A directivity-controlled virtual loudspeaker array for simulating sound sources in rooms

Densil Cabrera[1]**, Ken Stewart**[1] **and Luis Miranda**[1]

[1]Faculty of Architecture, Design and Planning, The University of Sydney, Sydney, Australia

ABSTRACT: This study examines the performance of a directivity-controlled loudspeaker designed for room acoustics measurements. The loudspeaker consists of a single driver mounted on one face of a sealed dodecahedral box. Measurements are made successively with the loudspeaker rotated onto each face (using an acoustically transparent suspension system), and the measured room impulse responses can then be combined *post hoc* to yield the effect of a source with arbitrary directivity (with an accuracy up to 2nd order spherical harmonics).

Conference theme: Computer Science
Keywords: Acoustics

1. INTRODUCTION

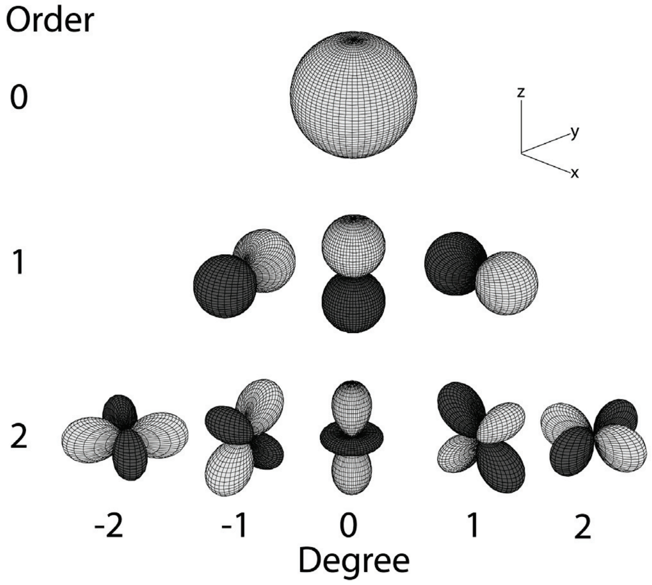
The spatial radiation patterns of real sound sources – such as the human voice – are often complex and can vary in time. Conventional room acoustics measurements use a sound source that is approximately omnidirectional. While such measurements can be used to roughly represent the effect of complex sound sources, there is a growing interest in simulating the directivity of real sound sources using a directivity-controlled measurement sound source. One approach to the simulation of source directivity is to build a physical model of particular sound sources, and an example of this is the head and torso simulator that includes a mouth simulator (i.e., mouth-shaped orifice as part of a full scale physical model of a human, from which sound is radiated). The advantage of this fixed modelling approach is that a particular sound source can be modelled quite accurately, but the device cannot be used for modelling sound sources with other directivity patterns. Furthermore, a fixed physical model is not capable of simulating the *time-varying* directivity of the sound sources that they are emulating. The alternative is to create a generic sound source from an array of independently controlled elements (i.e. loudspeaker drivers). Usually these elements are mounted on the surface of a sphere, and with appropriate control, they can create a wide variety of directivity patterns (including time-varying), within the resolution constraints of the array design (e.g., Avizienis *et al.* 2006; Noisternig *et al.* 2011; Pasqual *et al.* 2010; Pollow and Behler 2009; Rafaely 2009; Zotter *et al.* 2007; Zotter 2009; Zotter and Pasqual 2011). The design and implementation of such sound sources can be quite involved, and in this paper we examine a simple and inexpensive alternative to this problem: creating the array using successive measurements from a single element.

That directivity is an issue in room acoustics can be easily appreciated by comparing the intelligibility of a person's speaking voice when they are either facing towards or away from oneself. The human voice is not very directional (Chu & Warnock 2002), but it is sufficiently so for there to be a dramatic degradation in intelligibility as a talker turns away in some room acoustical environments (note that visual cues from seeing a talker's mouth and face also contribute to intelligibility). Using modelling based on statistical room acoustics, Stewart and Cabrera (2009) examined the extent to which objectively estimated intelligibility is sensitive to source directivity in a range of acoustic environments. They showed that, in noise-free rooms, intelligibility is maximally sensitive to source directivity at a distance around the critical distance (which is the distance for which the direct and diffuse soundfields have equal energy). This distance depends on room volume (larger rooms have greater critical distances) and reverberation time (longer reverberation time is associated with shorter critical distance, but the effect of reverberation time is greater than indicated by critical distance alone).

