19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

TOWARDS PHYSICS-BASED RE-SYNTHESIS OF WOODWIND TONES

PACS: 43.75.Pq

Van Walstijn, Maarten; De Sanctis, Giovanni
Sonic Arts Research Centre, School of Electronics, Electrical Engineering, and Computer Science, Queen's University Belfast.

ABSTRACT
In this paper, we propose a new approach to physics-based re-synthesis of tones produced with acoustic instruments, with particular application to woodwinds. The main concept is to sense the pressure at different positions along the main bore, separating the propagating waves, which allows to characterise the driver and the resonator separately. The focus of the paper is on the development of a suitable wave separation technique; results are given for computer-simulated sensor data.

INTRODUCTION

Physical modelling has become an increasingly popular technique for sound synthesis in recent decades; while traditional synthesis techniques aim to produce or mimic a particular waveform or spectrum, physical modelling aims to simulate the actual sound-producing mechanism. In addition to a much increased realism (including natural-sounding transients), the parameters of such models have a physical meaning, thus providing an intuitive understanding of the relationship between the control parameters and the resulting sound output. Up till now, the emphasis in this area has been on pure sound synthesis, i.e. translating a set of physical parameters into audio. However in order to fully exploit the potential of physical modelling with regard to music and audio applications, methods for performing the inverse process, i.e. estimating the physical parameters from player-generated instrument oscillations, are required. This way one captures information not only about the instrument, but also about how the player controls the instrument in order to produce a specific sound. This approach offers some important new possibilities. Firstly, the direct measurement of player-controlled parameters is often extremely difficult. For example, how to directly measure the lumped parameters of a reed-mouthpiece-lip model? Extraction of these parameters from the audio/oscillatory level via inverse modelling would provide a much-needed alternative parameter determination method. Secondly, the inverse-modelling approach opens up the possibility of physics-based re-synthesis, which would provide radical new ways of editing sound produced with acoustic instruments.

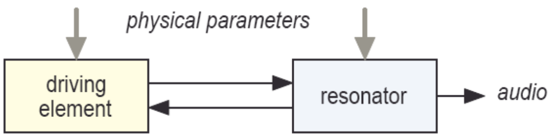

Figure 1. General block diagram for physics-based simulation of a musical instrument. The feedback between main resonating element (air column, string, drum membrane) and driving element (reed, bow, mallet/hammer) is continuous for a sustained sound, but occurs only during the short attack part of a percussive sound.

While physics-based synthesis has been explored extensively, the inverse process of physics-based analysis (i.e. parameter extraction) has received very little attention so far, and is still in its infancy as a research area. A few promising results have recently been obtained with inverse-modelling of percussive sounds, including realistic re-synthesis and transformations (see for example [1]). In comparison, parameter extraction from sustained musical sounds, such as produced with wind and bowed string instruments, is notoriously difficult, despite the relatively low modelling complexity required for realistic simulation of the main resonating element for these types of instruments. This is because, unlike with percussive sounds, the interaction between the resonator and the driver (e.g. the bow, the reed) is continuous (see Figure 1). In this scenario, and with only the output audio as available data, the resonator and the driver cannot be characterised independently. As a consequence, model inversion is only possible under the assumption of an instantaneous, time-independent mapping between oscillation states and control parameters (see for example [2,3]); this assumption is invalid for most wind instruments, and precludes precise re-synthesis of transients. Scavone and Cook [4] have attempted identification of the sound generation loop of a (simplified) clarinet, but employed a sensing method that does not provide adequate information for precise identification and parameter estimation.

We aim to establish a new approach to extraction of physical model parameters from the air vibrations inside wind instruments. The proposed methodology is to first identify the pressure waves travelling inside a wind instrument by (post)-processing the internal pressure signals sensed at different points in the instrument air column. This enables to characterise the driver and the resonator separately, avoiding the ambiguity problems inherent to concurrent estimation. The first step in this research is the development of suitable wave separation techniques, which is the main focus of the next section.

WAVE SEPARATION

Estimation of travelling waves from 2 microphones

Guérard en Bouttillion [8] developed a wave separation method using 5 microphones, based on using finite-difference approximations of spatial pressure and velocity derivatives. In our approach, we aim at develop an estimation method that avoids any intrinsic numerical approximation errors.

