PARAMETRIC AUDIO CODING OF BASS GUITAR RECORDINGS USING A TUNED PHYSICAL MODELING ALGORITHM

Jakob Abeßer, Patrick Kramer, Christian Dittmar, Gerald Schuller
Fraunhofer IDMT
Ilmenau, Germany
jakob.abesser@idmt.fraunhofer.de

ABSTRACT

In this paper, we propose a parametric audio coding framework that combines the analysis and re-synthesis of electric bass guitar recordings. In particular, an existing synthesis algorithm that incorporates 11 playing techniques is extended by two calibration algorithms. Both the temporal and spectral decay parameters as well as the inharmonicity coefficient are set according to the fretboard position on the instrument. Listening tests show that there is still a gap in perceptual quality between real-world instrument recordings and the re-synthesized versions. Due to this gap, the perceived improvement due to the model calibration is only small. Second, the listening tests reveal that the plucking styles are more important towards realistic synthesis results than expression styles.

1. INTRODUCTION

1.1. Analysis & Synthesis of Musical Instrument Recordings

In the past decades, various publications focused on the *analysis* and *synthesis* of musical instrument recordings. These recordings can be described as a sequences of musical note events. Multiple algorithms for automatic music transcription, i.e., the extraction of *score-related* note parameters such as pitch, onset, and duration, were proposed so far. Similarly, methods for estimating *instrument-related* parameters such as the fretboard position (string and fret number) [1] or the applied playing techniques [2] were presented for different string instruments.

For the purpose of audio synthesis, algorithms exist to mimic the sound production of string instruments such as the guitar, the violin, or the piano. These *physical modeling* algorithms are commonly based on a simple model of the string vibration and include model extensions to simulate different aspects of sound production such as the instrument body, coupling effects between strings, or different playing techniques on the instrument [3, 4, 5].

The representation of instrument tracks based on note-wise parameters has many advantages and application scenarios: First, the estimated parameters can be used to automatically generate written music notations such as scores or tablatures, which can be utilized in different music education scenarios. Second, the parametric representation allows for expressive modifications of the musical content such as changing the rhythmic structure (note onsets and offsets) or changing the applied playing techniques. Finally, a significantly lower bit-rate can be achieved for the transmission of encoded audio recordings in comparison to conventional audio coding schemes.

In this paper, we propose to combine both analysis and synthesis algorithms to a *parametric audio coding framework*. In particular, we focus on a sound synthesis algorithm for bass guitar recordings that includes 11 playing techniques and can be tuned to the sonic properties of a particular instrument.

1.2. The Bass Guitar

1.2.1. Fretboard Position

The bass guitar—as well as other string instruments—allows to play notes within a certain pitch range at different *fretboard positions*. The fretboard position is defined by the string number and the fret number. Notes played on different strings not only vary in pitch but also in terms of timbre-related properties. The four bass guitar strings have diameter values between 0.45 mm and 1.05 mm, different inharmonicity coefficients β, and time-varying magnitude relationships between the overtones as shown in [1]. These sonic properties must be considered in a bass guitar sound synthesis algorithm.

1.2.2. Playing Techniques

Table 1: *Proposed taxonomy of bass guitar playing techniques.*

Plucking Style		**Expression Style**	
Finger-Style	(FS)	Normal	(NO)
Picked	(PK)	Harmonics	(HA)
Muted	(MU)	Dead-notes	(DN)
Slap-Thumb	(ST)	Vibrato	(VI)
Slap-Pluck	(SP)	Bending	(BE)
		Slide	(SL)

We model the performance of the bass guitar player by means of two consecutive playing gestures: The *plucking style* describes the initial plucking of a string using the playing hand; the *expression style* describes how the fretting hand is used to manipulate the string vibration. Table 1 gives an overview over the 11 playing techniques included in the proposed sound synthesis model.

