APPLICATIONS OF ENVIRONMENTAL SENSING FOR SPHERICAL LOUDSPEAKER ARRAYS

Adrian Freed
Center for New Music and Audio Technologies
University of California, Berkeley
1750 Arch Street, Berkeley, CA, USA
email: adrian@cnmat.berkeley.edu

Andy Schmeder
Center for New Music and Audio Technologies
University of California, Berkeley
1750 Arch Street, Berkeley, CA, USA
email: andy@cnmat.berkeley.edu

Franz Zotter
Institut für Elektronische Musik und Akustik
Universität für Musik und darstellende Kunst
Inffeldg. 10, 8010 Graz, Österreich
email: zotter@iem.at

ABSTRACT
This paper surveys emerging applications of spherical loudspeaker arrays facilitated by environmental sensing.

KEY WORDS
Spatial audio, inertial sensing, environmental sensing, location-aware signal processing, loudspeaker arrays

1 Introduction

Signal processing for loudspeaker arrays can be used to control complex spatial distributions of acoustic energy. Because of the expense and challenges of providing individual control over a unique sound source for each driver, the first generation of spherical loudspeaker arrays were mainly used to create approximations of omnidirectional sound sources (e.g. Tarnow [1]). The availability of compact Class-D amplifier chipsets and the computational and I/O throughput of modern computers now makes it possible to individually process and deliver audio streams to and from individual drivers/transducers in large arrays. Spherical and planar arrays of hundreds of drivers are the current state of the art, and prototypes are in development for thousands of drivers.

The small size and portability of spherical arrays, flexible mounting options, and their ability to control acoustic energy with uniform accuracy in all directions enables them to be employed in a wide variety of acoustic applications. The incorporation of sensing technologies can extend and enhance these applications as well as improve the user-interface and runtime reliability.

This paper proceeds with a description of the three sensing domains of interest–internal, local and global external environments, followed by an overview of the applications enabled by sensing in these areas. A theoretical section describes at a high level the operation and control of a spherical loudspeaker array. The integration of these ideas in practical terms is then described, followed by a discussion of extensions to other audio technologies.

2 Sensing Modalities

2.1 Global Environment

Inertial, magnetic and global position sensing can be employed to establish the orientation and location of loudspeaker arrays with respect to the space they are installed in. This allows spatial steering parameters to be expressed in world coordinates–a more intuitive control space for users of the room.

2.2 Internal Environment

Increases in channel count have significant benefits for reproduction of highly detailed spatial radiation patterns (e.g., as demonstrated by Kassakian [2]), however the probability of mechanical or electrical faults increases proportionally with the number of elements. In addition, time-dependent properties of loudspeakers such as thermal limits cannot be precomputed for causal realtime applications. Therefore, the integration of intrinsic environmental sensors such as per-channel monitoring of power supply current, voltage and driver temperature is crucial for the proper operation of large arrays.

2.3 External Environment

The incorporation of external environmental awareness enables adaptive signal processing in response to estimations regarding the direction and distance of objects and people relative to the array.

3 Applications Overview

Spherical arrays in particular are optimized for applications involving uniform spatial selectivity over all directions from a central point. With individual digital control over each array element, signal processing algorithms

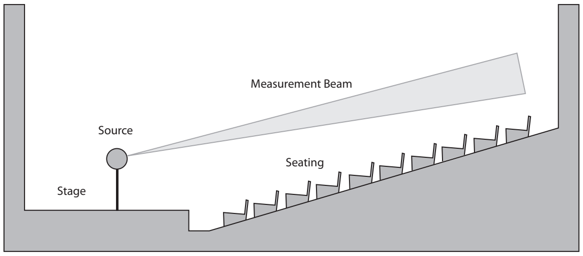

Figure 1. Probe measurement beam in a concert hall

can control directivity patterns in realtime at audio rates. Applications that exploit this feature include:

- **Spatial pattern reproduction:** Reconstruction and simulation of time and frequency dependent directitvity of acoustic sources, e.g., musical instruments, especially woodwinds, gongs, strings (Weinreich [3], Warusfel [4], Zotter [5], Pollow [6]).

