# Gesture Control of Sound Spatialization for Live Musical Performance

Mark T. Marshall, Joseph Malloch, and Marcelo M. Wanderley

Input Devices and Music Interaction Laboratory
Centre for Interdisciplinary Research in Music Media and Technology
Music Technology Area - McGill University
Montreal, QC, Canada
{mark.marshall,joseph.malloch,marcelo.wanderley}@mcgill.ca
http://idmil.org

**Abstract.** This paper presents the development of methods for gesture control of sound spatialization. It provides a comparison of seven popular software spatialization systems from a control point of view, and examines human-factors issues relevant to gesture control. An effort is made to reconcile these two design- and parameter-spaces, and draw useful conclusions regarding likely successful mapping strategies. Lastly, examples are given using several different gesture-tracking and motion capture systems controlling various parameters of the spatialization system.

## 1 Introduction

Sound spatialization has been explored as a musical tool by a many composers (see [1] for a review) and a number of systems have been developed for controlling the spatialization of sound. These systems have offered control which has varied from simple multi-channel fader arrangements, to software-based mixer systems, to software containing graphical representations of the sound sources in space. Even in the simplest implementations, sound spatialization is a multidimensional system (the basic case being 2-dimensional position and volume for each sound source) and with the development of more complex systems incorporating modelling of sound source and room parameters the dimensionality of these systems has increased dramatically. Surprisingly, the interfaces for controlling these systems have remained generally unchanged, and so in most cases, the large numbers of spatialization parameters are handled by pre-programmed cues, envelopes, and trajectories. Using gesture[1] to control spatial parameters is generally restricted to low-dimensional tasks, such as placing a sound source in 2-dimensional space. Yet, gesture offers more interesting methods for control of sound spatialization, in which a performer may control large parameter-spaces in an intuitive manner. This paper describes ongoing development of a number

[1] There are a number of different definitions of the term gesture in the literature. For this work, we use the term to mean a separable segment of the continuous, multidimensional physical movements of the performer.

M. Sales Dias et al. (Eds.): GW 2007, LNAI 5085, pp. 227–238, 2009.

of examples of gesture control of sound spatialization. These include the development of standard gesture tracking systems, e.g. data gloves and position trackers, and also novel methods of control based on musical performance gestures. In addition to producing specific implementations, the final goal of this work is to develop general methods for gesture-controlled spatialization systems which can be used in real-time during concert performances. Since this work is taking place within the framework of a larger project on the compositional uses of gesture-controlled sound spatialization, much of the work is heavily influenced by the goals of the composer. We begin by examining the types of control parameters which are made available by existing spatialization systems, in order to determine the parameters of which we may wish to enable control using gesture. This is followed by a brief overview of control issues relevant to the task, and an examination of a number of systems which we have been developing that allow for different forms of gesture-controlled sound spatialization.

## 2 Related Work

Gesture control of sound spatialization for music has been investigated as early as 1951 with the development of the *potentiomètre d'espace* by Pierre Schaeffer in collaboration with composer Pierre Henry[2]. This system allowed the performer to route prerecorded audio signals through multiple speakers in a concert performance. This aim of positioning sounds in space has become the main focus of many works on gesture controlled spatialization.

The *Zirkonium* software described in [3] allows for the control of sound source positions in 3 dimensions, together with a control over the size of the sound source. Control of these parameters may be through a graphical interface using the mouse, using any HID (Human Interface Device) controller such as a joystick, or remotely from other software or hardware using messages sent by Open Sound Control (OSC).

One interesting implementation of gesture-controlled spatialization is given in [4]. In this case, the overall system is an immersive environment for the creation of sound and music. The system allows users to create, manipulate and position a large number of different sound objects (including audio files, real-time midi input, midi arpeggiators, synthesized rhythm loops, etc.) in a virtual environment. The sound objects are then spatialized through the audio system to match their position in the virtual environment. Manipulation of the object is performed by the user through a 6 degree-of-freedom hand tracking system along with handheld mice.

The idea of using gesture control together with immersive virtual environments and sound spatialization has been addressed in a number of other works also. [5] describes a system for collaborative spatial audio performance in a virtual environment. Multiple users may interact with the system at the same time to manipulate the position of objects in 3D space. Interaction is performed using hand gestures through tracked data gloves together with stereoscopic display projection.

A system is presented in [6] which offers a navigable 3D immersive virtual space which contains both sound sources and sound sinks. The sources and sinks can be combined and manipulated to create immersive soundscapes which are conveyed to the user using sound spatialization. Interaction can take place using a number of different devices for gesture control, including datagloves, cameras and HIDs.

