Sound source localisation using a single acoustic vector sensor and multichannel microphone phased arrays

Wen-Qian JING[1,2]; Daniel FERNANDEZ COMESAÑA[1]; David PEREZ CABO[1]

[1] Microflown Technologies, the Netherlands

[2] Hefei University of Technology, China

ABSTRACT

In recent years, there has been growing interest in the development of noise prediction and reduction techniques. The ability to localise problematic sound sources and determine their contribution to the overall perceived sound provides an excellent first step towards reducing noise. Several well-known methods can be applied in order to achieve a detailed acoustic assessment using microphone phased arrays. However, pressure-based solutions encounter difficulties assessing low frequency problems and their performance is often limited by spatial coherence losses. Alternatively, the use of acoustic vector sensor (AVS) offers several advantages in such conditions due to their vector nature. Each AVS is comprised of a pressure microphone and three orthogonal particle velocity sensors, allowing for the sound direction of arrival to be determined at any frequency within the audible frequency range. Sound localisation techniques using AVS are evaluated in this paper, comparing the characteristics of this innovative solution with respect to traditional microphone phased arrays.

Keywords: acoustic vector sensor, particle velocity, source localisation, beamforming, microphone phased arrays.

I-INCE Classification of Subjects Number(s): 74.6

1. INTRODUCTION

There are many applications which require the utilisation of microphone arrays in order to localise sound sources. However, the number of sensors and the size of microphone arrays required to achieve reliable results is often prohibitive, particularly if the frequency range of interest is wide. Furthermore, the measurement resolution would depend upon the number of sensors used and their respective positions (the geometry of the array). If the array consists of too many sensors, it becomes acoustically significant, biasing the characterisation of the sound field. In contrast, an Acoustic Vector Sensor (AVS) integrates a sound pressure microphone with three orthogonally placed particle velocity sensors to provide the sound Direction Of Arrival (DOA). Figure 1 shows a picture of a particle velocity sensor together with an AVS, also known as "3 dimensional intensity probe".

The acquisition of a vector quantity possess a number of advantages over conventional measurements of the (scalar) sound pressure (1). This topic was first covered in detail from a theoretical point of view by Nehorai and Paldi in 1994, introducing the signal model of a vector sensor into the field of signal processing (2). AVS were later applied to sound source localisation in air in 2002 (3), where sound intensity was used to localise a monopole source. In 2009 (4) a single AVS was utilized for locating two incoherent sound sources by using the MUSIC algorithm (multiple signal classification). Later, Wind et al. evaluated the performance of an AVS array for aeroacoustic applications, reviewing its practical advantages over microphone arrays (5, 6).

In this paper, a single AVS and several microphone phased arrays are evaluated by localising sound sources in three dimensional space and far field conditions, focusing upon the results achieved in the audible frequency range. The Delay-And-Sum (DAS) algorithm and the Capon algorithm are both used to study the performance of the different systems in terms of localisation accuracy and spatial resolution.
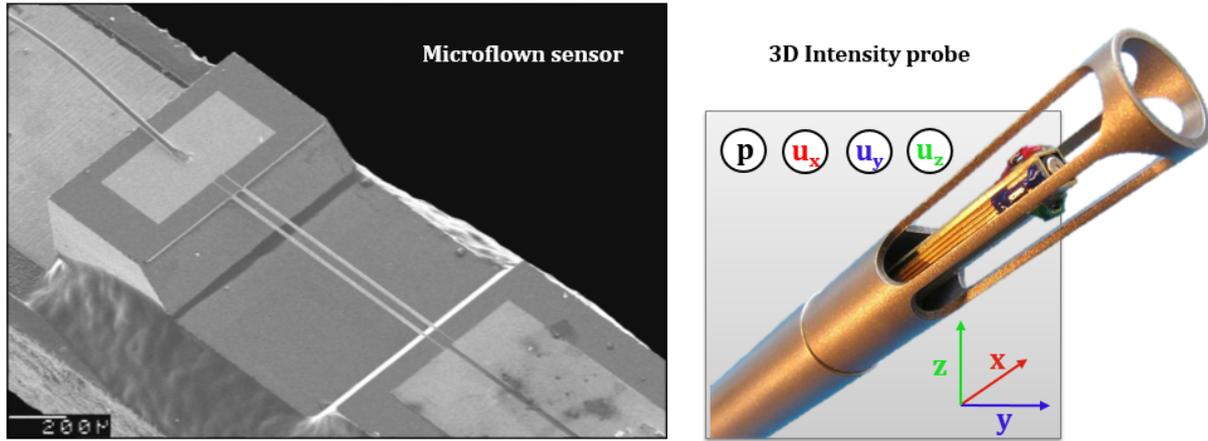
Figure 1 – Acoustic particle velocity sensor or Microflown (Left) and 3 dimensional intensity probe (Right).

