the altered frequency and hence α as the doppler factor. The speed of one horn aperture relative to the listener can be calculated as

$$v_x = v_{\text{source}} \cdot \cos(\varphi). \tag{13}$$

Setting α equal to the resampling factor of phase modulation the depth of modulation is found to be approx. 0.5 msec. The simulation of the Doppler effect of two opposite horns is done by the use of two delay lines modulated with 180° phase shifted signals in vibrato configuration (see Fig. 9).

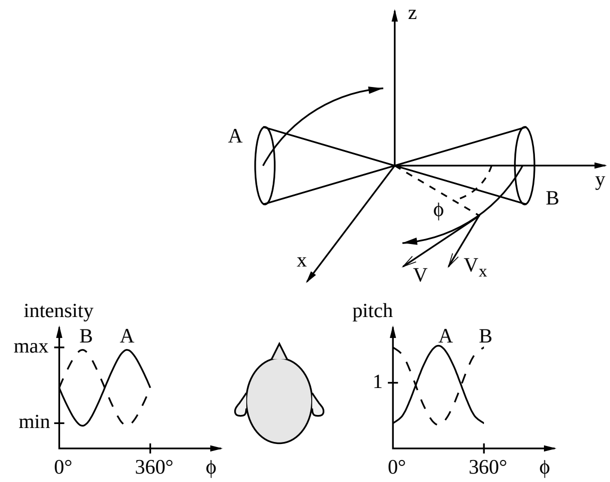

Figure 8: Model of rotating speaker effect.

Directional sound: A directional sound characteristic similar to rotating speakers can be achieved by amplitude modulating the output signal of the delay lines. The modulation is synchronous to the delay modulation in a manner, that the back moving horn has lower pitch and decreasing amplitude. At the return point the pitch is unaltered and the amplitude is minimum. The movement in direction to the listener causes a raised pitch and increasing amplitude.

Stereo effect: A stereo rotary speaker effect is perceived due to unequal mixing of the two delay lines to the left and right channel output. The topology in Fig. 9 shows the simulation of a rotating double horn.

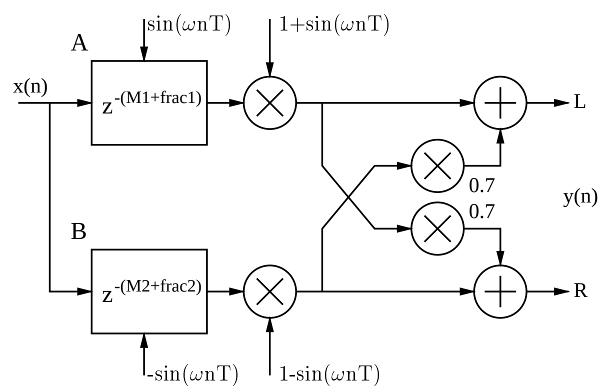

Figure 9: Simulation of rotating speakers.

Two speeds: The switching between the two rotation frequencies creates a continuous frequency sweep because of the inertia of the moved masses. This behavior is simulated with a ramp signal applied to the voltage controlled oscillator (VCO) in case of changing the rotation frequency.

Rotating horns and cylinder: The effect of splitting off the signal in a rotating horn part and a rotating cylinder part is achieved by separately processing a high and low frequency portion of the input signal. The processing is not synchronous and the VCO

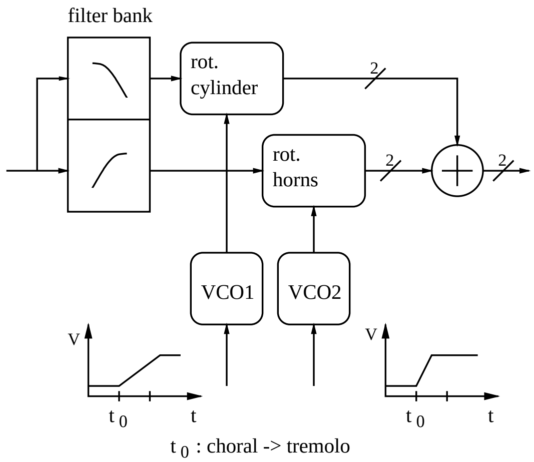

Figure 10: Rotary speaker simulation.

ramps have different time constants to simulate the different masses of the horns and the cylinder. Figure 10 shows the entire application.

