Kees van den Doel
kvdoel@cs.ubc.ca
Department of Computer Science
University of British Columbia
Vancouver, Canada

Dinesh K. Pai
pai@cs.ubc.ca
Department of Computer Science
University of British Columbia
Vancouver, Canada

The Sounds of Physical Shapes*

Abstract

We propose a general framework for the simulation of sounds produced by colliding physical objects in a virtual reality environment. The framework is based on the vibration dynamics of bodies. The computed sounds depend on the material of the body, its shape, and the location of the contact.

This simulation of sounds allows the user to obtain important auditory clues about the objects in the simulation, as well as about the locations on the objects of the collisions.

Specifically, we show how to compute (1) the spectral signature of each body (its natural frequencies), which depends on the material and the shape, (2) the "timbre" of the vibration (the relative amplitudes of the spectral components) generated by an impulsive force applied to the object at a grid of locations, (3) the decay rates of the various frequency components that correlate with the type of material, based on its internal friction parameter, and finally (4) the mapping of sounds onto the object's geometry for real-time rendering of the resulting sound.

The framework has been implemented in a Sonic Explorer program which simulates a room with several objects such as a chair, tables, and rods. After a preprocessing stage, the user can hit the objects at different points to interactively produce realistic sounds.

1 Introduction

What information is conveyed by the sound of a struck object? Before reading further, you may want to try the following informal experiment. Tap on the physical objects around you like tables and file cabinets and listen to the sounds produced. First tap an object lightly and then try hitting it harder. Tap on wooden objects and metal objects. Tap on small objects like cups and telephones and large objects like tables and doors. Finally, for each object, try tapping it near the center, around the edges, from the side, and at other locations on the object. Most people can hear clear differences among the sounds in each of these cases.

When an object is struck, the energy of impact causes deformations to propagate through the body, causing its outer surfaces to vibrate and emit sound waves. The resulting sound field propagates through and also interacts with the environment before reaching the inner ear where it is sensed. Real sounds therefore provide an important ''image'' of various physical attributes of the object, its environment, and the impact event, including the force (or energy) of the impact, the material composition of the object, its shape and size, the

Presence, *Vol. 7, No. 4, August 1998,* 382–395

*This work was supported in part by grants from the Institute for Robotics and Intelligent Systems, the BC Advanced Systems Institute, and NSERC.

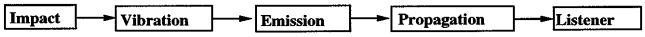

Figure 1. *Sound Pipeline.*

place of impact on the object, and finally the location and environment of the object.

In this paper we show how to synthesize realistic sounds from physical models of the geometry and material properties of objects.

The cognitive importance of realistic sounds is well known in the entertainment industry where sampled sound effects are added to enhance realism. As simulations become more interactive—for instance, in large architectural walkthroughs (Airey, Rohlf, and F. P. Brooks, 1990) and virtual reality—synthesizing realistic object sounds directly from physical models and rendering them in real time will be increasingly important.

The generation of sounds can be characterized as shown in Figure 1, which depicts the process as a pipeline similar to the sound-rendering pipeline of Takala and Hahn (1992). While this is a simplification, it indicates the major computational tasks in going from a collision event to the sound heard by a human ear. The focus of this paper is the initial stage of this pipeline: the computation of the sounds produced by vibrations of the object, which depend on the geometry of the object, its material properties, and the characteristics of the impact. For completeness we also briefly describe other parts of the pipeline such as impact dynamics and environment modeling.

1.1 Related Work

A review of the scientific and technological issues of auditory displays can be found in Durlach and Mavor (1995). An overview of the physical factors involved in producing natural sounds was presented in Gaver (1993b). Several synthesis methods for impact, scraping, and composite sounds were described in Gaver (1993a). However, no method was given to compute the free parameters of the synthesis methods, such as the set of eigenfrequencies, the relative amplitudes of the partials, and the bandwidths of the frequencies. With the methods described in this paper, the parameters of these synthesis methods can be computed.

Takala and Hahn introduced the concept of sound rendering for computer animation (Takala & Hahn, 1992). They associated a characteristic sound with each object that could then be rendered after filtering the sound to model the environmental effects. Recently they proposed ''timbre trees,'' which, like shade trees, provide a scene description language for sounds (Hahn et al., 1992). While Takala and Hahn (1992) indicated that the collision sounds could, in principle, be generated from vibration analysis, they were concerned mainly with the modulation of sound due to material properties. They did not synthesize sounds that account for the shapes of the colliding objects or the location of the collision on the objects.

Much progress has been made recently with the simulation of the filtering of sound by the human ear and head, giving the illusion of sound coming from various directions. Off-the-shelf audio hardware and software is already available to ''place'' a sound in space, using the ''head-related transfer function.'' For a review of this topic we refer to Begault (1994).

Offline computation of acoustical properties of performance halls, in the context of graphical visualization techniques, was investigated in Stettner and Greenberg (1989).

Wildes and Richards (1988) and Krotkov and Klatzky (1995) investigated the problem of recovering the material type from impact sounds, and it was proposed to use the internal friction parameter, which is an approximate material property, as a characteristic signature of the material. In this paper we invert this approach by relating the material properties of the synthesized sounds to the decay rates of the partials with the internal friction parameter.

For a more mathematically oriented example of the relation between shape and sound for membranes, we refer to the long-standing open problem ''Can one hear the shape of a drum?'', which was posed in 1966 (Kac, 1966). Recently counterexamples have been found (Gordon, Webb, & Wolpert, 1992).

A standard work on acoustics is Morse (1976). For a

Figure 2. *Waveform of a daf.*

book on vibration analysis we refer to Fahy (1985). For an application of vibration analysis to animation, see Pentland and Williams (1992). A survey of the use of the auditory channel in virtual reality is given in Durlach and Mavor (1995, Chapter 3). More general treatises on sound and hearing are Moore (1986) and Bregman (1990).

1.2 Overview

In this article we investigate the computation and rendering of sounds emitted by colliding bodies. When we strike an object such as a metal bar, we hear a brief transient or click, with a complex frequency spectrum, followed by a more-or-less sustained sound, which decays. In Figure 2 we show the waveform of a sound of a daf, which is a large drum. It shows the noisy onset of the sound, followed by a smoother part. The click or onset has some role in identifying the sound. For example, try listening to a recording of a flute played backwards. The sound is no longer as clearly recognizable as a flute, even though the sustained part of the sound is unchanged. See also Bregman (1990) and Moore (1986).

