Chapter 12

Devices for manipulation and control of sounding objects: the *Vodhran* and the *InvisiBall*

Roberto Bresin, Kjetil Falkenberg Hansen and Sofia Dahl
Kungl Tekniska Högskolan – Department of Speech, Music, and Hearing
Stockholm, Sweden
roberto.bresin@speech.kth.se, hansen@speech.kth.se,
sofia.dahl@speech.kth.se

Matthias Rath
Università di Verona – Department of Computer Science
Verona, Italy
rath@sci.univr.it

Mark Marshall, Breege Moynihan
University of Limerick – Interaction Design Centre
Limerick, Ireland
Mark.T.Marshall@ul.ie, Bridget.Moynihan@ul.ie

12.1 Introduction

In the case of physics-based sound models, the control of the synthesized sound is straightforward. As shown in previous chapters, these models offer direct access to sound source characteristics. In this chapter two applications that focus on direct control of sounding objects are described, the *Vodhran* and the *Invisiball*. These two applications have a common characteristic: they realize the control of a sound model by mean of physical gestures. This characteristic facilitated the design of the applications and the interface between the control model and the sound model.

The first application is the implementation of an augmented version of the *Bodhran*, called the *Vodhran*. The *Bodhran* is an Irish traditional percussive instrument. It is a relative simple instrument but characterized by good, even complex, expressive possibilities, and therefore it is a good starting point for the design of a control model. In the next section we illustrate control techniques based both on data obtained from measurements of gestures of drum player, and on the use of gesture controllers.

The second application presented in this chapter is based on a game, called the *Invisiball*. Users can move an invisible ball placed on an elastic surface by pressing the surface with a finger. The ball will roll towards the position occupied by the finger. The rolling of the ball is communicated to users by a rolling sound. The sound stops as soon as the ball has reached the finger position.

12.2 The virtual *Bodhran*: the *Vodhran*

In order to have a successful interface it is widely known that it is preferred to employ a metaphor that the end user of the artifact is familiar with. In the following pages we illustrate an application that aims to provide users with an expressive virtual musical instrument, based on the traditional *Bodhran*: the *Vodhran*. This is not designed entirely to simulate the *Bodhran*, but to provide an instrument which can be played in a similar way, and which creates a recognizable Bodhran-like sound. This instrument is to be an extension of the original, allowing for additional playing techniques and styles, which could not be accomplished with the real instrument.

In the *Bodhran*, sound is emitted in response to a stick (beater) beating the skin, generally controlled with the right hand. Sound is modulated/damped by pressing the inside of the *Bodhran* with the left hand. The beater is held

Figure 12.1: Traditional *Bodhran*.

loosely in the hand and is moved/controlled primarily by wrist action (rotation). The contact with the *Bodhran* is made with alternative ends of the beater in rapid succession. The damping factor depends on the left-hand pressure, and often a dynamic/colorful pitch range can be achieved by continuous damping control obtained by putting the left hand in counter direction to the beater. There is a variety of different applications of the damping factor employed by the musicians, e.g., fingers only, side of hand only and so on (see Figure 12.1).

In the following sections four control modalities are discussed. In all of them the same sound model implementing the *Bodhran* was used. This sound model is described in the next section.

12.2.1 Sound generation

The sound generation mechanism for the *Vodhran* is based on the (approximate and simplified) modal description of the drum and on the robust numerical solution of a nonlinear stick-membrane interaction model [11].

This approach aims at an integrated "sound object", oriented at the real drum, including different forms of a player's interference, rather than a perfectly realistic reproduction of isolated signals. The superordinate term of "modeling" points out this difference to sample-based sound production.

Resonator-Membrane

The technique of modal synthesis [2] forms a well-suited basis for our efforts for several reasons (see Appendix 12.A for a description of the modal synthesis technique):

- real-time implementation requires a sound synthesis technique that delivers convincing results even under preferably low computational expenses - as opposed to, e.g., waveguide techniques;
- at the same time the possibility of dynamical interactions with the player, during changing position/velocity/direction of the stroke or variable damping gestures, must be provided (this demand addresses the basic drawbacks of sample playback techniques);
- the synthesis parameters should be comfortably estimable under physical and perceptional specifications as, e.g., tuning or intensity of damping. Modal parameters are particularly close to terms of auditory perception and can be estimated from guidelines to the aspired sound properties.

The practical procedure of defining and adjusting a modal synthesis unit modelled on the *Bodhran*

The sound of a *Bodhran*, struck at 6 equally spaced positions from the centre to the edge, was recorded. A relatively hard wooden stick with a rounded tip (resulting in a very small area of contact) was used and the strike was performed approximately perpendicular to the membrane with least possible external force applied by the player in the moment of contact (loosely held, resembling a pendulum motion).

The subsequent interaction between the stick and the membrane can be approximately seen as a one-dimensional impact, and the resulting curve of interaction force is close to a fed-in ideal impulse. Each excited membrane movement is therefore with good approximation handled as an impulse response of the resonator at the point of striking; its transform in turn approximates the frequency response. The modal parameters where finally extracted from these "frequency responses" according to the theory of the modal description [11]: Peaks in the magnitude response curves mark frequencies of resonant modes, decay factors are calculated from Short Time Fourier Transform (STFT) values at two different temporal points; the position dependent weighting factors [11] of the modes are given by the different peak levels (at each resonant frequency).