2.   DATA MODEL

A simplified data frequency domain model is introduced in the following sections. The covariance matrix and the steering vectors of both pressure-based and velocity-based systems used in the simulations are hereby described.

2.1   Sensor array and covariance matrix

The signals perceived by N sensors can be expressed in matrix form as

$$\mathbf{y}(\omega) = [y_1(\omega), y_2(\omega), ..., y_N(\omega)]^T. \tag{1}$$

The covariance (or cross-spectral) matrix of the data can be then formulated as

$$\mathbf{R} = [\mathbf{y}\mathbf{y}^H]. \tag{2}$$

In the particular case of a set of sound pressure microphones, the above expressions can be re-formulated as such

$$\mathbf{y}_p(\omega) = \mathbf{p}(\mathbf{x}_n, \omega) = [p(\mathbf{x}_1, \omega), p(\mathbf{x}_2, \omega), ..., p(\mathbf{x}_N, \omega)]^T. \tag{3}$$

The covariance matrix of a microphone array is therefore composed by the cross-spectral terms of different sensor positions $\mathbf{x}_n$, i.e.

$$\mathbf{R}_p = [\mathbf{y}_p\mathbf{y}_p^H] = \begin{bmatrix} p(\mathbf{x}_1, \omega)\,p(\mathbf{x}_1, \omega)^* & p(\mathbf{x}_1, \omega)\,p(\mathbf{x}_2, \omega)^* & \dots & p(\mathbf{x}_1, \omega)\,p(\mathbf{x}_N, \omega)^* \\ p(\mathbf{x}_2, \omega)\,p(\mathbf{x}_1, \omega)^* & p(\mathbf{x}_2, \omega)\,p(\mathbf{x}_2, \omega)^* & \dots & p(\mathbf{x}_2, \omega)\,p(\mathbf{x}_N, \omega)^* \\ \vdots & \vdots & \ddots & \vdots \\ p(\mathbf{x}_N, \omega)\,p(\mathbf{x}_1, \omega)^* & p(\mathbf{x}_N, \omega)\,p(\mathbf{x}_2, \omega)^* & \dots & p(\mathbf{x}_N, \omega)\,p(\mathbf{x}_N, \omega)^* \end{bmatrix}. \tag{4}$$

On the other hand, for an AVS comprised of a pressure microphone and three orthogonal particle velocity sensors, the signal matrix can be expressed as

$$\mathbf{y}_v(\omega) = [p(\mathbf{x}, \omega), u_x(\mathbf{x}, \omega), u_y(\mathbf{x}, \omega), u_z(\mathbf{x}, \omega)]^T. \tag{5}$$

The covariance matrix of a single AVS could be then expressed as

$$\mathbf{R}_v = [\mathbf{y}_v\mathbf{y}_v^H] = \begin{bmatrix} p\,p^* & p\,u_x^* & p\,u_y^* & p\,u_z^* \\ u_x\,p^* & u_x\,u_x^* & u_x\,u_y^* & u_x\,u_z^* \\ u_y\,p^* & u_y\,u_x^* & u_y\,u_y^* & u_y\,u_z^* \\ u_z\,p^* & u_z\,u_x^* & u_z\,u_y^* & u_z\,u_z^* \end{bmatrix}. \tag{6}$$

2.2   Plane wave model and steering vector

If the sensor system is placed sufficiently far from the sound source, the microphone array or the vector sensor will be exposed to incident plane waves, as shown on the left hand side of Fig. 2. In the particular case of the microphone array, each sensor position determines the time delay between the sound emission and reception. Traditional beamforming techniques steer a beam to a particular direction by computing a properly weighted sum of the individual sensor signal. As such, this procedure results in the addition of signals coming from the direction of focus, maximising the energy of the beamformer output whilst sound waves from other directions are attenuated. A set of time delays $\tau_m(\kappa)$ can be computed from the scalar product between the sensor position $\mathbf{r}_m$ and a unitary vector κ which aims the direction of interest, i.e.

