Designing Acoustic Vector Sensors for localisation of sound sources in air

M. Shujau, C. H. Ritz and I. S. Burnett
School of Electrical, Computer, and Telecommunications Engineering
University of Wollongong, Wollongong NSW Australia
[ms970, critz]@uow.edu.au

School of Electrical and Computer Engineering
RMIT University, Melbourne, VIC, Australia
ian.burnett@rmit.edu.au

ABSTRACT

Abstract— This paper investigates the design and application of an Acoustic Vector Sensor (AVS), traditionally used for underwater applications, for localisation of sound sources in air. The paper investigates the relationship between the design factors and the accuracy of Direction of Arrival (DOA) estimates for sources of varying frequency and compares the performance of an AVS with a Uniform Linear Array (ULA) of comparable size. Results show that the design of the AVS is critical in achieving polar responses that result in DOA estimates of high accuracy. For our proposed design, DOA estimates were found to be more accurate for a range of source frequencies when compared to an existing AVS design and were significantly more accurate than those obtained from a ULA of comparable size.

1. INTRODUCTION

Acoustic Vector Sensors (AVS) arrays have been studied extensively for underwater applications including sound source localisation [1], but very little work exists for applications in air. An AVS consists of three orthogonally mounted acoustic particle velocity sensors and one omni-directional acoustic pressure microphone, allowing the measurement of scalar acoustic pressure and all three components of acoustic particle velocity [2]. Compared to linear microphone arrays, AVS's, consisting of collocated microphones, are significantly more compact (typically occupying a volume of 1 cm^3) [3] and can be used to record audio signals in both the azimuth and elevation plane without modifying the orientation of the array. Typical applications of an AVS include speech enhancement for hands free communication [3], sound source localization in air and detection of seismic activities.

Directional information from an AVS can be extracted by examining the relationship between the 4 microphone channels. Accurate Direction of Arrival (DOA) estimates are dependent upon placement of the microphones, the structure that holds the microphones and the polar patterns generated by each microphone. To investigate these effects we will consider alternative AVS designs and compare the resulting polar responses and DOA estimates with those obtained from a benchmark AVS. Section 2 of this paper proposes an AVS design that results in improved polar responses compared to an existing AVS design [4]. Section 3 compares the polar responses of these three AVS designs and examines the dependency of DOA estimate on both the design and source frequency. Section 4 compares results for DOA estimates for sound sources recorded using the three AVS designs as well as a traditional Uniform Linear Array (ULA). Conclusions are presented in Section 5.

2. ACOUSTIC VECTOR SENSORS: THEORY AND DESIGN

2.1 Source Localization using AVS

The output of the AVS consists of two components: an acoustic particle velocity and acoustic pressure component. This can be expressed in vector form as:

$$\mathbf{y}(t) = [p(t), v_x(t), v_y(t), v_z(t)]^T \quad (1)$$

Where $p(t)$ represents the acoustic pressure component and $v_x(t)$, $v_y(t)$, $v_z(t)$ represents that the velocity gradient of the x, y and z axis. The relationship between the acoustic pressure and the particle velocity is given by [4]:

$$v(r,t) = f\big(p(r,t)\big)\mathbf{u} \quad (2)$$

Where $v(r,t) = [v_x(t) , v_y(t) , v_z(t)]$ represents the acoustic particle velocity vector and f is a function of the acoustic pressure gradient and where:

$$\mathbf{u} = \begin{bmatrix}\cos\theta\cos\phi & \sin\theta\cos\phi & \cos\phi\end{bmatrix}^T \quad (3)$$

is the source bearing vector with θ representing the azimuth and ϕ the elevation [4].

