Controlling Material Properties in Physical Models of Sounding Objects

Federico Avanzini[1], Davide Rocchesso[2]
[1]Dip. di Elettronica ed Informatica - Università di Padova, [2]Dip. di Informatica - Università di Verona
email: <avanzini@dei.unipd.it> <rocchesso@sci.univr.it>

Abstract

Contact sounds can provide information for material perception, and according to previous studies decay seems to be one of the most important cues. This information can be used to tune a physical model to represent various materials. This work deals with a specific class of contact sounds, i.e. collision sounds. A hammer-resonator physical model is developed, and subjective experiments are performed in which listeneres are asked to indicate what material each synthesized sound is coming from. Experiments show that the model is able to convey information about material of the sound source.

1 Introduction

The sound produced by an acoustic resonator depends on a large number of factors, including shape, material, excitation. According to ecological acoustics (Gaver 1993), these can be grouped into two broad cathegories, namely *structural invariants* (specifying individual properties of *objects* such as size, shape, material) and *transformational invariants* (characterizing *interactions*, i.e. the way objects are played). A still open problem is how structural invariants are conveyed into sound signals, and to what extent they can be recovered by a listener. This is referred to as the problem of sound-source determination, and is a fundamental question for the sonification of multimedia environments and for the design of auditory icons (Gaver 1994).

At present, few studies have investigated what acoustic cues (if any) are exploited by the auditory system in order to recognize materials of sound sources. Based on theoretical considerations, Wildes and Richards (1988) suggested the overall decay time as a significant cue, since it is a direct measure of internal friction in a given material; however, this is only true when a standard anelastic linear solid model is assumed. Two recent studies with listening subjects provided some experimental basis to this conjecture, but results were not in accordance. Lutfi and Oh (1997) found that changes in the decay time are not easily perceived by listeners, while changes in the fundamental frequency seem to be a more salient cue. On the other hand, Klatzky, Pai, and Krotov (2000) showed that decay plays a much larger role than pitch in affecting judgement.

Even less clear is how to incorporate material properties in synthesis algorithms. Physically based synthesis is a natural approach for dealing with such a problem, since it provides control parameters that are directly related to physical reality. Djoharian (2000) showed that finite difference models of resonators can be covered by a “viscoelastic dress” to fit a given frequency-damping characteristic, which is taken as the sound signature of the material. This approach relies on a low-level physical description and, as a result, very accurate yet computationally expensive algorithms are obtained.

Alternatively, one can use a much simpler model, that allows for control over significant acoustic cues (namely, pitch and decay) and neglects to a certain extent other features of the sound source. In this paper we develop a simple hammer-resonator physical model and show that even such an over-simplified model can be used to convey information about materials and to synthesize “cartoon” sounding objects made of various materials. A similar approach has been used for incorporating shape information into physically based sound models (Rocchesso 2001).

The structure of the paper is as follows: Sec. 2 discusses the physical model and the numerical implementation; Sec. 3 describes subjective experiments performed with synthetic stimuli obtained from the model. In Sec. 4, results from the experiments are discussed. The model and sound examples are available at `www.soundobject.org`.

2 The model

Table 1 summarizes the main variables and parameters used throughout the paper. Since the main signal features we are interested in are pitch and decay time, a second order oscillator is a suitable structure for describing the acoustic resonator. In order to achieve realistic simulation of the excitation mechanism, a hammer model is needed as well: we assume the hammer to be a lumped mass moving freely. During contact, interaction between hammer and resonator is modeled through the force f_h, acting on both.

quantity	symbol	unit
Oscillator position	x_o	[m]
Hammer position	x_h	[m]
Penetration	x	[m]
Oscillator mass	m_o	[Kg]
Osc. el. constant	k_o	[N/m]
Osc. damp. coeff.	r_o	[N·s/m^2]
Hammer mass	m_h	[Kg]
f_h el. constant	k_h	[N/m$^{3/2}$]
f_h damp. weight	λ_h	[N·s/m$^{5/2}$]
f_h exponent	α	[adim]

Table 1: Variables, parameters and constants used throughout the document.

