DESIGN AND EVALUATION OF OPEN SPHERICAL MICROPHONE ARRAYS

Chaoran Du, Quentin Leclere, Baojiang Li

24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

ICSV24
23-27JULY 2017
LONDON CALLING

DESIGN AND EVALUATION OF OPEN SPHERICAL MICROPHONE ARRAYS

Chaoran Du
OROS, F-38944, Meylan, France
email: chaoran.du@oros.com

Quentin Leclere
University of Lyon, INSA-Lyon, LVA EA677, F-69621, Villeurbanne, France

Baojiang Li
TStech, Beijing, China

Beamforming with a spherical microphone array has become an attractive tool for sound source localization in cabin environments. To provide a reliable localization solution with high performance, the first step is to optimize the array design. In this paper, the spatial resolution and dynamic range of an open spherical array with various microphone arrangements are evaluated, which are used as metrics for optimal array design. Several popular approaches for evenly distributing points on a sphere are investigated such that the array response is as isotropic as possible. In addition to the conventional beamforming, a state-of-the-art algorithm called functional beamforming is also implemented. The objective is twofold: first is to improve spatial resolution and dynamic range significantly; and second is to verify that array configurations have a consistent effect on localization performance, regardless of beamforming algorithms.

Keywords: spherical microphone array, array design, beamforming, localization performance evaluation

1. Introduction

Spherical microphone arrays have been widely used in the automotive industry for sound source localization in cabin environments. Generally speaking, there are two main ways of building a spherical array: placing microphones on the surface of a solid sphere or an open sphere. A solid sphere is easier in manufacturing and provides a housing for all electronic components and wiring. However, the solid sphere disturbs the sound field around its body due to the scattered sound waves. In order to take into account the scattering effects in the propagation model, the calculation of acoustic maps has to be done in the frequency domain with a high computational cost. Another disadvantage of the solid sphere configuration is the limited array size. Under certain circumstances, *e.g.*, at low frequencies, large arrays are preferable. A large solid sphere might be difficult to handle in practice due to its weight. Furthermore, scattering of the incident waves from a large solid sphere might be reflected back into the measurement region by surrounding objects, modifying the measured sound field [1]. Therefore, the open sphere configuration is selected in our study due to its acoustic transparency, flexibility in beamforming algorithm choices, light weight and the possibility of building large arrays for better performance.

Different schemes for placing microphones on a sphere have been presented and analyzed [1]. In order to have an isotropic array response, strictly speaking, each microphone has to have the same number of neighbors at the same distance, equally spaced around itself. However, there are only a very limited number of configurations based on the five Platonic solids offer such an uniform distribution. The problem of distributing many points on a sphere in a nearly-uniform manner has drawn attention in many fields, and a wide range of approaches have been proposed in the literature [2, 3, 4, 5]. Most analysis was done in terms of how to choose the sampling points and their corresponding weights such that the integral of an order-limited continuous function on the sphere can be well approximated by the weighted summation of the samples of the function. However, the influence of different microphone arrangements on the localization performance of an open array is unclear. Therefore, in this paper, the spatial resolution and dynamic range of various spherical arrays are evaluated, which are used as metrics to find the array design offering the optimal localization performance.

The data acquired by a spherical array are processed with a beamforming algorithm to generate acoustic maps, based on which the locations and strengths of sound sources are estimated. Two beamforming algorithms are implemented in this paper. One is the conventional frequency domain beamforming (FDBF), whihc is one of the simplest and most robust beamforming algorithms and widely used in many fields. The other is a state-of-the-art algorithm called functional beamforming, which was recently proposed by Dougherty [6]. It has been shown that functional beamforming provides enhanced dynamic range and spatial resolution, which is comparable with that of other advanced beamforming algorithms, *e.g.*, the deconvolution methods. More importantly, the required computational time of functional beamforming is approximately the same as that of the conventional beamforming, which is substantially shorter than the deconvolution methods. This is of great significance to the fast and accurate identification of sound sources in 3D cabin environments.