All of the commonly used room acoustical parameters (described in ISO3382-1) can be sensitive to source directivity, some more so than others. Most obviously, clarity index (which is the ratio of early to late sound energy in a room impulse response, expressed in decibels) and the closely related parameter 'definition' (which is the ratio of early energy to the total energy of the room impulse response, expressed as a percentage) may be substantially influenced by the prominence of the direct sound relative to the subsequent reflections and reverberation. Greater source directivity (towards the receiver) yields greater values for clarity index and definition. Directivity also affects the relative intensity of particular early reflections, which can also influence these parameters – for example if sound from the source is directed towards a nearby surface, the early-reflected energy can be greater than if sound from the source is directed away from this surface. The parameters 'strength factor' (which is an indicator of the loudness of a room, considering the source and receiver positions) and 'early decay time' (which is an indicator of reverberance) are, similarly, likely to be significantly affected by directivity in typical circumstances. Both are likely to be most sensitive to directivity when the reverberant energy is relatively weak – for example when the source and receiver are close to each other, or the room is large or with a high average absorption coefficient. On the other hand, reverberation time is likely to be relatively insensitive to source directivity, because the calculation of reverberation time excludes the first 5 decibels of the reverberant decay (including the direct sound). Reverberation time would be more sensitive to directivity in rooms that have highly uneven spatial distribution of absorption. Interaural parameters (especially early interaural cross correlation) can be highly sensitive to directivity (*cf.* Cabrera, Lee *et al.* 2011).

There have been several studies examining how acoustical parameters could be adapted for directivity-controlled microphone arrays (e.g., Abdou & Guy 1996, Bassuet 2009, Cabrera *et al.* 2010). Even using only first order directivity patterns, such studies show that directivity control can be used to help reveal the relationship between architectural features and acoustic features, and to model the acoustics in terms of human auditory spatio-temporal sensitivity. A directivity-controlled source, such as that considered by the present paper, has similar potential for room acoustics characterization – and in combination with a directivity-controlled microphone it could provide a substantially more powerful description of an acoustic environment than is currently the norm.

Whether it is a source or a receiver, it is possible to describe a compact transducer's directivity pattern in terms of spherical harmonics. The advantage of this approach is that arbitrary spatial patterns can be described around a point by a fixed set of simple functions, each with its respective weight (i.e., this approach is a spatial extension of conventional Fourier analysis). The first nine spherical harmonics are illustrated in Figure 1. These are referred to by their order (order 0 is a unit sphere, order 1 is a cosine or sine function, and order 2 has a higher combination of sine and cosine functions). For each order, there is a set of degrees (two additional degrees with each increment of order). While the pyramid in Figure 1 could be extended indefinitely, in practice the maximum order of a transducer is limited by, among other things, the number of elements in the source or receiver.



