A SEMI-VIRTUAL VIOLIN FOR INVESTIGATIONS INTO SOUND QUALITY AND MUSICIAN-INSTRUMENT INTERACTION

Friedrich v. Türckheim, Thorsten Smit, and Robert Mores
University of Applied Sciences Hamburg
Department of Media Technology, Hamburg, Germany
friedrich.tuerckheim@haw-hamburg.de

ABSTRACT

This paper presents a semi-virtual violin research platform which consists of a silent generator violin and a modifiable virtual resonance body. The real-time platform is developed with particular emphasis on highly realistic sound behaviour, haptic experience, and visual properties. Thus, the platform allows for experiments on perceived violin sound quality and musician to instrument interaction together with professional violinists. Binaural transfer functions of real violins measured at the player’s hearing position serve as starting point for spectral manipulations. The virtual body can be modified in terms of the positions and amplitudes of individual resonances and resonance areas. A filtering technique is described which enables changes of the magnitude spectrum by using a system of second-order peak filters. In addition, other components along the signal chain are described, such as a specific impulse response measurement technique, a generator violin, and an equalization process for achieving authentic and traceable sound properties.

1. INTRODUCTION

The sound properties of a violin are primarily characterized by the nature of the resonance body. The latter is a complex system of resonances and formants, the position and intensity of which vary according to the quality of the instrument. The body can broadly be described as a linear time-invariant system and, thus, it can be characterized by its frequency response which is also known as the violin resonance profile. The relationship between a resonance profile and the quality of a violin has been investigated repeatedly, e.g. in [1], [2]. Scientists and violin makers are also able to infer physical properties from an instrument’s resonance profile [3], [4]. The sound quality of musical instruments, however, is not only the result of their physical properties, but also the result of the musician’s intensive and nuanced work. Which sound characteristics facilitate technical and interpretative subtleties? To what extent are musicians able to compensate for weak properties of an instrument?

In theory, synthetic musical instruments permit investigations on perceived sound quality and musician to instrument interaction. Simulating a variety of sound properties, the instruments allow inferring physical properties or, at least, preferred sound properties from experiments with musicians or listeners. Different sophisticated synthesis techniques for string instruments have been available for several years; each of them has its own advantages. For example, physical modelling techniques focus on the sound production mechanism [5], [6], [7], whereas body modelling techniques are aimed at less expensive computational effort at acceptable accuracy [8], [9]. Besides, various efficient modelling techniques are known, which focus on the strings, e.g. such as Digital Waveguide Filters [10], [11], the Functional Transformation Method [12], or combining methods [13]. Other research studies involve performance-based instruments or interfaces, e.g. in [14], [15], and [16].

In summary, one can state that a great deal of work in the field of synthesized string instruments has been done in recent years.

This paper aims at extending previous work with a violin model which is strictly developed for the use in experiments with professional violinists. In case of musician-based studies, some main requirements are mandatory in order to draw reliable conclusions: realistic sound, real-time capability, and parameterizability. Furthermore, the interface, i.e. the control features of the model, has to be as familiar as possible in order to allow the player to act in a realistic way. To meet all these requirements at the same time, the sound generation process of the violin has to be divided into two aspects: (i) the excitation signal (source) which is the natural excited string signal and (ii) the resonator (filter) which is the virtual body\(^1\). This approach is made in many previous attempts, e.g. in [17] or in [18].

In the following sections, the method and the main components of the research platform are described. Sections 2.1 - 2.3 describe the silent generator violin, the method for achieving binaural violin impulse responses and a filtering technique for modifying the resonance profile. Section 2.4 explores the equalization of the generator violin signal and the compensation of the headphone/dummy head trans-

---

\(^1\)The term ‘semi-virtual’ is used in the present work to underline the fact that only the body of the instrument is realized virtually.
fer functions. Finally, in Section 3, a real-time implementation using MATLAB and an external signal processor will be briefly described.

2. METHOD AND MAIN COMPONENTS OF THE INSTRUMENT

Figure 1 shows the block diagram of the violin model, including the musician, a specifically designed silent generator violin which is the interface of the model, and the virtual resonance body. The dashed box contains the computer-based elements. The important feedback loop between musician and violin model consists of two components: the vibrotactile (haptic) feedback and the acoustic feedback which is realized by means of closed headphones. At this point in time, the room block in the signal flow graph represents an optional standard reverberation algorithm realized with typical room impulse responses.

