Development and Calibration of a Guitar Synthesizer*

VESA VÄLIMÄKI, *AES Member,* **AND TERO TOLONEN,** *AES Student Member*

Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, FIN-02015 HUT, Finland

A specific efficient implementation for the synthesis of acoustic guitar tones is described which is based on recent digital waveguide synthesis techniques. Enhanced analysis methods are used for calibrating the synthesis model based on digitized tones. The lowest resonances of the guitar body are implemented with separate parametric second-order resonators that run at a lower sampling rate than the string model.

0 INTRODUCTION

Digital sound synthesis of musical instruments using physics-based signal-processing techniques has become a popular approach during recent years. Physical-modeling synthesis techniques have been proposed for many musical instruments, such as string and woodwind instruments. The most successful approach to physical-modeling synthesis is the *digital waveguide synthesis* technique developed by Smith [1]. For a review of recent developments in this field, see [2].

In 1983 Jaffe and Smith [3] introduced several extensions to the Karplus–Strong algorithm [4] that permit high-quality synthesis of plucked string tones. These extensions include the use of a one-pole filter in the feedback loop to account for the frequency-dependent damping of the string, a first-order all-pass filter to fine-tune the pitch, simulation of the dependence of the brightness of the tone on the dynamics, and simulation of sympathetic coupling between strings, among others. Certain extensions of the Karplus–Strong algorithm, such as the "pick position" parameter, were guided by physical insights made possible by the physical interpretation of the Karplus–Strong algorithm as a special case of digital waveguide synthesis [5]. This extended model was the starting point of digital waveguide synthesis. More details on the initial ideas of string instrument modeling can be found in [5]. Karjalainen et al. have described in detail how *single-delay loop* synthesis models of the extended Karplus–Strong type can be derived from digital waveguide models of plucked string instruments [6].

Afterward many improvements and further extensions have been introduced to model-based guitar synthesis. These include the use of Lagrange interpolation for fine-tuning the pitch and producing smooth glissandos [7] and all-pass filtering techniques to simulate dispersion due to string stiffness [5], [8], [9]. The commuting of the string and body models has been proposed to enable computationally efficient and accurate inclusion of body resonances in the excitation signal [10]–[12]. A transient-elimination technique was devised to enable smooth glissandos and vibrato when all-pass fractional delay filters are used for fine-tuning the pitch [13]. Välimäki et al. proposed a parameter calibration technique based on the short-time Fourier transform (STFT) and showed that the plucked string model is suitable for analysis-based synthesis of many plucked string instruments [14].

The current paper describes our recent work on guitar synthesis and extends the results presented in [14] by further elaborating on the parameter calibration method. Furthermore, we calibrate a complete synthesis model for one instrument, the classical acoustic guitar. The basic rules in this study have been that the synthesis model must be kept as simple as possible to be feasible for real-time implementation whereas the calibration method must be accurate to enable high-quality resynthesis of the tones analyzed.

In Section 1 we describe a minimal synthesis model for acoustic guitar tones. Further simplification of this model seems difficult since some essential features of the tones would then not be incorporated and thus not parameterized either. The model is based on the commuted waveguide synthesis algorithm for plucked strings, or "commuted waveguide strings" (CWS) for simplicity, developed by Smith [10] and Karjalainen et al. [11], [12]. The least damped resonances are extracted from the excitation signal as suggested previously in

* Presented at the 103rd Convention of the Audio Engineering Society, New York, 1997 September 26–29; revised 1997 October 28 and 1998 May 13.

[14] and [15] and synthesized as interpolated low-frequency resonators [16], [17]. The two main parts of the algorithm used in the present study are thus 1) production of harmonics of string tones using CWS models and 2) production of the lowest resonances of the guitar body using second-order resonators which are realized efficiently at a low sampling rate.

Section 2 concentrates on analysis methods for the resynthesis of plucked string tones. A simple method for extracting the harmonics of a digitally recorded sound is discussed. A sinusoidal representation of the harmonics is subtracted from the original tone, and the remaining residual signal is further processed in order to whiten its spectrum. It is also ensured that the harmonics of a resynthesized tone have the same initial amplitudes as the original, analyzed tone.

Experiments on calibrating the acoustic guitar model to the timbre of a particular acoustic instrument are reported in Section 3. Our aim is high-quality resynthesis of the tones analyzed, but also data reduction. The parameters of the string model are extracted for all strings and frets of the instrument. Low-order polynomial approximations are given for the loop-filter parameters of the synthesis model. Finally, Section 4 concludes the paper.

1 GUITAR SYNTHESIS MODEL

The plucked string synthesis model used in this work is a generalization of the *commuted waveguide string* (CWS) algorithm. The model employs the principle of commutativity of linear and time-invariant elements such that the models of the body and the string are commuted. Thereafter the response of the instrument body has been incorporated in the input signal, together with the pluck excitation [10]–[12]. A library of excitation signals corresponding to different pluck types and instrument bodies can be stored in wavetables.

1.1 String Model

It is known that when a guitar string is plucked, the excitation pulse first radiates efficiently from the body of the instrument, and at the same time standing waves start to form in the string. Thus the attack part of the string tone contains a burst with decaying resonances of the body and growing harmonics. The resonances of the body decay within about 100 ms, but the decay of harmonics is very slow and can take 10–20 seconds. The harmonics, which correspond to the resonances of the string, are, strictly speaking, not exactly harmonically related and should thus be called "partials" instead. For simplicity, we use the term "harmonic" here.

The synthesis model should imitate this phenomenon accurately and efficiently. The harmonics of the string can be generated efficiently with the string model illustrated in Fig. 1 [3], which has the following transfer function:

$$S(z) = \frac{1}{1 - z^{-L_I} F(z) H_l(z)} \tag{1}$$

where L_I is the integral part of string length L (in samples),

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

is the *loop filter*, and $F(z)$ is a *fractional delay filter* used for fine-tuning the pitch. The fractional delay filter may be a first-order all-pass filter, as originally suggested by Jaffe and Smith [3], or a Lagrange interpolation FIR filter [7]. A review of fractional delay filter design methods has been presented in Laakso et al. [18].

The string-model transfer function $S(z)$ is completely determined by the following three parameters: loop delay L, loop-filter gain g, and loop-filter cutoff parameter a. The values of these parameters must be determined for each tone to be synthesized, since they are functions of fundamental frequency and fret number. This issue is discussed further in Section 3.

1.2 Structure of the Guitar Synthesizer

Since 1993 the body resonances have been incorporated into the excitation signal of the synthesizer. This is the principle of commuted synthesis introduced in [10] and independently in [11], [12]. It is a clever way to reduce the computational load of plucked string synthesis drastically since the instrument body need not be simulated with a separate filter, which would have to be of a very high order (see, for example, [14], [19]). The derivation of the CWS model for a single string of the guitar is shown in Fig. 2, where the excitation, the vibrating string, and the radiating body are presented as linear filters with transfer functions $E(z)$, $S(z)$, and $B(z)$, respectively. Since the system is linear and time invariant, we can interchange the order of the components, as shown in the middle of Fig. 2. In practice, it is useful to convolve the impulse responses $b(n)$ and $e(n)$ of the body and excitation models into an aggregated excitation signal denoted by $x_p(n)$ at the bottom of Fig. 2.

The drawbacks of commuting the string and the body are that the body resonances are not explicitly parameterized and that the excitation signal must sometimes be quite long (about 100 ms or more) not to truncate the decay of the least damped body resonances. Thus a large amount of fast memory is required since there must be several excitation signals available (for example, one for each different string and pluck type), and every excitation signal needs more than 2000 words at the sampling rate of 22 kHz.

Recent work has considered further modeling of the

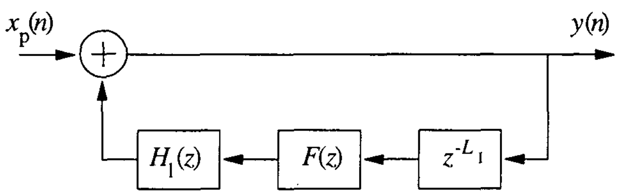


Fig. 1. Block diagram of plucked-string model used in this work [3].

excitation signal, which contains the impulse response of the body. One approach that was proposed in [14] and [19] is to extract the least damped resonances from the excitation signal and implement them with second-order recursive filters. The least damped resonances are in practice the lowest modes of the body. The main advantages of this approach are that the excitation signal becomes shorter and that the lowest body resonances are fully parameterized [14], [19]. In addition Karjalainen and Smith have pointed out that, for example, the signal-to-quantization-noise ratio of the excitation signal becomes larger and that many output points become available when some resonances are realized separately [19].