Nevertheless, most information about the nature of the object is present in the sustained part. To obtain this, we need to compute the vibrations of an object when it is struck, and compute the resulting sound emitted.

The sound made by a struck object depends on many factors, of which we consider the following:

1. *The shape of the object.* The propagation of vibrations in the object depends on its geometry. This is why a struck gong sounds very different from a struck piano string, for example. Shape and material together determine a characteristic frequency spectrum.
2. *The location of the impact.* The timbre of the sound—i.e., the amplitudes of the frequency components—depends on where the object is struck. For example, a table struck at the edges makes a different sound than when struck at the center.
3. *The material of the struck object.* The harder the material, the brighter the sound. We also relate the material to the decay rate of the frequency components of the sound through the internal friction parameter. (See below.)
4. *The force of the impact.* Typically, the amplitude of

Table 1. *Graphics-Audio Analogies*

Sound	Graphics
Material	Color
Spatialization	Perspective projection
Shape	Shape
Impact	Light Ray
Reverberation	Raytracing and radiosity

the emitted sound is proportional to the square root of the energy of the impact.

All these factors give important cues about the nature of an object and about what is happening to an object in the simulation.

Based on material and shape properties, we do a precomputation of the relevant characteristic frequencies of each object in Section 2. In Section 3 we then divide the boundary of the object into small regions and determine the amplitudes of the excitation modes if an impulsive force is applied to a point in this region.

This is similar to the tesselation of a surface for graphics rendering. The whole procedure is analogous to assigning a color to a surface and rendering it with some shading model. Although the role of sound synthesis differs in many aspects from the role of graphics, there are some useful analogies that can be drawn between the two. We show one possible correspondence in Table 1. Shape, impact location, and material are the focus of this paper.

In Section 4, we normalize the energies of the vibrations associated with the different impact points to some constant value, and scale them proportional to the impact energy when rendered.

The decay rate of each mode is assumed to be determined by the internal friction parameter, which is an approximate material property (Wildes & Richards, 1988; Krotkov & Klatzky, 1995). In effect, the decay rate of a component is assumed to be proportional to the frequency, with the constant determined by the internal friction parameter.

Besides the natural frequencies there is a brief transient, a ''click,'' mentioned before, which we model by a short burst of white noise.

After the preprocessing, a sound map is attached to an object, allowing us to render sounds resulting from impacts on the body. We discuss the structure of this map and a possible approach to reduce its storage requirements in Section 5.

We have constructed a testbed application, called the ''Sonic Explorer,'' which demonstrates the level of reality that can be achieved within this model. The Sonic Explorer is currently set up to precompute the impact sounds of several types of objects, incorporate them in a real-time interactive simulation, and render them in real time, using the audio hardware. This is described in Section 6. A picture of a virtual room environment is given in Figure 3.

2 Vibrating Shapes from Impact

We now introduce the framework for modeling vibrating objects. We will illustrate it with a rectangular membrane, but the framework is quite general; we have used it to generate sounds of strings, bars, plates, and other objects. The framework is based on the well-developed models in the literature on vibration or acoustics (e.g., Morse (1976)). For the calculus involved we refer to Strang (1986).

2.1 Vibration Modes from Shape

The vibration of the object is described by a function $\mu(x, t)$, which represents the deviation from equilibrium of the surface, defined on some region S, which defines the shape of the object. We assume that μ obeys a wave equation of the form

$$\left(A - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mu(x, t) = F(x, t) \qquad (1)$$

with c being a constant (related to the speed of sound in the material), and A a self-adjoint differential operator (Rudin, 1973), under the boundary conditions on ∂S. In the following we shall assume that the external force,