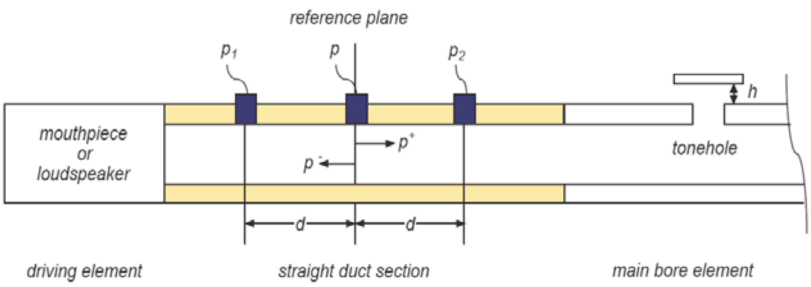

Figure 2: Schematic depiction of the wave separation device.

Consider the 3-sensor arrangement in Figure 2. Under the assumption of linearity, the following frequency-domain relationships hold:

$$P_1 = \cosh(\Gamma d)\, P + \sinh(\Gamma d)\, ZU \,, \qquad P_2 = \cosh(\Gamma d)\, P - \sinh(\Gamma d)\, ZU \qquad \text{(Eqs 1)}$$

where Z and Γ are the characteristic impedance and propagation constant in the cylindrical duct, respectively, and U is the volume velocity at the reference plane. Using these equations, the pressure and normalised volume velocity at the reference plane can be estimated from the sensor signals as:

$$\hat{P} = \frac{(P_1 + P_2)}{2\cosh(\Gamma d)} \,, \qquad Z\hat{U} = \frac{(P_1 - P_2)}{2\sinh(\Gamma d)} \qquad \text{(Eqs 2)}$$

The forward- and backward-travelling waves at the reference plane, as shown in Figure 2, can be then derived from the standard acoustic relationships:

$$P = P^+ + P^- \,, \qquad ZU = P^+ - P^- \qquad \text{(Eqs 3)}$$

A discrete-time formulation of this procedure can be can be derived as follows. Knowing that

$$2\cosh(\Gamma d) = e^{+\Gamma d} + e^{-\Gamma d}, \qquad 2\sinh(\Gamma d) = e^{+\Gamma d} - e^{-\Gamma d}, \qquad \text{(Eqs 4)}$$

equations (4) can be re-written as:

$$\hat{P} = \frac{H(\omega)}{1+H(\omega)}\left(P_1 + P_2\right), \qquad Z\hat{U} = \frac{H(\omega)}{1-H^2(\omega)}\left(P_1 - P_2\right) \qquad \text{(Eqs 5)}$$

Where $H(\omega) = e^{-\Gamma d}$ is a frequency-dependent expression that represents the propagation of a wave travelling over a distance d. In discrete-time, this expression can be modelled as an FIR filter, denoted with $H(z)$, which we will refer to as the "propagation filter". We may now write the *z*-transforms of the estimates as:

$$\hat{P}(z) = H(z)\left(P_1 + P_2\right) - H^2(z)\hat{P}(z) \qquad \text{(Eq 6)}$$

$$Z\hat{U}(z) = H(z)\left(P_1 - P_2\right) + H^2(z)Z\hat{U}(z) \qquad \text{(Eq 7)}$$

Now taking the inverse *z*-transforms gives the discrete-time formulations of the wave separator:

$$p(n) = h(n) * \left(p_1(n) + p_2(n)\right) - h(n) * \left(h(n) * p(n)\right) \qquad \text{(Eq 8)}$$

$$Z\hat{u}(n) = h(n) * \left(p_1(n) - p_2(n)\right) + h(n) * \left(h(n) * Z\hat{u}(n)\right) \qquad \text{(Eq 9)}$$

where the operator '$*$' denotes discrete-time convolution. The signal flow diagram of the estimation procedure, which we will refer to as 'delay-loop filtering' is depicted in Figure 3.

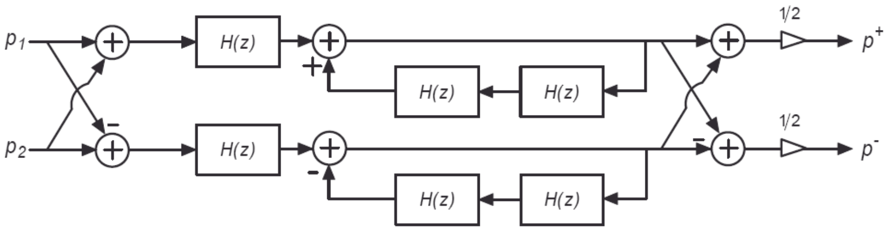

Figure 3. Signal flow diagram for wave separation.

We note that a delay-free loop is avoided only for $h(0) = 0$, i.e. when the distance *d* is at least $c \cdot T$ long, where $T = f_s^{-1}$ is the sample period.

