

# Physically Informed Signal Processing Methods for Piano Sound Synthesis: A Research Overview

## Balázs Bank

*Department of Measurement and Information Systems, Faculty of Electronical Engineering and Informatics, Budapest University of Technology and Economics, H-111 Budapest, Hungary  
Email: bank@mit.bme.hu*

## Federico Avanzini

*Department of Information Engineering, University of Padova, 35131 Padua, Italy  
Email: avanzini@dei.unipd.it*

## Gianpaolo Borin

*Dipartimento di Informatica, University of Verona, 37134 Verona, Italy  
Email: borin@prn.it*

## Giovanni De Poli

*Department of Information Engineering, University of Padova, 35131 Padua, Italy  
Email: depoli@dei.unipd.it*

## Federico Fontana

*Department of Information Engineering, University of Padova, 35131 Padua, Italy  
Email: fontana@sci.univr.it*

## Davide Rocchesso

*Dipartimento di Informatica, University of Verona, 37134 Verona, Italy  
Email: rocchesso@sci.univr.it*

*Received 31 May 2002 and in revised form 6 March 2003*

This paper reviews recent developments in physics-based synthesis of piano. The paper considers the main components of the instrument, that is, the hammer, the string, and the soundboard. Modeling techniques are discussed for each of these elements, together with implementation strategies. Attention is focused on numerical issues, and each implementation technique is described in light of its efficiency and accuracy properties. As the structured audio coding approach is gaining popularity, the authors argue that the physical modeling approach will have relevant applications in the field of multimedia communication.

**Keywords and phrases:** sound synthesis, audio signal processing, structured audio, physical modeling, digital waveguide, piano.

## 1. INTRODUCTION

Sounds produced by acoustic musical instruments can be described at the *signal* level, where only the time evolution of the acoustic pressure is considered and no assumptions on the generation mechanism are made. Alternatively, *source* models, which are based on a physical description of the sound production processes [1, 2], can be developed.

Physics-based synthesis algorithms provide *semantic* sound representations since the control parameters have a straightforward physical interpretation in terms of masses,

springs, dimensions, and so on. Consequently, modification of the parameters leads in general to meaningful results and allows more intuitive interaction between the user and the virtual instrument. The importance of sound as a primary vehicle of information is being more and more recognized in the multimedia community. Particularly, source models of sounding objects (not necessarily musical instruments) are being explored due to their high degree of interactivity and the ease in synchronizing audio and visual synthesis [3].

The physical modeling approach also has potential applications in *structured audio* coding [4, 5], a coding scheme

where, in addition to the parameters, the decoding algorithm is transmitted to the user as well. The structured audio orchestral language (SAOL) became a part of the MPEG-4 standard, thus it is widely available for multimedia applications. Known problems in using physical models for coding purposes are primarily concerned with parameter estimation. Since physical models describe specific classes of instruments, automatic estimation of the model parameters from an audio signal is not a straightforward task: the model structure which is best suited for the audio signal has to be chosen before actual parameter estimation. On the other hand, once the model structure is determined, a small set of parameters can describe a specific sound. Casey [6] and Serafin et al. [7] address these issues.

In this paper, we review some of the strategies and algorithms of physical modeling, and their applications to piano simulation. The piano is a particularly interesting instrument, both for its prominence in western music and for its complex structure [8]. Also, its control mechanism is simple (it basically reduces to key velocity), and physical control devices (MIDI keyboards) are widely available, which is not the case for other instruments. The source-based approach can be useful not only for synthesis purposes but also for gaining a better insight into the behavior of the instruments. However, as we are interested in efficient algorithms, the features modeled are only those considered to have audible effects. In general, there is a trade-off between the accuracy and the simplicity of the description. The optimal solution may vary depending on the needs of the user.

The models described here are all based on *digital waveguides*. The waveguide paradigm has been found to be the most appropriate for real-time synthesis of a wide range of musical instruments [9, 10, 11]. As early as in 1987, Garnett [12] presented a physical waveguide piano model. In his model, a semiphysical lumped hammer is connected to a digital waveguide string and the soundboard is modeled by a set of waveguides, all connected to the same termination.

In 1995, Smith and Van Duyne [13, 14] presented a model based on commuted synthesis. In their approach, the soundboard response is stored in an excitation table and fed into a digital waveguide string model. The hammer is modeled as a linear filter whose parameters depend on the hammer-string collision velocity. The hammer filter parameters have to be precalculated and stored for all notes and hammer velocities. This precalculation can be avoided by running an auxiliary string model connected to a nonlinear hammer model in parallel, and, based on the force response of the auxiliary model, designing the hammer filters in real time [15].

The original motivation for commuted synthesis was to avoid the high-order filter which is needed for high quality soundboard modeling. As low-complexity methods have been developed for soundboard modeling (see Section 5), the advantages of the commuted piano with respect to the direct modeling approach described here are reduced. Also, due to the lack in physical description, some effects, such as the restrike (*ribattuto*) of the same string, cannot be precisely modeled with the commuted approach. Describing the com-

mutated synthesis in detail is beyond the scope of this paper, although we would like to mention that it is a comparable alternative to the techniques described here.

As part of a collaboration between the University of Padova and Generalmusic, Borin et al. [16] presented a complete real-time piano model in 1997. The hammer was treated as a lumped model, with a mass connected in parallel to a nonlinear spring, and the strings were simulated using digital waveguides, all connected to a single-lumped load. Bank [17] introduced in 2000 a similar physical model, based on the same functional blocks, but with slightly different implementation. An alternative approach was used for the solution of the hammer differential equation. Independent string models were used without any coupling, and the influence of the soundboard on decay times was taken into account by using high-order loss filters. The use of feedback delay networks was suggested for modeling the radiation of the soundboard.

This paper addresses the design of each component of a piano model (i.e., hammer, string, and soundboard). Discussion is carried on with particular emphasis on real-time applications, where the time complexity of algorithms plays a key role. Perceptual issues are also addressed since a precise knowledge of what is relevant to the human ear can drive the accuracy level of the design. Section 2 deals with general aspects of piano acoustics. In Section 3, the hammer is discussed and numerical techniques are presented to overcome the computability problems in the nonlinear discretized system. Section 4 is devoted to string modeling, where the problems of parameter estimation are also addressed. Finally, Section 5 deals with the soundboard, where various alternative techniques are described and the use of the multirate approach is proposed.

