Jukka Rauhala, Heidi-Maria Lehtonen, and Vesa Välimäki

Toward Next-Generation Digital Keyboard Instruments

Generating natural-sounding instrument tones algorithmically without a large sample database

Acoustic keyboard instruments, such as the piano and the harpsichord, are particularly interesting for sound synthesis, because they are large in size and are prone to wear and tear. The high cost of acoustic keyboard instruments makes modeling them financially attractive. They may require amplification during performances, which causes difficulties related to the use of microphones. Digital versions of these instruments benefit from the fact that Musical Instrument Digital Interface (MIDI) keyboard controllers are commonly available. Digital pianos imitating grand pianos are currently among the most popular electronic musical instruments.

Commercial products implementing synthesis of keyboard instrument sounds are often based on sampling and wavetable synthesis techniques. These methods employ recorded acoustic waveforms of instrument sounds. In extreme cases, tones of all the keys of the instrument are sampled at several velocity levels to cover the whole dynamic range, and these samples are as long as necessary, even about a minute each for the low piano tones. Very high sound quality can be achieved, but it is dependent on the size of the sample memory. Limitations in memory reduce the obtainable quality, because samples must be shortened or their bit rate must be compressed.

In this article, alternative approaches to digital keyboard instrument synthesis are looked into. Physics-based sound synthesis, which aims at generating natural-sounding musical instrument tones algorithmically without using a large sample database, is a promising approach. It would provide high-quality

1053-5888/07/$25.00©2007IEEE

music synthesis to systems that cannot afford a large memory, such as mobile phones and portable electronic games. The realistic parametric synthesis of musical instrument sounds is still a challenge, but physical modeling techniques introduced during the last few decades can help to solve it [1], [2]. Recently, the first commercial products have been introduced, for example, by Pianoteq [3]. Three keyboard instruments, the clavichord, the harpsichord, and the grand piano, are focused on here. The sound production principles and acoustics of these instruments are first discussed. Then, the previous parametric synthesis algorithms developed for these instruments are reviewed. The remaining part of this article concentrates on new signal processing methods for parametric synthesis of the piano.

ACOUSTICS OF KEYBOARD INSTRUMENTS

Before considering the modeling of keyboard instruments, the acoustics of the keyboard instruments are presented.

CLAVICHORD

The clavichord is a struck string instrument with a long history dating back to the 12th century. The sound of a clavichord is soft, and it was used mainly as a solo instrument. During the 18th century it was totally dwarfed by the piano, which was able to produce a louder sound. The playing range of a clavichord is usually four octaves covering the notes from C_2 to C_6. A typical clavichord has a small, rectangular box, and the strings are organized in pairs. The strings in each pair are slightly detuned, which causes beating and a two-stage decay [4]. The strings are positioned in the long direction of the box, and they are terminated with hitch pins at one end and with tuning pins at the other end.

When the player presses down a key, a metal tangent, attached to the other end of the key, strikes against a pair of strings. Actually, this mechanism does not only set the strings into vibration but also defines the speaking length of the strings to be the length between the tangent and the bridge. The other part of the strings, that is, the stretch from the hitch pin to the tangent, is damped by a piece of felt. With this mechanism, it is possible to use a single pair of strings for producing several, usually two or three, notes.

The advantage of the clavichord is its expressiveness. The tone can be modified in loudness as well as in terms of vibrato. When the tangent still remains in contact with the string after depressing the key, the player can vary the tension of the string by varying the pressure on a key, and thus execute a pitch vibrato. Figure 1 shows the structure of the string register and the keyboard of a clavichord that was used in [5]. A more exhaustive overview of the acoustics of the clavichord is given in [6].

HARPSICHORD

The harpsichord is known especially from the baroque era, when it was used both as a solo instrument and as an important part of the baroque orchestra. It was widely used until the 19th century, when it was eclipsed by the piano, as the dynamical range of the harpsichord was not enough to compete with the loudness of a large symphony orchestra. However, its dynamical output is still greater than that of a clavichord.

The playing range of a harpsichord covers four to five octaves, depending on the instrument. The shape of the instrument is triangular, and one of its sides is curved. The soundboard is made of thin wood, such as spruce, and it is stiffened by light ribs. A harpsichord may have one to four sets of strings, which are usually called choirs or registers. The basic set is called the 8′ (8 foot) register, borrowing from organ terminology. One of the string sets may be tuned an octave higher (4′) and one octave lower (16′). A harpsichord may have one or two manuals, which can be used to control different registers in order to vary the loudness and timbre of the instrument. When a player depresses a key, a jack is raised, which, in turn, causes a plectrum to pluck the string. Unlike in the clavichord and the piano, the strings are single. Figure 2 shows the manuals and the string register of the harpsichord that was used in [7]. A more comprehensive description of the acoustics of the harpsichord is given in [6].

[FIG1] A clavichord showing the tangents that hit the string groups, and the damping felts are visible under the strings. The hitch pins and the tuning pins to which the strings are attached are visible on the left and right side of the strings, respectively.

[FIG2] A harpsichord being tuned. This instrument has two manuals and three sets of string choirs.

PIANO

The piano, which is nowadays probably the most popular musical instrument, has a more complex structure and sound than its predecessors. It has a wide dynamic range, and its playing range is more than seven octaves. The roots of the modern piano go back to the beginning of the 18th century, when Bartolomeo Christofori of Florence modified the harpsichord by replacing the jacks with hammers. He called the new instrument the "gravicembalo col piano et forte," because it was capable of dynamical variations in tone [6]. During the last 3 centuries, the instrument has evolved into two distinct instruments: the grand piano and the upright piano. This article concentrates on the acoustics and modeling of the grand piano, but the same principles are applicable to the upright piano as well. A grand piano, opened for its structure to be seen, is shown in Figure 3.

