Efficient Asynchronous Resampling of Audio Signals for Spatial Rendering of Mirror Image Sources

Christian Borß

Institute of Communication Acoustics, Ruhr-Universität Bochum, 44780 Bochum, Germany, Email: christian.borss@rub.de

Abstract

In this article we present a computationally efficient resampling method for rendering mirror image sources. In a first stage, the proposed sample-frequency converter transforms the input signal into a parameter space, in which a second stage can compute sub-sample values with few operations. For this, we propose transforming the oversampled input signal into a sequence of polynomial coefficients and a Farrow structure for the implementation of the second stage.

Introduction

Mirror image models are widely used for auditory virtual environments (AVEs) to render early reflections. Among other things, the simulation of this early reverberation takes the sound propagation delay of the mirror image sources into account. For a moving receiver or moving sound sources, the propagation delays are a function of time. An asynchronous sample-frequency converter can be used to implement this variable propagation delay and thus can also take the Doppler shift into account. Although the Doppler shift may be below the perception threshold for slow movements, an asynchronous resampling of all mirror image signals is necessary to avoid processing artifacts. Such artifacts occur when a block-based signal processing model is used and the propagation delay is determined only once per block.

Signal Distortion

For the implementation of the fractional delay a variety of interpolation methods exists [1, 2, 3, 4, 5]. Interpolation based on Lagrange polynomials provides a good compromise between signal distortion, computational complexity, and implementation effort for the intended system [2]. Signal distortion is caused by the interpolation filter and can be quantified for a sinusoidal signal s_f with frequency f and the interpolated output signal $\hat{s}_f$ by the signal-to-noise ratio

$$\frac{S_f}{N_f} = 10\log_{10}\frac{E\{s_f^2\}}{E\{(s_f-\hat{s}_f)^2\}} \quad . \tag{1}$$

This formula can be evaluated for uniformly distributed subsample positions λ and phases φ by integration over all phases and all sub-sample positions. We approximate the signal-to-noise ratio for 3$^{\text{rd}}$ order Lagrange interpolation by

$$\frac{S_f}{N_f} \approx 10\log_{10}\frac{LM}{\sum_{l=1}^{L}\sum_{m=1}^{M}\left(s_f\left(\frac{l}{L},\frac{m\pi}{M}\right)-\hat{s}_f\left(\frac{l}{L},\frac{m\pi}{M}\right)\right)^2} \tag{2}$$

for the sinusoidal $s_f(\lambda,\varphi)=\sqrt{2}\cos(\lambda 2\pi f/f_s+\varphi)$, where f_s denotes the sampling frequency, $\hat{s}_f(\lambda,\varphi)$ the interpolated output signal, and L and M the granularity of λ and φ, respectively. For the determination of the polynomial

$$\begin{aligned}\hat{s}_f(\lambda,\varphi) &= c_3\lambda^3+c_2\lambda^2+c_1\lambda+c_0\\ c_0 &= x(1)\\ c_1 &= x(2)-x(0)/3-x(1)/2-x(3)/6\\ c_2 &= [x(0)+x(2)]/2-x(1)\\ c_3 &= [x(3)-x(0)]/6+[x(1)-x(2)]/2\end{aligned} \tag{3}$$

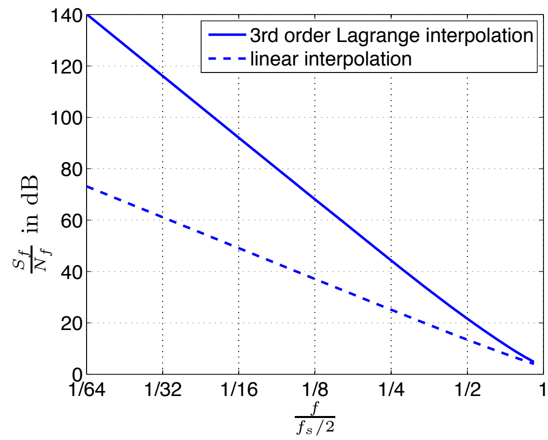

Figure 1: Signal-to-noise ratio of an interpolated discrete sinusoidal signal with the frequency f and the sampling frequency f_s, using linear and 3$^{\text{rd}}$ order Lagrange interpolation

the samples $x(n) = s_f(n-1,\varphi)$ are used. We reform these equations to yield a form which is optimal for implemention due to a minimum number of multiplications. Figure 1 shows the resulting signal-to-noise ratio for $L = 100$ and $M = 100$.

