Audio Engineering Society

Convention Paper 10555

Presented at the 152nd Convention
2022 May, In-Person and Online

This convention paper was selected based on a submitted abstract and 750-word precis that have been peer-reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This convention paper has been reproduced from the author's advance manuscript with editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library, http://www.aes.org/e-lib. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

MMAD 3D Audio – Designing for Height – Practical Configurations

Michael Williams, Sounds of Scotland, Le Perreux sur Marne, France
Correspondence should be addressed to mike@soundsscot.com

ABSTRACT

Although the basic philosophy behind the design of Microphone arrays, for 3D audio recording and reproduction, has been described in previous AES papers[1][2][3][4], no specific examples have been given with respect to various Surround Sound arrays and the corresponding height arrays (1st layer of height array microphones and the Zenith microphone). This paper gives four examples of complete 3D Audio arrays with perfect critical linking. Examples include same microphone directivity arrays, as well as hybrid arrays. Suitable steering functions are discussed, and specific values are given for each array combination, so as to obtain perfect critical linking.

Introduction

There has been much confusion over the last few years concerning the vocabulary that has been used to describe the various ways of producing, or reproducing, spatial or height information, in the plethora of sound recording systems that have been introduced recently. Wilfred Van Baelen (Galaxy & Auro3D, Mol, Belgium) introduced the term 'Immersive' at the AES International Conference in Tokyo in 2018, and this was later adopted into a SMPTE standard. But this can be contrasted with my preferred term of '**3D Audio**' recording and reproduction. The term 'Immersive' englobes any process where there is an impression of space in the reproduction of an audio recording, good or bad.

On the other hand the term that I prefer, for my own work on height reproduction, is '3D Audio', hence the title of this paper 'MMAD 3D Audio'.

I am specifically researching into the accurate reproduction of sound source localization (or as near as is possible), its associated acoustic architecture, as well as the natural acoustic environment in the complete hemisphere (or sphere) around us. This, to me, is the summum of good quality spatial sound recording. Many pseudo height microphone array recording systems have been created recently, specifically for use in the context of the cinema or home-cinema industry. These pseudo height systems are simply a combination of two surround sound systems placed one above the other - essentially two layers of surround sound

reproduction. It must be understood that these systems have absolutely nothing to do with creating any real impression of height. Although the result in reproduction is far from satisfactory for the discerning listener of 3D, it seems to have been accepted, in the context of cinema or home-cinema reproduction, in giving an impression to the public of something happening above their heads – the listener gets the impression of what is now being called 'an immersive' soundfield.

The MMAD 3D Audio system, on the other hand, is designed to generate a realistic virtual 3D Audio localization of sound in the upper hemisphere, and in both upper hemisphere and in the lower hemisphere for the most recently developed systems.

It is not possible to present a wide selection of microphone arrays in the limited space of an AES paper, but a few representative designs can illustrate the approach to both same microphone directivity arrays as well as a few specific hybrid arrays (arrays using different microphone directivities in each layer of the complete array).

The full array can be divided into 3 sections, the surround sound array in the horizontal plane around the listener, the 1st height layer at around 35° to the horizontal plane, and the Zenith microphone at the top of the complete array.

1 Design procedure

Although the division of the upper hemisphere into two segments of 45° would seem logical at first, hindsight has shown that this will cause considerable difficulty when we come to linking with the upper Zenith Zone. However, a division of 35° for the Height segment (the segment between the Surround Sound Array and the Height Array), and 55° for the Zenith Zone (the segment between the Height Array and the upper Zenith microphone) will be nearer the mark. We will need to analyse each of these segments separately at first.

1.1 All Cardioid Array - The Height Segment

As most recording engineers are well equipped with Cardioid microphones, the first worked example concerns an all-Cardioid array. First of all, looking at the initial Height segment of 35°, i.e. 0° to 35° of elevation. The starting point for this configuration is two parallel Cardioid microphones one above the other at a distance of 126cm., as show in Figure 1 - the value of 126cm is highlighted in red in the SCA Diagram.

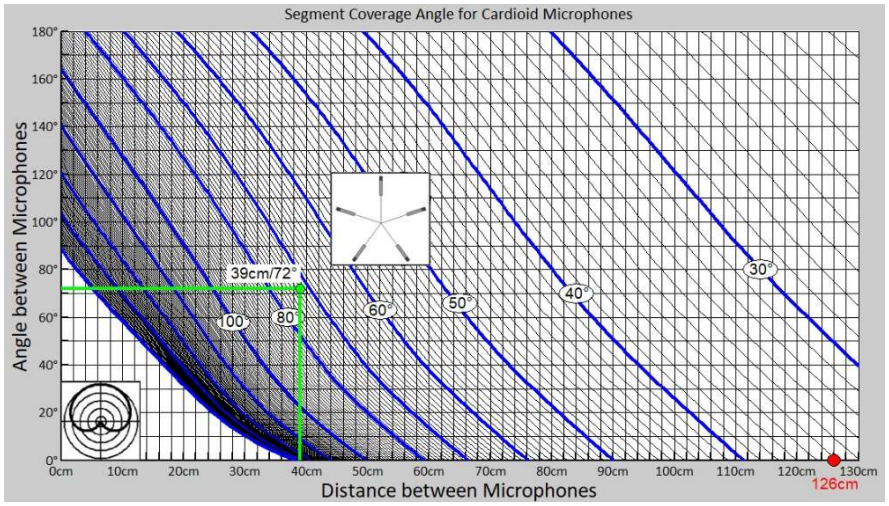


FIGURE 1 – SCA DIAGRAM FOR CARDIOID MICROPHONES

It is important to emphasize that this configuration uses only ***Time Difference*** to generate height information. In the multiple listening tests that have been carried out to study the validity of this method of generating height information, Time Difference generation has proved to be absolutely fundamental in the process of being able to record and reproduce viable height restitution. It is unfortunate that this important parameter of Time Difference has been completely neglected in most other experimentation in height localisation systems. It must also be noted that the microphone array is of course only part of the complete process of a 3D audio system, the loudspeaker configuration must also follow some basic rules in relation to this type of recording system, for instance: univalent microphone/loudspeaker channels and especially isosceles triangle loudspeaker structure.