The resonances can be implemented as a cascade or parallel combination of second-order resonators. The cascade resonators can also be connected in parallel with the string model [17]. The model structures for these choices are depicted in Fig. 3. In [19] a cascade realization is proposed whereas in [15] a parallel resonator bank is used (consisting of fourth-order filters, however). The choice of realization structure is related to the extraction of the remaining excitation signal: inverse filtering of the resonances from the excitation signal corresponds to a cascade realization of Fig. 3(a), which then reconstructs the original excitation signal; subtraction of the resonances (in the time or frequency domain) calls for a parallel realization of Fig. 3(c), which then effectively adds back the resonances. The hybrid model of Fig. 3(b) is a compromise between the cascade and parallel structures: the body resonances are extracted via subtraction, but the model only needs one excitation signal for the resonators. Some of the generality of the control is lost, however, since the excitation character of the two resonators is difficult to control separately with a single excitation signal. In this work we use parallel resonators since the resonances are subtracted from the excitation signal and also since the multirate structure used for implementing these resonators is suited for parallel implementation. The parallel resonator is also easy to use since the maximum amplitude of each resonator is independent of the others in practice.

In summary, the guitar synthesis algorithm used in this work is divided into two parts: harmonics and body resonances. While part of the body response (including the lowest resonances with large Q value) is now produced with separate resonators, the rest of it (resonances with small Q values) is still incorporated in the excitation signal. The model structure is illustrated in Fig. 4. It consists of wavetables of excitation signals for each of the six string models $S_1(z)$, $S_2(z)$, . . . , $S_6(z)$ as well as for the two resonators $R_1(z)$ and $R_2(z)$, which model the most prominent body resonances. Notice that the two body resonators are shared by all the string models. The following section describes a fully parametric but extremely efficient implementation of the resonances.

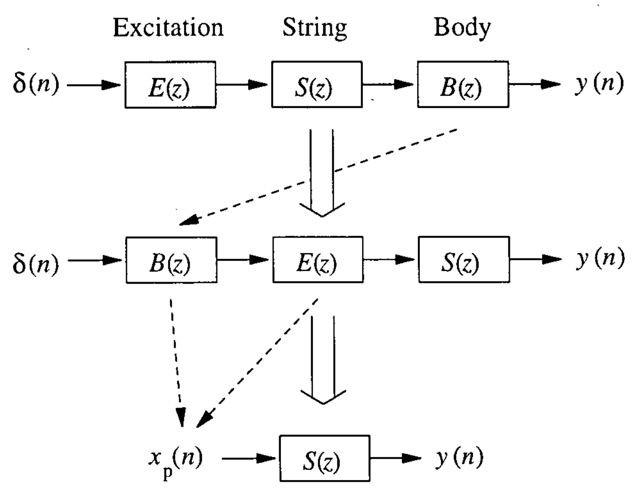


Fig. 2. Derivation of commuted guitar model.

1.3 Efficient Reproduction of Low-Frequency Body Resonances

Since the least damped body resonances that are to be removed from the excitation signal have very low center frequencies with respect to the sampling frequency, it is unnecessary to synthesize them at the full

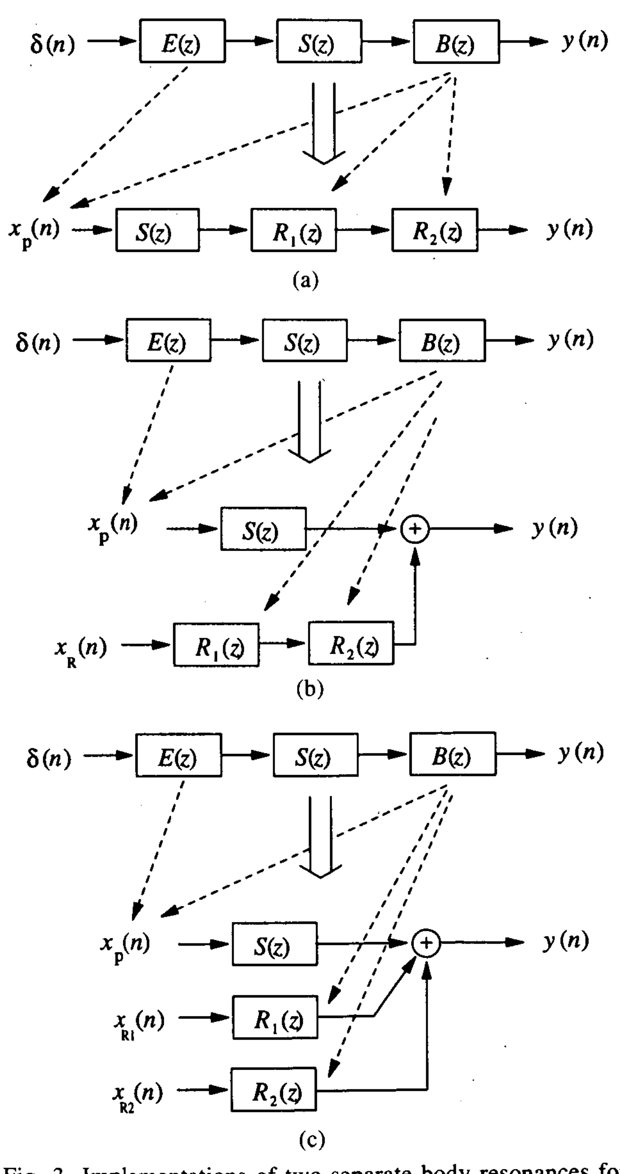


Fig. 3. Implementations of two separate body resonances for commuted waveguide synthesis model [17]. (a) Cascade. (b) Hybrid. (c) Parallel. Dashed arrows indicate where each part of original model (top) contributes in derived model (bottom).

output sampling rate. Instead we use a multirate scheme where the resonators run at a much lower sampling rate than the string model [16], [17]. We lower the sampling rate of the resonators by the factor M, which may be about 5 to 10, then upsample the output of the resonators with a computationally cheap method, and finally suppress the imaging distortion (that is, aliasing) with an interpolator.

Since the center frequencies of the lowest body resonances are very low (typically less than 1% of the sampling frequency at 22 kHz), the interpolator only needs to suppress images near the multiples of the lowered sampling rate, where the images of frequency components near 0 Hz appear. The interpolator may then be a *recursive-running-sum* (RRS) filter [20], which can be implemented as a zeroth-order hold in this case (that is, we simply use the latest output sample to fill the $M - 1$ missing samples). The magnitude response of such a simple interpolator is

$$I(e^{j\omega}) = \left| \frac{\sin(\omega M/2)}{\sin(\omega/2)} \right| . \tag{3}$$

It may be necessary to use more effective suppression than this to make the image components inaudible. One way is to use more RRS filters after the zeroth-order hold operation. These additional interpolators may be implemented with a delay line of M samples and one addition and one subtraction [20]. The transfer function of this filter is

$$I(z) = \frac{1 - z^{-M}}{1 - z^{-1}} . \tag{4}$$

We call the combination of a downsampled resonator and interpolators an *interpolated low-frequency resonator* (ILFR) [16]. The ILFRs are based on second-order peak filters. There are several second-order filter designs available. Since we want to connect the resonators in parallel, we have chosen the following peak filter, which has a small-magnitude response far away from the center frequency:

$$R(z) = (1 - b)\frac{1 - z^{-2}}{1 - 2b\cos(\omega_0)z^{-1} + (2b - 1)z^{-2}} \tag{5}$$

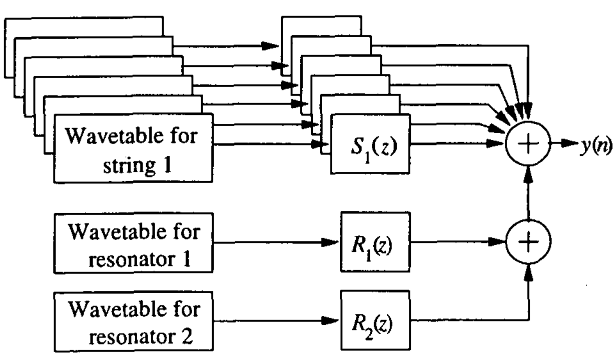


Fig. 4. Structure of guitar model used in this work.

with

$$b = \frac{1}{1 - \tan(\Delta\omega/2)} \tag{6}$$

where ω_0 is the center frequency and $\Delta\omega \approx 2(1 - r)$ is the 3-dB bandwidth of the resonance, with r being the pole radius. For a more detailed discussion on this digital resonator, see Orfanidis [21, pp. 248–253, 583–590].

The peak filter for the ILFR is designed as follows:

1) Design a second-order peak filter to match the desired resonance.