If the *finger-style* technique is used, the strings are alternately plucked by the index and the middle finger. Alternatively, the string can also be plucked using a plastic pick (*picked*). The plucking style *muted* describes the plucking of the string using the thumb of the playing hand while simultaneously damping the vibrating string by using the inner hand side. The amount of damping effectively shortens the note decay part. The two techniques *slap-thumb* and *slap-pluck* describe the striking of the string using the thumb and the picking of a string using either the index or the middle finger, respectively. Both plucking techniques cause the string to hit the higher frets on the instrument neck due to the high deflection of the string. Hence, the two slap techniques result in a typical metallic sound.

The most common way to play the bass guitar is to do so without any expression style (which will be hereafter referred to as *normal*). Here, one of the fingers of the fretting hand is first located at a defined fretboard position that corresponds to the desired note pitch. The *dead-note* technique describes the damping of the string using the fretting hand, which leads to a percussively sounding note with almost no harmonic components. Similarly, the fretting hand can be used to softly dampen the string vibration at integer fractions of the string length (denoted as *harmonics*). Then, a standing wave with a node at the damping point is excited on the string. Hence, all vibration modes with an anti-node at the damping point cannot be excited. Frequency modulation can be applied to achieve arbitrary (continuous) pitch alterations. After plucking the string with the plucking hand, the fretting hand can be used to bend the string up- and downwards once (*bending*) or periodically (*vibrato*). Instead of playing two consecutive notes with two plucking gestures, the musician often plays the first note and *slides* upwards or downwards to the next note without a second note pluck.

1.3. Goals

Our overall goal is to develop a parametric audio coding framework that is tailored towards bass guitar recordings. The framework consists of an analysis and synthesis stage as shown in Figure 1. The analysis stage can be thought of as an extended bass transcription algorithm, which extracts not only the score-related note event parameters pitch, onset, and offset, but also the instrument-specific parameters plucking style, expression style, string number, and fret number. These parameters are transmitted to the synthesis stage and converted to a synthetic audio signal using a physical modeling algorithm. So far, we presented algorithms for estimating bass guitar playing techniques [6] as well as the fretboard position [1] from isolated notes.

In this paper, we will focus only on the synthesis stage of the coding framework. Therefore, we use all note parameters from the (correct) reference annotations of our data-set (see Section 4.1) and feed them to the synthesis stage. The influence of transcription errors, which are inevitable in state-of-the-art automatic transcription algorithms, towards the synthesis quality will not be discussed in this paper.

To summarize, we follow two goals in this paper. First, we will propose two methods to adapt an existing (generic) bass guitar synthesis algorithms [7] to the sonic properties of a particular bass guitar. Second, based on a MUSHRA listening test, we will evaluate, which note parameters are perceptually more significant than others in order to achieve a realistic re-synthesis of bass guitar recordings.

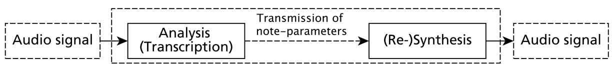


Figure 1: *Proposed parametric audio codec framework.*

1.4. Outline

This paper is structured as follows: We will specify our research objectives followed in this paper in Section 1.3. After a review of related work in Section 2, we will describe an algorithm for physical modeling of bass guitar recordings. More specifically, we will propose two algorithms to calibrate the synthesis model towards the sonic properties of a particular bass guitar in order to achieve better sound quality results in Section 3.2. In Section 4.1, we introduce a novel public dataset that was created for evaluating the system for real-world bass guitar recordings including various playing techniques. We describe the applied listening tests in Section 4.2 and discuss the obtained results in Section 5.

2. PREVIOUS WORK

Parametric audio coding has a long tradition. For instance, speech coding uses an extensive source model for compression [8]. The vocal cord is modeled with an FIR filter and the excitation signal of the Glottis are approximated using pulses or noise. This way, very low bit-rates can be obtained. Since this model also has a physical meaning, its parameters are also used in speech recognition applications.

For more general audio signals, the aforementioned approach is less effective because of the variety of possible audio sounds or musical instruments. However, if we restrict ourselves to a specific instrument, we can use a detailed model of the sound production to efficiently encode and reproduce its sound, and also to modify or classify its sound in a musically meaningful way. As examples, in [9] an approach to encode or represent a clarinet signal using a physical model is presented, and [10] presents a parametric instrument codec for guitars. An important part of these approaches is to design a model with parameters which can be readily extracted from the instrument sound.