THE PIANO HAS BEEN CONSIDERED A PARTICULARLY INTERESTING INSTRUMENT FROM THE MODELING POINT OF VIEW, SINCE IT IS A PROMINENT INSTRUMENT IN WESTERN MUSIC AND IT HAS A COMPLEX STRUCTURE.

The grand piano consists of five main parts: the keyboard, the action, the strings, the soundboard, and the frame. From the keystroke the information message is transmitted to the action, which controls the hammer. The hammer hits the string and sets it into vibration. The kinetic energy of the hammer is transformed into vibrational energy, which is stored in the normal modes of the string. This energy is transmitted to the soundboard, the main radiating part of the piano, via the bridge. The soundboard is a thin, wooden plate positioned under the frame. The cast-iron frame, positioned at the upper part of the wooden case, keeps the instrument together, and it is designed to withstand the high tension of the strings. The strings are attached to the tuning pins at the player end and to the hitch-pin rail at the other end. The speaking length of the string, however, is restricted to the bridge.

[FIG3] The keyboard and the string register of a grand piano. The cast iron frame encases the string register, under which is the soundboard. The contrast between the lengths of the massive bass and thinner treble strings is clearly visible.

The most important characteristics of the piano sound are inharmonicity, complicated decay, and beating. A concert grand piano has 243 steel strings. The lowest strings are long and massive (length of even 2 m) while the strings corresponding to the highest keys are thin and short (approximately 5 cm). The first eight strings are single strings and the rest of the strings, corresponding to the 80 highest keys, are in groups of two or three strings, depending on the instrument.

The inharmonicity in the piano strings is caused by stiffness. It makes the higher partials travel faster in the piano string, which means that their frequencies are a little higher compared to those of an ideal string. In the spectrum, the partial components are slightly shifted making the series of overtones "stretch" upward. Usually, strong inharmonicity is considered to be an undesired feature of the piano sound. On the other hand, it is not desired to get totally rid of this inharmonicity, since a slight inharmonicity adds warmth to the sound [8].

The decay process of a piano tone is very complicated. The decay rate is two-fold; the tone begins to decay fast, but after a few seconds the decay rate changes and becomes slower. This is due, among other factors, to the change in the predominant vibration of the strings. The vertical (perpendicular to the soundboard) vibration decays rapidly whereas the horizontal (parallel to the soundboard) vibration decays slowly [4]. In addition, the partials decay at different rates. Some of them may sound even dozens of seconds whereas others decay in a few seconds. The spectrum varies over time and differs from key to key; at the bass end over 50 partials can be extracted while at the treble end the corresponding number is only about 3 or 4.

Another important phenomenon in the piano sound is beating, which results from the vibration of unison groups of strings. When the hammer excites a tricord, that is, a set of three unison strings, the strings begin to vibrate in the same phase. Due to small differences in frequency between the strings in the tricord, the tone starts to beat soon. The strings can be considered by no means as independent, since they are coupled to the bridge. This coupling allows energy leakage between the strings resulting in a highly complicated system.

PHYSICS-BASED SYNTHESIS ALGORITHMS FOR KEYBOARD INSTRUMENTS

The earlier physics-based synthesis models for the clavichord, the harpsichord, and the piano are briefly reviewed in the following sections.

CLAVICHORD

A physics-based synthesis model for the clavichord has been previously developed by Välimäki et al. [5]. The commuted waveguide synthesis method [9], [10] is applied by using inverse-filtered recorded clavichord tones as excitation for the synthesis model. Inverse filtering here refers to the processing of a signal with the inverted transfer function of a waveguide string model [12]. In this case, the inverse filtering essentially cancels the partials of a recorded tone.

The structure of the clavichord synthesis model is shown in Figure 4. The excitation signals, which are truncated to a length of about 0.5 s, are stored in a database from which they are retrieved as input signals for the synthesizer. One such excitation signal is used for each key. The effect of coupling of the two basic string models $S_1(z)$ and $S_2(z)$ is simulated with the unconditionally stable technique suggested in [11]: the output of only one of the string models is fed to the input of the other and hence there is no feedback and there can be no stability problems. In practice, the coupling coefficient g_c is selected to have a small value.

Two additional sample databases are needed for realistic reproduction of the soundbox response and the percussive noise caused by key release. The reverberation caused by the soundbox is incorporated in a simplified way by triggering a soundbox response sample at a low level each time any note is played. This sample must be at least 5 s long so that it provides the reverberant character of the clavichord. This is particularly important for short notes, such as staccato playing, for which the output signal would otherwise stop suddenly in an unnatural manner.

HARPSICHORD

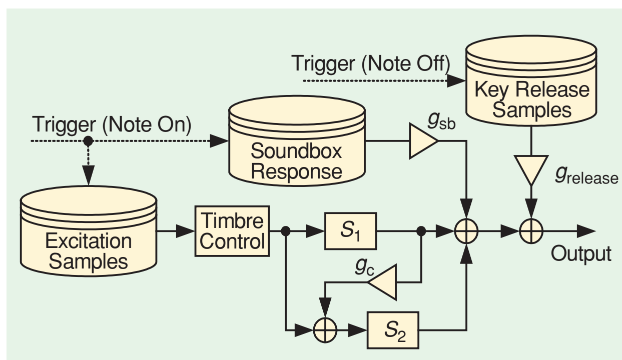
Figure 5 shows the block diagram of the harpsichord synthesis model developed by Välimäki et al. [7]. The algorithm structure has been modified from the clavichord synthesizer discussed above. A version of the commuted waveguide synthesis approach is used, where each tone is generated with a parallel combination of the string model $S(z)$ and a second-order resonator $R(z)$ that are excited with a common excitation signal. The second-order resonator, previously proposed for this purpose by Bank [13], approximately simulates the beating effect appearing in many harpsichord tones. In this approach, a resonator is slightly detuned compared to the beating partial, which produces a perceptually realistic beating effect.