## 2. ACOUSTICS AND MODEL STRUCTURE

Piano sounds are the final product of a complex synthesis process which involves the entire instrument body. As a result of this complexity, each piano note exhibits its unique sound features and nuances, especially in high quality instruments. Moreover, just varying the impact force on a single key allows the player to explore a rich dynamic space. Accounting for such dynamic variations in a wavetable-based synthesizer is not trivial: dynamic postprocessing filters which shape the spectrum according to key velocity can be designed, but finding a satisfactory mapping from velocity to filter response is far from being an easy task. Alternatively, a physical model, which mimics as closely as possible the acoustics of the instrument, can be developed.

The general structure of the piano is displayed in Figure 1a: an iron frame is attached to the upper part of the wooden case and the strings are extended upon this in a direction nearly perpendicular to the keyboard. The keyboard-side end of the string is connected to the tuning pins on the pin block, while the other end, passing the bridge, is attached to the hitch-pin rail of the frame. The bridge is a thin wooden bar that transmits the string vibration to the soundboard, which is located under the frame.

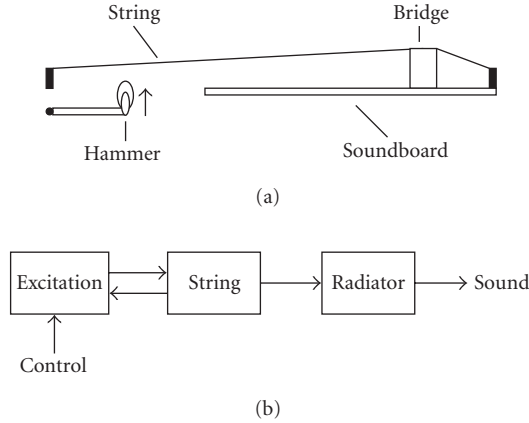


FIGURE 1: General structures: (a) schematic representation of the instrument and (b) model structure.

Since the physical modeling approach tries to simulate the structure of the instrument rather than the sound itself, the blocks in the piano model resemble the parts of a real piano. The structure is displayed in Figure 1b. The first model block is the excitation, the hammer strike. Its output propagates to the string, which determines the fundamental frequency of the tone. The quasiperiodic output signal is filtered through a postprocessing block, covering the radiation effects of the soundboard. Figure 1b shows that the hammer-string interaction is bidirectional since the hammer force depends on the string displacement [8]. On the other hand, there is no feedback from the radiator to the string. Feedback and coupling effects on the bridge and the soundboard are taken into account in the string block. The model differs from a real piano in the fact that the two functions of the soundboard, namely, to provide a terminating impedance to the strings and to radiate sound, are located in separate parts of the model. As a result, it is possible to treat radiation as a linear filtering operation.

### 3. THE HAMMER

We will first discuss the physical aspects of the hammer-string interaction, then concentrate on various modeling approaches and implementation issues.

#### 3.1. Hammer-string interaction

As a first approximation, the piano hammer can be considered a lumped mass connected to a nonlinear spring, which is described by the equation

$$F(t) = -m_h \frac{d^2 y_h(t)}{dt^2}, \quad (1)$$

where  $F(t)$  is the interaction force and  $y_h(t)$  is the hammer displacement. The hammer mass is represented by  $m_h$ . Experiments on real instruments have shown (see, e.g., [18, 19, 20]) that the hammer-string contact can be described by the following formula:

TABLE 1: Sample values for hammer parameters for three different notes, taken from [19, 20]. The hammer mass  $m_h$  is given in kg.

	$C_2$	$C_4$	$C_6$
$p$	2.3	2.5	3
$k$	$4.0 \times 10^8$	$4.5 \times 10^9$	$1.0 \times 10^{12}$
$m_h$	$4.9 \times 10^{-3}$	$2.97 \times 10^{-3}$	$2.2 \times 10^{-3}$

$$F(t) = f(\Delta y(t)) = \begin{cases} k\Delta y(t)^p, & \Delta y(t) > 0, \\ 0, & \Delta y(t) \leq 0, \end{cases} \quad (2)$$

where  $\Delta y(t) = y_h(t) - y_s(t)$  is the compression of the hammer felt,  $y_s(t)$  is the string position,  $k$  is the hammer stiffness coefficient, and  $p$  is the stiffness exponent. The condition  $\Delta y(t) > 0$  corresponds to the hammer-string contact, while the condition  $\Delta y(t) \leq 0$  indicates that the hammer is not touching the string. Equations (1) and (2) result in a nonlinear differential system of equations for  $y_h(t)$ . Due to the nonlinearity, the tone spectrum varies dynamically with hammer velocity. Typical values of hammer parameters can be found in [19, 20]. Example values are listed in Table 1.

However, (2) is not fully satisfactory in that real piano hammers exhibit hysteretic behavior. That is, contact forces during compression and during decompression are different, and a one-to-one law between compression and force does not correspond to reality. A general description of the hysteresis effect of piano felts was provided by Stulov [21]. The idea, coming from the general theory of mechanics of solids, is that the stiffness  $k$  of the spring in (2) has to be replaced by a time-dependent operator which introduces memory in the nonlinear interaction. Thus, the first part of (2) (when  $\Delta y(t) > 0$ ) is replaced by

$$F(t) = f(\Delta y(t)) = k[1 - h_r(t)] * [\Delta y(t)^p], \quad (3)$$

where  $h_r(t) = (\epsilon/\tau)e^{-t/\tau}$  is a *relaxation function* that accounts for the “memory” of the material and the  $*$  operator represents convolution.

Previous studies [22] have shown that a good fit to real data can be obtained by implementing  $h_r$  as a first-order low-pass filter. It has to be noted that informal listening tests indicate that taking into account the hysteresis in the hammer model does not improve the sound quality significantly.

#### 3.2. Implementation approaches

The hammer models described in Section 3.1 can be discretized and coupled to the string in order to provide a full physical description. However, there is a mutual dependence between (2) and (1), that is, the hammer position  $y_h(n)$  at discrete time instant  $n$  should be known for computing the force  $F(n)$ , and vice versa. The same problem arises when (3) is used instead of (2). This implicit relationship can be made explicit by assuming that  $F(n) \approx F(n - 1)$ , thus inserting a fictitious delay element in a delay-free path. Although this approximation has been extensively used in the literature (see, e.g., [19, 20]), it is a potential source of instability.

The theory of wave digital filters addresses the problem of noncomputable loops in terms of wave variables. Every component of a circuit is described as a scattering element with a reference impedance, and delay-free loops between components are treated by “adapting” reference impedances. Van Duyne et al. [23] presented a “wave digital hammer” model, where wave variables are used. More severe computability problems can arise when simulating nonlinear dynamic exciters since the linear equations used to describe the system dynamics are tightly coupled with a nonlinear map. Borin et al. [24] have recently proposed a general strategy named “K method” for solving noncomputable loops in a wide class of nonlinear systems. The method is fully described in [24] along with some application examples. Here, only the basic principles are outlined.