One technique for estimating the source position is to find the ratios of the gradient microphone signals. However, this assumes perfectly orthogonal sensors with matching frequency responses, which was found to not be true in practice. A significantly more reliable and efficient method for finding the DOA estimate is the Multiple Signal Classification (MUSIC) method of Schmidt [5]. The MUSIC algorithm allows for the estimation of the DOA using the eigenvalues and eigenvectors of the covariance matrix formed from the recorded signals [5] [6]. The MUSIC algorithm can be expressed as:

$$P(\theta_i) = \frac{1}{\sum\left|\mathbf{V}_i^H \mathbf{h}(\theta_i)\right|^2} \quad (4)$$

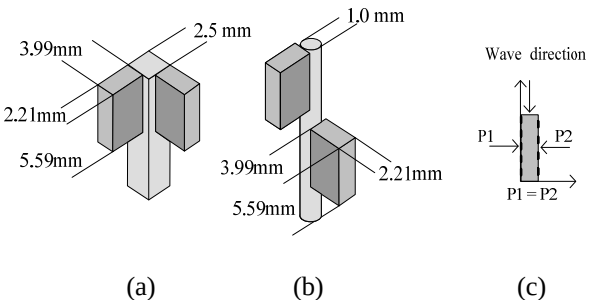

Figure 1 - 3D View of (a) AVS I (b) AVS II (O and Z not shown) (c) Pressures P1 and P2 at the two inlets

where $\mathbf{V}_i^H$ is the smallest eigenvector and $\mathbf{h}(\theta_t)$ is the steering vector for the AVS and where $\theta_t \in (-\pi, \pi)$. For sources with an elevation of 0° relative to the AVS (assumed in this work), the steering vector [6] [7] can be described as a function of the azimuth as:

$$\mathbf{h}(\theta_i) = \left[\cos(\theta_i) \;\; \sin(\theta_i) \; 1\right] \qquad (5)$$

which is formed from the x and y components of (3) with $\phi = 0°$ and where 1 represents the omni-directional microphone. The peaks of $P(\theta_t)$ represent the DOA estimate for that source.

2.2 Design of the AVS

The AVS's constructed for the present work were built using one Knowles EK-3132 omni-directional pressure microphones and three Knowles NR-3158 pressure gradient microphones, similar to those used in [3]. Three designs were investigated. The first design (AVS I) is similar to that reported by Lockwood et al. [3] and is illustrated in Fig. 1 (a) and pictured in Fig. 2 (a).

Initial results found that the accuracy of a (DOA) estimate is highly dependent on the design and placement of the microphones. It is proposed that these differences are due to the effect of shadowing by the sensor on the opposite axis and the size of the aluminium pole used. The area blocking the path is approximately (3.99×5.59) 22.36 mm^2 as illustrated in Fig. 1 (a) which is the surface area of the sensor on the opposite axis. As the frequency of the audio signal increases, the shorter wave lengths result in increased reflection and blocking of the sound waves by the sensors, wires, and the metal support. High frequencies can not move around objects easily due to the short wave lengths hence they are more reflected [8]. This results in an artificial increase in pressure on one side of the sensor; which must be minimized for efficient use of the AVS. Hence, this work focuses on design changes of the AVS to minimise these effects. An alternative might be to predict the effect of the shadowing but this is a highly variable and complex problem.

To test the effects of the sensor placement, AVS I was altered such that the sensors located on the opposite axis were placed with an offset as pictured in Fig. 2(b). An offset of approximately 0.5 cm was chosen; this distance represents the approximate one-quarter wavelength for a 15 kHz source (the maximum frequency that the sensors can record). To test the effect of the metal support and the sensor placement a