Adaptive calibration

In principle, the expression $H(\omega) = e^{-\Gamma d}$ is well known from theory [7], and thus an appropriate FIR filter can be pre-calculated. However, Γ is dependent on the thermodynamic constants, such as air density and wave velocity, which in turn depend on the actual conditions during sensing, i.e. the temperature and humidity. With a player driving the signals, these conditions tend to vary over time, and are therefore the calibration would have to be able to 'follow' these changes.

In order to address this problem, the third microphone (positioned perfectly at the mid-point between the other 2 microphones) is used to provide a reference pressure, enabling the calibrate the system adaptively. Taking equation (6) and substituting the estimated pressure $\hat{P}(z)$ with the actual sensed pressure $P(z)$, one obtains:

$$P(z)(1 + H^2(z)) = H(z)\left(P_1 + P_2\right). \qquad \text{(Eq 10)}$$

Using equation (10) to define an error signal, the estimation of $H(z)$ can now be formulated as an adaptive estimation problem (see Figure 4). An optimum filter can be obtained by minimising the least-square error $\xi(n) = E\left[e^2(n)\right]$.

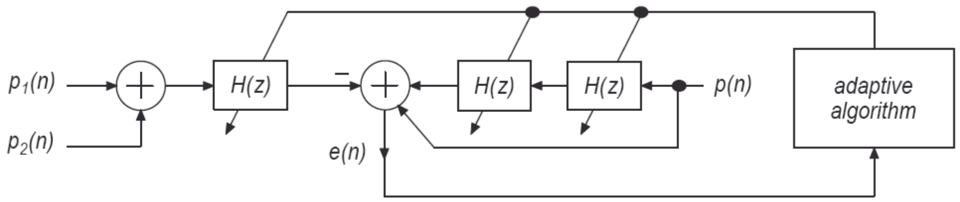

Figure 4. Adaptive estimation of the propagation filter. The adaptive algorithm updates the propagation filter coefficients on basis of the error-signal *e(n)*. Successful estimation requires *e(n)* to become small.

In first instance, our approach has been to first estimate the propagation filters as an FIR filter with filter coefficients h_j , using a steepest-descent LMS adaptive filtering approach, where the *j*th element of the gradient was derived to be:

$$\frac{\partial e(n)}{\partial h_j} = -p_1(n-j) - p_2(n-j) + 2\sum_{m=1}^{M} h_m \, p(n-j-m) \qquad \text{(Eq 11)}$$

Note that the gradient differs significantly from that of the standard LMS algorithm, due to the fact that we are dealing with a system configuration in which the sensed signals are passed once as well as twice through the filter that characterises the channel. Figure 5 shows a typical comparison with the target response using a 256-tap filter at a 100 kHz sample rate.

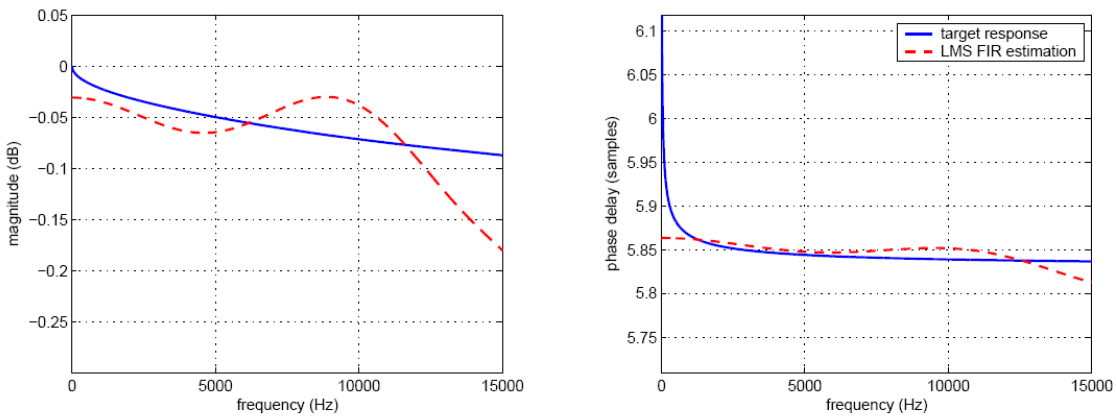

Figure 5. A comparison between the ideal (target) response and the LMS FIR estimation of the propagation filter.