Whichever the discretization method is, the hammer compression  $\Delta y(n)$  at time  $n$  can be written as

$$\Delta y(n) = p(n) + KF(n), \quad (4)$$

where  $p(n)$  is the linear combination of past values of the variables (namely,  $y_h$ ,  $y_s$ , and  $F$ ) and  $K$  is a coefficient whose value depends on the numerical method in use. The interaction force  $F(n)$  at discrete time instant  $n$ , computed either by (2) or (3), is therefore described by the implicit relation  $F(n) = f(p(n) + KF(n))$ . The  $K$  method uses the implicit function theorem to solve the following implicit relation:

$$F = f(p + KF) \xrightarrow{K_{\text{meth.}}} F = h(p). \quad (5)$$

The new nonlinear map  $h$  defines  $F$  as a function of  $p$ , hence instantaneous dependencies across the nonlinearity are dropped. The function  $h$  can be precomputed and stored in a lookup table for efficient implementation.

Bank [25] presented a simpler but less general method for avoiding artifacts caused by fictitious delay insertion. The idea is that the stability of the discretized hammer model with a fictitious delay can always be maintained by choosing a sufficiently large sampling rate  $f_s$  if the corresponding continuous-time system is stable. As  $f_s \rightarrow \infty$ , the discrete-time system will behave as the original differential equation. Doubling the sampling rate of the whole string model would double the computation time as well. However, if only the hammer model operates at double rate, the computational complexity is raised only by a negligible amount. Therefore, in the proposed solution, the hammer operates at twice sampling rate of the string. Data is downsampled using simple averaging and upsampled using linear interpolation. The multirate hammer has been found to result in well-behaving force signals at a low-computational cost. As the hammer model is a nonlinear dynamic system, the stability bounds are not trivial to derive in a closed form. In practice, stability is maintained up to an impact velocity ten times higher than the point where the straightforward approach (e.g., used in [19, 20]) turns unstable.

Figure 2 shows a typical force signal in a hammer-string contact. The overall contact duration is around 2 ms and the pulses in the signal are produced by reflections of force

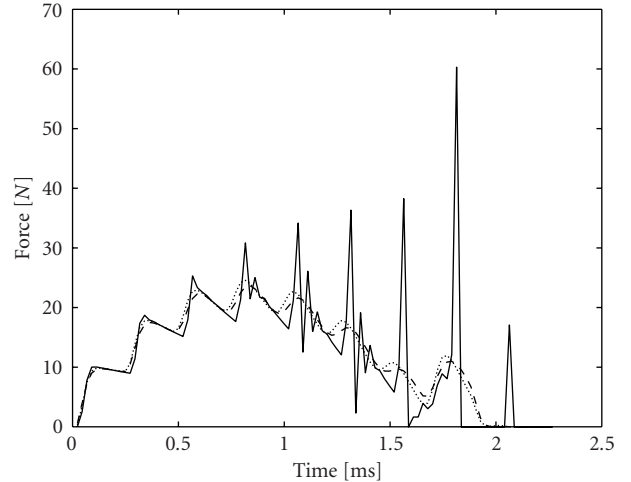


FIGURE 2: Time evolution of the interaction force for note  $C_5$  (522 Hz) with  $f_s = 44.1$  kHz, and hammer velocity  $v = 5$  m/s, computed by inserting a fictitious delay element (solid line), with the  $K$  method (dashed line), and with the multirate hammer (dotted line).

waves at string terminations. The  $K$  method and the multirate hammer produce very similar force signals. On the other hand, inserting a fictitious delay element drives the system towards instability (the spikes are progressively amplified). In general, the multirate method provide results comparable to the  $K$  method for hammer parameters realistic for pianos, while it does not require that precomputed lookup tables be stored. On the other hand, when low-sampling rates (e.g.,  $f_s = 11.025$  kHz) or extreme hammer parameters are used (i.e.,  $k$  is ten times the value listed in Table 1), the system stability cannot be maintained by upsampling by a factor of 2. In such cases, the  $K$  method is the appropriate solution.

The computational approaches presented in this section are applicable to a wide class of mechanical interactions between physical objects [26].

#### 4. THE STRING

Many different approaches have been presented in the literature for string modeling. Since we are considering techniques suitable for real-time applications, only the digital waveguide [9, 10, 11] is described here in detail. This method is based on the time-domain solution of the one-dimensional wave equation. The velocity distribution of the string  $v(x, t)$  can be seen as the sum of two traveling waves:

$$v(x, t) = v^+(x - ct) + v^-(x + ct), \quad (6)$$

where  $x$  denotes the spatial coordinate,  $t$  is time,  $c$  is the propagation speed, and  $v^+$  and  $v^-$  are the traveling wave components.

Spatial and time-domain sampling of (6) results in a simple delay-line representation. Nonideal, lossy, and stiff strings can also be modeled by the method. If linearity and time invariance of the string are assumed, all the distributed losses



Another approach was proposed by Bank [17] who suggested the transformation of the specification. As the goal is to match decay times, the magnitude specification  $g_k$  is transformed into a form  $g_{k,\text{tr}} = 1/(1 - g_k)$  which approximates  $\tau_k$ , and a transformed filter  $H_{\text{tr}}(z)$  is designed for the new specification by least squares filter design. The loss filter  $H_l(z)$  is then computed by the inverse transform  $H_l(z) = 1 - 1/H_{\text{tr}}(z)$ .

Bank and Välimäki [32] presented a simpler method for high-order filter design based on a special weighting function. The resulting decay times of the digital waveguide are computed from the magnitude response  $\hat{g}_k = |H(e^{j\theta_k})|$  of the loss filter by  $\hat{\tau}_k = d(\hat{g}_k) = -1/(f_0 \ln \hat{g}_k)$ . This function is approximated by its first-order Taylor series around the specification  $d(\hat{g}_k) \approx d(g_k) + d'(g_k - g_k)$ . Accordingly, the error with respect to decay times can be approximated by the weighted mean square error

$$e_{\text{WLS}} = \sum_{k=1}^K w_k (H_l(e^{j\theta_k}) - g_k)^2, \quad (11)$$

$$w_k = \frac{1}{(g_k - 1)^4}.$$

The weighted error  $e_{\text{WLS}}$  can be easily minimized by standard filter design algorithms, and leads to a good match with respect to decay times.

All of these techniques for high-order loss filter design have been found to be robust in practice. Comparing them is left for future work.

Borin et al. [16] have used a different approach for modeling the decay time variations of the partials. In their implementation, second-order FIR filters are used as loss filters that are responsible for the general decay of the note. Small variations of the decay times are modeled by connecting all the string models to a common termination which is implemented as a filter with a high number of resonances. This also enables the simulation of the pedal effect since now all the strings are coupled to each other (see Section 4.3). An advantage of this method compared to high-order loop filters is the smaller computational complexity. On the other hand, the partial envelopes of the different notes cannot be controlled independently.