In [7] the author proposes a hierarchical system for the control of sound source positions in space which allows a combination of different control methods at different levels. The use of small joysticks and graphics tablets is recommended for the control of 2-dimensional positions but thought is also given to the use of tracked datagloves and haptic devices for more complex control.

## 3 Comparing Spatialization System Control Parameters

In general, existing spatialization systems are considered to offer two groups of parameters which can be controlled, namely *sound source position and orientation parameters* and *room model parameters*. However, a number of newer systems allow for the control of sound source (and sink) characteristics, so it may be more intuitive to consider 3 levels of parameters:

- sound source/sink position and orientation
- sound source/sink characteristics
- environmental and room model parameters.

**Table 1.** Comparison of spatialization system control parameters

| | VBAP[8] | Spat[9] | SSP[3] | audio TWIST[6] | Ambisonics[12] | WFS[11] | ViMiC[10] |
|---|---|---|---|---|---|---|---|
| | Sound Source Position and Orientation | | | | | | |
| Position (X,Y,Z) | √ | √ | (X,Y) | √ | √ | √ | √ |
| Azimuth, Elevation | - | √ | - | √ | - | √ | √ |
| | Sound Source Characteristics | | | | | | |
| Size | - | - | √ | - | - | √ | - |
| Directivity | - | √ | - | √ | - | √ | √ |
| Presence/Distance | - | √ | - | - | - | - | - |
| Brilliance/Warmth | - | √ | - | - | - | - | - |
| | Room Parameters | | | | | | |
| Size | - | √ | - | √ | - | √ | √ |
| Presence | - | √ | - | - | - | √ | - |
| Early Reflections | - | √ | - | √ | - | √ | √ |
| Reverberation | - | √ | - | √ | - | √ | √ |
| Reverb. Cutoff Freq. | - | - | - | - | - | - | √ |
| Doppler Effect | - | √ | - | √ | - | - | - |
| Equalization/Filtering | - | √ | - | √ | - | - | - |
| Air Absorption | - | √ | - | √ | - | - | - |
| Distance Decay | - | √ | - | √ | - | - | √ |
| Mic/Sink/Speaker Pos. | √ | √ | - | √ | - | - | √ |
| Mic/Sink Directivity | - | - | - | √ | - | - | √ |
| Heaviness/Liveness | - | - | - | - | - | - | √ |

The extra level (sound source/sink characteristics) includes aspects of the sound sources other than their position or orientation in space. Parameters such as source/sink size or directivity are important descriptors in some systems and cannot be accurately described as room model parameters. Environmental and room model parameters include room parameters such as the room size and amount of room reverberation, along with environmental parameters such as the energy absorption factor of the air or additional equalization.

In an attempt to include a broad range of possible parameters, we examined a number of available spatialization systems, including those which are currently most used and a number of more recently developed systems. These systems range in complexity from sound source panning systems such as VBAP [8], to complex positioning and room models such as Spat [9] and ViMiC [10]. Table 1 shows the results of this comparison of systems.

## 4 Control Issues

A number of issues are of crucial importance when using gesture to control sound spatialization parameters. Careful consideration of these issues, and carefully matching parameters in the mapping process, will allow not just intuitive and powerful control of sound spatialization, but yield rewarding performances for performer and audience alike. These issues must also inform choices in sensors, sampling rates, A/D converters, transmission bandwidth, and communication protocols.

**Discrete vs. continuous parameters.** Depending on the desired spatialization effects, some parameters, such as the number of sound sources, are usually approached as discrete values, while others, such as source volume, air absorption, or spatial position, usually require continuous control for common spatialization effects. In order to successfully map gestures for the control of these parameters, movement and sensing should match the discrete or continuous nature of the control. Discrete values can often be extracted from continuous signals, allowing the detection of discrete signs, but the opposite is not usually true: continuous parameters are best controlled by gestures that are continuously executed and sensed.

**Resolution.** The resolution required for controlling a given spatialization parameter depends on the desired spatialization effect and on the parameter's perceptual impact. Listeners are not likely to notice centimeter-scale changes in sound source location, especially for low-pitched sounds, but small changes in filter frequencies or air absorption parameters may be more easily perceived. On the gesture side, resolution can be limited by sensors or acquisition hardware, but is also naturally limited by human ability. As anyone who has tried to control musical pitch using finger pressure can attest, certain movements and certain body parts are unable to exert fine control, a problem that is compounded when the performer is not given adequate feedback regarding the state of the system they are controlling.