- **Spatial motion and effects:** Simulations of acoustic objects undergoing rotation, e.g. simulations of rotating loudspeakers (Henricksen [7]) and recreation of kinesthetic dynamics of human performers who intentionally move instruments to engage direction-specific room reflections (Wanderley [8]) and doppler-shifts.

- **Room acoustics measurement:** The directional steering of a narrow beam serves as an acoustic probe for measurement of concert halls, depicted in Figure 1 (Schwenke [9]). Additionally, calibrated spherical arrays can produce omnidirectional sources with better spatial uniformity than traditional impulse-response generators such as the balloon-pop and starter pistol.

- **Controlled acoustic reflection:** An acoustic reflector (such as a hard, flat surface) can be used to transform the basic energy distribution from spherical to another geometry, e.g., to planar by placement of the array at the focus of a parabola, or to hemispherical by placement next to a wall. Knowledge about the geometry of a room can be leveraged to control acoustic energy in terms of sound-quality metrics such as clarity (the ratio of direct to indirect sound).

- **Microphone feedback control:** Positioning of a steerable null in the direction of a microphone can provide control over feedback in sound reinforcement applications.

- **Location-based entertainment:** Acoustic beams can be targeted towards listeners in specific areas of an environment, possibly reacting in response to estimates of their current position (Momeni [10]).

4 Spherical Array Control Theory

The theoretical development in this section is presented in a simplified form–the intent is to establish a basic technical language for describing applications of the spherical array. For a complete mathematical treatement see Teutsch [11].

Euler's equation of momentum and the acoustic wave equation can be generalized to any orthonormal coordinate system; spherical coordinates being the optimal representation for spherical loudspeaker arrays. The solution to the wave equation in spherical coordinates gives rise to the spherical harmonics transform.

4.1 Pattern Analysis and Reconstruction

Angular Radiation Patterns Let $g(\mathbf{\Omega})$ represent an angular radiation pattern in amplitude vs spherical angle $\mathbf{\Omega} = (\phi, \theta)$. The spherical harmonics transform of g yields the expansion coefficients g_{nm} by convolution with spherical harmonics of order n, degree m, Y_n^m.

$$g_{nm} = \mathcal{SHT}\left\{g(\mathbf{\Omega})\right\}_n^m, \tag{1}$$

$$\mathcal{SHT}(g(\mathbf{\Omega}))_n^m := \iint_{\mathbb{S}^2} g(\mathbf{\Omega}) \cdot Y_n^m(\mathbf{\Omega})\, \mathrm{d}\mathbf{\Omega}, \tag{2}$$

$$n = 0 \ldots \infty, m = -n \ldots n \tag{3}$$

The inverse transform is:

$$\mathcal{SHT}^{-1}(\boldsymbol{g}) = \sum_{n=0}^{\infty} \sum_{m=-1}^{n} g_{nm} \cdot Y_n^m(\mathbf{\Omega})^\star \tag{4}$$

Loudspeaker Patterns For a discrete array of L elements with angular positions $\mathbf{\Omega}_l$, $l = 1 \ldots \mathrm{L}$, the bandwidth of an angular pattern is limited to spherical harmonics less than order N proportional to $\sqrt{L}$, assuming the angular positions are near-uniformly distributed over the surface of the sphere. In this situtation, a discrete approximation to $\mathcal{SHT}$ is used to control angular patterns over the given arrangement of loudspeakers. In order to address these coefficients by a single index, let $p = n^2 + n + m + 1$ and $Y_n^m = Y_p$. Note that $1 \leq p \leq (N+1)^2$