As for f_h we use a model originally developed by Marhefka and Orin (1999) for simulating contact of robotic systems with environment. They assume f_h to depend both on penetration $x = x_h - x_o$ and on penetration velocity $\dot{x}$ (elastic and dissipative components):

$$f_h(x, \dot{x}) = \begin{cases} k_h x^\alpha + \lambda_h x^\alpha \dot{x} & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x \leq 0 \end{cases} \qquad (1)$$

where α is given the value $3/2$ in (Marhefka and Orin 1999). The continuous-time equations for the hammer-resonator system are therefore:

$$\begin{cases} m_o \ddot{x}_o + r_o \dot{x}_o + k_o x_o = f_h(x, \dot{x}) \\ m_h \ddot{x}_h = f_h(x, \dot{x}) \end{cases} \qquad (2)$$

System (2) is discretized using the bilinear transformation:

$$s = 2F_s \frac{1 - z^{-1}}{1 + z^{-1}}, \qquad (3)$$

where F_s is the sampling rate. The numerical system is then obtained after some calculations:

$$\begin{array}{l} \boldsymbol{x_o}(n) = \bar{\boldsymbol{A}}_{\boldsymbol{o}} \boldsymbol{x_o}(n-1) + \bar{\boldsymbol{b}}_{\boldsymbol{o}}[f_h(n) + f_h(n-1)] \\ \boldsymbol{x_h}(n) = \bar{\boldsymbol{A}}_{\boldsymbol{h}} \boldsymbol{x_h}(n-1) + \bar{\boldsymbol{b}}_{\boldsymbol{h}}[f_h(n) + f_h(n-1)] \end{array} \qquad (4)$$

where

$$\boldsymbol{x_o} = \begin{bmatrix} x_o \\ \dot{x}_o \end{bmatrix}, \quad \boldsymbol{x_h} = \begin{bmatrix} x_h \\ \dot{x}_h \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x \\ \dot{x} \end{bmatrix},$$

$$\bar{\boldsymbol{A}}_{\boldsymbol{o}} = \frac{1}{\Delta_o} \begin{bmatrix} \Delta_o - 2k_o & 4F_s m_o \\ -4k_o F_s & 8F_s^2 m_o - \Delta_o \end{bmatrix},$$

$$\bar{\boldsymbol{A}}_{\boldsymbol{h}} = \frac{1}{\Delta_h} \begin{bmatrix} \Delta_h & 4F_s m_h \\ 0 & \Delta_h \end{bmatrix}, \qquad (5)$$

$$\bar{\boldsymbol{b}}_{\boldsymbol{o}} = \frac{1}{\Delta_o} \begin{bmatrix} 1 \\ 2F_s \end{bmatrix}, \qquad \bar{\boldsymbol{b}}_{\boldsymbol{h}} = -\frac{1}{\Delta_h} \begin{bmatrix} 1 \\ 2F_s \end{bmatrix},$$

$$\Delta_o = 4F_s^2 m_o + 2F_s r_o + k_o, \qquad \Delta_h = 4F_s^2 m_h,$$

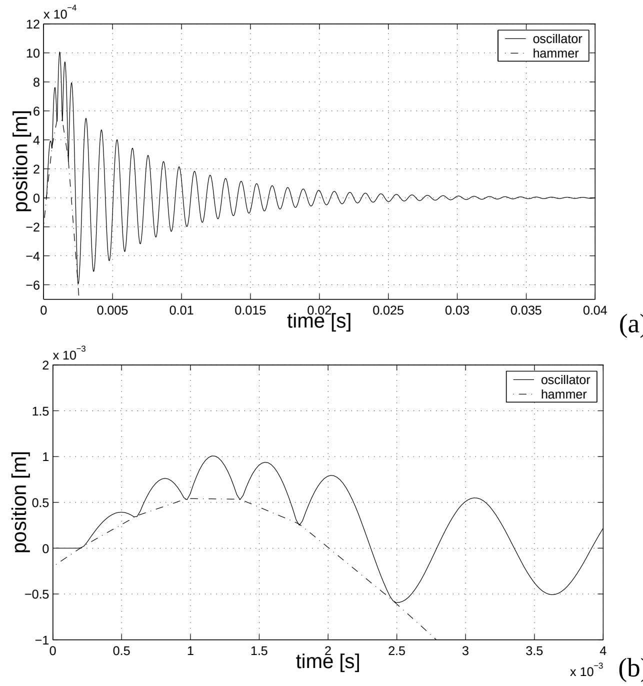

Figure 1: A sample from the model (hammer and resonator displacements): (a) waveform and (b) zoom on interaction during attack transient.