Although optimizing the loss filter with respect to decay times has been found to give perceptually adequate results, we remark that the loss filter design can be helped via perceptual studies. The audibility of the decay-time variations for the one-pole loss filter was studied by Tolonen and Järveläinen [33]. The study states that relatively large deviations (between  $-25\%$  and  $+40\%$ ) in the overall decay time of the note are not perceived by listeners. Unfortunately, theoretical results are not directly applicable for the design of high-order loss filters as the tolerance for the decay time variations of single partials is not known.

#### 4.2. Dispersion simulation

Dispersion is due to stiffness which causes piano strings to deviate from ideal behavior. If the dispersive correction term in the wave equation is small, its first-order effect is

to increase the wave propagation speed  $c(f)$  with frequency. This phenomenon causes string partials to become inharmonic. If the string parameters are known, then the frequency of the  $k$ th stretched partial can be computed as

$$f_k = k f_0 \sqrt{1 + Bk^2}, \quad (12)$$

where the value of the inharmonicity coefficient  $B$  depends on the parameters of the string (see, e.g., [34]).

Phase delay specification  $D_d(f_k)$  for the dispersion filter  $H_d(z)$  can be computed from the partial frequencies:

$$D_d(f_k) = \frac{f_s k}{f_k} - N - D_l(f_k), \quad (13)$$

where  $N$  is the total length of the waveguide delay line and  $D_l(f_k)$  is the phase delay of the loss filter  $H_l(z)$ . The phase specification of the dispersion filter becomes  $\phi_{\text{pre}}(f_k) = 2\pi f_k D_d(f_k)/f_s$ .

Van Duyne and Smith [35] proposed an efficient method for simulating dispersion by cascading equal first-order all-pass filters in the waveguide loop; however, the constraint of using equal first-order sections is too severe and does not allow accurate tuning of inharmonicity.

Rocchesso and Scalcon [36] proposed a design method based on [37]. Starting from a target phase response,  $l$  points  $\{f_k\}_{k=1,\dots,l}$  are chosen on the frequency axis corresponding to the points where string partials should be located. The filter order is chosen to be  $n < l$ . For each partial  $k$ , the method computes the quantities

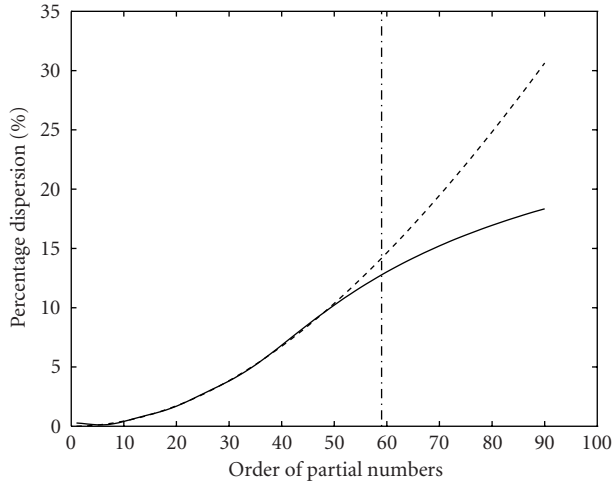
$$\beta_k = -\frac{1}{2}(\phi_{\text{pre}}(f_k) + 2n\pi f_k), \quad (14)$$

where  $\phi_{\text{pre}}(f)$  is the prescribed allpass response. Filter coefficients  $a_j$  are computed by solving the system

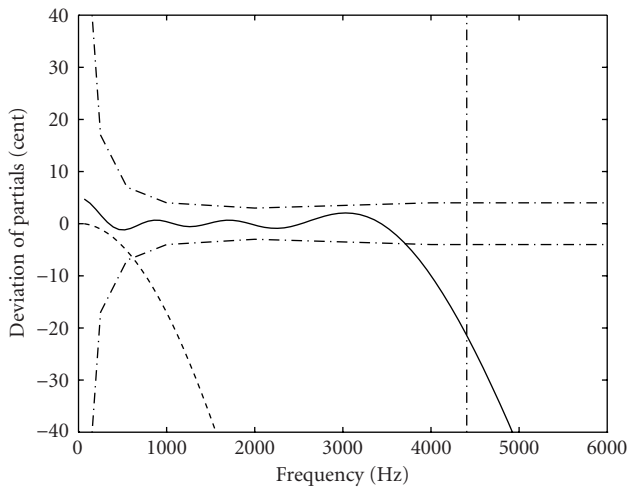
$$\sum_{j=1}^n a_j \sin(\beta_k + 2j\pi f_k) = -\sin(\beta_k), \quad k = 1, \dots, l. \quad (15)$$

A least-squared equation error (LSEE) is used to solve the overdetermined system (15). It was shown in [36] that several tens of partials can be correctly positioned for any piano key, with the allpass filter order not exceeding 20. Moreover, the fine tuning of the string is automatically taken into account in the design. Figure 4 plots results obtained using a filter order of 18. Note that the pure tone frequency JND (just noticeable difference) has been used in Figure 4b as a reference as no accurate studies of partial JNDs for piano tones are available to our knowledge.

Since the computational load for  $H_d(z)$  is heavy, it is important to find criteria for accuracy and order optimization with respect to human perception. Rocchesso and Scalcon [38] studied the dependence of the bandwidth of perceived inharmonicity (i.e., the frequency range in which misplacement of partials is audible) on the fundamental frequency by performing listening tests with decaying piano tones. The bandwidth has been found to increase almost linearly on a



(a)



(b)

FIGURE 4: Dispersion filter (18th order) for the  $C_2$  string: (a) computed (solid line) and theoretical (dashed line) percentage dispersion versus partial numbers and (b) deviation of partials (solid line). Dash-dotted vertical lines show the end of the LSEE approximation; dash-dotted bounds in (b) indicate the pure tone frequency JND as a reference; and the dashed line in (b) is the partial deviation from the theoretical inharmonic series in a nondispersive string model.

logarithmic pitch scale. Partial above this frequency band only contribute some brightness to the sound, and can be made harmonic without relevant perceptual consequences.

Järveläinen et al. [39] also found that inharmonicity is more easily perceived at low frequencies even when coefficient  $B$  for bass tones is lower than for treble tones. This is probably due to the fact that beats are used by listeners as cues for inharmonicity, and even low  $B$ 's produce enough mistuning in higher partials of low tones. These findings can help in the allpass filter design procedure, although a number of issues still need further investigation.