The next step in the calibration procedure is to fit a theoretical model to the obtained FIR impulse response. Starting with the formulations of wall losses in ducts by Pierce [7], the propagation over a distance *d* can be formulated as a function of two parameters:

$$H(\omega) = e^{-j\omega\tau - g\tau\sqrt{j\omega}} \qquad \text{(Eq 12)}$$

where $\tau = d/c$ is the travelling time and $g = a^{-1}\sqrt{\eta/\rho}\left(1+(\gamma-1)/\upsilon\right)$ is a convenient parameter related to the attenuation, and depends on the viscosity η, the ratio of specific heats γ, and the is the square root of the Prandtl number υ.The fitting routine estimates τ and g using a standard least-square nonlinear fitting routine. Finally, these two parameters are used to design an appropriate FIR filter. In the design procedure, the target response $H(\omega)$ is first evaluated for a bandwidth ten times the Nyquist frequency, then low-pass filtered using a linear-phase filter, and finally transformed back to the time domain. This approach avoids problems due to frequency truncation, and for high FIR orders is typically accurate within 0.2 dB up to near Nyquist.

RESULTS USING SIMULATED DATA

Wave separation

The wave separation method was analysed using simulated measurement data as follows. A pressure wave p_1^+ at microphone 1 is defined as the sum of a low-pass-filtered square wave and white noise (40 dB less strong compared to the square wave), and a simple low-pass filter reflection is used to represent the source tube termination. The backward-wave and the waves at the other positions are then calculated using plane wave theory. The 'sensed pressures' at each of the three microphone positions are simulated by simply by adding some white noise (-40 dB) to the sum of the travelling waves. The propagation over distance *d* is simulated using the FIR design method mentioned in the previous section. Figure 6 compares the originally simulated pressure waves the backward-travelling wave with its estimate. The dotted vertical lines in (b) indicate the positions of singular frequencies, which occur at integer multiples of $f_c = c/(4d)$. Because half the wavelength equals the distance between the two microphones at these frequencies, the two microphone signal effectively provide the same information, and the problem becomes ill-conditioned. Hence for $d = 20$ mm, the system only provides accurate wave separation data up to about $0.9f_c \approx 3.9$ kHz. Zero-frequency also presents a singularity, and results below $0.1f_c \approx 0.4$ kHz tend to be noisy.

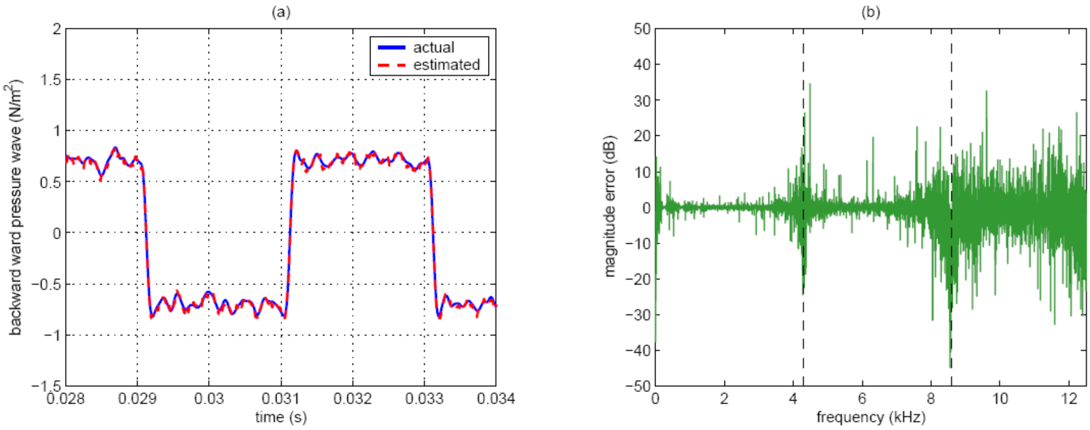

Figure 6. (a): a comparison between the simulated and the estimated backward-travelling pressure wave (b) the frequency-domain estimation magnitude error.