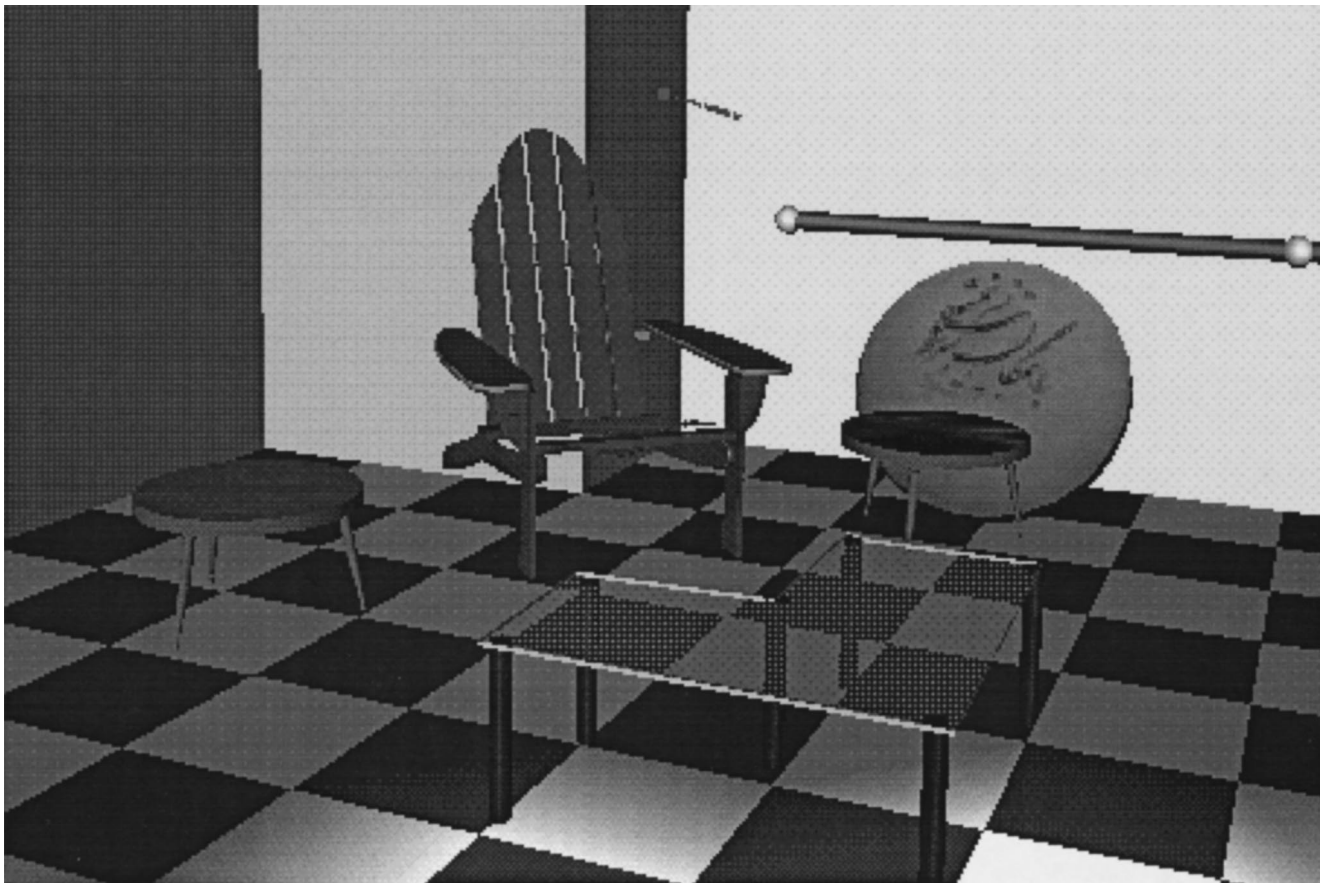

Figure 3. *A room modeled with the Sonic Explorer.*

$F(x, t)$, is zero. This type of equation is well known for various objects, and can be solved by analytic methods for simple geometries with homogeneous materials, as we will do below. See for example Morse (1976) and Lamb (1910).

Example: For the rectangular membrane we consider a rectangle $[0 - L_x, 0 - L_y]$ spanned by a membrane under uniform tension. For this case the operator A is given by

$$A = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

The boundary conditions are that $\mu(x, y, t)$ is fixed on the boundary of the membrane, i.e., the membrane is attached to the rectangular frame.

We will take the following initial value conditions:

$$\mu(x, 0) = y_0(x),$$

i.e., the surface is initially in configuration $y_0(x)$, and

$$\frac{\partial \mu(x, 0)}{\partial t} = v_0(x),$$

where $v_0(x)$ is the initial velocity of the surface.

The solution to equation (1) is written as

$$\mu(x, t) = \sum_{n=1}^{\infty} (a_n \sin(w_n ct) + b_n \cos(w_n ct))\Psi_n(x), \qquad (2)$$

where a_n and b_n are arbitrary real numbers to be determined by the initial value conditions. The values of w_n are related to the eigenvalues of the operator A under

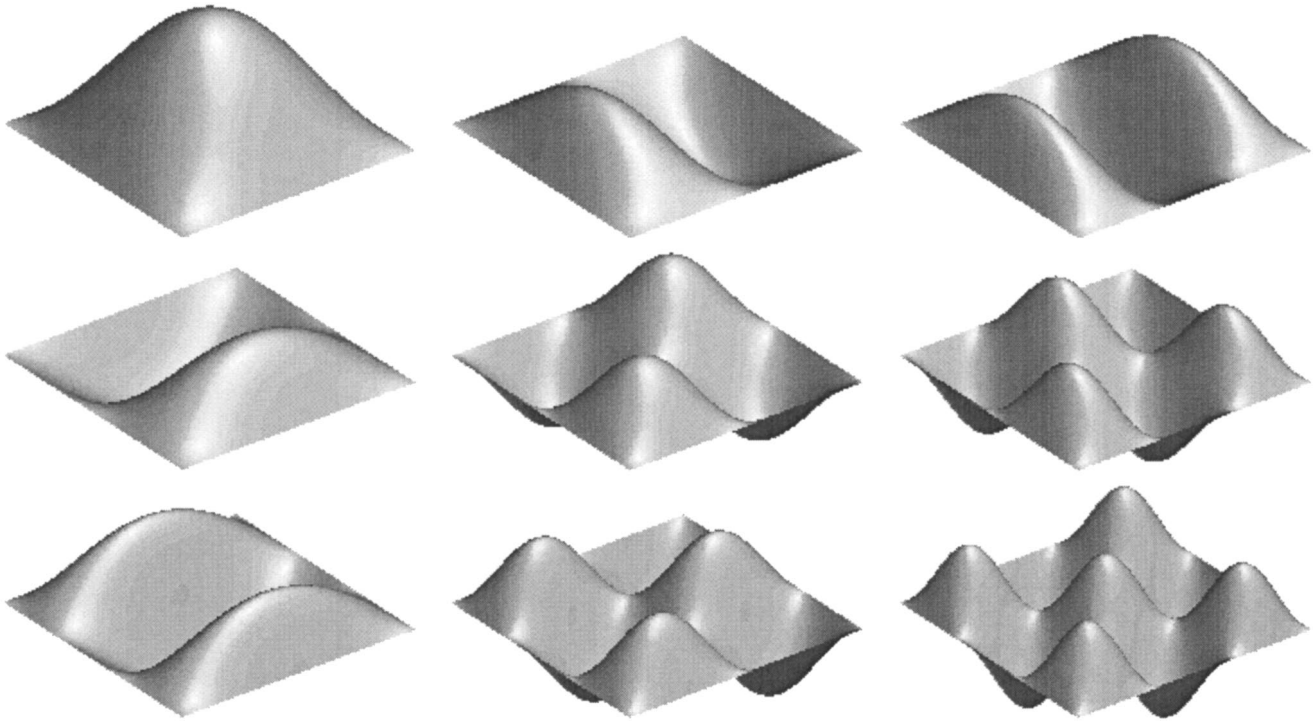

Figure 4. *First nine eigenfunctions of a square membrane.*

the appropriate boundary conditions (which we specify below), and the functions $\Psi_n(x)$ are the corresponding eigenfunctions. Thus we have

$$(A + w_n^2)\Psi_n(x) = 0. \tag{3}$$

The spectrum of a self-adjoint operator A is discrete, and the eigenfunctions are orthogonal. Their norm is written as

$$\alpha_n = \int_S \Psi_n^2(x) dV.$$

Example: For the rectangular membrane of dimension $L_x \times L_y$, the eigenfunctions and eigenvalues are most naturally labeled by two positive integers, n_x and n_y, and are given by

$$\Psi_{n_x n_y}(x, y) = \sin(\pi n_x x / L_x)\sin(\pi n_y y / L_y),$$

and

$$w_{n_x n_y} = \pi \sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2}$$

In Figure 4, we show the first nine eigenfunctions on a square membrane.