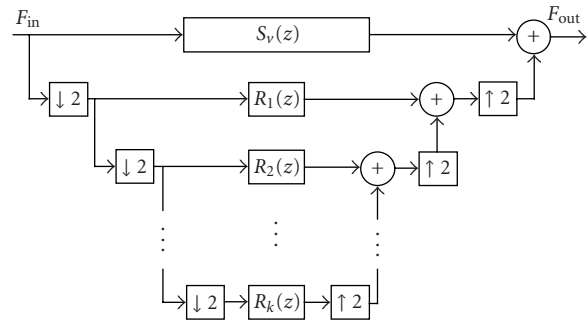


FIGURE 5: The multirate resonator bank.

As high-order dispersion filters are needed for modeling low notes, the computational complexity is increased significantly. Bank [17] proposed a multirate approach to overcome this problem. Since the lowest tones do not contain significant energy in the high-frequency region anyway, it is worthwhile to run the lowest two or three octaves of the piano at half-sampling rate of the model. The outputs of the low notes are summed before upsampling, therefore only one interpolation filter is required.

### 4.3. Coupled piano strings

String coupling occurs at two different levels. First of all, two or three slightly mistuned strings are sounded together when a single piano key is pressed (except for the lowest octave) and complicated modulation of the amplitudes is brought about. This results in beating and two-stage decay, the first refers to an amplitude modulation overlaid on the exponential decay, and the latter means that tone decays are faster in the early part than in the latter. These phenomena were studied by Weinreich as early as in 1977 [40]. At the second level, the presence of the bridge and the action of the soundboard is known to originate important coupling effects even between different tones. In fact, the bridge-soundboard system connects strings together and acts as a distributed driving-point impedance for string terminations.

The simplest way for modeling beating and two-stage decay is to use two digital waveguides in parallel for a single note. Varying by the used type of coupling, many different solutions have been presented in the literature, see, for example, [14, 41].

Bank [17] presented a different approach for modeling beating and two-stage decay, based on a parallel resonator bank. In a subsequent study, the computational complexity of the method was decreased by an order of ten by applying multirate techniques, making the approach suitable for real-time implementations [42]. In this approach, second-order resonators  $R_1(z) \cdots R_k(z)$  are connected to the basic string model  $S_v(z)$  in parallel, rather than using a second waveguide. The structure is depicted in Figure 5. The idea comes from the observation that the behavior of two coupled strings can be described by a pair of exponentially damped sinusoids [40]. In this model, one sinusoid of the mode pair is simulated by one partial of the digital waveguide and the other

one by one of the resonators  $R_k(z)$ . The transfer functions of the resonators are as follows:

$$R_k(z) = \frac{\operatorname{Re}\{a_k\} - \operatorname{Re}\{a_k \overline{p_k}\} z^{-1}}{1 - 2 \operatorname{Re}\{p_k\} z^{-1} + |p_k|^2 z^{-2}}, \quad (16)$$

$$a_k = A_k e^{j\varphi_k}, \quad p_k = e^{j(2\pi f_k/f_s) - 1/f_s \tau_k},$$

where  $A_k$ ,  $\varphi_k$ ,  $f_k$ , and  $\tau_k$  refer to the initial amplitude, initial phase, frequency, and decay-time parameters of the  $k$ th resonator, respectively. The overline stands for complex conjugation,  $\operatorname{Re}$  indicates the real part of a complex variable, and  $f_s$  is the sampling frequency.

An advantage of the structure is that the resonators  $R_k(z)$  are implemented only for those partials whose beating and two-stage decay are prominent. The others will have simple exponential decay, determined by the digital waveguide model  $S_v(z)$ . Five to ten resonators have been found to be enough for high-quality sound synthesis. The resonator bank is implemented by the multirate approach, running the resonators at a much lower-sampling rate, for example, the 1/8 or 1/16 part of the original sampling frequency.

It is shown in [42] that when only half of the downsampled frequency band is used for resonators, no lowpass filtering is needed before downsampling. This is due to the fact that the excitation signal is of lowpass character leading to aliasing less than  $-20$  dB. As the role of the excitation signal is to set the initial amplitudes and phases of the resonators, the result of this aliasing is a less than 1 dB change in the resonator amplitudes, which has been found to be inaudible. On the other hand, the interpolation filters after upsampling cannot be neglected. However, they are not implemented for all notes separately; the lower-sampling rate signals of the different strings are summed before interpolation filtering (this is not depicted in Figure 5). Their specification is relatively simple (e.g., 5 dB passband ripple) since their passband errors can be easily corrected by changing the initial amplitudes and phases of the resonators. This results in significantly lower-computational cost, compared to the methods which use coupled waveguides.

Generally, the average computational cost of the method for one note is less than five multiplications per sample. Moreover, the parameter estimation gets simpler since only the parameters of the mode pairs have to be found by, for example, the methods presented in [17, 41], and there is no need for coupling filter design. Stability problems of a coupled system are also avoided. The method presented here shows that combining physical and signal-based approaches can be useful in reducing computational complexity.

Modeling the coupling between strings of different tones is essential when the sustain pedal effect has to be simulated. Garnett [12] and Borin et al. [16] suggested connecting the strings to the same lumped terminating impedance. The impedance is modeled by a filter with a high number of peaks. For that, the use of feedback delay networks [43, 44] is a good alternative. Although in real pianos the bridge connects to the string as a distributed termination, thus coupling different strings in different ways, the simple model of Borin et al. was able to produce a realistic sustain pedal effect [45].

## 5. RADIATION MODELING

The soundboard radiates and filters the string waves that reach the bridge, and radiation patterns are essential for describing the “presence” of a piano in a musical context. However, now we are concentrating on describing the sound pressure generated by the piano at a certain locus in the listening space, that is, the directional properties of radiation are not taken into account. Modeling the soundboard as a linear postprocessing stage is an intrinsically weak approach since on a real piano it also accounts for coupling between strings and affects the decay times of the partials. However, as already stated in Section 2, our modeling strategy keeps the *radiation* properties of the soundboard separated from its *impedance* properties. The latter are incorporated in the string model, and have already been addressed in Sections 4.1 and 4.3; here we will concentrate on radiation.

A simple and efficient radiation model was presented by Garnett [12]. The waveguide strings were connected to the same termination and the soundboard was simulated by connecting six additional waveguides to the common termination. This can be seen as a predecessor of using feedback delay networks for soundboard simulation. Feedback delay networks have been proven to be efficient in simulating room reverberation since they are able to produce high-modal density at a low-computational cost [43]. For an overview, see the work of Rocchesso and Smith [44]. Bank [17] applied feedback delay networks with shaping filters for the simulation of piano soundboards. The shaping filters were parametrized in such a way that the system matched the overall magnitude response of a real piano soundboard. A drawback of the method is that the modal density and the quality factors of the modes are not fully controllable. The method has proven to yield good results for high piano notes, where simulating the attack noise (the knock) of the tone is the most important issue.

