J. Rauhala and V. Välimäki, Parametric excitation model for waveguide piano synthesis, in Proceedings of the 2006 IEEE International Conference on Acoustics, Speech, and Signal Processing, Toulouse, France, 2006, pp. 157-160.

PARAMETRIC EXCITATION MODEL FOR WAVEGUIDE PIANO SYNTHESIS

Jukka Rauhala and Vesa Välimäki

Helsinki University of Technology
Laboratory of Acoustics and Audio Signal Processing
P.O. Box 3000, FI-02015 TKK, Espoo, Finland
jukka.rauhala@acoustics.hut.fi

ABSTRACT

In this paper, a method providing an excitation signal for the waveguide piano synthesis is presented. The waveguide synthesis string model needs an excitation signal, which stimulates the model to resonate at the partial frequencies. This signal simulates the force pulse, which occurs in the piano when the hammer hits the string. In the proposed method, the excitation signal is produced by using additive synthesis with matching partial amplitudes and frequencies, and by adding bandlimited white noise into the signal. The excitation model takes into account the velocity at which the piano key is pressed, using bandstop and lowpass filtering. The proposed method is suitable for real-time piano synthesis, as it is controllable and computationally efficient.

1. INTRODUCTION

In digital waveguide synthesis [1], an excitation signal is critical to the perceived timbre of the tone. The signal must have energy at the desired partial frequencies so that the string model resonates at the partial frequencies. Moreover, the amount of energy at a certain partial frequency contributes together with the string model gain to the resulting partial amplitude. In piano synthesis, the excitation signal corresponds to the hammer action in the physical world. When the hammer hits the string, it passes on an amount of energy, which transforms into vibration in the string.

Several issues have to be taken into account when designing a model simulating the excitation signal. First, the excitation signal should be calibrated by using recorded tones, and it should produce tones with perceptually realistic timbre. Moreover, the excitation must depend on the key velocity in order to implement the dynamics of the piano. Another important requirement for the excitation model is that it must not cause any audible undesirable tones. It should also be independent of the tuning and the inharmonicity. For example, when the inharmonicity coefficient is controllable, it implies that the excitation method must take into account that the partial frequencies have been changed. In addition, the method should be computationally efficient in order to be used in real-time applications.

In [2]-[4], the excitation signal has been produced by modeling the hammer through using physics-based methods. Although this is a very interesting approach from the research point of view, these models fail to produce piano tones of perceptually perfect quality. The commuted waveguide synthesis approach [5] has been successful in modeling of several musical string instruments, such as the acoustic guitar [6] and the harpsichord [7], but it has some disadvantages in application to the piano, such as lack of control over individual partial amplitudes and frequencies. In addition, Bensa et al. [8] have proposed an excitation method based on subtractive synthesis.

In the proposed method, the excitation signal is produced by using additive synthesis and an equalizing filter, which is controlled by the hammer velocity. Moreover, the partials at high frequencies can be replaced with filtered white noise.

2. THE EXCITATION MODEL

An overview of the proposed method is shown in Figure 1. The method consists of five blocks: an additive synthesis generator, a white noise generator, an equalizing filter, a one-pole lowpass filter, and a shaping window generator. First, the source signal is generated with additive synthesis and is windowed with a shaping window. Then, the resulting signal is filtered with the equalizing filter, which attenuates the higher frequencies depending on the key velocity. When the key number is between 1-49 (key A_0 is denoted as key number 1), bandlimited white noise, which is windowed with the shaping window and is filtered with a velocity-dependent one-pole lowpass filter, is added into the signal.

The piano hammer not only excites the strings to vibrate, but also produces an audible impact sound, which contains additional frequency peaks. Hence, in order to simulate realistic piano tones, it is not enough to model only

1-4244-0469-X/06/$20.00 ©2006 IEEE

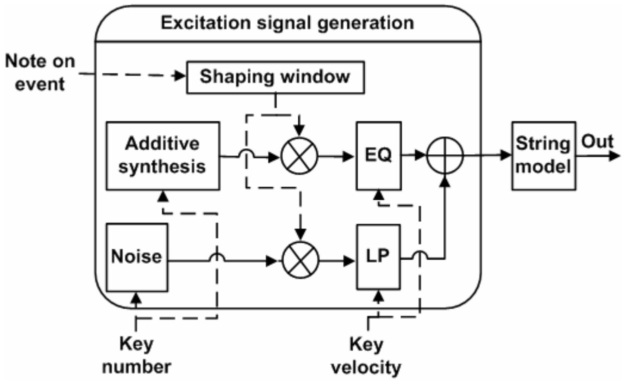

Fig. 1. An overview of the excitation signal generation method.