2. Array design

In this section, we introduce the microphone arrangements in a spherical array. Five of the most popular schemes for evenly distributing points on a sphere are investigated in this paper, and their design criteria are briefly described as below.

- Spherical t-design
  This approach was extent from the t-design based on the five Platonic solids to a larger set of configurations with more sampling points, offering an asymptotically uniform distribution. Precisely speaking, Hardin and Sloane [2] found the coordinates of N points via numerical optimization, such that the integral of any polynomial of degree up to t over the sphere is equal to the average value of the polynomial over the N points.

- Packing
  The packing problem, which is also known as the Tammes's problem, is to find the centres of N identical non-overlapping spherical caps such that their common radius is maximized. A spherical cap with radius r is defined as the set of all points on the sphere whose distance from the centre is no more than r. The packing scheme distributes the N points on a sphere by maximizing the smallest distance among them.

- Covering
  The covering problem is to find the centres of N identical spherical caps which cover the sphere completely such that their common radius is minimized. This scheme arranges the N points on a sphere by minimizing the greatest distance of any point from its nearest neighbour.

- Minimal energy
  This approach, which is also known as the electrostatic repulsion problem, finds the positions of *N* points on a sphere by minimizing the sum of the inverse of the distances between them. If the points are considered as charged particles repelling each other, minimizing such a sum physically represents minimizing the potential energy of the particles.

- Spiral points
  Inspired by the sunflower seed pattern occurring naturally, a simple approach was proposed which arranges points on spirals wrapping around a sphere [5, 7]. This method cuts the sphere with *N* evenly apart horizontal planes, and assigns one point to each latitude. The difference in longitudes of successive points is selected to be the *golden angle* so that no two points in nearby bands come too close to each other in longitude. The cylindrical coordinates of the *k*-th spherical spiral point can be written as

$$\begin{aligned} \psi_k &= k \cdot (3 - \sqrt{5})\pi \\ \rho_k &= \sqrt{1 - z_k^2} \\ z_k &= (1 - \frac{1}{N}) - \frac{2k}{N} \end{aligned} \tag{1}$$

  where $k = 0, 1, ..., N-1$. Due to practical concerns, the extreme points are offset in order to avoid using North and South poles as start and end points.

Note here that although the objective functions of the packing, covering, and minimal energy problems are simply defined, solving them is nontrivial and there is no known general solution. The coordinates of the sampling points are thus calculated via numerical optimization, and various solutions have been reported in a number of articles. In this paper, we utilize the putatively optimal solutions provided by Sloane, Hardin, and Smith, which are available online [8].

3. Spherical array beamforming

In order to estimate the locations and strength of sound sources of interest, beamforming algorithms are widely used to process the microphone array data, generating acoustic maps. In this section, we will briefly revise the conventional FDBF and a state-of-the-art functional beamforming.

3.1 FDBF

Assume that the spherical array has N microphones and there are Q mutually incoherent point sources with strength s_q and position vector $\mathbf{r}_q^{\mathbf{s}} \in \mathbb{R}^3$, respectively. Taking the array centre as the system origin, the source position can be described by the spherical coordinates $\mathbf{r}_q^{\mathbf{s}} = (r_q, \theta_q, \psi_q)$, where r, θ, ψ denote the radius, inclination angle, and azimuth angle, respectively. The $N \times N$ cross-spectral matrix (CSM) of the measured pressures can be expressed as

$$\mathbf{C} = \mathrm{E}\left[\mathbf{p} \cdot \mathbf{p}^H\right] = \sum_{q=1}^{Q} s_q \cdot \mathbf{g}_q \mathbf{g}_q^H \tag{2}$$

where the superscript H denotes Hermitian transpose, and the vector $\mathbf{p} = \mathbf{p}(f) \in \mathbb{C}^{N \times 1}$ contains the Fourier transform of the recorded sound pressure at each microphone at a frequency f. The vector $\mathbf{g}_q \in \mathbb{C}^{N \times 1}$ denotes the steering vector corresponding to the source position $\mathbf{r}_q^{\mathbf{s}}$, whose n-th element

is given by

$$g_{q,n} = \frac{\exp(-ik \parallel \mathbf{r}_n - \mathbf{r}_q^{\mathbf{s}} \parallel)}{\parallel \mathbf{r}_n - \mathbf{r}_q^{\mathbf{s}} \parallel} \tag{3}$$

