ICSV14
Cairns • Australia
9-12 July, 2007

14th International Congress on Sound & Vibration

LIMITATIONS AND ERROR ANALYSIS OF SPHERICAL MICROPHONE ARRAYS

Thushara D. Abhayapala[1]*, Michael C. T. Chan[2]

[1]Research School of Information Sciences and Engineering
Australian National University
Canberra ACT 0200 Australia
*Thushara.Abhayapala@anu.edu.au

Abstract

There has been a strong interest in theory and design of spherical microphone arrays in recent years due to their possible use in recording directional sounds for surround sound creation. The theory of such arrays is based on decomposition of sound fields to spherical harmonics which are the natural basis functions for valid sound fields over three dimensional space. However, due to physical size of practical spherical arrays, number of extractable spherical harmonic components are limited to the third order. This paper investigates the various design issues including the inherent limitations of the spherical microphone, discretization of the theoretical continuous spherical microphone into a microphone array and associated spatial aliasing problems, calibration errors of these microphones and signal processing issues. A fourth order microphone design was presented and analysed, which allowed the verification, integration and evaluation of the design issues mentioned earlier, in the context of this design. Overall, the design was capable of recording a frequency range of $[340; 3400]$Hz. The work presented in this paper has made the following main contributions to sound field recording: (i) Analysis of the role of rigid spheres in improving microphone design; (ii) Analysis of various microphone arrangements, which are applicable to microphone array design; (ii) A model for analysing the error due to inexact positioning of microphones in a spherical array; (iii) A set of error measures for error analysis of spherical microphones; and (iv) Design and analysis of a fourth order spherical microphone.

1. INTRODUCTION

Spherical microphone arrays have been introduced [1, 2] to use in spatial sound recording and beamforming applications. Over the last 5 years, there has been a strong interest [3, 4] in theory and design of spherical microphone arrays due to their possible use in recording directional sounds for surround sound creation. The theory of such arrays is based on decomposition of sound fields to spherical harmonics which are the natural basis functions for valid sound fields over three dimensional space. This paper investigates the various design issues including the inherent limitations of the spherical microphone, discretization of the theoretical continuous spherical microphone into a microphone array and associated spatial aliasing problems, and other errors. We also propose a fourth order design to span a frequency range $[340; 3400]$Hz.

2. SOUNDFIELD AROUND A RIGID SPHERE

In the literature, spherical microphone arrays has been proposed based on soundfield analysis around rigid and open (free space) spheres. In this section, we outline the general soundfield theory as applicable to rigid and open spheres.

Let the centre of the sphere of radius R be the arbitrary chosen origin of the co-ordinate system, then an arbitrary soundfield at a point $\boldsymbol{x} = (r, \theta, \phi; r \geq R)$ from the origin is given by

$$S(\boldsymbol{x}; k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \gamma_{nm}(k) b_n(k\|\boldsymbol{x}\|) Y_{nm}(\hat{\boldsymbol{x}}), \tag{1}$$

where $\hat{\boldsymbol{x}} = \boldsymbol{x}/|\boldsymbol{x}\|$, $n(\geq 0), m$ are integers, $k = 2\pi f/c$ is the wavenumber, f is the frequency, c is the speed of wave propagation, $Y_{nm}(\cdot)$ are the spherical harmonics, $\gamma_{nm}(k)$ are complex valued coefficients of the soundfield for wavenumber k, and

$$b_n(k\|\boldsymbol{x}\|) = \begin{cases} j_n(k\|\boldsymbol{x}\|) - \frac{j_n'(kR)}{h_n^{(1)'}(kR)} h_n^{(1)}(k\|\boldsymbol{x}\|), & \text{for rigid sphere}, \\ j_n(k\|\boldsymbol{x}\|) & \text{for open sphere}, \end{cases} \tag{2}$$