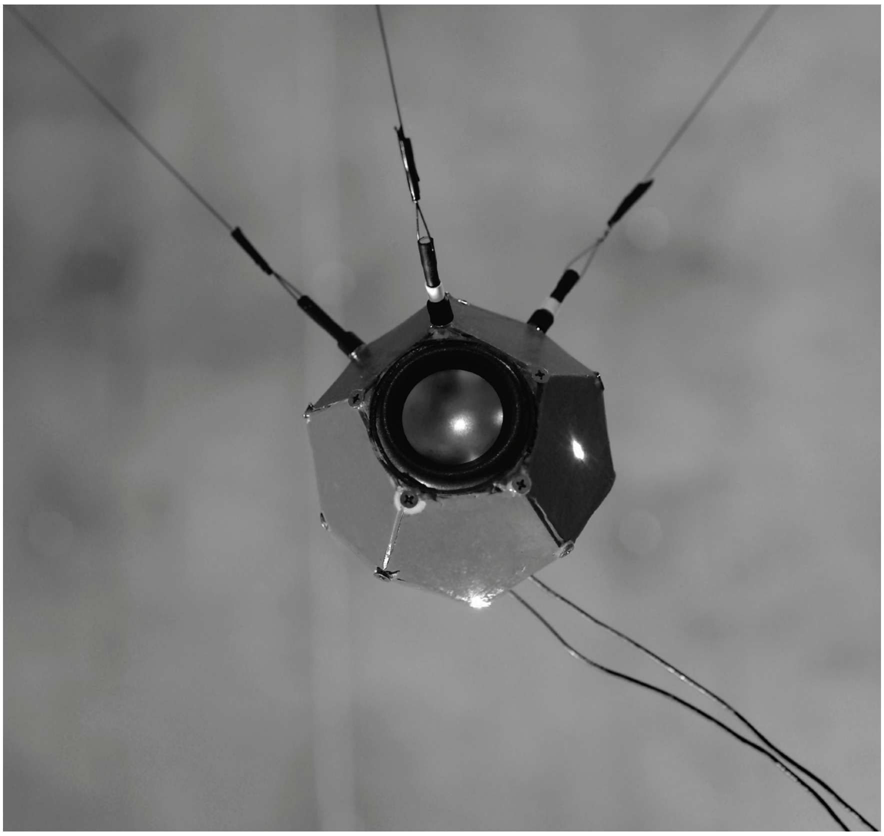
Source: (Author 2011)

Figure 1: Illustration of spherical harmonics up to second order. Darker surfaces have negative coefficients, while lighter surfaces have positive coefficients.

There are many issues that are encountered in designing spherical loudspeaker arrays. The spacing between the centres of loudspeakers influences the highest frequency that can be spatially controlled, following sampling theory. The size of the radiating surface influences the radiation efficiency of the device, and this is a function of frequency. The number of elements determines the spatial complexity of the radiation pattern. There are other practical issues, such as the limitations of loudspeaker drivers, the effect of the loudspeaker enclosure on the drivers, and the effect of enclosure irregularities.

A dodecahedral sound source was built with the ultimate goal of providing the ability to create directivity patterns from an inexpensive loudspeaker enclosure with one driver and multiple measurements. In an accompanying paper (Cabrera *et al.*, 2011), the authors describe the device characteristics and construction. In this paper, the authors also explored the ability of the loudspeaker presented to accurately produce spherical harmonic patterns up to 2nd order at different frequency ranges. A computer simulation model was created to compare the expected theoretical response of a loudspeaker with the characteristics of the device being presented, and the measured response of the device. Twenty-four individual measurements were made and, due to the symmetrical characteristics of a dodecahedron, these measurements were mirrored to produce seventy-two measurement points around a sphere. Correlation coefficient was used as a measure of how well the simulated and measured device could produce first and second order spherical harmonics. When comparing to a perfect spherical harmonic directional pattern, the results show that, for the simulated measurements, the range of frequencies that can be used with certainty to produce first order spherical harmonics spans from the 250 Hz octave band to 4 kHz octave. For the second order spherical harmonics case the usable range drops to the frequencies ranging from the 1 kHz octave band to the 4 kHz octave band. When doing the comparison with the measured results, the first order spherical harmonics' responses are well matched to the simulated results, however when examining the second order spherical harmonics' results, there is a noticeable drop in correlation at all frequency bands. It is important to keep this in mind, as this will influence the beamforming abilities of the loudspeaker.

Figure 2 shows a photograph of the constructed loudspeaker. It has a small driver (49 mm diameter) which is mounted in a sealed dodecahedral enclosure. The loudspeaker is supported by three neodymium magnets, providing an acoustically transparent support system, and allowing the loudspeaker to be repositioned on to each of its twelve faces. Lasers are used to confirm that the loudspeaker repositioning is accurate. A succession of twelve room impulse responses is recorded in each of the twelve loudspeaker orientations, creating a virtual array of twelve independently controlled elements.



Source: (Author 2011)

Figure 2: Prototype loudspeaker, showing the magnetic suspension system and laser-guided positioning (the two bright spots).