A modification to the loss filter of the waveguide string model $S(z)$ was introduced in the harpsichord synthesizer [7]. It allows more flexible control of decay rates of partials than is possible with a one-pole digital filter, which is a usual choice for the loss filter. The characteristic key-release thump terminating harpsichord tones is reproduced by triggering a sample that has been extracted from a recording. A digital filter model for the soundboard has been designed based on recorded bridge impulse responses of the harpsichord. The output of the string models is injected into the soundboard filter that imitates the reverberant nature of the soundbox and the ringing of the short parts of the strings behind the bridge.

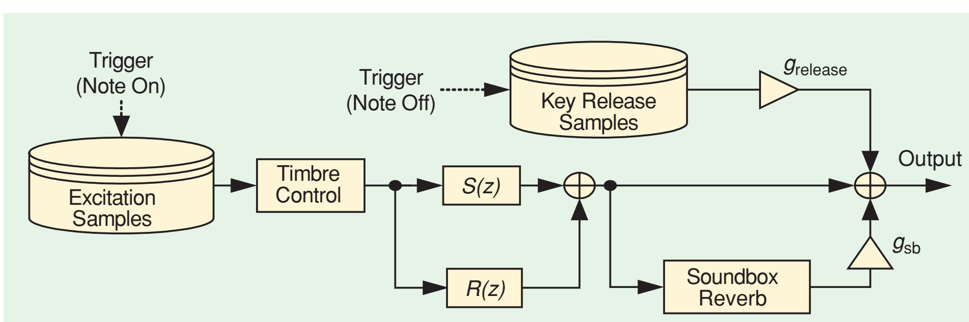
DUCASSE RECENTLY SHOWED THAT DETAILED WAVEGUIDE MODELING OF DISPERSIVE PIANO STRINGS IS A MORE COMPLICATED TASK THAN THOUGHT PREVIOUSLY.

PIANO

The piano has been considered a particularly interesting instrument from the modeling point of view, since it is a prominent instrument in Western music and it has a complex structure. The main components that have to be taken into account in the modeling process are the string, the hammer, the soundboard, and the pedals, especially the sustaining pedal, which is used in every professional piano performance. Several physics-based sound synthesis models for the piano have been proposed (see, e.g., [13]–[16]). This article emphasizes developing further the piano synthesis model based on the digital waveguide technique [1], [2], which the authors consider the most appropriate for this application. Nevertheless, Ducasse recently showed that detailed waveguide modeling of dispersive piano strings is a more complicated task than thought previously [17].



[FIG4] **Block diagram of the clavichord synthesis algorithm for one key (adapted from [5]). The sample databases are common for all keys.**



[FIG5] **Block diagram of the harpsichord synthesis algorithm for a single string (adapted from [7]). The release sample database and the soundbox reverb are common to all strings of all keys.**

The string model contains four major parts: the delay line, the tuning filter, the loss filter, and the dispersion filter. The integer delay line determines the pitch of the tone, and the tuning filter, which is traditionally designed as an all-pass fractional delay filter, is used to fine-tune the pitch in those cases where the length of the delay line corresponding to the desired pitch is not an integer. The loss filter, which models the complicated, frequency-dependent decay of the partials, can be designed as a lowpass filter [15] or as a multiripple filter, where several feedforward paths are added in cascade with a one-pole filter [18], [19].

The dispersion phenomenon is usually modeled with a cascade of low-order all-pass filter sections. As the target is to make the higher partials stretch upward, a filter with a proper phase response is needed. The design methods can employ standard or, in some cases, custom-made filter design techniques. An excellent overview of design methods is given in [20].

The interaction between the hammer and the string is highly nonlinear due to the felt covering the hammer. Although the hammer-string interaction can be described with a simple formula (see, e.g., [20]), there is a mutual dependence between the hammer position and the interaction force; the hammer position should be known before computing the interaction force and vice versa. However, the implicit relation between the hammer position and the interaction force can be made explicit by inserting a fictitious delay element in the model. This kind of approach is widely used in literature, e.g., [21] and [22], although it can be a possible source of instability. Van Duyne and Smith [23] described the problem in terms of wave variables and presented the wave digital hammer model, which is based on the theory of wave digital filters. By appropriately choosing the model parameters, they were able to avoid the fictitious delay element in the model. Borin et al. presented a method called the "K method," which maps the interaction force as a function of the linear combination of the past values of the string and hammer positions as well as the interaction force [24]. The advantage is that the instantaneous dependencies of the variables are dropped. Bank introduced a multirate hammer model that overcomes the stability problem by doubling the sample rate in order to achieve smaller changes in the variables of interest [13]. Smith and Van Duyne [14] came up with the idea that the hammer-string interaction consists of a few discrete events during the hammer strike. These events can be approximated with one or more impulses that are lowpass filtered. In a more recent study, Bensa et al. presented a source-resonator model for the hammer-string interaction, where the resonator is modeled as a digital waveguide and the source is modeled using a subtractive signal model [25].

ONE GREAT ADVANTAGE OF THE PHYSICS-BASED MODELING TECHNIQUE COMPARED TO THE SAMPLING TECHNIQUE IS THAT IT OFFERS WAYS TO CONTROL THE SYNTHESIS MODEL WITH PHYSICAL AND NONPHYSICAL PARAMETERS.