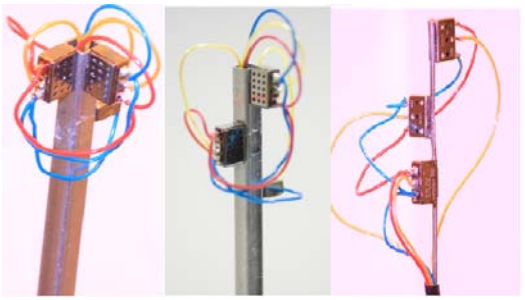

Figure 2: (a) AVS I front, (b) AVS I Offset (c) AVS II

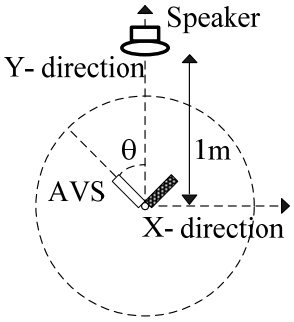

Figure 3: Setup for characterization of the AVS

second AVS was built (AVS II) with an offset between the two sensors in the X and Y axis as illustrated in Fig. 1 (b) and pictured in Fig. 2 (c). To avoid shadowing, the omni-directional microphone was also located at a similar distance from the gradient microphones as used for AVS I with offset. A thin cylindrical metal rod of cross sectional area of 0.79 mm^2 is used (Fig. 2 (c)) instead of the square aluminium pole which had a cross sectional area of (2.5×2.5) 6.25 mm^2 as shown in fig. 1(a). The use of a cylindrical metal rod allows the smooth flow of air particles thus reducing reflections [9].

3. CHARACTERIZATION OF THE AVS

The characterization of the AVS sensors was performed by measuring their polar responses and Direction of Arrival (DOA) estimates for sources located directly in line with the microphones.

3.1 Experimental Setup

Recordings were made in an anechoic room at the University of Wollongong [10]. For testing, the experimental setup of Fig. 3 was used, where the AVS was mounted on a custom built rotating platform (to allow positioning of the microphones relative to the source) and a self powered speaker (Genelec 8020A) was placed in front of the AVS at a distance of 1 m with an elevation of 0°. A series of monotone signals each 2 seconds long and of equal energy were played with frequencies ranging from 100 Hz to 10 kHz. Recordings of 2 seconds long were made at 5° intervals and signals were sampled at 48 kHz.

3.2. Polar Responses

The polar response was measured by finding the average recorded signal energy for each source location. Analysis

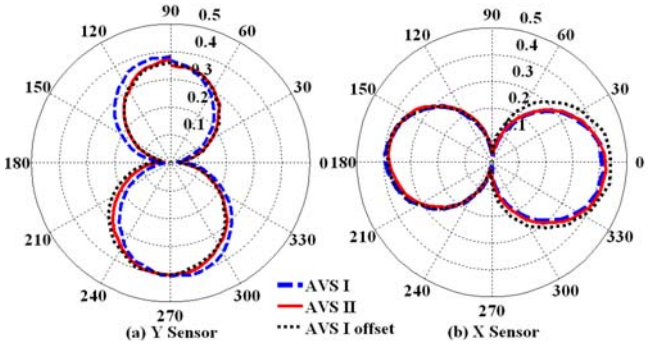

Figure 4: X and Y Sensor Polar Plots for AVS I , AVS I with offset and AVS II at 3 kHz

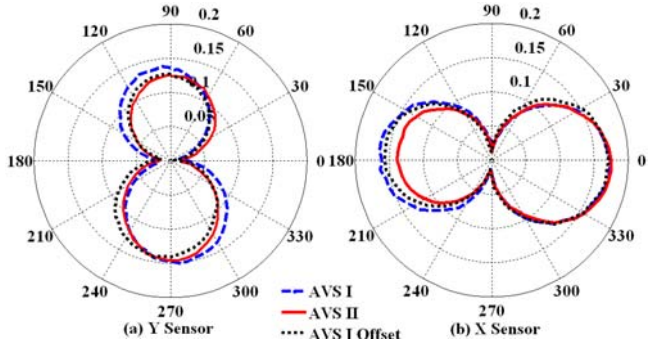

Fig 5: X and Y Sensor Polar Plots for AVS I , AVS I with offset and AVS II at 5 kHz