The problem of soundboard radiation can also be addressed from the point of view of filter design. However, as the soundboard exhibits high-modal density, a high-order filter has to be used. For  $f_s = 44.1$  kHz, a 2000 tap FIR filter was necessary to achieve good results. The filter order did not decrease significantly when IIR filters were used.

To resolve the high-computational complexity, a multi-rate soundboard model was proposed by Bank et al. [46]. The structure of the model is depicted in Figure 6. The string signal is split into two parts. The part below 2.2 kHz is downsampled by a factor of 8 and filtered by a high-order filter  $H_{\text{low}}(z)$  precisely synthesizing the amplitude and phase response of the soundboard for the low frequencies. The part above 2.2 kHz is filtered by a low-order filter, modeling the overall magnitude response of the soundboard at high frequencies. The signal of the high-frequency chain is delayed by  $N$  samples to compensate for the latency of decimation and interpolation filters of the low-frequency chain.

The filters  $H_{\text{low}}(z)$  and  $H_{\text{high}}(z)$  are computed as follows. First, a target impulse response  $H_t(z)$  is calculated by



measuring the force-pressure transfer function of a real piano soundboard. Then, this is lowpass-filtered and downsampled by a factor of 8 to produce an FIR filter  $H_{\text{low}}(z)$ . The impulse response of the low-frequency chain is now subtracted from the target response  $H_t(z)$  providing a residual response containing energy above 2.2 kHz. This residual response is made minimum phase and windowed to a short length (50 tap). The multirate soundboard model outlined here consumes 100 operations per cycle and produces a spectral character similar to that of a 2000-tap FIR filter. The only difference is that the attack of high notes sounds sharper since the energy of the soundboard response is concentrated to a short time period above 2.2 kHz. This could be overcome by using feedback delay networks for  $H_{\text{high}}(z)$ , which is left for future research.

The parameters of the multirate soundboard model cannot be interpreted physically. However, this does not lead to any drawbacks since the parameters of the soundboard cannot be changed by the player in real pianos either. Having a purely physical model, for example, based on finite differences [47], would lead to unacceptable high-computational costs. Therefore, implementing a black box model block as a part of a physical instrument model seems to be a good compromise.

## 6. CONCLUSIONS

This paper has reviewed the main stages of the development of a physical model for the piano, addressing computational aspects and discussing problems that not only are related to piano synthesis but arise in a broad class of physical models of sounding objects.

Various approaches have been discussed for dealing with nonlinear equations in the excitation block. We have pointed out that inaccuracies at this stage can lead to severe instability problems. Analogous problems arise in other mechanical and acoustical models, such as impact and friction between two sounding objects, or reed-bore interaction. The two alternative solutions presented for the piano hammer can be used in a wide range of applications.

Several filter design techniques have been reviewed for the accurate tuning of the resonating waveguide block. It has been shown that high-order dispersion filters are needed for accurate simulation of inharmonicity. Therefore, perceptual issues have been addressed since they are helpful in optimizing the design and reducing computational loads. The requirement of physicality can always be weakened when the effect caused by a specific feature is considered to be inaudible.

A filter-based approach was presented for the soundboard model. As such, it cannot be interpreted as physical, but this does not influence the functionality of the model. In general, only those parameters which are involved in block interaction or are influenced by control messages need to have a clear physical interpretation. Therefore, we recommend synthesis structures that are based on building blocks with physical input and output parameters, but whose inner

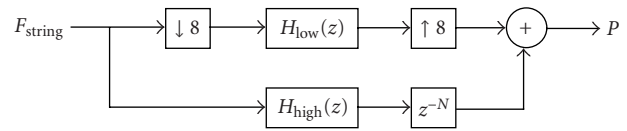


FIGURE 6: The multirate soundboard model.

structure does not necessarily follow physical model. In other words, the basic building blocks are black box models with the most efficient implementations available, and they form the physical structure of the instrument model at a higher level.

The use of multirate techniques was suggested for modeling beating and two-stage decay as well as the soundboard. The model can run at different sampling rates (e.g., 44.1, 22.05, and 11.025 kHz) or/and with different filter orders implemented in the digital waveguide model. Since the stability of the numerical structures is maintained in all cases, the user has the option of choosing between quality and efficiency. This remark is also relevant for potential applications in structured audio coding. In cases when instrument models are to be encoded and transmitted without the precise knowledge of the computational power of the decoder, it is essential that the stability is guaranteed even at low-sampling rates in order to allow graceful degradation.

## ACKNOWLEDGMENTS

Work at CSC-DEI, University of Padova, was developed under a Research Contract with Generalmusic. Partial funding was provided by the EU Project “MOSART,” Improving Human Potential, and the Hungarian National Scientific Research Fund OTKA F035060. The authors are thankful to P. Hussami and to the anonymous reviewers for their helpful comments, which have contributed to the improvement of the paper.

## REFERENCES

- [1] G. De Poli, “A tutorial on digital sound synthesis techniques,” in *The Music Machine*, C. Roads, Ed., pp. 429–447, MIT Press, Cambridge, Mass, USA, 1991.
- [2] J. O. Smith III, “Viewpoints on the history of digital synthesis,” in *Proc. International Computer Music Conference (ICMC '91)*, pp. 1–10, Montreal, Quebec, Canada, October 1991.
- [3] K. Tadamura and E. Nakamae, “Synchronizing computer graphics animation and audio,” *IEEE Multimedia*, vol. 5, no. 4, pp. 63–73, 1998.
- [4] E. D. Scheirer, “Structured audio and effects processing in the MPEG-4 multimedia standard,” *Multimedia Systems*, vol. 7, no. 1, pp. 11–22, 1999.
- [5] B. L. Vercoe, W. G. Gardner, and E. D. Scheirer, “Structured audio: creation, transmission, and rendering of parametric sound representations,” *Proceedings of the IEEE*, vol. 86, no. 5, pp. 922–940, 1998.
- [6] M. A. Casey, “Understanding musical sound with forward models and physical models,” *Connection Science*, vol. 6, no. 2-3, pp. 355–371, 1994.