The other important configuration parameter is the distance and angle between the microphones, this again can be determined from Figure 1, and is highlighted in green – both the 5 channel surround array and the 5 channel height array are 39cm / 72°, i.e. 39cm between the microphone capsules and 72° between the axes of microphone directivity. A plan and elevation view of this array is shown in Figure 2.

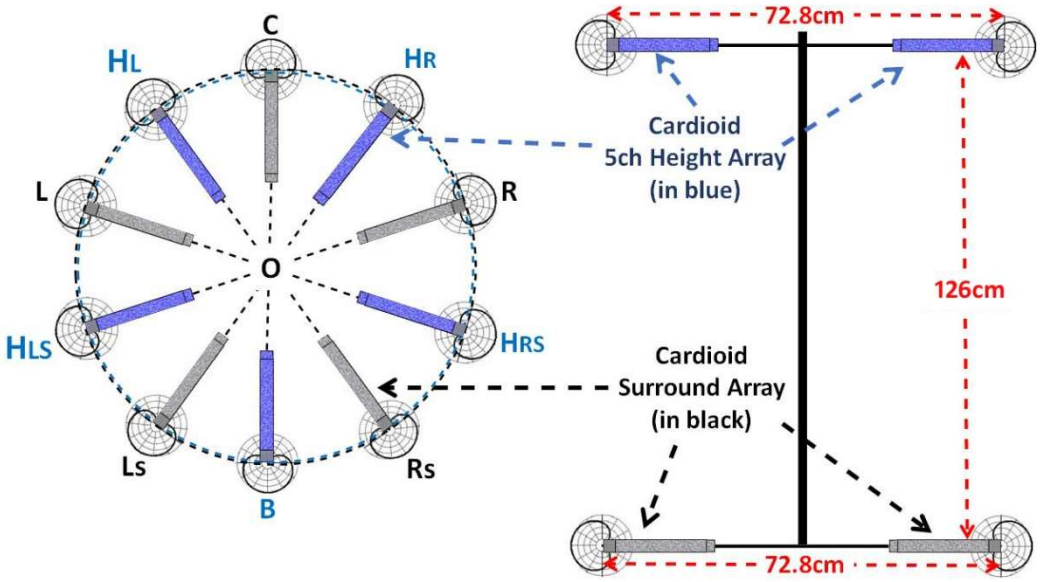


FIGURE 2 - 5 CHANNEL CARDIOID SURROUND SOUND ARRAY
PLUS 5 CHANNEL CARDIOID HEIGHT ARRAY WITH ROTATION OF 36°

The surround sound array is therefore formed by Cardioids in a 5 Channel configuration (drawn in black) with a wingspan of 72.8cm. The Height array is also formed by Cardioids in a 5-channel configuration (drawn in blue), again with a wingspan of 72.8cm. The first elevation array is rotated through 36° degrees so that individual microphones in the elevation array form isosceles triangles with respect to the lower surround sound array microphone pairs.

As the wingspan of the Height Array is the same as the Surround Sound Array, there is no initial Microphone Position Offset, as shown in Figure 3.

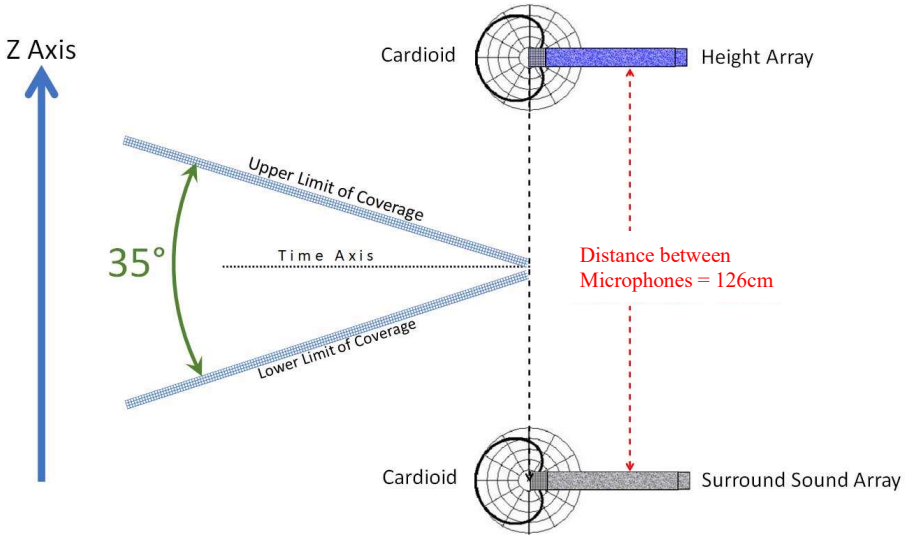


FIGURE 3 - 5 CHANNEL CARDIOID SURROUND SOUND ARRAY
PLUS 5 CHANNEL CARDIOID HEIGHT ARRAY WITH NO SEGMENT STEERING

In order to align the lower coverage limit with the horizontal reference plane, which will rotate the upper coverage limit to 305°, as shown in Figure 5, we need to apply clockwise steering of 17.5°. This is achieved by applying 1mS Electronic Time Offset to the Height Array i.e. the first elevation array microphones have to be delayed by 1mS with respect to the lower surround sound array microphones.

