Audio Engineering Society

Convention Paper 9569

Presented at the 140th Convention
2016 June 4–7 Paris, France

Microphone Array Design applied to Complete Hemispherical Sound Reproduction – from Integral 3D to Comfort 3D.

Michael Williams, “Sounds of Scotland”, Le Perreux sur Marne, France

mike@soundsscot.com

ABSTRACT

This paper describes the parameters that need to be taken into account in the design of a 13 channel microphone array recording system for reproduction also in a 13 loudspeaker hemispherical configuration. Both the microphone array and the loudspeaker array use 8 channels in the horizontal reference plane, 4 channels in the 45° elevation plane, and a Zenith channel at the top (at 90° elevation). This paper will also describe the various stages of advancement to complete 3D coverage (Integral 3D), and the logical development of this type of array to a new format proposition - the 16 channel ‘Comfort 3D format’.

Introduction

Previous papers[1,2,3,4] have described some basic 8 channel surround sound configurations, and also the Isosceles triangle microphone array structure and loudspeaker configuration that is necessary to obtain optimum reproduction definition and localization accuracy in the first 45° of elevation.

The development of each stage of height recording and reproduction will be described, both with the successes and failures. Each failure being just as much a contribution to the final design process as the successes. It must be said that this study makes use of a microphone array structure and natural sound sources. It is possible that the use of an artificial sound source, such as a Gaussien envelope noise burst, or time and/or level difference, using a time difference pan-pot and/or a level difference pan pot, will not produce the same localization results, or in some cases mediocre localization, or possibly no virtual localization at all.

1 Initial development

The first stage is obviously based on the horizontal, or surround sound, microphone array structure, which has been described in previous papers. But in the end the MAGIC 8 channel array surround sound structure[2] was adopted as usual in order to maintain a maximum number of configurations that were downwards compatible with existing formats for surround sound recording and reproduction. However the basic principles of design for height reproduction are applicable to most of the surround sound microphone array structures that make use of both Time Difference and Level Difference, to generate virtual surround sound localization in the horizontal plane.

The research into elevation reproduction can be divided into two segments:

- the first 45° of elevation,
- the last 45° of elevation, from +45° to +90°, 90° being immediately above the listener.

There is no reason why this could not be extended to 3 segments or more, but economy of channels is a major consideration, and localization with 2 segment coverage is perfectly satisfactory. In the reproduction of virtual sound it is not true to say that the more channels we have, the better is the 3D sound reproduction. This may be the case in object oriented mixing but not in microphone array recording and reproduction.

The research for localization within the first 45° of elevation can itself be divided into two phases:

- the research necessary to establish the relative contribution of Time Difference and Level Difference to localization in elevation
- the optimum position of loudspeakers in the reproduction configuration, necessary to reproduce the most accurate localization.

It was also considered necessary to maintain the same type of univalent structure in relation to the microphone array recording system, and the loudspeaker reproduction set up.

The addition of a Zenith microphone produces optimum results in the upper segment from 45° to 90°, but is not absolutely necessary if we are only concerned by the reverberation field around and above us, and not interested in the precise and accurate localization of direct sound in the complete hemisphere above the listener. But with the introduction of a Zenith channel and with close similarity between the recording and reproduction configurations, we will of course satisfy both 3D (or hemispherical) virtual localization in the reverberant environment, as well as the recording and reproduction of direct sound sources in the elevation (for example for birds or helicopters, or even large musical instruments with a strong height component such a church organ, etc. etc.).