The signal distortion caused by the interpolation can be reduced by prior oversampling of the input signal. According to Figure 1, a signal-to-noise ratio of ∼70dB (at the Nyquist frequency) or greater (at lower frequencies) can be achieved with 8x oversampling and 3$^{\text{rd}}$ order Lagrange interpolation. Figure 2 shows the spectrogram and selected short-time spectra of the output signal of the implemented interpolation filter for a moving sound source. We observe that the noise level is even below -78dB as approximated by (2), because its energy is spread over several discrete frequencies.

Computational Complexity

The determination of the polynomial coefficients - like in (3) for a Lagrange polynomial - is based on a weighted sum of consecutive samples. This can be seen as a convolution which results in the polynomial coefficient sequence $c_k(n)$. Because all mirror image sources use the same input signal [6], the computational complexity of the interpolation filter can be reduced by applying this convolution to the oversampled input signal $x(n)$ once for all mirror image sources. This constitutes a transformation into a parameter space in which sub-sample values can be computed with few operations. Using a Farrow structure as shown in Figure 3, only K multiplications per sample are necessary for each mirror source for an interpolation based on a K^{th} order polynomial [7].

Based on these considerations we propose a two-stage procedure for an efficient realization of the sample frequency conversion. The first stage which consists of an oversampling unit - with a constant factor - and a polynomial coefficient transformation has a computational complexity of the order $\mathcal{O}(1)$. The second stage is applied to all N mirror image