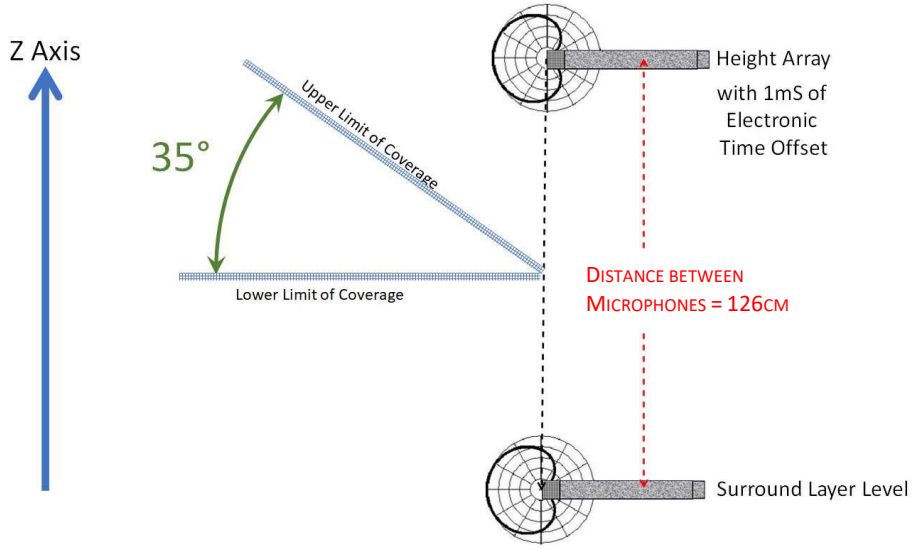


FIGURE 4 - 5 CHANNEL CARDIOID SURROUND SOUND ARRAY
PLUS 5 CHANNEL CARDIOID HEIGHT ARRAY WITH 1mS ETO

1.2 All Cardioid Array - The Zenith Zone

If we now switch to the Zenith Zone Coverage, we are looking for a coverage of +/-55°, i.e. 55° on each side of the zenith microphone. Zenith is also commonly called 'the Top'. We can read off the SCA diagram below (Figure 6), when the Segment Coverage is 55° (highlighted by a green dot), the array is 51cm/90° (51cm between the microphone diaphragms and 90° between the axes of directivity).

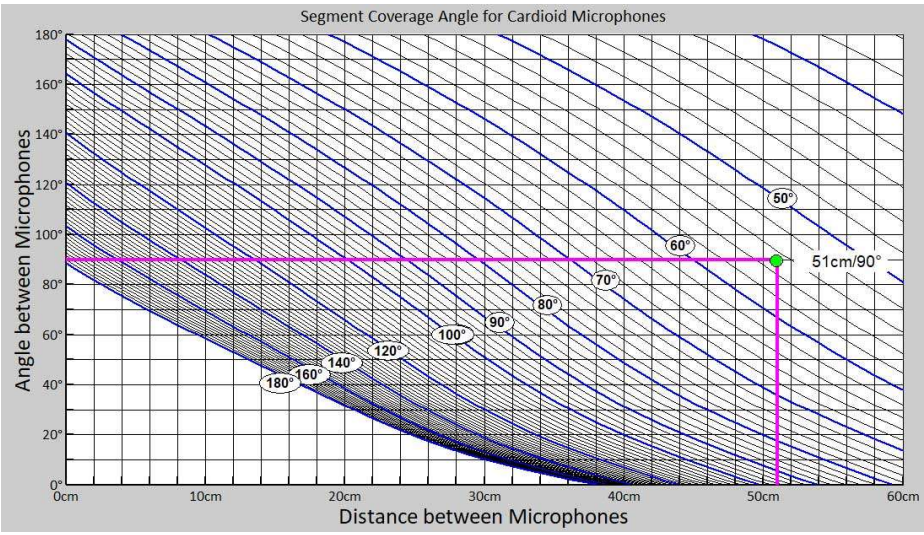


FIGURE 5 - SEGMENT COVERAGE ANGLE DIAGRAM FOR CARDIOID MICROPHONES

The elevation plan of this configuration is shown in Figure 6

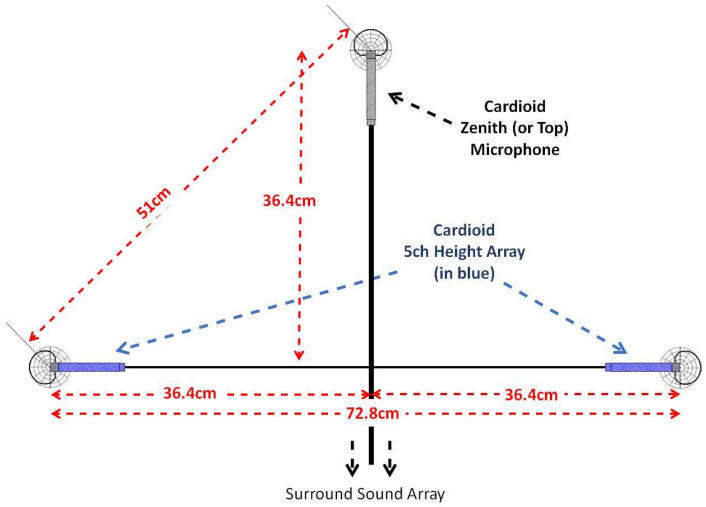


FIGURE 6 - ZENITH ZONE ARRAY USING CARDIOID MICROPHONES

Coverage of each part of this Zenith Array is shown in Figure 7.