1.1 The Psychoacoustic Parameters for Vertical Virtual Localization in the first 45° Segment of elevation

A previous paper[4] has described the choice of parameters that are necessary for good localization in the first 45° of elevation. We can consider that two basic characteristics must be considered:

- Height information is mainly captured by a Time Difference system
- An Isosceles array structure with respect to the horizontal microphones and the first elevation level of microphones will produce optimum results as shown for a recording array in Fig 1, and the loudspeaker configuration for reproduction in Figure 2

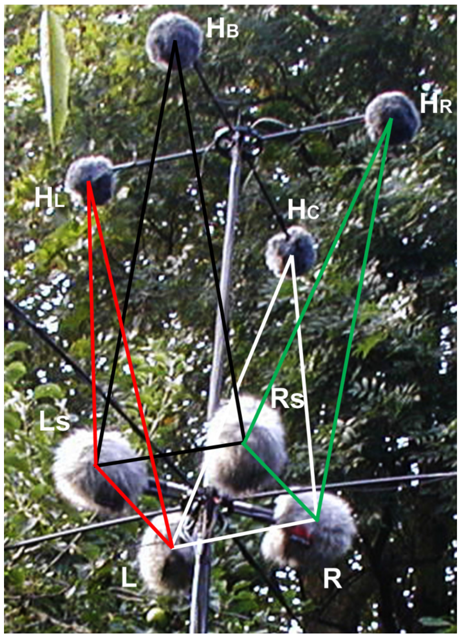

Figure 1 - Isosceles Triangle Structure of the Recording Array

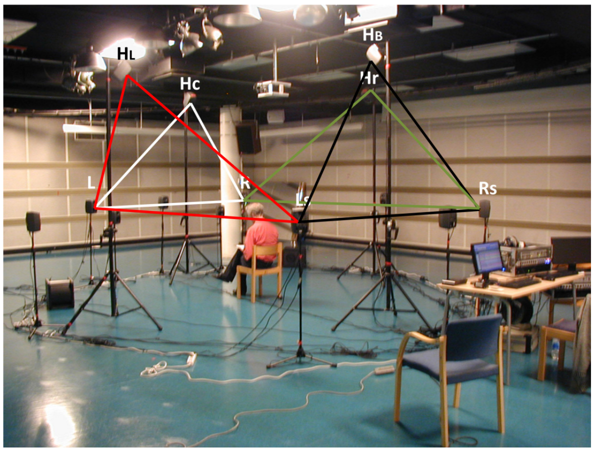

Figure 2 – Reproduction with Loudspeakers in Isosceles Triangle Structure.

Vertical Level Difference information was found to be almost without interest at all, and produced either no height localization at all, or a very mediocre result compared with Time Difference information[3] shown in Figures 3 & 4 – Localisation Performance w.r.t. Loudspeaker position.

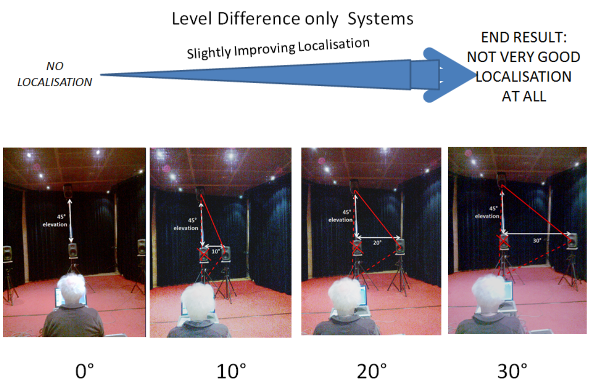

Figure 3 – Localisation performance with Level Difference only

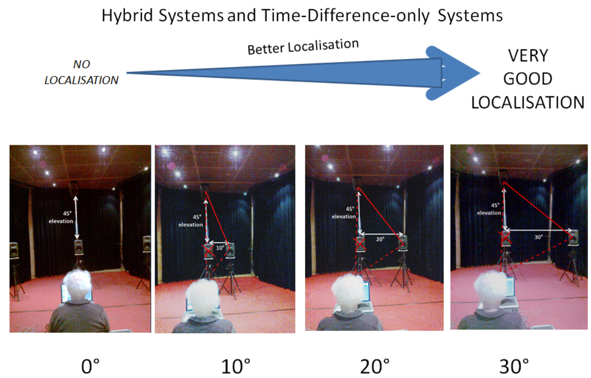

Figure 4 – Localisation Performance with Level Difference and time Difference.

