AMBISONICS SYMPOSIUM 2009
June 25-27, Graz

AN EXPLORATION OF DESIGN PARAMETERS FOR HUMAN-INTERACTIVE SYSTEMS WITH COMPACT SPHERICAL LOUDSPEAKER ARRAYS

Andrew Schmeder
Center for New Music and Audio Technologies
Department of Music, UC Berkeley
1750 Arch Street, Berkeley CA 94720, USA
andy@cnmat.berkeley.edu

Abstract: *The exploration and evaluation of design parameters is introduced for applications of compact spherical loudspeaker arrays in the context of human-interactive systems. Applications include: electro-acoustic music–live and reproduced, multi-observer interactive auditory displays, communications, and acoustical engineering tasks. We describe the basic situational operation of an array of this type: its perceptual effects and their magnitude, and how to control the system in the context of spatial auditory display design. Several interactive paradigms are explored including: 3D pointing and selection, person-tracking, 3D haptic controls, and live audio.*

Key words: compact spherical loudspeaker array, auditory display, auditory perception, human-computer interaction

1 INTRODUCTION

This paper describes recent research in human-computer interactive real-time audio systems using compact spherical loudspeaker arrays. These systems have many audio engineering applications, especially for room measurement. However, the original motivation that led to experiments with spherical loudspeaker arrays was to address the complaint that traditional loudspeaker systems (typically consisting of cabinet-style forward-firing loudspeakers) produced sound lacking certain "lifelike" features required for the continuity of perceived agency between a live performance and its reproduction [1].

1.1. Compact Spherical Loudspeakers

Over the course of many years, the research group at CNMAT [2] created a series of compact multi-directional loudspeaker prototype systems, beginning with the simple platonic-solid arrangements, cube and dodecahedron (Figure 1, left and center) (also [3] [4] [5] [6] [7] [8] [9]). These early systems already showed potential for an increased sense of sound dynamics, however theoretical analysis showed that these systems are actually quite limited in directional control [10]. A simple explanation of this fact is that the order of directivity control for an array of N elements scales approximately by $\sqrt{N}$. Recently, the construction of a 120-channel spherical array [11] (Figure 1, right) was completed, with real-time computer-driven control of directivity up to 9th order spherical harmonic. This array is comprised of 2.5cm dome transducers and has an approximately spherical shape with diameter 20cm.

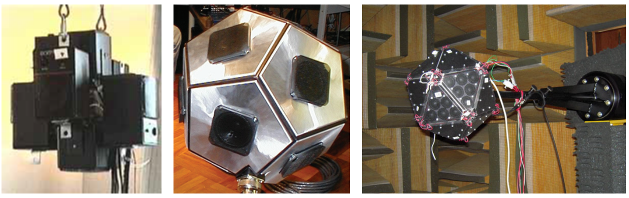

Figure 1: 6, 12 and 120-channel spherical speaker arrays

1.2. Parametric Design Exploration

The present work returns to the original motivation of perceptual sound quality.

In order to establish a means to evaluate such properties, this paper works towards the establishes a framework of descriptive terminology for spatial auditory scenes including directional sources, along with a set of parametric "effects" based on source directivity manipulation that influence auditory scene perception. Through interactive exploration of the design parameters and gesture-interface prototyping the space of possibilities can be covered rapidly. It is also possible to measure the effect of these parameters on a perceptual scale using the method of adjustment.

1.3. Thinking by Doing and User Experience Design

The design of interactive auditory displays is ultimately concerned with the user experience [12] [13]. A visual depiction of an interactive auditory system with a "human in the loop" is shown in Figure 2. Interactivity and prototyping are important in this context because auditory perception cannot always be predicted. Interaction enables direct phenomenal experience of the correlation between manipulation of a parameter and the magnitude of its percept [14]. Furthermore, auditory perception is not a static sensory system: not only is there neurological and physiological adaptation, it includes kinetic effects such as head motion. Kinetic cues conveyed by auditory features such as doppler and early-reflection structure are of particular importance for spatial perception. Kinetic models and methods of probing the universe are a primary cognitive function that emerges in early childhood development and remains through many aspects of human interaction [15]. For this reason the present work emphasizes use of tangible, physical and haptic controls.

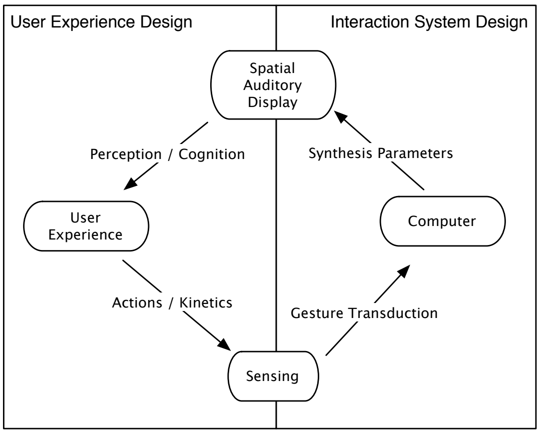

Figure 2: Human interaction with an auditory display

2 DESCRIBING SPATIAL AUDITORY SCENES

A spatial auditory scene description is an abstraction for describing an acoustic field using simple terminology that is related to the cognitive experience induced by the phenomena of immersion in that field. Such scenes may be the result of extant physical processes (reality based), or may be synthesized (virtual reality), or formed from a mixture of real and synthetic elements (augmented reality) [16]. Whereas wavefield arrays typically are used to produce a virtual scene, situated compact spherical arrays produce an augmented reality.

Taking as a starting point the terms and structure given in [17], a new set of descriptors, shown in Table 1, is our proposal for an improved set that generalizes to the case of three-dimensional sound fields containing omnidirectional, directional and even volumetric sources.

The necessity of presenting this taxonomy stems from problematic use of terminology in other literature. For example, extensive research based on lateral arrays has resulted in a preponderance of one-dimensional qualifiers for source dimensional attributes such as "width". More importantly, the conventional emphasis on positional control of sources in periphonic audio systems omits entirely the attribute of source directivity. Understandingly, these systems are not optimized for directional control. However, periphonic wave-field arrays are capable of reproducing source directivity [18] [19], albeit limited to relatively low-order patterns.

