Acoustic Based Safety Emergency Vehicle Detection for Intelligent Transport Systems

Bruno Fazenda[1], Hidajat Atmoko[1], Fengshou Gu[1], Luyang Guan[1] and Andrew Ball[1]
[1]Diagnostics Engineering Research Group, School of Computing and Engineering, University of Huddersfield, Huddersfield, UK
(Tel : +441484471893; E-mail: b.fazenda@hud.ac.uk)

Abstract: A system has been investigated for the detection of incoming direction of an emergency vehicle. Acoustic detection methods based on a cross microphone array have been implemented. It is shown that source detection based on time delay estimation outperforms sound intensity techniques, although both techniques perform well for the application. The relaying of information to the driver as a warning signal has been investigated through the use of ambisonic technology and a 4 speaker array which is ubiquitous in most modern vehicles. Simulations show that accurate warning information may be relayed to the driver and afford correct action.

Keywords: Safety; Source detection; Warning systems

1. INTRODUCTION

Currently, in-car environments are highly insulated against external noise. It is therefore difficult for the driver to detect the presence of vehicles on emergency travel (ambulances, police, fire-fighting trucks), especially when all car windows are fully closed and the car's audio system is playing audio programme. As inner city roads become ever busier it becomes important to design and implement systems that alert the driver of the proximity of incoming emergency vehicles and, as importantly, which direction they are approaching from.

Previous approaches to solve this problem have taken the form of IR radio signals being emitted by the emergency vehicle and a common radar detector installed in the passenger's vehicle. This demands the installation of a radio device in each emergency vehicle and a sensor in each passenger vehicle. It also implies that radio frequencies need to be reserved for emergency vehicles.

Hitherto there appears to be no evidence of a detection system relying on the ubiquitous acoustic information generated by most emergency vehicles – the siren. This is the oldest and most common means of detecting the presence of emergency vehicles and works well both for drivers as well as pedestrians. The work presented here describes the concept and initial development of a safety device to be installed on vehicles that detects the incoming direction of sound waves emitted by another vehicle on emergency travel and relays this into meaningful alert information for the driver via the existing audio reproduction system in the vehicle. The system should relay to the driver the general incoming direction of the emergency vehicle and, as such, allow the driver to take the necessary action.

The system being developed consists of three stages, Acoustic Detection, Direction Extraction, Acoustic Information Relay, and this paper will be sectioned accordingly with the discussion of results before the concluding remarks.

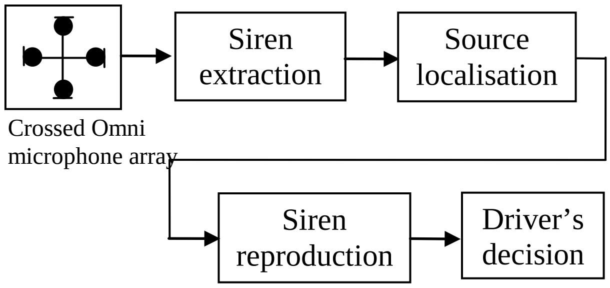

Fig. 1 - Emergency Vehicle Detection system

2. ACOUSTIC DETECTION-SOUND SOURCE LOCALISATION

The detection of the travelling vehicle is based on acoustic detection of the siren. In order to detect this, a set of microphones are placed on the exterior shell of the passenger vehicle. Directional information is extracted using a microphone array formed of pressure transducers (Omni directional) with a known configuration – in this case a crossed array.

2.1 Time Delay Estimation Method

The time delay of arrival for the signal arriving at each microphone is extracted to determine the direction of the incoming source. The technique for determining the direction of a sound source is established using the generalised cross-correlation (GCC) methods [1] as follows.

The relative time difference or delay between microphones 1 and 2, for instance, is defined as

$$\tau_{12} = \tau_1 - \tau_2 \qquad (1)$$

and the signal at each microphone is modelled as

$$x_i(n) = \alpha_i s(n - \tau_i) + \eta_i(n) \qquad (2)$$

where τ_i is the propagation time from the unknown source $s(n)$ to microphone i (i=1,2) and α is an attenuation factor due to sound propagation. $\eta_i(n)$ is assumed as an uncorrelated additive stationary Gaussian random noise. Equation (1) is solved by determining the cross-spectrum between signals x_1 and x_2 with a weighting function $\Phi(\omega)$. The maximum value of the function coefficient is then extracted using Equations (3) and (4):

$$\Phi(\omega) = \frac{1}{|S_{x1x2}(\omega)|} \quad (3)$$

$$\tau_{12} = \arg\max_{\tau}\left(\Phi(\omega)S_{x_1x_2}(\omega)\right) \quad (4)$$

The weighting function $\Phi(\omega)$ in this case takes the form of the phase transform (PHAT) which has been reported to have greater immunity to the effects of reverberation and performs more consistently for source signals that change over time [2].