Suppose that the velocity pattern at the surface of the array for each element is is well approximated by bandlimited Dirac delta distributions at the angular positions $\mathbf{\Omega}_l$, $l = 1 ... \mathrm{L}$ of the elements. In spherical harmonics these deltas can be gathered into the *loudspeaker encoding* matrix $\boldsymbol{C}$

$$\boldsymbol{C} = [\boldsymbol{c}_\mathbf{1}, \ldots, \boldsymbol{c}_\mathbf{L}], \tag{5}$$

$$\boldsymbol{c}_l = [\mathcal{SHT}(\delta(\mathbf{\Omega} - \mathbf{\Omega}_l))] = [Y_p(\mathbf{\Omega}_l)]. \tag{6}$$

Optimal Angular Reproduction An optimal decoder matrix $\boldsymbol{D}$ for the reproduction of angular patterns on the array surface is given by the pseudo-inverse of $\boldsymbol{C}$

$$\boldsymbol{D} = \boldsymbol{C}^{\boldsymbol{t}} \left(\boldsymbol{C}\boldsymbol{C}^{\boldsymbol{t}}\right)^{-1}. \tag{7}$$

It contains a set of real-valued weights for reproduction of $g\left(\boldsymbol{\Omega}\right)$ using the transducer signals $\boldsymbol{y}$ and the input signal x.

$$\boldsymbol{y} = \boldsymbol{D}\ \boldsymbol{g}\ x. \tag{8}$$

Optimal Radial Reproduction For non-omnidirectional patterns, the dispersion of acoustic energy is dependent on wavelength, giving rise to the near-field and far-field effects. Compensation for this effect requires an equalization filter per-order $\boldsymbol{H}_n\left(r\right)$ at every frequency to ensure the desired spectral balance is achieved at the target radius r (Zotter [5]). It is convenient to notate the radially-equalized transducer signals and the input x in the frequency domain

$$\tilde{\boldsymbol{y}} = \boldsymbol{D}\ \boldsymbol{H}\left(r\right)\ \boldsymbol{g}\ x. \tag{9}$$

Note, however, that $\boldsymbol{H}_n\left(r\right)$ is efficiently and accurately implemented with an IIR filter.

Angular Aliasing The finite realization of $\mathcal{SHT}$ requires a low-pass filter to prevent aliasing of patterns that exceed the reproduction capability of the array. The design of windowing function $\boldsymbol{W}_N$ with coefficients $w(n)$ attenuates terms of increasing order n and influences the tradeoff between ripple and main-lobe width in the spatial domain. Reconstruction on the array is then given by

$$\tilde{\boldsymbol{y}} = \boldsymbol{D}\ \boldsymbol{W}_N\ \boldsymbol{H}\left(r\right)\ \boldsymbol{g}\ x \tag{10}$$

4.2 The Steerable Beam

Given a 1-dimensional signal x, a beam having the maximum possible angular selectivity transmitting x in the direction $\boldsymbol{\Omega}_d$ can be constructed using the order-N-limited spherical harmonics expansion of a spherical dirac-delta function, $\delta\left(\phi - \phi_d, \theta - \theta_d\right)$.

$$\tilde{\boldsymbol{y}}_{beam} = \boldsymbol{D}\ \boldsymbol{H}\left(r\right)\ \boldsymbol{W}_N\ \mathcal{SHT}(\delta\left(\boldsymbol{\Omega} - \boldsymbol{\Omega}_d\right)) \cdot x \tag{11}$$

Figure 2 shows an example of radial amplitude as a function of θ for different choices of window function $\boldsymbol{W}$. The half-power beam-width boundary is at the point of intersection with the dashed line.