Category	**Term**
Scene Entities	Source Surface Shell Field
Dimensional Attributes	Position Volume Directivity
Immersive Attributes	Presence Immersion
Spatialization Schemata	Object Path Rotation Scale Containment

Table 1: Terminology for auditory scene description

2.1. Directivity

A source without directivity, also called a monopole or omni-directional source, radiates acoustic energy equally in all directions. A source with directivity has a radiation pattern specified with respect to spherical angle around a point. Directed sources can also be formed from the superposition of monopole sources, for example a dipole is formed by the colocation of two identical but phase-inverted sources. Most importantly, directed sources have a sound-power loss that is frequency dependent, resulting in the near-field and far-field effects. Note that because the power loss is analytically predictable, directional systems can be controlled as to have a flat response at a specific distance [20].

2.2. Presence

Presence is a component of perception related to the early reflections generated by a contained source. For an omnidirectional source, the location of first reflections can be predicted from the spherical expanding wavefront. For a directed source, the position of early reflections is dependent on the angular distribution of energy, and also undergo motion effects when the source is rotated.

2.3. Envelopment

Envelopment describes the quality of the late-reflections of sound in an enclosed space. Reverberation can be described statistically for simple room geometries. In a spatial context, envelopment includes the distribution over spherical angle of incident sound around the listener (i.e., diffuse field).

3 USE OF SPHERICAL LOUDSPEAKERS

3.1. Spatial Distribution

The spherical array geometry has uniform control of directivity with high accuracy over the entire sphere. This provides a unique capability among auditory display geometries in that the sound field is correct for all observers regardless of location (i.e., there is no sweet spot or anticausal listening region). However, this choice to prioritize uniformity makes a tradeoff against directional precision. A geometry with negative curvature (such as parabolic) can produce directivity with a narrower angle, but with restricted steering. The other significant compromise of the compact spherical geometry is the extremely limited positional control. However, the array does have positional control over sources created by means of specular reflection when situated in an enclosing space.

3.2. Beam-forming Directivity Control

Several methods are available to control the reproduction of directivity. The method employed in this work is a beam-forming algorithm based on the analytical inverse spherical harmonics transform [21] [20] [22]. Beam-forming can be interpreted in a sampling-theory framework as a band-limited dirac impulse. It is possible to steer many such beams in real-time, each having its own output signal. For the 120-channel 9th order array, the smallest beam width (defined as the angle where sound power is 6 dB down from its maximum) is approximately 35 degrees.

3.3. Perception of the Direct Beam Path

Here we consider the perception of beam directivity when the listener is situated within the main lobe of a beam (Figure 3).

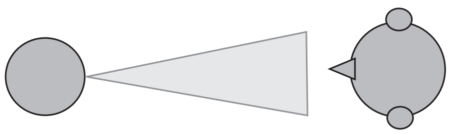

Figure 3: Direct path observation of beam

3.3.1 Motion Effects

When the beam is in motion and sweeps across the listener, a large change in amplitude as well as spectrum is observed. There is a strong and immediate sense of being "hit" by focused source of acoustic energy.

3.3.2 Near-Field Localization

Given that a compact spherical loudspeaker array is only an approximation of a point source, its non-zero volume causes a small, unintentional translation of the beam when it is rotated. If the listener is looking towards the array, localization accuracy from ILD/ITD cues is accurate within one degree. Therefore, unless the listener is in the far field, the perception of motion may be attributed to the translation rather than the beam rotation (Figure 4). For the array used in this work, having a diameter of 0.02m, this effect is significant for listeners located within 1 meter of the array.

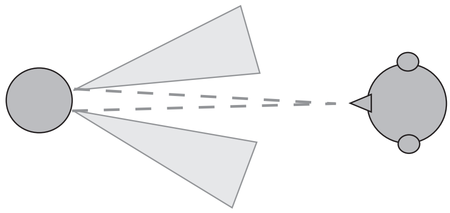

Figure 4: Unintended translation of apparent position

3.3.3 Very-Near Field Body Effects

At a distance less than 0.5m from the array surface, beam forming can induce an audio-powered "spotlight" that falls on a particular sub-region of the users own body (Figure 5). For example, the sweeping of a beam of noise over the face "reveals" the spectral response in the sound of the user's individual HRTF. Thus, in the very-near field the spectral flux induced by partial HRTF response provides an additional cue to beam direction. Placement of hands into this area of the display gives an interesting demonstration of frequency-dependent diffraction.

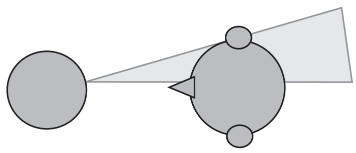

Figure 5: Very-near field partial HRTF

3.4. Perception of the Early Reflections

Suppose the beam is directed so that the listener is outside of the main lobe. Assume also that the beam has ideal directivity such that there is no energy outside of the main lobe. In an anechoic environment, no acoustic information will reach the ears of the listener. If the beam strikes a reflective surface, the listener will hear the location of reflection to be the origin of the source. (Figure 6)

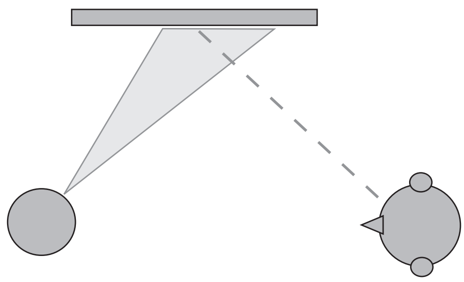

Figure 6: Localization of first reflection

3.4.1 Reflection Localization Sensitivity

Using beam steering, it is possible to place reflections at arbitrary positions in the room. However, the sensitivity to localization of reflections is not uniform. For the typical case of a listener looking towards the array, the expected positions in the lateral plane where localization sensitivity is at extremes is shown in Figure 7. For presence enhancement effects, angles resulting in reflection positions with high localization sensitivity should be preferred. For immersive effects, the opposite strategy results in greater perceived diffusion.