2.2 Sound Intensity Method

An alternative method has been investigated comprising of a combination of the microphone signals to form a 2 dimensional orthogonal acoustic probe. This method extracts the incoming angle of a sound source by determining sound intensity vectors in the 2 orthogonal directions (Fig. 2). This can be done in the time-domain or using the direct method such as in [3], expressed as

$$I_r = -\frac{1}{2\rho_0\Delta r}\frac{1}{T}\int_0^T (p_A(t)+p_B(t))\int_{-\infty}^{t}(p_B(\tau)-p_A(\tau)\cdot d\tau\cdot dt \quad (5)$$

where I_r is the sound intensity at the point half way between the line connecting the two microphones *A* and *B*. Δr is the distance between microphones. ρ_0 and *T* are the medium density and period of the measurement respectively. In this work, microphones A and B can either be microphone pairs M1/M3 or M2/ M4.

3. MEASUREMENT

Fig. 2 shows the microphone array configuration with a radius of 134 mm. The line connecting the pair M2/ M4 line is established as the array's primary axis - X. All source directions are determined referenced to this axis with angle increasing anti-clockwise.

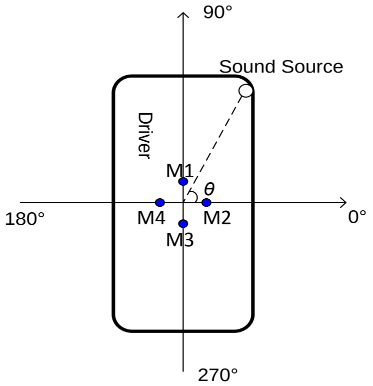

Fig. 2 - Plan view of crossed microphone array position and orientation on car.

Two types of sound sources, a random and a typical siren noise, were used for the experimental tests. The source was modelled with a small speaker of 60 mm diameter and its position was fixed at a known distance to the microphone array centre point. The microphone array was rotated to acquire the emitted sound for 8 positions representing the main regions of interest. The acquired signals were then passed through the siren signal extraction process to separate the siren from the surrounding noise.

3.1 Detection of siren signal

To extract the siren signal, the so-called *adaptive predictor noise canceller* scheme employing the least mean squared (LMS) algorithm such as in [4] has been adopted and incorporated into the system. Fig. 3 demonstrates that the method is capable of extracting the siren signal (Fig. 3b) from the background noise.

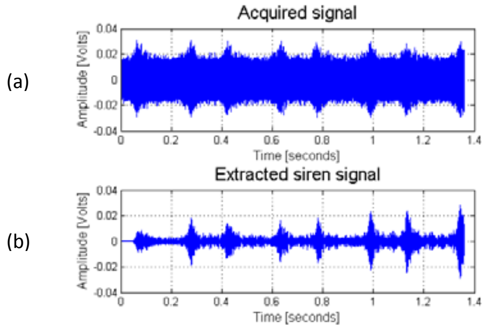

Fig. 3 - The acquired signal from the microphone array. (a) the raw signal and (b) its extracted siren signal

The adaptation process was setup with the following parameters: 10 samples delay, 128 FIR filter taps and 0.25 adaptation step coefficient. Results show that with this particular setup the adaptation process has not fully converged until 1 second has elapsed. However, as the typical signal features of the siren are formed as early as 0.1 seconds, the adaptive method is sufficient to detect the siren source.

The extracted siren signals shown were further processed to obtain the source directions using both the sound intensity and the time delay methods described in the following sections.

3.2 Time Delay Estimation Method

Fig. 4 shows results obtained using the time delay GCC method for both random and siren noise sources. As is evident, the accuracy of source direction identification is quite good with all directions tested being identified within a range smaller than ±5 degrees.