Here, the position vector $\mathbf{r}_n \in \mathbb{R}^3$ represents the location of the n-th microphone, $k = 2\pi f/c$ is the wavenumber for frequency f and sound speed c, and $\parallel \cdot \parallel$ stands for Euclidean norm. For array beamforming, a set of L focus points covering the scene of interest are defined, containing all the potential sound source positions. The FDBF output at the l-th focus point located at position $\mathbf{r}_l^{\mathbf{f}} \in \mathbb{R}^3$ can be written as

$$b(\mathbf{r}_l^{\mathbf{f}}) = \frac{\mathbf{g}_l^H \cdot \mathbf{C} \cdot \mathbf{g}_l}{\parallel \mathbf{g}_l \parallel^2 \cdot \parallel \mathbf{g}_l \parallel^2} \tag{4}$$

where $\mathbf{g}_l \in \mathbb{C}^{N\times 1}$ is the steering vector corresponding to the point position $\mathbf{r}_l^{\mathbf{f}}$, whose definition is similar to Eq. (3). In addition, it has been demonstrated that CSM is mainly contaminated by microphone self-noise on the diagonal, and removing the diagonal of the CSM in the conventional beamforming tends to substantially improve the appearance of the acoustic maps [9]. The output of FDBF with diagonal removal is given by

$$b(\mathbf{r}_l^{\mathbf{f}}) = \frac{\mathbf{g}_l^H \cdot \mathbf{C}_{\text{diag=0}} \cdot \mathbf{g}_l}{\mathbf{v}_l^T \cdot \mathbf{1}_{\text{diag=0}} \cdot \mathbf{v}_l} \tag{5}$$

where the superscript T denotes the transpose and $\mathbf{v}_l \equiv \left[\mid g_{l,1} \mid^2 \quad \mid g_{l,2} \mid^2 \quad \ldots \quad \mid g_{l,N} \mid^2\right]^T$. The matrices $\mathbf{C}_{\text{diag=0}}$ and $\mathbf{1}_{\text{diag=0}}$ denote the modified CSM and unity matrix whose diagonal elements are replaced by zeros, respectively.

3.2 Functional beamforming

In spite of its simplicity and robustness, the conventional beamforming produces acoustic maps with poor dynamic range and limited spatial resolution. Functional beamforming, developed by Dougherty recently [6], provides significantly higher dynamic range and better spatial resolution than FDBF. The functional beamforming is based on FDBF, and its beamformer output of order ν at the focus point location $\mathbf{r}_l^{\mathbf{f}}$ can be defined as

$$b_\nu(\mathbf{r}_l^{\mathbf{f}}) = \frac{1}{\parallel \mathbf{g}_l \parallel^2}\left[\frac{\mathbf{g}_l^H \mathbf{C}^{\frac{1}{\nu}} \mathbf{g}_l}{\parallel \mathbf{g}_l \parallel^2}\right]^\nu = \frac{1}{\parallel \mathbf{g}_l \parallel^2}\left[\frac{\mathbf{g}_l^H \mathbf{U}\Sigma^{\frac{1}{\nu}}\mathbf{U}^H \mathbf{g}_l}{\parallel \mathbf{g}_l \parallel^2}\right]^\nu, \quad \nu \geq 1 \tag{6}$$

where $\mathbf{U} \in \mathbb{C}^{N\times N}$ is an unitary matrix whose columns $[\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_N]$ are the eigenvectors of $\mathbf{C}$, and Σ is a diagonal matrix whose diagonal elements $[\sigma_1, \sigma_2, ..., \sigma_N]$ are the eigenvalues of $\mathbf{C}$. The performance of functional beamforming is determined by the exponent parameter ν, which needs to be set by users. Note here that functional beamforming becomes FDBF when $\nu = 1$.

