VSOUND: REAL-TIME DIGITAL EMULATION OF THE ACOUSTIC VIOLIN

P Gaydecki School of EEE, University of Manchester, Manchester, UK
H Johannsson Johannsson Violins, Ingólfsstræti, 10, Reykjavik 101, Iceland
S Ismail School of EEE, University of Manchester, Manchester, UK
A Barton School of EEE, University of Manchester, Manchester, UK
T Lloyd School of EEE, University of Manchester, Manchester, UK

1 INTRODUCTION

A device named *vSound* is described, which has been developed to process the raw electrical output from an electric violin to produce an output signal which, when fed to an amplifier and loudspeaker, approximates closely the timbre of an acoustic instrument. The system comprises a high impedance preamplifier, a 24-bit sigma delta codec and a digital signal processor (DSP) operating at 590 million multiplications-accumulations per second (MMACs). The device holds in its memory up to sixteen far-field impulse responses of wooden instruments, any one of which may be convolved in real-time with the input signal to synthesize the modified output signal. This first stage convolution is realized as a finite impulse response (FIR) structure. In addition, the device incorporates a graphic arbitrary equalizers, a blender, bass-cut filter unit and a gain adjustment function. The user interface includes a set of navigator-style buttons and a simple display screen. The device may be connected to a computer or tablet and controlled from a separate software package. Initial tests using an impulse response obtained from a Guarneri del Gesù violin constructed in the early-mid 18th century suggest that the system significantly enhances the tonal quality and timbre of an electric instrument.

Real-time signal processing based on both general purpose microprocessors and fast digital signal processors (DSPs) is a technique that emerged over twenty years ago, and is now widely considered one of the fastest growing application areas in the field of digital technology. Typically, the analogue waveform is first digitized by an analogue to digital converter (ADC), and the binary values are transmitted to a DSP device that performs a real-time convolution operation in discrete space using either a finite impulse response (FIR) or infinite impulse response (IIR) algorithm. The processed data are then sent to a digital to analogue converter (DAC) that outputs a filtered analogue signal.

It is important to recognize however, the limitations of this approach to electric violin enhancement. The sound radiation pattern emanating from a wooden body is spatially dependent: for example, the frequency response from the side of the instrument is quite distinct from the same measurement taken from the front or rear sound plates. The system described here only manipulates the far-field response. Perhaps most important, the hardware does not provide any form of feedback to the player from the string to bow.

2 BACKGROUND

2.1 Basic DSP filter theory

The real-time filtering system employs a combination of both finite impulse response (FIR) and infinite impulse response (IIR) filters to affect the processing of the input force signal. Specifically, an FIR algorithm convolves the incoming signal with the body impulse response, and an IIR structure is used for both the parametric equalizer and the reverberation stage. In general, the convolution integral in continuous space is expressed as:

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau \tag{1}$$

where $y(t)$ is the output (filtered) signal, $x(t)$ is the incoming signal, τ is the time-shift operator and $h(\tau)$ is the impulse response of the filter. In discrete space, this equation may be implemented using either an FIR or IIR solution. In the former case, the infinite response is truncated, which yields an expression of the form:

$$y[n] = \sum_{k=0}^{M} h[k]x[n-k] \tag{2}$$

with the z-transform of the impulse response, i.e. the transfer function $H(z)$, being given by:

$$H(z) = \frac{Y(z)}{X(z)} = \sum_{n=0}^{\infty} h[n]z^{-n} \tag{3}$$

In contrast, IIR filters rely on recurrence formulae, where the output signal is given by:

$$y[n] = \sum_{k=0}^{N} a[k]x[n-k] - \sum_{k=1}^{M} b[k]y[n-k] \tag{4}$$

and the transfer function is given by:

$$H(z) = \frac{a[0] + a[1]z^{-1} + ... + a[m]z^{-m}}{1 + b[1]z^{-1} + ... + b[n]z^{-n}} = \frac{\sum_{m=0}^{M} a[m]z^{-m}}{1 + \sum_{n=1}^{N} b[n]z^{-n}} \tag{5}$$

There are important consequences associated with these two approaches to filtering; one of the most important criteria in assessing the performance of a filter is its stability. As Equations (2) and (3) show, FIR filters are unconditionally stable since there is no feedback in the convolution process. In contrast, IIR filters always feedback a fraction of the output signal, which necessitates careful attention to design if stability is to be ensured. This may be viewed another way: Equation (5) shows that the transfer function is the ratio of two polynomials in ascending negative powers of z. Thus high-order polynomials are associated with very small denominator terms and hence the risk of an ill-conditioned division. It is for this reason that IIR filters are sensitive to the word-length of the DSP device. In general, the higher the order of the filter, the greater the risk of instability, so high-order IIR filters are often designed by cascading together several low-order sections.