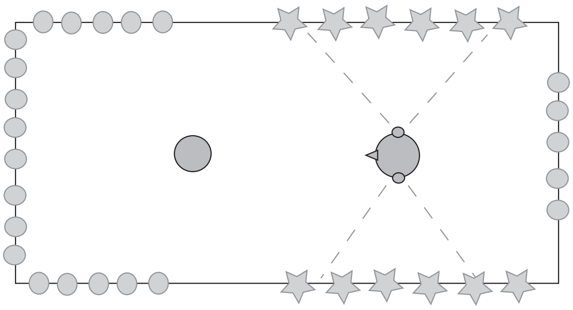

Figure 7: Expected sensitivity of reflection localization. Circles at maximum, stars at minimums. Dashed line shows listener's cone of confusion.

3.5. Motion of Reflections

When a directed source is rotated, its reflections undergo a translation and therefore a doppler frequency shift is observed. For the case of a single reflection, the doppler is coherent and moving in one direction. For the more typical case of multiple reflections, approximately half of the reflections are moving towards and half away from the listener. This causes a symmetric spectral smearing, also called spectral line broadening. In the case of musical instrument timbre synthesis, spectral line broadening is a quality known to increase perceived realism [23] [24].

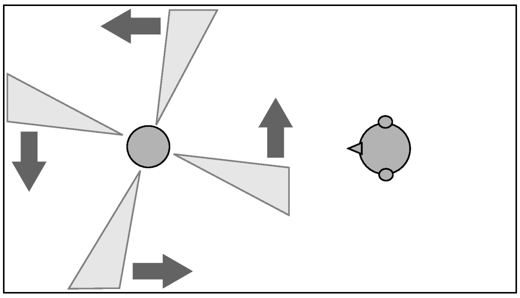

Figure 8: Multidirectional motion in first reflections of a rotating directional source

3.6. Apparent Source Volume

The perception of apparent source width (more generally, source volume) is known to be correlated with the structure of early reflections. Using directivity control it is possible to manipulate the time-structure of the early reflections within a room relative to the time of arrival of a direct path beam. When the direct path beam has zero time offset to the early reflections, the listener perceives the source to be small in size. When the direct path beam is delayed, the listener perceives that the source has a larger diameter. The amount of delay possible before the source image decoheres may depend on the size of the room and its architectural features.

Note that this approach to source width control is drastically different from the multi-channel spreading strategy used in amplitude-panning. It is also more effective in simulating the desired result.

4 DIRECTIVITY EFFECT PERCEPTION

In this section an informal study is described of perceptual effects resulting from source directivity control. The goal is to establish the useful range and perceptual sensitivity for each effect in terms of its parameterization. The data in this section was obtained by the author, using the method of adjustment. The listening conditions include the situation of the array within an ordinary rectangular room approximately 5x7 meters.

4.1. Directivity Rotation: Generalizing the Leslie Effect

The Leslie Cabinet is an early example of an audio system featuring a source directivity effect. The appeal of the Leslie sound arises from its enhanced sense of presence as its rotating horn creates a kinetic aura in the first reflections and motion-induced spectral smear [25] [26]. In early experiments with real-time directivity steering, a rotating beam effect was identified as a simple and effective means to create an interesting and obvious effect. In this case the enclosing room replaces the cabinet.

4.1.1 Parameterization

The rotation effect applied to a beam is parameterized by angular rate of change. This rotation does not necessarily occur in the horizontal plane, although from prior work on localization we expect the lateral direction to be the most effective.

At extremely slow speeds, the rotation effect appears to be stationary. And at very high speeds, the beam at various angles fuses temporally into a coherent source with an apparent amplitude modulation but no perceived rotational component.

4.1.2 Measurements

The rotation effect is known to work well with a continuous broad band white noise as the source. A further hypothesis was made that this effect might vary by frequency, so the stimulus in this experiment was reduced to a narrowband filtered noise centered at the given frequency.

Since we do not consider the stationary AM state to be a directional effect, by method of bracketing adjustment, the maximum rate-of-rotation was found, prior to angular fusion. Again by method of adjustment, a just-noticable difference in rotation frequency was estimated, which is given in terms of octaves relative to the reference frequency.

The preliminary findings (Table 2) show that effect is remarkably robust to spectral region, although it shows a slight drop off in very high frequencies where other factors are likely at play such as reduced hearing sensitivty. In the lateral plane, a maximum rate of 5Hz and a JND of .18 octaves gives about 30 JND steps from zero rotation to angular fusion, although speeds below 2Hz are more perceptually obvious. The reduced sensitivity to vertical

Noise Notch Frequency	Max Horizontal (Hz)	Max Vertical	JND Rotation @ 1.58Hz	@ 2Hz	@ 2.8Hz	@ 4Hz (octaves)
500	5.2	1.9	.16	.12	.12	.12
1000	5.7	2.0	.2	.09	.11	.14
2000	5.2	1.9	.18	.14	.12	.16
4000	4.7	1.5	.2	.14	.18	.14
8000	4.5	1.3	.18	.12	.14	.16

Table 2: Maximum angular frequency before fusion for a rotating beam in horizontal, vertical planes. Just-noticable difference for horizontal rotation at various frequencies (vertical JND is not significantly different so omitted).

rotation is not surprising.

4.2. Directivity Flux

The directivity flux experiment is an attempt to measure the perception of changing directivity without any rotational (doppler inducing) component.

The beam forming system is configured with 20 beams, distributed uniformly over the sphere. Each beam transmits a broadband noise, while each beam also undergoes a random amplitude modulation.

4.2.1 Parameterization

To induce directivity flux, a set of uniform random $(0, 1)$ gains is chosen every 20 msec and applied to the beams. To increase directivity flux, the random gains are raised to a power y, which is parameterized in dB, and then normalized. A secondary parameter is estimated as the number of effective beams firing, which is 20 in the zero flux case and 1 in at maximum.

4.2.2 Measurements

Zero directivity flux is obtained (approximately) at the power setting $y = -24$. The first step of noticable increase in directivity flux is found at $y = -4$. Table 3 lists each step of just noticable step in increasing flux.

Directivity Flux Power (dB)	Number of Effective Beams
-24	20
-4	17
3	13
9	9
16	6
21	4
36	2
45	1

Table 3: Power and number of effective beams active for each JND step from zero to maximum flux

There are a number of interesting aspects of the directivity flux effect: firstly, the effect is strongly affected by head orientation. Looking away from the array causes a drop in sensitivity of about 12 dB, and about 6 dB looking to the side. Second, any slight motion of the head causes a momentary but dramatic increase in sensitivity to the phenomena. Overall, the effect has a small perceptual resolution, having only about 6 JND steps from minimum to maximum.