here is restricted to the gradient microphone (polar responses for the omnidirectional microphone were found to be constant for a wide range of frequencies). Figs. 4 and 5 show results for 3 kHz and 5 kHz tones for the gradient microphones recording both the x and y directions (see Fig. 1 (b)). All AVS designs display a typical figure-of-eight polar pattern [11]. For all AVS's examined, the lobe on one half of the figure 8 pattern was larger than the other; this was found to be a result of the characteristics of the Knowles gradient microphones, which had differing numbers of air-inlets on each side. Compared with the results for AVS I, the polar plots for AVS I with offset and AVS II are more symmetrical. The loss of symmetry for AVS I about the Y axis is particularly noticeable for the Y sensor as the source frequency increases (compare Figs. 4 and 5), while for AVS I with offset and AVS II, symmetry of the polar response is maintained as source frequency increases.

3.3. Response to source perpendicular to sensor inlets

An ideal gradient microphone should have an output of 0 for sources located perpendicular to the sensor inlets; this is because the pressure will be identical on either side of the microphone as illustrated in Fig. 1 (c) and hence the pressure gradient (or difference) should be zero.

To analyse this characteristic, the Average Angular Error (AAE) of signals impinging on the AVS at 0° was measured for AVS I, AVS I with offset and AVS II for the frequencies from 1 kHz to 10 kHz using:

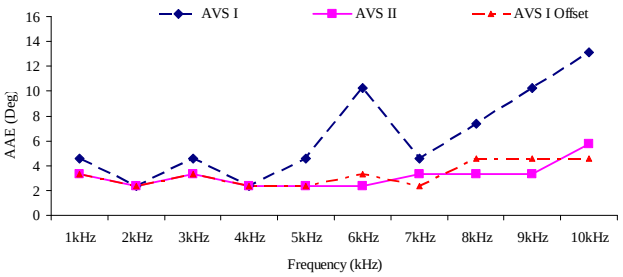

Figure 6: The AAE for output at 0^0 for frequencies 1 kHz to 10 kHz

$$AAE = \frac{1}{N}\sum_{n=1}^{N}\left|\theta_{n,m} - \theta_{n,a}\right| \qquad (6)$$

Where N is number of sources (tones) and $\theta_{n,m}$ and $\theta_{n,a}$ are the measured (m) and actual (a) DOAs, respectively, for source n. Fig. 6 shows the difference error for sources located at 0° to the Y axis; the error for AVS I is much higher than that compared to AVS I with offset sensors and AVS II. Furthermore the results show that that the difference in error increases as the frequency of the source signal increases.

In the case of AVS I when the source is at 0° to the Y sensor, air particles flowing on the X sensor side are disrupted creating a slight increase in pressure on one side. This disruption of the air flow causes an error in the output and this error increases as the frequency of the signal increases. When the sensors are placed at offset in the AVS I the error is reduced significantly and for frequencies up to 5 kHz the error is the same as that for AVS II. The improved result is due to the offsetting of the sensors which causes minimum disruption of the flow of the air particles creating an output which is more accurate than the output of AVS I.

4. LOCALISATION EXPERIMENTS

Results for localization were obtained using the same experimental rig, recording environment and sound sources described in Section 3. As well as the three AVS configurations, recordings were also made with a 3 element Uniform Linear Array (ULA), which was chosen so that the number of sensors matched the AVS. The ULA was designed using the same Knowles EK-3132 omnidirectional microphones as used in the AVS and using a spacing of 21 mm; this results in an array of approximately 42 mm long. Localization was performed using the DOA estimate described in Section 2. All the results presented in this section are for confidence intervals of 95 % and for frequencies 1 kHz to 10 kHz in 1 kHz intervals and for the first two quadrants i.e. $\theta \in (0, \pi)$. The results for the 3rd and 4th quadrants are the same as those for the 1st and the 2nd quadrants; hence for brevity they have been omitted.