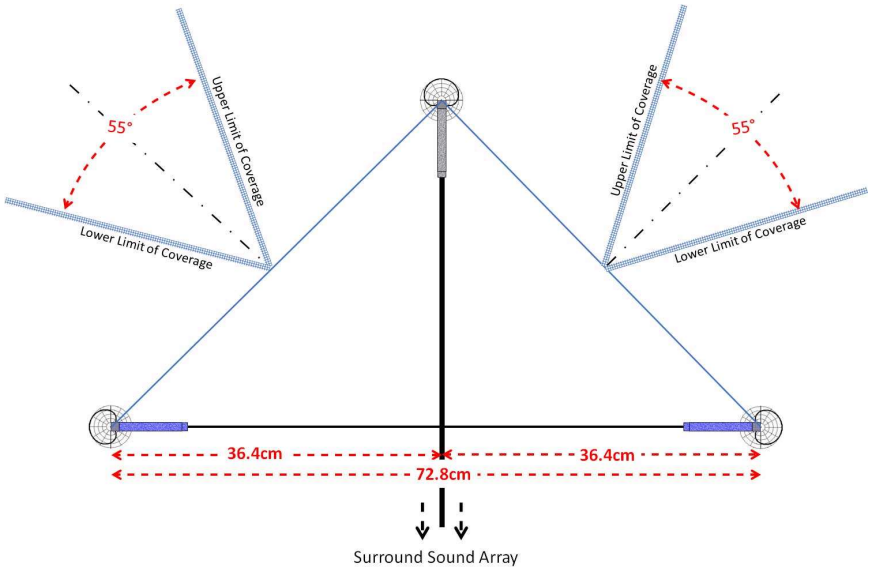


FIGURE 7 - SEGMENT COVERAGE ANGLE FOR A ZENITH ZONE SEGMENT OF 55°
FOR A 36.4CM/90° RIGHT-ANGLED ISOSCELES TRIANGLE ZENITH ZONE ARRAY

If we apply Electronic Time Offset of 17.5° or 1mS to the Zenith Zone Segment Coverage, we can steer the Zenith Zone coverage segments so that they are aligned with the vertical Zenith microphone axis, as shown in Figure 9
.

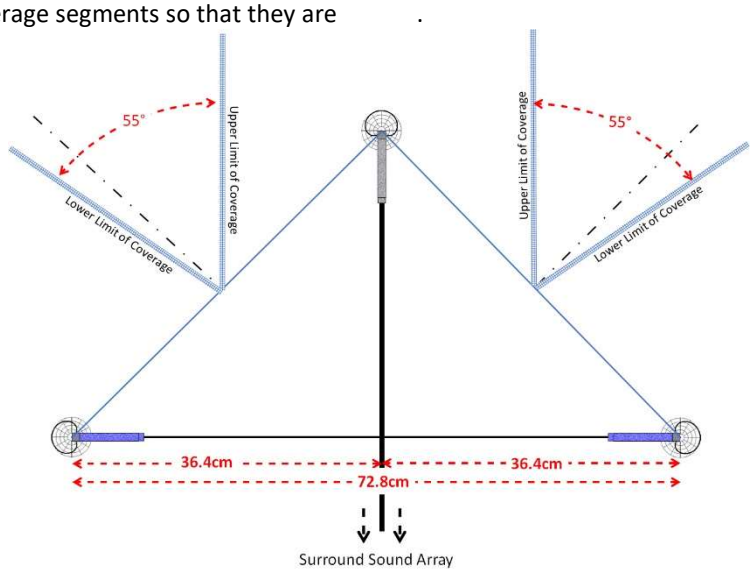


FIGURE 8 – UPPER LIMIT OF COVERAGE ALIGNED WITH ZENITH MICROPHONE AXIS.

1.3 The Complete All Cardioid Array

The final stage is just to join the height array to the zenith zone array to assemble to complete cardioid 3D audio array, as shown in Figure 10.