4.3. Source Volume

Source volume modulation (described in Section 3.6), is achieved by time-delay of the direct beam relative to first reflections. The listener stands in the direct beam path, and beams to create first reflections are pointed along all other axial directions (including straight up and down). The same signal is sent to all beams, using recorded music as a source. The signal going to the direct beam is then delayed in units of milliseconds. Table 4 lists each just-noticable step from zero delay to maximum.

Direct Beam Delay (msec)
0
4
10
17
42

Table 4: JND steps from zero to maximum for source volume effect

The perceptual scaling is approximately logarithmic, doubling with each JND step. Delays longer than 10 milliseconds correspond to a source volume larger than the containing room. The interpretation of this case is therefore debatable. In general this effect is quite small having only a few JND steps over its useful range.

5 INTERACTION CASE STUDIES

Here a series of gesture interfaces is described, and used to enable real-time response between the user and the auditory display system. Several of these interfaces are DIY prototypes constructed by the author [27] and others are common and commercially available.

5.1. Personally Fixated Source

A 3D position sensor was constructed by attaching a Wii Remote to a MadCatz Gametrak absolute position sensor (Figure 9). This device provides absolute position sensing inside a cone of about +/- 30 degrees and three meters height [28]. Typically the origin of the cone is situated above the loudspeaker array (ceiling mounted).

The positioner is a graspable, movable object and provides data with a 1:1 correspondence with real space. Real-time software transforms the positioner coordinates to spherical

Figure 9: 3D position tracking with overhead Gametrak

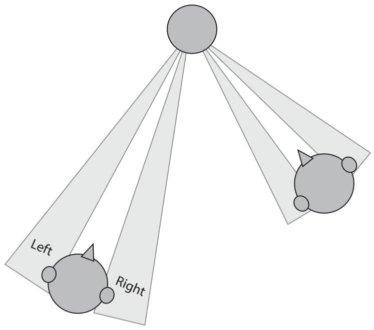

Figure 11: Multiple simultaneous personally fixated binaural sources

coordinates with respect to the center of the array. A beam is then produced at that angle through which an audio signal is transmitted (Figure 10). Numerous signals can be used, although typically broadband noise is useful in order to make an obvious effect. The beam can be retargeted through any point within the reach of the user, enabling the partial HRTF illumination effect when the user is in the very-near field (see Section 3.3.3). When the user walks around the array while holding the positioner, an unusual (anti-physical) effect is observed whereby the source appears to be completely static because it is fixated on the user position.

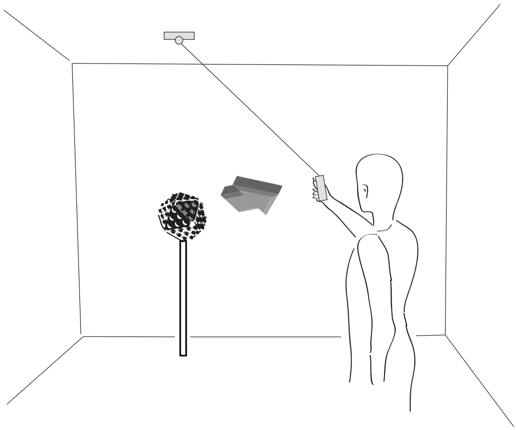

Figure 10: Beam steering towards the positioner

This demonstration can be enhanced by using a stereo music file as the audio source, and steering a pair of beams separated by a few degrees. This creates a binaural image for the user resulting in a noticeable enhancement of source width. The sensing can be further extended using multiple position sensors enabling simultaneous fixation for multiple listeners (Figure 11).

5.2. Reflection Steering with 3D Pointer

The Wii Remote mentioned previously was further enhanced by the addition of a 3-axis tilt-compensated compass for absolute orientation sensing [29]. Using the orientation data, it is possible to steer a beam along the same direction. When absolute orientation is combined with absolute position sensing, a pointing gesture is possible.

A limitation of the personally fixated source is that only a small proportion of possible beam angles are within arm-span of the listener. Pointing enables intuitive access to distant locations and also keeps other interaction modalities available [30].

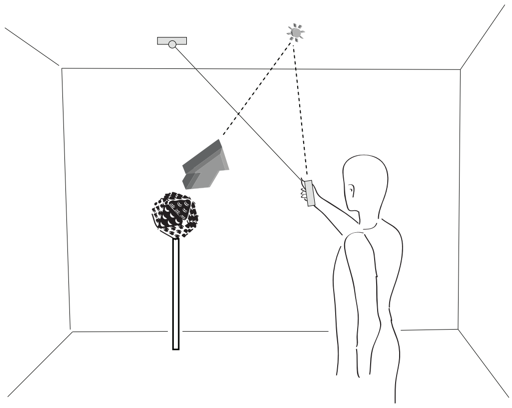

Figure 12: Positioning of reflections by pointing

In combination with an impulsive sound, this gesture interface is particularly useful for rapid assessment of reflective properties of various acoustic surfaces in a concert hall [31] (Figure 13).

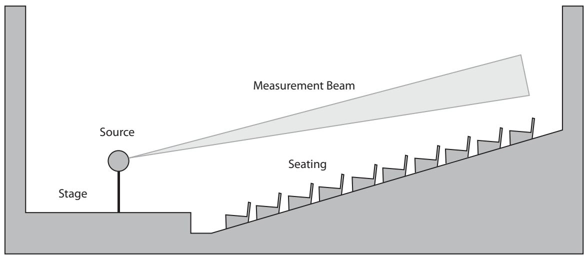

Figure 13: Acoustic probing by reflection steering

5.3. Beam Steering with a Spherical Slider

In stereo mixing consoles, a 1-dimensional slider can be used to pan sources to the left or right. For surround sound, a knob is typically used to pan to a lateral angle. These interfaces support static positioning, and also are contained easily within the postural comfort zone [32]. Directivity control requires setting of two angles simultaneously. To support this gesture in an analogous manner, a 3D haptic controller (the NovInt Falcon shown in Figure 14) is used to implement a spherical slider.