4.1 Direction of Arrival for AVS I , AVS I with offset and AVS II

The average angular error for sources located in the 1st quadrant and averaged over all recorded source frequencies is shown in Fig. 7. On average, the AAE is 1.5° for AVS II, for

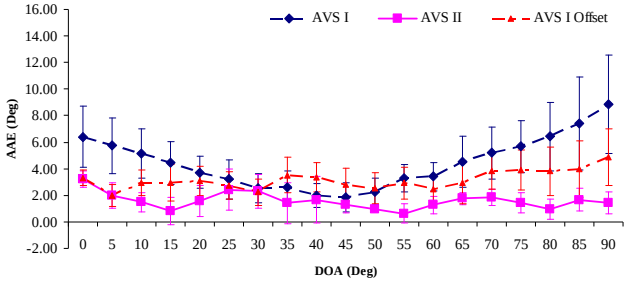

Figure 7: AAE of the DOA estimates for 1st quadrant. Error bars represent 95 % confidence intervals

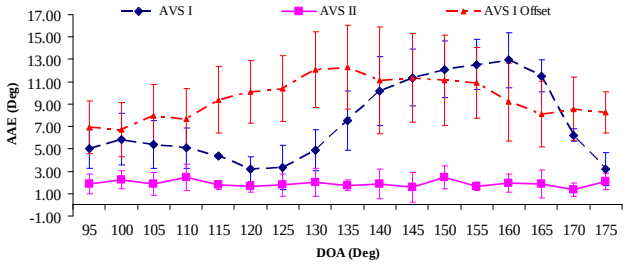

Figure 8: AAE of the DOA estimates for 2nd quadrant. Error bars represent 95 % confidence intervals

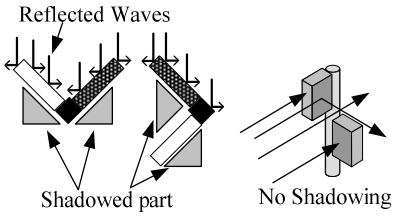

Figure 9: (a) AVS I Reflected waves creates Shadowed Regions (b) AVS II No shadowing due to offset

AVS I with the offset the AAE is 3.2° and 4.6° for AVS I. The results of Fig. 8 are for the AAE of the DOA estimates for the second quadrant. On average, the AAE for AVS II is 1.9 while for AVS I with offset is 9.5 and for AVS I it is 7.3. Compared to the results from the 1st quadrant the overall error for all AVS's is seen to have increased significantly for the second quadrant. The results for second quadrant are for frequencies 1 kHz to 6 kHz as the error from frequencies including 7 kHz and above were statistically not reliable for AVS I.

For the first quadrant it is proposed that the increase in error for AVS I is caused by acoustic shadowing by the sensor on the opposite axis. A significant improvement in error is seen when the sensors are at offset (see AVS I offset results and AVS II). It is proposed that this improvement is due to reduced blocking from any object to the flow of the air particles; hence the sensor readings are more accurate.

For the second quadrant, there are two factors that contribute to the errors: (a) reflection from the edges of the square pole at high frequencies and (b) shadowing or blocking by the sensor on the opposite axis at high frequencies. It is believed that the reflections and blocking from the square pole has a greater significance in the error as suggested by

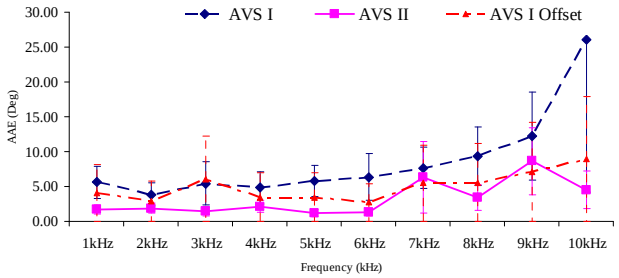

Figure 10: AAE for each frequency Band vs Frequency. Error bars represent 95 % confidence intervals (top half of error bar for 10kHz removed for clarity)