2. BEAMFORMING FROM SPHERICAL HARMONICS

There are significant advantages in using spherical harmonics to create beam patterns from a multi-source loudspeaker with individual driver gain control (see Williams 1999). The process of transforming spherical harmonic patterns to a desired beam pattern of arbitrary shapes can be done with simple mathematic operations, as described below, within the constraints of the source's maximum spherical harmonic order.

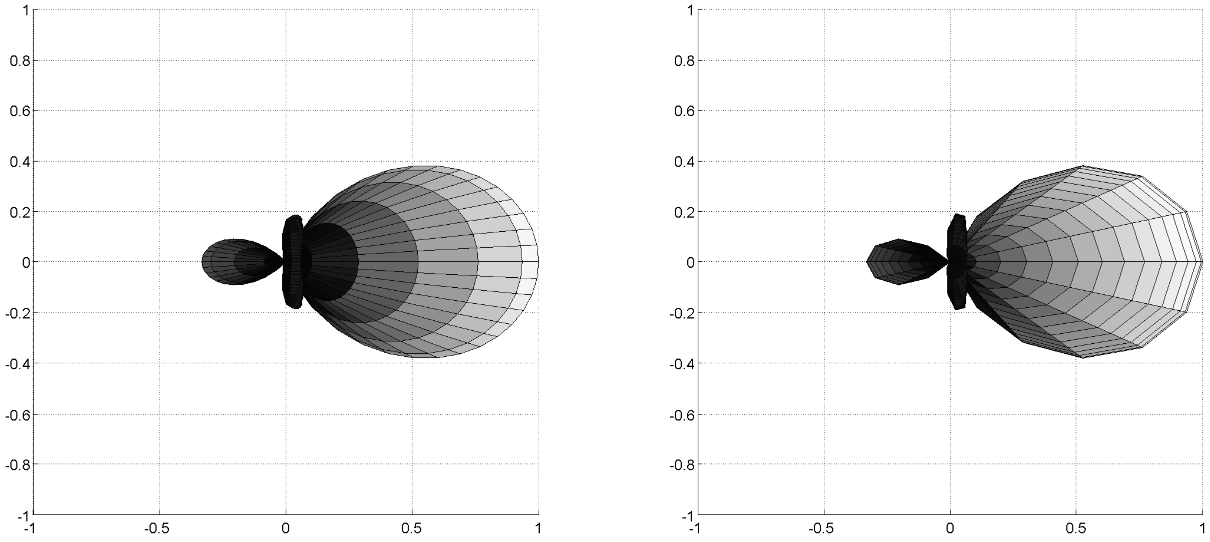
The first step to creating the desired beams is converting the loudspeaker signals into spherical harmonic signals. This is achieved by scaling each signal by applying the gain associated with the spherical angle of emission (Daniel, 2003). For this paper only this initial stage is considered. Further refinement of the spherical harmonic response could be achieved by taking into account the frequency dependent interaction of the loudspeaker enclosure and correcting for it using digital signal processing.

After encoding the signals to spherical harmonic signals by applying the gains described before, the desired beamforming is achieved from the expected beam shape. This is achieved by solving a system of linear equations of the form $Ax = b$.

$$\begin{bmatrix} a_{1,1}x_1 & a_{1,2}x_2 & . & . & . & a_{1,n}x_n \\ a_{2,1}x_1 & . & & & & \\ . & & . & & & \\ . & & & . & & \\ . & & & & . & \\ a_{m,1}x_1 & & & & & a_{m,n}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ . \\ . \\ . \\ b_m \end{bmatrix} \quad (1)$$

In equation 1, *b* is a column vector of the known desired response at the angles of interest, *A* is a matrix of spherical harmonic response values at the same angles as the column vector *b* and *x* is the weight for each spherical harmonic to obtain the desired response in *b*.

Using the technique described above, we examine the theoretical directivity response achievable from second order spherical harmonics, first at one angle and then trying to replicate the directional characteristics of a human talker. To understand the ability of our loudspeaker to create directional beams, we direct a beam at one angle by setting the column vector *b* to zeros, except for the angle that we desire to direct the loudspeaker beam to, where we assign it a value of 1. This directional beam should hold for the frequency range where our loudspeaker is capable of producing second order spherical harmonics. The theoretical directional beam using second order spherical harmonics is shown in Figure 3.