where $j_n(\cdot)$ are the spherical Bessel functions, $h_n^{(1)'}(\cdot)$ are the spherical Hankel functions of the second kind, and $(\cdot)'$ denotes the derivative of a function with respect to its argument. Using the orthogonal properties of spherical harmonics, we write the analysis equation of the soundfield as

$$\gamma_{nm}(k) = \frac{1}{b_n(k\|\boldsymbol{x}\|)} \int S(\boldsymbol{x}; k) Y_{nm}^*(\hat{\boldsymbol{x}}) d\hat{\boldsymbol{x}} \tag{3}$$

where the integration is over the unit sphere. Analogous to the Fourier series coefficients for a periodic temporal signal, the modal coefficients $\gamma_{nm}(k)$ in (1) contain all the information about an arbitrary sound field. Therefore, if we can record them, it is theoretically possible to reconstruct the acoustic environment using an array of loudspeakers [5, 6].

It is possible to capture the modal coefficients of a sound field by evaluating this integral over an arbitrary surface, for each n and $-n \leq m \leq n$ desired. This means that it is only necessary to evaluate the sound field over this surface rather than at each and every point in the sound field.

3. DESIGN ISSUES OF SPHERICAL MICROPHONES

Theory of array of spherical microphone arrays [1, 2] are based on capturing soundfield pressure $S(\boldsymbol{x}; k)$ over all points on the surface of a sphere; i.e., a continuous spherical microphone. In practice, this is achieved using only a limited number of microphones on the surface. In this section, we analyse the inherent limitations of recording from a continuous spherical microphone. Because such a microphone cannot be practically implemented, we look at issues of discretizing it by using a microphone array, which introduces problems associated with insufficient spatial sampling, and the calibrating errors associated with inexact positioning.

3.1. Finite Order Design

In practice, radius of a spherical microphone is small (i.e., up to 20 cm or so). For such a small radius, lower order soundfield components are dominant and higher orders contribute very little. This can be observed by studying the nature of the functions $b_n(\cdot)$, which are high pass for larger n (see Fig. 1). Conversely, it is difficult to capture higher order components using a small sphere as the measured soundfield pressure consists less energy from higher orders, thus enhancing noise. Existing spherical arrays can only record up to a third order of components

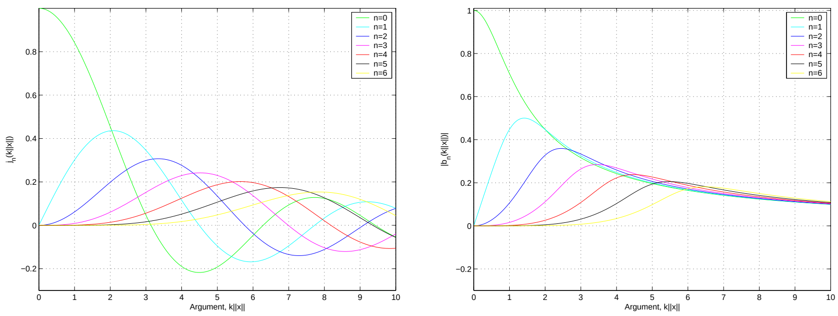

Figure 1. Plots of spherical Bessel functions, (a) $j_n(k\|\boldsymbol{x}\|)$, and (b) $b_n(k\|\boldsymbol{x}\|)$ for various orders n.

for entire band of audio frequencies.

3.2. Nature of Spherical Microphone

3.2.1. Open Sphere Configuration

Consider the situation where a continuous spherical microphone of radius $\|\boldsymbol{x}\| = r$ use to record the sound field. The work [1] assumes that the continuous spherical microphone is transparent with respect to the sound field, such that it does not disturb the sound field in any way. This is called the open sphere configuration. Note that in the case of open sphere, $b_n(k\|\boldsymbol{x}\|) = j_n(k\|\boldsymbol{x}\|)$ in the expression of (1). This is plotted in Fig. 1. Notice that for small values of frequency k, the values of $j_n(kr)$ for $n \geq 2$ are small. In addition, $j_n(kr)$ has zeros for all

$n \geq 0$. At the corresponding frequencies k, the modal components of the corresponding orders n, will be "perceived" by the continuous spherical microphone to be small or equal to zero, respectively. Therefore, they will be difficult to record at those frequencies. Thus, the presence of these zeros in (1) means that we are greatly restricted in the range of frequencies for which an accurate recording can be made.