**Control rates.** The desired rate of change for the parameter must also be considered. Many spatialization parameters are typically considered to be static and are set only once before a performance, while others are constantly changing within a performance. Static parameters usually include room dimensions, wall reflectivity, and speaker positions; source position, orientation, and volume are usually changeable. Gestures used to control these parameters must obviously match or exceed the needed rate.

This description reflects a *current control* (or continuous control) approach to gesture control. However, another approach is familiar from percussion performance. A *ballistic* movement (such as a drum strike) may be used to determine the parameters of a sound spatialization event, after which no further control is exerted. In this case, all of the control information must be included in the launching gesture, or encoded in preparatory movements.

**Integrality and separability.** The relationships between spatialization parameters are also an essential consideration [13]. Certain parameters, such as the components of source position and orientation in each dimension, are closely related and deserve special attention during the mapping process. Similarly, the spatial position and orientation of the performers body, or parts thereof, are also related and should not be considered separately. In musical performance especially instrumental performance this is further complicated by the goal-directed manipulation of an external object (the instrument). In this way the posture and movement of individual body parts become integral to some extent.

**Conscious vs. non-conscious control.** An especially important issue when considering performance in the musical sense of gesture controlled sound spatialization is that of *cognitive load*, which is the total amount of mental activity imposed on working memory at an instance in time [14]. Considering that not all performers have spare bandwidth for controlling something other than their instrument [15], three main approaches may be taken: a separate performer may control *only* spatialization, a performer may be tasked with controlling spatial parameters in addition to their instrument, or spatialization parameters may be controlled by the performers gestures *without their conscious control.* If conscious control is desired, gestures must be chosen such that they can be performed without disturbing the instrumental performance, and it is assumed that the performer has spare attention for this task. For non-conscious control, mapping is more of a compositional process rather than performer interpretation, and raises some interesting questions: Will the performer feel that they are not in control of the sound? One partial solution is to attempt to map resting body states to neutral spatialization states.

## 5 Gesture for Control of Spatialization

In this section we describe our implementations of systems for gesture-controlled sound spatialization. In keeping with the three types of spatialization system parameters discussed in Sec. 3 we have identified three specific roles for control of spatialization. These roles are:

**Spatial Performer** - performs with sound objects in space by moving sound sources in real-time using gesture

**Instrumental Performers** - indirectly manipulate parameters of their own sound sources through their performance gestures on their own acoustic instrument

**Spatial Conductor** - directly controls large-scale (room and environment) parameters of the spatialization system using gesture

These roles derive from an examination of the types of parameters available to control, as well as the specific needs and goals of the composer for the musical portion of the project.

### 5.1 Manipulating Sound Source Positions: The Spatial Performer

One of the most obvious applications for gesture control of sound source position would be to make use of *direct manipulation*, or in this case *virtual direct manipulation.* In our specific implementation, we make use of a pair of custom-built datagloves, which allow measurement of the posture of the performer's hands using bend sensors mounted on the fingers. Along with a Polhemus Liberty magnetic position tracker to track the position and orientation of the hands, this system allows the performer to directly position multiple sound sources in space [16]. In order to take into account the integrated nature of the 3 positional parameters of the sound source objects, an integral 3-dimensional tracker is used to track the hand positions [13]. Tracking both hands (a bimanual system) allows for individual repositioning of two sources simultaneously.

To allow for a comparison between current control and ballistic control of sound sources, we have also implemented a ballistic sound source control system. This system allows the performer to launch sound sources by striking pads on a Roland V-Drum midi drumkit. Each pad provides a different angle of launch in the horizontal plane and the velocity of the strike controls the force of the launch. Multiple sound sources can be launched in this way and will move around the virtual room, pulled by gravity and potentially colliding with objects in its path. As can be imagined, this offers a much different form of interaction with the system than the more traditional direct manipulation-like system already described. Figure 1 shows this system in action.

Another example of a basic control system for sound source positioning uses the idea of "steering". In this case, "steering" forms a metaphor for interaction with the system [17]. It is possible to steer a single sound source around the space in a number of ways. For example, a joystick interface has been used to move the source in 2 dimensions, a task that could be performed similarly using a hand-held two-axis tilt sensor, allowing the user to steer by tilting their hand. In our implementation a weight-sensitive floor was developed, which allowed the user to steer a sound source in 2 dimensions by shifting their center of mass.