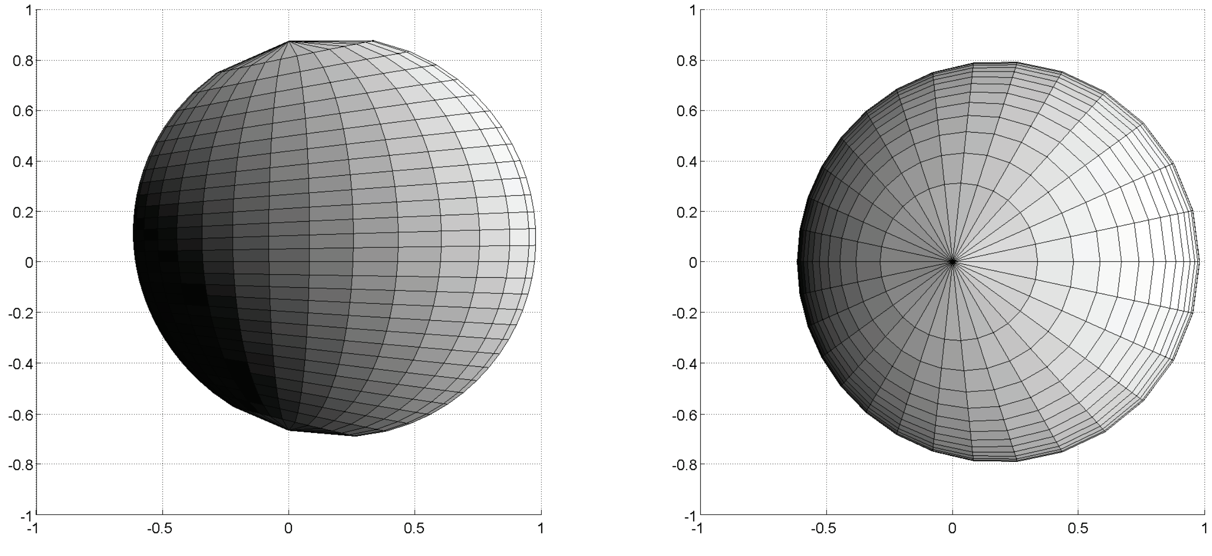


Side View

Top View

Figure 3: Beamforming to one angle from second order spherical harmonics. Values are normalized farfield pressure, with the beam directed along the horizontal axis.

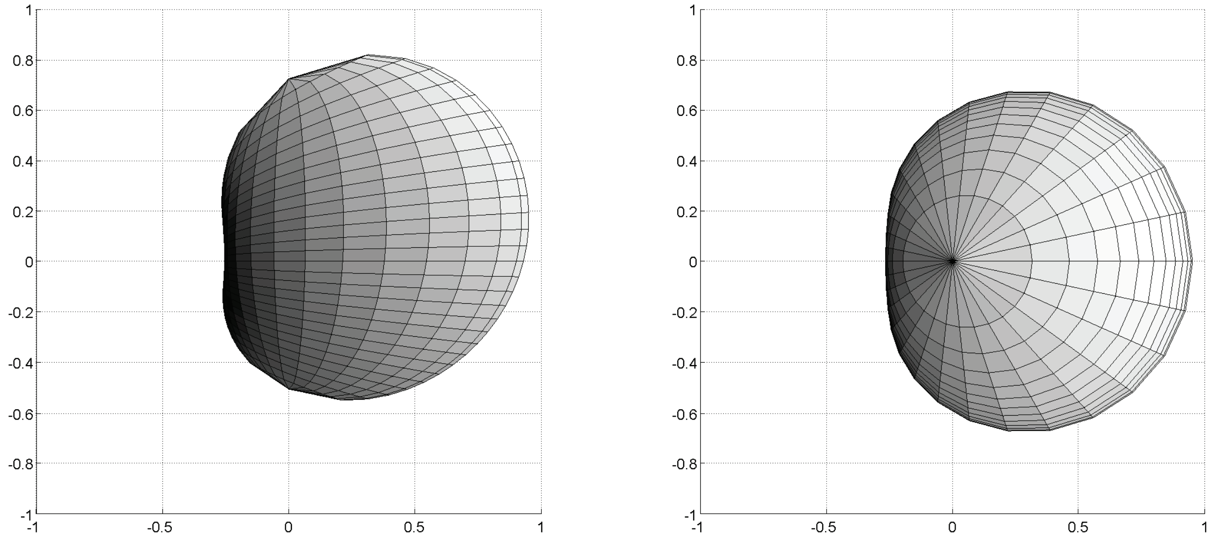
An approximation of the long term directivity of a human talker based on beamforming with second order spherical harmonics at different frequencies is presented in Figures 4 to 6.



