

# Multirate Extensions for Model-Based Synthesis of Plucked String Instruments

Vesa Välimäki and Tero Tolonen

Helsinki University of Technology  
Laboratory of Acoustics and Audio Signal Processing  
P.O.Box 3000, FIN-02015 HUT, Finland  
vesa.valimaki@hut.fi, tero.tolonen@hut.fi  
<http://www.hut.fi/HUT/Acoustics/>

## Abstract

We describe new developments for model-based sound synthesis of plucked string instruments. The synthesis model is based on the extended Karplus–Strong algorithm. A new implementation approach where part of the excitation signal is modeled with interpolated low-frequency resonators and the decay part of the tones using a multirate string model is introduced. The resonators allow for a full parametrization of the lowest body modes. The computational burden and memory requirements of plucked string synthesis are decreased using the proposed techniques without degrading the sound quality.

## 1 Introduction

Almost 15 years ago, Jaffe and Smith introduced several extensions to the Karplus–Strong algorithm that enabled high-quality synthesis of plucked string tones [1]. This model launched the boom of physical modeling of musical instruments. More details on the physical modeling aspects of string instruments can be found in [2]. Afterwards, many improvements and further extensions have been introduced. These include the commutation of the string and body models to enable computationally efficient and accurate inclusion of body resonances in the excitation signal [3], [4], an analysis technique based on the short-time Fourier transform [5], and a transient-elimination technique to enable smooth glissandos and vibratos when allpass fractional delay filters are used for fine-tuning the pitch [6].

In this paper we describe new developments for the plucked string synthesis algorithm. The guitar is used as a representative of the plucked string instrument family. The common factor in the proposed methods is the use of multirate DSP principles. The lowest resonances of the body of the guitar are synthesized using interpolated low-frequency resonators which allow cheap and fully parametric control of the resonances and enable shortening of the excitation wavetables. Furthermore, the decay part of string tones is generated at a lower sampling rate than the attack. This idea has been proposed in some former publications [4], [5] but to our knowledge it has not been implemented before.

## 2 The String Model

The plucked string synthesis model used in this work is a generalization of the extended Karplus–Strong algorithm. It uses the principle of *commuted synthesis* where

the response of the instrument body has been incorporated in the input signal, together with the pluck excitation [3], [4]. A library of excitation signals for different pluck types and instrument bodies can be stored in wavetables. The string model itself may consist of one or two generalized Karplus–Strong models, each of which simulates one polarization of a vibrating string.

The transfer function of the string model can be expressed as

$$S(z) = \frac{1}{1 - z^{-L_1} F(z) H_1(z)} \quad (1)$$

where  $L_1$  is the integer part of string length,

$$H_1(z) = \frac{g(1+a)}{1+az^{-1}} \quad (2)$$

is called the *loop filter*, and  $F(z)$  is a *fractional delay filter* used for fine-tuning the pitch. The transfer function  $S(z)$  is completely determined by the following three parameters: delay-line length  $L$ , loop filter gain  $g$ , and loop filter cutoff parameter  $a$ . The model is applicable to the synthesis of many members of the string instrument family, such as guitars, the banjo, and the mandolin. For a more detailed description of the synthesis model and its parameters, see [5]. In [7] we explain how to calibrate the parameters of the string model.

## 3 Efficient Modeling of Resonances

This section describes a new technique for reducing the memory requirements of plucked string synthesis. The excitation signal is shortened by removing the lowest body resonances which are reproduced with resonators that can be implemented efficiently.

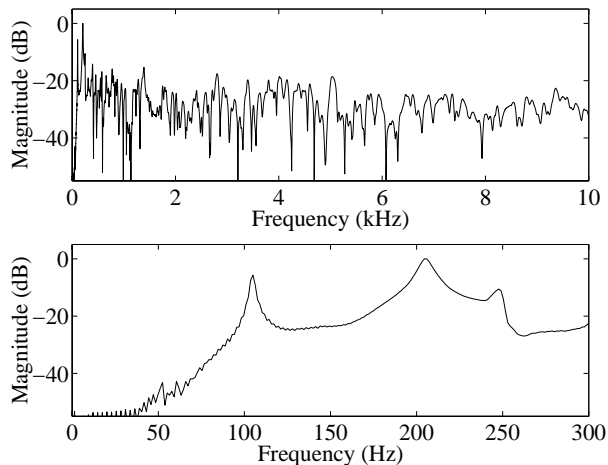


Figure 1: The magnitude response of a guitar body (top) and a zoom to low frequencies (bottom) where the lowest resonances are visible.