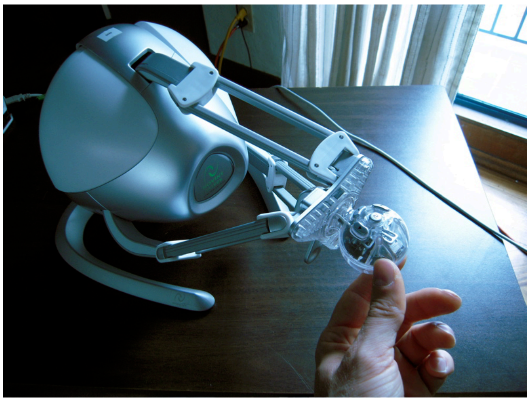

Figure 14: 3D force-feedback haptic controller

A force-feedback system is programmed so that the manipulator is constrained by a spring force onto the surface of a sphere. The range of mechanical motion limits the sphere to a diameter of about 5cm. Viscous damping plus a gravity-countering buoyant force is used to hold the manipulator in a static position when it is released.

This gesture interface supports not only beam steering, but also gain manipulation by means of pushing towards or pulling away from the sphere surface. Localized force detents can also be added to the model to mark cardinal positions, front, left, right, etc. Most interestingly, the interface has a dynamic property as the user can easily find and trace out various stable orbits (Figure 15). These trajectories can be followed in real-time or recorded for later reproduction. Unlike an actual sphere, the virtual haptic sphere provides the user with equal ease of access to all sides of the object.

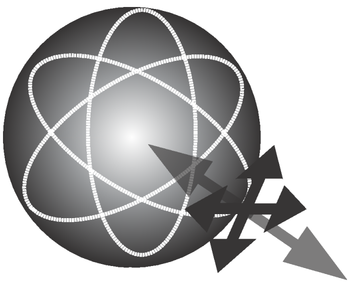

Figure 15: Kinematics of spherical surface constraint: angular steering, stable orbit paths and spring-resisting radial push-pull

A drawback of the haptic controller is that it is difficult to ascertain the state of the spherical slider by visual inspection. The user must grasp the manipulator and probe the force system to determine the location. A 3D visualization provided a visual representation of current state of the beam steering (Figure 16).

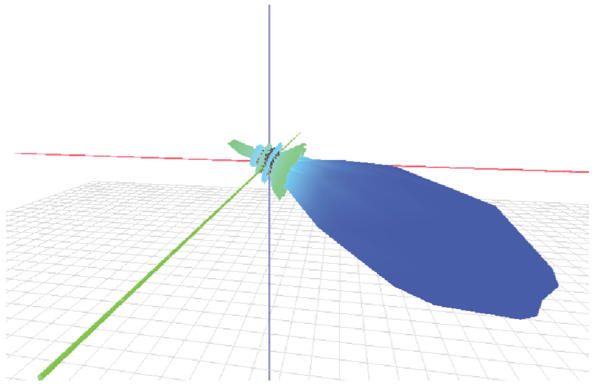

Figure 16: Beam steering 3D visualization

5.4. Kinetic Interactions: Lighthouse, Sprinkler, Chimes

When used with overhead mounting, the Gametrak positioning sensor supports pendulum-like kinetic interactions (also called "pendaphonics" when used for interactive audio applications [28]). By tuning of the weight of the bob, the system has neutral buoyancy so that the pendulum can be "parked" at various levels. Using the position-based steering described in Section 5.1, this interface supports the throw and catch gestures. By controlling the direction and force of throw, the user can setup various stable oscillations of the bob by influencing swing amplitude, eccentricity and procession. These oscillations have a typical period of about 0.5 to 1Hz resulting in excellent perception of beam directivity rotation. This interface supports steering of the beam away from the user (similar to pointing), but not static positioning.

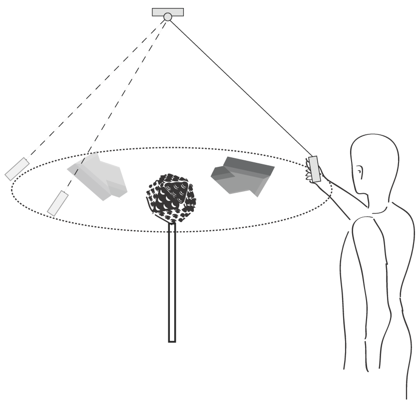

Figure 17: Kinetic beam-steering with a position-sensing pendulum

The so-called lighthouse effect sends a constant beam of noise, an audio-powered illuminator, that can be thrown around the room. In an expo attended by the general public, the author found that this effect was a simple and effective means to demonstrate the auditory percept of directional steering. A humorous extension to this demonstration uses gated noise with two different speeds of AM, depending on clockwise or counter-clockwise azimuthal rotation, sounding like a typical angle-sweeping lawn sprinkler.

The chimes demonstration is a 16-step sequencer where the bob crossing through an azimuthal angle (quantized to 22 degrees) triggers playback of a wind-chime sample to a beam pointing in that direction. The eccentricity of the pendulum swing then corresponds with a monotonic time-distortion in the sequencer rhythm. Catching and throwing in the opposite direction induces a time-reversal.

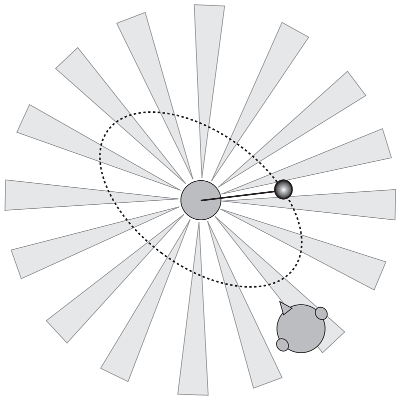

Figure 18: Kinetic interaction with angle-crossing triggered sequencer

5.5. Auto-kinetic Interactions: Gong

This interaction study attempts to link physical model synthesis with a sense of spatial physicality. A radio drum baton [33] is used to excite a waveguide mesh simulation of a gong. Using a mass-spring model, a kinetic simulation causes the gong synthesis signal, reproduced in a dipole pattern, to undergo directional oscillation depending on the strike velocity. The parameter scaling of this interaction requires some care as there is no natural mapping from strike velocity to directivity rotation. Careful measurement of the range and non-linear scaling of strike velocity is required, which is then transformed to the log-scale rotation rates that retain the coherent sense of motion (see Section 4.1).