1.2 The Psychoacoustic Parameters for Vertical Virtual Localization in the 45° Segment of elevation from +45° to +90°

Two basic options are available to experiment the quality of localization in this last segment:

- A microphone array structure using only Time Difference information – nicknamed "The Witches Hat Approach or Witches Hat Localization"

- A microphone structure that uses both Level Difference and Time difference information – nicknamed "The Top Hat Approach or Top Hat Localization".

It is hoped that the image that is projected by these two approaches is one of

- Time Difference only, in both the lower elevation and upper elevation segments for the Witches Hat.

- a combination of Time Difference in the first elevation segment, coupled with Time and Level Difference for the second upper segment of elevation.

1.2.1 The "Witches Hat" Localization system

"Witches Hat" Localization uses only Time Difference information for localization in both upper segments, from 0° to +45°, and from +45° to 90° as shown in Figure 5.

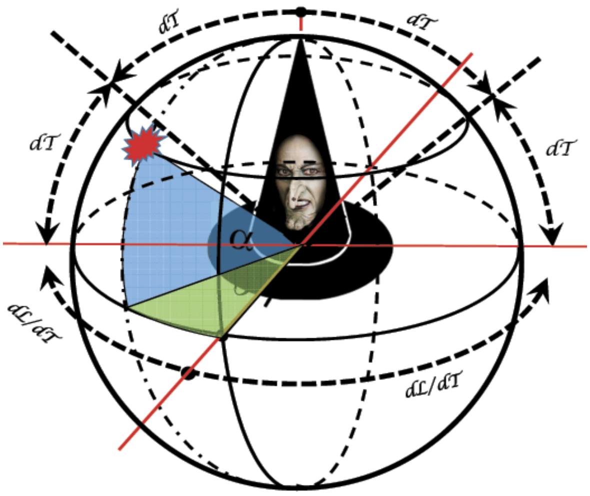

Figure 5 – Witch's Hat Localisation.

This is of course valid only for a microphone array structure using a horizontal array, a first elevation array and a top Zenith microphone. The height of the Witches Hat is meant to represent the distance that must separate the horizontal array and the first elevation array to produce localization in the first elevation segment (Time Difference parameter is predominant), and the distance that must separate the first elevation array from the top Zenith microphone (again Time Difference parameter is predominant). Figures 6 shows a practical 8+4+1 microphone array under test.

Figure 6 – 8+4+1 Array in Witch's Hat configuration

Both laboratory tests and on-location recording were carried out to study the characteristics of this type of microphone array structure. Recordings were made

- in a studio environment with a loudspeaker as a sound source and a Witches hat microphone array structure.
- on-location, with a the occasional helicopter passing overhead, and general environmental sound. These recordings were made on a small island, so there is also the sound of barges passing by, in the horizontal plane.
- On location during the Lossiemouth Airshow in Scotland with aeroplanes of all types passing around and above the microphone recording array.

It was found after multiple listening tests, that this design structure gives very good localization in the first elevation segment, but very poor results in the top elevation segment.

If the Zenith microphone is muted, then first elevation segment localization is excellent but the top localization follows reproduction in the square generated but the first elevation array structure. A helicopter flying above the array will give the impression of approaching the array system from the desired direction, but immediately it passes overhead it will follow one side of the square and then the other, and then will give a correct direction restitution as the helicopter flies away in the opposing direction. This effect goes completely unnoticed when recording music, as we are not really looking for perfect localization of the reverberant field above.

1.2.2 The "Top Hat" Localization system

Top Hat Localization uses Time Difference only information for localization in the first elevation segment from 0° to 45°. Whereas the second elevation segment uses a hybrid system, where both Time Difference and Level Difference contribute to localization as shown in Figure 7.