2) Choose the downsampling factor M.

3) Modify the parameters of the second-order resonator to restore the impulse response at every Mth sample (multiply ω_0 by M and raise the pole radius r to the Mth power, that is, r^M).

All the downsampled resonators can share a common interpolator for upsampling their output signals. 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, for example, 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, suitable downsampling factor are $M = 5$ for two resonators or $M = 7$ for three resonators.

An example of generating a resonance with a low center frequency using an ILFR with $M = 7$ is illustrated in Fig. 5. A sharp resonance at 200 Hz is generated at the sampling rate of 22 kHz. The magnitude response of a downsampled resonator is plotted in Fig. 5(a) together with the response of a single RRS interpolator. The resonator has first been downsampled by the factor 7 and then upsampled using zero fill. Fig. 5(b) shows the magnitude response of the ILFR when two RRS filters are used [that is, zeroth-order hold and a filter implemented with the transfer function of Eq. (4)]. The dashed line in Fig. 5(b) shows the magnitude response of a second-order resonator with the same center frequency and Q value. The two responses in Fig. 5(b) are almost identical in the vicinity of the center frequency, but the ILFR response has suppressed image components at high frequencies. If the center frequency is lower than 200 Hz, the image distortion components are attenuated more. If the center frequency of the resonator is higher, the performance of the ILFR is worse than in this example. The ILFR is only useful for implementing resonators whose center frequency is very small with respect to the sampling rate.

1.4 Missing Features and Possible Extensions

It is easily observed that the synthesis model described in previous sections is a simplified one, and it could be extended in a number of ways, many of which have been discussed in the earlier literature. In this section we review some possibilities and reasons for omitting them.

Sympathetic coupling of the guitar strings refers to the phenomenon where the vibration of one string excites the others through the bridge. It can be included in the synthesis model by feeding the output signal (multiplied with a very small gain factor) of each string into the input of the other strings [3]. A new structure for implementing sympathetic coupling has been proposed in [6] and [22]. The analysis of the gain factors for sympathetic coupling simulation is possible, but should be done separately for each fret position of all strings. In principle, it would be necessary to know the transfer function from each string to all the other strings, but in practice it seems sufficient to measure the magnitude of the transfer function for every note. This experiment is left for future work.

The loop filter in our model is of first order and has only one pole (plus a trivial zero at $z = 0$). Such a filter is merely capable of simulating roughly the average decay rate of harmonics. A higher order loop filter could be used if it were of interest to imitate the decay rate of individual harmonics more carefully. The quality of synthesized tones has been considered very good, even with a first-order loop filter (see, for example, [14]), and thus this filter is still used in this study. Also the phase response (not only the magnitude) of a higher order loop filter could be optimized to account for the frequency-dependent propagation speed (dispersion) of vibration along the string. While this is feasible and filter design methods are readily available for such an approximation (see [23]), we have noticed earlier that usually the slight inharmonicity of the partials of guitar tones is not very important perceptually.

The horizontal and vertical modes of transverse string vibration can be simulated by using two string models that are linearly coupled [10]. When the string lengths of these models are not exactly the same, as is appropriate for acoustic string instrument models (since the bridge is stiffer for horizontal than for vertical waves), beats will be observed in the output signal. This technique is easy to use and produces very realistic sounds. The proper detuning can be chosen by experiment.

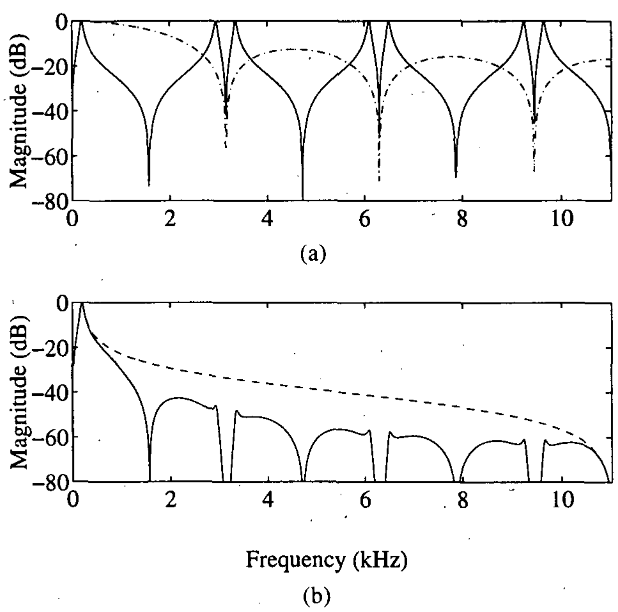


Fig. 5. Magnitude responses. (a) Resonator downsampled by a factor $M = 7$ (solid line), and RRS interpolator (dash-dot line). (b) Interpolated LF resonator with two RRS filters (solid line), and response of conventional, full-rate second-order resonator (dashed line).

The plucking point is one feature that is often included in physical models of strings. It consists of a feedforward comb filter [3]. We have not included it here since a robust analysis technique for the plucking point is not available. Attempts to extract this feature from digitized plucked string tones have been discussed in [14], [15], but in our experiments the results have not been reliable. Thus the influence of the plucking point is now fixed since it is incorporated in the excitation signal.

One of the most fascinating possibilities in string synthesis and physical modeling in general is the use of nonlinearities. Pierce and Van Duyne have proposed a passive nonlinear filter for simulating a nonlinearly behaving termination of a vibrating string [24]. However, techniques for analyzing the parameters of such a model are yet to be developed. Nonlinear extensions are thus not included in the model that we use for resynthesis here.

2 PARAMETER ESTIMATION TECHNIQUES

In this section we present analysis and parameter estimation methods for the simple guitar model described in Section 1. Techniques to calibrate the two parts of the synthesis model (the string model and the resonator bank) are given. In both cases the procedure starts with a time-varying spectrum analysis. Inverse filtering and digital filter design techniques are also applied.

2.1 Analysis of Harmonics

In [14], time-varying analysis of the partials was achieved with an STFT-based method with a peak picking algorithm. In this study we use the heterodyne filtering technique [25] to extract the partials of digitized guitar tones. This technique has been used previously for estimating the frequency and amplitude tracks for additive synthesis [25], [26]. It is much easier to implement and is computationally less intensive than the STFT-based method. While the technique is not as general as the STFT-based one, it is good enough for our purposes, namely, to analyze the amplitude variations of the partials of nearly harmonic tones.

The analysis procedure is pictured in Fig. 6. First, the fundamental frequency is estimated using one of the existing techniques (such as the well-known technique based on the autocorrelation function, which was also used in [14]). The nominal string length in samples is computed as

$$L = f_s / \hat{f}_0 \tag{7}$$

where f_s is the sampling rate and $\hat{f}_0$ the estimated fundamental frequency. The partials of guitar tones are not exactly harmonically related, but still we may apply the heterodyne filter around the multiples of the fundamental frequency to extract the partials. A better approximation of the frequencies of the partials for heterodyning is obtained by detecting peak locations in a discrete Fourier transformation. This detection procedure uses integral multiples of the fundamental frequency estimate $\hat{f}_0$ as initial guesses for frequencies of the partials.

In the next step, the frequency estimates are used as center frequencies of the heterodyne filtering procedure. In heterodyne filtering the signal is multiplied with a complex exponential, and this operation is repeated for each partial to be extracted. This shifts the desired frequency band around the frequency of the partial so that the center frequency becomes zero. The complex-valued signal is low-pass filtered to suppress other frequency components. We have used a low-pass filter with a cutoff frequency depending on the fundamental frequency of the tone. The low-pass filter should be narrow, but still wide enough to accept slight variations in the frequency track of the partial. A second-order Butterworth low-pass filter was used to filter the signal twice, first in a conventional manner and then backward (as implemented in MATLAB's filtfilt.m function), so that the filter effectively has a zero-phase fourth-order transfer function. The amplitude envelope of the partial is obtained by taking the absolute value of the low-pass-filtered signal. The phase track is computed by unwrapping the imaginary part of the logarithm of the low-pass-filtered signal. A sinusoidal model of the analyzed signal may then be computed based on the amplitude and phase tracks as presented in [25], [26].

The next step in the analysis is the detection of the decay rate of the partials. The amplitude envelopes are smoothed using the backward integration method of Schroeder [27] as generalized to an STFT representation by Jot [28]. The use of this technique has been suggested earlier by Laroche and Meillier in a related study [29]. An example is shown in Fig. 7, where the amplitude trajectories of three partials are smoothed. The level of the trajectories has been normalized after the backward integration.