Side View

Top View

Figure 4: Approximation of the directivity of a human speaker at 250 Hz from second order spherical harmonics beamforming. Values are normalized farfield pressure, with the beam directed along the horizontal axis.



Side View

Top View

Figure 5: Approximation of the directivity of a human speaker at 1000 Hz from second order spherical harmonics beamforming. Values are normalized farfield pressure, with the beam directed along the horizontal axis.

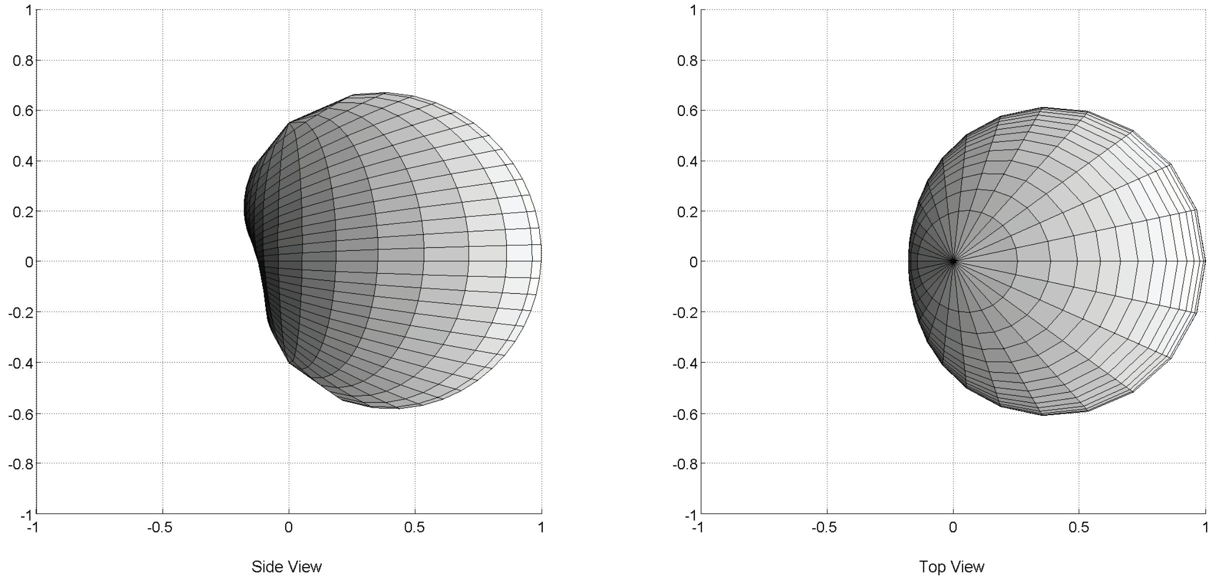


Figure 6: Approximation of the directivity of a human speaker at 4000 Hz from second order spherical harmonics beamforming. Values are normalized farfield pressure, with the beam directed along the horizontal axis.