As equation (3) is linear, we can normalize the eigenfunctions $\Psi_n(x)$ such that α_n is independent of n, which often simplifies some of the algebra. Using the orthogonality of the eigenfunctions we can find the coefficients in the expansion given in equation (2) as

$$a_n = \int_S \frac{v_0(x)\Psi_n(x)}{c\alpha_n w_n} d^k x, \tag{4}$$

and

$$b_n = \int_S \frac{y_0(x)\Psi_n(x)}{\alpha_n} d^k x. \tag{5}$$

The time-averaged energy of the vibration is given by

$$E = \text{constant} \times \left\langle \int_S \left(\frac{\partial \mu(x, t)}{\partial t}\right)^2 \rho(x) d^k x \right\rangle, \tag{6}$$

where $\rho(x)$ is the mass density of the vibrating object. The $\langle\rangle$ indicates an average over time. If the mass is

distributed uniformly, we have

$$E = \text{constant} \times \sum_{n=1}^{\infty} \alpha_n w_n^2 (a_n^2 + b_n^2). \tag{7}$$

2.2 Mode Amplitudes from Impact Location

Next we compute the vibrations resulting from an impact at some point p, when the body is initially at rest.

The initial value conditions are taken to be

$$y_0(x) = 0, \tag{8}$$

and

$$v_0(x) = \delta(x - p), \tag{9}$$

with $\delta(x)$ the k-dimensional Dirac delta function.

We note that equation (9) is not strictly correct as an initial value condition. The reason is that the expression for the energy given in equation (6) involves the square of the time derivative of $\mu(x, t)$. But the integral of the square of the Dirac delta function is infinite. One symptom of this is that the infinite sum appearing in equation (2) does not converge. A mathematically more correct method would replace the delta function in the initial value conditions by some strongly peaking external force function, representing the impact on a small region of the object over a finite region and over a small but finite extension in time. However, this would complicate things quite a bit, and we would gain little in terms of more-realistic sounds. Therefore, we shall just assume an appropriate frequency cutoff in the infinite sum appearing in equations (7) and (2). Typically, we will use only the frequencies in the audible range. For more details and a more rigorous treatment of this problem for the special cases of the ideal string and the circular membrane, see Morse (1976).

Using equations (8) and (9), and substituting them in equations (4) and (5), we obtain the amplitudes of the vibration modes as a function of the impact location as

$$a_n = \frac{\Psi_n(p)}{c\alpha_n w_n}, \tag{10}$$

and

$$b_n = 0.$$

The energy of the vibration is determined by the impact strength. It will be used to scale the amplitudes of equation (10). The energy is given by

$$E = \text{constant} \times \sum_{n=1}^{n_f} \frac{\Psi_n^2(p)}{\alpha_n},$$

where n_f is determined by the frequency cutoff mentioned above.

Example: In Figures 5 to 7 we show the amplitudes a_n, graphed against the frequency of the modes (i.e., w_n) for a square membrane struck at the points (0.1, 0.1), (0.1, 0.4), and (0.5, 0.4). We use a coordinate system in which (0, 0) corresponds to the lower-left corner and (1, 1) corresponds to the upper-right corner of the membrane. We have taken the lowest frequency to be 500 Hz and taken the first 400 modes into account. We can see clearly that the higher frequencies become relatively more excited for strike points near the boundary of the membrane. In other words, the membrane sounds dull when struck near the center, and bright (or sharp) when struck near the rim.

The method outlined above is very general, and allows the computation of the vibrations under impact of any object governed by a differential equation of the form given in equation (1). This covers all vibrating structures, at least for small vibrations. For large external forces, non-linear effects will come into play, and the linear approximation will break down. For example, certain types of cymbals and Chinese gongs are designed explicitly to exhibit nonlinear behaviour as discussed by Fletcher and Rossing (1991).

The frequency spectrum w_n and the eigenfunctions $\Psi_n(x)$ can be computed analytically in a number of cases. In general, one has to resort to numerical methods. For membranes, the problem reduces to the solution of the Laplace equation on a given domain, which is a well-studied problem. We mention the method of particular