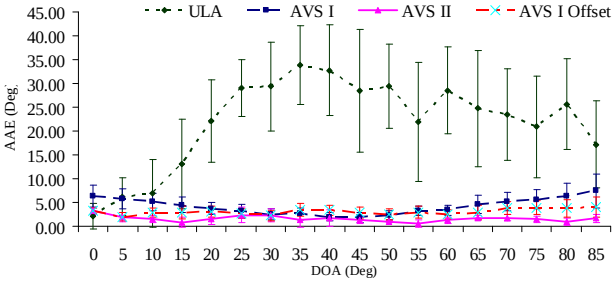

Figure 11: ALE for DOA estimates for AVS I, AVS II and ULA. Error bars represent 95 % confidence intervals.

the results in Fig. 8. By removing the square pole and replacing it with the thin metal pole there is a significant reduction in error for AVS II. In Fig. 9, it can be seen that the impinging sound wave hits the sensor and the square pole and the waves are reflected creating regions of attenuations or shadowed parts. These cause incorrect readings of pressure difference that produce an error in the output for AVS I, and for the AVS I with offset the square aluminium pole cause reflections which create errors in the output. In contrast, for AVS II the metal pole is much smaller than the sensors, which are also offset; this results in a pressure difference and an output with minimum error.

The relationship between DOA estimation and the geometry of the sensors has been previously investigated by Hawkes et al. [12]. Here, since the aluminium pole used for AVS I is square the accuracy with which the sensors are attached to form 90° between the sensors is high compared to that of the metal rod used in the construction of AVS II, the results of Figs. 7 and 8, show much higher errors in the estimated source location for AVS I. This indicates that the acoustic shadowing by the sensors and reflections by the aluminium pole is much more significant and creates larger errors than those produced by the misalignment of sensors.

4.2 DOA estimates Vs Frequency for AVS

Results in Fig.10 show that as the frequency of the source increases the error in the DOA estimate also increases. The average AAEs for source frequencies from 1 kHz to 10 kHz are approximately: 6.1° for AVS I, 5.0° for AVS I with offset; and 2.4° for AVS II. For AVS I, the AAE versus source frequency is approximately constant up to 5 kHz, increases by approximately 2° per kHz between 6 and 9 kHz before in-

creasing sharply to 20° at 10kHz. The AAEs for AVS I with the offset remain below 9° for source frequencies up to 10 kHz, while for AVS II the maximum error is 8° (except at 9 kHz).

This result shows that by offsetting the sensors and reducing the surface area of the structure holding the AVS microphones, more consistent and accurate DOA estimates can be obtained for all source frequencies tested. By first offsetting the sensors, a reduced DOA estimate error is achieved for high frequencies (as seen by the results for AVS I offset). Replacing the square pole with a cylindrical pole of much smaller area leads to further reductions in the DOA estimate errors for all frequencies.

4.3 Comparison of DOA estimates for AVS and ULA

The results in Fig. 11 are a comparison of the DOA error produced by a ULA to that of the AVS. The ULA is the simplest and most common type of microphone array. For optimum performance of a ULA the spacing between the microphones has to be set logarithmically, but since they are only 3 microphones they have been attached with the same separation. The steering vector (also known as the array response vector) for the ULA used with the equation 4 is described as:

$$\mathbf{h}(\theta_i) = \frac{1}{\sqrt{M}}\left[1 \;\; e^{-j2\pi[(d\sin\theta_i)/\lambda]} \ldots e^{-j2\pi[(d\sin\theta_i)/\lambda](M-1)}\right]^T \quad (6)$$

Where M is the number of microphones in the array, $\theta \in (0, \pi)$, d is the microphone spacing and λ is the wave length of the signal.

The results show that at a distance of only 1 m from the source AVS II has an average error of 1.6°compared to that of ULA which has and average error of 21.8°. AVS I has an average error of 4.5° which is the worst result for all AVS's but is still approximately 4 times better than the average error produced from a ULA with same number of microphones and comparable size. This results show that the performance of AVS's are much better than ULA's of comparable size.