It is to be noted, that the described procedure shows of course many inaccuracies. In addition to the mentioned idealizations (spatially distributed strike interaction is not an impulse at one point, stick is not a point mass, not absolutely free from external forces) the signal recorded through a microphone does not match the membrane movement at one point, and peak values in the frequency response do not immediately display modal frequencies and weighting factors.

Our overall goal, though, is not a perfect imitation of the *Bodhran* sound, but a practically functioning expressive sound object inspired by the *Bodhran* in its main behaviour and sound characteristics. Under these premises even a first implementation sketch, under evaluation of only the lowest 16 frequency peaks, gave an agreeable result (b.t.w., also to the probably less computer euphoric ears of interviewed *Bodhran* players). An important final step consists of the tuning of synthesis parameters controlled by ear (in a final implementation together with sensory feedback), which remains the ultimate judging instance in our case.

The impact model

Our model of the impact interaction assumes a point mass with a certain inertia representing the stick, and an interaction force that is a non-linear function of the distance between stick and membrane [11].

The instantaneous cross relationship between the variables of the modal states (displacements and velocities), the state of the striker, and the force (expressed by an equation that contains the mode weights depending on the strike position), can—due to its "high degree" non-linearity—not be resolved analytically. Instead of inserting an artificial delay to solve this non-computability on the cost of new introduced errors (as commonly done), we are using an iterative procedure to numerically approximate the implicit variable dependency[1] [24].

Implementation.

The sound engine for the virtual *Bodhran* is implemented as a module for real-time sound processing software PureData (`pd`) [197]. Control messages (as incoming user "commands") are handled with the temporal precision of an "audio buffer" length; practical values for the `pd` internal buffer size are, e.g.,

[1] A (computationally) low cost version of the impact also exists; here the interaction force term is linear—excluding stickiness phenomena—and the resulting equation is solved explicitly, leading to a faster computation of variables.

32 or 64 samples, providing a temporal accuracy of 32/44100 s ≈ 0.75 ms or $64/44100$ s ≈ 1.5 ms. This is perfectly acceptable even for this sensitive practical realization of a percussion instrument; latencies of sound cards/drivers and gesture control systems are much more problematic.

We end up with a flexible expressive algorithm that models interaction properties, i.e., the weight of the stick, hardness, elasticity and "stickiness" of the contact surfaces, as well as player's controls, i.e., strike-velocity and -position and damping, and efficiently runs in real-time in a standard environment.

12.2.2 Controlling the Vodhran

The sound model described in the previous section was controlled with four different modalities. These involve both software and hardware controllers (see Figure 12.2) , and are presented in the following.

Control by modelling real players

A `pd` patch implementing the model of professional percussionists was used for controlling the sound model of the *Bodhran*. This control model was obtained by the measurements described in chapter 7. The measured striking velocities of two players (subjects S1 and S2) were used as excitation velocities for the sound model. The `pd` patch, a "metronome rhythm generator", plays metronome beats with a striking velocity that is affected by dynamic level (*pp*, *mf*, or *ff*), chosen playing style (symphonic or drumset) and tempo. Any time an accent can be generated, and depending on the current tempo, dynamic level and playing style, the striking velocity exciting the model will be changed accordingly. The shift between the two playing styles can also be generated in real-time.

To enable a swift change between the different playing styles and dynamic levels these control parameters were connected to an external custom made switch board. The switches were connected to the PC through a device to take analogue signals through the parallel port, the `pico` [1].

The control for tempo was mapped to the Korg KAOSS pad, a MIDI controller with a touch pad. The x-axis on the touch pad controls the tempo between 100 and 250 beats per minute. The y-axis controls the fundamental frequency of the physical model. By using the touch sensitivity, from the touch pad it is possible to start the metronome by touching the pad and stopping it on release. In this way the onset and offset of beats becomes immediate and very

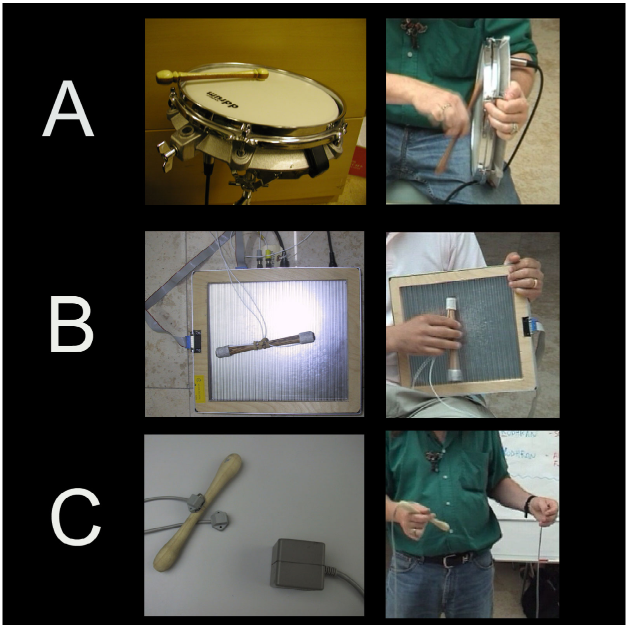

Figure 12.2: The three hardware controllers used to implement the *Vodhran*. (A) The ddrum, (B) the Radio Baton, and (C) the Polhemus Fastrack.

flexible, resulting in less resemblance to a metronome. The rhythm generator was played in a live concert at Centro Candiani in Mestre-Venezia, Italy, on the 20th of June 2002.