4.3 The Steerable Null

A steerable null can be convolved with any radiation pattern g to minimize energy transfer in the direction of $\boldsymbol{\Omega}_d$. For this we use the spherical harmonics transform of $1 - \delta$.

$$\tilde{\boldsymbol{y}}_{null} = \boldsymbol{D}\ \boldsymbol{H}\left(r\right)\ \boldsymbol{W}_N\ \mathcal{SHT}(1 - \delta\left(\boldsymbol{\Omega} - \boldsymbol{\Omega}_d\right)) \cdot x \tag{12}$$

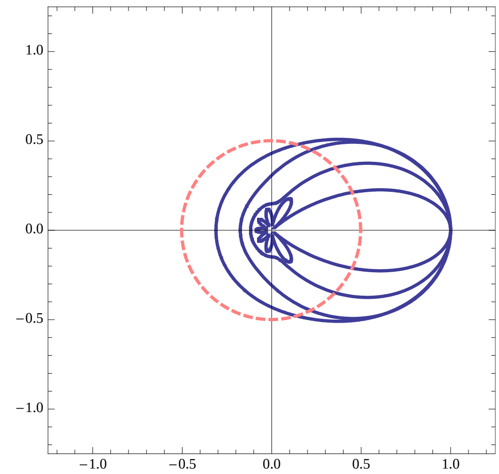

Figure 2. Main-lobe width vs ripple for various windows and finite order expansion, normalized to unity. The dashed line marks half-power.

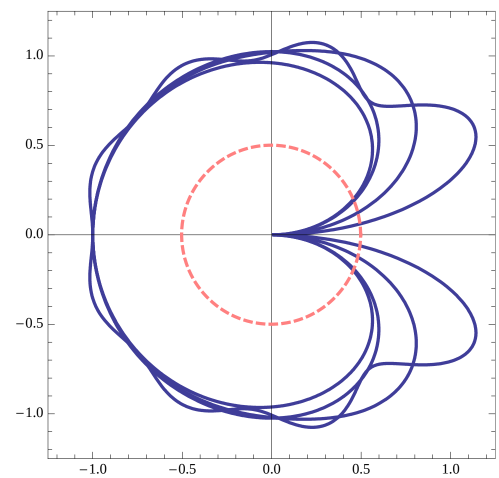

Figure 3. Main-notch width vs ripple for various windows and finite order expansion, normalized to unity. The dashed line marks half-power.

The cutoff and slope of the window function again controls the half-power beam angular size of the notch, shown in Figure 3 for various choices of window.

5 Spherical Array Practice

The use of the spherical array control theory in practical terms is the reproduction and steering in realtime of sound radiation patterns by the coordinated adjustment of real-valued gains on each channel, the adjustment of filter coefficients to compensate for distance-dependent spectral equalization, and the superposition of multiple patterns by time-domain addition. The signal processing chain can be implemented on a general purpose computer or directly on the control-board FPGA inside the array.