The modeling of the soundboard is often implemented with a reverberation algorithm. This kind of approach was taken, among others, by Bank [13]. In addition to the soundboard, the sustaining pedal can be modeled with a reverberation algorithm [26]. Generally, the sustaining pedal has an effect on the sound in two ways; it increases the beating present in the tone and enriches the sound, and, thus, a properly designed reverberation algorithm is suitable for the simulation task.

[TABLE 1] TYPICAL LOW-LEVEL PARAMETERS FOR A SINGLE KEYBOARD INSTRUMENT STRING MODEL.

FUNDAMENTAL FREQUENCY
INHARMONICITY COEFFICIENT VALUE
PARTIAL AMPLITUDES
KEY VELOCITY GAIN
PARTIAL DECAY RATES
PARTIALS WITH BEATING EFFECT
PARTIAL BEATING EFFECT FREQUENCIES
PARTIAL BEATING EFFECT AMPLITUDES
PHANTOM GAIN

TUNABLE KEYBOARD INSTRUMENT SYNTHESIS MODEL

One great advantage of the physics-based modeling technique in comparison to the sampling technique is that it offers ways to control the synthesis model with parameters, both physical and nonphysical. The latter part of this article considers how the parameterization can be taken into account in the string model and the excitation model of a piano synthesis model, which can be applied to other keyboard instruments as well.

TUNABLE PARAMETERS FOR KEYBOARD INSTRUMENT SYNTHESIS

The synthesis parameters can be separated into two categories from the synthesis model point of view: low-level and high-level parameters. The low-level parameters are directly used for designing the synthesis model, whereas the high-level parameters, such as the size of the piano frame, require complex rules in order to be mapped onto the design parameters. This article focuses on the low-level parameters.

A list of the low-level keyboard instrument synthesis parameters for a single string is given in Table 1. When these parameters are considered in the design of the synthesis blocks, it is observed that the fundamental frequency and the inharmonicity coefficient value must be taken into account in every block, whereas the other parameters are related to only one synthesis block.

The fundamental frequency is practically equivalent to the frequency of the first partial, which determines the perceived pitch of a single string. The real-time implementation of this parameter allows the synthesizer to be tuned in real-time, string by string—similar to the real instrument. Dispersion is an important phenomenon in the piano that is represented usually with the inharmonicity coefficient value B [6]. In the piano tones, dispersion is audible particularly in the bass range making the timbre of the tones warmer and richer [6]. It is interesting to note that if the material parameters are excluded, the fundamental frequency and dispersion parameters are not completely independent of each other, because both depend on the tension of the string.

THE CONTROL OVER THE FUNDAMENTAL FREQUENCY AND INHARMONICITY COEFFICIENT PARAMETERS MUST BE TAKEN INTO ACCOUNT IN THE DESIGN OF THE SYNTHESIS MODEL.

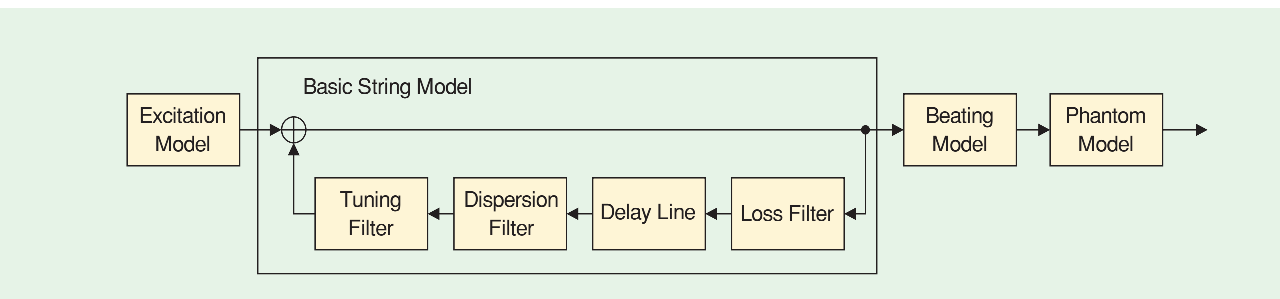
CONSIDERATIONS IN THE SYNTHESIS MODEL

The control over the fundamental frequency and inharmonicity coefficient parameters must be taken into account in the design of the synthesis model. A diagram of the keyboard instrument synthesis model is shown in Figure 6. A summary of requirements for each block is presented in Table 2.

The delay line and tuning filter blocks can be easily parameterized, whereas the parameterization of the loss filter is quite difficult; none of the current implementations offer closed-form design formulas except for a first-order filter [13]. However, it is unclear whether the loss filter coefficients should be modified when the fundamental frequency or dispersion is changed, since the human ear is insensitive to small changes in the partial decay times [27]. On the other hand, the excitation model, the dispersion filter, the phantom model, and the beating model are blocks that require more effort in the design phase in order to be parameterized.

EXCITATION MODEL

The main purpose of the excitation model is to produce energy distributed through the partial amplitudes of the partial frequencies that forces the waveguide string model to resonate. The excitation model should be flexible enough to be controlled via the fundamental frequency parameter and the dispersion parameter, which is not a trivial task. For instance, the traditional commuted waveguide synthesis model uses excitation signals that are inverse-filtered from real instrument tones. Hence, the real-time modification of the fundamental frequency parameter or the dispersion parameter is practically impossible without affecting the partial amplitude levels.