The Radio Baton as control device

In the same concert, the metronome rhythm generator was used together with a more direct control device for playing, Max Mathew's Radio Baton [25].

The set-up used in the concert is shown in Figure 12.3. The set-up used the radio baton with the two radio transmitters at each end of a *Bodhran* stick, and the antenna was played with the stick as a "normal" *Bodhran* by the player, Sandra Joyce. The position of each end of the stick versus the antenna controlled the playing position of the sound model. During the concert the dampening and the fundamental frequency of the model was also controlled through

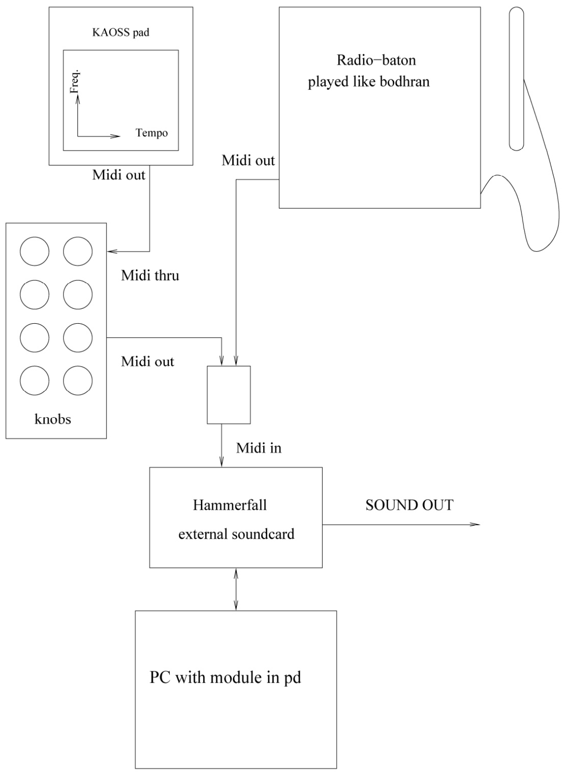

Figure 12.3: Set-up used for the concert at Centro Candiani in Mestre-Venezia, Italy, on the 20th of June 2002. A player used a *Bodhran* stick, modified with cables to transmit radio signals, and played the antenna like a *Bodhran*. While the position across the antenna controlled the striking position of the model, the fundamental frequency and dampening was manipulated by another person through the Korg KAOSS pad and a set of knobs. The same person also played the "metronome rhythm generator" and then used the x-axis of the KAOSS pad to control the metronome tempo.

devices such as the KAOSS pad and Doepfer pocket dial knobs. During the concert this was not done by the player herself, although this could be remedied by using foot switches that would leave the hands free for playing.

The ddrum as control device

In addition to the control devices described above, the ddrum [61] was used to play the sound model of the *Bodhran*. The ddrum is a commercial available electronic drumpad and the MIDI velocity out from the control unit was used to excite the sound model. For the pad used there was also MIDI poly-aftertouch, which was used to control the dampening of the model. The ddrum is a nice interface to play the model because of it's tactile feedback to the player and the lack of cables to the sticks used for playing.

Control by tracking gestures in real time

A fourth modality of control was based on tracking users body gestures. The process of tracking users movements requires a means of capturing gestures in real time, and extracting the relevant features of the gesture. This requires some form of input device which can take a gesture as input and extract the relevant characteristics of this gesture.

Traditionally, the use of gesture as input involves the use of computer vision techniques, such as the recording of gesture with a camera, and the tracking of the movement of a person over time. These systems have a number of limitations, for instance they can be too slow for real-time use, and do not cope well with tracking more than one user. Also, they might not be accurate enough for present requirements.

So, it was decided to use a Polhemus Fastrak[2] device in order to track the users movement. This system tracks the position of up to four small receivers as they move in the 3D space, with respect to a fixed-point electromagnetic transmitter. Each sensor returns full six degree-of-freedom measurements of position (X, Y, and Z Cartesian coordinates) and orientation (azimuth, elevation, and roll). The device connects to a computer through its RS-232 port, and operates at a speed of up to 115.2 Kbaud, with an accuracy of 0.03" (0.08 cm) RMS for the X, Y, or Z receiver position, and 0.15° RMS for receiver orientation, and a resolution of 0.0002 inches/inch of range (0.0005 cms/cm of range), and 0.025°.

[2]Polhemus Fastrak `http://www.polhemus.com/fastrak.htm`

The software system. The software architecture for the control of the Fastrak is made up of two layers. For the *Vodhran*, these layers have been implemented as `pd` externals.

At the bottom layer there is a driver, which allows the applications to communicate with the system, and to get the bare position and orientation information from each sensor.

This layer communicates with the Fastrak over the RS-232 connection using a binary protocol, in which all the information received from the Fastrak is encoded in binary format. This allows encoding of the complete position and orientation data for a receiver in just 14 bytes, and allows for receiving 1000 position readings per second. In conjunction with the Fastrak's update rate of 8 ms per receiver, this means that applications will receive every update from each sensor, and so record all available data on the gesture. This high update rate also means that applications should not have to wait for data at any stage, so that latency is minimized.