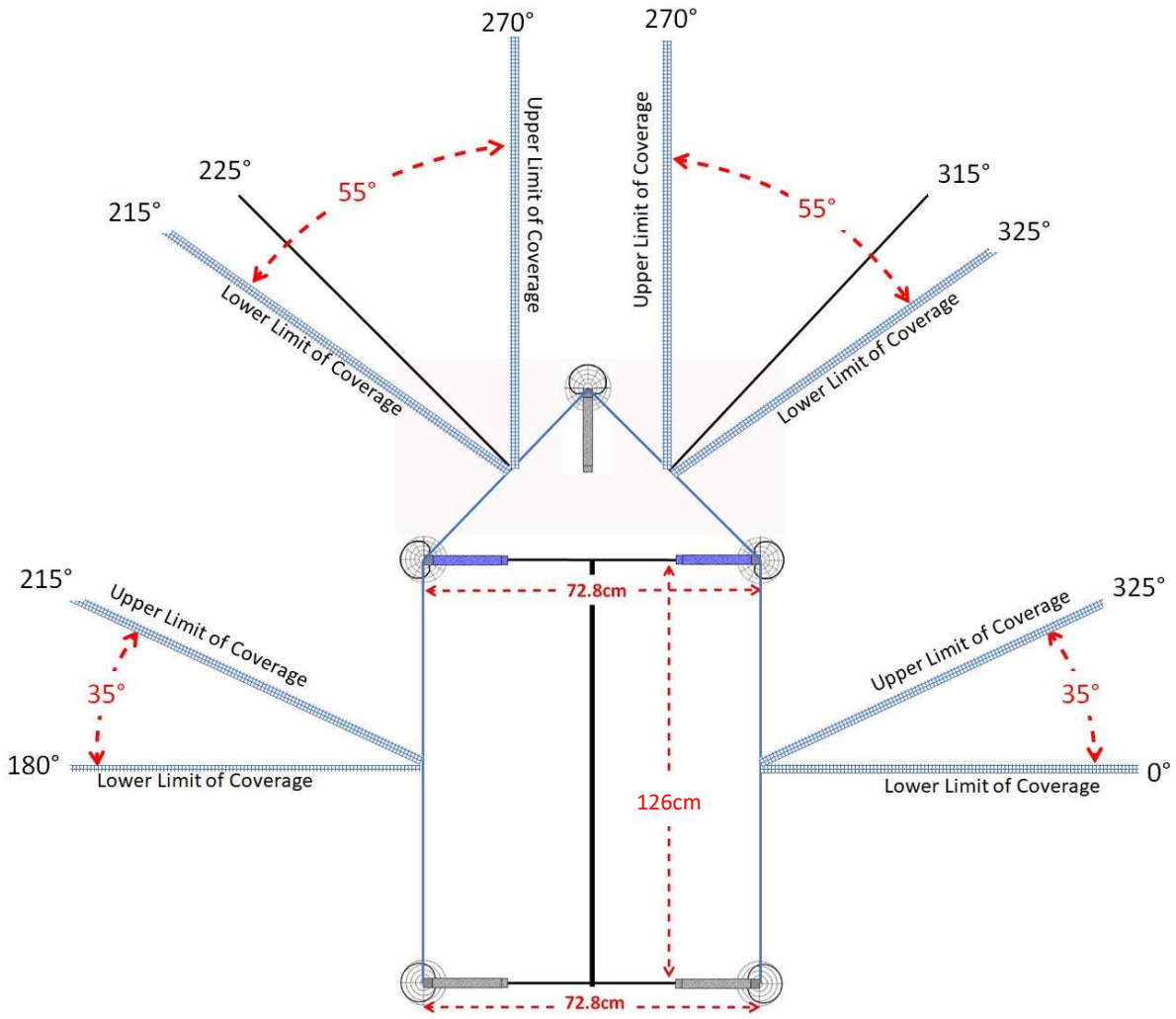


FIGURE 9 - TOTAL COVERAGE OF A COMPLETE 5CH CARDIOID / CARDIOID ARRAY WITH ETO BEING APPLIED TO EACH SEGMENT.

As you can imagine this is just one possible 3D Audio array design from a multitude of other solutions.

Here are three other solutions using various microphone directivity combinations:

- Figure 10a,10b & 10c – an all Hypocardioid Array
- Figure 11a, 11b a 11c
  A combined Hypocardioid/Supercardioid Array
- Figure 12a, 12b & 12c
  A combined Cardioid/Supercardioid Array

2 All Hypocardioid Array

This array has the advantage of better all round bass frequency response compared with an all Cardioid array.