In terms of sound synthesis, the digital waveguide algorithm proposed by Smith [11] is a computational efficient method to simulate the behavior of vibrating strings. It was developed further in numerous publications such as [12] and [13] for the synthesis of different string instruments. So far, only two publications proposed synthesis algorithms for bass guitar synthesis. Both Rank and Kubin [14], Trautmann and Rabenstein [15], as well as Janer et al. [16] presented algorithms to synthesize the slap technique (compare Section 1.2.2) by simulating the string collision with the fretboard.

In a similar fashion, models for the collision between strings and the fretboard of the guitar [3] and for the interaction between the player's fingers and the string [4] were proposed. In [5], Lindroos et al. extend an existing electric guitar synthesis model. The authors propose a three-part excitation function that allows to simulate different dynamic levels, different plucking angles of the plectrum[1], two-stage note decay, and the effect of magnetic pickups on the electric guitar. Concerning the expression techniques listed in Table 1, guitar synthesis models were proposed that incorporate frequency modulation techniqes such as vibrato [12], harmonics [17], and dead-notes [3].

Since commonly used digital representation formats such as MIDI and MusicXML are limited towards notating instrument performances, new notation methods and transmission protocols were proposed such as the Expressive Notation Package (ENP) [12], the

[1] This corresponds to the *picked* plucking style explained in Section 1.2.2.

PatchWork Graphical Language (PWGL) [18], the Music Parameter Description Language (MPDL) [19], or the Guitar Control Language [20].

3. PROPOSED SYSTEM

3.1. Physical Modeling Algorithm for Bass Guitar Synthesis

In this paper, we build upon a model previously presented in [7]. The synthesis algorithm is based on a waveguide model, which is simulating the sound of the bass guitar strings by decomposing the vibration of the string into two travelling waves. These waves are represented by two delay lines connected through inverting reflection terminations. A zero phase FIR filter and a damping factor are used to recreate the natural energy losses of the vibrating string. The sound is picked up in relation to the bass guitars electromagnetic pickup positions. Additionally, a filter is added to mimic the pickup's output frequency response.

The synthesis algorithm includes modular extensions to cover all bass guitar playing techniques listed in Table 1. Unique displacement functions are used for each plucking style to reproduce their characteristic attack phases. Additionally, a fretboard representation is implemented to simulate the string-fretboard collision for the two slap techniques ST and SP. The strong signal damping that is typical for the muted plucking style (MU) is achieved by adjusting the damping modules of the waveguide model accordingly.

The NO expression technique was realized by adjusting the delay line length using linear interpolation to produce the notes' exact fundamental frequency f_0. Depending on the fret position, the pluck and pickup position are adjusted for each string. The frequency modulating techniques BE, VI, and SL are realized by varying the delay line length according to a technique specific modulation function (f_0 course over time). Additionally, a module was added to simulate a punctual damping of the simulated string vibration to realize the HA and DN techniques. The interested reader is referred to [7] for further details.

3.2. Model Calibration

As detailed in Section 4, we evaluate the perceptual quality of the synthesis algorithm by comparing originally recorded bass lines with their re-synthesized counterpart. In order to tune the synthesis model to the sonic properties of the particular bass guitar[2] that was used for the recordings, we analyze isolated note recordings over the full fretboard range (all four strings, open string up to the 12th fret position). These notes were taken from the *IDMT-SMT-BASS* dataset (see Section 4.1). We propose three steps to tune the synthesis algorithm: tuning of the temporal loss parameter, tuning of the frequency loss filter, and tuning of the inharmonicity of the synthesized notes.