### 3.1 Extracting Lowest Body Resonances

The huge number of resonances in the response of the guitar body can be seen in the upper part of Fig. 1. The lowest two resonances are the Helmholtz or air resonance at approximately 100 Hz and the lowest mode of the top plate at around 200 Hz [8]. They are very sharp with a high Q value and are the main cause for the slow decay of the impulse response of the body. These modes may be removed from the excitation wavetable of the guitar model and synthesized separately using second-order resonators. The excitation wavetable is shortened considerably and thus a larger number of excitation signals than before can be stored in a computer memory.

In our example case the two lowest resonances occur at 104.7 and 205.2 Hz (see the lower part of Fig. 1). Also the third lowest resonance at 247.7 Hz has a high Q value, and it may be useful to remove it as well. The resonances may be removed from the residual for example using a second-order notch filter as proposed by Karjalainen and Smith [9]. Alternatively, they may be modeled with sinusoids using the McAulay–Quatieri (MQ) algorithm [10] and subtracted from the residual signal as proposed by Serra [11]. The Q values of the resonances are easily and accurately obtained from the amplitude envelopes of the sinusoidal representation [7].

Figure 2 illustrates the efficiency of the resonance extraction method in shortening the excitation signal. The residual signal of Fig. 2a has been obtained by subtracting the partials of a guitar tone which are modeled with the MQ algorithm. This approach has been found effective for obtaining the excitation signal for plucked string synthesis [7]; the formerly used inverse filtering technique had problems in some cases [5].

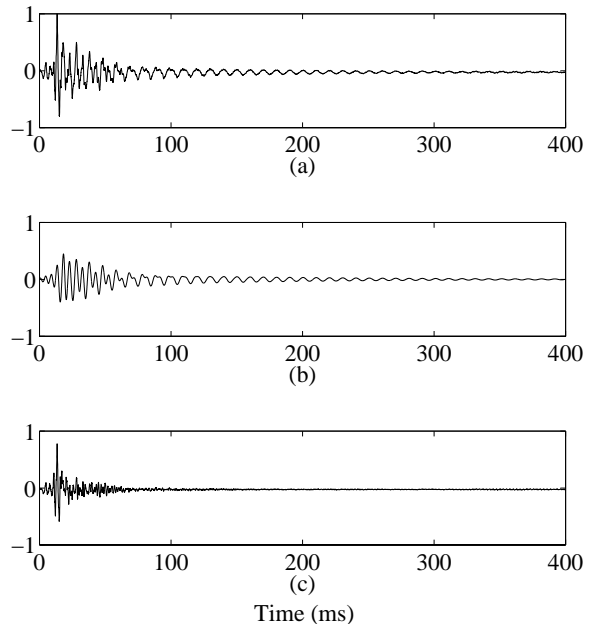


Figure 2: An example of extraction of the excitation signal: (a) the original residual signal, (b) the sinusoidal model of the two lowest body resonances, and (c) the excitation signal after the two resonances have been subtracted.

Figure 2b presents the sinusoidal model of the two lowest body resonances. When it is subtracted from the residual (Fig. 2a), the excitation signal of Fig. 2c is obtained. Note that this signal soon decays to small sample values. In practice we do not subtract all of the sinusoidal model but for example 90% of it, which corresponds to attenuating the resonances by 20 dB. This is enough for shortening the excitation signal but there is still energy left at the center frequency of the resonances so that partials which occur near those frequencies could be excited.

### 3.2 The Truncated Excitation Signal

The original residual of Fig. 2a would have been truncated or windowed to a length of 100 ms (2200 samples at 22 kHz) to include a significant part of the decay of the body resonances in the excitation signal [5]. The processed residual of Fig. 2c, on the other hand, may now be truncated to a length of about 50 ms (1100 samples). When the processed residual is truncated at 50 ms, almost the same relative amount of energy is included as in 100 ms in the case of the original response. It may be possible to use even shorter truncated residuals after the resonances have been removed since the lowest body modes are reproduced using resonators and they are thus not truncated at all. The use of very short excitation signals should be verified by listening to the synthesis results.

### 3.3 Interpolated Low-Frequency Resonators

Since the removed body resonances have very low center frequencies with respect to the sampling frequency, it is unnecessary to synthesize them at the full output sampling rate. Instead we propose a multirate scheme where the resonators run at a much lower sampling rate than the string model. We lower the sampling rate of the resonators by factor  $M$ , which may be about 5 to 10, and then suppress the aliased frequency components with an interpolator. Since the resonances are near to DC (0 Hz), the interpolator only needs to suppress aliasing near the multiples of the lowered sampling rate. The interpolator may then be a *Recursive-Running-Sum* (RRS) filter which can be implemented as zeroth-order hold. We call the combination of downsampled resonators and RRS filters *Interpolated Low-Frequency Resonators* (ILFR).