3.2.2. Rigid Sphere Configuration

The restriction by the zeros of the spherical Bessel functions can be alleviated by using a rigid spherical configuration, which means that the surface of the continuous spherical microphone of radius $\|\boldsymbol{x}\| = r$ coincides with a rigid spherical scatterer (ie. a rigid sphere). In this case, the rigid spherical scatterer interacts and alters the sound field.

Notice that there are no zeros in the functions, $b_n(k\|\boldsymbol{x}\|)$ for rigid spheres (see Fig. 1). The other noticeable advantage of these functions is that at lower values of k and for $n > 0$, $b_n(kr)$ is approximately 3dB greater than $j_n(kr)$. This is due to diffraction over the rigid sphere [2]. At higher values of frequency k, the scattering effects will become more prominent compared to diffractional effects.

Notice that in the presence of a rigid spherical scatterer, whether the continuous spherical microphone coincides with it or not, the first zero of the radial response $b_0(k|\boldsymbol{x}\|)$, is shifted to a larger value of $k\|\boldsymbol{x}\|$ than in the case of an open spherical configuration. In terms of design, this means that a larger range of frequencies can be accurately recorded when either a rigid or an intermediate spherical configuration is used, compared to an open spherical configuration.

3.3. Modal Aliasing

Modal aliasing occurs when the higher order modal components of the sound field are recognised as lower order modal components. This is analogous to temporal aliasing of signals, where higher frequency components are recognised as lower frequency components. Modal aliasing occurs due to insufficient spatial sampling when the theoretical continuous spherical microphone is approximated with a discrete microphone array [1, 2]. A detailed analysis is given in [7].

Aliasing are introduced, when the integral in (1) is approximated by Q omni-directional microphones. Assume that Q is sufficiently large and that the microphones are arranged in a way, such that this approximation is *exact* for resolving the modal coefficients of a sound field for orders $0 \leq n < N$. We denote this order truncated sound field that contains only these modal components as $S_{0:(N-1)}(\boldsymbol{x}; k)$. Then, the modal coefficients for orders $0 \leq n < N$ are

$$\gamma_{nm}(k) = \frac{1}{b_n(kR)} \sum_{q=1}^{Q} S_{0:(N-1)}(R\hat{\boldsymbol{x}}_q; k) Y_{nm}^*(\hat{\boldsymbol{x}}_q) w_q = \frac{1}{b_n(kR)} \int S_{0:(N-1)}(R\hat{\boldsymbol{x}}; k) Y_{nm}^*(\hat{\boldsymbol{x}}) d\hat{\boldsymbol{x}}. \tag{4}$$

However, when the microphone array is exposed to the total sound field $S(\boldsymbol{x}; k)$, the modal coefficients we are concerned with are corrupted by the sound field due to higher order modal components, which we denote as $S_{N:\infty}(\boldsymbol{x};k)$. Now, the recorded modal coefficients will be

$$\hat{\gamma}_{nm}(k) = \frac{1}{j_n(kR)} \sum_{q=1}^{Q} \big(S_{0:(N-1)}(R\hat{\boldsymbol{x}}_q;k) + S_{N:\infty}(R\hat{\boldsymbol{x}}_q;k)\big) Y_{nm}^{*}(\hat{\boldsymbol{x}}_q) w_q \quad (5)$$

$$= \gamma_{nm}(k) + \epsilon_{nm}(k), \quad (6)$$

where, $\epsilon_{nm}(k)$ represents the error due to modal aliasing [1]. This is given by

$$\epsilon_{nm}(k) = \frac{1}{j_n(kR)} \sum_{n'=N}^{\infty} \sum_{m'=-n'}^{n'} \gamma_{n'm'}(k) j_{n'}(kR) \times \sum_{q=1}^{Q} Y_{n'm'}(\hat{\boldsymbol{x}}_q) Y_{nm}^{*}(\hat{\boldsymbol{x}}_q) w_q. \quad (7)$$