We have also implemented a simple system for concert use, in which a tilt sensor was mounted on the head of a percussionist to enable manipulation of a sound source position while playing vibraphone, a task which requires both

**Fig. 1.** A performer using our ballistic sound source control system to launch multiple sound sources in the virtual space

hands and feet and leaves few gestures for independent control. Both this system and the floor-based system allowed for a coarse level of control of sound source position but not surprisingly proved not to be well suited for fine positioning tasks (in direct contrast to the hand-tracking system). Also, both systems allow for only a 2-dimensional control of sound source position and so do not allow as much control as the hand-tracking system.

### 5.2 Control of Sound Source Parameters: Instrumental Performers

While controlling source location may seem obvious, the control of other sound source parameters offers further interesting possibilities. One method is to use existing musical performance gestures to control these parameters. Useful gestures include excitation and modification gestures involved in performance as well as ancillary gestures made by the performer while playing [18]. While using low-level gesture data from live performance can be interesting, they are often so closely correlated to the acoustic sound that higher level parameters may be more useful for mapping. For instance, when tracking a cellists bowing gestures, the energy of the gesture can prove to be a more useful control parameter than the physical position of the bow. This energy can be calculated using the amplitude and velocity of the bow movement and mapped to, for example, the volume or directivity of a sound source. As the performer plays more energetically the sound source could become louder (or perhaps for an unusual compositional effect the opposite could be true). This type of mapping allows us to give some control of the spatialized sound to the performers, but without having to give them any explicit instructions or increasing their cognitive load.

Control of spatialization using existing performance gestures is an area which we are actively investigating. In one approach we use a motion capture system to extract motion data from performances and use these data to design gesture control systems which are controlled indirectly by the performer. This means

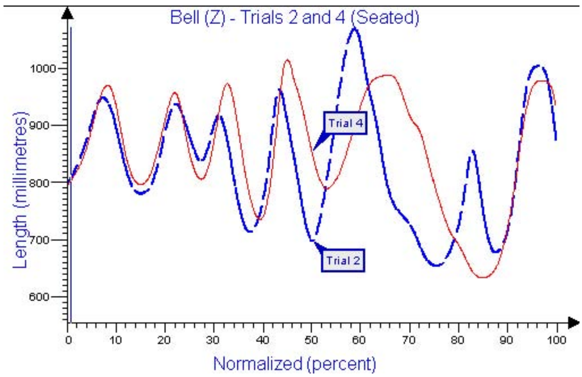


**Fig. 2.** Graph showing the vertical movement of the bell of a clarinet over two performances of the same piece by the same musician

that parameters which are extracted from the performer's movements are being used to control the system. For example, Fig. 2 shows a graph of the movement of the clarinet bell during two different performances of the same piece by a single performer. As can be seen from this graph, there is consistency between the performances, which would allow this data to be used to consistently control a parameter of the spatialization system, without requiring the performer to learn to use the system. As this movement is dependent on both the piece being played and the performer playing the piece, it allows for the control of spatialization parameters in a way which is directly tied to both the performer and the piece. The spatialization control becomes an intrinsic part of the performance. Similar consistancy of movement has also been found for changes in the center of mass of a clarinet performer and for the center of mass of a cello player, indicating that these could also be used to control the spatialization system in a similar way.

In order to allow us to make use of the performance gestures of instrumental performers, we have developed a wireless accelerometer-based system which can be either worn by the performer or directly attached to the instrument. The system tracks acceleration in 3 axes and can then be used to track a number of different gestures, including the clarinet bell movements discussed already. In particular we have found this system to be very useful in tracking aspects of cello performance gestures. When a cellist wears two of the accelerometer systems, one on each forearm, a number of interesting gesture parameters can be extracted. Thus far, we have managed to extract features such as bowing energy, relative position on the fingerboard and angle of bowing (which can allow us to determine which string is being bowed) [19]. Figure 3 gives some examples of these parameter signals. These gesture parameters can be used to subtly control sound source parameters allowing us to create interesting effects, such as mimicking changes in directivity which result from different playing techniques but which are lost when the sound is diffused using loudspeakers.