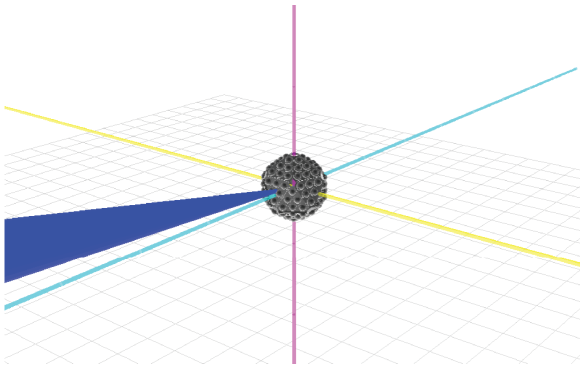

Figure 4. Screenshot of 3D Beam-steering User Interface

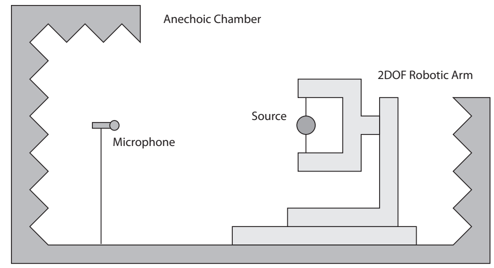

Figure 5. Array calibration with 2DOF robotic arm

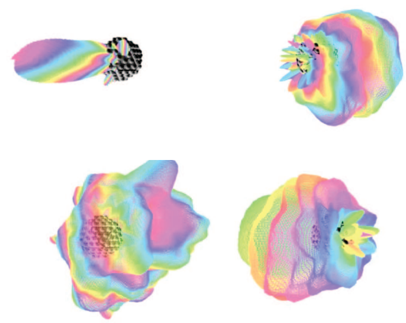

Figure 6. Measured acoustic intensity (displacement) and phase (color) as a function of spherical angle, high to low frequency (left to right, top to bottom).

The CNMAT 120-channel Spherical Array The applications discussed in this section draw from the authors experience in the design, construction of and experiments with a 120-channel spherical array of approximately 20cm diameter (Avizienis [12]).

5.1 Reference Frame Sensing

Spherical coordinates based at the center of an array are the natural representation for signal processing, however in practice these arrays are are used in spaces with natural horizontal and vertical reference points for which a geographic coordinate system is more useful. Room acousticians frequently reference room information provided from plan views in architectural drawings and may be reasoning about multiple arrays with respect to spatial reference points in the room such as the center of the stage or various audience listening points. Clearly, coordinate system transformations are a necessary part of the human-facing control structure of these arrays and these require a good estimate of the location of arrays with respect to points in the space they are installed in.

Frequently, access to rooms such as concert halls is expensive and time constrained so we have found it expedient to use MEMS accelerometers for orientation estimation. This allows arrays to be suspended from cables without requiring complex measurements and tuning of the rigging. Pole mounts can be employed without mechanical compensation for uneven floors. Additionally, in the authors' experience verification of proper orientation of the array can be extremely difficult to ascertain using only aural cues when the array is installed in a reverberant environment. A beam-steering user interface was developed using 3D visualization for this purpose, shown in Figure 4

Another application of orientation estimation involves measurement of an array's spatial scattering and angular crosstalk in an anechoic chamber. Figure 5 depicts an array mounted on a 2DOF robotic arm in an anechoic chamber. A fixed measurement microphone at 3 meters performs spectral analysis at each spherical angle as the array is rotated by the arm. Precision orientation measurements integrated into the array enable compensation for mechanical distortions due to flexibility and play in the arm and mounting hardware. Data collected from our array using this process is shown at various frequencies in Figure 6.

5.2 Fault Isolation and Thermal Protection

Figure 7 shows one of a pair of a new controller-board for our 120 driver spherical speaker array. It containts current, voltage sensors and associated analog-to-digital convertors. Additional data acquisition channels are implemented for temperature monitoring. These measurements are sent contiuously as Open Sound Control (OSC) packets [13] to the host via gigabit Ethernet. The measurements are used for fault isolation and health monitoring.

Thermal sensors are particularly important in room measurements applications where broadband acoustic output is required. When low frequency measurements are sought most energy sent to the speakers is converted to heat rather than radiated sound. With each driver handling 40 watts power input and the entire array occupying a sphere of approximately 20 cm diameter, thermal management is essential. We use a combination of prior measurements of the robustness of the drivers, real-time modeled estimates of cone excursion and temperature, and actual measurements of temperature and current consumption to keep the drivers within a safe operating zone.

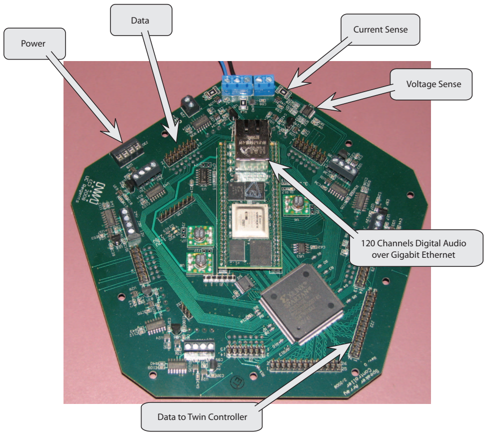

Figure 7. Array controller with intrinsic sensors

5.3 Extrinsic Environmental Sensing

The new controller of Figure 7 has sufficient I/O and processing capacity to process and deliver data from extrinsic environmental sensors, and we are experimenting with the incorporation of additional sensors to monitor objects, room structures and people in the space surrounding the array.