3.2.1. Tuning of the Loss Filters

As mentioned in 3.1, the simulation of the string vibration losses is realized by two components, a damping factor g and a FIR filter (see Losses in Fig. 2). The damping factor g is scaling the excitation signal over time and therefore responsible for the overall temporal decay of each note. To tune this parameter, the decay rate α was extracted using linear regression over the magnitude spectrogram (in dB) in the note decay part. This was done for each note in the dataset for every string and fret position. The corresponding damping factors were calculated for the waveguide model by $g = 10^{\frac{-\alpha NT}{20}}$ with N denoting the number of delay elements and T denoting the time of one sample.

The symmetrical FIR filter on the other hand serves as a zero phase lowpass filter. It introduces a faster decay to higher partials of the fundamental frequency to reproduce the natural faster damping of higher frequencies caused by string stiffness. Details on the applied filter can be found in [21]. It allows to adjust the raising decay rate towards higher frequencies with just one parameter. To tune the FIR filter, the time gradients of the decay rate of the magnitude spectrogram over frequency of every string-fret combination were estimated and the frequency loss filter was tuned accordingly by minimizing the distance of this gradient between synthesized and original decay.

3.2.2. Inharmonicity Tuning

Due to dispersive wave propagation within the vibrating string, the overtone frequencies of notes played on a real bass guitar deviate from the ideal integer relationship to the fundamental frequency, this effect is called inharmonicity [22]. We apply the inharmonicity tuning as post-processing step to single notes synthesized by the physical modeling algorithm. For each note event, the desired inharmonicity parameter β and the fundamental frequency f_0 are given. The inharmonicity coefficient has been extracted by polynomial interpolation as described in [1] and the fundamental frequency is chosen according to the reference note pitch. The center frequencies f_k of the harmonic series in the synthesized note are located at $f_k = kf_0$ with $k \geq 1$. Our goal is to enforce a frequency deviation on each harmonic according to $\hat{f}_k = kf_0\sqrt{1+\beta k^2}$ Thus, we extract the band signal of each individual harmonic by means of Short-Term Fourier Transform (STFT), binary spectral masking and inverse STFT. The binary mask for each harmonic is centered around the ideal f_k and has a bandwith of f_0. During the spectral masking, the mirror spectra are set to zero which yields the analytic signal [23] after inverse STFT. The analytic signal contains the original signal in its real part and a version of the original signal with constant phase shift of $90°$ in its imaginary part. We modulate the analytic signal with a complex exponential $\exp(j2\pi(\hat{f}_k - f_k))$. This procedure shifts the band signal upwards by the difference between inharmonic and ideal center frequency. This is repeated for all harmonics, and the individually frequency shifted band signals are superimposed. Thus, we get an inharmonic note signal with the original timbre qualities retained. For real-time purposes the waveguide model could alternatively be extended by an additional allpass filter [24].

4. EVALUATION

4.1. Dataset

In this publication, we introduce the novel *IDMT-SMT-BASS-SINGLE-TRACKS*[3] dataset. It comprises of 17 bass lines from different music styles (blues, rock, funk, bossa

[2] Fame Baphomet 4 NTB, string gauges 1.05 mm (E1), .85 mm (A1), .65mm (D2), and .45 mm (G2)

[3] see http://www.idmt.fraunhofer.de/en/Departments_and_Groups/smt/bass_lines.html