the partials. We suggest that the excitation signal can be separated into two parts: partials produced by the string, and additional sound produced by the hammer. The additional sound can be separated from recorded tones by filtering out the partials. Moreover, it can be reproduced, for example, by adding the hammer noise sample into the signal after the string model.

2.1. Source signal generation

A source signal can be generated accurately with the desired partial amplitudes and frequencies by using additive synthesis. Each partial is produced with a sine wave at a certain frequency and amplitude. In the end, all partial signals are summed into one signal. It is important to have variation in the phases of these partials in order to avoid cancellations between the partials. Instead of matching the estimated phases, it seems to be sufficient to use random phases. In addition, the length of the source signal should be one period of the fundamental frequency.

The number of generated partials with additive synthesis can be reduced by adding white noise into the source signal [9]. The noise must be highpass filtered to get rid of the undesired noise at low frequencies. We use a 50^{th}-order FIR filter in our model. In order to avoid too much extra calculation, we have decided to use four noise buffers, which have been filtered in advance with cutoff frequencies 2450 Hz, 3675 Hz, 5512.5 Hz, and 7350 Hz for key numbers 1-8, 9-23, 24-35, and 36-49, respectively (the sample rate used in this work is 44.1 kHz). Hence, the maximum number of partials generated with additive synthesis is reduced to 89.

The noise spectrum can be tilted by inserting a one-pole lowpass filter into the model, as proposed in [10]. The one-pole coefficient a_{op} was determined to be –0.998, and the one-pole gain G_{op} can be calculated in decibels as $G_{op} = [G_{min}-10(1-v)]$, where parameter G_{min} is the minimum value, which is shown in Figure 3(a) for different key numbers, and v is the key velocity.

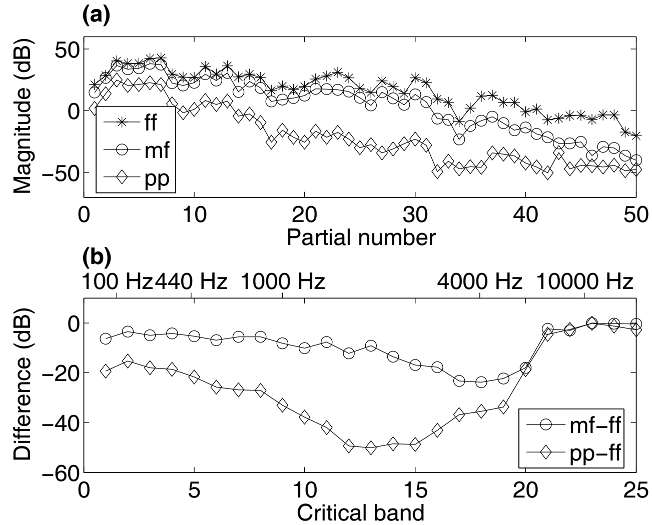

Fig. 2. (a) Partial magnitudes for key number 16 (C_2, f_0 = 65.1 Hz) corresponding to three different key velocities: forte (upper), mezzoforte (middle), and pianissimo (lower). (b) The attenuation magnitudes for the same key number at two velocities, mezzoforte and pianissimo shown on an auditory scale [11]. Magnitudes are obtained from recorded tones (samples from University of Iowa Electronic Music Studios, http://theremin.music.uiowa.edu).

In a real-time implementation, the computational load can be reduced by using pre-calculated source signals stored in the memory (the signals can be windowed with the shaping window before storing into the memory), because there is no need to change the partial amplitudes in real time. Moreover, if the user wants to modify the inharmonicity coefficient, which is not very common in normal use, the source signals can be recalculated. Hence, the proposed method is suitable for real-time applications.

