A PHYSICALLY-INFORMED AUDIO ANALYSIS FRAMEWORK FOR THE IDENTIFICATION OF PLUCKING GESTURES ON THE CLASSICAL GUITAR

Bertrand Scherrer, and Philippe Depalle
Sound Processing and Control Laboratory (SPCL), Center for Interdisciplinary Research in Music Media and Technology,
McGill University, 555 Sherbrooke Street West, Montréal, QC, Canada H3A 1E3
bertrand.scherrer@mcgill.ca

1. INTRODUCTION

This paper presents one part of a larger system whose ultimate purpose is to identify how a classical guitar was played from the analysis of audio recordings. Such a system could be used in a variety of musical tasks ranging from guitar pedagogy to audio effects. In this paper, the emphasis is put on one parameter guitarists can modify when plucking the string: the angle with which the string is released at the end of the finger-string interaction. While most guitarists may not be familiar with the *angle of release* (AOR) per se, they change it when choosing between a *rest stroke* (where the plucking finger rests on the next string) and a *free stroke* (where the plucking finger clears the next string).

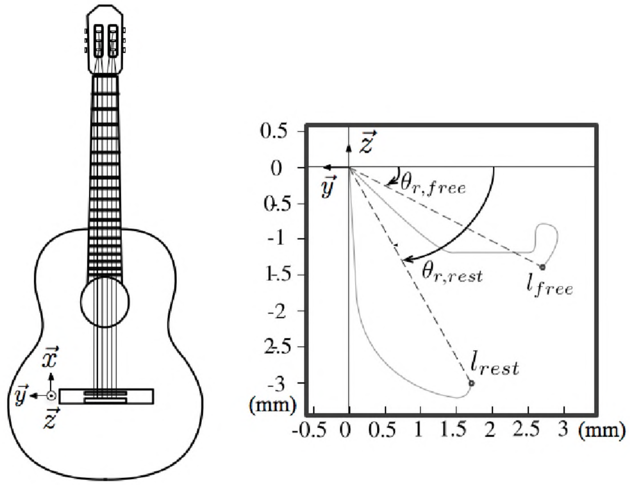

Figure 1: Coordinate system used in this paper (left), and measured trajectories of the string, near the plucking point, for rest and free strokes [1] (right).

The plot on the right of Figure 1, shows the difference in AOR between a free stroke and a rest stroke, respectively, as measured by [1]. The circled dots marked l_{free} and l_{rest} represent the points at which the string is released from the finger for the free and rest stroke.

2. MODELLING PLUCKING GESTURES

The Digital Waveguide (DW) paradigm [2] is used to devise a model for the vibration of one guitar string that takes into account the AOR. More specifically, a given string is represented by two bi-directional delay lines: the "*normal*" (z in Figure 1) and "*parallel*" (y in Figure 1) directions[1]. The coupling occurring between the two directions at the bridge is modelled using reflectance functions [2] as shown in Figure 2.

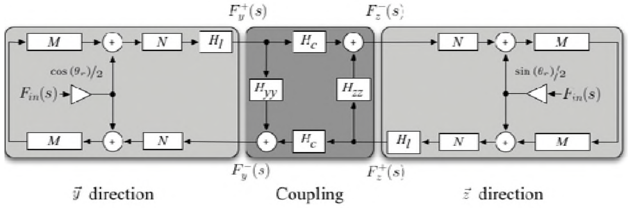

Figure 2: Diagram of the DW network simulating a string observed along the y and z directions, including the coupling at the bridge.

Force waves are used to study the mechanical behaviour of the coupled system. The AOR, θ_r, distributes the excitation, represented by its Laplace transform: $F_{in}(s)$, over the two transverse directions. The filter, H_l, accounts for the losses as well as potential phase delay during a round trip of the waves in the string. The terms of the 2x2 admittance matrix of the body at the bridge, as well as the characteristic impedance of the string are included in H_{yy}, H_{zz}, and H_c. The propagation of the waves from the plucking point to the bridge, and from the plucking point to the nut (or fret), are modelled using pure delay lines of lengths M and N, respectively.

Based on guitar radiation pattern measurements (e.g. [3]), it is assumed that the sound radiated by a guitar, and captured by a microphone positioned directly in front of it, only depends on the displacement of the bridge along z. The transfer function $T_z = F_z / F_{in}$ relating the input force to F_z, the force applied to the bridge along z, is given by:

$$T_z = \frac{e^{-Ns}(1+ e^{-2Ms})(\alpha_z + \beta_z e^{-Ds})}{2\left(1-[H_{yy}+H_{zz}]e^{-Ds}+[H_{yy}H_{zz}-H_c^2]e^{-2Ds}\right)}$$

where $\alpha_z = [1+H_{zz}]\sin\theta_r + H_c\cos\theta_r$ and $\beta_z = [H_c^2 - H_{yy} - H_{zz}H_{zz}]\sin\theta_r + H_c\cos\theta_r$. Due to the coupling at the bridge, T_z exhibits pairs of poles with slightly different frequencies [4]. Each pair of poles (and their conjugate counterparts) give rise to one string "partial"

---

[1] With respect to the guitar's top plate.

when observed with standard Fourier analysis. The factor $1 + e^{-2Ms}$ at the numerator of T_z is a feedforward comb filter that represents the effect of the plucking position on the sound. A method to estimate the value of M from audio signals can be found in [5]. The effect of the AOR on F_z is included in the terms α_z and β_z. This means that a change in the AOR will lead to a change in the amplitudes and phases of the 2D poles of T_z.