After smoothing, we use the same approach as in [14], where a straight line is fitted to the amplitude envelope of each partial on a logarithmic (dB) scale. This method has been found robust in practice. The loop gain of the string model at the frequencies ω_k of the harmonics may now be computed based on the slope β_k of these straight lines,

$$G_k = 10^{\beta_k L/20} \tag{8}$$

where k is the number of the harmonic component and L is the nominal string length in samples.

The G_k values form the prototype magnitude response of the loop filter [14]. The g and a parameters of a one-pole loop filter are fitted to this response. We use a weighted least-squares algorithm where the weight function is larger at the frequencies with a large loop gain G_k.

2.2 Excitation Signal

In commuted synthesis the instrument model is excited with a signal consisting of the pluck excitation and the response of the instrument body [10]–[12]. In [14] a recorded guitar tone was inverse filtered with the reciprocal of the string model. This time-invariant method is applicable to well-behaved signals while problems arise when nonexponential decay or shifts in the frequencies of the partials occur. In those cases the inverse-filtered signal is corrupted with "ringing" of the remaining partials that could not be removed. However, it should be noted that this method is optimal in the sense that it

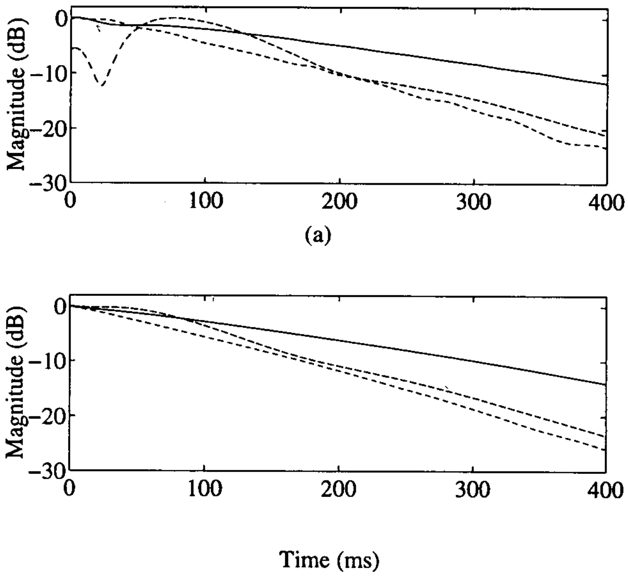


Fig. 7. (a) Amplitude trajectories of three partials obtained by heterodyne filtering. (b) Corresponding smoothed trajectories computed using Schroeder backward integration.

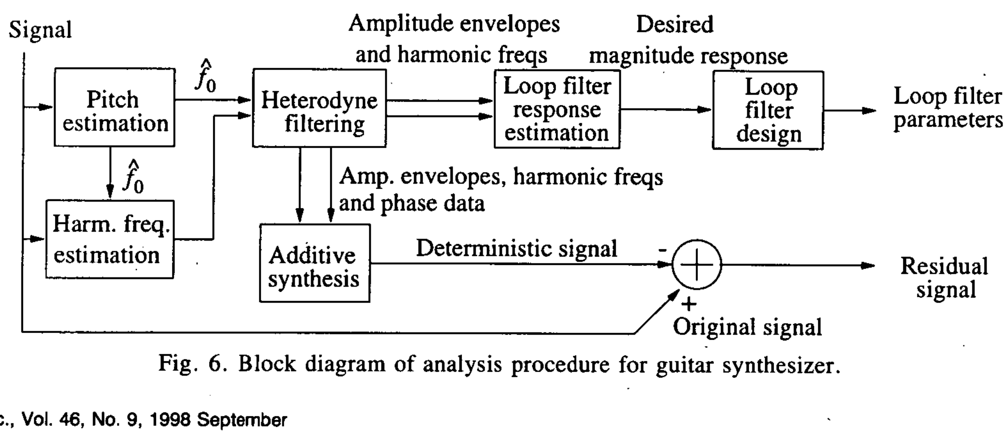


Fig. 6. Block diagram of analysis procedure for guitar synthesizer.

reproduces the original tone within limits of numerical accuracy if the string-model parameters are the same as those of the inverse string model. In general, this is not the case, but the model parameters change as the user controls the instrument. It would be particularly desirable to use a common excitation signal for several fundamental frequencies.

In Section 2.1 a sinusoidal model of the harmonics of a recorded signal was described for estimation of the string-model parameters. This model can also be used to obtain an excitation signal. The heterodyne filtering is very efficient in canceling the harmonics of the recorded signals. In fact, there is not enough energy left in the residual at the frequencies of the partials to excite the string model in a proper way. This can be avoided by postprocessing the residual signal, as explained in the following.

The residual is obtained by a time-domain subtraction as (see Fig. 6)

$$y_{\mathrm{res}}(n) = y(n) - y_{\mathrm{det}}(n) \tag{9}$$

where the sinusoidal model $y_{\mathrm{det}}(n)$ is called the *deterministic part* of the original signal, following the terminology of Serra and Smith [30]. The inverse filtering of [14] can now be represented as

$$Y_{\mathrm{exc}}(z) = Y(z)S^{-1}(z) = [Y_{\mathrm{res}}(z) + Y_{\mathrm{det}}(z)]S^{-1}(z) \tag{10}$$

where $Y_{\mathrm{exc}}(z)$ is the z transform of the excitation signal $y_{\mathrm{exc}}(n)$, $S^{-1}(z)$ is the reciprocal of the string-model transfer function $S(z)$, and $Y_{\mathrm{res}}(z)$ and $Y_{\mathrm{det}}(z)$ are the z transforms of the residual signal $y_{\mathrm{res}}(n)$ and the deterministic part $y_{\mathrm{det}}(n)$, respectively. In practice, we use the rightmost form of Eq. (10) and inverse filter the residual and the deterministic part separately with the inverse string model. The sum of these two signals is equivalent to the inverse-filtered signal, but we have now decomposed the residual signal into two components, which can be processed separately if desired. For example, one possibility is to truncate the inverse-filtered deterministic part before adding it to the excitation signal. Then the dips in the spectrum of the residual are "filled" in the attack part of the signal, but the "ringing" of the remaining higher partials is suppressed.

The most critical part of the signal to the perceived tone quality of plucked string tones is the attack. When the inverse-filtered sinusoidal model is truncated using windowing right after the attack and added to the inverse-filtered residual, an excitation signal that produces an exact copy of the attack is still produced. After the attack the reproduced signal is not an exact match to the original signal, but our informal listening tests suggest that the synthetic signal is in many cases indistinguishable from the original.

Let us consider an example where the excitation signal is computed. The original signal, a sinusoidal model representing its deterministic part, and their difference signal are presented in Fig. 8. The effect of noncausal zero-phase low-pass filtering can be observed in Fig. 8(b), where the signal is spread before the attack. This has to be compensated to retain the sharpness of the attack by applying a correction term. Fig. 9(a) shows the difference signal after it has been inverse filtered with the string model that has been designed as described in Section 2.1. Fig. 9(b) presents the deterministic part (the sinusoidal model) after inverse filtering. In Fig. 9(c) the inverse-filtered deterministic part has been windowed to a length of 500 samples (22.7 ms). The windowed inverse-filtered sinusoidal model is the correction that is used for equalizing the attack part of the excitation signal. It also restores the original attack in the inverse-filtered signal. In Fig. 9(d) the correction term has been added to the inverse-filtered residual signal. Notice the different amplitude scales in Fig. 9.