This bottom layer is constantly working, recording all current position and orientation data from each sensor, and making them available to the next layer up. This higher layer acts as a form of middleware for the system, taking the raw position data from the lower layer and extracting any necessary characteristics from this data. For instance, it calculates velocity, direction of movement, direction changes, etc. This allows applications to map specific characteristics of the gesture to parameters of the models directly.

Motion tracking. The Fastrak senses, for each activated receiver, two vectors:

- the position of a fixed point located inside the (receiver) sensor, referred to as receiver origin, is measured relative to the transmitter origin (which is analogously a fixed point inside the transmitter) in a system of (standard 3D) orthonormal coordinates (the transmitter system, which is in orientation rigidly connected to the transmitter);
- the orientation of the receiving sensor is expressed in term of the three angles (exactly their cosines) azimuth, elevation and roll. These values characterize three turning motions executed successively moving the transmitter system axes onto according receiver axes.

The coordinate change from receiver- to transmitter-system is accomplished by addition of the translation vector and multiplication with a 3D transformation matrix. While the translation vector is, of course, just the above position vector (or its negative), the transformation matrix has entries that are products of the direction cosines.

Calculations. For computational comfort the Fastrak allows immediate output of the transformation matrix entries themselves, so that the effort of their external calculation can be saved. What remains to be done is the matrix multiplication and (eventually) the translation (that is, an addition).

`pd` modules have been written, that execute matrix/vector multiplications and additions; they are combined in a `pd` patch to calculate the "absolute" position (i.e., relative to the transmitter) from receiver coordinates.

Geometry. In a first approach the receiver sensor has been rigidly fixed to a drumstick. The transmitter is used as an orientation system: It may be fixed to a frame indicating the imaginable or physical drum plane or might be possibly fixed to a part of the player's body that is not involved in the playing movement. Points on the drumstick can now be located within the transmitter system via the aforementioned `pd` patch. Of course the coordinates relative to the receiver, that is, dimensions of the stick and the fixing of the receiver, must be known to this end.

The tips of the stick are of course not of such small dimension that the point of contact with an assumed membrane (which would at first sight seem a necessary information) is the same for every stroke; it rather highly depends on the angle between stick and plane/direction of the stroke at contact time. A constant striking-point may not even approximately exist, since the head of a Bodhran stick may be around 2-3 cm in diameter. Tracking the position of a whole portion of the stick's surface and checking for distance from a striking-plane (or in a striking-direction) is of course computationally highly expensive in a real-time context.

It can though be noticed (see Figure 12.4) that for many Bodhran sticks (including the one we used) the tip is approximately spherical. As a consequence, a point inside the stick, at the center of an approximating sphere, is found at a nearly constant distance from a struck membrane for every stroke, independent from the striking-angle (and the actual point of contact). For all triggering strategies that we took into account, it suffices to track the position of such a point.

System design. In order to establish the requirements for the system, in terms of usability, methods of interaction and sound quality, a daylong workshop was held. Three expert *Bodhran* players, each one with their own distinct and different styles and techniques, took part. The players also had various amounts of experience in the use of technology in performance.

The players were asked to perform different traditional rhythms with the

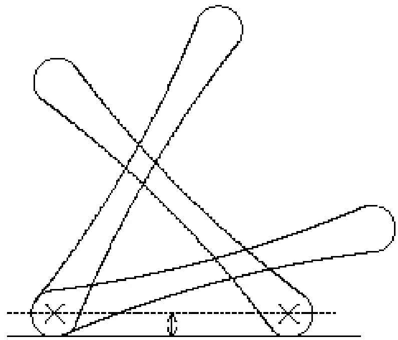

Figure 12.4: *Bodhran* stick hitting a membrane at different angles. Note that the touching point varies, while the center point of the sphere forming the surface of the stick lies at the same distance from the plane for each stroke.

sensor attached to the beater (see Figure 12.5), and the results were recorded in order to further analyse the gestural patterns involved in *Bodhran* playing. Video data were also gathered for further analysis.

The results of the analysis of this data are being used to determine any common movements, gestures, or techniques performed by the players, so that the control parameters of the model may be extended in order to allow more interaction for the players.

By examining the data in this way, and continuing to extend the model, we ensure that the overall goal of the project is met, in that a virtual instrument is created, which can be played like a *Bodhran*, but is not limited to the physical characteristics of a *Bodhran*.

Analysis results

The analysis of the workshop has led to a number of additional features in the design of the system. The nuances of how a *Bodhran* is played are very complex, involving both the left and right hands, with the right hand beating the skin, while the left hand can be used to damp certain modes of the skin.

Damping. Left-handed damping was used by all players, and in some cases was used to produce very complex tone changes, even to provide a melody from the rhythmic beating of the skin.

Figure 12.5: Workshop session. Sandra Joyce playing the *Bodhran* with a sensor attached to the beater.

This damping effect has a major role during the playing of the *Bodhran* and, as such, must be entirely incorporated into the system. Currently the sound model does contain a facility to damp the skin, but only at a single point at any given time. Also, the damping portion of the model would have to be coupled to the Fastrak hardware, and a Fastrak sensor attached to the left hand of a player, to allow them to damp the virtual Bodhran in a similar way to the real instrument.

Tactile feedback. During the course of the workshop, when the players were asked to use the beater from the virtual *Bodhran* to create a simple beat by striking a virtual plane, it was noticed that some players required more tactile feedback from the system than others.