2.1. Generator Violin

The source signal $x(n)$, i.e. the string oscillation, is generated on a silent violin. Taking into account the relevance of an unfamiliar vibrotactile feedback of commercially available (bodyless) silent violins [19], a generator violin has been designed to achieve the mechanical impedance and, at the same time, the lowest acoustic radiation.

In the present case, a conventional violin (in high quality workmanship) is specifically filled with a combination of light polyurethane foam and silicone. Thus, the instrument features the outer appearance and the natural haptic properties of traditional violins (Figure 2) but at the same time incorporates controlled mechanical impedance. The bridge is loaded with piezo-electric force sensors which provide high-impedance signals. These signals are routed separately to an active impedance converter within the body (Figure 3). The circuit board is externally powered by means of phantom power in order to avoid a weight increase of the instrument. Due to the remaining modes of the damped body, the string signal has to be equalized to obtain a body-independent input signal. The equalization procedure is described in Section 2.4.

2.2. Impulse Response Measurement

Due to the complex radiation characteristics of a violin and the asymmetric playing position, the sound at the musician’s left ear differs from that at the right ear. For this reason, binaural impulse responses are measured at the player’s hearing position by means of a dummy head microphone which is suitable for near-field recordings. Figure 4 shows the arrangement of the dummy head and a violin in an anechoic chamber.

Usually, the transfer functions of violins are measured by exciting the instruments at the side of the bridge, e.g. with an impulse or an MLS Signal [8], [18], [20]. These techniques do not take into account the bridge motion in longitudinal direction and the torsional vibrations which occur due to the string deflection. These bridge motions do not decisively influence body radiation; however, they result in an acoustic radiation of the bridge itself. This again affects

\footnote{Parallel to the neck, normal to the bridge plane.}
A new body impulse response measurement technique is used which stimulates the violin bridge not only in the bridge plane but also in longitudinal direction, similar to the natural excitation, i.e. the bowed or plucked string oscillation. Therefore, the e-string is automatically pulled aside with a thin copper wire at the bowing position until the wire breaks. A similar method is used in [21] in order to pluck guitars in a natural way. The impulse is highly reproducible, since the level of stress which lets the wire break is always the same. All the strings are damped, so that the fundamental frequency of the 'plucked' string takes place beyond 10 kHz and therefore is filtered out by a subsequent low-pass filter (Section 2.4). Exactly the same excitation procedure is done on a specific steel quadrochord, i.e. a non-vibratory string instrument with same dimensions as a real violin (Figure 5). This quadrochord is fitted with piezoelectric sensors in the same way as the generator violin. The input impulse signal \( x_\delta(n) \), obtained on the quadrochord, is required for calculating the binaural frequency responses

\[
V_{l,r}(e^{i\Omega}) = \frac{\text{DFT} \{y_{l,r}(n)\}}{\text{DFT} \{x_\delta(n)\}} \frac{Y_{l,r}(e^{i\Omega})}{X_\delta(e^{i\Omega})},
\]

with \( \Omega = \omega_T = 2\pi f_s / f_s \), where \( f_s \) is the sampling frequency. \( y_l(n) \) and \( y_r(n) \) are the left and right dummy head signals at the violinist’s hearing position.

As an example, Figure 6 shows both the left and right ear magnitude spectra of a mid-priced violin.

2.3. Virtual Resonance Body

Violin impulse responses typically have a duration of about 100–150 ms. It is evident that the most natural sound of the semi-virtual instrument is obtained by convolving the source signal with the pure body impulse response. A computationally less expensive way is to filter the source signal with a model of the body, e.g. an all-pole model\(^3\). Taking into account the natural reverberation characteristics of

\(^3\)Theoretically, any other body modelling technique is also applicable here.
a violin resonance body, this is not only a matter of computation power but also of natural sound characteristics [9]. Here, both methods—convolving the source signal with the pure impulse response as well as filtering it with an all-pole model—are implemented. The audible differences are slight; however, in upcoming listening and playing tests, the methods will be compared to each other in order to find out the relevance of natural reverberation for perceived sound quality. In both cases, the subsequent filtering process which modifies the body characteristics is the same. So, in the following, the all-pole method is described only.