- [7] S. Serafin, J. O. Smith III, and H. Thornburg, "A pattern recognition approach to invert a bowed string physical model," in *Proc. International Symposium on Musical Acoustics (ISMA '01)*, pp. 241–244, Perugia, Italy, September 2001.
- [8] N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, Springer-Verlag, New York, NY, USA, 1991.
- [9] J. O. Smith III, *Techniques for digital filter design and system identification with application to the violin*, Ph.D. thesis, Department of Music, Stanford University, Stanford, Calif, USA, June 1983.
- [10] J. O. Smith III, "Principles of digital waveguide models of musical instruments," in *Applications of Digital Signal Processing to Audio and Acoustics*, M. Kahrs and K. Brandenburg, Eds., pp. 417–466, Kluwer Academic, Boston, Mass, USA, 1998.
- [11] J. O. Smith III, *Digital Waveguide Modeling of Musical Instruments*, August 2002, <http://www-ccrma.stanford.edu/~jos/waveguide/>.
- [12] G. E. Garnett, "Modeling piano sound using waveguide digital filtering techniques," in *Proc. International Computer Music Conference (ICMC '87)*, pp. 89–95, Urbana, Ill, USA, September 1987.
- [13] J. O. Smith III and S. A. Van Duyne, "Commutated piano synthesis," in *Proc. International Computer Music Conference (ICMC '95)*, pp. 335–342, Banff, Canada, September 1995.
- [14] S. A. Van Duyne and J. O. Smith III, "Developments for the commuted piano," in *Proc. International Computer Music Conference (ICMC '95)*, pp. 319–326, Banff, Canada, September 1995.
- [15] B. Bank and L. Sujbert, "On the nonlinear commuted synthesis of the piano," in *Proc. 5th International Conference on Digital Audio Effects (DAFx '02)*, pp. 175–180, Hamburg, Germany, September 2002.
- [16] G. Borin, D. Rocchesso, and F. Scalcon, "A physical piano model for music performance," in *Proc. International Computer Music Conference (ICMC '97)*, pp. 350–353, Thessaloniki, Greece, September 1997.
- [17] B. Bank, "Physics-based sound synthesis of the piano," M.S. thesis, Department of Measurement and Information Systems, Budapest University of Technology and Economics, Budapest, Hungary, May 2000, published as Tech. Rep. 54, Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, Helsinki, Finland.
- [18] D. E. Hall, "Piano string excitation VI: Nonlinear modeling," *Journal of the Acoustical Society of America*, vol. 92, no. 1, pp. 95–105, 1992.
- [19] A. Chaigne and A. Askenfelt, "Numerical simulations of piano strings. I. A physical model for a struck string using finite difference methods," *Journal of the Acoustical Society of America*, vol. 95, no. 2, pp. 1112–1118, 1994.
- [20] A. Chaigne and A. Askenfelt, "Numerical simulations of piano strings. II. Comparisons with measurements and systematic exploration of some hammer-string parameters," *Journal of the Acoustical Society of America*, vol. 95, no. 3, pp. 1631–1640, 1994.
- [21] A. Stulov, "Hysteretic model of the grand piano hammer felt," *Journal of the Acoustical Society of America*, vol. 97, no. 4, pp. 2577–2585, 1995.
- [22] G. Borin and G. De Poli, "A hysteretic hammer-string interaction model for physical model synthesis," in *Proc. Nordic Acoustical Meeting (NAM '96)*, pp. 399–406, Helsinki, Finland, June 1996.
- [23] 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. International Computer Music Conference (ICMC '94)*, pp. 411–418, Århus, Denmark, September 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, and Audio Processing*, vol. 8, no. 5, pp. 597–605, 2000.
- [25] B. Bank, "Nonlinear interaction in the digital waveguide with the application to piano sound synthesis," in *Proc. International Computer Music Conference (ICMC '00)*, pp. 54–57, Berlin, Germany, September 2000.
- [26] F. Avanzini, M. Rath, D. Rocchesso, and L. Ottaviani, "Low-level models: resonators, interactions, surface textures," in *The Sounding Object*, D. Rocchesso and F. Fontana, Eds., pp. 137–172, Edizioni di Mondo Estremo, Florence, Italy, 2003.
- [27] D. A. Jaffe and J. O. Smith III, "Extensions of the Karplus-Strong plucked-string algorithm," *Computer Music Journal*, vol. 7, no. 2, pp. 56–69, 1983.
- [28] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy, "Physical modeling of plucked string instruments with application to real-time sound synthesis," *Journal of the Audio Engineering Society*, vol. 44, no. 5, pp. 331–353, 1996.
- [29] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay—tools for fractional delay filter design," *IEEE Signal Processing Magazine*, vol. 13, no. 1, pp. 30–60, 1996.
- [30] V. Välimäki and T. Tolonen, "Development and calibration of a guitar synthesizer," *Journal of the Audio Engineering Society*, vol. 46, no. 9, pp. 766–778, 1998.
- [31] C. Erkut, "Loop filter design techniques for virtual string instruments," in *Proc. International Symposium on Musical Acoustics (ISMA '01)*, pp. 259–262, Perugia, Italy, September 2001.
- [32] B. Bank and V. Välimäki, "Robust loss filter design for digital waveguide synthesis of string tones," *IEEE Signal Processing Letters*, vol. 10, no. 1, pp. 18–20, 2002.
- [33] T. Tolonen and H. Järveläinen, "Perceptual study of decay parameters in plucked string synthesis," in *Proc. AES 109th Convention*, Los Angeles, Calif, USA, September 2000, preprint No. 5205.
- [34] H. Fletcher, E. D. Blackham, and R. Stratton, "Quality of piano tones," *Journal of the Acoustical Society of America*, vol. 34, no. 6, pp. 749–761, 1962.
- [35] S. A. Van Duyne and J. O. Smith III, "A simplified approach to modeling dispersion caused by stiffness in strings and plates," in *Proc. International Computer Music Conference (ICMC '94)*, pp. 407–410, Århus, Denmark, September 1994.
- [36] D. Rocchesso and F. Scalcon, "Accurate dispersion simulation for piano strings," in *Proc. Nordic Acoustical Meeting (NAM '96)*, pp. 407–414, Helsinki, Finland, June 1996.
- [37] M. Lang and T. I. Laakso, "Simple and robust method for the design of allpass filters using least-squares phase error criterion," *IEEE Trans. Circuits and Systems*, vol. 41, no. 1, pp. 40–48, 1994.
- [38] D. Rocchesso and F. Scalcon, "Bandwidth of perceived inharmonicity for physical modeling of dispersive strings," *IEEE Trans. Speech, and Audio Processing*, vol. 7, no. 5, pp. 597–601, 1999.
- [39] H. Järveläinen, V. Välimäki, and M. Karjalainen, "Audibility of the timbral effects of inharmonicity in stringed instrument tones," *Acoustic Research Letters Online*, vol. 2, no. 3, pp. 79–84, 2001.
- [40] G. Weinreich, "Coupled piano strings," *Journal of the Acoustical Society of America*, vol. 62, no. 6, pp. 1474–1484, 1977.
- [41] M. Aramaki, J. Bensa, L. Daudet, Ph. Guillemain, and R. Kronland-Martinet, "Resynthesis of coupled piano string vibrations based on physical modeling," *Journal of New Music Research*, vol. 30, no. 3, pp. 213–226, 2001.