To explain the sidelobe attenuation brought in by functional beamforming, we consider the scenario where a single point source of strength s_k exists at the location $\mathbf{r}_k^{\mathbf{s}}$, then

$$\mathbf{C} = s_k \mathbf{g}_k \mathbf{g}_k^H = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H \tag{7}$$

Since $\mathbf{C}$ has rank 1, the only one nonzero eigenvalue is $\sigma_1 = s_k \parallel \mathbf{g}_k \parallel^2$, whose corresponding eigenvector $\mathbf{u}_1 = \frac{\mathbf{g}_k}{\|\mathbf{g}_k\|}$. Recalling Eq.(6), the functional beamformer output for a single point source can be expressed as

$$b_\nu(\mathbf{r}_l^{\mathbf{f}}) = \frac{1}{\parallel \mathbf{g}_l \parallel^2}\left[\frac{\mathbf{g}_l^H \mathbf{u}_1 \sigma_1^{\frac{1}{\nu}} \mathbf{u}_1^H \mathbf{g}_l}{\parallel \mathbf{g}_l \parallel^2}\right]^\nu = s_k \frac{\parallel \mathbf{g}_k \parallel^2}{\parallel \mathbf{g}_l \parallel^2}\left[\frac{\mathbf{g}_l^H \mathbf{g}_k \mathbf{g}_k^H \mathbf{g}_l}{\parallel \mathbf{g}_k \parallel^2 \parallel \mathbf{g}_l \parallel^2}\right]^\nu \tag{8}$$

According to Cauchy-Schwarz inequality, the factor which is powered to the exponent ν satisfies the following relationship

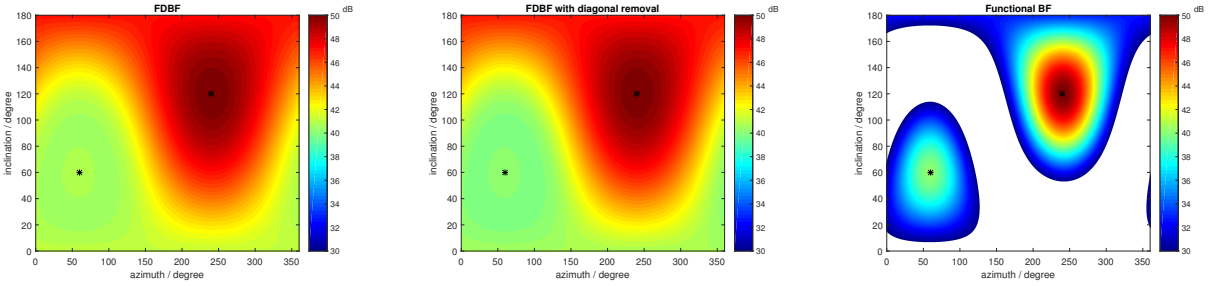
$$0 \leq \frac{\mathbf{g}_l^H \mathbf{g}_k \mathbf{g}_k^H \mathbf{g}_l}{\| \mathbf{g}_k \|^2 \| \mathbf{g}_l \|^2} \leq 1 \tag{9}$$

and reaches its maximum value 1 when $\mathbf{g}_l = \mathbf{g}_k$. Therefore, the functional beamformer output at the source location equals to the source strength irrespective of the value of ν, while the output at the sidelobes reduces significantly as ν increases, improving the dynamic range dramatically. At the focus points near the true source, the value of the factor in Eq. (9) is close to 1, and thus the effect of the exponent is not as remarkable. However, the mainlobe is still sharpened to some extent, improving the spatial resolution. Furthermore, functional beamforming attracts attention due to its low computational complexity, which is almost identical to the conventional beamforming. The only additional operation is the eigendecomposition of CSM, which is faster than the other steps in beamforming [6].