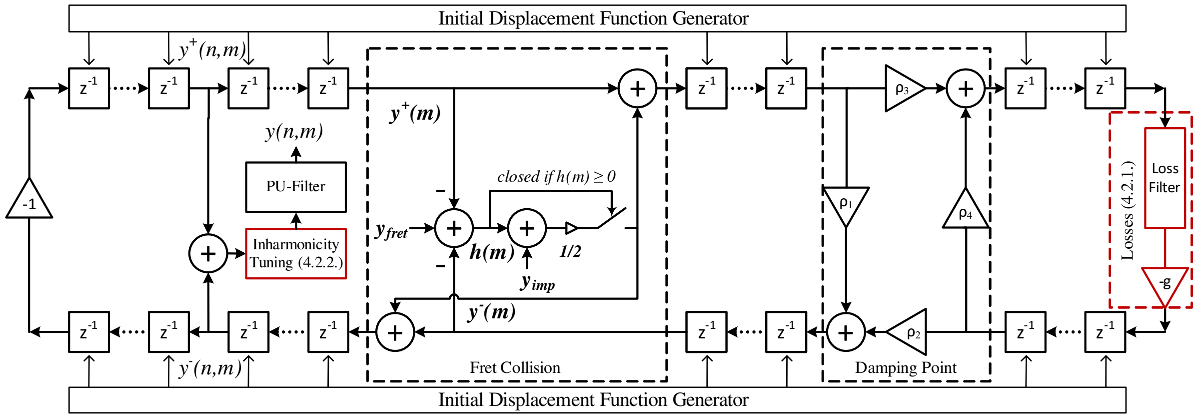


Figure 2: *The waveguide model presented in [7] including a collision and damping point. The number of the corresponding sections, where the model components are explained, are given in brackets.*

nova, and hip hop) that were recorded with an electric bass guitar. The notes in this dataset cover all playing techniques listed in Table 1 as well as all 4 strings of the bass guitar. The dataset contains extensive note-wise annotations (pitch, onset, offset, playing techniques, fretboard position) as well as structural annotations (repeating pattern appearances). Hence, it is designed to be applicable for different tasks such as music transcription, playing technique estimation, as well as retrieval of repeating bass patterns.

Table 2: *Distribution of all notes contained in the IDMT-SMT-BASS-SINGLE-TRACKS. Appearance of each class is given as absolute number of notes (N) as well as percentage value (%).*

Plucking Styles	**FS**	**MU**	**PK**	**ST**	**SP**	
N	419	216	138	114	54	
%	44.53	22.95	14.67	12.11	5.74	
Expression Styles	**NO**	**VI**	**BE**	**SL**	**DN**	**HA**
N	822	31	8	20	28	32
%	87.35	3.29	0.85	2.13	2.98	3.4
String	**E**	**A**	**D**	**G**		
N	234	313	256	138		
%	24.87	33.26	27.21	14.67		

All 17 bass lines were recorded using a Fame Baphomet 4 NTB bass guitar with 4 strings in standard tuning E1, A1, G2, and D2. The audio files are provided in the WAV format with 44.1 kHz and 16 bit. This is the same instrument and setting as it was used for the recording of the *IDMT-SMT-BASS* dataset[4] published in 2010 that contains single note bass guitar recordings with all of the above-mentioned plucking and expression styles.

4.2. Listening Test Methodology

In this paper, we aim to evaluate the influence of the model tuning and the selection of parameters to the perceptual quality of the re-synthesized bass lines. For this purpose, we conduct two MUSHRA tests (Multi Stimulus test with Hidden Reference and Anchor) as will be explained in the following two subsections. Table 3 gives an overview over the stimuli used in the listening tests.

[4]see `http://www.idmt.fraunhofer.de/en/Departments_and_Groups/smt/bass.html`

The original recordings of the bass lines are used as hidden reference and a low-pass filtered version (f_{cut} = 3.5 kHz) is used as hidden anchor. Nine participants—most of them being semi-professional musicians—took part in the listening tests.

4.2.1. Experiment 1 - Perceptual Influence of the Synthesis Model Tuning

In the first experiment, we aim to evaluate, how the model tuning described in Section 3.2 improves the perceptual quality of the re-synthesized bass lines: In addition to the hidden reference and anchor, we use three model configurations to synthesize the bass lines. The first configuration is the synthesis model used in [7]. Within the given scope, this model is considered as untuned model, as it was empirically adjusted and only separates between the four strings. The second model is tuned in such way that the filter parameters were optimized based on given single-note recordings as described in Section 3.2.1. The third model is the tuned model with an additional inharmonicity tuning as described in Section 3.2.2.