$$\tau_m(\kappa) = \frac{\kappa \cdot \mathbf{r}_m}{c}. \tag{7}$$

The unitary vector κ and the sensor position $\mathbf{r}_m$ can be respectively projected in three orthogonal components of a Cartesian axis as

$$\kappa = (k_x, k_y, k_z), \tag{8}$$

$$\mathbf{r}_m = (r_x, r_y, r_z). \tag{9}$$

The vector κ is related to the angle of azimuth θ and elevation φ of the propagating wavefronts as follows

$$\begin{aligned} k_x &= \cos(\varphi)\cos(\theta), \\ k_y &= \cos(\varphi)\sin(\theta), \\ k_z &= \sin(\varphi). \end{aligned} \tag{10}$$

The steering vector can be expressed as

$$\mathbf{a}(\theta, \varphi, \omega) = e^{j\omega\tau}. \tag{11}$$

Substituting the time delay τ in equation (7) to equation (11), one can get the steering vector as

$$\mathbf{a}(\theta, \varphi, \omega) = e^{j\omega\frac{\cos(\varphi)\cos(\theta)r_x + \cos(\varphi)\sin(\theta)r_y + \sin(\varphi)r_z}{c}}. \tag{12}$$

For an AVS, besides the plane wave model, the directivity of the AVS also has influence on the steering vector. In the right of Fig. 2, the directivity of an AVS is schematically shown.
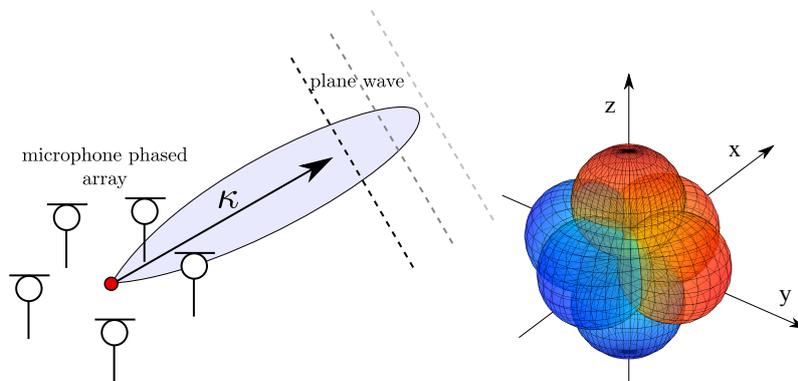
Figure 2 – Illustration of a microphone phased array aimed towards the sound direction of arrival (left) and directivity of an AVS (right).

Combined with the plane wave model in equation (12), we get the weights of the sensor elements as following:

$$\begin{aligned} w_p &= e^{j\omega\frac{\cos(\varphi)\cos(\theta)r_x + \cos(\varphi)\sin(\theta)r_y + \sin(\varphi)r_z}{c}}, \\ w_x &= \cos(\varphi)\cos(\theta)e^{j\omega\frac{\cos(\varphi)\cos(\theta)r_x + \cos(\varphi)\sin(\theta)r_y + \sin(\varphi)r_z}{c}}, \\ w_y &= \cos(\varphi)\sin(\theta)e^{j\omega\frac{\cos(\varphi)\cos(\theta)r_x + \cos(\varphi)\sin(\theta)r_y + \sin(\varphi)r_z}{c}}, \\ w_z &= \sin(\varphi)e^{j\omega\frac{\cos(\varphi)\cos(\theta)r_x + \cos(\varphi)\sin(\theta)r_y + \sin(\varphi)r_z}{c}}. \end{aligned} \tag{13}$$

So for a AVS, the steering vector can be expressed as

$$\mathbf{a}(\theta, \varphi, \omega) = [w_p, w_x, w_y, w_z]^T \tag{14}$$

3.   SOURCE LOCALISATION

One common application for acoustic sensor arrays is the Direction Of Arrival (DOA) estimation of propagating wavefronts for the localisation of noise sources. Generally, array geometry information is used in combination with the processed signals recorded by each sensor in order to create spatially discriminating filters (7). This spatial filtering operation is also known as beamforming.