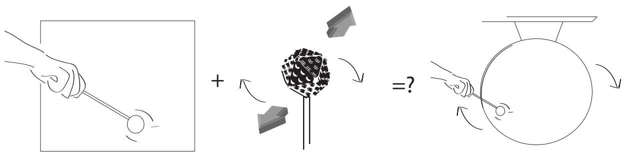

Figure 19: Augmenting physical model synthesis with spatial physicality

5.6. Live Instrument Reinforcement: Augmented Violin

Body-motion is known to be an important and non-incidental aspect of live musical performance, where musicians intentionally control the position of first reflections [34]. In this case study an attempt was made to ascertain if the spherical loudspeaker array could produce an enhanced sense of presence in response to the gestures of a performer.

A piezo contact microphone and a Wii Remote were attached to a violin. Using the tilt sensor, the range of easily accessible pitch and roll of the violin body was measured and found to be approximately +/- 30 degrees pitch and +/- 10 degrees roll. The tilt angles were scaled linearly to a beam steering angle of 180 +/- 90 degrees azimuth (away from the listener to ensure good reflection localization) and +/- 45 degrees elevation. The live signal from the piezo microphone was then transmitted through the beam. Toggling the directivity rotation effect on and off revealed a significant difference in perceived presence between static and dynamic sound reinforcement. A second signal from the microphone feed passing through a real-time granulator was used to diffuse additional sound at random angles for an enhanced sense of envelopment.

Figure 20: Violin with contact microphone and attached tilt-sensor

5.7. Touch the Sound

Two sources of inspiration contributed towards this interaction case study. One is a practical method for testing channel connectivity in a many-channel system where an object such as the hand or a sheet of paper is used to obscure the loudspeaker while it is playing a broadband noise signal. This induces a spectral change that is more easily distinguished (and safer) than probing with a level change.

A second source is the Sphere Spatializer developed for an immersive application in the UCSB Allosphere by Dan Overholt [35]. In the case of the compact spherical loudspeaker array, it was decided to integrate the sensors directly onto the object, rather than on a secondary representation. A continuous control was also desired rather than discrete buttons. Several infrared proximity sensors were installed on the surface of the array, using the gaps near corners of the icosahedron as mounting locations.

Interaction with this interface consists of moving the hands towards the spherical array (Figure 21). The user is able to "touch" the sound and shape its spectrum by occlusion and diffraction in the very-near field (see section 3.3.3). This is an unusual interaction modality for auditory display that emphasizes phenomenal embodiment.

5.8. Vanishing Point

This interaction study includes manipulation of the apparent source width and apparent source location by means of explicit control over position and time delay of first reflections

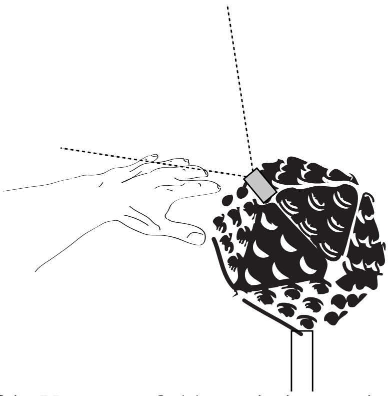

Figure 21: Very-near field proximity sensing and acoustic occlusion by the hand

and direct beam. The effect is modulated as a function of listener proximity, measured from a circular infrared ranging array mounted on the support pole just below the spherical loudspeaker (Figures 22 and 23).

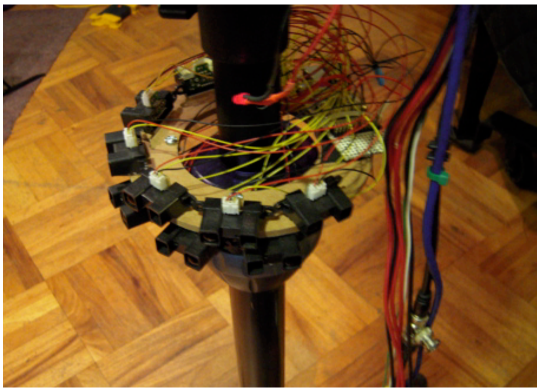

Figure 22: Circular proximity sensing array-active

The demonstration has two modes: exaggerated reality and anti-reality. In the exaggerated case, as the user moves away from the array, the source appears to shrink in volume and move into the distance faster than expected. In the anti-reality case, as the user steps towards the array, the source appears to grow in volume and move into the distance.

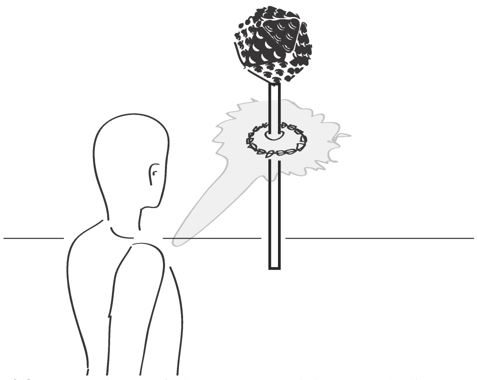

Figure 23: Sensing of listener position and distance in the plane

Because this audio effect manipulates reflections with respect to the listener position, it cannot be expected to work correctly for multiple simultaneous listeners.

6 CONCLUSION

6.1. On Experiential Interaction

The use of human-interaction for the exploration of auditory phenomena created by directivity control is important in accelerating the speed with which new ideas can be hypothesized and experientially tested. The use of physical controls such as absolute position and orientation sensing was found to be particularly useful, as the perception of certain configurations of beam steering are difficult to perceive without an external reference (e.g. reflections falling on confusion lines). Prior to this work, much of the beam steering control was carried out with numerical automation, which is a tedious process and makes difficult the simultaneous adjustment of multiple parameters.