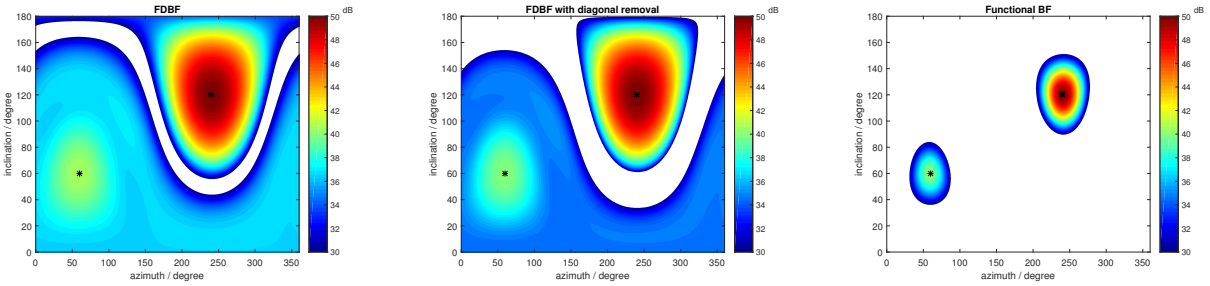
4. Simulation results

In this section, the localization performance of an open spherical array with various configurations is presented. Simulations were conducted using a 64-element spherical array with a radius of 15cm, and both the conventional and functional beamforming were applied to synthetic data. The mutually incoherent noise components were included and the resulting signal to noise ratio (SNR) was set as 20dB. The scene of interest was set as a sphere concentric to the array, and its radius equals to the distance between the sound source and array center. The azimuth and inclination angles of the focus points were 2° apart, respectively, and thus each acoustic map contained 180×91 points. Figure 1 shows examples of acoustic maps with 20dB display range generated by FDBF, FDBF with diagonal removal, and functional beamforming, respectively. The 64 microphones were distributed on the sphere following the spiral points scheme. It was assumed that two sound sources existed: one 50dB source at position $\mathbf{r}_1^{\mathbf{s}} = (1, 120^\circ, 240^\circ)$ and one 40dB source at position $\mathbf{r}_2^{\mathbf{s}} = (1, 60^\circ, 60^\circ)$, and the true source locations were marked with $*$ in the acoustic maps. The sources were assumed at 500 Hz, 1000 Hz, 3000 Hz and 7000 Hz, respectively. For functional beamforming, typical values of the exponent ν are in the range of 20-300 [6], and $\nu = 100$ was selected in the simulations as larger values do not improve the results significantly. Observing Fig. 1, it is obvious that, at low frequency, functional beamforming provides better spatial resolution than the other two methods. At mid-high frequency, compared with FDBF, FDBF with diagonal removal considerably enhances the appearance of the results. However, neither is able to locate the weaker source correctly due to the strong sidelobe contamination. The higher the frequency, the worse the contamination due to the spatial aliasing errors. In contrast, functional beamforming eliminates most artifacts, accurately localizing and quantifying both sources. The higher the frequency, the better the spatial resolution.