3.1   DAS beamformer

The conventional Delay-And-Sum (DAS) beamformer (8) maximizes the output power of a given sensor array for a certain input. With the convariance matrix $\mathbf{R}$ and steering vector $\mathbf{a}$, the pseudo-spectrum at a direction (θ, φ) can be expressed as

$$P_{\text{DAS}}(\theta, \varphi, \omega) = \mathbf{a}^H(\theta, \varphi, \omega)\mathbf{R}(\omega)\mathbf{a}(\theta, \varphi, \omega). \tag{15}$$

3.2   Capon beamformer

The Capon beamformer (9) (also known as minimum-variance distortionless response beamforming) is a high resolution algorithm that provide asymptotically unbiased estimations of source localisation. It is developed as a constrained optimisation problem that relies on the inversion of the data covariance matrix. The Capon's pseudo-spectrum of a given direction (θ, φ) is calculated with the convariance matrix $\mathbf{R}$ and steering vector $\mathbf{a}$ as follows

$$P_{\text{Capon}}(\theta, \varphi, \omega) = \frac{1}{\mathbf{a}^H(\theta, \varphi, \omega)\mathbf{R}^{-1}(\omega)\mathbf{a}(\theta, \varphi, \omega)}. \tag{16}$$

3.3   Error calculation

The error between theoretical and calculated positions can be undertaken as long as the position of the noise source is known. For 3D localisation techniques, error can be measured by calculating the euclidean norm of the error vectors of the sound source s in azimuth and elevation direction, i.e.

$$\|\mathbf{e_s}\| = \sqrt{e_{\theta_s}^2 + e_{\varphi_s}^2}, \tag{17}$$

where e_{θ_s} and e_{φ_s} are the errors between estimated and theoretical positions of the sound source s in azimuth and elevation direction respectively.

3.4   Spatial resolution

The spatial resolution of a sound localisation method determines the ability to distinguish two closely spaced noise sources. Usually the spatial resolution is represented by the -3 dB width of the main lobe. The geometry and number of channels are the main factors that determine the spatial resolution of an array. The effects of varying these parameters is studied in greater detail in a later section, comparing the perfomance achieved with both microphone phased arrays and an AVS.

4.   AN AVS VERSUS MICROPHONE PHASED ARRAYS

There are several commercial multichannel arrays that can provide reliable localisation of sound sources. In this paper, four common arrays are compared with an AVS, as described in Table 1.

Table 1 – Parameters of the sensor arrays used in the simulations

Number of sensors	Geometries of arrays	Measurement apertures (m)
4	AVS	0.01
32	Sphere	0.35
48	Star	3.4
90	Wheel	2.43
128	Circle	0.8

This section is divided in two parts: firstly, the error and spatial resolution are calculated to give a general comparison; next, the DOA maps at different frequency ranges are shown. All results are presented considering the two methods mentioned above, the DAS beamformer and the Capon beamformer, using a SNR of 30 dB.

4.1   Localisation accuracy and spatial resolution

First of all, the results obtained using the DAS beamformer with the different systems are compared in Figure 3. As shown on the left hand side of the figure, the error achieved with an AVS is small and very