6.2. Sensor Integration in Compact Spherical Loudspeaker Arrays

In prior work it was identified that the integration of environmental sensing extended the applications of compact spherical loudspeaker arrays [36]. Several of the case studies shown use sensors placed in close proximity to the loudspeaker, including directly on its surface. Furthermore, knowledge of the position and distance of walls is required to make full use of the reflection positioning capabilities of the array. This work extends those applications with more scenarios where comprehensive sensor integration is an important factor for the engineering of many-channel loudspeakers.

6.3. A Simplified Model for Angular Beam Steering in Rooms

A simplified model for beam steering can be extracted from this work, summarizing the perception of beam directivity in a typical rectangular room according to angle and distance. This model may be used as a basis for higher-level control systems that abstract over the beam-forming engine and present controls in terms of auditory scene attributes.

Schematically, the model is shown in Figure 24. The direct path beam points towards the listener and covers an azimuthal angle of about +/- 20 degrees. Behind and to the side of the array is the reflection-generating zone, where immersive presence effects are active. To the side of the listener (and also up towards the ceiling) is the diffusing zone where localization accuracy is poor and envelopment effects are active.

6.4. Connecting Schema to Scenes

Auditory spatialization schemas (path, containment, etc., see Section 2) are a type of design pattern for spatially manipulating auditory scenes. However, interface designers, sound designers and composers use design patterns that are specified in a generalized meta-domain (such as Icon, Progress [37]). The process of implementing high-level

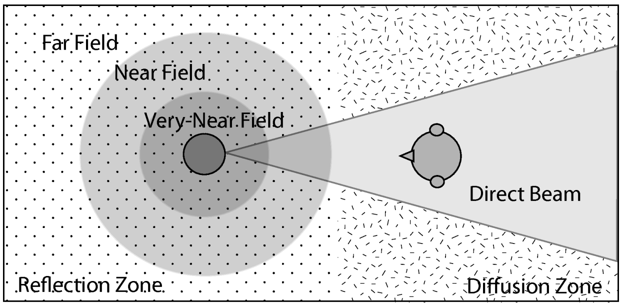

Figure 24: Simplified model of perception for a beam-forming spherical loudspeaker situated in a typical room

designs on a specific interface involves the specification and application of transformation rules [38] [39]. The final result of this process is the sonic rendering of data in a perceptual frame of reference.

The application of this process for auditory displays with spatial directivity control is the structural elaboration and measurement of connections between spatialization schema and the resulting dimensional and immersive attributes of the rendered auditory scene. To this end, the graph shown in Figure 25 illustrates the basic nature of this interplay.

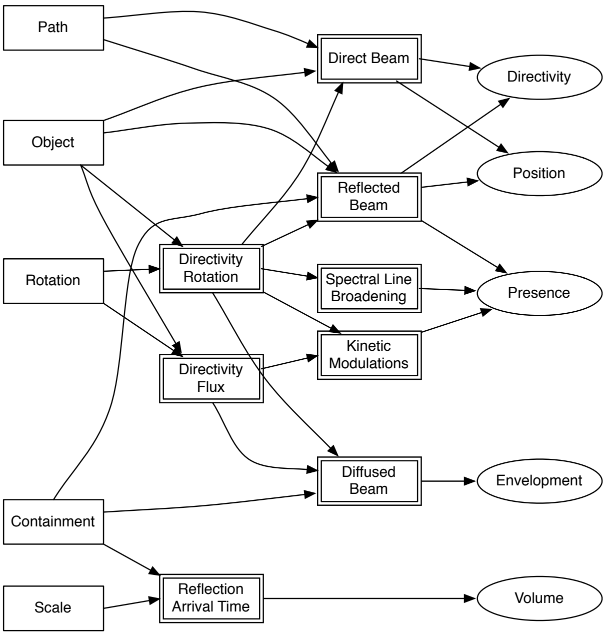

Figure 25: Structural connections between spatialization schema, parametric directivity effects, and auditory scene attributes

To interpret this graph, begin with a desired high-level schemata on the left, proceed through the directivity effects in the middle, and trace these to the low-level scene attributes on the right. For example, to invoke the concept of containment: beam steering is used to create coherent specular reflections and correlated diffuse reflections, resulting in the perception of presence and envelopment. However arrows only show enhancement, not prerequisites. For example the specular reflections alone may be sufficient. Working backwards, envelopment can be achieved without containment, in fact even without the concept of an object, by using diffuse beams in isolation.

7 ACKNOWLEDGEMENTS

Thanks to: Adrian Freed for guidance with this project and for ideas related to use of beam forming for early reflection control, to David Wessel and other researchers at CNMAT for pioneering work in theory and engineering of spherical loudspeaker arrays, to John MacCallum for assistance with programming and the sprinkler idea, and to Franz Zotter for contributions to the real-time beam steering system and the binaural beam idea. Thanks to Meyer Sound Labs, Inc. for financial support of this research.

REFERENCES

[1] D. Wessel, "Instruments That Learn, Refined Controllers, And Source Model Loudspeakers," *Computer Music Journal*, vol. 15, pp. 82–86, 1991. [Online]. Available: http://cnmat.berkeley.edu/publications/instruments_learn_refined_controllers_and_source_model_loudspeakers

[2] "The Center for New Music and Audio Technologies," http://cnmat.berkeley.edu, 1987.

[3] O. Warusfel and N. Misdariis, "Directivity Synthesis with a 3D Array of Loudspeakers, Application for Stage Performance," in *Proceedings of the DAFX*, 2001.

[4] V. Tarnow, "Sound Radiation from Loudspeaker Systems with the Symmetry of the Platonic Solids," Tech. Rep., 1974.

[5] G. Weinreich, "Directional Tone Color Loudspeaker," 2001.

[6] M. Pollow, "Variable directivity of dodecahedron loudspeakers," 2007.

[7] P. R. Cook, G. Essl, G. Tzanetakis, and D. Trueman, "N » 2: Multi-speaker Display Systems for Virtual Reality and Spatial Audio Projection," 1998. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=?doi=10.1.1.10.7241

[8] D. Lock and G. Schiemer, "Orbophone: a new interface for radiating sound and image," in *Proceedings of the 2006 conference on New interfaces for musical expression*. IRCAM—Centre Pompidou Paris, France, France, 2006, pp. 89–92.