5. PITCH TRANSPOSING

In this section an enhanced method for transposing audio signals is proposed. The method is based on an overlap-add scheme and does not need any base frequency estimation. The difference to other applications is the way the blocks are modulated and combined to the output signal. The enhanced transposing system is based on an overlap-add scheme with three parallel time varying delay lines. Figure 11 illustrates how the input signal is divided into blocks, which are resampled (phase modulation with a ramp type signal), amplitude modulated and combined to an output signal of the same length as the input signal. Adjacent blocks overlap with 2/3 of the block length. The modulation signals form a system of three 120° phase shifted raised cosine functions. The sum of these functions is constant for all arguments. Figure 12 shows the topology of the pitch transposer.

Since a complete cosine is used for modulation, the perceived sound quality of the processed signal is much better than in simple twofold overlap-add applications using several windows. The amplitude modulation only produces sum and difference frequencies with the base frequency of the modulation signal, which can be very low (6 - 10 Hz). Harmonics are not present in the modulation signal and hence cannot form sum or difference frequencies of higher order. The perceived artifacts are phasing like effects and are less annoying than local discontinuities of other applications based on twofold overlap-add methods.

An implementation of the pitch transposer has the following components: A low frequency oscillator with three 120° phase shifted cosine outputs and three synchronous sawtooth signals, three interpolating delay lines (one delay line with three interpolating outputs), three amplitude modulators and one mixing stage. On Motorola's DSP56002 a stereo implementation with 48 kHz sampling rate of this algorithm using third-order spline interpolation requires approx. 85% of the computational capacity.

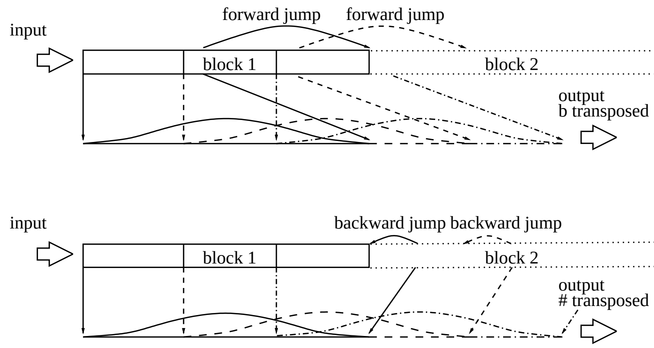

Figure 11: Pitch transposer.

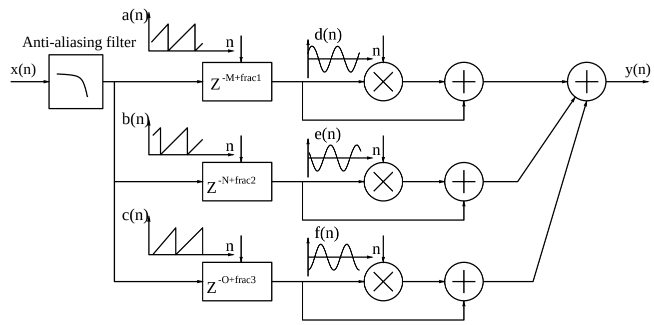

Figure 12: Pitch transposer with 3 delay lines.

6. CONCLUSION

In this paper we have shown the combination of phase modulation obtained by time varying delays and amplitude modulation to form digital audio effects. We presented the use of single sideband modulation/demodulation in combination with time varying delays for the simulation of a mechanical guitar vibrato (non harmonic vibrato). The combination of phase and amplitude modulation leads to the simulation of the rotating speaker effect and to a simple pitch transposing scheme.

7. REFERENCES

[1] J. Dattorro, *Effect Design - Part I + II*, J. Audio Eng. Soc., vol. 45, pp. 660-684 (Sept. 1997), pp. 764-788 (Oct. 1997)

[2] P. Dutilleux, *Filters, Delays, Modulations and Demodulations: A Tutorial*, Proc. DAFX98 Workshop on Digital Audio Effects, pp. 4-11, Barcelona 1998.

[3] T. I. Laakso, V. Välimäki, M. Karjalainen and U. K. Laine, *Splitting the unit delay*, IEEE Signal Process. Mag., vol. 13, pp. 30-60, 1996.

[4] S. Wardle, *A Hilbert-Transformer Frequency Shifter for Audio*, Proc. DAFX98 Workshop on Digital Audio Effects, , pp. 25-29, Barcelona 1998.

[5] U. Zölzer, *Digital Audio Signal Processing*, J. Wiley & Sons, Chichester 1997.