To produce results comparable to the AVS, the ULA would need to be placed much further (at least 2.5 m to 3 m) from the source and use more microphones (at least 5) with much larger separations [11]. Preferably, microphones should be separated logarithmically so that the array responds accurately to tones of different frequencies [11].

5. CONCLUSION

This paper has proposed an AVS design that ensures highly accurate estimates of DOA for in-air applications. The results obtained show that there is significant impact on the DOA estimates by acoustic shadowing and the shape and size of the metal support used for holding the sensors, and these effects are related to the frequency of the impinging sound wave. The results show that by placing the sensor such that the sensor on the off axis does not block the path of the adjacent sensor the result of the DOA estimates are improved significantly for the first quadrant.

Furthermore, the results show that by changing the shape of the support from square to cylindrical, which reduces the cross sectional area of the support by 5.46 mm^2, provides a significant improvement in the estimated DOA accuracy. The DOA estimates obtained from the new design have an average error of less than 2° for a range of source frequencies, compared with average errors of more than 4.5° for an alternative existing design. Furthermore it has been established that the accuracy of the DOA estimates generated by the AVS is much better than the estimates for a ULA with similar number of sensors and comparable size at close proximity to the target source. Future work will examine applications of this AVS for speech enhancement applications using e.g. beamforming and extensions to 3D source locations and alternative acoustic environments.

Acknowledgement

This project was partially supported by the Australian Research Council Grant DP0772004.

References

[1] Hawkes, M., Nehorai, A., "Hull-Mounted Acoustic Vector-Senor Processing", *Proc of the ASILOMAR – 29* ,1046-050, November 1996.

[2] Hawkes, M., Nehorai, A., "Bearing Estimates With Acoustic Vector Sensor Arrays", *American Institute of Physics Con. Proc*. Vol: 386, pp 345-358, April 1996.

[3] Lockwood, M. L. , Jones, D. L., "Beamformer Performance With Acoustic Vector Sensor In Air", *J. Acoust. Soc. Am*, 119, 608-619, January 2006

[4] Nehorai, A., "Acoustic Vector Sensor Array Processing", *IEEE Transaction on Signal Processing*, Vol. 42, No.9, 2481-2491, September 1994.

[5] Schmidt, R. O., "Multiple Emitter and Signal Parameter Estimation," *Proceedings,RADC Spectral Estimation Workshop*, 243-258, October 1979.

[6] Manolakis, D. G., Ingle, G. K., & Kogon, S. M., "Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing" Boston: Artech House, INC, 2005

[7] Mohan, S., Lockwood, M. E., Karmer, M. L., and Jones, D. L., "Localization of Multiple Acoustic Sources With Small Arrays Using a Coherence Test", *J. Acoust. Soc. Am*, 123, 2136-2147, April 2008.

[8] Traux, B., "Handbook for Acoustic Ecology**"** , Cambridge Street Publishing, 1999

[9] Svensson, U. P., Calamia, P. T., "Scattering from simple shaped objects using edge diffraction," *Proc. of AIJ/ASJ Symposium on computational acoustics*, March 2005.

[10] Ritz, C. H., Schiemer, G., Burnett, I., Cheng, E., Lock, D., Narushima, T., Ingham, S., Wood Conroy, D., "An Anechoic Configurable Hemispheric Environment for Spatialised Sound", *Proc. of the 2008 Australasian Computer Music Conference*, 10-12 July 2008.

[11] Eargle, J., "The microphone Book", Boston: Focal Press, 2001.

[12] Hawkes, M., Nehorai, A., "Effects of Sensor Placement on Acoustic Vector-Sensor Array Performance", *IEEE Journal of Oceanic Engineering*, Vol: 24 (1), 33-40, January 1999.