- [42] B. Bank, "Accurate and efficient modeling of beating and two-stage decay for string instrument synthesis," in *Proc. MOSART Workshop on Current Research Directions in Computer Music*, pp. 134–137, Barcelona, Spain, November 2001.
- [43] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. 90th AES Convention*, Paris, France, February 1991, preprint No. 3030.
- [44] D. Rocchesso and J. O. Smith III, "Circulant and elliptic feedback delay networks for artificial reverberation," *IEEE Trans. Speech, and Audio Processing*, vol. 5, no. 1, pp. 51–63, 1997.
- [45] G. De Poli, F. Campetella, and G. Borin, "Pedal resonance effect simulation device for digital pianos," United States Patent 5,744,743, April 1998, (Appl. No. 618379, filed: March 1996).
- [46] B. Bank, G. De Poli, and L. Sujbert, "A multi-rate approach to instrument body modeling for real-time sound synthesis applications," in *Proc. 112th AES Convention*, Munich, Germany, May 2002, preprint no. 5526.
- [47] B. Bazzi and D. Rocchesso, "Numerical investigation of the acoustic properties of piano soundboards," in *Proc. XIII Colloquium on Musical Informatics (CIM '00)*, pp. 39–42, L'Aquila, Italy, September 2000.

**Balázs Bank** was born in 1977 in Budapest, Hungary. He received his M.S. degree in electrical engineering in 2000 from the Budapest University of Technology and Economics. In the academic year 1999/2000, he was with the Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, completing his thesis as a Research Assistant within the "Sound Source Modeling" project. From October 2001 to April 2002, he held a Research Assistant position at the Department of Information Engineering, University of Padova within the EU project "MOSART Improving Human Potential." He is currently studying for his Ph.D. degree at the Department of Measurement and Information Systems, Budapest University of Technology and Economics. He works on the physics-based sound synthesis of musical instruments, with a primary interest on the piano.



**Federico Avanzini** received in 1997 the Laurea degree in physics from the University of Milano, with a thesis on nonlinear dynamical systems and full marks. From November 1998 to November 2001, he pursued a Ph.D. degree in computer science at the University of Padova, with a research project on "Computational issues in physically-based sound models." Within his doctoral activities (January to June, 2001), he worked as a visiting Researcher at the Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, where he was involved in the "Sound Source Modeling" project. He is currently a Postdoctoral Researcher at the University of Padova. His research interests include sound synthesis models in human-computer interaction, computational issues, models for voice synthesis and analysis, and multimodal interfaces. Recent research experience includes participation to both national ("Sound models in human-computer and human-environment interaction") and European ("SOB—The Sounding Object," and "MEGA—Multisensory Expressive Gesture Applications") projects, where he has been working on sound synthesis algorithms based on physical models.



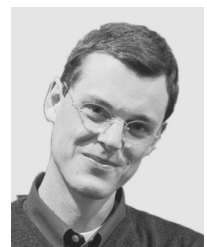
**Gianpaolo Borin** received the Laurea degree in electronic engineering from the University of Padova, Italy, in 1990, with a thesis on sound synthesis by physical modeling. Since then, he has been doing research at the Center of Computational Sonology (CSC), University of Padova. He has also been working both as a Unix Professional Developer and as a Consultant Researcher for Generalmusic. He is a coauthor of a Generalmusic's US patent for digital piano postprocessing methods. His current research interest include algorithms and methods for the efficient implementation of physical models of musical instruments and tools for real-time sound synthesis.



**Giovanni De Poli** is an Associate Professor in the Department of Information Engineering of the University of Padova, where he teaches classes of Fundamentals of Informatics II and Processing Systems for Music. He is the Director of Center of Computational Sonology (CSC), University of Padova. He is a member of the ExCom of the IEEE Computer Society Technical Committee on Computer Generated Music, Associate Editor of the International Journal of New Music Research, and member of the board of Directors of AIMI (Associazione Italiana Informatica Musicale), board of Directors of CIARM (Centro Interuniversitario Acustica e Ricerca Musicale), and Scientific Committee of ACROE (Institut National Polytechnique Grenoble). His main research interests are in algorithms for sound synthesis and analysis, models for expressiveness in music, multimedia systems and human-computer interaction, and preservation and restoration of audio documents. He is an author of several scientific international publications. He has also served in the Scientific Committees of international conferences. He is involved in the COST G6, MEGA, and MOSART-IHP European projects as a local coordinator. He is an owner of several patents on digital music instruments.



**Federico Fontana** received the Laurea degree in electronic engineering from the University of Padova, Padua, Italy, in 1996, and the Ph.D. degree in computer science from the University of Verona, Verona, Italy, in 2003. He is currently a Postdoctoral researcher in the Department of Information Engineering at the University of Padova, and collaborates with the Video, Image Processing and Sound (VIPS) Laboratory in the Dipartimento di Informatica at the University of Verona. From 1998 to 2000 he has been collaborating with the Center of Computational Sonology (CSC), University of Padova, working on sound synthesis by physical modeling. During the same period he has been a consultant of Generalmusic, Italy, and STMicroelectronics—Audio & Automotive Division, Italy, in the design and realization of real-time algorithms for the deconvolution, virtual spatialization and dynamic processing of musical and audio signals. In 2001, he has been visiting the Laboratory of Acoustics and Audio Signal Processing at the Helsinki University of Technology, Espoo, Finland. He has been involved in several national and international research projects. His main interests are in audio signal processing and physical sound modeling and spatialization in virtual and interactive environments and in multimedia systems.



**Davide Rocchesso** received the Laurea degree in electronic engineering and the Ph.D. degree from the University of Padova, Padua, Italy, in 1992 and 1996, respectively. His Ph.D. research involved the design of structures and algorithms based on feedback delay networks for sound processing applications. In 1994 and 1995, he was a visiting Scholar at the Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, Calif. Since 1991, he has been collaborating with the Center of Computational Sonology (CSC), University of Padova, as a Researcher and Live-Electronic Designer. Since 1998, he has been with the University of Verona, Verona, Italy, where he is now an Associate Professor. At the Dipartimento di Informatica of the University of Verona, he coordinates the project "Sounding Object," funded by the European Commission within the framework of the Disappearing Computer initiative. His main interests are in audio signal processing, physical modeling, sound reverberation and spatialization, multimedia systems, and human-computer interaction.