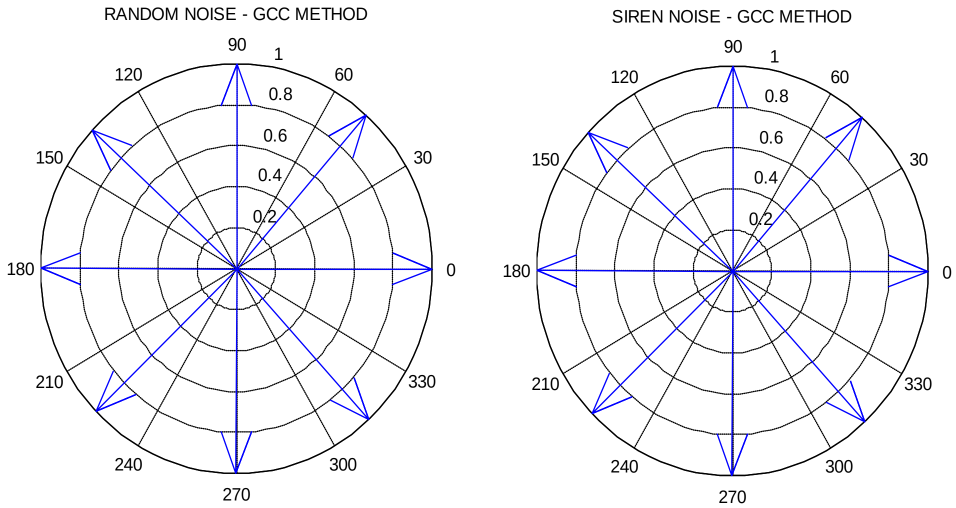

Fig. – 4 Sound source direction estimates utilising the GCC methods for a source of random noise on the left and a siren noise on the right.

3.3 Sound intensity probe method

Fig. 5 shows the results for the sound intensity method in detecting the direction of a random noise source starting at 0 degrees and spaced regularly at 45 degree steps around a full circle.

Results shown run in a clockwise direction starting at the top left plot where microphone M2 faces the source – see Fig. 2. The remaining plots show the result as the microphone is turned about its centre in 45 degree steps.

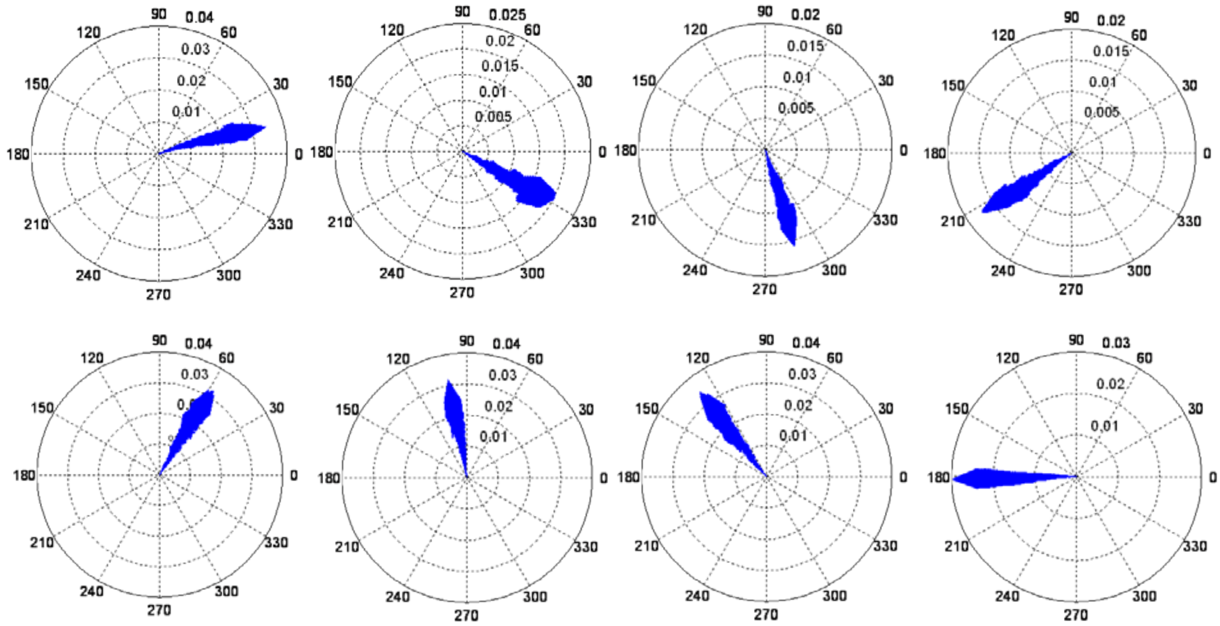

Fig. 5 - Sound intensity distributions at the array position for a source of random noise at 8 regular directions around a circle.