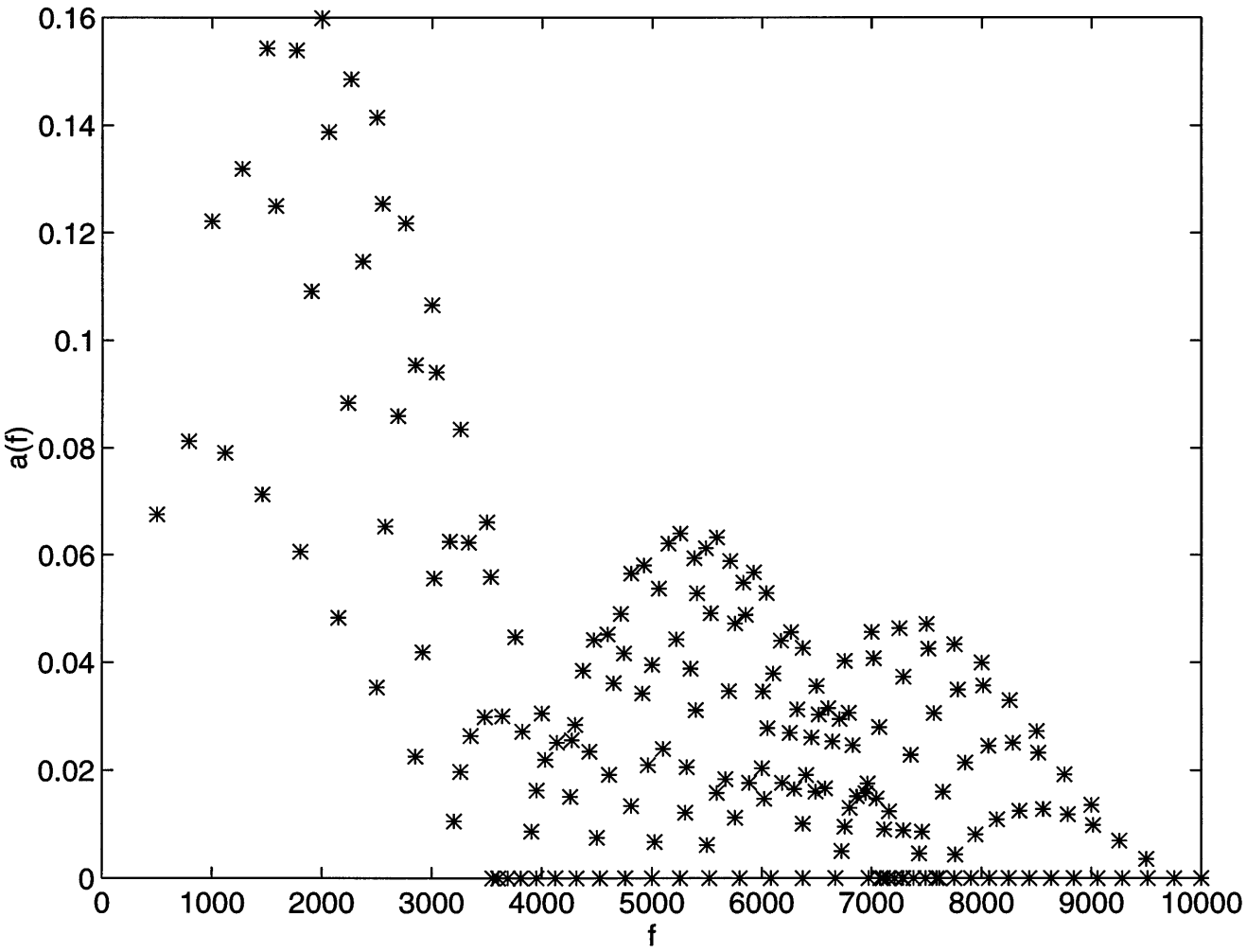

Figure 5. *Square membrane struck at (0.1, 0.1).*

solutions (Fox, Henrici, & Moler, 1967), which we have adapted for the example of the L-shaped membrane, described below in Section 6. For plates, the operator A is fourth order, and a more general finite element method can be used. See for example Johnson (1987).

3 Sound Sources from Vibrating Shapes

Suppose one has obtained the frequency spectrum and the eigenfunctions, as shown in Section 2.

What is the relation between the vibration of the object and the sound emitted? In general, the sound field around a vibrating body is very complicated and non-uniform. However, it is clear that the sound emitted can be described as a sum of monochromatic components with frequencies $w_n c$, and amplitudes a_n^S, which will depend on the location of the observer with respect to the object, as well as on the environment. Note that these are not identical to the amplitudes a_n of the vibration modes, which is why we distinguish them with the superscript ''S.''

As a first approximation, we will identify the coefficients a_n^S with the vibration amplitudes a_n, scaled with the inverse of the distance to the observer, as the amplitude decays inversely proportional to the distance.

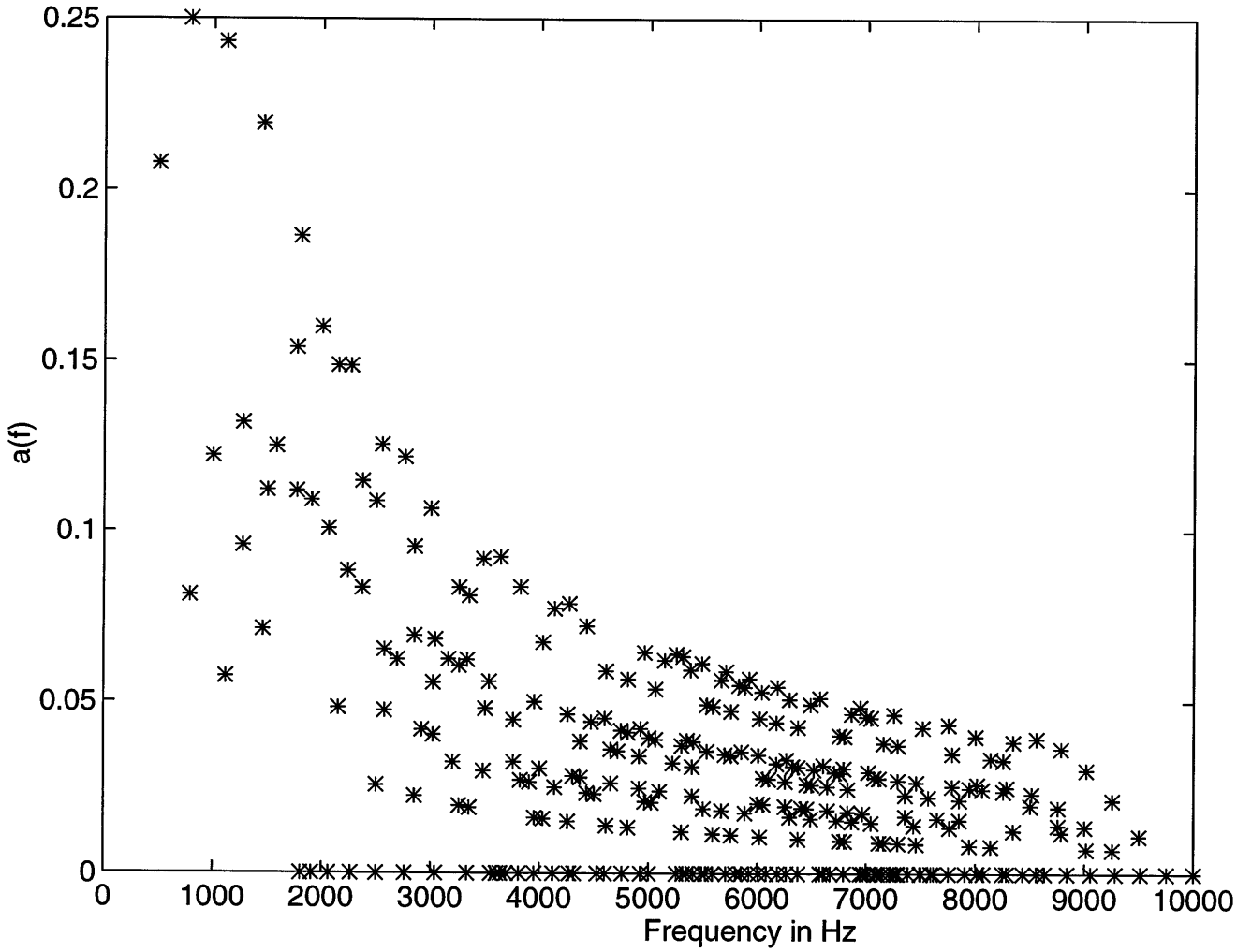

Figure 6. *Square membrane struck at (0.1, 0.4).*

This is not strictly correct, but we argue that it is reasonable as follows. Consider a vibrating plate. At some point above the plate, waves emerging from all locations on the plate arrive at this point. Some will be in phase, and some will be out of phase. This interference will depend very sensitively on the location of the observation point. However, in most real situations, the sound will not only arrive directly from the source, but also from reflections from the walls and other objects in the room. The total effect of this is to average out the phase differences, making the sound field less sensitive to the locations of the listener.

As a heuristic, we assume that the intensity (i.e., the energy) of the sound emitted in frequency w_n, I_n, is given by

$$I_n = E_n \int_S \Psi_n^2(x) = \text{constant} \times \Psi_n^2(p).$$

This seems reasonable, as it integrates the intensity of the vibration, but not the phase. This means we can identify a_n^S, the amplitudes of the heard sound, with the a_n given in equation (10), omitting the factor α_n. Note that since we assumed that the eigenfunctions are normalized so that the α_n are independent of n, this does not matter.