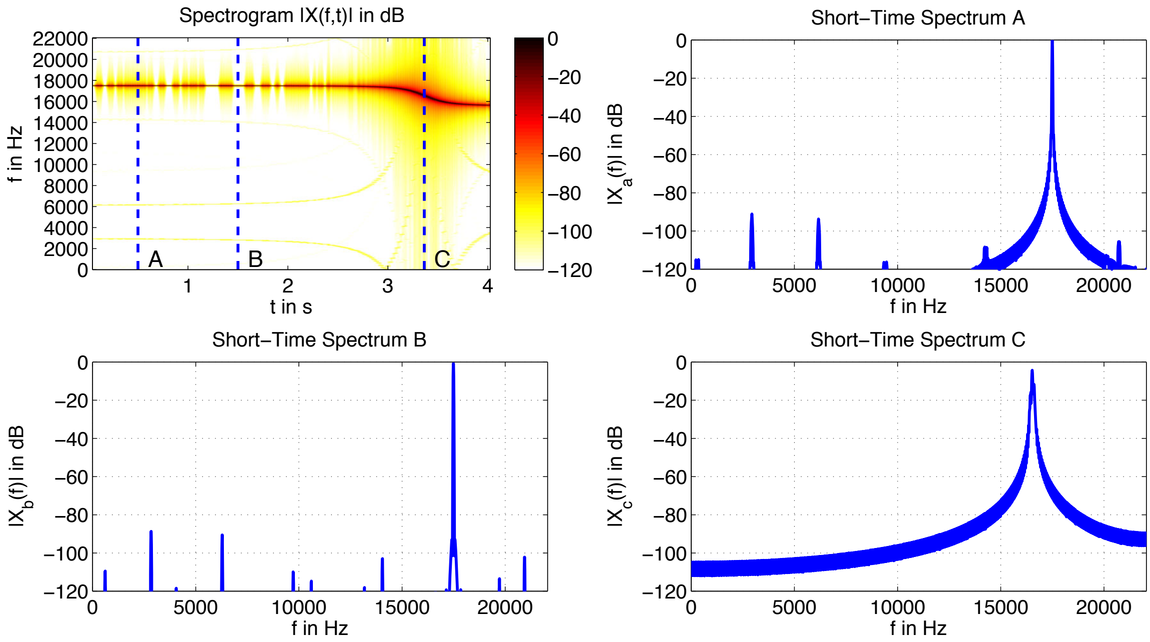

Figure 2: Spectrogram (top left) and selected short-time spectra (labeled as "A", "B", and "C") of an interpolated audio signal using the proposed settings (8x oversampling with a poly-phase lowpass filter and 3rd order Lagrange interpolation). A moving sound source is simulated which emits a sinusoidal signal (f = 16.5kHz, f_s = 44.1kHz), moves linearly with constant velocity (v = 72km/h), and passes the listener in a distance of 5m. The distance dependent attenuation of the transmitted signal was not simulated to yield a better comparability of the sub-figures.

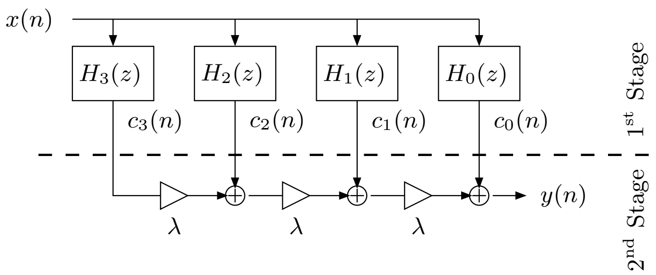

Figure 3: Farrow structure for polynomial interpolation at sub-sample position λ

sources and thus has a computational complexity of the order $\mathcal{O}(N)$. For a great number of mirror image sources the second stage has a major impact on the total computational complexity. Thus, the polynomial order used for the interpolation should be as low as possible. Due to symmetry reasons, an odd polynomial order is preferable. As shown in Figure 1, high oversampling factors are necessary for linear interpolation to yield low signal-to-noise ratios. This results in high memory requirements for the oversampled signal. A 3rd order Lagrange polynomial yields a signal-to-noise ratio of at least ~70dB with prior 8x oversampling and requires only 3 multiplications for the Farrow structure. Additionally, on modern IA-32 processors this can be implemented very efficiently using SIMD (Single Instruction, Multiple Data) instructions which simultaneously multiply a vector of 4 single precision floating point variables.

Conclusion

We proposed a computationally efficient resampling method for rendering moving mirror image sources based an a hybrid sample-frequency converter with two stages. We showed that a transformation of the input signal into a sequence of 3rd order Lagrange polynomial coefficients with prior 8x oversampling in the first stage and a Farrow structure in the second stage yields a good compromise between computational complexity and signal distortion.

References

[1] R. W. Schafer and L. R. Rabiner, "A Digital Signal Processing Approach to Interpolation," *Proceedings of the IEEE*, vol. 61, no. 6, pp. 692–702, Jun. 1973.

[2] T. I. Laakso et. al., "Splitting the Unit Delay," *IEEE Signal Processing Magazine*, vol. 13, no. 1, pp. 30–60, 1996.

[3] H. Strauss, "Implementing Doppler Shifts for Virtual Auditory Environments," in *104th AES Convention*, Amsterdam, The Netherlands, May 1998.

[4] G. Evangelista, "Design of Digital Systems for Arbitrary Sampling Rate Conversion," *EURASIP Journal on Applied Signal Processing*, vol. 83, no. 2, pp. 377–387, 2003.

[5] J. T. Olkkonen and H. Olkkonen, "Fractional Delay Filter Based on the B-Spline Transform," *IEEE Signal Processing Letters*, vol. 14, no. 2, pp. 97–100, Feb. 2007.

[6] A. Silzle, P. Novo, and H. Strauss, "IKA-SIM: A System to Generate Auditory Virtual Environments," in *116th AES Convention*, Berlin, Germany, 2004.

[7] C. W. Farrow, "A Continuously Variable Digital Delay Element," *IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 2641–2645, Jun. 1988.