While some players were able to hold the beat, using just the beater and listening to the generated sound, one in particular found this difficult. The addition of a physical plane of reference, which matches the virtual one, was found to alleviate this problem. This opens to some further investigation, to determine whether or not a physical reference is required, and, if so, the form of this reference.

Frame of reference. Another point which was raised by this workshop was that of a frame of reference for the instrument. Currently, the system uses a fixed reference point, which does not move with the player. In order for any virtual instrument to be internalised there needs to responsive in a non-arbitrary way and the modification was made for an extension to expressivity and also to allow deliberate musical control on the part of the musician in terms of control sensitivity and control selection. Development based on the GRIP instrument—"gradually expand and personalize their gestural 'vocabulary' without losing acquired motor skills and therefore gradually add nuances to their performances without needing to adapt to the instrument" [182].

To meet this requirement, the system must allow the players to move naturally, as they would while playing the actual instrument. This would allow players to add to their movement range, without infringing on their standard movements.

To enable this, the frame of reference for the system needs to move with the player, so that should they turn or lean as they play (which most players seem to do), the system will continue to function normally, in its new frame of reference.

12.2.3 Conclusions

The control systems that were described in this chapter cover three different kinds of interaction with a sound model based on different levels of abstraction; (1) algorithmic control, based on measurements of real players gestures, (2) interaction by mean of percussive instruments, that allow players using traditional gestures with haptic feedback, (3) use of a gesture tracking system providing continuous controls to the sound model. In particular, the system based on the Fastrack is a novel gesture-based interface to a sound model, which is used as a virtual musical instrument. By basing the design of the system on the real instrument, and by involving players in the analysis and design of the system, it was our hope to create an instrument, which captures the intrinsic characteristics of the real instrument.

However, by not restricting the system to just the characteristics of the real instrument, and by not necessarily tying the instrument to any physical presence, an instrument which allows the players to expand their playing vocabulary can be created.

12.3 The *InvisiBall*: Multimodal perception of model-based rolling

"Rolling" sounds form a category that seems to be characteristic also from the auditory viewpoint. Everyday experience tells that the sound produced by a rolling object is often recognizable as such, and in general clearly distinct from sounds of slipping, sliding or scratching interactions, even of the same objects.

Prior experiments showed that listeners can discriminate differences in size and speed of wooden rolling balls on the basis of recorded sounds [120]. Results from perceptual experiments, performed in the framework of the Sounding Object Project (see chapter 4), demonstrated that listeners are able to perceive the size and weight of a ball under different conditions. These conditions were: (1) steel ball falling on a wooden plate, (2) wooden ball falling on a ceramic plate.

12.3.1 Controlling the sound of an invisible ball

The impact model implemented in the framework of the Sounding Object Project (see chapter 8) was controlled by the mechanic equations of a rolling ball. The sound model was controlled by simply feeding it with the X, Y, and Z coordinates of a target position in a 3D space.

In the demonstrative experimental setup presented in this paper, the space is given by the dimensions of a haptic interface that was constructed for controlling the Max Mathews' Radio Baton [25]. The default baton-like controller was substituted by a rubber thimble-controller. For this purpose a "thimble-sender" device was constructed (see Figure 12.6). The radio signal is sent by a fingertip. The interface is made of stretching material and it was placed over the receiving antenna. Finger position in the 3D space is detected in real-time and it is used by the algorithm controlling the rolling movement of a ball. By pushing the membrane with the "thimble-sender", users can make the ball rolling towards the finger by moving it from rest position in the 3D space (see Figure 12.7). The position of the rolling ball as a projection on the XY plane, i.e. as seen from above, is visualized on the computer screen. The position is represented as a colored disk assuming colors in the red-range at high speed (hot ball) and blue-range at low speed (cold ball).

This new interface allows users an interaction by using three different types of feedback:

- Acoustic - the sound model of the rolling ball.

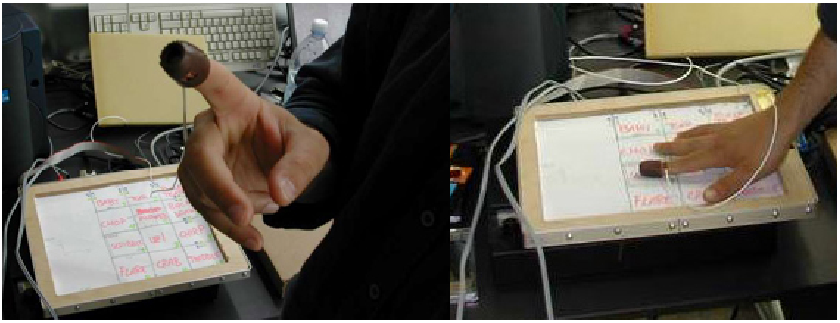

Figure 12.6: The “thimble-sender”; a controller for sending a radio signal from the fingertip to the receiving antenna.

- Haptic - control of the position of the ball by pressing the elastic membrane with a finger.
- Visual - graphical projection of position and speed of the ball.

12.3.2 Pilot experiment: multimodal perception test

The interface described in the previous section was used to test the quality of the sound model from a perceptual point of view. The main goal was to investigate the degree of reality of the rolling-ball sound when controlled with this interface. Since the interface allows three different feedback, acoustic, haptic and visual, three different experiments were conducted. The three experiments were run in parallel at the same time so that 3 subjects at a time could listen to exactly the same sound stimuli, in the way explained in the following.

