

The Vodhran – Design and Development of a Virtual Instrument

Mark Marshall, Breege Moynihan

Interaction Design Centre, University of Limerick
email: {mark.t.marshall, Bridget.Moynihan}@ul.ie

Matthias Rath

Università degli Studi di Verona, Dipartimento di Informatica
email: rath@sci.univr.it

Abstract

This paper introduces a subtle interface, which evolved from the design of an alternative gestural controller in the development of a performance interface. The conceptual idea used is based on that of the traditional Bodhran instrument, an Irish frame drum. The design process was user-centered and involved professional Bodhran players and through prototyping and user-testing the resulting Vodhran (Virtual Bodhran) emerged. The system is a joint hardware and software system, utilising a Polhemus Fastrak motion-tracking system, coupled with specially created drivers, and sound models. This system has been developed as part of the Sounding Object (SOB) project, which is part of the Disappearing Computer proactive initiative of the IST Future and Emerging Technologies.

1 Introduction

The **SOB project** aims at developing sound models that are responsive to physical interactions and are easily matched to physical objects. The sound models, being specified by physical descriptions and actions, will be ready to be integrated into artifacts that interact with each other and that can be accessed by direct manipulation. This involves developing sound and control models to characterize the production of a sound, from the physical characteristics of the object, and of the action producing it.

As a means of accomplishing these aims, work is taking place in the development of sound models, control models and demonstration applications. One of the main demonstration applications is that of the Vodhran. This application aims to provide users with an expressive virtual musical instrument, based on

the traditional Bodhran. 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 recognisable 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.

1.1 Description of how traditional Bodhran is played

Sound is emitted from the *Bodhran* in response to a stick (beater) beating the skin, generally executed by the right hand. The sound is modulated/dampened by the use of pressure placed on the inside of the *Bodhran* by the left hand. The beater is held 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 is discrete application of pressure, but more often than not a dynamic/colourful range in pitch can be achieved by continuous control over the damping achieved by its application in counter direction to the beater. There are a variety of different applications of the damping factor employed by the musicians e.g. fingers only, side of hand only (see Figure 1).

2 Model Development

This process 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.



Figure 1. Experts playing a Bodhran.

Traditionally, the use of gesture as input involves the use of computer vision techniques, such as 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 device in order to track the users movement. This system tracks the position of up to four small receivers as the move through 3D space, relative to a fixed 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 K baud, 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°.

2.1 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 externals for the real-time sound processing software PureData (PD) for Linux.

At the bottom layer there is a driver, which allows applications to communicate with the system, and to get the bare position and orientation information for 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 8ms 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, and so latency is minimized.

This layer of the system 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.

2.2 Sound Generation

The sound generation mechanism for the Vodhran is based on the (approximate and simplified) modal description of the drum and a robust numerical solution of a nonlinear stick--membrane interaction model (see Avanzini and Rocchesso 2001).

This approach aims at an integrated "sound object", oriented at the real drum, including different forms of a player's interference, rather than 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 (Adrien 1991) forms a well-suited basis for our efforts for several reasons (see Avanzini, Rath and Rocchesso 2002):

- 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, as 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

perceptual 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 occurring impact interaction of stick and 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 Fourier transform in turn approximates the frequency response. The modal parameters were finally extracted from these "frequency responses" according to the theory of the modal description (see Avanzini, Rath and Rocchesso 2001): Peaks in the magnitude response curves mark frequencies of resonant modes, decay factors are calculated from STFT (short time Fourier transform) values at two different temporal points; the position dependent weighting factors (Avanzini, Rath and Rocchesso 2001) of the modes are given by the different *peakslevels* (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 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 (Avanzini and Rocchesso 2001).

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 iterational procedure to numerically approximate the implicit variable dependency (see Borin, De Poli and Rocchesso). (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).

2.3 Implementation

The sound engine for the virtual *Bodhran* is implemented as a module for real-time sound processing software "PD". 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. 32 or 64 samples, providing a temporal accuracy of $32/44100s \approx 0.75ms$ or $64/44100s \approx 1.5ms$. 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.

2.4 Summary

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.

3 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 3dimensional) 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 change of coordinates from receiver- to transmitter-system is accomplished by addition of the translation vector and multiplication with a 3 dimensional 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.