3. MEASUREMENTS AND RESULTS

Measurements of the loudspeaker response were obtained at a distance of 1 m from the dodecahedron centre. Twenty-four measurements were made at the locations described by Cabrera *et al.* (2011) and then mirrored to seventy-two measurements around a sphere, based on the symmetrical characteristics of the dodecahedron. Impulse responses were derived from the driver to each measurement location. For the first test the impulse responses were processed to arrive at a beam with the characteristics of the beam shown on Figure 2. The beam was aimed at an angle of 0 deg azimuth and 0 deg elevation. The impulse responses were first weighted according to their corresponding dodecahedron face, with a weight value assigned based on the centre of the driver location in relation to a sphere and the spherical harmonic of n^{th} degree and order intended to be produced. Each dodecahedron face yields a matrix of 72 impulse responses by 9 spherical harmonics. Afterwards, each of the 9 spherical harmonic sets of impulse responses is scaled by a weight derived from the process described in section 2. The resulting impulse responses are added at each location, yielding a matrix of 1 impulse response per 72 measurement locations. An FFT of the impulse responses was performed and the magnitude of each octave band under test was calculated. The normalised magnitude (to a maximum value of 1) was compared to the normalised magnitude of the ideal beam using second order spherical harmonics at the same 72 measurement locations. The results are presented in Table 1.

Table 1: Correlation coefficients for loudspeaker beam directed at one point

Frequency	Correlation Coefficient
250	0.1924
500	0.1266
1000	0.4683
2000	0.5096
4000	0.2228

In order to test the ability of the loudspeaker to produce beam patterns other than the simple, maximum directivity index pattern presented above, the same process was followed, trying to optimise the loudspeaker to produce a pattern that changes over frequency and simulates the directivity of a human speaker. A matrix of impulse responses around a sphere from a human speaker was used to derive the frequency response at different frequencies around a sphere. Using the process described in section 2, a matrix of responses around a sphere was calculated for the octave bands going from 250 Hz to 4 kHz. The spherical harmonic weightings obtained were applied to the impulse responses and correlation coefficients were calculated, comparing the measurement results to a theoretical best approximation using second order spherical harmonics. The results are presented in Table 2.

Table 2: Correlation coefficients for beams emulating directivity of a human speaker

Frequency	Correlation Coefficient
250	-0.1336
500	-0.6396
1000	0.1362
2000	0.8378
4000	0.3984

4. DISCUSSION

The results obtained show several areas of improvement for the loudspeaker device in itself and its measurement. The correlation coefficients in both the directional to one point in space and human directivity emulation show large variation across the frequency range. In our accompanying paper the ability of the loudspeaker to create spherical harmonics was tested and the results also showed that the correlation coefficient had a drastic drop at the frequencies below the 1000 kHz octave band. As second order spherical harmonics are required to create the directional beam used for the test it is no surprise that the two lower octaves present lower correlation than the upper three. It should also be noted that the upper three octave bands do not posses high correlation coefficients either. The 4 kHz band in the directional beam exhibits very low correlation, almost as low as the lower two octave frequency bands.

For the human directivity emulation, the lower octave bands exhibit negative correlation while the upper two exhibit higher correlation than the directive beam case. The reason for this is could very well be that the beamforming achievable with first order spherical harmonics presents characteristics very similar to the directivity of a human speaker at this frequency range. It should also be noted that the 2 kHz frequency band had the highest correlation with the theoretically perfect first and second order spherical harmonics in our accompanying study.

There are several possible reasons for the poor performance of the loudspeaker, especially to create second order spherical harmonics. The rigid dodecahedral body could be creating an interference with the soundfield that is creating large deviations from the expected directivity of the loudspeaker. Another factor that could be influencing the results is the inherent frequency dependent directivity of the driver associated to its dimensions. Lastly the low radiation efficiency at lower frequencies could also be source of error where we would expect the omnidirectional characteristics of the loudspeaker to be more prominent (this could be ameliorated by creating a larger diameter loudspeaker).

5. CONCLUSION

An inexpensive loudspeaker and measurement system has been presented that has the ability to produce up to second order spherical harmonics directivity patterns. The prototype loudspeaker was tested to understand its ability to produce highly directional beam patterns and directional patterns that resemble that of a human speaker. The results show that several measures to improve the loudspeaker enclosure and measurement need to be taken to ensure that the loudspeaker operates accurately at an extended frequency range. However, given the favourable results in the 2 kHz octave band, the utility of a device such as this is proven. It is expected that by correcting the known issues with the loudspeaker, this device could provide analysis data of several arbitrary directivity patterns from one set of measurements.