Subjects and procedure There were twelve subjects, 6 female and 6 male. Their average age was 30. All subjects were researchers or students at the Speech Music Hearing Department of KTH, Stockholm.

Subjects listened to the sound stimuli individually over infrared headphones adjusted to a comfortable level. In this way three subjects at a time could take part in the experiment without seeing each other. The stimuli for each group of three subjects were organized as follows: (a) one subject could only listen to the sound stimuli, (b) another subject had both audio and visual stimuli, and (c) a third subject had both haptic and audio stimuli. Each subject was instructed to estimate the degree of realism for each sound stimulus. The responses were

Figure 12.7: Haptic interface for the control of the Max Mathews' Radio Baton. The interface is placed over the receiving antenna. Finger position in the 3D space is detected in real-time.

recorded on a scale on paper, from 0 to 10, with "unrealistic" and "realistic" as extremes. Stimuli were presented twice in a random order.

Experiment 1: haptic and acoustic feedback

Subjects. Subjects had an acoustic feedback through headphones and an haptic feedback from the finger-controlled interface presented in the previous section.

Stimuli. Nine sound model set-ups were used. They were obtained by combining 3 sizes of the ball with 3 damping values, thus producing 9 different set-ups. In this way the influence of both size and damping could be tested on the classification of the corresponding sounds stimuli as being realistic. These 9 set-ups were presented twice to the subjects. This correspond to a factorial design (3 sizes) x (3 damping) x (2 repetitions). By controlling the haptic interface with the "thimble controller" subjects produced 9 acoustic stimuli with 2 repetitions, for a total of 18 stimuli, of the duration of about 20 seconds each. The sound of each stimuli was that of a rolling ball.

Experiment 2: visual and acoustic feedback

Subjects. In this experiment subjects had an acoustic feedback through headphones and a visual feedback from a computer screen.

Stimuli. The acoustic stimuli were those produced simultaneously at the same time by the subject controlling the haptic controller in Experiment 1. In addiction a visual stimuli was synchronized with the acoustic stimuli. The visual feedback was presented on a computer screen and it represented the real-time moving position of the rolling ball in the 2D space with speed/colour as third dimension.

Experiment 3: acoustic feedback only

Subjects. In this third experiment subjects had only an acoustic feedback through headphones.

Stimuli. The acoustic stimuli were the same as those in Experiment1 and Experiment 2. They were produced simultaneously by the subject controlling the haptic controller in Experiment 1.

12.3.3 Results and discussion

In the following a preliminary analysis of the results from the three experiments is presented.

A repeated measurements ANOVA was conducted on answers collected in Experiment 1. There was no significant effect of either size or damping parameters. The sound stimuli with low and medium damping produced by the model of a rolling ball of medium size were classified as being more realistic (see Figure 12.8). A high subjective variability emerged for stimuli obtained with a medium value of the damping factor. Values ranging from 0 to 8 were recorded.

A repeated measurement ANOVA was conducted on answers collected in Experiment 2. There was no significant effect of either size or damping parameters. Nevertheless a closer observation of the results suggests that subjects tended to classify as more realistic the sound stimuli associated to low and medium damping and to large size of the rolling ball (see Figure 12.9). Also, subjects gave a higher preference rate to the most realistic stimuli as compared with best ratings given by subjects in Experiment 1. A high subjective variability emerged for stimuli obtained with a large value of the damping factor. Ratings ranging from 0 to 7 were recorded.

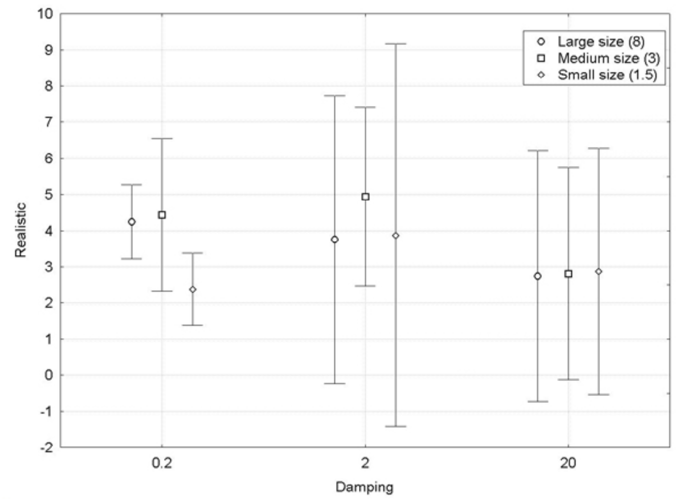

Figure 12.8: Effect of damping and size as resulted from the analysis of the responses in experiment 1.

A repeated measurement ANOVA was conducted on answers collected in Experiment 3. There was no significant effect of neither size or damping parameters. A preliminary observation of the results suggests that subjects with only the acoustic feedback tended to classify as more realistic the sound stimuli associated to low and medium damping and large size of the rolling ball (see Figure 12.10). Subjects gave a higher preference rate to the most realistic stimuli as compared with best ratings given by subjects in Experiment 1. These results are comparable to those obtained in Experiment 2. A high subjective variability emerged for stimuli obtained with a small value of the damping factor and large size of the ball. Values ranging from 0 to 9 were recorded.