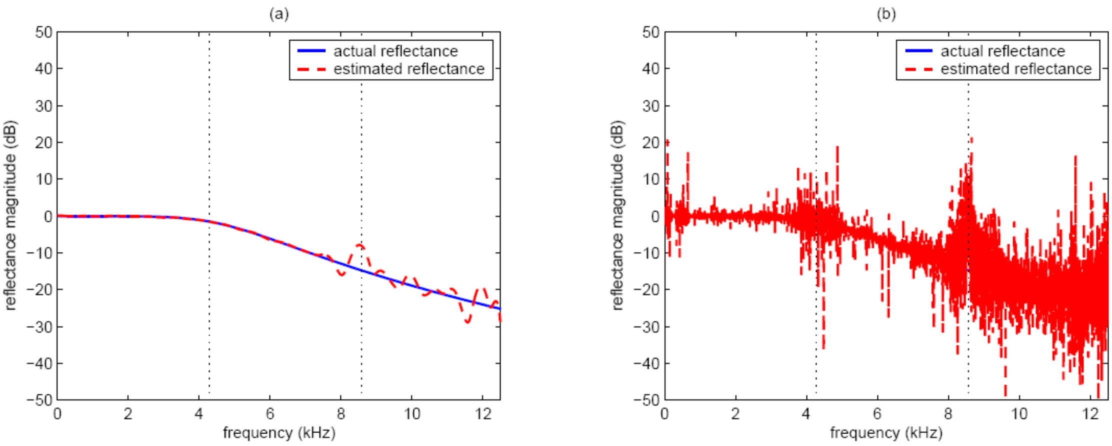

Figure 7. Bore reflectance estimation using (a) LMS, and (b) frequency-domain deconvolution.

Resonator reflectance estimation

Once the waves are estimated, they can be employed directly to characterise the bore reflectance as seen from the reference plane. In our simulation, the bore was set to be purely cylindrical, and terminated by a low-pass filter reflectance. The most straight-forward way of estimating the bore reflectance this is to perform deconvolution by dividing in the frequency domain. However, this leads to rather noisy results at frequencies at and near singular frequencies (see Figure 7b). Applying instead adaptive filtering with the standard LMS algorithm, where the FIR filter tap length is set to a length that could be expected to be the effective reflectance length, generally gives a much smoother result (see Figure 7a). So interestingly, the LMS estimation appears to 'overcome' the ill-conditioning problems around the first singular frequency $f_c \approx 4.3$ kHz.

CONCLUSIONS

In this paper we have outlined a strategy for sensing oscillations in a woodwind instrument as it is being played, with the purpose of separately characterising the bore (as a linear resonator) and the reed (as a non-linear driving element), as controlled by the player. It is envisaged that such a characterisation is an essential step forward in developing a system that allows precise physics-based re-synthesis of woodwind tones, which is the long-term objective of the authors.
A method was presented that estimates the forward- and backward-travelling wave in the bore using a delay-loop filtering structure, that is directly derived from the relationships between pressure and volume velocity at the three sensing points. An adaptive calibration of this method is possible, by using the mid-point sensor signal as a reference pressure, and optimising the wave propagation time between the sensors.

Results with computer-simulated data show that the waves can in principle be estimated accurately within a limited bandwidth; further estimation of the bore reflectance from the wave separation data indicates that the adaptive filtering approach can be advantageous in the sense of being less sensitive to noise than frequency domain deconvolution.

The described methodology relies on the assumption that the third microphone is positioned exactly between the two other microphones. However in practice this is difficult to ensure, since one tends not to know the exact 'acoustic center' of each microphone. A possible improvement would therefore be to allow and compensate for distance errors. A second possible improvement might be possible by direct adaptive estimation of τ from the data. Future research will focus on these possible improvements, as well as on the application to measurement data.

REFERENCES

[1] M. Aramaki, J. Bensa, L. Daudet: P. Guillemain, and R. Kronland-Martinet. Resynthesis of coupled piano string vibrations based on physical modeling. Journal of New Music Research, **30 No. 3** (2001) 213-226.

[2] T. Helie, C. Vergez, J. Levine, and X. Rodet: Inversion of a physical model of a trumpet. In Proceedings of the International Computer Music Conference (1999) 149-152.

[3] W. D'Haes and X. Rodet: A new estimation technique for determining the control parameters of a physical model of a trumpet. In 6th International Conference on Digital Audio Effects (DAFx03), London, England (2003).

[4] G. P. Scavone and P. R. Cook: Combined linear and non-linear prediction in calibrating models of musical instruments to recordings. In Proceedings of the1994 Int. Computer Music Conf., Aarhus, Denmark (1994), 433-434.

[5] T. Stilson: Forward-going wave extraction in acoustic tubes, in Proceedings of the 1995 International Computer Music Conference, Banff (1995), 517–520.

[6] S. Haykin: Adaptive Filter Theory, Prentice Hall, 1996

[7] A. D. Pierce. Acoustics: An Introduction to its Physical Principles and Applications. American Institute of Physics, New York, 1989.

[8] J. Guérard and X. Boutillion. Real time acoustic travelling waves separation in a tube. In Proc. 1997 Int. Symposium on Musical Acoustics, 1997.