The ILFRs are based on second-order peak filters, such as those described by Orfanidis [12] (pp. 583–590). All the ILFRs can share a common interpolator. The same signal value is observed at the output of the RRS interpolator for  $M$  consecutive sample cycles. This gives us  $M$  sample periods to compute the next output value. Hence it is advantageous to divide the calculations of the resonators so that the average computational cost per output sample is minimized. This can be done, e.g., by computing only the numerator or the denominator of one resonator at each output sample cycle. If we allow one output cycle for the update of the output value of the RRS interpolator, a suitable downsampling factor  $M$  is 5 for 2 resonators or  $M = 7$  for 3 resonators.

An example of generating a resonance using an ILFR with  $M = 5$  is illustrated in Fig. 3. A sharp resonance at 200 Hz is generated at the sampling rate of 11 kHz. The magnitude response of a downsampled resonator is plotted in Fig. 3a together with the response of an RRS interpolator. Fig. 3b shows their product, i.e., the magnitude response of the ILFR. The dashed line in Fig. 3b shows the magnitude response of a second-order resonator with the same center frequency and Q value. The two responses in Fig. 3b are almost identical in the vicinity of the center frequency but the ILFR response has aliased components at high frequencies. The worst case aliased component which occurs at 2.0 kHz has been attenuated by 19.6 dB. When  $M = 7$ , the strongest aliased component (at 1.4 kHz) would be attenuated by 16.7 dB. If the center frequency is lower than 200 Hz, the aliased components are attenuated more.

The response of ILFRs is summed to the output of the string model and this is why a modest attenuation of less than 20 dB is satisfactory: the aliased components are mixed with other body resonances at high frequencies and also with the partials of the synthetic string

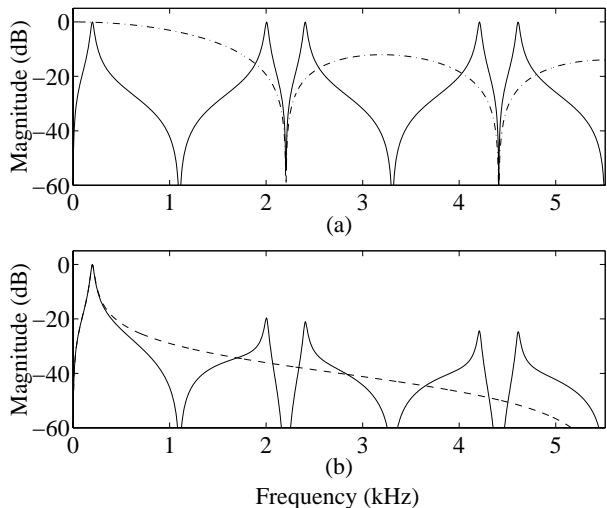


Figure 3: The magnitude responses of (a) a resonator downsampled by factor  $M = 5$  (solid line), an RRS interpolator of length 5 (dash-dot line), (b) an Interpolated LF Resonator (solid line), and a second-order resonator (dashed line) with center frequency 200 Hz and  $r = 0.99$ .

tone. The method described allows for a fully parametric but extremely efficient implementation of the most important body resonances.

### 4 Multirate Synthesis Algorithm

In this section we present a new multirate implementation structure for the string model. The synthesis model is divided into two parts: excitation and decay. The motivation is that the attack portion of a plucked string tone includes more energy at high frequencies than the decay part. Thus, the decay part may be synthesized at a lower sampling rate. This technique has been formerly proposed by Smith [4] and further elaborated by Välimäki *et al.* [5].

In practice, sampling rates of 22 kHz and 11 kHz appear adequate for the attack and decay parts, respectively, to produce synthetic plucked string tones which are virtually identical to the original ones. This may be demonstrated with a simple listening test: a guitar tone is filtered using a highpass and a lowpass filter with a sharp cutoff at 5 kHz. Listening to the filtered signals reveals that the attack portion requires a higher sampling rate than the decay part.

When the multirate realization with full-rate excitation signal and a half-rate decay part simulation is combined with the parallel ILFR bank, a new implementation structure of the string model is obtained (see Fig. 4). Since the full-rate excitation signal  $x_p(n)$  is directly fed into the output, it is necessary to modify the string model to have the following transfer function.