The body of a violin is, to a very close level of approximation, a linear system. Although its properties do alter over time (as the wood, varnish and adhesives age), these changes only become significant over years. Within any reasonable time frame, the body manifests the property of temporal invariance. Hence if the body is tapped at a particular position using an instrumented hammer, the sound that it produces will always retain the same spectral characteristics (although the amplitude may vary, depending on the strength of the strike). Non-linearities are often introduced during performance, but these are player-dependent and confined to manner in which the bow moves over the string, not the body. Although the near-field impulse response is spatially variable, in the far-field these variations are less significant; it the far-field which is of interest in this case. In this application, the impulse response of the body is realized as an FIR filter, i.e. Equation (2). The same approach is adopted for the equalizers. In contrast, the bass-cut filter, which requires selectivity but low computational overhead, employs and IIR architecture.

2.2 The violin as a linear system

The representation of a violin body as a linear system is pivotal since it allows the vibration behaviour to be characterized by the standard methods of linear acoustics, as has been explored by many researchers [1, 2, 3]. The most extensive attempts to date to use the linear model to explore the resonance characteristics of real stringed instruments have been conducted by Jansson and

Dünnwald [4, 5, 6]. Dünnwald measured resonance characteristics of about 700 violins and found that certain resonance features (such as the presence of strong resonances in the region 190-650 Hz and 1300-4200 Hz relative to resonances in the regions 650-1300 Hz and 4200-6500 Hz) were characteristic of 'Old Italian' instruments; it was concluded that these features were significant in determining the acceptability of violin sounds, but no empirical research was conducted to explore or support the perceptual validity of this claim.

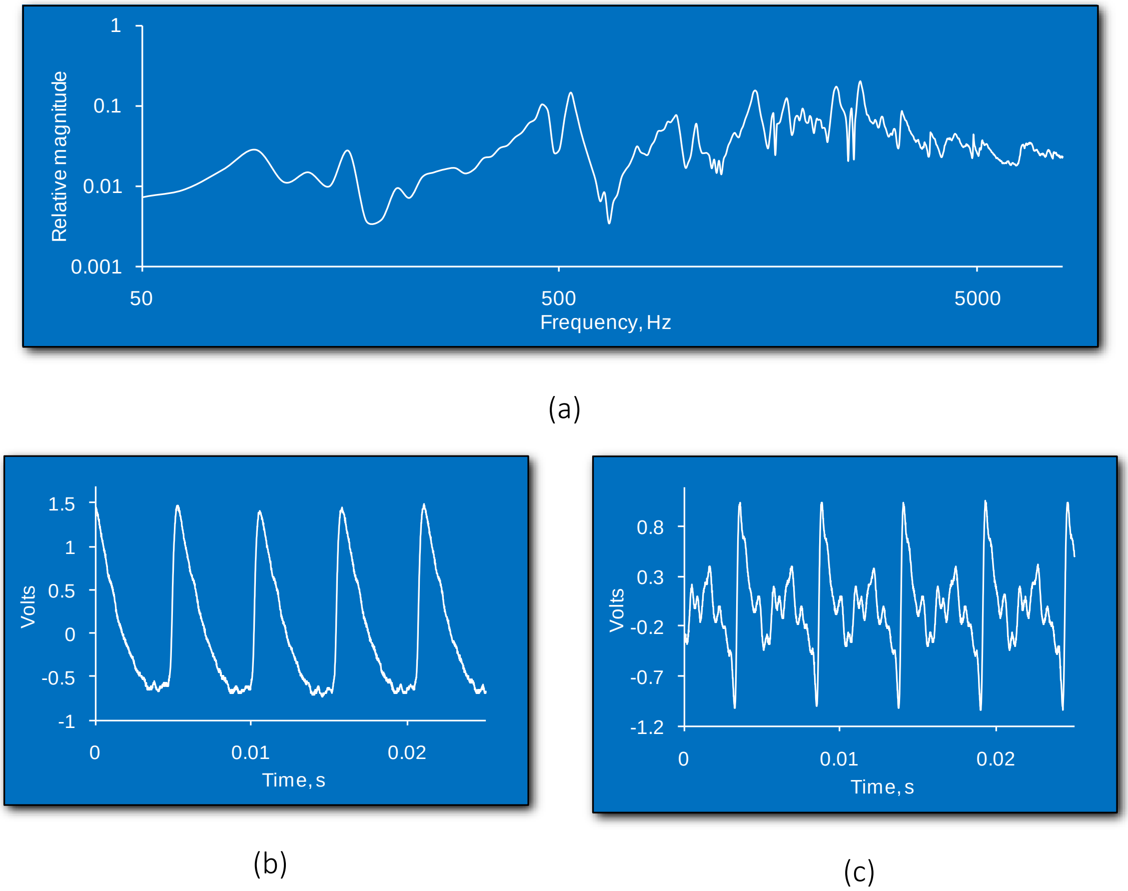


Figure 1. (a) Typical violin body frequency response; (b) force signal on bridge of violin; (c) force signal on bridge of violin mediated by body frequency response.