It is also instructive to look at the spectra of the signals of this example. Fig. 10 shows the magnitude spectrum of the original recorded guitar tone, where the harmonics are shown as peaks. The two peaks at low frequencies at the level of around 110 dB are the two lowest resonances of the guitar body. Fig. 11 is the spectrum of a residual obtained by the inverse-filtering method proposed in [14]. It is seen that some of the highest partials are not canceled accurately. Fig. 12 presents the residual spectrum when the deterministic part has been subtracted in the time domain. Now deep dips occur at the frequencies of the harmonics (most easily visible at low frequencies below 2 kHz). The magnitude spectrum of the inverse-filtered deterministic part is shown in Figs. 13 and 14 without and with windowing, respectively. Har-

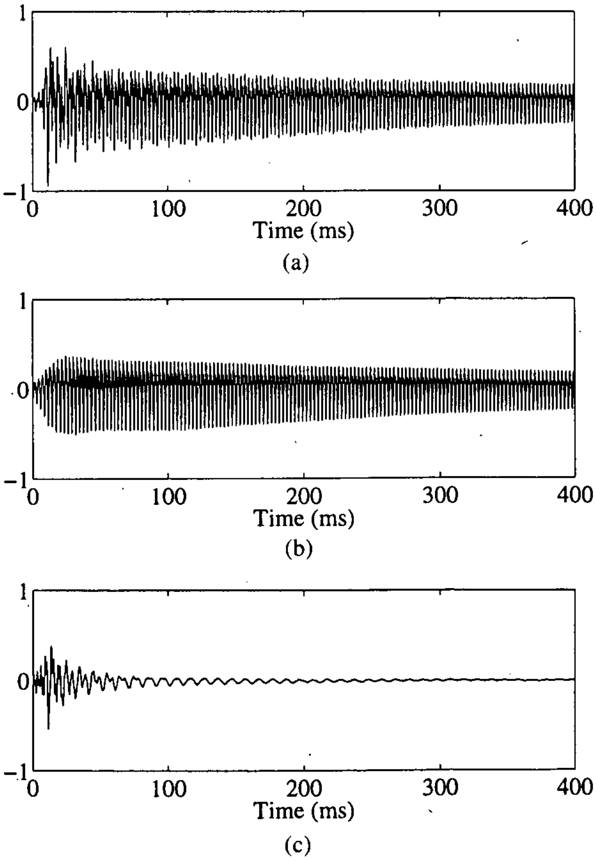


Fig. 8. (a) Original digitized guitar tone. (b) Its deterministic part (sinusoidal model). (c) Their difference.

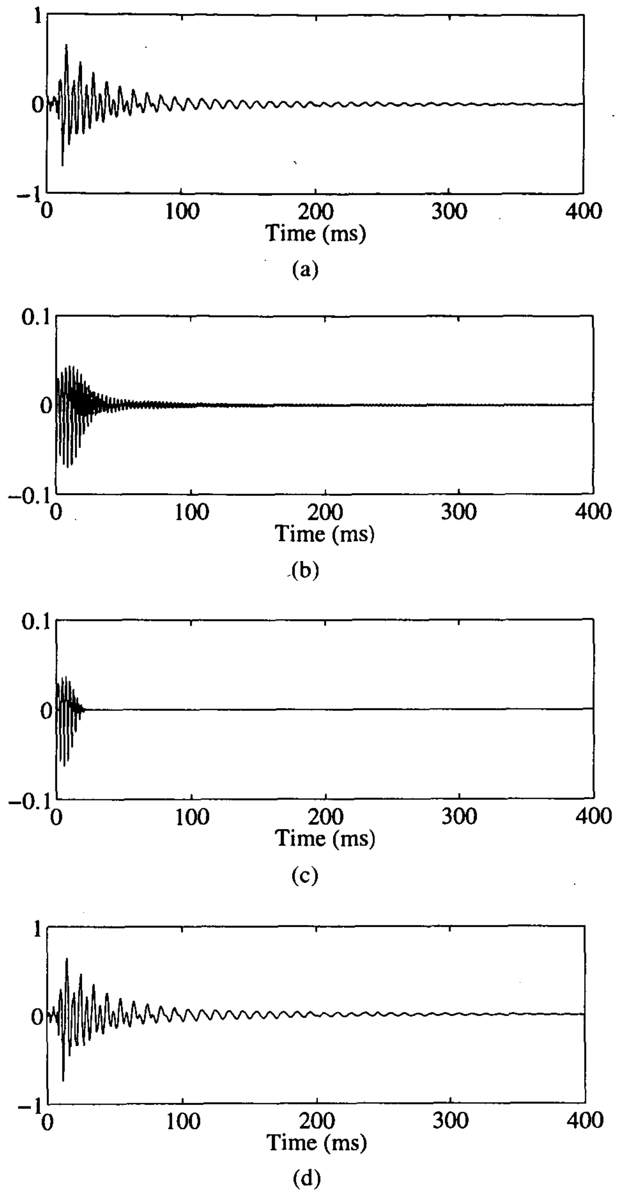


Fig. 9. (a) Inverse-filtered residual signal. (b) Inverse-filtered sinusoidal model. (c) Inverse-filtered and windowed (with 500-sample right wing of a Hanning window) sinusoidal model used for equalization. (d) Sum of signals (a) and (c).

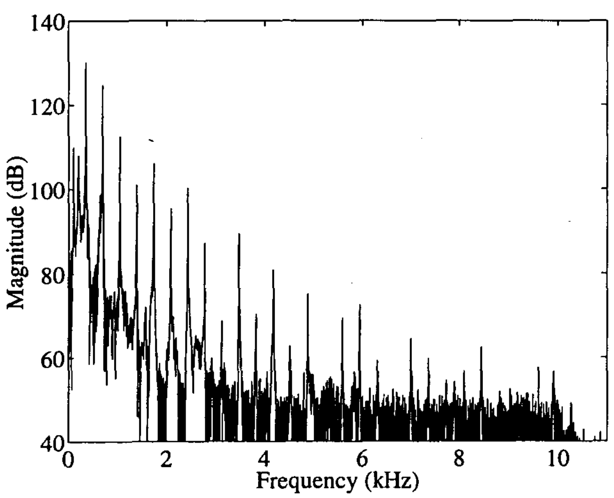


Fig. 10. Magnitude spectrum of recorded guitar tone (string 1, fret 1, f_0 = 347 Hz). [see Fig. 8(a)].

monics are still seen as sharp peaks in these figures. In Fig. 15 the residual has been corrected by adding the inverse-filtered and windowed (with the 500-sample right wing of a Hanning window) deterministic part. Now the dips have been removed but the partials have not reappeared.

2.3 Analysis of Lowest Body Resonances

There is a huge number of resonances in the response of the guitar body. An example frequency response is shown in [14, fig. 13]. The lowest two resonances of the guitar body are the Helmholtz or air resonance at approximately 100 Hz and the lowest mode of the top plate at around 200 Hz [31]. 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, as discussed in Section 1.

The resonances may be removed from the residual, for example, by using a second-order notch filter, as proposed by Karjalainen and Smith [19]. Alternatively, they may be modeled with sinusoids using the McAu-

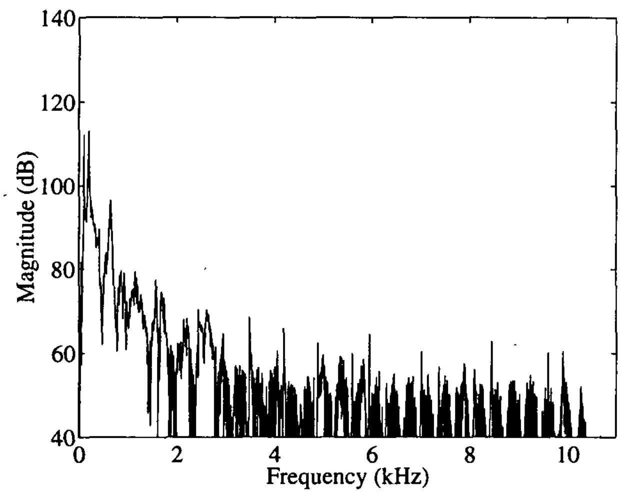


Fig. 11. Magnitude spectrum of inverse-filtered guitar tone obtained using inverse string model, as suggested in [14].

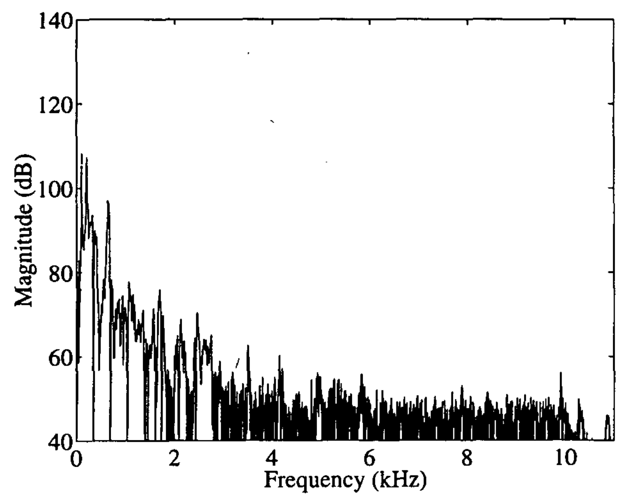


Fig. 12. Magnitude spectrum of residual signal after harmonics have been removed using heterodyne filtering [see Fig. 8(c)].