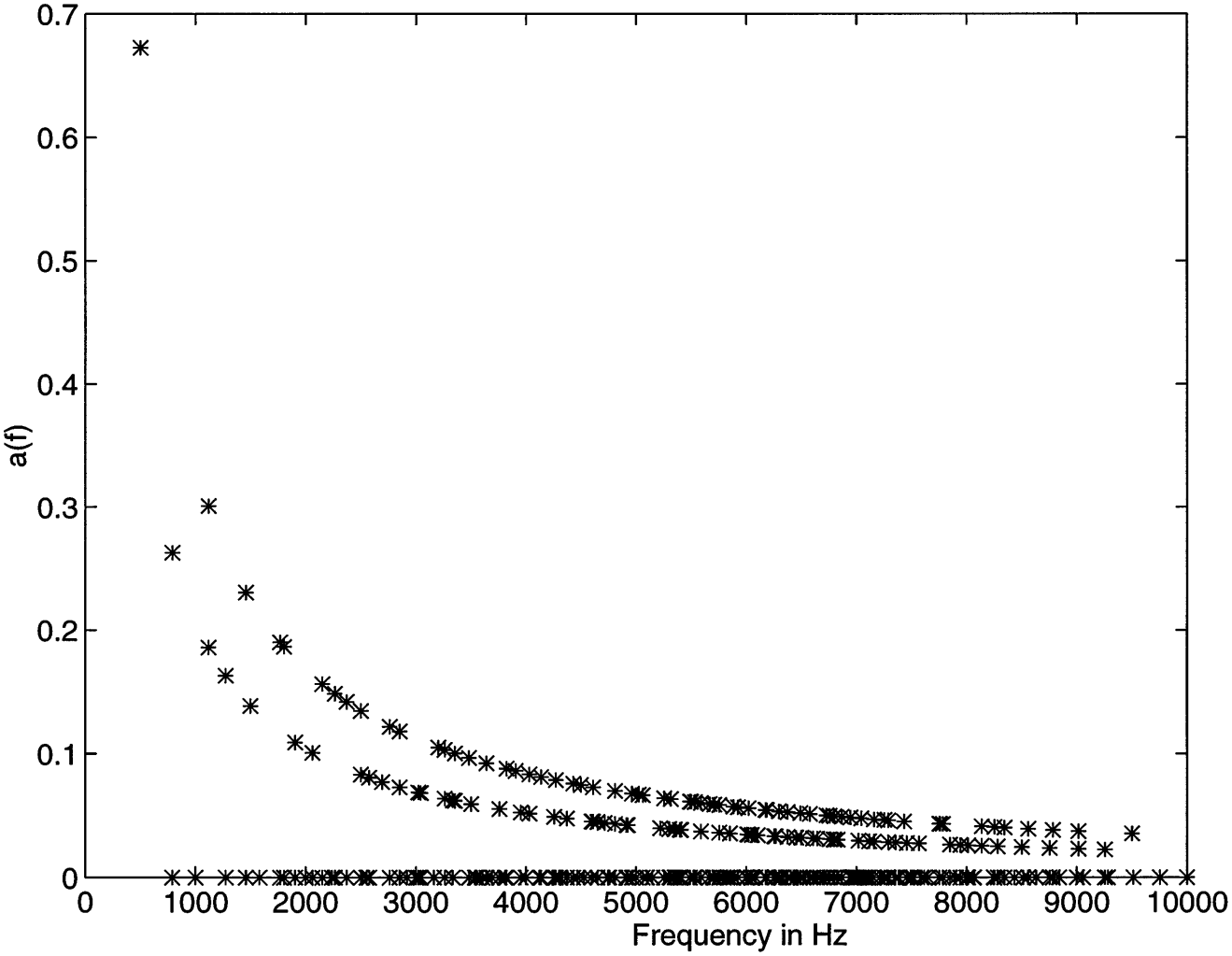

Figure 7. *Square membrane struck at (0.5, 0.4).*

Finally, we obtain the amplitudes a_n^S as

$$a_n^S = \frac{E_{\text{impact}} \Psi_n(p)}{w_n Q(p) d}, \tag{11}$$

with d the distance from the sound source, E_{impact} the energy of the impact, and

$$Q(p) = \sqrt{\sum_{i=1}^{n_f} \Psi_n^2(p)},$$

with n_f a suitable frequency cutoff. Of course, the a_n^S values are only defined up to a multiplicative constant (corresponding to the volume setting of the audio hardware).

For a more detailed treatment of the radiation of vibrating plates, we refer to books on vibration analysis (Shabana, 1991a, 1991b; Fahy, 1985).

4 Sounds and Material Properties

When the object is struck, each frequency mode is excited with an initial amplitude a_i, which depends on where the object is struck. The relative magnitudes of the amplitudes a_i determines the ''timbre'' of the sound. Each mode is assumed to decay exponentially, with de-

cay time

$$\tau_i = \frac{1}{\pi f_i \tan\phi}, \qquad (12)$$

where ϕ is the internal friction parameter. The internal friction parameter is roughly invariant over object shape, and depends on the material only. Wildes and Richards (1988) proposed a method to identify an object's material type from the sound it emitted when struck, by extracting the internal friction parameter of the material via equation (12). Such a model is also used by Takala and Hahn (1992) to simulate object sounds. Krotkov and Klatzky (1995) reported some experiments where it was concluded that a rough characterization of material was indeed possible. However, the internal friction parameter is only approximately invariant over object shape. See also Wert (1986).

To emulate external damping of the object, we add an overall decay factor of $e^{-t/\tau 0}$. This also allows us to adjust the length of the emitted sound, while maintaining its ''material character,'' which is determined by ϕ.

So we assume the sound wave $p_S(t)$ to be given for $t \geq 0$ (for $t < 0$ it is zero) by

$$p_S(t) = e^{-t/\tau_0} \sum_{i=1}^{n_f} a_i^S e^{-t f_i \pi \tan\phi} \sin(2\pi f_i t), \qquad (13)$$

with the amplitudes a_i^S given in equation (11), and

$$f_n = \frac{w_n c}{2\pi},$$

with w_n determined by equation (1).

5 The Sound Map

In the preprocessing stage we first compute the frequency spectrum f_i, and then the excitation spectrum a_i, under a suitably normalized impact (i.e., with fixed energy) for a number of locations on the surface. We can then compute digital samples of the sound wave $p_S(t) + p_C(t)$ for these locations and store them for playback during the real-time simulation. This is somewhat analogous to texture mapping in computer graphics. Alternatively, if appropriate real-time, sound-synthesis hardware is available, only the model parameters need to be stored, as the sound can be computed on the fly. An interpolation of the timbre spectrum a_i^S between precomputed locations is also obvious to implement.