Thus, we can see that modal aliasing is introduced when we use an approximation to the continuous spherical microphone. In the case of the continuous spherical microphone this error goes to zero. An in-depth analysis of modal aliasing is given in [7]. The constraint (called *orthonormality constraint* that microphone positions has to satisfy to eliminate modal aliasing is given by [2]

$$\sum_{q=1}^{Q} Y_{n'm'}(\hat{\boldsymbol{x}}_q) Y_{nm}^{*}(\hat{\boldsymbol{x}}_q) w_q = \delta_{nn'}\delta_{mm'}, \quad (8)$$

for $0 \leq n < N$. This constraint ensures that the Q microphones are able to resolve all $\gamma_{nm}(k)$ exactly for this range of n, hence the use of the term, orthonormality.

3.4. Microphone Arrangement

There are numerous ways to arrange a finite number of microphones in a spherical array to provide approximations to the continuous spherical microphone. The extent to which these arrangements satisfy (8) gives an indication as to the level of accuracy in which the modal coefficients for these orders can be resolved.

Possible array configurations are (i) Gaussian and trigonometric quadrature arrangements, (ii) truncated icosahedron arrangements, and (iii) cubature arrangements. Suitablity of each of these arrangements can be studied [7] in terms of scalability, physical realisability and efficiency. Reader is referred to [7] for an in-depth analysis.

4. FOURTH ORDER DESIGN

4.1. Design Specification

We would like to extend beyond the capabilities of existing designs [1, 2, 3, 4] by choosing to record the modal components of the sound field for orders $0 \leq n < 5$, i.e., a fourth order microphone. The desired frequency range is $f = [340, 3400]$Hz, or $[1:10]$ frequency band.

4.2. Error Measures

There are a few sources of error in which we would be interested in assessing a design. These are (i) Truncation error, (ii) Error of recording, and (iii) Aliasing error. All these errors characterise the spatial quality of the sound field as opposed to temporal quality and therefore, do not affect the intelligibility of the sound. Notice that all these error measures vary with the frequency component k of the sound field and the distance from the origin, $\|\boldsymbol{x}\|$.

Truncation Error: is an inherent limitation in any practical design. The fact that a design is specified to record modal coefficients of orders $0 \leq n < N$ means that the loss of information regarding the orders $n \geq N$ will introduce errors. The truncation error is given by

$$\epsilon_{\text{trunc}}(\|\boldsymbol{x}\|; k) = \frac{\int \left| S(\boldsymbol{x}; k) - S_{0:(N-1)}(\boldsymbol{x}; k) \right|^2 d\mathbf{x}}{\int \left| S(\boldsymbol{x}; k) \right|^2 d\mathbf{x}} = 1 - \frac{\sum_{n=0}^{N-1} \sum_{m=-n}^{n} \left| \gamma_{nm}(k) j_n(k|\boldsymbol{x}\|) \right|^2}{\sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left| \gamma_{nm}(k) j_n(k|\boldsymbol{x}\|) \right|^2}, \tag{9}$$

where $S(\boldsymbol{x}; k)$ is the original sound field and $S_{0:(N-1)}(\boldsymbol{x}; k)$ is the original sound field order truncated to include modal components of orders $0 \leq n < N$.