The results show that the intensity method can be used for the application, but source direction detection is less accurate when compared to the time delay estimation method. Source direction is detected within an error range of 15 degrees on the either side of the correct value.

Fig. 6 shows sound intensity distributions for a siren signal as the sound source. Similarly to Fig. 5, results are presented at the upper left corner for 0 degrees and running clockwise for 45 degree increments around the circle.

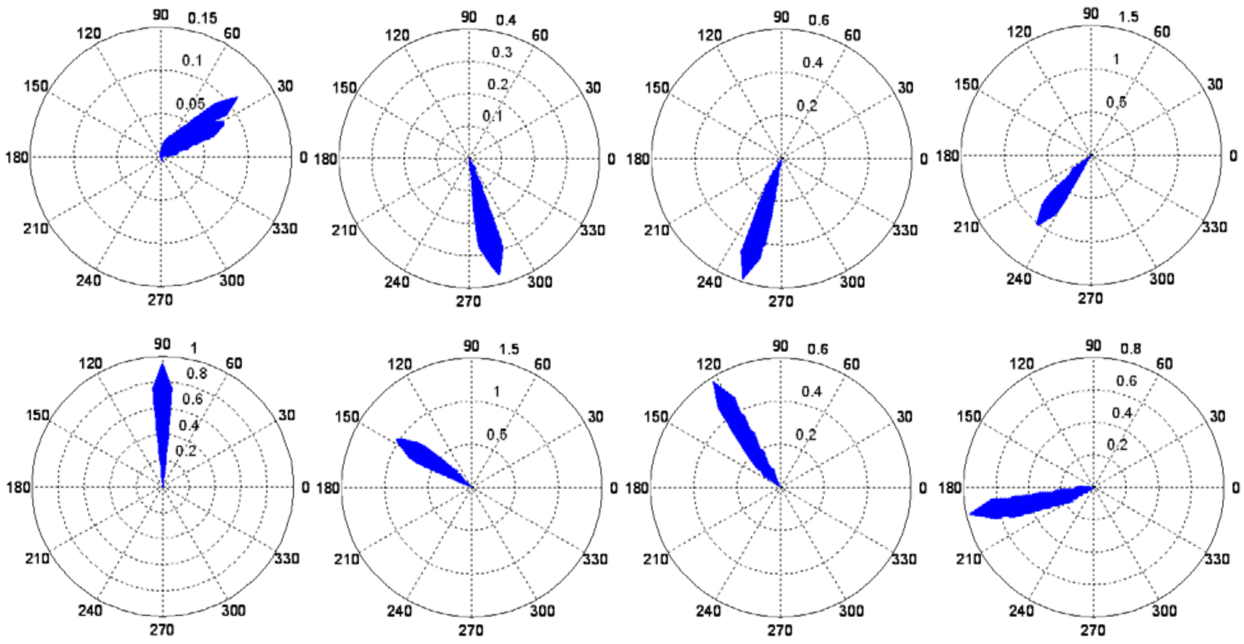

Fig. 6 - Sound intensity distributions for a source producing a siren signal at 8 regularly distributed directions along a circle.

It can be seen that the results are less accurate compared with the results for random noise. The on axis direction is detected as coming from almost 40 degrees to the left. Other positions are also detected with large error margins. Accuracy in the sound intensity system is very much dependent on an accurate calibration and behaviour of the microphone array which appears to affect the results significantly. Further work is being carried out to improve the calibration method and probe performance.

Another important aspect for accuracy in a sound intensity probe is the distance between elements. Its upper frequency range is limited by the spacing between the microphones and where this distance becomes comparable to half of the wavelength of source sound:

$$f_u = \frac{c}{2\Delta r} \qquad (6)$$

In the case presented, Δr=0.268m, which equates to a limiting frequency of around 640Hz for a typical speed of sound of 343m/s. Considering the relatively large microphone separation in the array, and the frequency range of the source signal which sweeps up to 10 kHz, the results are quite encouraging. A reduction in the microphone array would not only form a much more attractive, less bulky probe but would also improve results dramatically. The optimisation of the probe shape is a subject of further investigation by the authors.