A repeated measurement ANOVA was conducted on all answers collected in the three experiments. The main results suggest a significant effect of the ball size parameter, $F(2, 22)=6.6175$, $p=0.00562$, and a significant effect of the damping factor of the sound model, $F(2, 22)=5.5417$, $p=0.01124$ (see Figures 12.11 and 12.12). There was no significant interaction between the size of the ball and the damping parameter.

The average rate for each feedback modality across all stimuli is presented in Figure 12.13. The "acoustic & haptic" modality, Experiment 1, resulted to

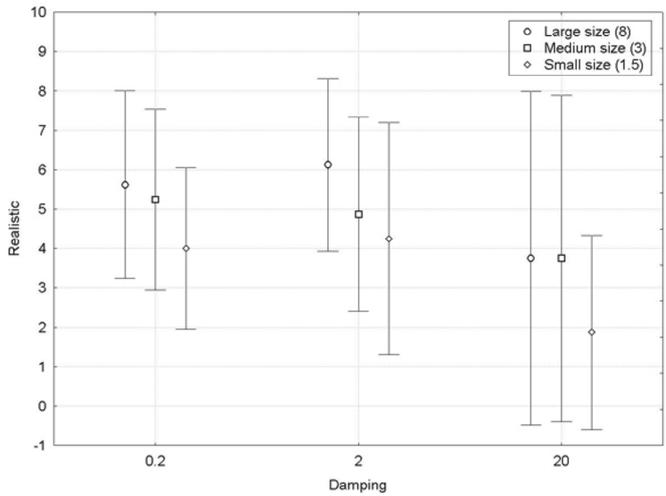

Figure 12.9: Effect of damping and size as resulted from the analysis of the responses in experiment 2.

be the worst one, and the "acoustic & visual", Experiment 2, was classified as best, according to the ratings given by the subjects in the three experiments. There is no significant difference between average ratings of the "acoustic" modality and of the "acoustic & visual" modality.

Responses to stimuli can be averaged through all three experiments, as shown in Figure 12.14. It can be observed that, in all damping conditions, stimuli with small size are classified as less realistic. In general, stimuli with both medium size and medium damping were classified as more realistic.

12.3.4 Conclusions

In this section we proposed a new device for controlling physics-based sound models through direct manipulation of an haptic interface, the thimble controller based on the Radio Baton system. This interface allows a gestural control of a sounding object in a 3D space by direct manipulation. The "realistic" property of sounds produced by acting on the interface was analyzed in three experiments.

Main result was that sound stimuli corresponding to balls with large and

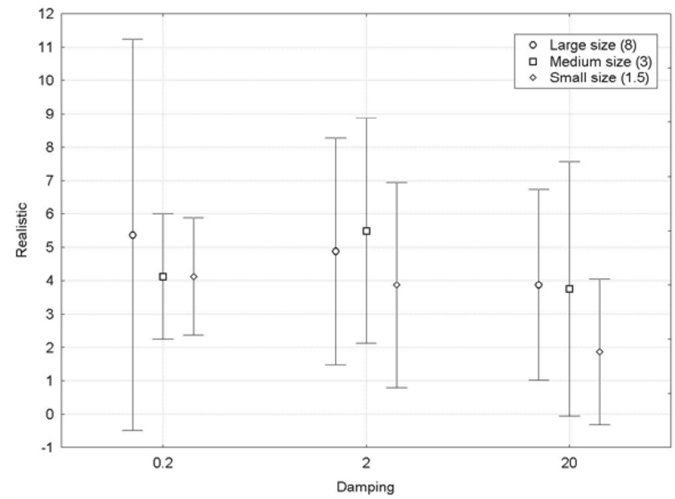

Figure 12.10: Effect of damping and size as resulted from the analysis of the responses in experiment 3.

medium size and low and medium damping were classified as more realistic.

As an overall result, sound stimuli were classified as more realistic by subjects using only “acoustic” feedback or “acoustic and visual” feedback. It seems that this is due to the difficulty in controlling the haptic interface and the sound metaphor associated to it. Some of the subjects reported that it was difficult to imagine the ball rolling towards their finger. Nevertheless some of the subjects in Experiment 1 gave high rates to stimuli corresponding to balls of medium size. These results suggests that the haptic controller and/or the testing application can be better designed.

12.4 Acknowledgments

Our thanks to those workshop participants who helped in the analysis and design of the system - Tommy Hayes, Robert Hogg and Sandra Joyce.

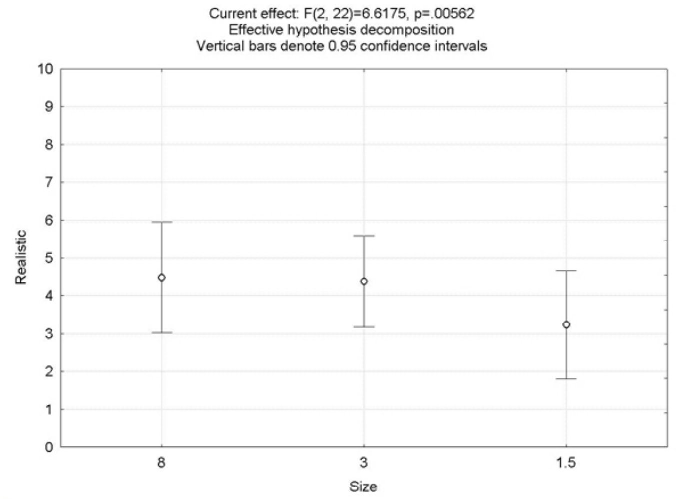

Figure 12.11: Effect of size obtained by averaging across all three experiments.