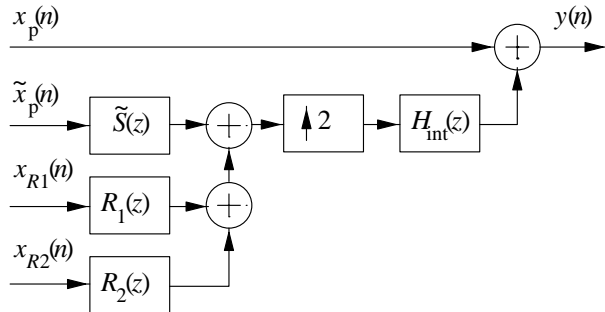


Figure 4: Block diagram of the multirate implementation of the guitar synthesis model.

$$\tilde{S}(z) = \frac{z^{-L_1} F(z) H_1(z)}{1 - z^{-L_1} F(z) H_1(z)} \quad (3)$$

The downsampled excitation signal is fed into the modified string model. Filters  $R_1(z)$  and  $R_2(z)$  which are the ILFRs for simulating the two lowest body modes have their own input signals. The RRS interpolator is included in these filters. After upsampling by factor 2, the output signal of the half-rate system must be lowpass filtered with a halfband filter  $H_{\text{int}}(z)$ .

The halfband filter and the resonators can be shared by all the string models (there may be 6 or 12 string models in a guitar synthesizer). Thus the computational cost per output sample of the complete string instrument model is decreased by almost 50% when using the proposed multirate implementation technique. Other advantages of the proposed technique are that the input wavetables require 25% less memory and the delay lines are 50% shorter than in the full-rate realization.

## 5 Conclusion

New multirate methods for improving the efficiency and parametrization and for reducing memory requirements of model-based plucked string synthesis have been presented. We introduced Interpolated Low-Frequency Resonators (ILFRs) for the synthesis of the most important body resonances. These resonances are removed from the excitation wavetable thereby shortening it. The ILFRs allow for a full parametric control of the lowest body resonances and save memory since the excitation wavetables become shorter. Each ILFR consists of a second-order resonator running at a low sampling rate, such as 10–20% of the sampling frequency of the string model, and a cheap interpolator that suppresses aliasing. In addition, the string model is implemented as a two-rate system, where the decay part of tones is synthesized at half the output sampling rate.

The computational load and memory requirements of guitar synthesis are decreased by almost 50% using the

proposed multirate techniques. In practice this may mean that 12 strings instead of 6 may be simulated in real time using a DSP processor without causing audible degradation in sound quality. In addition, more excitation wavetables may be stored into the on-chip memory of the processor. The model-based synthesis of plucked string tones becomes even more attractive than before.

## References

- [1] Jaffe, D., and J. O. Smith 1983. “Extensions of the Karplus–Strong Plucked-String Algorithm,” *Computer Music J.*, Vol. 7, No. 2, pp. 56–69.
- [2] Smith, J. O. 1983. *Techniques for Digital Filter Design and System Identification with Application to the Violin*. Ph.D. thesis. Report No. STAN-M-14, CCRMA, Stanford University, Stanford, CA.
- [3] Karjalainen, M., V. Välimäki, and Z. Jánosy 1993. “Towards High-Quality Sound Synthesis of the Guitar and String Instruments,” *Proc. ICMC*, pp. 56–63.
- [4] Smith, J. O. 1993. “Efficient Synthesis of Stringed Musical Instruments,” *Proc. ICMC*, pp. 64–71.
- [5] Välimäki, V., J. Huopaniemi, M. Karjalainen, and Z. Jánosy 1996. “Physical Modeling of Plucked String Instruments with Application to Real-Time Sound Synthesis,” *J. Audio Eng. Soc.*, Vol. 44, No. 5, pp. 331–353.
- [6] Välimäki, V., T. I. Laakso, and J. Mackenzie 1995. “Elimination of Transients in Time-Varying All-pass Fractional Delay Filters with Application to Digital Waveguide Modeling,” *Proc. ICMC*, pp. 327–334.
- [7] Tolonen, T., and V. Välimäki 1997. “Automated Parameter Extraction for Plucked String Synthesis,” *Proc. Int. Symp. Musical Acoustics*, Edinburgh, Scotland, Aug. 19–22, 1997.
- [8] Christensen, O., and B. B. Vistisen 1980. “Simple Model for Low-Frequency Guitar Function,” *J. Acoust. Soc. Am.*, Vol. 68, No. 3, pp. 758–766.
- [9] Karjalainen, M., and J. O. Smith 1996. “Body Modeling Techniques for String Instrument Synthesis,” *Proc. ICMC*, pp. 232–239.
- [10] McAulay, R. J., and T. F. Quatieri 1986. “Speech Analysis/Synthesis Based on a Sinusoidal Representation,” *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol. 34, No. 4, pp. 744–754.
- [11] Serra, X. 1989. *A System for Sound Analysis/Transformation/Synthesis Based on a Deterministic plus Stochastic Decomposition*. Ph.D. thesis. Report No. STAN-M-58, CCRMA, Stanford University, Stanford, CA.
- [12] Orfanidis, S. J. 1996. *Introduction to Signal Processing*. Englewood Cliffs, NJ: Prentice–Hall.