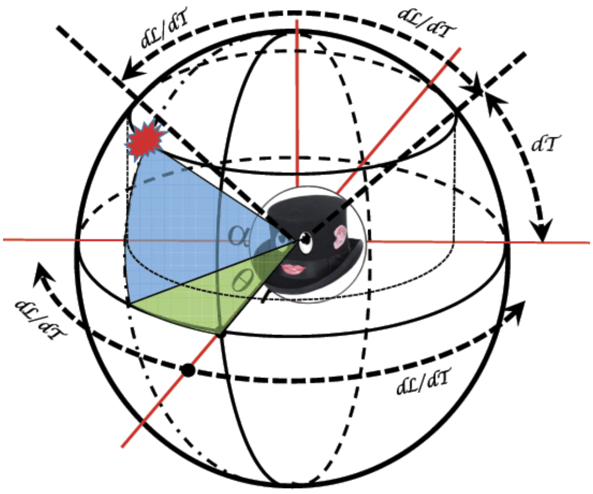

Figure 7 Top Hat Localisation

In other words the horizontal surround sound and the 2nd segment use the same parameters to obtain Localization from 45° to 90°, i.e. both Time Difference and Level Difference, as the horizontal surround sound layer. The two surfaces – the horizontal surround sound and the top elevation segment use the same parameters for localization, whereas the segment from 0 to 45° uses only Time Difference to obtain good localization. Figure 7 shows this type of array structure.

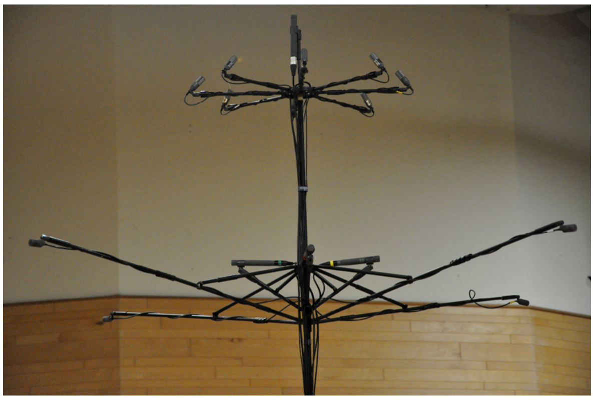

Figure 7 - a Top Hat 8+4+1 3D Microphone Array

This structure would seem logical if we consider that top elevation plane is also horizontal or more correctly, parallel to the horizontally plane – the brim of the hat and the top of the hat! Don't be deceived by the direction of the 1st elevation layer of the microphone in Figure 7, as the directivity axis is actually at 90° to the body of the microphones.

Figure 8 shows the test recording setup with the Blue Tooth loudspeaker as a sound source being moved around the microphone array. We found that a recording of Django Reinhardt produced the most reliable results. Other source material was tried but were not as satisfactory as the Django recordings.

Figure 8 – 360° Rail, Blue Tooth Loudspeaker Sound Source, and 13 channel Top Hat Array under test.

This Top Hat design of array structure gives excellent localization results in both segments of elevation.

We now have complete coverage, and quality localization, in the whole of the hemisphere above the microphone array – the name of Integral 3D has been given to this type of coverage. The laboratory tests indicate that the precision of localization in the complete hemisphere around the array and at various angles of elevation is also completely satisfactory.

The exact quantitive verification of this localization function during recording and reproduction presents a number of difficulties. This is due to the necessity to measure exactly the physical sound source position, but even more so, to measure the perceived localization position. An experiment to determine this correspondence function between the two situations is under consideration.

2.1 Integral 3D – Compatibility Tests

During tests carried out at the Chalmers University in Gothenberg in 2013, it was established that surround and elevation localization was completely acceptable if the horizontal part of the array structure was reduced to only 4 channels (a 4+4 array). On the other hand if compatibility with all other surround and 3D systems was required then the full 12 channel array structure (an 8+4 array) was necessary.

However this 4 channel horizontal array structure – leading to a 4+4+1 3D array calls for more analysis.