12.A Appendix – Modal synthesis

Modal Synthesis is based on the description of the behavior of a resonating object in coordinates that are not displacements and velocities (or other physical state variables, e.g., flow/pressure) at spatial positions, but in terms of modes (these discrete *modal coordinates* correspond to the eigenfunctions of the differential operator describing the system. In the case of a finite linear system of lumped elements the modal coordinates can be calculated from the matrix connected to the finite system of differential equations describing this finite case) [2].

While lacking the immediate graphic meaning of the spatial state variables, the modal description of a vibrating object is of strong advantage in certain respect. First of all, the development of the system along each modal axis is independent of its state and development along the other modal coordinates (the differential equation of the system splits into a series of independent equations).

The free oscillation (that is, the evolution without external) of each mode can be analytically described, even in simple form: Each mode follows the evolution of an exponentially decaying sinusoid of a fixed frequency. The

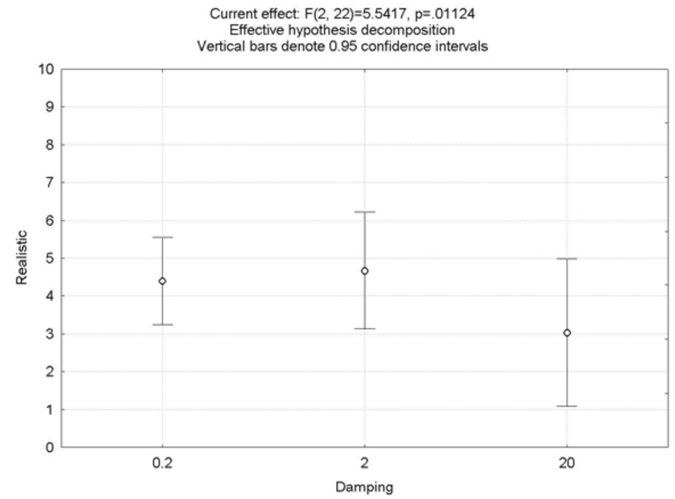

Figure 12.12: Effect of damping obtained by averaging across all three experiments.

corresponding resonance behavior (i.e. the frequency response) is that of a lowpass filter with a peak around this modal (or resonant) frequency (the bandwidth of this peak is proportional to the inverse of the mode's decay time).

The spatial state variables of the system can, of course, be reconstructed from the modal states through a linear transformation: Concretely, the movement of a specific "pickup point"—giving the sound picked up in that point—is a weighted sum of the movements of the modes; conversely, an external input to the system at the pickup point (i.e., an external force) is distributed to the distinct modes.

Summing up, the full modal description of the system reduces to a series of mode frequencies with according decay factors. A series of weighting factors represents each interaction point (or practically speaking, each interaction point of possible interest). The transform function of the system with specific interaction—or pickup points—is finally a weighted sum of the above described resonance filters (just as the impulse response is the weighted sum of the described sinusoids). This finally shows the immediate (acoustic) perceptual significance of the parameters of the modal description that we gain in trade for the missing ostensive meaning of the modal coordinates themselves.

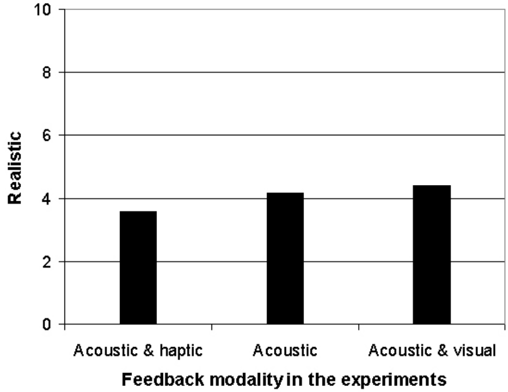

Figure 12.13: Effect of modality on the classification of the sound stimuli in the “realistic” scale.

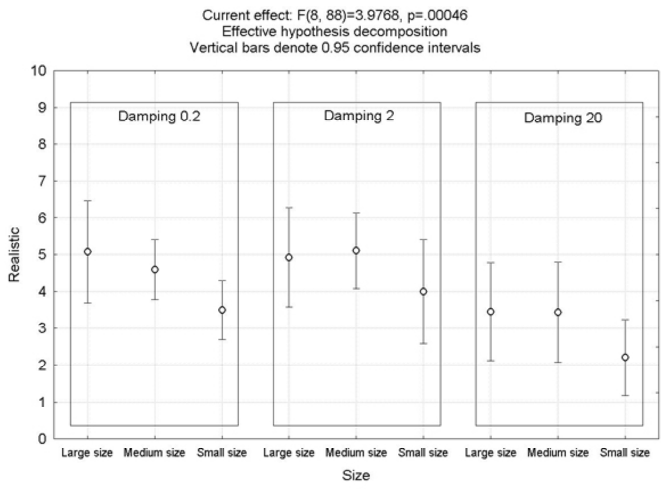

Figure 12.14: Average classification values for each sound stimulus used in the multimodal experiments.

Based on the clear acoustic meaning of the modal formulation, simplifications in the implementation of the system can be accomplished such that they introduce the smallest audible effect on the sound; or the acoustic response may even, along with implementational complexity and computational cost, be simplified in an aspired direction. The modal approach in this way supports well the idea of audio cartoonification [96].