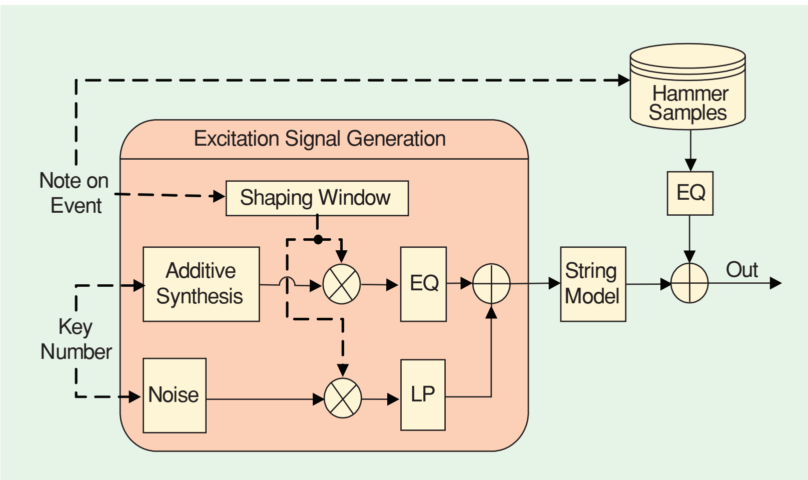


[FIG6] An advanced waveguide keyboard instrument synthesis model. The tuning filter and the dispersion filter are usually all-pass filters, whereas the loss filter is a finite input response/infinite input response filter. Excitation, beating, and phantom models are described in the text.

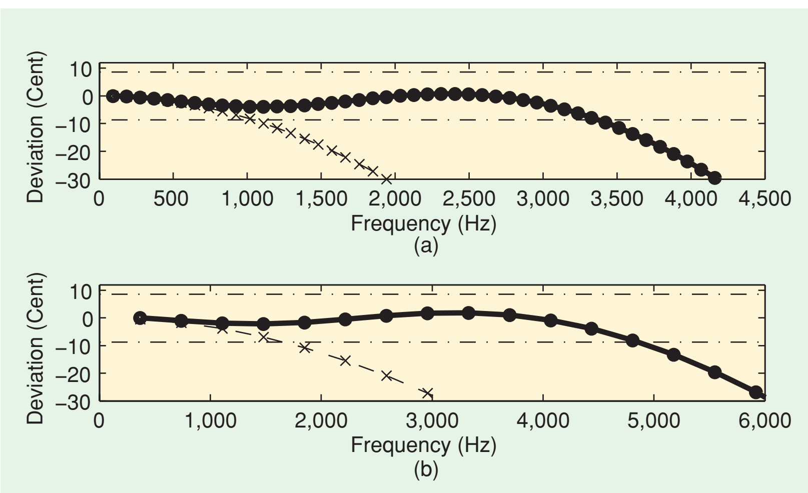
[TABLE 2] AN OVERVIEW OF THE REQUIREMENTS FOR THE KEYBOARD INSTRUMENT SYNTHESIS BLOCKS IN ORDER TO PROVIDE FUNDAMENTAL FREQUENCY AND DISPERSION REAL-TIME PARAMETERIZATION.

SYNTHESIS BLOCK	PRIMARY FUNCTION	REQUIREMENTS RELATED TO f_0 AND B PARAMETERS
EXCITATION MODEL	SET INITIAL LEVEL OF PARTIALS	THE PARTIAL FREQUENCIES MUST DEPEND ON f_0 AND B
LOSS FILTER	DETERMINE THE DECAY RATE OF PARTIALS	THE LOSS FILTER SHOULD HAVE THE DESIRED MAGNITUDE RESPONSE AT THE PARTIAL FREQUENCIES DEPENDING ON THE VALUES OF f_0 AND B.
DELAY LINE	TUNING OF THE PITCH	THE DELAY LINE LENGTH DEPENDS ON f_0. MOREOVER, THE EXTRA PHASE DELAY AT THE FUNDAMENTAL FREQUENCY PRODUCED BY THE DISPERSION FILTER MUST BE ACCOUNTED FOR.
DISPERSION FILTER	SET THE INHARMONICITY	THE FILTER MUST PRODUCE A PHASE DELAY RESPONSE ACCORDING TO THE DESIRED f_0 AND B VALUES.
TUNING FILTER	FINE-TUNING OF THE PITCH	THE REQUIRED FRACTIONAL DELAY TO BE PRODUCED BY THE TUNING FILTER DEPENDS ON f_0 AND ON THE EXTRA PHASE DELAY AT f_0 PRODUCED BY THE DISPERSION FILTER.
BEATING MODEL	INTRODUCE BEATING	THE PARTIAL FREQUENCIES MUST DEPEND ON f_0 AND B. MOREOVER, THE POSSIBLE INEXACTNESS OF THE PHASE DELAY RESPONSE OF THE DISPERSION FILTER MUST BE ACCOUNTED FOR.
PHANTOM MODEL	INTRODUCE NEW PARTIALS FOR FORTISSIMO TONES	B VALUE OF THE PHANTOM PARTIAL FREQUENCIES MUST DEPEND ON THE DISPERSION OF THE STRING. ADDITIONALLY, THE PHANTOM PARTIAL FREQUENCIES MUST DEPEND ON f_0.

In [28], this problem is solved by combining additive and subtractive syntheses to create a parameterized model. The block diagram of the excitation model is seen in Figure 7. The model consists of five blocks: additive synthesis block, noise generator block, equalizing filter block, one-pole filter block, and shaping window block. The source signal is generated with additive synthesis for the frequencies below the specific cutoff frequency, which depends on the key number, and with the noise generator for the high frequencies. The source signals are then filtered either with the equalizing filter (additive source signal) or with the one-pole filter (noise source signal), which depend on the key press velocity and, therefore, add dynamics to the excitation signal. Finally, the combined signal is windowed in order to avoid extra components in the resulting signal.



[FIG7] The excitation model.