and where $f_h(n)$ stands for $f_h(\boldsymbol{x}(n))$. Due to this non-linear interaction, computational problems occur in the numerical system (4); namely, at each time step n the three variables $\boldsymbol{x_o}(n)$, $\boldsymbol{x_h}(n)$ and $f_h(n)$ have instantaneous mutual dependence. Borin, De Poli, and Rocchesso (2000) proposed a general method that allows to solve such a non-computable loop in an efficient and accurate manner. We do not discuss details of the method; suffice it to say that $\boldsymbol{x}(n)$ can be written as

$$\boldsymbol{x}(n) = \boldsymbol{p}(n) + \boldsymbol{K} f_h(n), \qquad (6)$$

where

$$\boldsymbol{K} = -\left(\frac{1}{\Delta_h} + \frac{1}{\Delta_o}\right) \begin{bmatrix} 1 \\ 2F_s \end{bmatrix} \qquad (7)$$

and $\boldsymbol{p}(n)$ is a computable vector (i.e. it is a linear combination of past values of $\boldsymbol{x_o}, \boldsymbol{x_h}$ and f_h). Substituting equation (6) in the non-linear contact force (1) and applying the implicit function theorem we can find f_h as a function of $\boldsymbol{p}$. Such a function can be precomputed and stored in a look-up table for efficient implementation. In this work however we use a slightly different approach: at each time step we first compute $\boldsymbol{p}$ and then find $f_h(\boldsymbol{p})$ iteratively using the Newton-Raphson method. Experimental observations show that the number of iterations is never higher than ten; as a consequence, the model can be suitable for real-time implementation on an ordinary DSP. An example of output from the model is shown in Fig. 1.

Using an accurate yet efficient physical model for subjective tests is advantageous over using damped sinusoids or other signal-based sound models, in that realistic interactions can be reproduced. As a result, complex and realistic attack transients can be kept in the stimuli, thus eliminating possible biases due to oversimplified test sounds.

3 Subjective experiments

A Matlab©/Octave implementation of the model (available at `www.soundobject.org`) was developed and used for synthesizing acoustic stimuli. Tests were performed, in which subjects were asked to listen to 100 sounds and to indicate what material each sound was coming from, choosing from a set of four material classes: rubber, wood, glass and steel. Each sound was played once and followed by a pause in which subjects had to choose the corresponding material class. We used 22 subjects, both expert and non-expert listeners, all reporting normal hearing. Subjects were not paid.

Other authors have performed similar experiments (Gaver 1993; Klatzky, Pai, and Krotov 2000). However, they used additive synthesis models, that do not include interaction with the hammer. The physical model used in this work provides more realistic attack transients.

3.1 Stimuli

The contact sound produced by hammer-resonator interaction can give information on both hammer and resonator properties. This is known as *phenomenical scission* in experimental psychology, and indeed Freed (1990) showed that hammer hardness can be perceived from percussive sounds. Since we are here interested in the resonator side, all of the stimuli were synthesized using the same hammer, i.e. the same set of coefficients $m_h, k_h, \lambda_h, \alpha$. The impact velocity of the hammer was fixed as well, thus providing a constant excitation.

Two acoustic parameters were chosen for controlling synthesis of stimuli: pitch (corresponding to the center frequency $f_o = \sqrt{k_o/m_o}/2\pi$ of the resonator) and quality factor q_o. This relates to decay via the equation $q_o = \pi f_o t_e$, where t_e is the time for the sound to decay by a proportion $1/e$. We used five equally log-spaced pitches from 1000 to 2000 [Hz] and 20 equally log-spaced quality factors from 5 to 5000; these extremal q_o values correspond to typical values found in rubber and aluminium, respectively. In a recent study on plucked string sounds, Tolonen and Järveläinen (2000) found that relatively large deviations (between -25% and $+40\%$) in decay time are not perceived by listeners. With the values we chose, the relative lower/upper spacings between q_o values are $-31\%/+44\%$.