2.2. Velocity-controlled equalizing filter

When the key velocity is increased, two phenomena take place: the produced tone generally gets louder, and some of the higher frequencies become more audible, thus increasing brightness. This can be seen in Figure 2, which shows an example of the same key played at three different velocities. An interesting phenomenon is seen in Figure 2(b), which depicts the attenuation of the mezzoforte and pianissimo tones compared to the forte tone on an auditory scale [11], as the attenuation is strongest at the critical bands 10-16, which correspond approximately to frequencies 1000-3000 Hz. The overall noise in the signal has an effect on this, but it does not provide a comprehensive explanation. Hence, we suggest that it should be taken into account in the model.

The one-pole lowpass filter is unsuitable for filtering the additive source signal when we use the partial amplitudes obtained from forte tones, because it is unable to

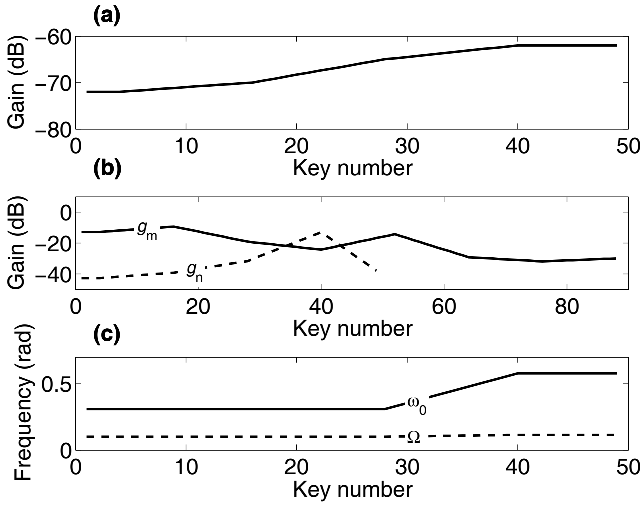

Fig. 3. (a) The one-pole lowpass filter parameter $G_{\min}$. (b) The minimum values for the equalizing filter notch gain g_n, and for the overall equalizing filter gain g_m. (c) The equalizing filter notch center frequency ω_0, and notch bandwidth Ω.

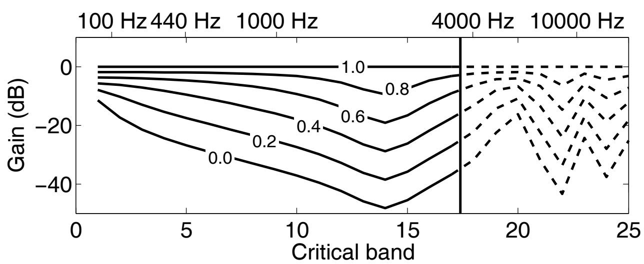

Fig. 4. The magnitude response on an auditory scale of the equalizing filter for key number 16 (f_0 = 65.4 Hz) with key velocities 0.0 (pianissimo), 0.2, 0.4, 0.6, 0.8, and 1.0 (forte). The vertical line denotes the cutoff frequency 3675 Hz.

simulate the attenuation phenomena seen in Figure 2(b). Another solution using a second-order LPC-based method was presented in [12]. In this work, a second-order equalizing filter was chosen to simulate the velocity effect. The transfer function of the equalizing filter [13] is $H_{EQ}(z) = c[1/2 + d/2 - (d/2 - 1/2)A(z)]$, where

$$A(z) = \frac{a - \cos(\omega_0)(1+a)z^{-1} + z^{-2}}{1 - \cos(\omega_0)(1+a)z^{-1} + az^{-2}}, \quad (1)$$

$a = [d - \tan(\Omega/2)]/[d + \tan(\Omega/2)]$, $c = 10^{(g_m(1-v))/20}$, $d = 10^{(g_n(1-v))/20}$, g_n is the minimum gain of the notch, g_m is the minimum overall gain, ω_0 is the center frequency in radians, and Ω is the normalized -3dB bandwidth in radians. Since the bandwidth of the tones produced by the additive synthesis is limited, a multi-rate approach can be used to ease the filter design. The sampling frequencies used in the filter design were chosen to be 11025 Hz for key numbers 1-35 (upsampling factor 4), and 14700 Hz for key numbers 36-49 (upsampling factor 3).