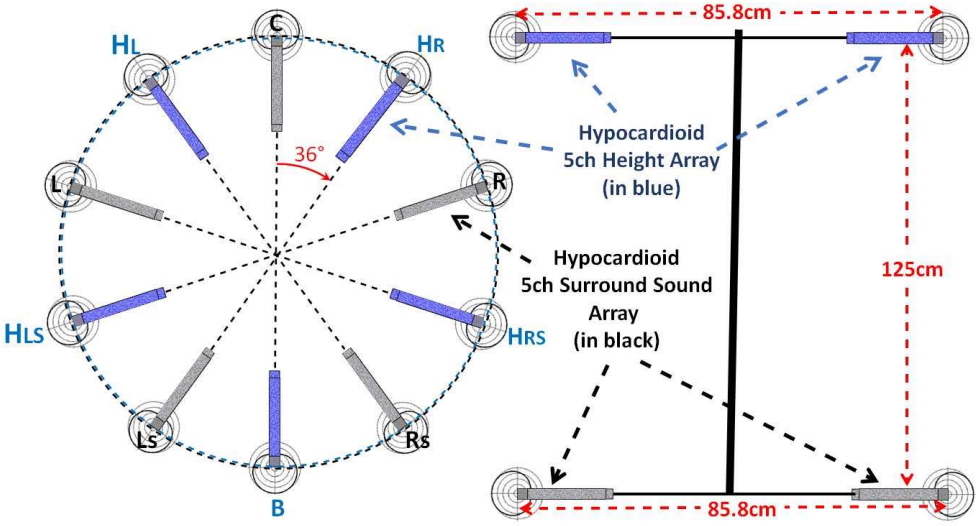


FIGURE 10A - 5 CHANNEL HYPOCARDIOID SURROUND SOUND ARRAY PLUS 5 CHANNEL HYPOCARDIOID HEIGHT ARRAY WITH ROTATION OF 36°

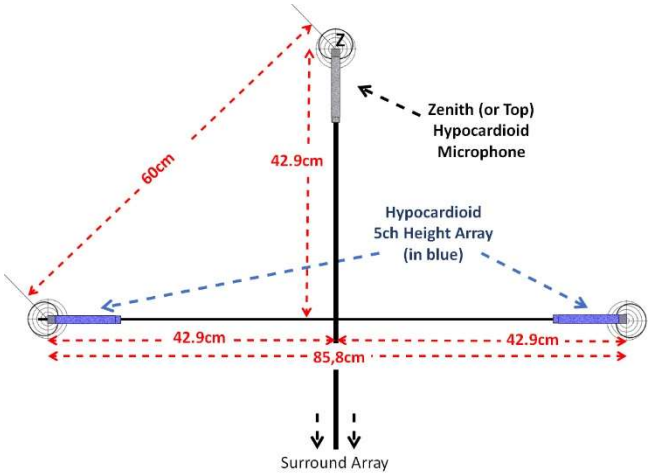


FIGURE 10B - ZENITH ZONE ARRAY USING HYPOCARDIOID MICROPHONES

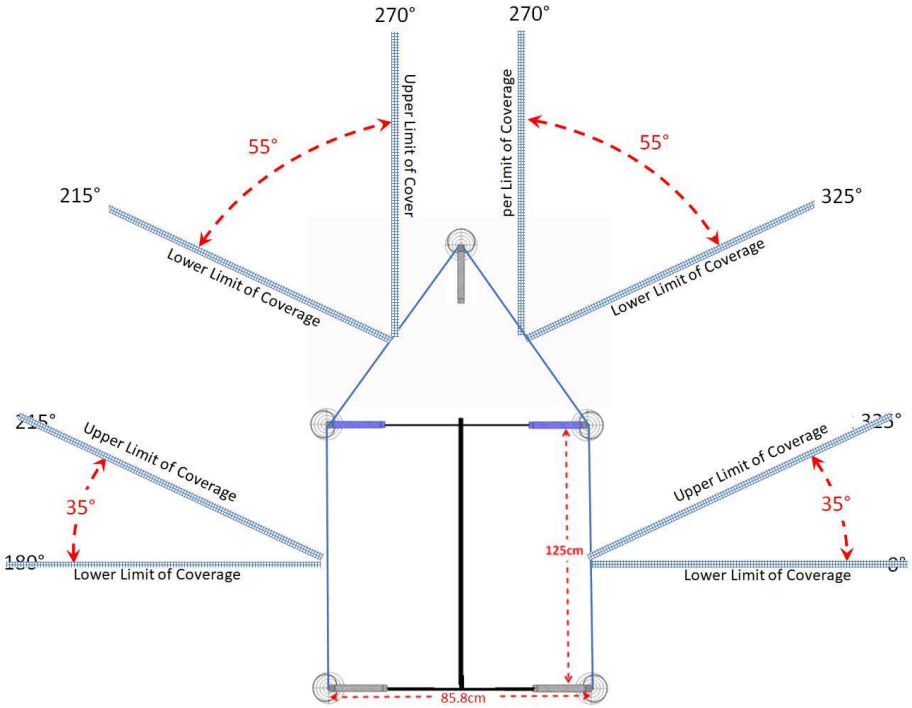


FIGURE 10C - CRITICAL LINKING BETWEEN ALL SEGMENTS OF A 3D HYPOCARDIOID ARRAY
17.5° OR 1MS ETO APPLIED TO HEIGHT ARRAY W.R.T. SURROUND ARRAY
17.5° OR 1MS ETO APPLIED TO ZENITH MICROPHONE W.R.T. HEIGHT ARRAY

4 Hypocardioid / Supercardioid Array

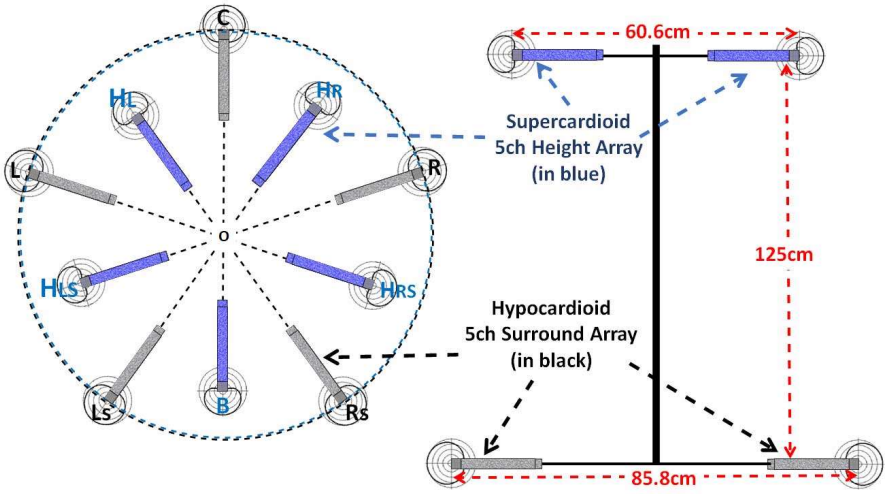


FIGURE 11A - 5 CHANNEL HYPOCARDIOID SURROUND SOUND ARRAY
PLUS 5 CHANNEL SUPERCARDIOID HEIGHT ARRAY WITH ROTATION OF 36°

As the wingspan of the Height Array is 60.6cm, there is already an initial Microphone Position Offset of 5.7°, as shown in Figure 12a