We all know that a quadraphonic surround sound array produces a more or less acceptable reproduction for surround sound, but it can be critized because the linearity of reproduction in each segment is far from satisfactory because each segment is quite wide at 90°. This is improved considerably by the use of a 5 channel surround sound array with 72°segments, as well as, of course, the 6 channel surround array. And nobody can deny the superiority of the 7 channel Blu-ray configuration. And if we add in a back channel to produce an eight channel surround array then the result is almost perfect, that is for surround sound.

In experiments during GOART Gothenberg project we were able to study the compatibility characteristics of a 4+4 microphone 3D array, with the 8 channel surround reproduction configuration. It was found that excellent surround sound could be obtained if the upper square of 4 microphone were positioned at 45° to the lower 4 channel array, and the upper 4 channel array was reproduced by routing each microphone to the lower 8 channel loudspeaker array.

In other words, if we consider that the quad array is made up of Left, Right and Ls and Rs, then:

- Height Centre (at 0° azimut and 45° elevation) is routed to the Centre of the 8 channel loudspeaker surround configuration
- The Height back (at 180° azimut and 45° elevation) is routed to the back of the 8 Channel loudspeaker surround configuration
- The Height Left Median channel (at 90° azimuth and 45° elevation) is routed to the Right Median loudspeaker of the 8 channel loudspeaker configuration
- The Height Left Median (at 270° azimuth and 45° elevation) is routed to the Right Median loudspeaker of the 8 channel loudspeaker configuration.

The compass angles rotate in a clockwise direction. This is very difficult to put into words but very simple as a configuration structure.

Figure 9 shows the 4+4 reproduction configuration whereas Figure 10 shows the projection of the upper channel signals onto the surround sound loudspeaker configuration.

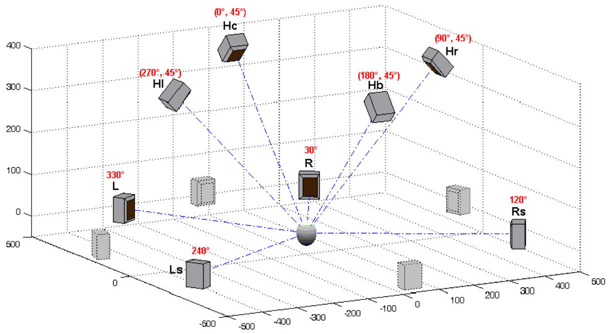

Figure 9 - 4 + 4 channel loudspeaker configuration

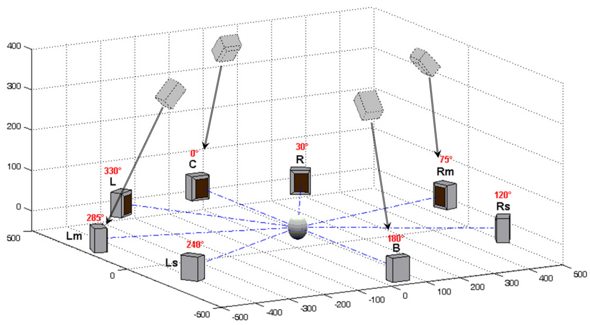

Figure 10 – '4 + 4' Vertical Projection to an 8 Channel Surround Sound Reproduction

In Figure 9 the Left, Right, Ls, Rs, Hc, Hl, Hr and Hb loudspeakers are active, whereas the Centre, Lm, Rm and Back loudspeakers are muted.

In Figure 10 the 4 + 4 3D configuration the Centre, Left, Right, Lm, Rm, Ls, Rs and Back, Loudspeakers are all active. Signals from Hc, Hl, Hr and Hb are folded into the Centre, Rm, Lm and Back loudspeakers of the surround sound configuration.

The extraordinary thing is that the 3D reproduction is quite satisfactory, and the surround sound configuration is almost perfect. The surround sound configuration is almost indistinguishable from the natural 8 channel surround sound recording and reproduction configuration.