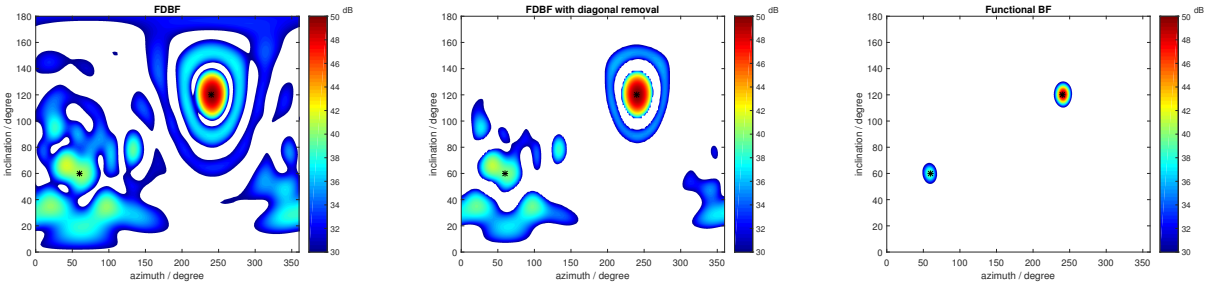
In order to optimize array design, the dynamic range and spatial resolution of various spherical arrays were evaluated as follows. For each realization, a single source was located at a random position, and an acoustic map is generated by using FDBF with diagonal removal or functional beamforming. For each acoustic map, the dynamic range was computed by finding the power difference in decibels between the peak of the mainlobe and the maximum of the strongest sidelobe. The spatial resolution was defined as the angular width of the mainlobe at 3 dB below its peak. The results were obtained by averaging over 200 realizations. Figure 2 shows the localization performance of the array, whose 64 microphones are distributed on the sphere surface according to the five schemes mentioned in Section 2. For the lowest frequencies, the dynamic range is undefined since there is no sidelobe. Sidelobes appear and become stronger as the frequency increases. Observing the results obtained by applying FDBF with diagonal removal, we notice that different microphone distributions lead to very similar spatial resolution. The spiral points distribution provides the best dynamic range, followed by *t*-desgin. Interestingly, such a conclusion is consistent with the one drawn in [10] for planar



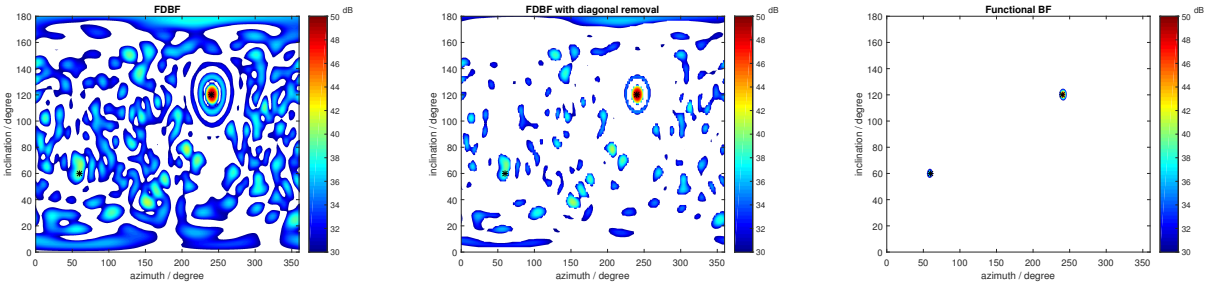
(a) 500 Hz



(b) 1000 Hz



(c) 3000 Hz



(d) 7000 Hz

Figure 1: Acoustic maps of two sources using different beamforming algorithms at (a) 500Hz, (b) 1000 Hz, (c) 3000 Hz, and (d) 7000 Hz.

microphone array design: The spiral arrangements are preferable to the other commonly used configurations. Among different kinds of spiral arrays, the one choosing golden angle as the angle interval between successive points performs the best. The same conclusion can be drawn from the figures corresponding to the functional beamforming, illustrating that array configurations have a consistent effect on localization performance, regardless of beamforming algorithms. Furthermore, we notice that, compared with conventional beamforming, functional beamforming brought in 20dB and 24dB improvement in dynamic range at low and high frequencies, respectively. It also notably enhances the spatial resolution, especially at low frequencies.