3.4 Summary

As is clear from observation of Figs. 4~6, the implementation of this particular probe on source localisation with either random or siren source signals achieves better performance if time delay estimation and the GCC method is used.

The specific distance between probes compromises the performance of each method inversely. A sound intensity probe which is accurate at higher frequencies requires a small spacing between microphones. Clearly, this affects the performance of a time delay estimation probe since it relies on the spacing between microphones to obtain accurate time delay estimates. There appears to be a great scope of investigation for the optimisation of source detection acoustic arrays that perform well using both methods.

4. RELAYING THE INFORMATION TO THE DRIVER

The information that has been obtained by the previous processes needs to be relayed to the passenger's cabin in a useful manner. It is practical to maintain this information in the acoustic domain since this auditory cue is what most drivers (or any road user) rely on to determine the presence and incoming direction of such a source. The in-car audio reproduction system is thus an ideal means of relaying this information to the driver. Most common in-car audio reproduction systems consist of at least four speakers arranged in a close to square shape as described in Fig. 7. The position of the driver is indicated with the symbol *D*.

The relay system proposed is based on a technique commonly used in the field of audio to create a given auditory impression to a listener sat within a known multi-channel loudspeaker system. The recreated sound source is positioned at the desired location by controlling the gains of the signal being sent to each loudspeaker in the system. In the system proposed, the desired source location is obtained from the detection stage as described in previous sections – a source position α is thus available to the relay system.

This information is passed on to a decoder that adjusts the gain of each speaker according to the desired location of the source.

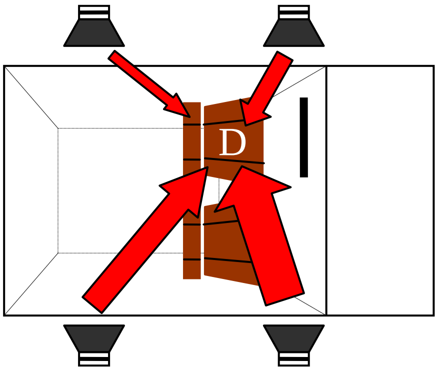

Figure 7 – The reproduction level at each speaker determines the necessary auditory cues to relay the correct directional information to the driver.

A system to relay the correct auditory cues to the driver is now modelled based on the required loudspeaker gains for a listener sitting at the centre point of a loudspeaker array, with all loudspeakers equidistant from the listener.

For a first order system with an evenly spaced speaker array, the gain for each speaker is expressed as [5]:

$$g_i = \frac{1}{N}(1 + 2\cos\alpha_i) \qquad (6)$$

where i denotes the i^{th} speaker, N is the number of speakers in the array and α_i is the angle between the loudspeaker and the desired source direction.

The performance of the system in terms of directional cues relayed to the driver of the vehicle may be measured in terms of *velocity* and *energy* vectors at the centre of the circle [13]:

$$r_v = \mathrm{Re}\frac{\sum_{i=1}^{N} g_i \hat{u}_i}{\sum_{i=1}^{N} g_i} \qquad (7)$$

$$r_e = \frac{\sum_{i=1}^{N} |g_i|^2 \hat{u}_i}{\sum_{i=1}^{N} g_i} \qquad (8)$$

Where u_i is a unit vector in the direction of the speaker. Typically, the velocity vector reflects performance in terms of low frequency localisation whereas the energy vector reflects high frequency localisation performance.

Four discrete source directions have been simulated at 10, 120, 185 and 300 degrees measured from the on axis direction of the speaker arrays (straight ahead).

Fig. 8 shows the resulting velocity and energy vectors for each source position. The amplitude and angle of the vectors indicate the performance of the system. Unity magnitude and angle matching between the desired source direction and that of the vectors reflect an optimal performance of the relay system.

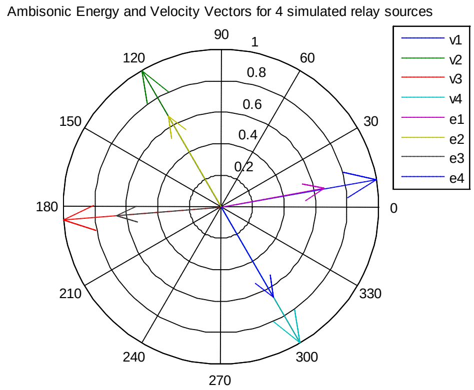

Fig. 8 Velocity (v) and Energy (e) vectors indicating performance of relay system for 4 discrete source warning directions at 10, 120, 185 and 300 degrees from the front.