4.2.2. Experiment 2 - Perceptual Influence of Playing Techniques

In the second experiments, we aim to evaluate the importance of the plucking and expression styles to the perceptual quality of the re-synthesized bass lines. Again, we use the hidden reference and anchor signal. Additionally, four synthesized versions are generated: The first version contains all plucking and expression styles as they were given by the ground truth annotations. In the second version, the plucking styles of all notes are neglected (set to FS) and the expression styles remain correct. In the third version, the expression styles of all notes are neglected (set to NO) and the plucking styles remain correct. Finally, in the fourth version, both the plucking and expression styles are neglected and set to FS and NO, respectively.

5. RESULTS

5.1. Results

5.1.1. Experiment 1 - Perceptual Influence of the Synthesis Model Tuning

The results of the first experiments are shown in Figure 3(a). While the reference was rated as excellent, the low-pass filtered anchor was rated as good. However, for the bass lines 7, 13, and 16, which were played using the slap techniques, the ratings were significantly lower. Since the slap technique is characterized by typical high frequency attack transients, the low-pass filtering impairs the audio quality more. It can be seen in the final column that the optimized synthesis algorithms with and without inharmonicity show no significant improvement to the baseline model.

5.1.2. Experiment 2 - Perceptual Importance of Playing Techniques

The results of the second experiments are illustrated in Figure 3(b). The ratings of the reference and the anchor are comparable to the first experiment. Again, the perceived audio quality of the slap bass lines is more strongly affected by the low-pass filtering. The averaged results in the final column indicate that changing the plucking style of all notes decreases the audio quality of the resynthesized bass lines more than changing the expression styles. The reason for that might be that most of the notes have no expression style (NO) at all, hence, removing all vibrato, bending, slide, harmonic, and dead note techniques does not change the bass line very much. On the other hand, if a bass line has a different plucking style than the finger-style technique (FS), changing the plucking style affects all notes in the bass line. The negative effect of changing the plucking style to "finger-style" is most prominent for the bass lines using either the slap-techniques ST and SP or the picked technique (PK).

6. CONCLUSIONS

In this paper, we proposed algorithms to calibrate a given bass guitar synthesis model to the sonic properties of a given instrument. The first listening test revealed that the improvements in the perceived audio quality of the resynthesized bass lines are only small. Interviews with the listening test participants confirm that the overall "synthetic" impression of the audio data still "masks" the perceptual improvements by the proposed tuning approaches. Future research must identify aspects of the instrument sound production that are still not captured by the synthesis model. The estimation of temporal and spectral decay parameters as well as the inharmonicity coefficient from single notes can be considered as less important to the overall perceptual quality of the synthesis. A second outcome of the listening tests was that the plucking styles have a higher importance for re-synthesizing string instrument recordings than the expression styles have. A reliable parametrization and modeling of the attack part of instrument notes (which is mainly influenced by the plucking style) is therefore crucial.

7. ACKNOWLEDGMENTS

The authors would like to thank the participants of the listening test and the anonymous reviewers.

8. REFERENCES

[1] Jakob Abeßer, "Automatic string detection for bass guitar and electric guitar," in *Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR)*, London, UK, 2012.

[2] Jakob Abeßer, Hanna Lukashevich, and Gerald Schuller, "Feature-based extraction of plucking and expression styles of the electric bass guitar," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2010, pp. 2290–2293.

[3] Gianpaolo Evangelista and Fredrik Eckerholm, "Player-instrument interaction models for digital waveguide synthesis of guitar: Touch and collisions," *IEEE Transactions on Audio, Speech & Language Processing*, vol. 18, no. 4, pp. 822–832, 2010.

[4] Giuseppe Cuzzucoli and Vincenzo Lombardo, "A Physical Model of the Classical Guitar, Including the Player's Touch," *Computer Music Journal*, vol. 23, no. 2, pp. 52–69, 1999.

[5] Niklas Lindroos, Henri Penttinen, and Vesa Välimäki, "Parametric Electric Guitar Synthesis," *Computer Music Journal*, vol. 35, no. 3, pp. 18–27, 2011.

[6] Jakob Abeßer, Christian Dittmar, and Gerald Schuller, "Automatic Recognition and Parametrization of Frequency Modulation Techniques in Bass Guitar Recordings," in *Proc. of the 42nd Audio Engineering Sociery (AES) International Conference on Semantic Audio*, Ilmenau, Germany, 2011, pp. 1–8.