In the present work, the body is synthesized using an LPC (Linear Predictive Coding) model of order 1024. Such a high order allows a sufficiently accurate modelling of the important low-frequency resonances, i.e. the Helmholtz resonance (at about 270 Hz) and the main body resonances (at about 450–650 Hz). Figure 7 shows the magnitude spectra of a body impulse response and the corresponding LPC model.

The modifications of the resonance profile can be made directly within the dB-scaled magnitude spectrum plot by defining new frequency points (see also Section 3). Afterwards, these frequency points are linearly interpolated and subtracted from the original resonance profile. The resulting difference spectrum is the aimed magnitude spectrum of the filter system which is described below.

A finite-duration impulse response (FIR) filter design, e.g. based on frequency sampling, is not suitable here due to the high order which is needed to capture the important low-frequency details. The symmetric impulse response of a high-order FIR filter would result in a perceptible latency and in ringing artefacts which are both non-acceptable in the present case. For this reason, a system of parametric second-order peak filters is used instead. In the following, the computation of the filters is described. Subsequently, an additional process which optimizes the bandwidth and gain parameters of the peak filter system is described briefly.

The peak filters used here are based on the direct all-pass decomposition which is described in [22]. Accordingly, a peak filter can be expressed as a second-order band-pass filter with an additional direct path. After applying the bilinear z-transformation and the all-pass decomposition, the transfer function of the filter can be put in the form

\[ H(z) = 1 + \frac{H_0}{2} \left( 1 - A(z) \right), \]

where \( H(z) = \frac{1}{2} \left( 1 - A(z) \right) \) is the all-pass decomposition of a band-pass with

\[ A(z) = \frac{-a_{B,C} + d(1 - a_{B,C}) \cdot z^{-1} + z^{-2}}{1 + d(1 - a_{B,C}) \cdot z^{-1} + a_{B,C} \cdot z^{-2}}. \]

The bandwidth \( B \) in Hz is included in the parameter

\[ a_B = \frac{\tan(\pi BT) - 1}{\tan(\pi BT) + 1} \]

for the boost case (\( T \) is the sampling period). For the cut case, the parameter is given by

\[ a_C = \frac{\tan(\pi BT) - V_0}{\tan(\pi BT) + V_0}, \]

with \( V_0 = |H(e^{j\Omega_c})| = 10^{G/20}; \Omega_c = 2\pi f_c \cdot T, G \) is the gain value in dB. The center frequency \( f_c \) is included in the parameter \( d = \cos(\Omega_c) \). The coefficient \( H_0 \) is given by \( H_0 = V_0 - 1 \). A system of \( N \) peak filters in parallel structure can then be written as a sum of individual second-order filters combined with a direct path:

\[ H(z) = 1 + \sum_{i=1}^{N} \left( 1 + a_{[B,C]} \cdot \frac{H_0}{2} \right) \frac{H_0}{2} \left( 1 - a_{[B,C]} \cdot z^{-1} + z^{-2} \right). \]

The center frequencies and gain values of the individual filters are defined by applying a simple peak detection algorithm which identifies the positions and amplitudes of the positive and negative peaks in the difference magnitude spectrum\(^4\). The bandwidth parameters \( B_i \) are defined by detecting the two frequency points around the center frequencies \( f_c \) where the magnitude spectrum takes a value of \( 0.8 \cdot G_i[\text{dB}] \). Experience has shown that this criterion is best suitable for estimating the \( Q \)-factors of peak filters in parallel structure. However, due to band interaction, the magnitude frequency response of the filter system does not satisfactorily match the desired filter curve. The result can be significantly improved by applying an additional optimization process. The optimization process finds a minimum of