The analysis framework presented in the following section aims at extracting the amplitude, phase, frequency and damping factor of the poles of T_z from a recording. The ratio of the complex amplitudes of the modes forming a string partial allows for the retrieval of the AOR.

3. ANALYSIS FRAMEWORK

The first step in the analysis chain is to pre-process the sound (high-pass filtering, DC-removal). Then, the signal is segmented to isolate plucks, and select the portion of the sound to analyze. In a global analysis step, the fundamental frequency is estimated on the audio segment. Then, a standard additive analysis is carried out (e.g. [6]) to identify the string partials. The frequency resolution is not sufficient, however, for the task of determining the frequency, damping factor, amplitude, and phase of the closely coupled modes in a given string partial.

Hence, the next step of the analysis chain consists of a focused analysis around each string partial. The frequencies and damping factors of the modes are estimated using the ESPRIT method [7]. The amplitudes and initial phases of the poles are estimated via the least squares method. To avoid numerical issues, a multi-band approach is adopted along with decimation. In other words, ESPRIT looks for K modes (modelled as exponentially decaying sinusoids) around each partial. The partials are isolated using a linear phase FIR band pass filter, and the resulting signal is then decimated.

Based on the model in Section 2, $K=2$ should suffice since we assume a string partial is the result of the coupling of two transverse directions only. With ESPRIT, as with such parametric methods however, an overestimated K is often advisable (K=5 in practice). Therefore, a pruning scheme is implemented in order to successfully identify the two modes of interest. Simple physical considerations are used in this process: a) the damping factor for the vibration along y will be lower than that of the vibration along z; b) the vibration along z will tend to have a slightly smaller effective vibrating length than the vibration along y, due to different terminations at the fret or nut [8]: we can expect $f_y > f_z$.

Also, for a given partial of frequency f_n, the components that are "too weak" in amplitude, "too far" from f_n, that are over-damped, or diverging, are discarded. The final choice of the coupled mode pairs is done based on "how well" each pair of components models the partial (using the Reconstruction Signal to Noise Ratio, a.k.a. RSNR.

4. RESULTS AND DISCUSSION

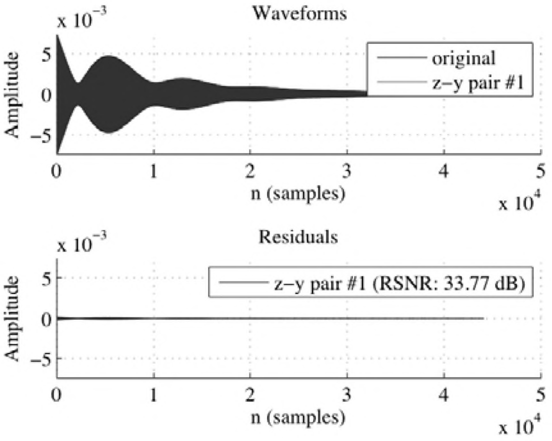

Figure 3: Waveforms of the 8th partial of a synthetic sound (top-black), of the 8th partial modelled by the identified z-y pair (top-grey) , and of the residual (original – zy_pair) (bottom).

An overview of an analysis framework of guitar sound has been presented. The physical model of Section 2 lead to the structure of the analysis chain presented in Section 3. The quality of the analysis of the framework is illustrated in Figure 3: the waveform of the 8th partial of a synthetic sound, and the waveform re-synthesized based on the estimated the z-y pair components are indistinguishable (RSNR of over 30dB). Future work includes the analysis of a database of plucks recordings with various AORs to evaluate the actual AOR estimation procedure.

REFERENCES

[1] Pavlidou, M., (1997). A physical model of the string-finger interaction on the classical guitar. Ph.D. Thesis, University of Cardiff.

[2] Smith, J. O., (2011). Physical audio signal processing. https://ccrma.stanford.edu/~jos/pasp/, accessed July 26th 2011.

[3] Hill, T. J. W., Richardson, and B. E., Richardson, S. J., (2004). Acoustical parameters for the characterisation of the classical guitar. Acustica - Acta Acustica, vol. 90(2), pp. 335–338.

[4] Gough, C. E. The theory of string resonances on musical instruments. Acustica, vol. 49, pp. 124–41.

[5] Traube, C., and Depalle, P., (2003). Deriving the plucking point along a guitar string from a least-square estimation of a comb filter delay, Proc. of IEEE CCECE.

[6] Smith, J. O., and Serra, X. (1987). PARSHL: An analysis/synthesis program for nonharmonic sounds based on sinusoidal representation, in Proc. ICMC.

[7] Roy, R., Paulraj, A. and Kailath, T. (1986). ESPRIT – a subspace rotation approach to estimation of parameters of cisoids in noise. IEEE Trans. on Acoustics, Speech, Signal Processing, vol. 34, pp. 1340–1342.

[8] Woodhouse, J. (2004). Plucked guitar transients: Comparison of measurements and synthesis. Acustica – Acta Acustica, vol. 90, pp. 945–65.