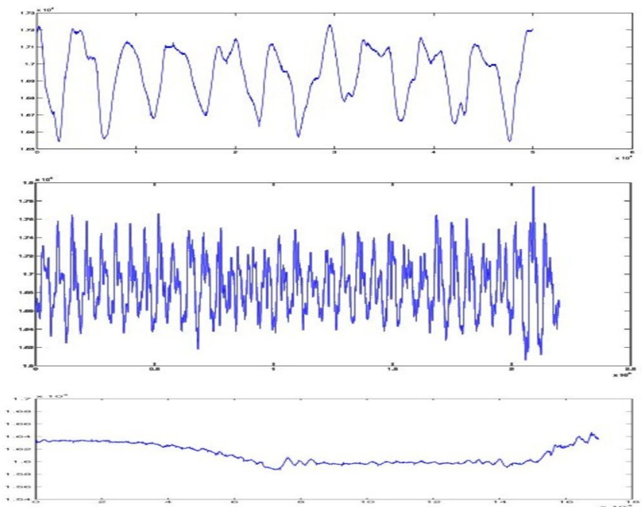

**Fig. 3.** Graphs of data captured from the accelerometer system worn by a cello performer. (a) bowing arm acceleration for regular bowing (b) bowing arm acceleration for tremolo bowing and (c) fingering arm rotation indicating relative position on the fingerboard.

### 5.3 Controlling Room and Environment Parameters: The Spatial Conductor

The control of room and environment model parameters offers scope for some unusual and large scale effects on the sound. By allowing the user in the role of spatial conductor to control these parameters, we allow them influence over the sound of the overall system and all the sounds within it. Simply by changing one parameter of the system, such as the virtual room size, the spatial conductor can make drastic changes to the overall system sound.

In order to implement a system for control of these parameters, we have once again used an instrumented data-glove system. Using a combination of bend sensors on the fingers and a 6 degree-of-freedom position tracker, we can recognize various hand gestures as a combination of hand postures (using the bend sensor signals) and movements (extracted from the position data). This allows us to evaluate the use of different gestures to control system parameters. Thus far, we have the ability to manipulate parameters such as room size (by grasping and stretching/shrinking the room), microphone ring diameter (through a measure of the openness of the hand) and reverb (through a hand posture together with a vertical position change). There remains some work in the evaluation of these gestures, along with the determination of an overall gesture vocabulary for these manipulations. Figure 4 shows a performer manipulating the size of the virtual room using an early prototype of the system.

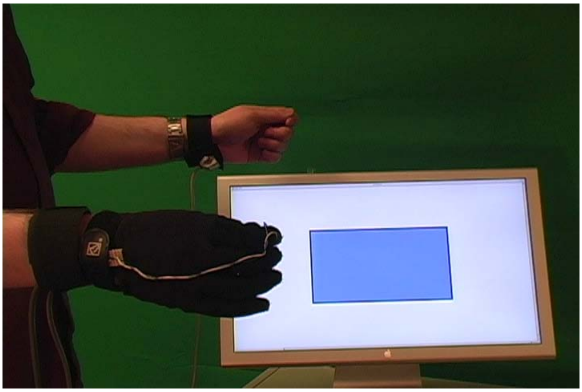

**Fig. 4.** A performer manipulating the size of the virtual room using simple hand postures and a stretching/shrinking motion

## 6 Discussion and Future Work

The primary aim of this work has been to investigate methods of gesture-based control of spatialization which have not so far seen much research. While systems such as our glove-based hand tracking system allow for a traditional direct manipulation-style interaction with a sound spatialization system, the other less traditional systems also offer some interesting forms of control.

Positioning sound sources using ballistic control methods or by steering using the performers center of mass offer a means in which a performers can spatialize sounds while still performing with their acoustic instruments. The spatialization becomes part of their performance. Using drum-pads for sound launching can allow a percussionist to both create and spatialize their own sounds, using gestures with which they are familiar. Similarily, a cello or clarinet performer can steer their own sound in the spatialization system while still playing their acoustic instrument. As has already been mentioned these systems offer a coarse level of control, rather than the fine level offered by the hand-tracking system. These systems could also place some extra level of cognitive load on the performer.

On the other hand, using data extracted from performers own gestures to control aspects of the spatialization system places less (in fact almost no) extra cognitive load on the performer. They can play their instrument and affect the spatialization system with little thought as to how the control is working. Choosing mappings such as one which increases the spatialized sounds brightness with increasing performer energy would mean that the performer would only have to play more or less energetically (and so think only in these terms rather than more complex ones) in order to affect the sound.

The division of the system control into the three primary roles already described also facilitates maximum control over the system.These roles also help distribute the cognitive load amongst the different performers based on the amount of work already placed on them by the performance. The roles themselves also inherit from the traditional roles of performer and conductor already present in orchestral performance.