[FIG8] Deviation of the partial frequencies for the proposed filter (solid line with circles at the partial frequencies) compared to the partial frequencies of an inharmonic tone, when (a) the inharmonicity coefficient value $B = 0.000080$, filter phase delay at dc $D = 14.57$, number of filters in cascade $M = 4$, delay line length $L_1 = 417$, tuning filter phase delay at dc $d_t = 1.36$, $f_0 = 92.5$ Hz (key F#2) and (b) $B = 0.00015$, $D = 10.91$, $M = 1$, $L_1 = 107$, $d_t = 1.25$, $f_0 = 370.0$ Hz (key F#4). The dashed line with crosses is the calculated deviation for a harmonic tone.

DISPERSION FILTER

In waveguide synthesis, the dispersion phenomenon is modeled with an all-pass filter inserted into the string model. The filter tries to produce an accurate phase-delay response according to the desired inharmonicity coefficient value. A common way to design the dispersion filter is to use a conventional filter design method. However, as these methods are computationally heavy, the filter cannot be controlled with the desired parameters in real time.

One solution to this problem is the tunable dispersion filter [29] that offers a closed-form formula to design the dispersion filter. The design method is based on the Thiran all-pass filter design method [30], which is commonly used for fractional delay filter design. The design of the tunable dispersion filter method is conducted in two phases. First, the delay value that produces a filter with the desired phase delay response according to the desired fundamental frequency and the desired inharmonicity coefficient value is approximated with closed-form formulas. Then, the actual filter coefficients are obtained by feeding the delay value into the Thiran formulas [30].

The tunable dispersion filter design includes two filters, a cascade of four second-order filters for the low fundamental frequencies and a single second-order filter for the high frequencies. Examples of the phase delay response errors produced with the filter are shown in Figure 8.

PHANTOM PARTIAL MODEL

Phantom partials [31] and longitudinal modes [32] are two phenomena occurring in fortissimo piano bass tones. It has been suggested that both phenomena are actually the same phenomenon and, hence, a single synthesis model is enough to model both of them [33]. The frequencies of the resulting spectral components depend on f_0 and B [33] (the inharmonicity coefficient value of the resulting partials is roughly one-fourth of the B of the string). Hence, the model for this phenomenon, denoted as the phantom model, should depend on these parameters.

Three different solutions for the phantom model in digital waveguide synthesis are proposed in [33] and [34]. The first solution is to filter the original signal produced by the string model with a low-Q comb filter and to add it to the original signal by using nonlinear mixing [34]. The other solutions use a parallel model (the phantom model shown in Figure 6 is a series model, as is the first solution) using either a second waveguide or a resonator bank [33]. All of these solutions can be controlled by f_0 and B in real-time. However, the model presented by Bensa and Daudet seems to have more advantages, as it is suggested to be simpler than the others [34].

BEATING MODEL

Another distinctive feature in piano tones is the beating phenomenon. It can be noticed especially on the low keys a couple of seconds after the attack of the tone. The reason for the phenomenon is the coupling of the piano strings. Moreover, it has been shown that even the lowest piano keys that do not have multiple strings produce this effect due to false coupling [35].

THIS ARTICLE PROPOSES A NEW AMPLITUDE MODULATION-BASED APPROACH FOR THE BEATING-EFFECT MODELING.

In the harpsichord model presented above, a resonator-based approach [13] is used for modeling the beating. However, one of the problems with this approach is that the exact frequencies have to be known in order to control the amplitude and the frequency of the beating. This may not be guaranteed, as the dispersion filter may not produce an accurate enough phase-delay response for all partials. Hence, it would require compensation and, thus, the phase delay response of the filter should be accurately parameterized.

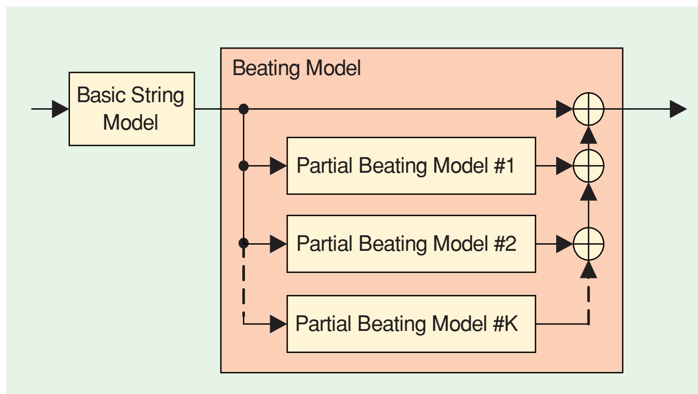
This article proposes a new amplitude modulation-based approach for the beating-effect modeling. The general principle of the method is shown in Figure 9. For each beating partial the model includes the partial beating model that is presented in Figure 10. The partial beating model includes two components, a bandpass filter H_{bp} and a modulator. First, the tone produced with the string model is filtered with H_{bp}, which has a peak at the estimated frequency of the desired beating partial. Then, the filtered signal is modulated with a low-frequency oscillator (LFO) block at the desired beat frequency. Finally, the modulated signal is added to the original signal. This can be done for all desired beating partials in parallel.

The model has three parameters: the notch frequency f_c of the bandpass filter H_{bp}, the beating frequency f_b, and the beating depth g_b. The bandpass filter is the filter presented in [36, pp. 126-129] with a large peak gain value K producing a very sharp peak. In order to compensate the notch gain, the filtered signal has to be multiplied by $g_c = g_b - K$. It should be noted that the beating depth value does not need to be accurate, as the beating effect is perceived as an on-off process [37]. In addition, the phase of the beating is not perceptually important, because the beating phenomenon is audible after a couple of seconds only.