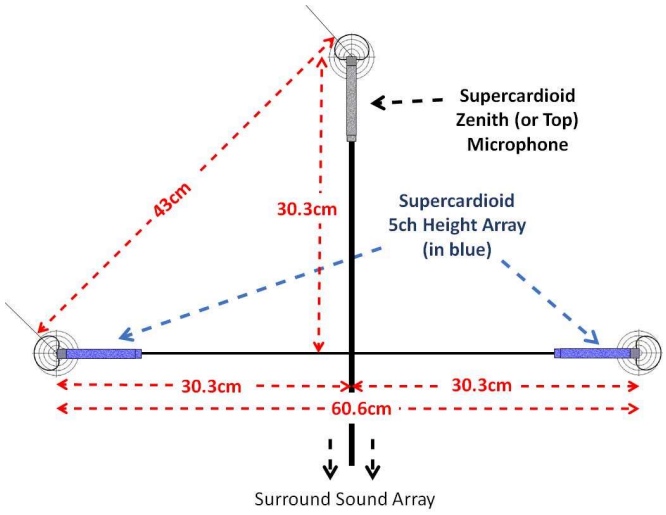


FIGURE 11B - ZENITH ZONE ARRAY USING SUPERCARDIOID MICROPHONES

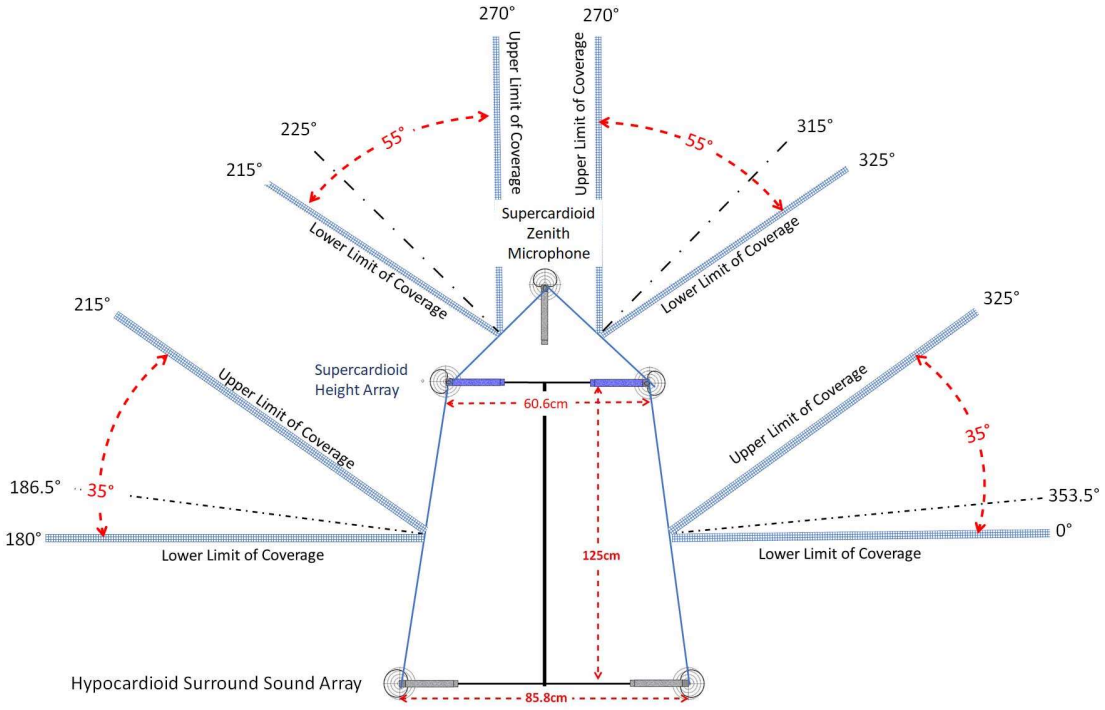


FIGURE 11C - TOTAL COVERAGE OF A COMPLETE HYPOCARDIOID / SUPERCARDIOID ARRAY WITH ETO OF 0.74MS (SIDE ELEVATION SEGMENTS) AND 1MS (ZENITH ZONE) BEING APPLIED