When performing on stage, violinists often use a microphone to amplify the sound from the violin; this has its disadvantages. The voice of the instrument is often not reproduced with fidelity since the microphone is in the near-field radiation zone of the transfer function. Additionally, interference is often introduced. From a pragmatic perspective, acoustic instruments are often very expensive – a typical Stradivarius will cost upwards of $5M.

An approximate violin bridge/body impulse response can be obtained by tapping the bridge with a small instrumented hammer and using a microphone and an ADC to record the signal. This impulse response is then processed with an FFT algorithm to obtain the frequency response. A typical violin body frequency response is shown in Figure 1a. The force signal from just a string may be obtained using miniature accelerometers mounted on the bridge; a typical signal is shown in Figure 1b. This is then processed digitally with the bridge/body filter to synthesize a convincing violin sound; the resulting waveform [7] appears in Figure 1c. An accurate measurement of the impulse response is technically challenging, and many publications have discussed this subject [8, 9].

3 SYSTEM DESCRIPTION

3.1 Hardware architecture

The hardware of vSound has been designed to perform the operations described above in real time. The output from the string pickup (normally a piezo-electric transducer) is fed to a high-impedance (1 MΩ) gain-switchable preamplifier and then to the analogue-to-digital converter section of a 24-bit codec sampling at 44.8 kHz. The output from the codec is then fed to the DSP device that performs all of the main-stage processing, i.e. body impulse response convolution, equalization, blending (dry/wet mixing), optional bass-cut filtering and final volume control. The processing chain is depicted in Figure 2.

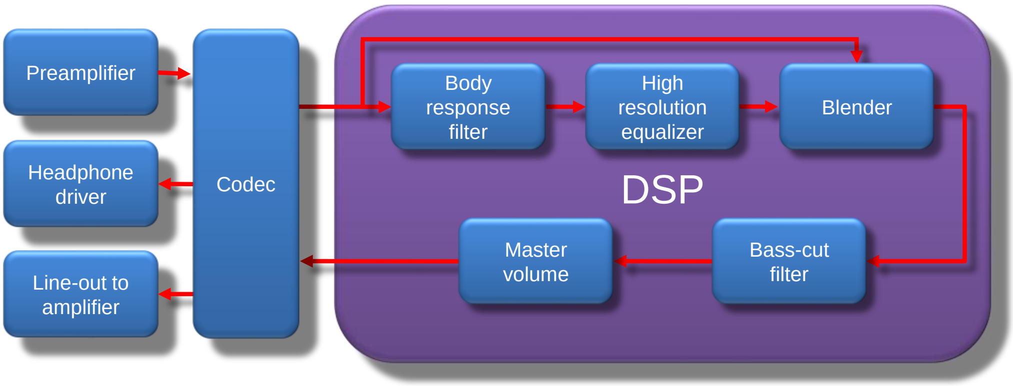


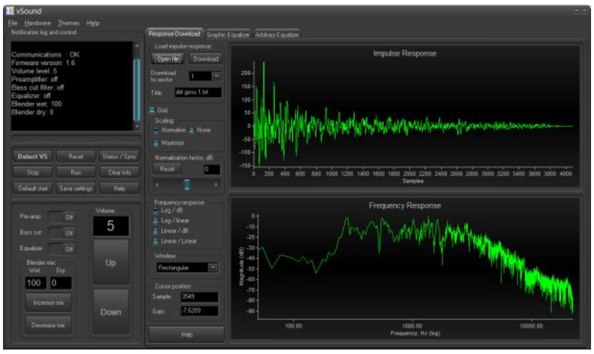
Figure 2. Main system components and signal processing chain.

After processing, the output signal is fed back to the digital-to-analogue section of the codec and from there to appropriate buffers to drive both high power audio amplifiers and headphones. Other sub-systems include a display, keypad and interface to allow the unit to be connected to and controlled from a computer. The final device is shown in Figure 3.

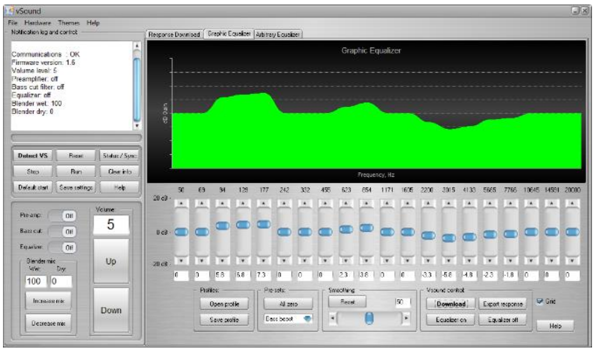
Figure 3. The vSound unit.