While we have already been examining these roles and testing our systems in laboratory situations, there now remains the task of evaluating their use in actual concert spaces. As such we have planned to perform a number of experiments in concert halls to examine the control methods, both from the point of view of the performers and also of the audience. Also planned are a number of concerts which will make use of our interfaces, both singly for small pieces and together for larger performances. These tests, together with the performances, should allow us to examine in detail the usefulness of these systems as part of spatial musical performance.

## 7 Conclusion

This paper discussed our ongoing research into gesture control of sound spatialization. This work has involved the examination of a number of different existing spatialization systems in order to determine the types of parameters used in such a system which might be controlled using gesture. We have also examined a number of methods for gesture control from the human-computer interaction literature which might be useful in our pro ject, as well as examining a number of mapping issues which impact the design of our systems.

Some examples of systems which we have developed and used for control of spatialization have also been discussed. We have been concentrating on three main roles for control of spatialization, which have been developed in conjunction with a composer to allow for good compositional use of gesture controlled spatialization. Results from informal use of the systems have been promising and now more formal evaluation is planned. Following on from this we will be holding a number of performances which will make use of these control interfaces to provide an aspect of gesture controlled sound spatialization.

## Acknowledgements

This work has taken place with the support of the *Natural Science and Engineering Research Council of Canada* and the *Canada Council for the Arts*. Thanks to Chloe Dominguez for taking part in the motion capture sessions and to Nils Peters for feedback on the survey of spatialization systems.

## References

1. Zvonar, R.: A history of spatial music (2006), `http://cec.concordia.ca/econtact/Multichannel/spatial_music.html`
2. Chadabe, J.: Electric Sound: The Past and Promise of Electronic Music. Prentice-Hall, Englewood Cliffs (1997)

3. Ramakrishnan, C., Goßmann, J., Brm̈mer, L.: The zkm klangdom. In: Proceedings of the 2006 Conference on New Interfaces for Musical Expression (2006)
4. Valbom, L., Marcos, A.: Wave: Sound and music in an immersive environment. Computers and Graphics 29, 871–881 (2005)
5. Naef, M., Collicott, D.: A vr interface for collaborative 3d audio performance. In: Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME 2006), pp. 57–60 (2006)
6. Wozniewski, M., Settel, Z., Cooperstock, J.: A system for immersive spatial audio performance. In: Proceedings of the 2006 Conference on New Interfaces for Musical Expression (2006)
7. Schacher, J.: Gesture control of sounds in 3d space. In: Proceedings of the 2007 Conference on New Interfaces for Musical Expression (NIME 2007), New York, USA, pp. 358–362 (2007)
8. Pulkki, V.: Spatial Sound Generation and Perception by Amplitude Panning Techniques. PhD thesis, Helsinki University of Technology (2001)
9. IRCAM: Spatializateur website, `http://forumnet.ircam.fr/356.html`
10. Braasch, J.: A loudspeaker based sound projection using virtual microphone control (vimic). In: 118th Convention of the AES, Barcelona, Spain (2005)
11. Spors, S., Teutsch, H., Rabenstein, R.: High-quality acoustic rendering with wave field synthesis. Vision, Modeling, and Visualization, 101–108 (2002)
12. Malham, D.: Spatial hearing mechanisms and sound reproduction. Technical report, Music Technology Group, University of York (1998)
13. Jacob, R.J.K., Sibert, L.E., McFarlane, D.C., Mullen, M.P.: Integrality and separability of input devices. ACM Transactions on Computer-Human Interaction 1(1), 3–26 (1994)
14. Cooper, G.: Research into cognitive load theory and instructional design at unsw (December 1998), `http://education.arts.unsw.edu.au/staff/sweller/clt/index.html`
15. Cook, P.: Principles for designing computer music controllers. In: ACM CHI 2001 Workshop on New Interfaces for Musical Expression (2001)
16. Marshall, M.T., Peters, N., Jensenius, A.R., Boissinot, J., Wanderley, M.M., Braasch, J.: On the development of a system for the gesture control of spatialization. In: Proceedings of the International Computer Music Conference, pp. 260–266 (2006)
17. Baecker, R., Grudin, J., Buxton, W., Greenberg, S.: Readings in Human-Computer Interaction: Toward the Year 2000., 2nd edn. Morgan Kaufmann, San Francisco (1995)
18. Cadoz, C., Wanderley, M.: Gesture - music. Trends in Gesture Control of Music (2000)
19. Marshall, M., Malloch, J., Wanderley, M.: Using the 'energy' of performance gestures as a control for a spatialization system. Technical Report MUMT-IDMIL-07-08, Input Devices and Music Interaction Laboratory, McGill University (May 2007)