consistent in the evaluated frequency range. In contrast, the localisation error obtained with the microphone arrays is fairly large at low frequency range, especially with the star array and wheel array, probably due to the low microphone spatial density. Among the microphone arrays, the sphere array which is a 3D array and has high spatial density of microphones gives the best accuracy. In addition, the resolution presented on the right of the figure shows that an AVS gives a very stable resolution versus frequency, though its value is relatively large. For a single AVS, the pressure and particle velocity sensors are placed at the same position and therefore instead of delaying the signals for achieving strong signal amplification and cancellations, the amplitude is modulated producing smooth localisation maps frequency invariant. The reason why the value is large may be that there is no time delay between sensors because only one AVS is used. Then observing the resolution achieved with the microphone arrays, it can be seen that their values are very large at low frequency range and become much smaller when the frequency increases. And it can be found that the star array and wheel array can give better resolution whose measurement apertures are very large, while the sphere array and circle array give much worse resolution whose measurement apertures are much smaller.
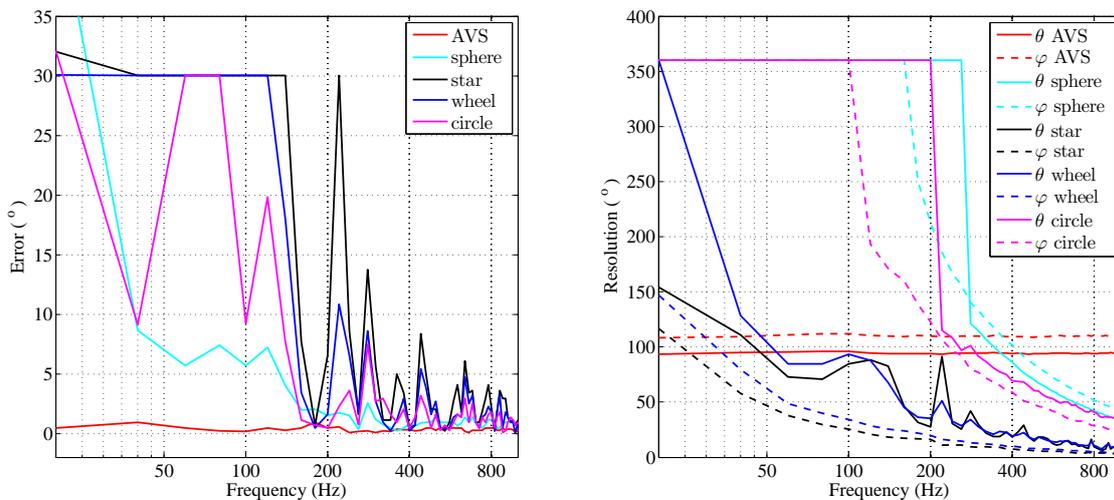
Figure 3 – The properties of the DAS beamformer: Error (Left) and Resolution (Right).

Also, the Capon beamformer is utilized to estimate the DOA, and the error and the resolution versus frequency are presented in Fig. 4. From the figure, it is obvious to find that an AVS gives a very good accuracy and resolution and these properties are very stable with frequency. But the microphone arrays have very bad accuracy and resolution results at low frequency range, though these properties become much better at mid-high frequency.
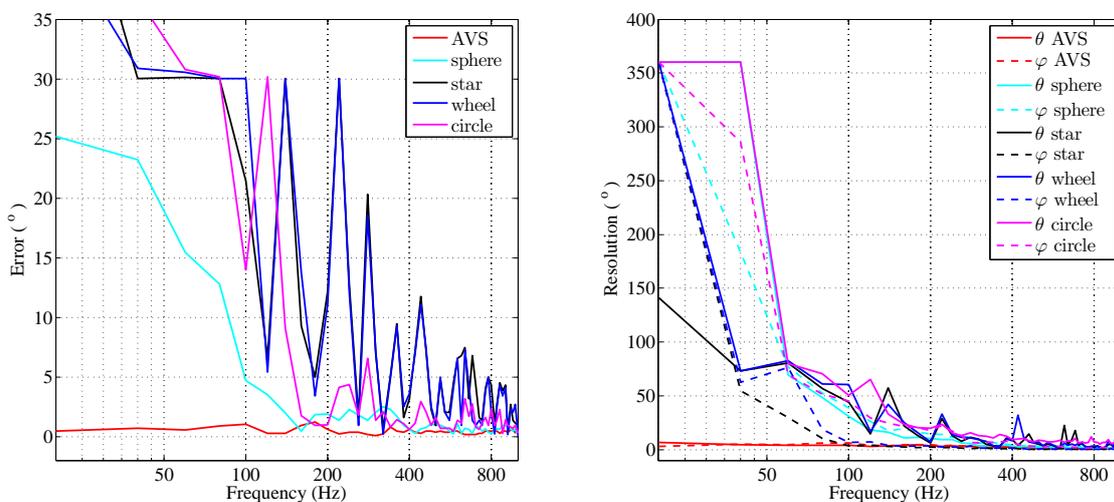
Figure 4 – The properties of the Capon beamformer: Error (Left) and Resolution (Right).