As digital sound samples take a lot of space, we do not want to store any more than we need. So one may ask how many points on the surface need to be computed. In general, the timbre of the sound changes non-uniformly over the surface. For example, a string sounds ''dull'' when plucked near the center, and becomes dramatically brighter when excited near the endpoints. In this case, one would need a denser set of points near the ends.

Given two sounds S_1 and S_2, a measure $d(S_1, S_2)$ is needed, that tells us how different sound S_1 ''sounds'' from sound S_2, such that if $d(S_1, S_2) < d_0$, with d_0 a threshold (depending on the individual), S_1 and S_2 cannot be distinguished. Perception of timbre is a complex subject—see for example Bregman (1990) and Moore (1986) for a discussion—so we cannot expect to be able to formulate such a sound-distance measure easily and accurately.

As an initial proposal we take the sonic distance $d(S_1, S_2)$ between two sounds to be

$$d^2(S_1, S_2) = \sum_{i=1}^{n_f} S(f_i)(H(\log(E_i^1 / E_0)) - H(\log(E_i^2 / E_0)))^2,$$

where E_i^r denotes the energy contribution of the ith mode of sound r (=1, 2), i.e., $E_i^r = (a_i^{Sr})^2 f_i^{\,2}$. We take the logarithm of the energy, as the human ear is sensitive to the logarithm of intensity (measured in decibels). The function $H(x)$ is zero for $x < 0$, and x otherwise. The constant E_0 represents the lowest sound level that can be heard, so the term $H(\log(E_i^r / E_0)$ vanishes if $E_i^r < E_0$. The function $S(f)$ models the sensitivity of the ear to frequency. Without loss of generality we take $0 \leq S(f) \leq 1$. The function $S(f)$ has to be determined through psychoacoustic experiments.

We have ignored the ''masking'' effect, which changes the sensitivity curve of the ear in the presence of other

stimuli. One could also argue that the threshold energy E_0 depends on frequency.

We leave a refinement of the measure d as a topic for future research.

6 Results

For a number of cases, the vibration equations can be solved exactly, and the sound-model parameters can be found in analytic form. Some of these models are already very useful in giving a realistic ''feel'' of the collision sounds.

We have solved a number of configurations (using well-known solutions to the corresponding wave equation), and have computed the sounds for some instances of these problems. They are:

1. The taut string. This is the simplest example of a vibrating system. The eigenfunctions are simple sine functions. The sound becomes brighter for impacts near the ends of the string. The frequency spectrum is harmonic, i.e., all frequencies are integer multiples of the lowest (fundamental) frequency. The amplitudes a_n are inversely proportional to n, for large n, in contrast to a plucked string, considered in Takala and Hahn (1992), where they decay as $1/n^2$. This is one factor accounting for the difference between a piano and a guitar sound, for example.
2. The rigid bar. For the rigid bar, the operator A appearing in equation (1) is given by

$$A = -\frac{\partial^4}{\partial x^4}\cdot$$

   As A is a fourth-order operator, we need to specify four boundary conditions. We have computed a clamped-clamped bar, i.e., the bar is rigidly attached at both ends. The boundary conditions are

$$\Psi_n(0) = \Psi_n(1) = \left(\frac{d\Psi_n}{dx}\right)_{x=0} = \left(\frac{d\Psi_n}{dx}\right)_{x=1} = 0. \quad (14)$$

   The frequency spectrum is less dense than for the string, and it is not harmonic. This is due to the different nature of the restoring forces on a bar.
3. The rectangular membrane. This geometry gives the simplest solution to the wave equation for a two-dimensional geometry. The sound spectrum is extremely dense, giving a rich complex sound.
4. The circular membrane. This corresponds to the vibrations of a drum, ignoring the effects of the surrounding air on the drum membrane. The eigenfunctions are Bessel functions, and the eigenfrequencies can be computed as the zeros of Bessel functions.
5. The circular plate. This is one of the few cases where the two-dimensional plate equations can be separated, which allows an analytic solution. The eigenfunctions are a combination of Bessel functions and modified Bessel functions. We have considered a plate clamped rigidly at the boundary. The spectrum is much less dense than for the circular membrane. As with the bar, this is due to the larger restoring forces in a plate as compared to a membrane.
6. The L-shaped membrane. A membrane supported by a domain consisting of three unit squares in the shape of an L does not allow an analytic solution of the wave equation. This problem has received some attention in the literature, as the resulting boundary value problem requires some refined numerical methods. We have computed the eigenfunctions and the spectrum with an adaptation of the method of partial solutions (Fox et al., 1967). As an aside, we note that the first eigenfunction features prominently on the cover of the MATLAB reference guide (Matlab-reference, 1992).

We have implemented a Sonic Explorer, which allows us to create a graphical scene with objects of the above types. The sounds associated with a grid of points on each object are precomputed and stored as digital samples. For this we have created a synthesis tool that implements various vibrational objects with relatively simple geometry. More-complicated geometries would demand a finite element method, for which an interface to some existing finite element modeling could be cre-

ated. A problem with this approach is that for complicated materials and geometries there may not be enough information available about the elastic properties of the material and the geometry to allow a precomputation of the vibration modes from a model.

Although designed for precomputation, the precomputation of the sounds can be done almost in real time, depending on how many vibration modes we wish to incorporate. On a 200 Mhz SGI Indy, we can do roughly 40 modes in real time. We are currently investigating a real-time synthesis approach that will dispense with the need to precompute samples.

By clicking on a point on a surface, a drumstick hits the object at the specified point and the sound is rendered. The purpose of this Explorer is to investigate the realism of the sounds by integrating them in a graphics environment, so the user can integrate visual and audio cues. Our implementation allows the optional use of a set of HRTF filters to spatialize the sounds completely in three dimensions using the public-domain Kemar HRTF filters (Gardner & Martin, 1994). However, as we have no specialized hardware, this introduces unacceptable delays. A simple left-right localization based on interaural delays (Begault, 1994) does perform in real time.

7 Discussion and Conclusions

We have developed a framework to add an auditory component to real-time simulation environments. We have focused on aspects of sounds that enhance the level of realism and on effects that provide the user with useful auditory clues about the simulated environment. Within our framework different materials such as wood or metal have a recognizable signature, determined through the internal friction parameter, which determines the decay rate of the different frequency components.