Good methods are available for people-tracking using multiple overlapping camera views, and laser range finders (Park [14]). This work could readily be adapted to spherical arrays that integrate cameras. We are focusing initially on simpler, cheaper approaches using low-cost sensors because of size constraints and to explore interesting interactive applications as early as possible.

An interesting design decision arises in the mechanical construction of spherical arrays with respect to the distribution of sensors around the sphere and the resulting displacement of loudspeakers from their ideal positions which may result. In the pictured array, pentagonal areas at the 12 vertices of the truncated icosahedron are available for the incorporation of environmental sensors.

The angular density of directional sensors necessary for sufficient sampling of the extrinsic environment is determined by both the field-of-view of each sensor as well as the angular specificity of the array's reproduction capability. In the case of the pictured array having an upper limit of degree 9 spherical harmonic, the best case half-power beam width is approximately 30 degrees, setting an upper-limit on the useful angular resolution of the external sensors.

5.4 Feedback Elimination

The 4-point tracking IR imager tracks IR sources placed on a microphone. (Figure 9). A steerable null is convolved with the desired spatial radiation pattern to minimize the pickup of energy by the microphone.

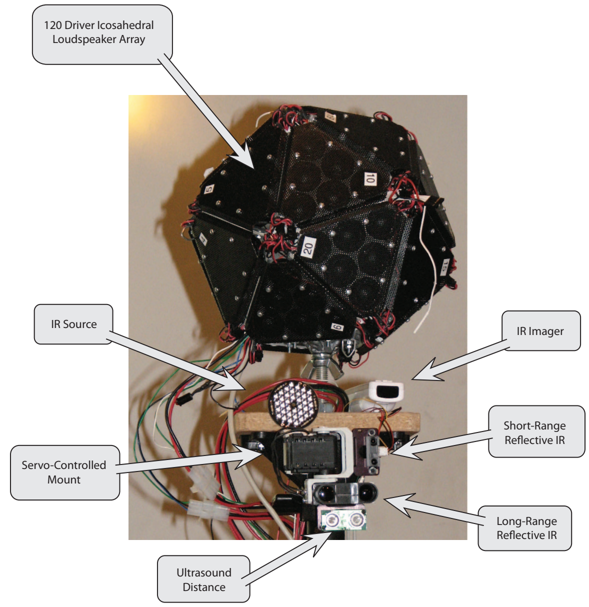

Figure 8. Extrinsic sensors mounted under loudspeaker array

5.5 Location-Based Entertainment

Infrared-reflective dots placed on the shoulders of people moving around the array are illuminated by a bright IR source and tracked by the IR imager shown in Fig. 8. Multiple distance sensors (short and long range IR, ultrasound) are also used to estimate the radial distance between each person and the array. Adaptive signal processing provides each person in the environment with a pair of "binaural beams" (Figure 10) targeted towards their angular position and spectrally equalized according to their distance.

6 Extension to Microphone Arrays

The acoustic theory of the spherical harmonics transform employed in loudspeaker applications also applies to spherical microphone arrays. Beam-forming and null steering have obvious application in sound recording situations particularly in ensemble recording situations. Unlike the interactive applications we have described for loudspeakers, there are applications of microphone arrays where the beam-forming computations can be done off-line (e.g. in recording mastering). In this case it is essential that the sensor's data samples are locked to the audio samples. This is achieved in our controllers by embedding both sensor and audio in streams of time-tagged ethernet packets using the OSC protocol [13].