\(^4\)When using the pure impulse response instead of an LPC-model, the difference spectrum is obtained from the dB-scaled magnitude FFT spectrum of the impulse response. Therefore, the magnitude frequency curve is smoothed in order to decrease the number of detected peaks.
the squared magnitude response error by tuning the gain values and the bandwidth parameters. With a vector
\[ \mathbf{p} = [B_1, ..., B_N, G_1, ..., G_N], \]
containing the parameters \( B_i \) and \( G_i \) of each filter, the optimization problem can be expressed as
\[ \min_{\mathbf{p}} \left\| \mathbf{w}(\Omega) \cdot \left( \frac{H(e^{j\Omega}, \mathbf{p})}{H_{\text{diff}}(e^{j\Omega})} \right)^2 \right\|_2, \]
where \( H(e^{j\Omega}) \) is the frequency response of the filter system and \( H_{\text{diff}}(e^{j\Omega}) \) is the difference spectrum. The optional vector \( \mathbf{w}(\Omega) \) is a weighting vector. Here, bark-dependent weighting factors are used in order to consider the human auditory frequency resolution. The least-square minimization problem in (8) is solved in MATLAB by means of the \textit{lsqnonlin} command of the optimization toolbox (Levenberg-Marquardt algorithm).

Figure 8 shows one possible result of the filtering process. In this case, the main wood resonance (at about 480 Hz) is damped and slightly shifted. The corresponding system of optimized peak filters is shown in Figure 9. Another example is shown in Figure 10 where the so-called ‘bridge hill’ at about 2.5 kHz is shifted and boosted.\(^5\)

2.4. Input Signal Equalization

The remaining modes of the damped generator violin body are still enclosed in the piezo input signal. These have to be removed in order to obtain a body-independent source signal; otherwise, the sound output of the tool would still include the properties of the damped violin body, effectively resulting in two audible violins. This equalization is done by an inverse filtering process. Therefore, the bridge admittance, i.e. the frequency dependent bridge mobility, of the silent violin is measured by exciting the silent violin as described in Section 2.2. The bridge motion is recorded by an accelerometer (having a weight of 0.3 g) mounted at the side of the bridge. The input impulse signal \( x_g(n) \) is measured on the vibration-free quadrochord as described in Section 2.2. The admittance function is the quotient of the bridge motion signal and the input impulse signal in the frequency domain:
\[ G(e^{j\Omega}) = \frac{Y_g(e^{j\Omega})}{X_g(e^{j\Omega})}. \]

Figure 11 shows the bridge admittance curve of the generator violin. For comparison, the bridge admittance curve of a conventional violin is represented, too.

\(^5\)For the sake of completeness, it should be mentioned that, in addition to the peak filters, another filtering method is implemented in the violin platform: Band-shelving filters with arbitrary edge frequencies allow for optional amplification or attenuation of certain frequency bands of the body impulse responses. As this is a common audio equalizer function, details are omitted here. The shelving filters used here are described in [23].

Due to the damped body, the bridge motion impulse response of the generator violin has a very short duration of about 1.5 ms. Thus, the inverse filtering can be done by means of a low-order FIR filter without taking account of an interfering reverberation part. The inverted magnitude spectrum is combined with the filter of the virtual resonance body which is described in Section 2.3. Due to the natural band-pass behaviour of the violin body, the inverse frequency response results in an increased noise floor below 100 Hz and above 8 kHz. To avoid this, the string signal is subsequently filtered with a band-pass filter \( H_{\text{BP}}(e^{j\Omega}) \) with an upper edge frequency of 8 kHz. This filter also removes the undesirable string harmonics which arise due to the impulse measurement technique (see Section 2.2).

Figure 12a shows the elements which influence the frequency response. In addition to the bridge admittances of the generator instrument and the filterings within the model, the signal meets the frequency response of the headphones \( A(e^{j\Omega}) \) and the frequency response of the violin player’s pinna and auditory canal \( E_{\text{PI}}(e^{j\Omega}) \). Since the binaural impulse responses already include the transfer function of the dummy head, it has to be compensated in order to cancel the duplicate frequency characteristics of the ear canal.\(^6\) The frequency response of the binaural body impulse response

\(^6\)It is assumed that the characteristics of the artificial head’s pinna and auditory canal are identical to those of the violinist.
can be written as

\[
V(e^{j\Omega}) = V'(e^{j\Omega}) \cdot H_{DH}(e^{j\Omega}) \cdot E_{DH}(e^{j\Omega}),
\]

where \(V'(e^{j\Omega})\) is the frequency response of the resonance body and \(H_{DH}(e^{j\Omega})\) is the direction-dependent part of the dummy head transfer function (i.e. torso and nose reflections as well as interaural time differences). \(E_{DH}(e^{j\Omega})\) is the pinna/auditory canal transfer function of the dummy head. The latter is measured together with the headphone transfer function by recording logarithmic sweeps which are reproduced by the headphones with the dummy head microphone. Several measurements with different headphone positions are done to obtain an averaged transfer function. Figure 12b shows the implemented equalization elements of the violin.