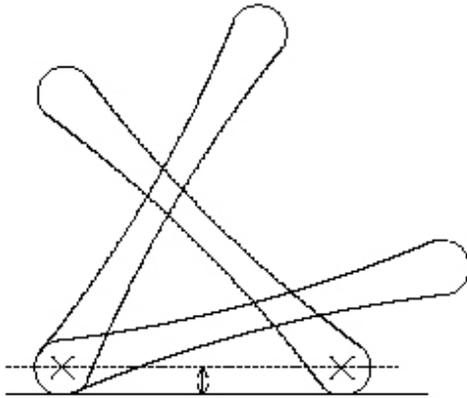


Figure 2: Bodhran stick hitting a membrane at different angles. Note that the touching point varies, while the center point of the sphere forming the surface of the stick is in equal distance from the plane for each stroke.

3.1 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 is the execution of the matrix multiplication and (eventually) the translation (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.

3.2 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 possibly be 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 2) 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.

4 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 with their own distinct and different styles and techniques took part. The players also had varying amounts of experience in the use of technology in performance.

The players were asked to perform different traditional rhythms with the sensor attached to the beater (see Figure 3), the results were recorded in order to further analyse the gestural patterns involved in *Bodhran* players. Video data was also gathered for further analysis.

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



Figure 3: Workshop session

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.

4.1 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 dampen certain modes of the skin.

Damping. This left-handed damping was used by all of the 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.

This damping effect is a major part of how the Bodhran is played, and as such will have to be incorporated fully 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 being asked to use the beater from the virtual Bodhran in order to create a simple beat by striking a virtual plane, it was noticed that some players require more tactile feedback from the system than others.

While some of 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 area will require 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 that 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 be 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” (Mulder 1996).

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.

5 Future Developments

Over the course of the development of this system so far, a number of problems have been encountered. These problems arise from the complexity of the task involved in mapping the intricate gestures used by Bodhran players to parameters of a sound model.

The movements are very small, and fast, and thus can cause difficulties for the system tracking them. As a result of this, the latency, speed and refresh rate of the hardware are currently being examined, to determine if the hardware interface itself is suited to the task. If it is found to be lacking then a new hardware device will have to be found, which offers similar tracking facilities, but at a higher update rate.

Further examination of the playing data from the workshop is also taking place, to determine if there are any more gestures or interaction methods for the Bodhran, which have not yet been incorporated into the current system.

Work is also taking place into examining the use of gesture to control a percussion instrument, when there is no reference plane, real or virtual. Players are being asked to play rhythms using the beater, and the resulting data is being analysed in order to extract the key data that identifies a strike, such as changes in velocity, changes in direction, and local minima in three dimensional movement curves. It is hoped that a means will be discovered to identify a striking action, not from it's passing a particular point in space, but from the characteristics of the movement itself.

6 Conclusions

The system that was described here 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 is hoped 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.

7 Acknowledgments

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

A 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.)

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 movement (that is the development without external interference) of each mode is analytically describable, moreover of a simple form: Each mode performs the movement of an exponentially decaying sinusoid of a fixed frequency. The corresponding resonance behavior (i.e. the frequency response) is that of a lowpass filter with a peak around this mode- (or resonance-) 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 the basis transformation: Concretely, the movement of a specific "pickup point" - giving the sound picked up here - is a weighted sum of the movements of the modes; conversely, an exterior input to the system at a "pick point" ("a 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 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.

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 (Gaver 1993).

References

F. Avanzini and D. Rocchesso, "Modeling Collision Sounds: Non-linear Contact Force", in *Proc. COST-G6 Conf. Digital Audio Effects (DAFX-01)*, Limerick, Dec. 2001, pp. 61-66, Available at <http://www.soundobject.org>.

J. M. Adrien, "The Missing Link: Modal Synthesis," in *Representations of Musical Signal*, G. De Poli, A. Piccialli, and C. Roads, Eds., pp. 269-297. MIT Press, 1991.

F. Avanzini, M. Rath, and D. Rocchesso, "Physically-based audio rendering of contact," in *(ICME-O2)*, Lausanne, Aug. 2002, Available at <http://www.soundobject.org>.

G. Borin, G. De Poli, and D. Rocchesso, "Elimination of delay-free loops in discrete-time models of nonlinear acoustic systems", in *IEEE Transactions on Speech and Audio Processing, Vol. 8, No. 5*.

Mulder, A, "Getting a grip on alternate controllers", in *Leonardo Music Journal*, 1996.

Gaver, W. W., "How do we hear in the world? Explorations of ecological acoustics". *Ecological Psychology*, 5(4): 285 – 313, 1993.