The mapping from the two acoustic parameters and the physical parameters of the resonator was chosen as follows:

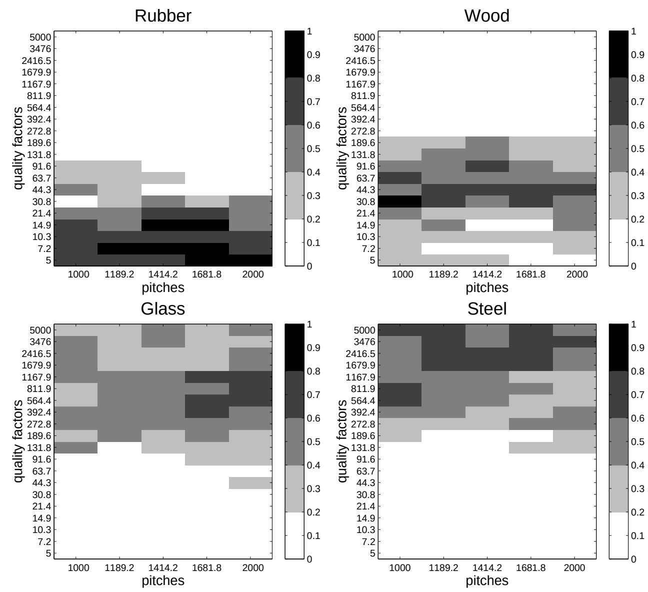

Figure 2: Proportion of subjects who recognized a certain material for each sound example.

- k_o was given a fixed value for all the stimuli, so that for each f_o the corresponding mass was computed as $m_o = k_o/(2\pi f_o)^2$.
- For each q_o, the corresponding damping coefficient was computed as $r_o = 2\pi f_o m_o/q_o$.

3.2 Results

Figure 2 summarizes results for the 22 listening subjects: it shows the proportion of subjects who assigned each sound to a given material cathegory, as a function of the two acoustic cues (pitch and quality factor). The intersubject agreements (proximity of the reponse proportions to 0 or 1) are qualitatively consistent with indications given by Wildes and Richards (1988), namely (1) q_o tends to be the most significant cue and (2) q_o is in increasing order for rubber, wood, glass and steel. A slight dependence on pitch can be noticed: rubber and glass tend to be preferred at high pitches, while wood and steel are more often chosen at low pitches.

Table 2 collects the q_o ranges for each material, each one computed as the minimum/maximum values where more than 50% of the subjects chose that material. The corresponding ranges for t_e are also given.

From both Fig. 2 and Table 2, the regions corresponding to rubber and wood appear clearly, while glass and steel are not well discriminated. Indeed, many subjects reported that the indication “glass” was not not immediately clear to them, since they could not guess what sound is produced by a bar made of glass. Another possible explanation has to do with the synthesis model: for long decay times (such as those of

Material	q_o [adim]	t_e [s]
rubber	$[5, 44.3]$	$[8 \cdot 10^{-4}, 1.41 \cdot 10^{-2}]$
wood	$[14.9, 131.8]$	$[2.3 \cdot 10^{-3}, 3.53 \cdot 10^{-2}]$
glass	$[189.6, 5000]$	$[4.34 \cdot 10^{-2}, 1.1254]$
steel	$[272.8, 5000]$	$[4.34 \cdot 10^{-2}, 1.5915]$

Table 2: Minimum and maximum values for q_o and t_e where more than 50% of the subjects chose a given material.

glass and steel) an exponential decay envelope is probably a too poor approximation, and more accurate description of the decay envelope is needed. Fig. 3 plots the same data as in Table 2 on the $q_0/(2\pi f_o), t_e$ plane, thus allowing direct comparison with the qualitative plot reported in Wildes and Richards (1988). Again, it can be noticed that rubber and wood are better discriminated, while glass and steel ranges are largely overlapping.