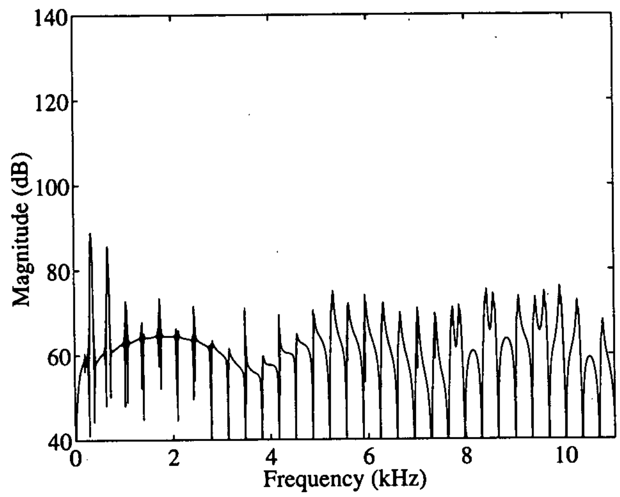


Fig. 13. Magnitude spectrum of inverse-filtered deterministic part [see Fig. 9(b)].

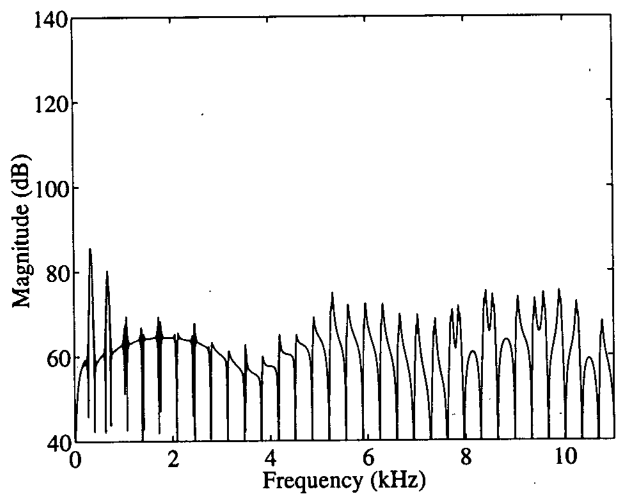


Fig. 14. Magnitude spectrum of inverse-filtered deterministic part after windowing with right half of a Hanning window [see Fig. 9(c)].

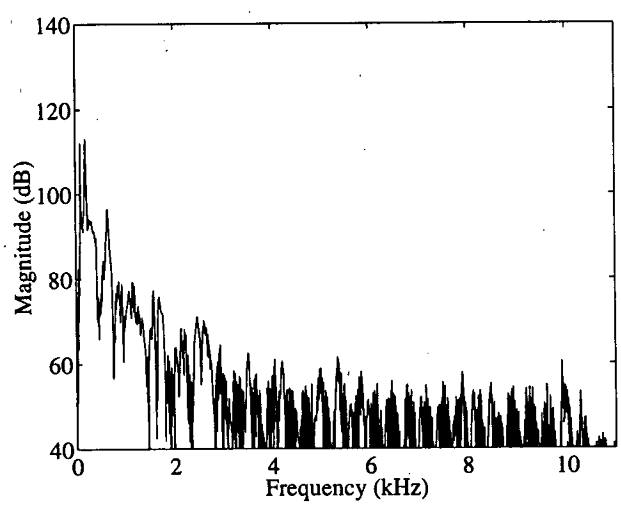


Fig. 15. Magnitude spectrum of corrected residual signal of Fig. 9(d) [that is, signal of Fig. 9(c) has been added to signal of Fig. 9(a)].

lay–Quatieri (MQ) algorithm [32] and subtracted from the residual signal, similar to what is done in spectral modeling synthesis of Serra and Smith [30], [33]. The Q values of the resonances are obtained easily and accurately from the amplitude envelopes of the sinusoidal representation [34].

Fig. 16(a) illustrates the efficiency of the resonance extraction method in shortening the excitation signal. The signal of Fig. 16(a) has been obtained by subtracting the sinusoidal models of the two lowest body resonances from the signal of Fig. 9(c). Note that the signal in Fig. 16(a) soon decays to small sample values. It is now possible to use an excitation signal of about 50 ms in length. When all the body resonances are included, the length of the excitation signal must be about 100 ms [14]. Thus the removal of the lowest resonances helps to save about 50% of memory in commuted waveguide synthesis of the guitar.

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 that occur near those frequencies could be excited. Fig. 17 shows the magnitude spectrum of the excitation signal when the body resonances have been extracted.

The synthetic body resonances (center frequencies around 100 and 200 Hz) reproduced using the ILFR

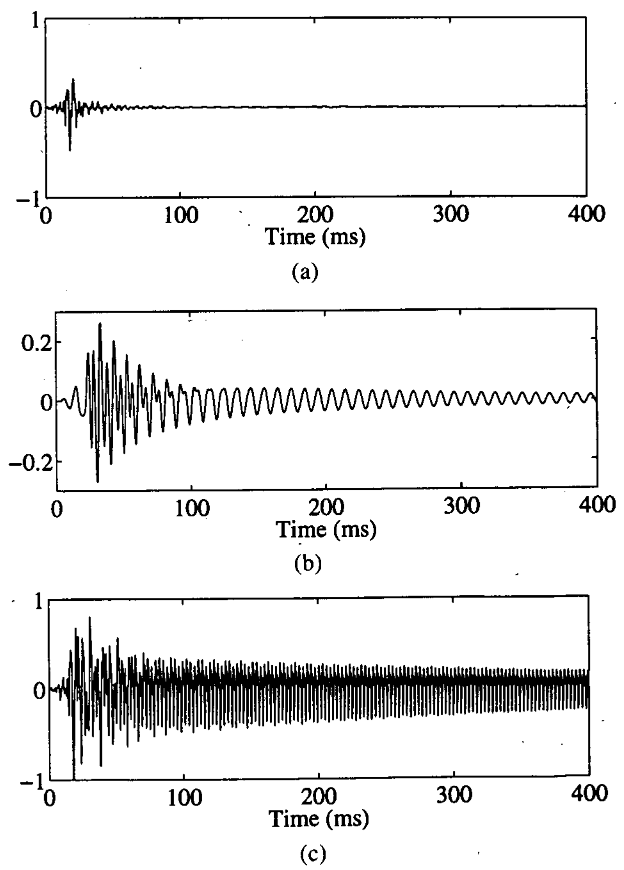


Fig. 16. (a) Excitation signal. (b) Synthetic responses of lowest two body resonances. (c) Resynthesized signal to be compared with Fig. 8(a).

technique are shown in Fig. 16(b). In this case the resonator excitation signals have been obtained by inverse filtering the sinusoidal models of the body resonances. The inverse filter has been designed so that it has a similar frequency response in the neighborhood of the resonance frequencies as the inverse transfer function of the resonators, which are asymptotically unstable, however. Details of this method have been presented in [17]. The excitation signal derived in this way brings about a typical slow attack of the body resonances, as seen in Fig. 16(b).

Fig. 16(c) shows the waveform of the resynthesized tone, which should be compared with Fig. 8(a). It can be seen that the attack parts of these signals are nearly identical, and also the decay of the synthetic signal corresponds well to the original.

3 CALIBRATION OF THE STRING-MODEL PARAMETERS

This section describes an experimental study of the calibration of the string-model parameters. A professional guitar player played several tones on all frets of every string in an anechoic chamber. These sounds were digitally recorded and are used for calibrating the model.

The string model presented in Fig. 1 is fully described by three parameters, namely, the delay-line length L, the loop-filter gain g, and the loop-filter cutoff parameter a. Starting with the most fundamental of these parameters, the delay-line length L determines the fundamental frequency of the tone. The mapping of the fundamental frequency to the delay-line length is given by Eq. (7).

The loop-filter parameters a and g determine the quality of the tone decay. They correspond to physical losses, such as the friction with air, nonrigid string terminations, and internal viscoelastic losses [35]. The model presented here does not attempt to model these physical phenomena separately. Furthermore, the simple one-pole loop filter is not capable of producing exact matches of decay rates of every partial of a tone. Nevertheless, the loop filter produces a perceptively good match of the overall decay characteristics of the instrument at every fret position. This suggests that the parameter values obtained by the analysis procedure described in the previous section could be approximated by mathematical formulas. In this section an example of creating polynomial approximations for the loop-filter parameters is presented.

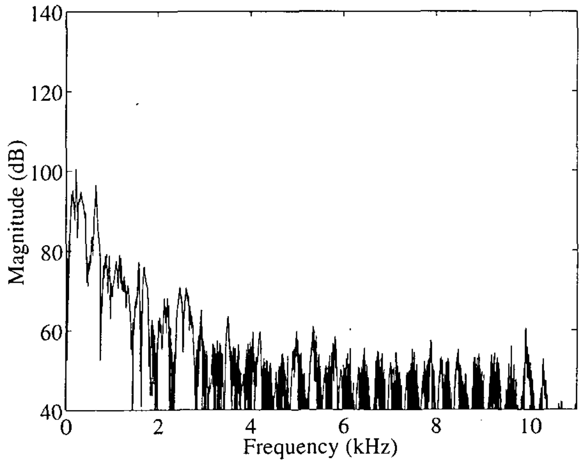


Fig. 17. Spectrum of excitation signal of Fig. 16(a). Harmonics have been removed and equalized, and body resonances extracted.