6. REFERENCES

Aarts, RM & Janssen, AJEM (2011) ‘Comparing sound radiation from a loudspeaker with that of a flexible spherical cap on a rigid sphere’, *Journal of the Audio Engineering Society*, vol. 59, no. 4, pp. 201-212.

Abdou, A & Guy, RW (1996) ‘Spatial information of sound fields for room-acoustics evaluation and diagnosis’, *Journal of the Acoustical Society of America*, vol. 100, pp.3215-3226.

Avizienis, R, Freed, A, Kassakian, P and Wesel, D (2006) ‘A compact 120 independent element spherical loudspeaker array with programmable radiation patterns’, in *Proceedings of the 120th Audio Engineering Society Convention*, Paris, France.

Bassuet, A (2010) ‘New acoustical parameters and visualization techniques to analyse the spatial distribution of sound in music spaces’, in *Proceedings of the International Symposium on Room Acoustics*, Melbourne, Australia.

Cabrera, D, McCarthy, R, Bassett, J & Gilfillan, D (2010) ‘The measurement of directional stage support in a drama theatre: Quantifying the effect of set’, *Journal of Building Acoustics*, vol. 17, pp. 1-14.

Cabrera, D, Stewart, K & Miranda, L (2011) ‘Acoustic source beam control using a compact virtual loudspeaker array’ submitted to *Acoustics 2011 The 2011 Conference of the Australian Acoustical Society,* Surfers Paradise, Queensland, Australia.

Cabrera, D, Lee, D Collins, R, Hartmann, B, Martens, WL, & Sato, H (2011) ‘Variation in oral-binaural room impulse responses for horizontal rotations of a head and torso simulator’, *Journal of Building Acoustics*, vol. 18, pp. 801-811.

Chu, WT & Warnock, ACC (2002) *Detailed directivity of sound fields around human talkers*, National Research Council of Canad IRC-RR-104.

Daniel, J, Nicol, R & Moreau, S (2003) 'Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging', presented at the *114th AES Convention*, Amsterdam, Then Netherlands.

Noisternig, N, Zotter, F, Katz, BFG (2011) 'Reconstructing sound source directivity in virtual acoustic environments', Suzuki, Y, Brungart, D, Iwaya, Y, Iida, K, Cabrera, D & Kato, H (eds), *Principles and applications of spatial hearing*, World Scientific, New Jersey, pp. 357-372.

Pasqual, AM, de França Arruda, JR & Herzog, P (2010) 'A comparative study of platonic solid loudspeakers as directivity-controlled sound sources', *Proceedings of the second international symposium on Ambisonics and spherical acoustics*, Paris, France.

Pasqual, AM, Herzog, P, Arruda, JRF (2010) 'Theoretical and experimental analysis of the electromechanical behavior of a compact spherical loudspeaker array for directivity control', *Journal of the Acoustical Society of America*, vol. 128, pp. 3478–3488.

Pollow, M & Behler, GK (2009) 'Variable directivity for Platonic sound sources based on spherical harmonics optimization', *Acta Acustica united with Acustica*, vol. 95, pp. 1082–1092.

Rafaely, B. (2009) 'Spherical loudspeaker array for local active control of sound', *Journal of the Acoustical Society of America*, vol. 125 pp. 3006–2017.

Stewart, K & Cabrera, D (2009) 'Effect of acoustic environment on the sensitivity of speech transmission index to source directivity', *Architectural Science Review*, vol. 52, pp. 71-78.

Williams, EG (1999) *Fourier acoustics: sound radiation and nearfield acoustical holography*. Academic Press: London.

Zotter, F, Sontacchi, A & Höldrich, R (2007) 'Modeling a spherical loudspeaker system as multipole source', in *Proceedings of the 33rd German Annual Conference on Acoustics*, Stuttgart, Germany, pp. 1–2.

Zotter, F (2009) *Analysis and Synthesis of Sound-Radiation with Spherical Arrays*, Doctoral Thesis, Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz.

Zotter, F and Pasqual, AM (2011) 'Radiation modes of t-design and extremal-points compact spherical loudspeaker arrays', *Fortschritte der Akustik*, DAGA, Düsseldorf.