It can be seen from Fig. 8 that both the energy and velocity vectors accurately match the desired angle direction of the source. The magnitude of the velocity vectors is very close to 1 for all cases, indicating a good localisation cue. The magnitude of the energy vectors is around 0.6 which suggests that localisation cues according to this measure will not be as effective. Improvement to such a system should be enabled by the use of 2nd order ambisonic codec equations which are known to perform better [12,13]. However, the required performance of the system needs to be assessed in a real case scenario in order to determine if the extra cost and complexity is necessary in the application.

Further experimentation and testing of the interaction between relay system and driver is a subject of further investigation being undertaken by the authors. Furthermore, improvements to this system are being investigated for an *off*-centre listening position as is most common in passenger vehicles.

5. CONCLUSION

A system for the warning of an incoming emergency vehicle to the driver of a passenger vehicle has been investigated. The system is designed to integrate with the existing sound reproduction system in the car and make use of the common auditory cues that warn of a vehicle in emergency travel.

It has been demonstrated that a cross array of pressure transducers is capable of determining the incoming direction of a siren as a sound source. For the suggested array radius, methods based on time delay estimation outperform those based on calculating the intensity at the microphone array. This is due to the required distance between units in the array which is conflicting for each method studied. Notwithstanding, both methods have the potential of affording source localisation which is sufficient for the type of application proposed since such a system would possibly benefit from a truncation of the available and unambiguous warning directions relayed to a driver – eg: front, back, left or right.

The post processing of microphone data allows the relay of warning information to the driver inside the vehicle through the existing audio equipment. The required subjective performance of such a system is now under investigation by the authors.

Future developments will concentrate on a fully operational unit, where the system takes control over the audio amplifier whenever an emergency signal is detected. The audio programme is dimmed and the decoded emergency signal is replayed through the audio system. An ancillary visual indicator displaying the approaching direction and speed of the emergency vehicle are also being developed.

REFERENCES

[1] C.H. Knapp and G.C. Carter, “The generalized correlation method for estimation of time delay”, *IEEE Trans. Acoust.*, *Speech, Signal Processing*, Vol. 24, pp. 320-327, 1976.

[2] JM Valin, F. Michaud, J. Rouat, “Robust localisation and tracking of simultaneous moving sound sources using beam forming and particle filtering”, *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* 2006.

[3] Musha T, Mukagawa T, “Intensity analysis methods for transient signals”, Appl Acoust, (2006), doi:10.1016/j.apacoust.2006.08.011

[4] B. Widrow and S.D. Stearns, “Adaptive signal processing”, *Englewood Cliffs NJ: Prentice-Hall*., 1985.

[5] V. Pulkki, “Spatial sound generation and perception by amplitude panning techniques”, *Doctoral Thesis, Helsinki University of Technology,* 2001.

[6] M. A. Gerzon, “Design of ambisonic decoders for multispeaker surround sound”. *AES, Nov*. 1977.

[7] J. Benesty, “Adaptive eigenvalue decomposition algorithm for passive acoustic source localization”, *J. Acoust. Soc. Am.* 107 (1), 2000.

[8] YC Choi and YH Kim, “Near field impulsive source localization in a noisy environment”, *J. Sound and Vibration,* 303, pp. 209-220, 2007.

[9] J. Chen, L Yip, J. Elson, H. Wang, D. Maniezzo, R.E. Hudson, K. Yao, D. Estrin, “Coherent acoustic array processing and localisation on wireless sensor networks”, *IEEE Proceedings,* vol. 91, issue 8, pp. 1154 -1162, Aug 2003.

[10] L. Guan ., F. Gu., B. Fazenda, A. Ball, “Delay of Arrival Estimation in Car for Abnormal Sound Localisation”, *DAMAS 2009*, Beijing, China

[11] F, Bruno, “Localisation of sound sources using coincident microphone techniques”, *Proceedings of the Institute of Acoustics*, 29 (7). pp. 106-115. 2008.

[12] M. Gerzon, “Periphony: With-height sound reproduction”, *JAES Jan/Feb*., Vol.21, No:1, 1973

[13] Benjamin, E., Lee, R., Heller, A., “Localisation in Horizontal-only Ambisonic Systems”, presented at the 121st Convention of the Audio Engineering Society, held October 5 – 8, 2006 in San Francisco, CA USA