3.2 Real-time firmware

As discussed above, convolution of the incoming signal with the impulse response is realized as an FIR structure. This is also the case for the equalizer (both graphic and arbitrary). The spectral resolution of the arbitrary filter is 20 Hz across the entire system bandwidth, allowing the user to introduce very subtle alterations in the tonal quality of the final audio response. Following the equalizer stage, the blender allows the player to mix the original (dry) signal with the processed (wet) signal, in any proportion, using the navigator keys on the top panel of the unit or, alternatively, from the computer using the software provided. Blending is easily implemented in software, although it is essential to incorporate a delay function in the dry signal path to compensate for the delay resulting from the various processing stages. The function of the bass cut filter, again selectable from the unit or the software, is primarily intended to minimize mains power hum (50/60 Hz), which is a problem often associated with high-impedance pickups.



(a)



(b)

(c)

Figure 4. The user software: (a) impulse response download tab; (b) graphic equalizer; (c) arbitrary equalizer.

3.3 User software

The user software allows the user to download impulse responses into the flash memory of the device, adjust the parametric and graphic equalizer and the blender ratio. Further it implements control functions such as volume setting, enabling or disabling the pre-amplifier, equalizer and bass-cut filter.

The device connects to the computer via a USB interface and is automatically for configured for communication via the software (no user interaction is required). A sync facility synchronizes the software to the operational state of the device. Typical screenshots are provided in Figure 4.

4 PERFORMANCE

The acoustic performance of the device has been evaluated using a number of derived impulse responses, most notably that of a Guarneri del Gesù violin constructed in the early-mid 18th century. Accurate measurement and digital translation of the impulse response is a task that must be undertaken with considerable care, involving excitation of the bridge/body with an instrumented hammer, recording of the response in the far field using a calibrated microphone, and subsequent deconvolution of the hammer strike from the recorded signal to yield the final impulse response. The process is not described in detail here, but is discussed in a separate publication [10]. It is perhaps important to note that the response is spatially variant, at least in the far-field, and perceptually is different for the player and the listener. Although no formal listening tests have yet been conducted, even casual scrutiny confirms a pronounced enhancement to the timbre and tonal quality of an electric instrument processed by the vSound unit. In simple terms, the harsh, nasal resonance is replaced by an authentic acoustic or wooden voice.

5 CONCLUSION

The vSound device is intended to approximate, in real-time, the far-field impulse response of an acoustic violin. It comprises a high-speed DSP core and codec in combination with software that convolves the incoming signal with a measured far-field radiative response. Other features include a software selectable preamplifier, user-controlled equalizer, blender unit, bass-cut filter and volume control. Preliminary tests suggest that the device may enhance significantly the voice of an electric instrument; additionally, it represents an ideal research tool due to its flexibility and the ease with which the entire convolution operation may be reconfigured.

6 REFERENCES

1. L. Cremer, The physics of the violin: MIT Press, Cambridge, Mass. (1985)
2. K. D. Marshall, Modal analysis of a violin: J. Acoust. Soc. Amer. 77, 695–709 (1985)
3. C. M. Hutchins, Research papers in violin acoustics 1975–1993: Acoustical Society of America, Woodbury NY. (1997)
4. E. Jansson and B. Niewczyk, On the acoustics of the violin: bridge or body hill: CAS Journal Series 2(3) 23-27 (1999)
5. E. Jansson, Admittance measurements of 25 high quality violins: Acustica - Acta Acustica 83 337–341 (1997)
6. H. Dünnwald, Deduction of objective quality parameters on old and new violins: J. Catgut Acoust. Soc. Series 2(1) 1–5 (1991)
7. P. Gaydecki J. Woodhouse and I. Cross, Advances in audio signal enhancement and stringed instrument emulation using real-time DSP Proc. Forum Acusticum Sevilla 2002 (on CD; ISBN: 84-87985-06-8) (2002)
8. P. Cook and D.Trueman, A database of measured musical instrument body radiation impulse responses, and computer applications for exploring and utilizing the measured filter functions, Proc. Int. Symp. Musical Acoust. (Acoustical Society of America, Woodbury, NY), 303–308 (1998)
9. C. Fritz, I. Cross, B. Moore and J. Woodhouse, Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin, J. Acoust. Soc. Am. 122, 3640–3650 (2007)
10. T. Lloyd, P. Gaydecki and H. Johannsson, This deconvolution business: an insight into violin characterisation and emulation (2015) (in prep.)