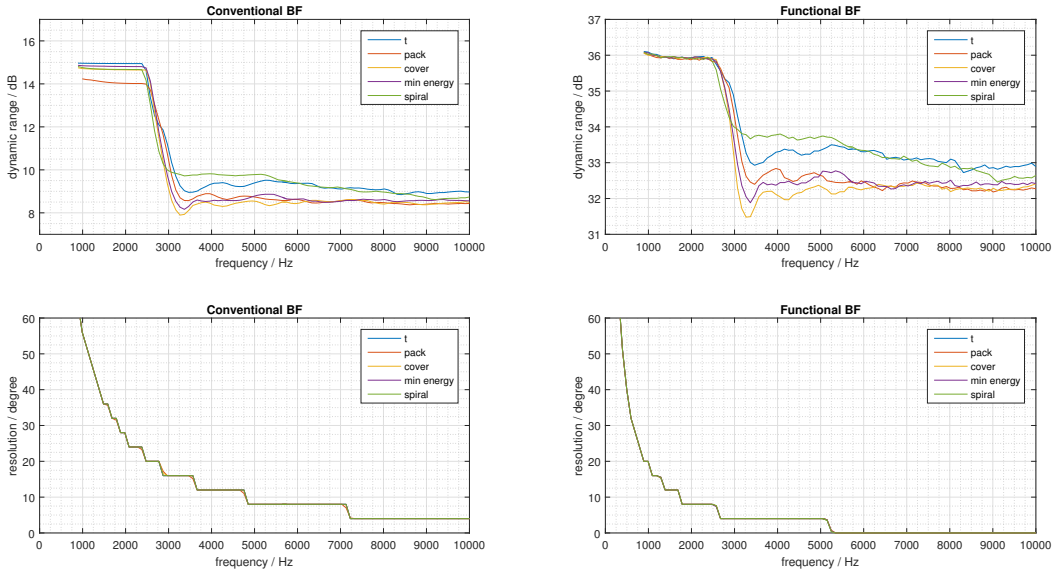


Figure 2: Localization performance of spherical array with different designs.

5. Conclusion

In this paper, the dynamic range and spatial resolution of an open spherical array with various microphone arrangements were evaluated with both conventional beamforming and functional beamforming. Five popular approaches evenly distributing points on a sphere were investigated. Simulation results showed that the scheme which arranges points on spirals wrapping around the sphere provides the best sound source localization performance. Functional beamforming offers significantly enhanced dynamic range and spatial resolution, enabling quick and accurate sound source identification in 3D environments.

REFERENCES

1. Rafaely, B., *Fundamentals of spherical array processing*, vol. 8, Springer (2015).

2. Hardin, R. H. and Sloane, N. J. A., McLaren's improved snub cube and other new spherical designs in three dimensions, *Discrete and Computational Geometry*, **15** (4), 429–441, (1996).

3. Womersley, R. S. and Sloan, I. H., How Good can polynomial interpolation on the sphere be?, *Advances in Computational Mathematics*, **14** (3), 195-226, (2001).

4. Fliege, J. and Maier, U., The distribution of points on the sphere and corresponding cubature formulae, *IMA Journal of Numerical Analysis*, **19** (2), 317-334, (1999).

5. Saff, E. B. and Kuijlaars, A. B. J., Distributing many points on a sphere, *The Mathematical Intelligencer*, **19** (1), 5–11, (1997).

6. Dougherty, R. P., Functional beamforming, *Proceedings of the 5th Berlin Beamforming Conference*, Berlin, Germany, 19–20 February, (2014).

7. Swinbank, R. and Purser R. J., Fibonacci grids: A novel approach to global modelling, *Quarterly Journal of the Royal Meteorological Society*, **132** (619), 1769-1793, (2006).

8. Sloane, N. J. A., Hardin, R. H., and Smith, W. D., Tables of spherical codes, [Online.] available: `http://neilsloane.com/index.html#TABLES`

9. Dougherty, R. P., Cross spectral matrix diagonal optimization, *Proceedings of the 6th Berlin Beamforming Conference*, Berlin, Germany, 29 February – 1 March, (2016).

10. Sarradj, E., A generic approach to synthesize optimal array microphone arrangements, *Proceedings of the 6th Berlin Beamforming Conference*, Berlin, Germany, 29 February – 1 March, (2016).