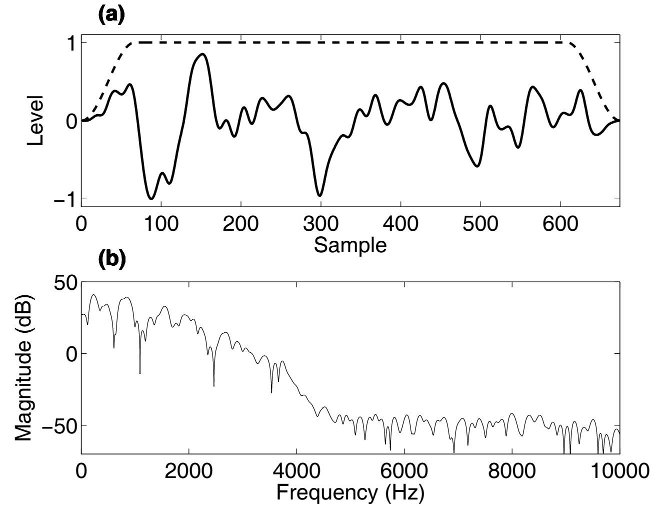

Fig. 5. An excitation signal (a) waveform (solid), shaping window function (dashed), and (b) spectrum for key number 16 (f_0 = 65.4 Hz, inharmonicity coefficient = 0.000173, forte) produced by the proposed method. The signal includes 48 partials produced with additive synthesis and highpass-filtered noise with cutoff frequency 3675 Hz. The period length is 674.2 samples.

The determined filter parameters are shown in Figure 3 as a function of key number. Since the attenuation notch is not visible with high key numbers, the equalizing filter can be replaced with the gain g_m for key numbers larger than 49. Figure 4 shows an example of how the filter magnitude response depends on the key velocity.

2.3. Pulse shaping

If the excitation signal were fed to the string model without any kind of windowing in the time domain, it would cause undesirable high-frequency components in the produced tone. In order to avoid this, the signal is multiplied with a shaping window generator output, which corresponds to the desired envelope of the excitation signal. Moreover, the windowing should be done before filtering with the equalizing filter (or the one-pole filter for the noise signal), because otherwise the sidelobes of the short signal can cause problems.

In this work, a Hanning window of length $L/5$ is used as the windowing function, where L is the fundamental frequency period length in samples. Instead of storing the entire Hanning window, memory can be saved by storing only the first half of the window and using its mirror image for the end of the signal. Hence, the transfer function of the windowing function is defined as:

$$H_w(z) = \sum_{i=0}^{L/10-1} w_{i+1} z^{-i} + \sum_{i=L/10}^{L-L/10-1} z^{-i} + \sum_{i=L-L/10}^{L-1} w_{L-i} z^{-i}, \quad (2)$$

where w is the $L/5$ samples long Hanning window.

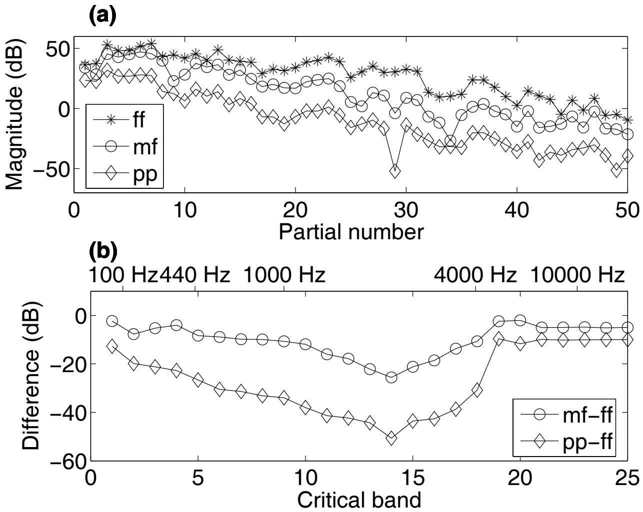

Fig. 6. (a) The spectrum of a piano tone produced by a synthesis model including the proposed method at different velocity levels (key number 16, C_2). (b) The magnitude differences compared to forte for the same key number at two velocities, mezzoforte and pianissimo shown on an auditory scale.