4.2   Source localisation maps

As shown in the previous section, results obtained with an AVS are independent of frequency. For the sake of clarity, only the DOA maps at a low frequency range are displayed. Figure 5 presents the maps obtained by using both the DAS and the Capon beamformers. On the left of Figure 5, it can be seen that a single AVS is capable of localising the source accurately using the DAS beamformer, but the focusing spot is relatively large. However, when the Capon beamformer is used, the performance of a single AVS improves remarkably, localising the source accurately with a very sharp focusing spot.
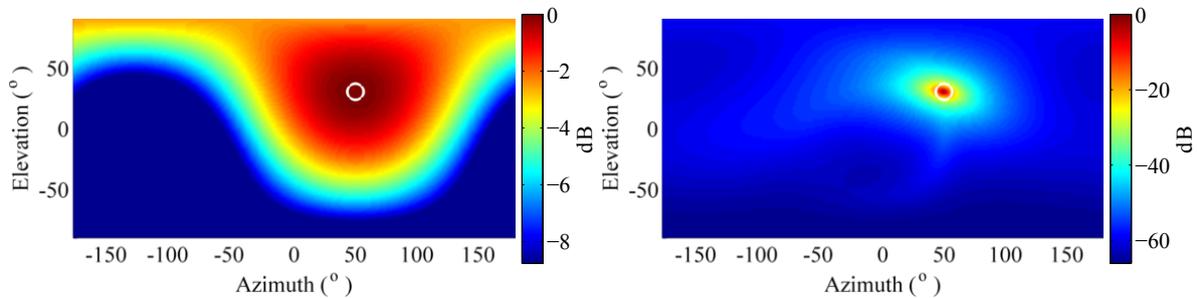
Figure 5 – DOA maps of a single AVS at 10-100Hz: DAS beamformer (left) and Capon beamformer (right).

Furthermore, Figure 6 shows the DOA maps achieved with four microphone array systems using the DAS beamformer at low, mid and high frequency ranges. On the left hand side of the figure, it can be seen that the sphere array can localise the source accurately at low and middle frequency range. Nonetheless, the sphere array cannot give good resolution at low frequencies due to its small array dimensions. On the other side of
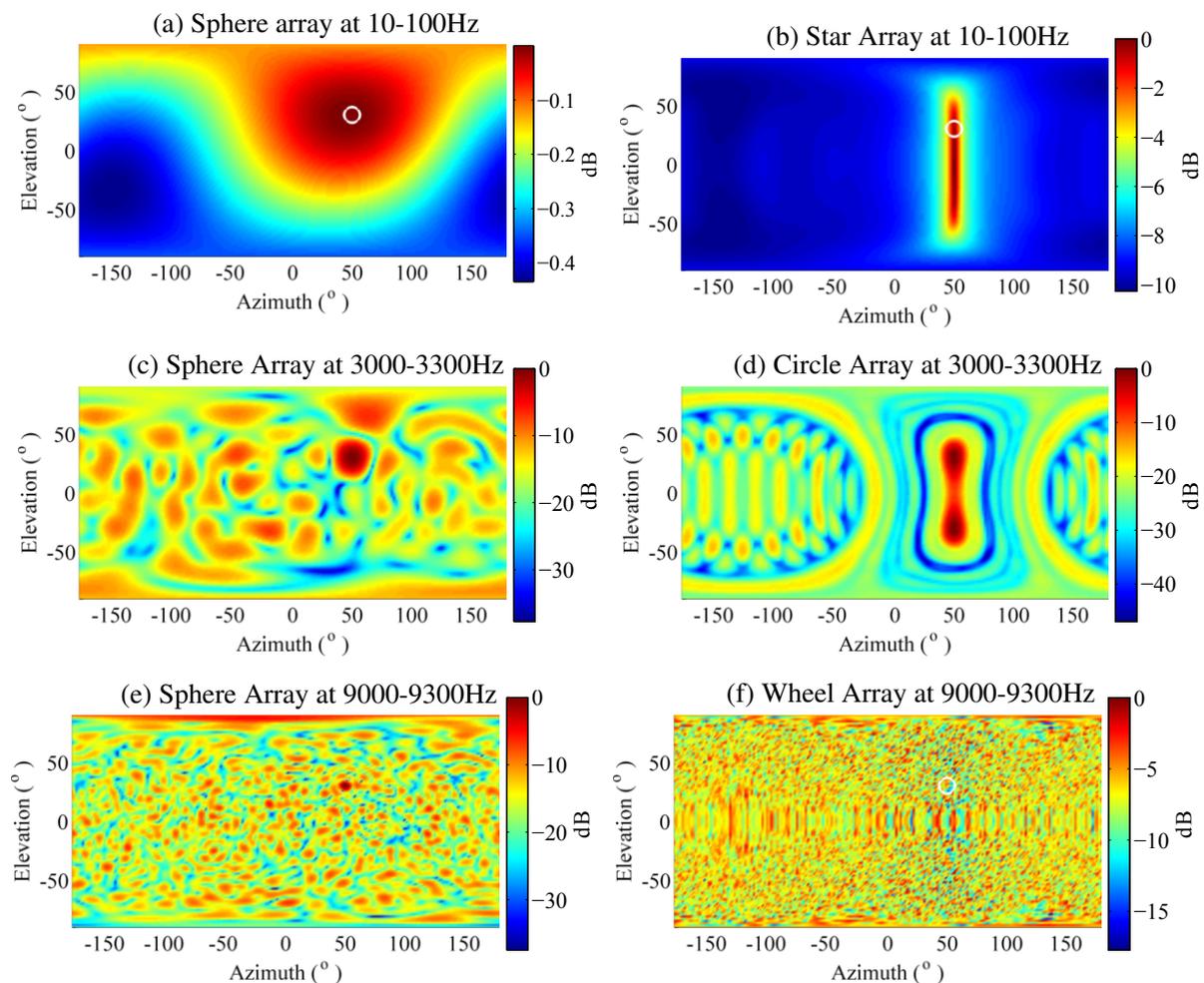
Figure 6 – DOA maps obtained with microphone arrays by using the DAS beamformer.