There are significant engineering differences between speakers and microphone arrays relevant to sensor deployment. An important part of speaker array optimization is maximimizing the proportion of the structure that is radiating acoustic energy. Microphone arrays on the other

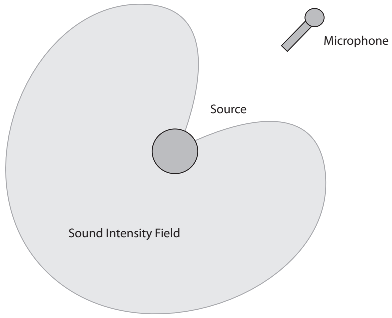

Figure 9. Directed null for feedback control

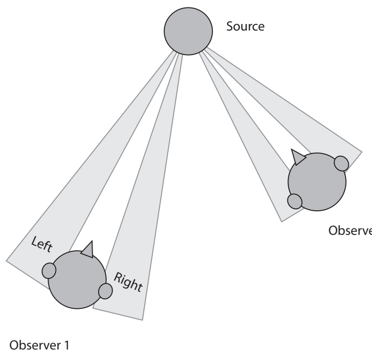

Figure 10. Binaural beams to multiple listeners

hand can afford more surface area for sensors. They also have greater volume available for the associated electronics as they don't need to house large power amplifiers or power supply components. The power components and driver magnets in speakers may also interfer with magnetic field sensors. A challenge with microphone arrays is obtaining sufficient signal to noise ratio so that careful design of the sensor systems is needed to minimize sources of acoustic or electrical interference.

7 Conclusion

Integration of sensors into loudspeaker arrays has many interesting applications for environmentally-dependent acoustical processing, array reliability and measurement accuracy. In addition, the portability of spherical arrays makes more compelling the integration of sensors directly into the array: in deployment scenarios, all cabling is contained within the speaker and sensed data is obtained in a convenient frame of reference. In addition, when sensors are distributed with the same basic geometry as the array itself, the sampled data has the same spatial resolution as the reproduction capabilities of the array. This strategy may be considered advantageous from a cost-perspective, but is not necessarily optimal for applications where sensor data requires a different sampling distribution, or where the sensing modality requires an alternative viewpoint.

Acknowledgements

This work was made possible in part by the generous support of Meyer Sound Laboratories, Inc.

References

[1] V. Tarnow, Sound radiation from loudspeaker systems with the symmetry of the platonic solids, tech. rep., 1974.

[2] P. Kassakian and D. Wessel, Characterization of spherical loudspeaker arrays, San Francisco, CA, p. Convention paper #1430, 2004.

[3] G. Weinreich, Directional tone color loudspeaker, 2001.

[4] O. Warusfel, P. Derogis, and R. Causse, Radiation synthesis with digitally controlled loudspeakers, in *Proceedings of the 103rd AES Convention*, 1997.

[5] 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.

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

[7] C. A. Henricksen, Unearthing the mysteries of the leslie cabinet, *Recording Engineer/Producer Magazine*, 1981.

[8] M. Wanderley, P. Depalle, and O. Warusfel in *Proc. of the 1999 International Computer Music Conference*, ICMA, 1999.

[9] 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.

[10] A. Momeni, What you say, how you say it, 2006.

[11] H. Teutsch, *Modal Array Signal Processing: Principles and Applications of Acoustic Wavefield Decomposition*. Springer-Verlag Berlin Heidelberg, 2007.

[12] R. Avizienis, A. Freed, P. Kassakian, and D. Wessel, A compact 120 independent element spherical loudspeaker array with programmable radiation patterns, Paris, France, p. Convention Paper No. 6783, 2006.

[13] M. Wright, Open sound control 1.0 specification, 2002.

[14] S. Park and M. Trivedi, Multi-person interaction and activity analysis: a synergistic track- and body-level analysis framework, *Machine Vision and Applications*, vol. 18, pp. 151–166, 2007.