3.1 Polynomial Approximation of Loop-Filter Coefficients

Fig. 18 shows an example of the measured loop-gain parameters g for string 1 of the acoustic guitar. We have fitted a first-order polynomial to these data as a function of the fret number m_{fret}. In this case the straight line is logarithmic as a function of frequency, since the frequency obtained at a given fret is approximately

$$f = f_{\text{fs}} 2^{m_{\text{fret}}/12} \tag{11}$$

where f_{fs} is the fundamental frequency of the free string whose fret number is 0. The variation of the loop gain, on the other hand, depends mostly on the internal losses in the string while the losses due to end supports remain almost constant. Thus there is a constant part in the loop gain which is multiplied by a varying part that is proportional to the string length and thus to the fret number. The first-order polynomial is shown in Fig. 18 by a dashed line.

Table 1 presents the polynomials for the loop-gain coefficients of the six strings of the acoustic guitar. The loop gain is obtained by evaluating the polynomial as

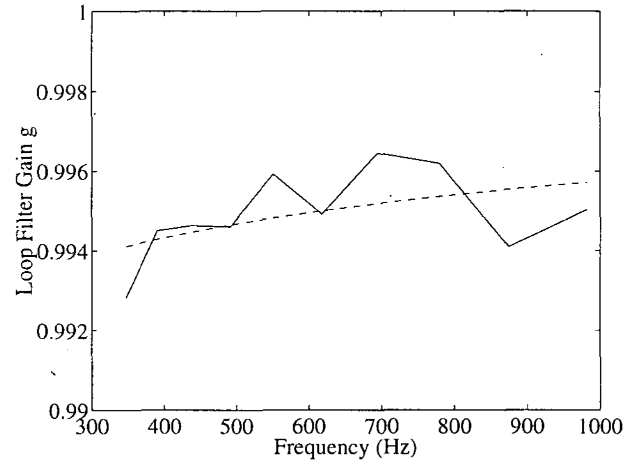


Fig. 18. Loop gain g for modeling string 1 (solid line) and first-order polynomial fit (dashed line) as a function of fret number m_{fret}.

Table 1. Coefficients for first-order polynomial fit to loop-gain data.

	$g_{\text{pol}}(0)$	$g_{\text{pol}}(1)$
String 1	0.99402123928178	0.00008928138142
String 2	0.99247813966550	0.00012644399078
String 3	0.99012478445221	0.00025250158133
String 4	0.98780640700360	0.00037712305083
String 5	0.98347976839019	0.00040239847018
String 6	0.97816203269973	0.00061375406757

follows:

$$g = g_{pol}(0) + g_{pol}(1)m_{fret}\,. \qquad (12)$$

The dependence of the loop-filter coefficient a on frequency is more complicated than this. The coefficient value depends on the frequency-dependent characteristics of string losses and also on the loop-filter design method. Anyway, we have experimentally observed that a first-order polynomial approximation method used for the loop-gain parameter is also applicable here. Fig. 19 illustrates an example of estimated a parameters (solid line) and a first-order polynomial fit (dashed line) to these data. Table 2 gives the polynomial coefficients for the loop-filter coefficients. The coefficient is now evaluated in the same way as the loop-gain parameter, that is,

$$a = a_{pol}(0) + a_{pol}(1)m_{fret}\,. \qquad (13)$$

4 CONCLUSIONS

In this paper we have presented an analysis and synthesis procedure for producing guitar tones. A simple guitar model based on commuted waveguide synthesis is used. The parameters of the synthesis model include the delay-line length and two loop-filter coefficients of the string model, coefficients of two resonators used for modeling the lowest body resonances, and excitation signals for the string and the resonance models. The two lowest body resonances have been implemented as interpolated low-frequency resonators for increased computational efficiency.

An alternative analysis method to extract the parameter values from a recorded tone of the acoustic guitar has been introduced. The technique is based on sinusoidal modeling of guitar tones. The loop-filter coefficients are optimized to simulate the frequency-dependent decay rate of the analyzed tone. The excitation signal of the string model is obtained by subtracting the sinusoidal model from a recorded signal and by correcting the spectral content of the resulting signal corresponding to the attack. This approach improves the quality of the excitation signal used in commuted guitar synthesis. Furthermore, polynomial approximations of the loop-filter parameters have been presented to enable easy calibration of the synthesizer. The methods described in this paper enable an efficient and accurate synthesis of guitar tones.

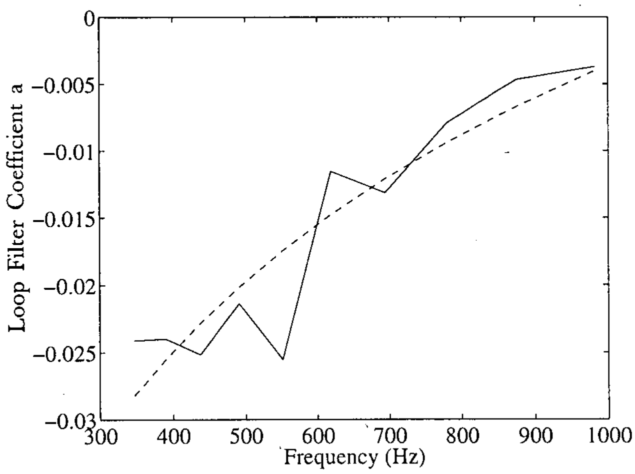


Fig. 19. Loop-filter coefficient a for string 1 (solid line) and first-order polynomial fit (dashed line) as a function of fret number m_{fret}.

Table 2. Coefficients for first-order polynomial fit to filter coefficients.

	$a_{pol}(0)$	$a_{pol}(1)$
String 1	−0.02955827361150	0.00134421335136
String 2	−0.03042891937178	0.00113090288951
String 3	−0.03840938807507	0.00081125415233
String 4	−0.06091679973956	0.00298025530804
String 5	−0.05928143968051	0.00171045642780
String 6	−0.08135045114297	−0.00085796015850

5 ACKNOWLEDGMENT

The authors wish to thank Julius O. Smith (Stanford University, Stanford, CA) and Matti Karjalainen (Helsinki University of Technology, Espoo, Finland) for helpful discussions and criticism. The guitar tones used for calibrating the synthesizer were played by Jukka Savijoki (Sibelius Academy, Helsinki, Finland). The help of Jyri Huopaniemi (Stanford University, Stanford, CA) in anechoic recordings is also acknowledged.

6 REFERENCES

[1] J. O. Smith, "Acoustic Modeling Using Digital Waveguides," in C. Roads, S. T. Pope, A. Piccialli, and G. De Poli (eds.), *Musical Signal Processing* (Swets & Zeitlinger, Lisse, The Netherlands, 1997), pp. 221–263.

[2] J. O. Smith, "Physical Modeling Synthesis Update," *Computer Music J.*, vol. 20, no. 2, pp. 44–56 (1996).

[3] D. A. Jaffe and J. O. Smith, "Extensions of the Karplus–Strong Plucked-String Algorithm," *Computer Music J.*, vol. 7, no. 2, pp. 56–69 (1983).

[4] K. Karplus and A. Strong, "Digital Synthesis of Plucked-String and Drum Timbres," *Computer Music J.*, vol. 7, no. 2, pp. 43–55 (1983).

[5] J. O. Smith, "Techniques for Digital Filter Design and System Identification with Application to the Violin," Ph.D. thesis, Rep. STAN-M-14, CCRMA, Department of Music, Stanford University, Stanford, CA (1983 June).

[6] M. Karjalainen, V. Välimäki, and T. Tolonen, "Plucked-String Models: from the Karplus–Strong Algorithm to Digital Waveguides and Beyond," *Computer Music J.*, vol. 22, no. 3 (1998).

[7] M. Karjalainen and U. K. Laine, "A Model for Real-Time Sound Synthesis of Guitar on a Floating-Point Signal Processor," in *Proc. 1991 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 5 (Toronto, ON, Canada, 1991), pp. 3643–3656.

[8] A. Paladin and D. Rocchesso, "A Dispersive Resonator in Real Time on MARS Workstation," in

Proc. 1992 Int. Computer Music Conf. (San Jose, CA, 1992 Oct.), pp. 146–149.