the figure, the star array, the circle array and the wheel array all encounter image source problem, showing symmetry along the elevation direction because they all are 2D arrays, they use planar geometries. These geometries can create ghost sources between the real source and its image source, especially at low frequency ranges. At high frequency range, above 9 kHz, all microphone-based arrays encounter difficulties due to spatial aliasing.

In addition, the Capon DOA maps achieved with the four microphone arrays studied above are shown in Figure 7. Compared to the maps obtained by using the DAS beamformer, it is easily to find that the Capon beamformer always yields maps with a larger dynamic range, but both techniques provide similar location estimations. As a result, the sphere array is capable of localising the source accurately at low-middle frequency range with much smaller sidelobes than with DAS; however, the three 2D arrays still encounter image source problems. All the microphone array system studied still show aliasing phenomenon at high frequencies.
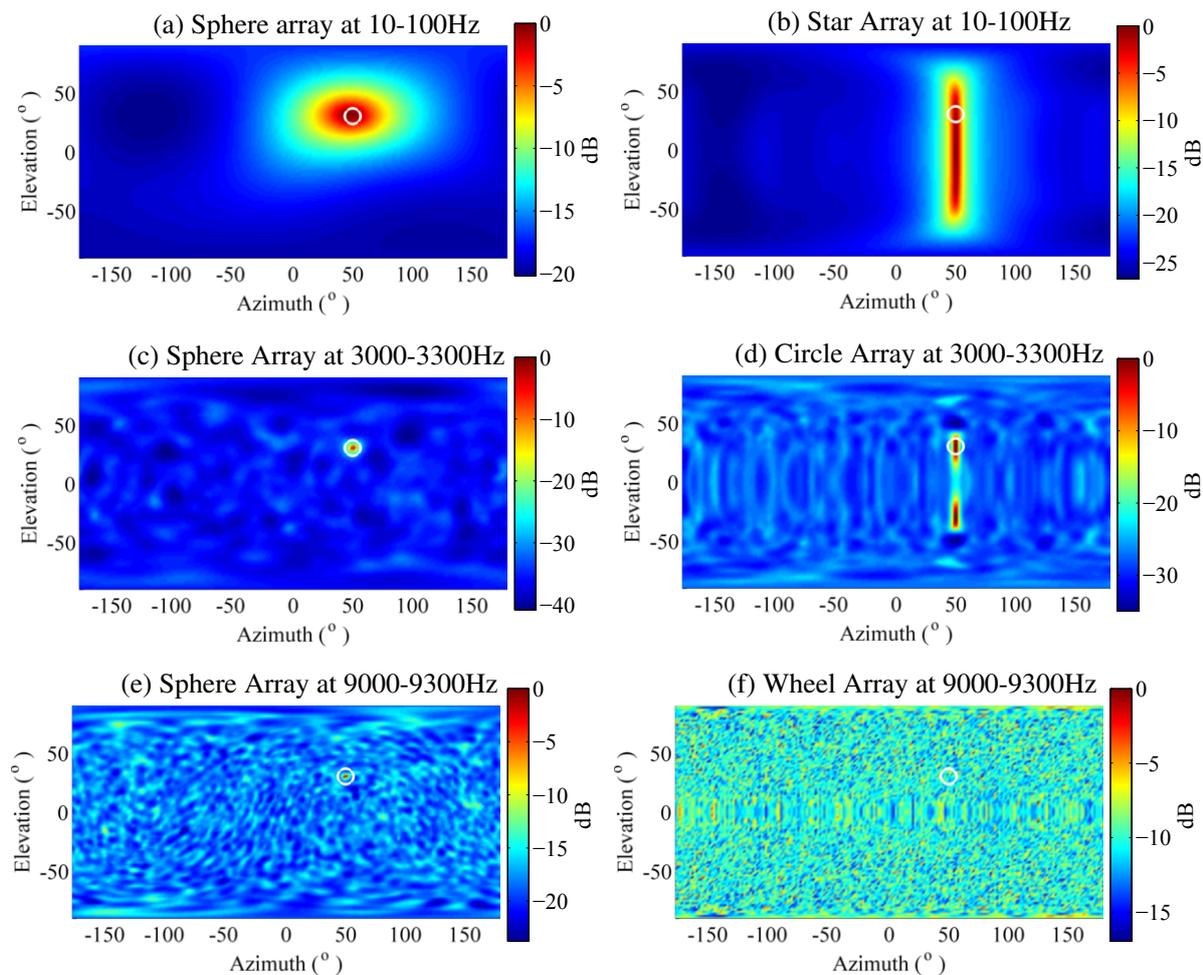
Figure 7 – DOA maps obtained with microphone arrays by using the Capon beamformer.