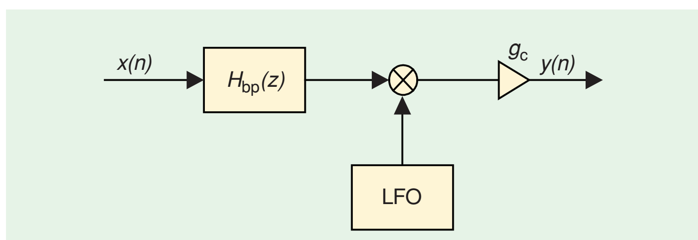
The beating model can be easily controlled with the fundamental frequency parameter and the dispersion parameter; when either one is changed, the equalizing filter parameters should be recalculated. An advantage of this approach is that it does not need accurate estimates for the partial frequencies, since the bandpass filter notch is not extremely sharp. Figure 11 shows envelopes of the second partials of three tones: a recorded piano tone, a synthesized tone without the beating model, and a synthesized tone with the beating model. It can be seen that the beating model is able to produce a realistic beating effect for a synthesized tone. Sound examples are available at the http://www.acoustics.hut.fi/demos/piano-beating/.

CONCLUSIONS

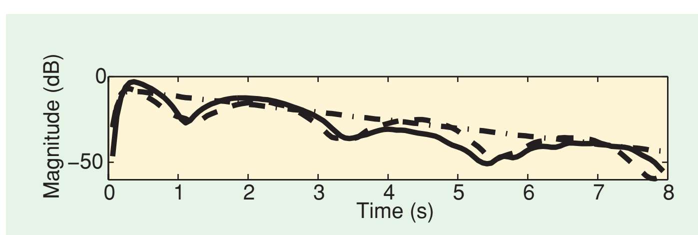
Physics-based modeling techniques enable the development of new kinds of digital keyboard instruments. This article gave a



[FIG9] The working principle of the beating model.



[FIG10] The partial beating model.



[FIG11] The envelope of the second partial (key E_2) of the original recorded piano tone (solid line), the synthesized tone without the beating model (dash-dotted line), and the synthesized tone with the beating model (dashed line). The beating model parameter values were $K = 100$ dB, $f_c = 0.2 f_0$, $f_b = 0.45$ Hz, and $g_c = -2$ dB. (Recorded tone obtained from University of Iowa Electronic Music Studios, http://theremin.music.uiowa.edu.)

short overview on how keyboard instrument sounds can be synthesized by using physics-based signal processing techniques. By using the proposed new solutions, it is possible to implement an advanced keyboard instrument synthesis string model that can be controlled via parameters, such as fundamental frequency and dispersion, in real time. Moreover, these techniques can be applied in synthesis of other keyboard instruments as well.

ACKNOWLEDGMENTS

This work was funded by the Academy of Finland (project no. 104934). Jukka Rauhala was supported by the Nokia Foundation and Helsinki University of Technology, and Heidi-Maria Lehtonen was supported by Tekniikan edistämissäätiö and the Emil Aaltonen Foundation.

AUTHORS

Jukka Rauhala (jukka.rauhala@acoustics.hut.fi) is a doctoral student at TKK (Helsinki University of Technology), Espoo, Finland. He is currently with the TKK Laboratory of Acoustics and Audio Signal Processing. He received his M.S. from TKK in 2003. His research interests include sound synthesis of musical instruments and audio signal processing.

Heidi-Maria Lehtonen (heidi-maria.lehtonen@tkk.fi) is a doctoral student at TKK (Helsinki University of Technology), Espoo, Finland. She has worked as a researcher at the TKK Laboratory of Acoustics and Audio Signal Processing since 2004. She received her M.S. in acoustics and audio signal processing from TKK in 2005. Her research interests include physics-based sound synthesis of musical instruments and musical acoustics.

Vesa Välimäki (vesa.valimaki@tkk.fi) is a professor at the Laboratory of Acoustics and Audio Signal Processing, TKK (Helsinki University of Technology), Espoo, Finland. He received the M.S., the Licentiate of Science, and the Doctor of Science degrees in technology, all in electrical engineering, from TKK in 1992, 1994, and 1995, respectively. His research interests include music synthesis, acoustics of musical instruments, and digital filter design. He is a Senior Member of the IEEE.

REFERENCES

[1] J.O. Smith III, "Physical modeling using digital waveguides," *Computer Music J.*, vol. 16, no. 4, pp. 74–91, 1992.

[2] V. Välimäki, J. Pakarinen, C. Erkut, and M. Karjalainen, "Discrete-time modelling of musical instruments," *Rep. Prog. Phys.*, vol. 69, no. 1, pp. 1–78, Jan. 2006.

[3] Pianoteq Web site [Online]. Available: http://www.pianoteq.com.

[4] G. Weinreich, "Coupled piano strings," *J. Acoust. Soc. Am.*, vol. 62, no. 6, pp. 1474–1484, 1977.

[5] V. Välimäki, M. Laurson, and C. Erkut, "Commuted waveguide synthesis of the clavichord," *Computer Music J.*, vol. 27, no. 1, pp. 71–82, 2003.

[6] N.H. Fletcher and T.D. Rossing, *The Physics of Musical Instruments*. New York: Springer-Verlag, 1991.

[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. Applied Signal Processing*, vol. 2004, no. 7, pp. 934–948, June 2004.

[8] H. Fletcher, E.D. Blackham, and R. Stratton, "The quality of piano tones," *J. Acoust. Soc. Am.*, vol. 34, no. 6, pp. 749–761, 1962.

[9] J.O. Smith, "Efficient synthesis of stringed musical instruments," in *Proc. Int. Computer Music Conf.*, Tokyo, Japan, pp. 64–71, 1993.

[10] M. Karjalainen, V. Välimäki, and Z. Jánosy, "Towards high-quality sound synthesis of the guitar and string instruments," in *Proc. Int. Computer Music Conf.*, Tokyo, Japan, pp. 56–63, 1993.