3. RESULTS

The proposed method was implemented in our waveguide piano model [14] by using Matlab [15]. An example of an excitation signal produced by the proposed method can be seen in Figure 5. Figure 6 shows an example of the produced partial amplitudes at three different key velocities. Comparison against Fig. 2 demonstrates that the proposed method is able to excite the string model, producing realistic partial amplitudes. Moreover, the method simulates the key velocity effect well. Sound examples are available at http://www.acoustics.hut.fi/demos/param-exc/.

4. CONCLUSION

In this paper, a new method for providing an excitation signal for the waveguide piano model is presented. The method is suitable for real-time piano synthesis. First, it enables control over individual partial amplitudes and frequencies. Second, it can produce perceptually realistic piano tones, without any undesirable elements in the tone caused by the excitation signal. Third, the method provides a velocity-controlled excitation signal. Finally, the parameters are easy to obtain from recorded piano tones.

5. ACKNOWLEDGMENTS

This work was supported by the Academy of Finland (projects no. 53537 and no. 104934).

6. REFERENCES

[1] J.O. Smith III, “Physical Modeling Using Digital Waveguides,” *Computer Music Journal*, vol. 16, no. 4, pp. 74-91, 1992.

[2] S.A. Van Duyne, J.R. Pierce, and J.O. Smith III, “Traveling Wave Implementation of a Lossless Mode-Coupling Filter and the Wave Digital Hammer,” in *Proc. Int. Computer Music Conf.*, Århus, Denmark, pp. 411-418, September 1994.

[3] G. Borin, G. De Poli, and D. Rocchesso, “Elimination of Delay-free Loops in Discrete-Time Models of Nonlinear Acoustic Systems,” *IEEE Trans. on Speech and Audio Processing*, vol. 8, no. 5, pp. 597-605, September 2000.

[4] B. Bank, “Nonlinear Interaction in the Digital Waveguide with the Application to Piano Sound,” in *Proc. Int. Computer Music Conf.*, Berlin, Germany, pp. 54-57, September 2000.

[5] J.O. Smith III and S.A. Van Duyne, “Commuted Piano Synthesis,” in *Proc. Int. Computer Music Conf.*, Banff, Canada, pp. 335-342, September 1995.

[6] M. Karjalainen and V. Välimäki, “Model-based Analysis/Synthesis of the Acoustic Guitar,” in *Proc. Stockholm Music Acoustics Conf.*, Stockholm, Sweden, pp. 443-447, 1993.

[7] V. Välimäki, H. Penttinen, J. Knif, M. Laurson, and C. Erkut, “Sound Synthesis of the Harpsichord Using a Computationally Efficient Physical Model,” *EURASIP J. App. Sig. Proc.,* vol. 4, no. 7, pp. 934-948, 2004.

[8] J. Bensa, K. Jensen, and R. Kronland-Martinet, “A Hybrid Resynthesis Model for Hammer-String Interaction of Piano Tones,” *EURASIP J. App. Sig. Proc.*, vol. 4, no. 7, pp. 1021-1035, 2004.

[9] P.A.A. Esquef, V. Välimäki, and M. Karjalainen “Restoration and Enhancement of Solo Guitar Recordings Based on Sound Source Modeling,” *J. Audio Eng. Soc.*, vol. 50, no. 4, pp. 227-236, 2002.

[10] D.A. Jaffe and J.O. Smith III, “Extensions of the Karplus-Strong Plucked-String Algorithm,” *Computer Music J.*, vol. 7, no. 2, pp. 56-69, 1983.

[11] Zwicker E. and R. Feldtkeller, *The Ear as a Communication Receiver*, Acoustical Society of America, 1999.

[12] M. Laurson, C. Erkut, V. Välimäki, and M. Kuuskankare, “Methods for Modeling Realistic Playing in Acoustic Guitar Synthesis,” *Computer Music J.*, vol. 25, no. 3, pp. 38-49, 2001.

[13] Zölzer U., *Digital Signal Processing*, Wiley & Sons, 1997.

[14] J. Rauhala, H.-M. Lehtonen, and V. Välimäki ”Multi-ripple Loss Filter for Waveguide Piano Synthesis,” in *Proc.Int. Computer Music Conf.*, Barcelona, Spain, pp. 729-732, September 2005.

[15] Matlab version 7.0, Mathworks Inc. 1984-2004.