The shape of the object and its structure (such as a plate versus a membrane) determines a characteristic frequency spectrum. This frequency signature facilitates the recognition of an object by its sound. A metal bar, for example, rings with a sparse nonharmonic spectrum, of which the higher modes decay rapidly. This is how we recognize its sound.

The sound of an impact also depends on where an object is hit, and this also provides useful information about the environment. Though the frequency signature is the same over the object, the relative amplitudes change. Generally, an object sounds brighter (i.e., more upper partials are excited) when struck near an edge than when struck near the center. Recall for example Sherlock Holmes, who taps on the walls to find a secret compartment. By precomputing a grid of sounds for each object, we incorporate this location information.

We have derived some general formulas to determine the spectrum from the shape, material, and amplitudes as a function of the impact site. We have considered systems that can be described by a linear wave equation on some domain, which covers all vibrating solids for sufficiently small vibrations.

We have not taken the directionality of the sound emitted into account. To do so involves radiation theory—see for example Fahy (1985)—but this falls in the next stage of the pipeline depicted in Figure 1.

These ideas were implemented in a Sonic Explorer, which allows the user to explore the environment by hitting various objects with a virtual drumstick. We find that it is essential to listen to the synthesized sounds in a visual and tactile context to judge their realism. So the Sonic Explorer is intended both as a research tool for the creation of synthesized sounds as well as a demonstration of the enhanced realism that can be achieved by adding an auditory component to the scene that consists of more than the playback of canned sounds.

An interesting direction for future research is to integrate the Sonic Explorer with a haptic interface (Durlach & Mavor, 1995, Chapter 4). We expect that a virtual reality environment with three sensory feedback channels (sight, hearing, and touch) will provide a significant enhancement.

As digital samples tend to take large amounts of storage space, the question of how many sounds need to be stored for each object comes to mind. For this we need to have some criterion to determine if two similar sounds can be distinguished by the user. The sounds associated with different impact sites are very similar and

differ only in the relative amplitudes of their frequencies, and it is not clear how to construct such a measure. We have made a tentative proposal which needs to be refined with psychoacoustic experiments before it can be used to optimize the sound grids associated with the bodies. It is also possible to store just the model parameters (the partials, their decay rates, and the amplitudes as a field on the object surface) and use these to synthesize the sounds in real time. This is currently under investigation by us.

Continuous sounds such as scraping, sliding, and rolling can be obtained with a generalization of our method, as the relevant quantities, the eigenfrequencies, the amplitudes as a function of location, and the decay rates of the partials determine the response to such interactions in principle. Taking into account the finite duration in time, and the extension over a finite space, of impacts is another important topic for further research.

References

Airey, J., Rohlf, J., & Brooks, F. P. (1990). Towards image realism with interactive updates in complex virtual building environments. *Computer Graphics: Proceedings of 1990 Symposium on Interactive 3D Graphics, 24*(2), 41–50.

Begault, D. R. (1994). *3-D sound for virtual reality and multimedia.* London: Academic Press.

Bregman, A. S. (1990). *Auditory scene analysis.* London: MIT Press.

Durlach, N. I., & Mavor, A. S. (Eds.). (1995). *Virtual reality, scientific and technological challenges.* Washington, D.C.: National Academy Press.

Fahy, F. (1985). *Sound and structural vibration. radiation, transmission and response.* London: Academic Press.

Fletcher, N. H., & Rossing, T. D. (1991). *The physics of musical instruments.* New York: Springer-Verlag.

Fox, L., Henrici, P., & Moler, C. (1967). Approximations and bounds for eigenvalues of elliptical operators. *SIAM J. Num. Analy., 4,* 89–102.

Gardner, B., & Martin, K. (1994). *HRTF measurements of a KEMAR dummy-head microphone* (Tech. Rep. No. 280). MIT Media Lab Perceptual Computing.

Gaver, W. W. (1993a). Synthesizing auditory icons. *Proceedings of the ACM INTERCHI '93* (pp. 228–235).

Gaver, W. W. (1993b). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Psychology, 5*(1), 1–29.

Gordon, C., Webb, D., & Wolpert, S. (1992). Isospectral plane domains and surfaces via Riemannian orbifolds. *Invent. Math., 110,* 1–22.

Hahn, J. K., Geigel, J., Lee, J. W., Gritz, L., Takala, T., & Mishra, S. (1992). An integrated approach to motion and sound. *Journal of Visualization and Computer Animation, 6*(2), 109–123.

Johnson, C. (1987). *Numerical solutions of partial differential equations by the finite element method.* Cambridge: Cambridge University Press.

Kac, M. (1966). Can one hear the shape of a drum? *Mer. Math. Mon., 73*(II), 1–23.

Krotkov, E., & Klatzky, R. (1995). Robotic perception of material: Experiments with shape-invariant acoustic measures of material type. *Preprints of the fourth international symposium on experimental robotics, ISER '95.* Stanford, California.

Lamb, H. (1910). *The dynamical theory of sound.* London: Edward Arnol.

Matlab reference guide. (1992). The MathWorks, Inc.

Moore, B. C. J. (1986). *An introduction to the psychology of hearing.* London: Academic Press.

Morse, P. (1976). *Vibration and sound* (4th ed.). American Institute of Physics for the Acoustical Society of America.

Pentland, A., & Williams, J. (1992). Good vibrations: Modal dynamics for graphics and animation. *Proc. SIGGRAPH'89, ACM Computer Graphics, 23*(3), 215–222.

Rudin, W. (1973). *Functional analysis.* New York: McGraw-Hill.

Shabana, A. A. (1991a). *Theory of vibration, Volume I: An introduction.* London: Springer-Verlag.

Shabana, A. A. (1991b). *Theory of vibration, Volume II: Discrete and continuous systems.* London: Springer-Verlag.

Stettner, A., & Greenberg, D. P. (1989). Computer graphics visualization for acoustic simulation. *Proceedings of SIGGRAPH'89, Computer Graphics, 23*(3), 195–206.

Strang, G. (1986). *Introduction to applied mathematics.* Wellesley-Cambridge Press.

Takala, T., & Hahn, J. (1992). Sound rendering. *Proceedings of SIGGRAPH'92, ACM Computer Graphics, 26*(2), 211–220.

Wert, C. A. (1986). Internal friction in solids. *Journal of Applied Physics, 60*(6), 1888–1895.

Wildes, R. P., & Richards, W. A. (1988). Recovering material properties from sound. In W. Richards (Ed.), *Natural computation.* Cambridge, Massachusetts: The MIT Press.