[11] M. Karjalainen, V. Välimäki, and T. Tolonen, "Plucked-string models: From the Karplus-Strong algorithm to digital waveguides and beyond," *Computer Music J.*, vol. 22, no. 3, pp. 17–32, 1998.

[12] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy, "Physical modeling of plucked string instruments with application to real-time sound synthesis," *J. Audio Eng. Soc.*, vol. 44, no. 5, pp. 331–353, May 1996.

[13] B. Bank, "Physics-based sound synthesis of the piano," M.Sc. thesis, Budapest Univ. of Technology and Economics, Budapest, Hungary, 2000, published as Report 54, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland. Available: http://www.acoustics.hut.fi/publications/.

[14] J.O. Smith III and S.A. Van Duyne, "Commuted piano synthesis," in *Proc. Int. Computer Music Conf.*, Banff, Canada, pp. 335–342, Sept. 1995.

[15] S.A. Van Duyne and J.O. Smith III, "Developments for the commuted piano," in *Proc. Int. Computer Music Conf.*, Banff, Canada, pp. 319–326, Sept. 1995.

[16] J. Bensa, "Analysis and synthesis of piano sounds using physical and signal models," Ph.D. dissertation, Université de la Méditérranée, Marseille, France, 2003. Available: http://www.lma.cnrs-mrs.fr/~bensa/.

[17] É. Ducasse, "On waveguide modeling of stiff piano strings," *J. Acoust. Soc. Am.*, vol. 118, no. 3, pp. 1776–1781, 2005.

[18] 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, Sept. 2005.

[19] H.-M. Lehtonen, J. Rauhala, and V. Välimäki, "Sparse multi-stage loss filter design for waveguide piano synthesis," in *Proc. 2005 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, pp. 331–334, 2005.

[20] B. Bank, F. Avanzini, G. Borin, G. De Poli, F. Fontana, and D. Rocchesso, "Physically informed signal processing methods for piano sound synthesis: A research overview," *EURASIP J. Applied Signal Processing*, vol. 2003, no. 10, pp. 941–952, 2003.

[21] A. Chaigne and A. Askenfelt, "Numerical simulations of piano strings. I. A physical model for a struck string using finite difference methods," *J. Acoust. Soc. Am.*, vol. 95, no 2, pp. 1112–1118, 1994.

[22] A. Chaigne and A. Askenfelt, "Numerical simulations of piano strings. II. Comparisons with measurements and systematic exploration of some hammer-string parameters," *J. Acoust. Soc. Am.*, vol. 95, no. 3, pp. 1631–1640, 1994.

[23] S. Van Duyne and J.O. Smith, "The wave digital hammer: A traveling wave model of the piano hammer and the felt mallet," in *Proc. 1994 Int. Computer Music Conf.*, Århus, Denmark, pp. 411–418, 1994.

[24] G. Borin, G. De Poli, and D. Rocchesso, "Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems," *IEEE Trans. Speech Audio Processing*, vol. 8, no. 5, pp. 597–605, Sept. 2000.

[25] 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.

[26] H.-M. Lehtonen, "Analysis and parametric synthesis of the piano sound," M.Sc. thesis, Helsinki Univ. of Technology, Espoo, Finland, Oct. 2005. Available: http://www.acoustics.hut.fi/publications/.

[27] H. Järveläinen and T. Tolonen, "Perceptual tolerances for the decaying parameters in string instrument synthesis," *J. Audio Eng. Soc.*, vol. 49, no. 11, pp. 1049–1059, 2001.

[28] J. Rauhala and V. Välimäki, "Parametric excitation model for waveguide piano synthesis," in *Proc. 2006 IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Toulouse, France, pp. 157–160, May 2006.

[29] J. Rauhala and V. Välimäki, "Tunable dispersion filter design for piano synthesis," *IEEE Signal Processing Lett.*, vol. 13, no. 5, pp. 253–256, May 2006.

[30] T.I. Laakso, V. Välimäki, M. Karjalainen, and U.K. Laine, "Splitting the unit delay—Tools for fractional delay filter design," *IEEE Signal Processing Mag.*, vol. 13, no. 1, pp. 30–60, Jan. 1996.

[31] H.A. Conklin, "Generation of partials due to nonlinear mixing in a stringed instrument," *J. Acoust. Soc. Am.*, vol. 105, no. 1, pp. 536–545, Jan. 1999.

[32] B. Bank and L. Sujbert, "Generation of longitudinal vibrations in piano strings: From physics to sound synthesis," *J. Acoust. Soc. Am.*, vol. 117, no. 4, pp. 2268–2278, Apr. 2005.

[33] B. Bank and L. Sujbert, "Modeling the longitudinal vibration of piano strings," in *Proc. Stockholm Music Acoustics Conf.*, Stockholm, Sweden, pp. 143–146, Aug. 2003.

[34] J. Bensa and L. Daudet, "Efficient modeling of "phantom" partials in piano tones," in *Proc. Int. Symp. Musical Acoustics*, Nara, Japan, pp. 207–210, 2004.

[35] B. Capleton, "False beats in coupled piano string unisons," *J. Acoust. Soc. Am.*, vol. 115, no. 2, pp. 885–892, Feb. 2004.

[36] U. Zölzer, *Digital Audio Signal Processing*. New York: Wiley, 1997.

[37] H. Järveläinen and M. Karjalainen, "Perception of beating and two-stage decay in dual-polarization string models," in *Proc. Int. Symp. Musical Acoustics*, Mexico City, 2002. Available: http://lib.tkk.fi/Diss/2003/isbn9512263149/article6.pdf, pp. 1–10.