[9] S. A. Van Duyne and J. O. Smith, "A Simplified Approach to Modeling Dispersion Caused by Stiffness in Strings and Plates," in *Proc. 1994 Int. Computer Music Conf.* (Aarhus, Denmark, 1994 Sept.), pp. 407–410.

[10] J. O. Smith, "Efficient Synthesis of Stringed Musical Instruments," in *Proc. 1993 Int. Computer Music Conf.* (Tokyo, Japan, 1993 Sept.), pp. 64–71.

[11] M. Karjalainen and V. Välimäki, "Model-Based Analysis/Synthesis of the Acoustic Guitar," in *Proc. Stockholm Music Acoustics Conf.* (Stockholm, Sweden, 1993 July 28–Aug. 1), pp. 443–447.

[12] M. Karjalainen, V. Välimäki, and Z. Jánosy, "Towards High-Quality Sound Synthesis of the Guitar and String Instruments," in *Proc. 1993 Int. Computer Music Conf.* (Tokyo, Japan, 1993 Sept.), pp. 56–63.

[13] V. Välimäki, T. I. Laakso, and J. Mackenzie, "Elimination of Transients in Time-Varying Allpass Fractional Delay Filters with Application to Digital Waveguide Modeling," in *Proc. 1995 Int. Computer Music Conf.* (Banff, AB, Canada, 1995 Sept.), pp. 327–334.

[14] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy, "Physical Modeling of Plucked String Instruments with Application to Real-Time Sound Synthesis," *J. Audio Eng. Soc.*, vol. 44, pp. 331–353 (1996 May). Sound examples available at <URL: http://www.acoustics.hut.fi/>.

[15] K. Bradley, M. Cheng, and V. L. Stonick, "Automated Analysis and Computationally Efficient Synthesis of Acoustic Guitar Strings and Body," in *Proc. 1995 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics* (New Paltz, NY, 1995 Oct.).

[16] V. Välimäki and T. Tolonen, "Multirate Extensions for Model-Based Synthesis of Plucked String Instruments," in *Proc. 1997 Int. Computer Music Conf.* (Thessaloniki, Greece, 1997 Sept.), pp. 244–247.

[17] T. Tolonen, "Model-Based Analysis and Resynthesis of Acoustic Guitar Tones," Master's thesis, Rep. 46, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland (1998 Jan.). A Postscript version available at <URL: http://www.acoustics.hut.fi/publications/>.

[18] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the Unit Delay—Tools for Fractional Delay Filter Design," *IEEE Signal Process. Mag.*, vol. 13, pp. 30–60 (1996 Jan.). MATLAB programs available at <URL: http://www.acoustics.hut.fi/>.

[19] M. Karjalainen and J. O. Smith, "Body Modeling Techniques for String Instrument Synthesis," in *Proc. 1996 Int. Computer Music Conf.* (Hong Kong, 1996 Aug.), pp. 232–239.

[20] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1975).

[21] S. J. Orfanidis, *Introduction to Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1996).

[22] T. Tolonen, V. Välimäki, and M. Karjalainen, "A New Sound Synthesis Structure for Modeling the Coupling of Guitar Strings," in *Proc. 1998 IEEE Nordic Signal Processing Symp.* (Vigsø, Denmark, 1998 June).

[23] D. Rocchesso and F. Scalcon, "Accurate Dispersion Simulation for Piano Strings," in *Proc. 1996 Nordic Acoustical Meeting* (Helsinki, Finland, 1996 June), pp. 407–414.

[24] J. R. Pierce and S. A. Van Duyne, "A Passive Nonlinear Digital Filter Design which Facilitates Physics-Based Sound Synthesis of Highly Nonlinear Musical Instruments," *J. Acoust. Soc. Am.*, vol. 101, pp. 1120–1126 (1997 Feb.).

[25] J. A. Moorer, "Signal Processing Aspects of Computer Music: A Survey," *Proc. IEEE*, vol. 65, pp. 1108–1137 (1977 Aug.). A revised and updated version appears in J. Strawn (ed.), *Digital Audio Signal Processing: An Anthology* (Kaufmann, Los Altos, CA, 1985), pp. 149–220.

[26] Y. Ding and X. Qian, "Processing of Musical Tones Using a Combined Quadratic Polynomial-Phase Sinusoid and Residual (QUASAR) Signal Model," *J. Audio Eng. Soc.*, vol. 45, pp. 571–584 (1997 July/Aug.).

[27] M. R. Schroeder, "A New Method of Measuring Reverberation Time," *J. Acoust. Soc. Am.*, vol. 37, pp. 409–412 (1965).

[28] J. M. Jot, "An Analysis/Synthesis Approach to Real-Time Artificial Reverberation," in *Proc. 1992 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 2 (San Francisco, CA, 1992), pp. 221–224.

[29] J. Laroche and J. L. Meillier, "Multichannel Excitation/Filter Modeling of Percussive Sounds with Application to the Piano," *IEEE Trans. Speech Audio Process.*, vol. 2, pp. 329–344 (1994 Apr.).

[30] X. Serra and J. O. Smith, "Spectral Modeling Synthesis: A Sound Analysis/Synthesis System Based on a Deterministic Plus Stochastic Decomposition," *Computer Music J.*, vol. 14, no. 4, pp. 12–24 (1990). See also the SMS Web site at <URL: http://www.iua.upf.es/~sms/>.

[31] O. Christensen and B. B. Vistisen, "Simple Model for Low-Frequency Guitar Function," *J. Acoust. Soc. Am.*, vol. 68, pp. 758–766 (1980).

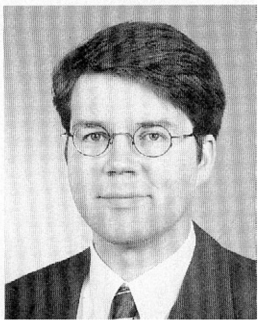
[32] R. J. McAulay and T. F. Quatieri, "Speech Analysis/Synthesis Based on a Sinusoidal Representation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, pp. 744–754 (1986).

[33] X. Serra, "A System for Sound Analysis/Transformation/Synthesis Based on a Deterministic plus Stochastic Decomposition," Ph.D. thesis, Rep. STAN-M-58, CCRMA, Department of Music, Stanford University, Stanford, CA (1989 Oct.).

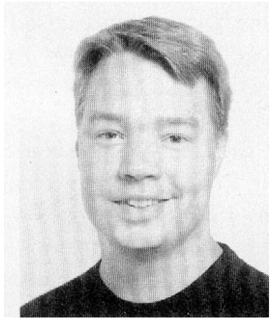
[34] T. Tolonen and V. Välimäki, "Automated Parameter Extraction for Plucked String Synthesis," *Proc. Inst. Acoust.*, vol. 19, pt. 5, book 1, pp. 245–250 (1997 Aug.).

[35] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments* (Springer, New York, 1991), pp. 50–53.

THE AUTHORS



V. Välimäki



T. Tolonen

Vesa Välimäki was born in Kuorevesi, Finland, in 1968. He studied acoustics and audio signal processing at Helsinki University of Technology (HUT), Espoo, Finland, and received the master of science, the licentiate in technology, and the doctor of technology degrees in electrical engineering in 1992, 1994, and 1995, respectively. His doctoral thesis dealt with sound synthesis using physical models of wind instruments. He worked at the HUT Laboratory of Acoustics and Audio Signal Processing from 1990 until 1995. In 1996 he was a postdoctoral research fellow in the Signal Processing Group of the School of Electronic and Manufacturing Systems Engineering (now the Department of Electronic Systems), University of Westminster, London, UK.

Dr. Välimäki then returned to Helsinki University of Technology, where he is currently a senior assistant in acoustics and audio signal processing. His research interests include music synthesis, active control of sound, audio coding, and digital filter design. He is author or coauthor of over 50 technical papers published in international journals and conferences. Many of his publications deal with physics-based sound synthesis. Dr. Välimäki is a member of the AES, the IEEE, the International Computer Music Association, and the Acoustical Society of Finland. His email address is: Vesa.Valimaki@hut.fi

•

Tero Tolonen was born in Oulu, Finland, in 1972. He majored in acoustics and audio signal processing and received the master of science degree in electrical engineering at Helsinki University of Technology (HUT) in January 1998. He has been working at the HUT Laboratory of Acoustics and Audio Signal Processing since 1996. He is currently a postgraduate student at the Graduate School in Electronics, Telecommunications and Automation (GETA) and he conducts research at the HUT Laboratory of Acoustics and Audio Signal Processing. His research interests include model-based audio representation and coding, physical modeling of musical instruments, and digital audio signal processing. Mr. Tolonen is a member of the AES, the IEEE, and the International Computer Music Association. His email address is: Tero.Tolonen@hut.fi