Error of Recording: includes the inherent limitations of the use of a single spherical microphone array and the error due to the integration approximation of the continuous spherical microphone. This is given by

$$\epsilon_{\text{rec}}(\|\boldsymbol{x}\|; k) = \frac{\int \left| S_{0:(N-1)}(\boldsymbol{x}; k) - \tilde{S}_{0:(N-1)}(\boldsymbol{x}; k) \right|^2 d\mathbf{x}}{\int \left| S_{0:(N-1)}(\boldsymbol{x}; k) \right|^2 d\hat{\boldsymbol{x}}} = \frac{\sum_{n=0}^{N-1} \sum_{m=-n}^{n} \left| \big(\gamma_{nm}(k) - \tilde{\gamma}_{nm}(k)\big) j_n(k\|\boldsymbol{x}\|) \right|^2}{\sum_{n=0}^{N-1} \sum_{m=-n}^{n} \left| \gamma_{nm}(k) j_n(k\|\boldsymbol{x}\|) \right|^2}, \tag{10}$$

where $\tilde{S}_{0:(N-1)}(\boldsymbol{x}; k)$ is the perfectly reconstructed sound field from the modal coefficients recorded from $S_{0:(N-1)}(\boldsymbol{x}; k)$.

Aliasing Error: is a measure of the extent of modal aliasing, and is given by

$$\epsilon_{\text{alias}}(\|\boldsymbol{x}\|; k) = \frac{\int \left| \tilde{S}_{0:(N-1)}(\boldsymbol{x}; k) - \hat{S}_{0:(N-1)}(\boldsymbol{x}; k) \right|^2 d\mathbf{x}}{\int \left| S_{0:(N-1)}(\boldsymbol{x}; k) \right|^2 d\hat{\boldsymbol{x}}} = \frac{\sum_{n=0}^{N-1} \sum_{m=-n}^{n} \left| \big(\tilde{\gamma}_{nm}(k) - \hat{\gamma}_{nm}(k)\big) j_n(k\|\boldsymbol{x}\|) \right|^2}{\sum_{n=0}^{N-1} \sum_{m=-n}^{n} \left| \gamma_{nm}(k) j_n(k\|\boldsymbol{x}\|) \right|^2}, \tag{11}$$

where $\hat{S}_{0:(N-1)}(\boldsymbol{x};k)$ is the perfectly reconstructed sound field from the modal coefficients recorded from the original sound field, $S(\boldsymbol{x};k)$.

4.3. Detailed Design

We use a concentric pair of rigid spherical configuration and a open sphere configuration. This will enable us to cover the entire frequency band of interest.

We use the truncated icosahedron arrangement since this arrangement satisfies the orthonormality constraint of (8) for $N = 5$ and possesses the ability to prevent aliasing from 5th order modal components [7]. Furthermore, it is the most efficient of the options available, requiring only 32 microphones.

To specify the level of accuracy desired, we will set two constraints. For a given microphone radius r, these constraints limit the range of frequencies k which are recorded. We use subscripts u and l to denote the upper bound and the lower bound of a quantity respectively. The constraints are

1. $(kr)_u$ is assigned the value given in Table 3.1 of [7] for $N = 6$ to minimise the modal aliasing.
2. $(kr)_l$ is assigned the minimum value of kr such that $\epsilon_{\text{rec}}(\|\boldsymbol{x}\|;k) \leq \epsilon_{\text{rec,max}}$ holds, to limit the error of recording. For argument sake, let $\epsilon_{\text{rec,max}}$ =5%.

These constraints imply a level of accuracy of recording the modal coefficients for the frequency range, [k_l,k_u] or equivalently, [f_l,f_u].

The analysis of the first constraint directly affects the design of the radius of the microphone. To minimise modal aliasing from orders $n \geq 6$, we chose $(kr)_u = 3.87$ (which corresponds to $N = 6$). Therefore, $r = 3.87c/(2\pi f_u)$. Substituting, $f_u = 4300$Hz gives us $r = 6.2$cm. The application of this constraint means that an upper bound has been placed on the aliasing error within the frequencies in $[0, f_u]$.

The second constraint will specify a value of f_l. The error of recording was calculated for this design as a function of f, and is plotted in Fig. 2(a) at various radii from the origin.