[9] F. Zotter and A. Sontacchi, "Icosahedral Loudspeaker Array," IEM, Tech. Rep., 2007.

[10] P. Kassakian and D. Wessel, "Characterization of Spherical Loudspeaker Arrays," San Francisco, CA, 2004, p. Convention paper #1430.

[11] R. Avizienis, A. Freed, P. Kassakian, and D. Wessel, "A Compact 120 Independent Element Spherical Loudspeaker Array with Programmable Radiation Patterns," Paris, France, 2006, p. Convention Paper No. 6783.

[12] W. Buxton, *Sketching user experiences: getting the design right and the right design*. Amsterdam;

Boston: Elsevier/Morgan Kaufmann, 2007.
[13] Z. Bilda, L. Candy, and E. Edmonds, "An Embodied Cognition Framework for Interactive Experience," *CoDesign*, vol. 3, no. 2, pp. 123–137, 2007.
[14] A. Hunt, T. Hermann, and S. Pauletto, "Interacting with Sonification Systems: Closing the Loop," *Information Visualisation, International Conference on*, vol. 0, pp. 879–884, 2004.
[15] S. R. Klemmer, B. Hartmann, and L. Takayama, "How bodies matter: five themes for interaction design," in *DIS '06: Proceedings of the 6th conference on Designing Interactive systems*. New York, NY, USA: ACM, 2006, pp. 140–149.
[16] P. Milgram and H. C. Jr., "A Taxonomy of Real and Virtual World Display Integration," 1999.
[17] G. Kendall and M. Ardila, *The Artistic Play of Spatial Organization: Spatial Attributes, Scene Analysis and Auditory Spatial Schemata*, 2008, pp. 125–138. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-85035-9_8
[18] E. Corteel, "Synthesis of directional sources using wave field synthesis, possibilities, and limitations," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, pp. 188–188, 2007.
[19] T. Caulkins, E. Corteel, and O. Warusfel, "Analysis of certain challenges for the use of wave field synthesis in concert-based applications," in *Proceedings of the DAFx*, 2004, pp. 250–255.
[20] F. Zotter and M. Noisternig, "Near and farfield beamforming using spherical loudspeaker arrays," in *Proceedings of the 3rd AAAA Congress*. Alps-Adria Acoustics Association, 2007.
[21] F. Zotter, "Acoustic Directivity Control for Spherical Loudspeaker Arrays," Ph.D. dissertation, Institute of Electronic Music and Acoustics, To appear.
[22] P. Kassakian, "Magnitude Least-Squares Fitting via Semidefinite Programming with Applications to Beamforming and Multidimensional Filter Design," Philadelphia, Pennsylvania, 2005.
[23] A. Freed, "Spectral Line Broadening with Transform Domain Additive Synthesis," in *International Computer Music Conference*. Beijing, China: International Computer Music Association, 1999, pp. 78–81. [Online]. Available: http://cnmat.berkeley.edu/publications/spectral_line_broadening_transform_domain_additive_synthesis
[24] J.-C. Risset and D. Wessel, *Exploration of Timbre by Analysis Synthesis*, 2nd ed. San Diego: Academic Press, 1999, pp. 113–169. [Online]. Available: http://cnmat.berkeley.edu/publications/exploration_timbre_analysis_synthesis
[25] R. Kronland-Martinet and T. Voinier, "Real-time perceptual simulation of moving sources: application to the Leslie cabinet and 3D sound immersion," *EURASIP J. Audio Speech Music Process.*, vol. 2008, pp. 1–10, 2008.
[26] C. A. Henricksen, "Unearthing the mysteries of the Leslie Cabinet," *Recording Engineer/Producer Magazine*, 1981.
[27] A. Schmeder and A. Freed, "A low-level embedded service architecture for rapid diy design of real-time musical instruments," in *New Interfaces for Musical Expression*, 04/06/2009 2009. [Online]. Available: http://cnmat.berkeley.edu/node/6902
[28] A. Freed, D. McCutchen, A. Schmeder, A.-M. Skriver Hansen, and W. J. C. A. Overholt, Dan;Burleson, "Musical Applications and Design Techniques for the Gametrak Tethered Spatial Position Controller," in *SMC 2009*, 2009.
[29] A. Schmeder, "Hacking a Wii Remote for Full 3D Absolute Position and Orientation Sensing," http://cnmat.berkeley.edu/node/7649, 2009.
[30] R. Bolt, "Put-that-there: Voice and gesture at the graphics interface," in *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*. ACM New York, NY, USA, 1980, pp. 262–270.
[31] R. Schwenke, F. Zotter, A. Schmeder, and D. Wessel, "Room Acoustics Measurements with an Approximately Spherical Source of 120 Drivers," in *ASA-EAA Joint Conference*, Paris, FR, June 29 2008.
[32] M. Kolsch, A. Beall, and M. Turk, "The postural comfort zone for reaching gestures," in *HFES Annual Meeting Notes*, 2003.
[33] M. Mathews, "Three dimensional baton and gesture sensor," Dec. 25 1990, US Patent 4,980,519.
[34] M. Wanderley, P. Depalle, and O. Warusfel, "Improving Instrumental Sound Synthesis by Modeling the Effects of Performer Gesture," in *Proc. of the 1999 International Computer Music Conference*. ICMA, 1999.
[35] G. Wakefield, D. Overholt, L. Putnam, W. Smith, M. Novak, and J. Kuchera-Morin, "The AlloBrain: an interactive stereographic, 3D audio immersive environment," in *CHI Conference Workshop on Sonic Interaction Design*, 2008.
[36] A. Freed, A. Schmeder, and F. Zotter, "Applications of Environmental Sensing for Spherical Loudspeaker Arrays," in *IASTED Signal and Image Processing*, Hawaii, 2008.
[37] M. Adcock and S. Barrass, "Cultivating design patterns for auditory displays," in *Proc. ICAD*, vol. 4, 2004.
[38] C. Frauenberger, T. Stockman, V. Putz, and R. Holdrich, "Design patterns for auditory displays," *PEOPLE AND COMPUTERS*, vol. 19, p. 473, 2005.
[39] S. Daudé, L. Nigay, and C. IIHM, "Design process for auditory interfaces," in *International Conference on Auditory Display, Boston, USA*, 2003.