[7] Patrick Kramer, Jakob Abeßer, Christian Dittmar, and Gerald Schuller, "A digital waveguide model of the electric bass guitar including different playing techniques," in *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Kobe, Japan, 2012.

[8] N. Jayant and Peter Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice-Hall, Englewood Cliffs, 1984.

[9] Mark Sterling, Xiaoxiao Dong, and Mark Bocko, "Representation of Solo Clarinet Music by Physical Modeling Synthesis," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, USA, 2008, pp. 129–132.

[10] Mirko Arnold and Gerald Schuller, "A Parametric Instrument Codec for Very Low Bitrates," in *Proceedings of the 125th Audio Engineering Society (AES) Convention*, San Francisco, CA, USA, 2008.

[11] Julius O. Smith, "Physical modeling using digital waveguides," *Computer Music Journal*, vol. 16, pp. 74 – 91, 1992.

[12] Mikael Laurson, Jarmo Hiipakka, Cumhur Erkut, Matti Karjalainen, Vesa. Välimäki, and Mika Kuuskankare, "From Expressive Notation to Model-Based Sound Synthesis: a Case Study of the Acoustic Guitar," in *Proceedings of the 1999 International Computer Music Conference*, 1999, pp. 1–4.

[13] Vesa Välimäki, Matti Karjalainen, Zoltàn Jànosy, and Unto K. Laine, "A real-time DSP implementation of a flute model," in *International Conference on Acoustics, Speech and Signal Processing*, 1992, pp. 249–252.

[14] Erhard Rank and Gernot Kubin, "A waveguide model for slapbass synthesis," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1997, pp. 443–446.

[15] Lutz Trautmann and Rudolf Rabenstein, "Stable systems for nonlinear discrete sound synthesis with the functional transformation method," in *Proceedings of the 27th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Orlando, USA*, 2002, vol. 2, pp. 1861–1864.

[16] Jordi Janer, "Voice-controlled plucked bass guitar through two synthesis techniques," in *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, 2005, pp. 132–135.

[17] Jyri Pakarinen, "Physical modeling of flageolet tones in string instruments," in *Proceedings of the 13th European Signal Processing Conference*, 2005, pp. 4–8.

[18] M. Laurson, M. Kuuskankare, and V. Norilo, "An Overview of PWGL, a Visual Programming Environment for Music," *Computer Music Journal*, vol. 33, no. 1, pp. 19–31, 2009.

[19] Keith McMillen, David Wessel, and Matthew Wright, "The ZIPI music parameter description language," *Computer Music Journal*, vol. 18, no. 5, pp. 52–73, 1994.

[20] Matti Karjalainen, Teemu Mäki-Patola, Aki Kanerva, Antti Huovilainen, and Pekka Jänis, "Virtual Air Guitar," in *Proceedings of the 117th Audio Engineering Society (AES) Convention*, San Francisco, CA, USA, 2004, pp. 1–19.

[21] Maarten van Walstijn, "Parametric fir design of propagation loss filters in digital waveguide string models," *IEEE Signal Processing Letters*, vol. 17, no. 9, pp. 795 – 798, 2010.

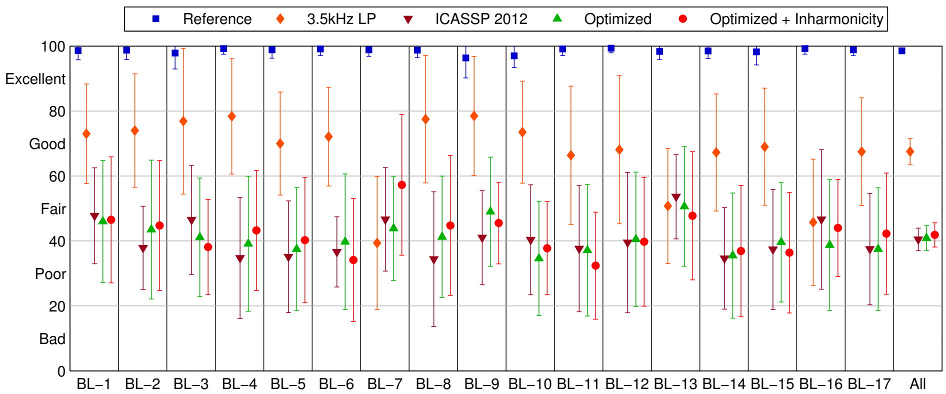
[22] Neville H. Fletcher and Thomas D. Rossing, *The Physics Of Musical Instruments*, Springer, New York, London, 2 edition, 1998.