The peaking behaviour of the curves can be explained by the fact that as we move towards lower frequencies, the magnitude of $b_n(k\|\boldsymbol{x}\|)$ for $n \geq 2$ becomes small. By considering this figure and the fact that only a spherical region of radius 0.5m is considered, this allows us to assign a frequency f_l in which to ensure that the second constraint is satisfied. By observation, we chose $f_l = 1100$Hz. Therefore, the microphone design satisfies the specifications and constraints outlined above for the frequencies in the range $f = [f_l, f_u] = [1100, 3400]$Hz.

Thus far, the design does not satisfy the intended frequency range. The approach to solving this problem is to employ a second microphone array at a larger radius in order to capture the lower frequencies in $f = [340, 1100]$Hz. Thus, we use an open spherical microphone, which encases the inner microphone array. We use the truncated icosahedron arrangement for the open configuration as well. Again, we use the above two constraints with appropriate values to design the outer (open) microphone array with radius of 15.5cm covering a frequency range of $[0, 1100]$Hz.

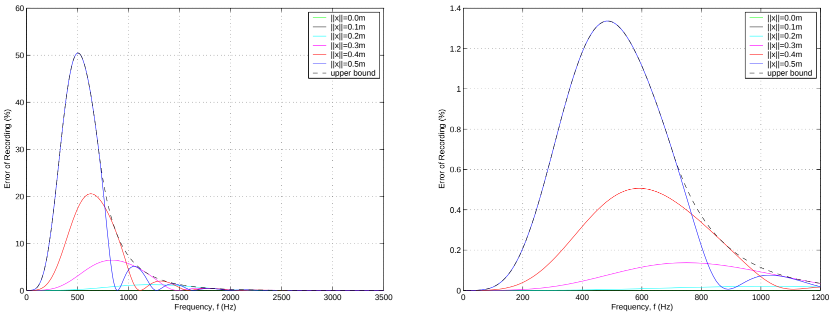

Figure 2. (a) Error of recording at various radii, for microphone array at $r = 6.2$cm with coincident rigid spherical scatterer. The dotted black line indicates the upper bound to error for $0 \leq \|\boldsymbol{x}\| \leq 0.5$m; (b) Error of recording at various radii, for microphone array at $r' = 15.5$cm with a rigid spherical scatterer at $R = 6.2$cm. The dotted black line indicates the upper bound to error for $0 \leq \|\boldsymbol{x}\| \leq 0.5$m.

In the double array microphone a total of 64 microphones are placed in two spherical locations. The frequency response of the system satisfies the specified range of $[340, 3400]$Hz. Reader is referred to [7] for more detail of the design and simulation results.

REFERENCES

[1] T.D. Abhayapala and D.B. Ward, "Theory and design of higher order sound field microphones using spherical microphone array," in *Acoustics, Speech and Signal Processing, IEEE International Conference*, Orlando, FL, USA, May 2002, vol. 2, pp. 1949–1952.

[2] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Acoustics, Speech and Signal Processing, IEEE International Conference*, Orlando, FL, USA, May 2002, vol. 2, pp. 1781–1784.

[3] B. Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Acoust. Speech Sig. Proc.*, vol. 13, pp. 135–143, Jan. 2005.

[4] L. Zhiyun, R. Duraiswami, and L.S. Davis, "Flexible layout and optimal cancellation of the orthonormality error for spherical microphone arrays," in *Acoustics, Speech and Signal Processing, IEEE International Conference*, Orlando, FL, USA, May 2004, vol. 4, pp. iv–41–44.

[5] T.D. Abhayapala and D.B. Ward, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 6, pp. 697–707, Sept. 2001.

[6] D. Excell, "Reproduction of a 3d sound field using an array of loudspeakers," Honours Thesis (http://www-sigproc.eng.cam.ac.uk/ dae30/pubs/thesis-ug-dexcell-2003.pdf), June 2003.

[7] M.C.T. Chan, "Theory and design of higher order sound field recording," Honours Thesis (http://users.rsise.anu.edu.au/ thush/ugstudents/MCTChanThesis.pdf), June 2003.