The complete filtering process along the signal chain—which begins with the string oscillations and ends with the perception of the violinist—can be written as:

\[
H_{all}(e^{j\Omega}) = \frac{V(e^{j\Omega}) \cdot H_{peak}(e^{j\Omega}) \cdot H_{BP}(e^{j\Omega})}{G(e^{j\Omega}) \cdot E_{DH}(e^{j\Omega}) \cdot A(e^{j\Omega})}
\]

(11)

With (1) and (9) the filtering process can be simplified by using the dummy head microphone signals and the bridge mobility signal of the generator violin:

\[
H_{all}(e^{j\Omega}) = \frac{Y_{lr}(e^{j\Omega}) \cdot H_{peak}(e^{j\Omega}) \cdot H_{BP}(e^{j\Omega})}{Y_{B}(e^{j\Omega}) \cdot E_{DH}(e^{j\Omega}) \cdot A(e^{j\Omega})}
\]

(12)

Except for the binaural impulse response and the peak filters, the frequency responses can be combined to a single filter with the magnitude frequency response

\[
|H(e^{j\Omega})| = \left| \frac{H_{BP}(e^{j\Omega})}{Y_{B}(e^{j\Omega}) \cdot E_{DH}(e^{j\Omega}) \cdot A(e^{j\Omega})} \right|
\]

(13)

Here, the frequency response is realized by means of an FIR filter of order 256 by using the frequency sampling method.

3. IMPLEMENTATION

In the present case, the filtering process is outsourced to a Texas Instruments TMS320C6416 DSP platform in order to achieve real-time sound processing. Due to that, a processing latency of less than 5 ms of the overall system is reached. The computation of the band-shelving and peak filter coefficients as well as the parameter optimization proceed in MATLAB. The communication with the external processor takes place via real-time data exchange channels (RTDX). As mentioned above, there is a simple convolution reverb effect implemented using typical room impulse responses.

The desired resonance profile is defined within the magnitude spectrum plot of a graphical user interface using the MATLAB command `ginput`. The equalizer function is implemented by drag and drop of ‘gain bars’ within the same plot. Modifications can be done either by changing each of the binaural impulse response spectra separately or by changing an interpolated resonance profile. In the latter case, the generated single filter is applied to both of the binaural spectra. Other adjustments can also be set in the software interface such as filter order, sampling frequency or the selection of different room impulse responses used for the reverberation process. Sliders control the ‘wet/dry’ ratio of the reverb effect as well as the main volume level.
The ‘reference violins’ can be chosen by means of a pop-up menu (Figure 13).

4. SUMMARY AND FUTURE WORK

In this paper, the main components of a semi-virtual violin research platform have been presented. The instrument is specifically designed in terms of control and sound quality in order to improve the quality of experimental results when used in sessions with musicians. The sound of the instrument is very close to that of real violins. Besides, due to the real-time implementation on an external DSP platform, the latency of the instrument is sufficiently low for the use with professional violinists.

In upcoming playing experiments, the presented platform will be used to investigate the conscious and unconscious responses of musicians to specific violin sound properties. One item is, for example, the relevance of single prominent and narrow resonances which sometimes occur in resonance profiles of real violins. Additionally, in concurrent research work, the authors investigate the relationship between perceptible vowel quality in violin sounds and the quality of the instruments. Informal tests have shown, for example, that violinists prefer ‘e’-vowel sound characteristics to ‘i’-vowel sound characteristics on the upper e-string\(^7\).

So, the violin tool will be used to simulate different timbre characteristics which resemble vowel sounds to a greater or lesser extent. Future work will also include the extension of the impulse response library for a wider range of reference resonance profiles. Implementing a more realistic room simulation is another challenging task to give greater consideration to the true radiation characteristics of violins.

\(^7\)The German vowel pronunciation is meant here. The German ‘i’ is equal to the English ‘e’, e.g. as in ‘feel’, the German ‘e’ is pronounced similar like the vowel in ‘say’.
5. ACKNOWLEDGEMENTS

The authors thank the German Federal Ministry of Education and Research for funding.

6. REFERENCES