4 Cardioid / Supercardioid Array

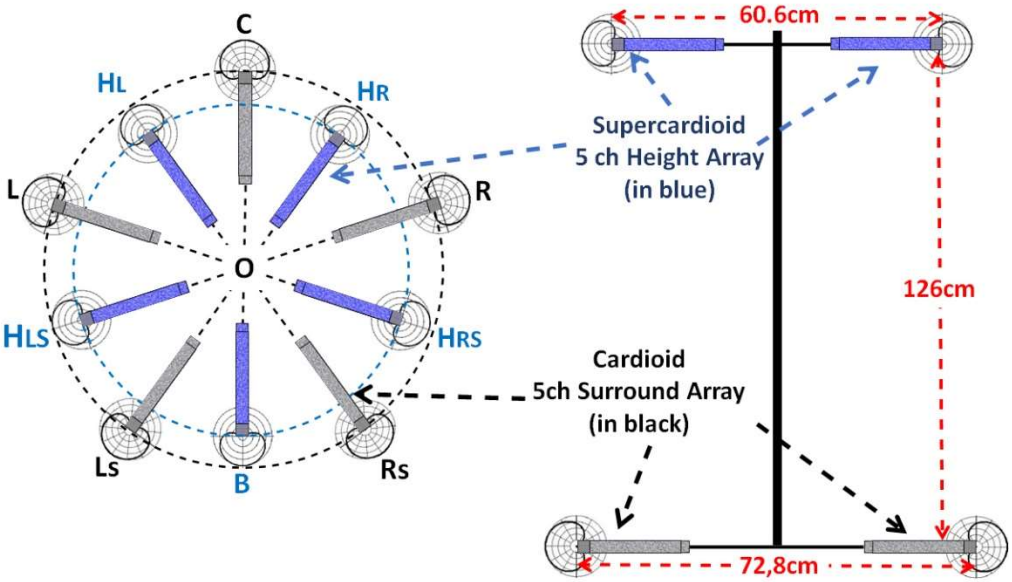


FIGURE 12A - 5 CHANNEL CARDIOID SURROUND SOUND ARRAY
PLUS 5 CHANNEL SUPERCARDIOID ELEVATION ARRAY WITH ROTATION OF 36°

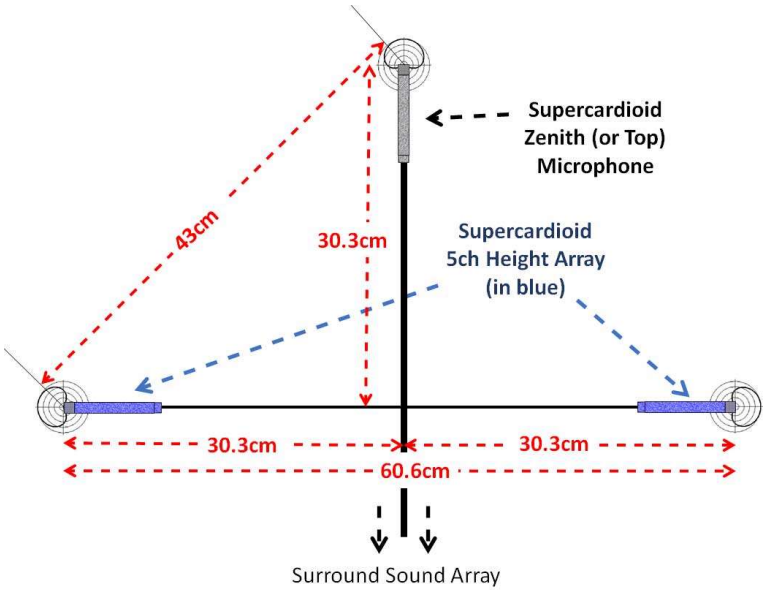


FIGURE 12B - ZENITH ZONE ARRAY USING SUPERCARDIOID MICROPHONES

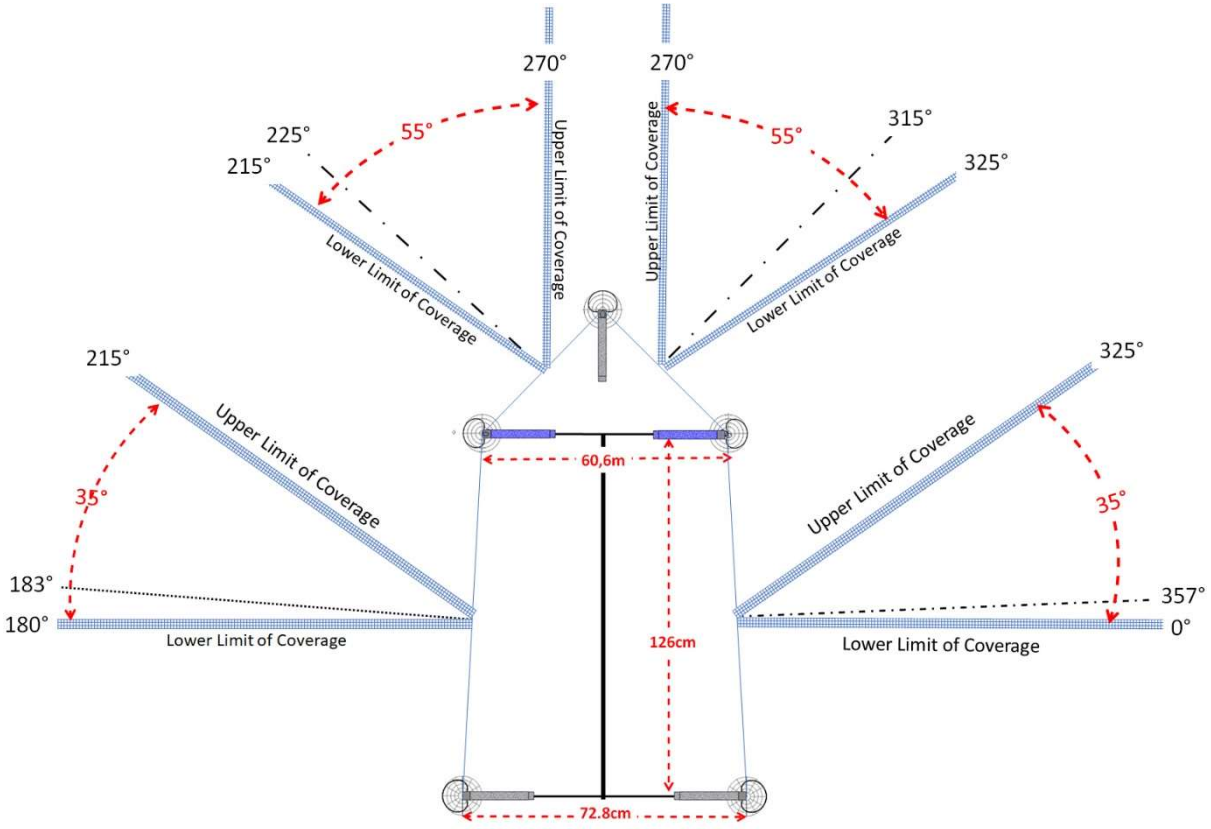


FIGURE 12C - TOTAL COVERAGE OF A COMPLETE CARDIOID / SUPERCARDIOID ARRAY WITH ETO OF 0.94MS (SIDE ELEVATION SEGMENTS) AND 1MS (ZENITH ZONE) BEING APPLIED

5 Conclusion

The array solutions presented in this paper are, of course, just half of the total system for complete 3D Audio recording and reproduction. Without equal attention being paid to the loudspeaker configuration we cannot expect satisfactory results.

Most of the precautions that need to be respected in the loudspeaker configuration have been presented in previous AES papers[1][2][3][4], and an on-line workshop was presented by the author at the 152nd AES Convention entitled "MMAD 3D Audio – An On-Line Experiment in Microphone Array Design Compatibility". During this presentation the advantages and disadvantages of various loudspeaker configurations were discussed. A video recording of this session should be made available on the AES Website, otherwise it can be downloaded from the author's own website: www.williamsmmad.com or www.mmad.info

This is the first time that I have been able to obtain permission to make available files from a recent recording of a concert of Trumpet and Organ by Pierre Kumor and Christian Bacheley, at the St Just Church in Arbois, in the Jura area of France. These files are the complete set of recordings made with a 'Comfort Array' i.e. a total of 19 individual files. A link to enable downloading of these files will be indicated on author's website.

During this on-line presentation, attendees were invited to download sets of files corresponding to their own individual loudspeaker configurations, so as to be able to experience the remarkable compatibility possibilities of this master recording array system.

6 References

[1] M. Williams, “Microphone Array Design for Localization with Elevation Cues 132nd AES Convention, Budapest (2012) - Preprint 8601

[2] M. Williams,” The Psychoacoustic Testing of the 3D Multiformat Microphone Array Design, and the Basic Isosceles Triangle Structure of the Array and the Loudspeaker Reproduction Configuration” 134th AES Convention, Rome (2013) - Preprint 8839

[3] M. Williams, Downward Compatibility Configurations when using a univalent 12 Channel 3D Microphone Array Design as a Master Recording Array “, 137th AES Convention, Los Angeles, USA (2014) - Preprint 9186

[4] M. Williams, “Microphone Array Design applied to Complete Hemispherical Sound Reproduction - from Integral 3D to Comfort 3D”, 140th AES Convention, Paris (2016) - Preprint 9569