5. COMPARISON OF THE AVS WITH MICROPHONE PHASED ARRAYS

From the simulation results presented above, it can be concluded that a single AVS has several advantages over the conventional multichannel microphone phased arrays for locating a single dominant noise source:

- sound localisation maps preserve the same spatial resolution and accuracy properties for all frequencies.
- there is no sidelobes in the DOA maps
- single sensor position, this potentially can simplify problems where spatial coherence is key.

One of the main limitations of a single AVS is the number of sound source: the maximum number of sources that can be localised with the Capon beamformer depends upon the rank of the covariance matrix, for a single AVS it has a maximum of 4 channels. This limit could be enhanced by using information from several sensors. Wind et al. have made several study for multiple sources localisation (more information in (5)).

6.   CONCLUSIONS

The sound source localisation performances of multichannel microphone phased arrays and a single AVS have been compared. As shown, AVS have a series of advantages over microphone phased array systems, especially at low and high frequency ranges. The frequency independent spatial resolution, the absence of ghost sources and the lack of spatial aliasing are the main advantages of the AVS compared to traditional microphone arrays.

REFERENCES

1. Hawkes M, Nehorai A. Acoustic vector-sensor beamforming and Capon direction estimation. Signal Processing, IEEE Transactions on. 1998 Sep;46(9):2291–2304.

2. Nehorai A, Paldi E. Acoustic vector-sensor array processing. Signal Processing, IEEE Transactions on. 1994;42(9):2481–2491.

3. Raangs R, Druyvesteyn E. Sound Source Localization Using Sound Intensity Measured by a Three Dimensional Pu-probe. In: Audio Engineering Society Convention 112; 2002. .

4. Basten T, de Bree H, Druyvesteyn W, Wind J. Multiple incoherent sound source localization using a single vector sensor. In: ICSV16, Krakow, Poland; 2009. .

5. Wind J, Tijs E, Yntenna D. Acoustic vector sensors for aeroacoustics. In: CEAS; 2009. .

6. Wind J, de Bree HE, B X. 3D sound source localization and sound mapping using a p-u sensor array. In: Proceedings of CEAS-AIAA; 2010. .

7. Manolakis DG, Ingle VK, Kogon SM. Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing. McGraw-Hill series in electrical and computer engineering: Computer engineering. McGraw-Hill; 2000.

8. Krim H, Viberg M, et al. Sensor array signal processing: two decades later. 1995;.

9. Capon J. High-resolution frequency-wavenumber spectrum analysis. Proceedings of the IEEE. 1969 Aug;57(8):1408–1418.