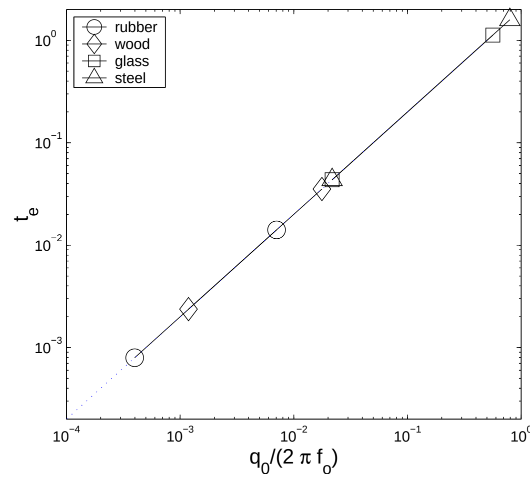

Figure 3: Distribution of materials on the $q_0/(2\pi f_o), t_e$ plane.

4 Discussion

Our findings allow to conclude that decay (or quality factor q_o) plays a much larger role than pitch f_o in material perception. Moreover, material classification by subjects is qualitatively in accordance with reported measures of internal friction coefficients for these material classes. This indicates that even an extremely simple physical model can elicit perception of material provided that it allows for control over salient acoustical cues.

However, intersubject agreement measures have shown that classification is unaccurate for high quality factors (glass and steel), thus suggesting that the overall decay time does not fully account for material properties and that control on decay shape would be needed in order to allow for a more accurate description of reality. Analysis and subjective experiments with real sounds have to be performed, in order to understand how classification is improved and to investigate wheter decay shape can play a role in helping material perception. This information can then be integrated in the model by treating the resonator as a non-linear oscillator: if the damping coefficient r_o is chosen not to be a constant parameter, and is instead taken as a function of the oscillator displacement, then non-exponential decay envelopes can be obtained.

5 Acknowledgments

The authors wish to thank Giorgia Tosco, who conducted the listening tests and helped in analysing the data. This work has been supported by the European Commission under contract IST-2000-25287 ("SOb - the Sounding Object": `www.soundobject.org`).

References

Borin, G., G. De Poli, and D. Rocchesso (2000). Elimination of Delay-free Loops in Discrete-Time Models of Nonlinear Acoustic Systems. *IEEE Trans. Speech Audio Process. 8*(5), 597–606.

Djoharian, P. (2000). Shape and Material Design in Physical Modeling Sound Synthesis. In *Proc. of the International Computer Music Conference*, Berlin, pp. 38–45.

Freed, D. J. (1990). Auditory Correlates of Perceived Mallet Hardness for a Set of Recorded Percussive Events. *J. Acoust. Soc. Am 87*, 311–322.

Gaver, W. W. (1993). How Do We Hear in the World? Explorations in Ecological Acoustics. *Ecological Psychology 5*, 285–313.

Gaver, W. W. (1994). Using and Creating Auditory Icons. In G. Kremer (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces*, pp. 417–446. Addison-Wesley.

Klatzky, R. L., D. K. Pai, and E. P. Krotov (2000). Perception of Material from Contact Sounds. *Presence 9*(4), 399–410.

Lutfi, R. A. and E. L. Oh (1997). Auditory Discrimination of Material Changes in a Struck-Clamped Bar. *J. Acoust. Soc. Am. 102*(6), 3647–3656.

Marhefka, D. W. and D. E. Orin (1999). A Compliant Contact Model with Nonlinear Damping for Simulation of Robotic Systems. *IEEE Trans Systems, Man and Cybernetics-Part A 29*(6), 566–572.

Rocchesso, D. (2001). Acoustic Cues for 3-D Shape Information. In *Proc. of the International Conference on Auditory Display*, Espoo, Finland.

Tolonen, T. and H. Järveläinen (2000). Perceptual Study of Decay Parameters in Plucked String Synthesis. Presented at the AES 109th Int. Conv., preprint no. 5205.

Wildes, R. P. and W. A. Richards (1988). Recovering Material Properties from Sound. In W. A. Richards (Ed.), *Natural Computation*, pp. 357–363. MIT Press.