[23] Frederick W. King, *Hilbert Transforms:*, Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2009.

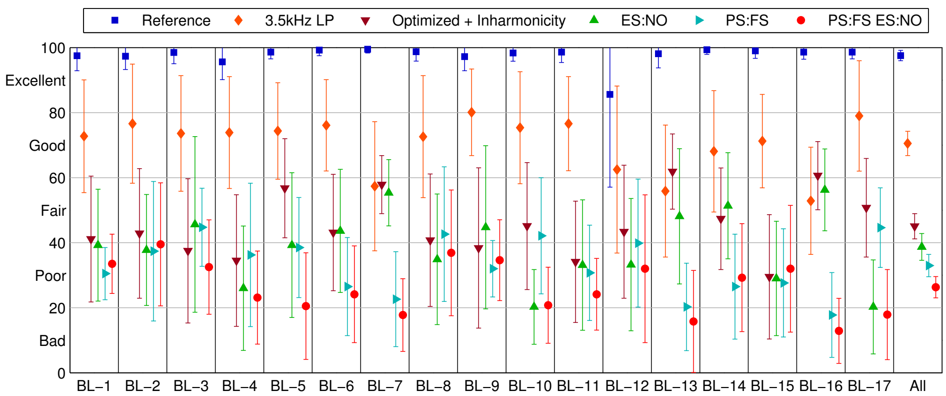
[24] Jukka Rauhala and Vesa Välimäki, "Dispersion modeling in waveguide piano synthesis using tunable allpass filters," in *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06)*, 2006, pp. 71–76.

Table 3: *Stimuli in the MUSHRA listening tests in the two listening experiments*

Experiment	Original / Synthesis	Low-pass filtered	Tuning	Inharmonicity	Plucking & expression styles	Label
1	Original				all PS, all ES	Reference
1	Original	x			all PS, all ES	3.5kHz LP
1	Synthesis				all PS, all ES	ICASSP 2012
1	Synthesis		x		all PS, all ES	Optimized
1	Synthesis		x	x	all PS, all ES	Optimized + Inharmonicity
2	Original				all PS, all ES	Reference
2	Original	x			all PS, all ES	3.5kHz LP
2	Synthesis		x	x	all PS, all ES	Optimized + Inharmonicity
2	Synthesis		x	x	PS = FS, all ES	PS:FS
2	Synthesis		x	x	all PS, ES = NO	ES:NO
2	Synthesis		x	x	PS = FS, ES = NO	PS:FS ES:NO



(a) Experiment 1



(b) Experiment 2

Figure 3: *The results of the MUSHRA test for the different stimuli described in Table 3 for the 17 bass lines. Mean ratings with 95%-confidence intervals are given.*

Table 4: *Plucking styles (PS) and expression styles (ES) used in the bass lines of the IDMT-SMT-BASS-SINGLE-TRACKS dataset*

BL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
PS	MU	FS	MU	FS	PK	PK	ST, SP	FS	MU	FS	MU	FS	ST, SP	PK	FS	SP, ST	FS
ES	VI, NO, BE, DN	VI, NO	NO, DN, SL	NO, VI, SL, BE	NO, DN, SL	NO, VI	NO, DN, VI	NO	BE, NO, VI	NO, SL	NO	NO, SL	NO, SL	NO	NO, VI	NO, DN	NO, HA