This is the case whether we work with a 4+4 array, a 5+5 array, a 6+6 array, etc., as long as the microphones in elevation have an isosceles triangle structure in relation to the horizontal array structure.

Most people nowadays have opted for a 5 channel structure for the horizontal array. It would seem logical to adopt a 5+5 3D array structure to help compatibility from one context to another.

The addition of a Zenith microphone to the previous microphone array structure does not change the compatibility characteristics, except of course that the Zenith microphone is eliminated (muted) from the projection that is made onto the horizontal array structure.

The full dual set of quintuple arrays plus the Zenith channel becomes a 5+5+1 integral 3D. This type of array will therefore be compatible with any of the lower order 3D and surround sound formats.

2.2 From Integral 3D to Comfort 3D.
This development comes from one remark made during the listening tests at Galaxy Studios in 2015. One of the listening panel said, quite spontaneously, that reproduction of the sound source at 15° and 30° degrees of elevation seemed more comfortable than reproduction in the horizontal plane. This remark deserves very much deeper analysis.

Mono sound recording restricts the sound reproduction to one loudspeaker. And we do not look for any space around that loudspeaker.

However in Stereo Sound reproduction we are so taken by the stereophonic spread of sound that we do not easily realize that sound that comes from outside the Stereophonic Recording Angle (SRA) is still being reproduced as mono sound on either the left or right loudspeakers – in most cases reverberation from all around the stereo pair. Sound from above and below the stereo array is either spread out between the two loudspeakers, or again reproduced as mono sound on the left and right loudspeakers. When we move to surround sound, then again the horizontal spread of sound is very satisfactory, but does not remove the fact that sound from above and below the array is condensed onto the horizontal spread of sound.

When we move on to the so called 3D reproduction, we again are so taken with the 3D reproduction spread or more correctly the hemispherical reproduction of sound, that we do not perceive easily the fact that the lower sound field is again being projected onto the horizontal surround sound spread.

The remark that sound is more comfortable at 15° or 30° degrees elevation then takes on a new meaning. It is simply that the indirect sound architecture is being distributed both above and <u>below</u> the direct sound source. The sound source is generally distributed over 10° or15° above and below the horizontal plane of recording. Therefore the recording of the direct sound which is usually in the horizontal plane of the array system, does not have a component below the horizontal plane. It is obvious that one cannot record and then reproduce the direct sound at an angle of around 10° or 15° of elevation as we are too used to perceiving this sound along the horizontal plane.

However why not introduce an array that will record and reproduce sound both above and below the horizontal plane. We do not need to consider the sound from directly under the array as there will be no significant sound from that direction, so the 'Voice of the Devil' channel is not required.

And so we have the extension to 'Integral 3D', that now becomes 'Comfort 3D', the lower segment under the horizontal plane can now record and reproduce the full architecture of the sound source, to the left, to the right, to the top, AND to the bottom. We then obtain 'comfortable' sound reproduction – the sound is reproduced with its full architecture, or the full extent of the surface of acoustic radiation. The lower elevation array and the upper elevation array are, of course, both oriented at 45° to the horizontal plane array.

The story is not completely finished, as we still have to rely on the brain to reconstruct depth from various acoustic clues, like direct-to-reverberant sound ratio, or the expected timbre variations with distance, etc. But at least we have at least 75% of the 3D sound field covered.

There are practically no reflections in the lower quarter sphere when recording music and natural on-location sound, and until nature endows us with wings, I don't think that this last 25% of the sphere even needs to be considered – but who knows!

1. ACKNOWLEDGEMENTS

I would like to thank again Wilfred van Baelen, (CEO of Galaxy Studios and Auro3D) for allowing me to carry out all these experiments in their magnificent studio complex, and participating in the listening tests. Also Thanks to Joris Van den Eynden (Technical Engineer at Auro Technologies) for all his help in setting up these recording and reproduction systems, and also for his participation in the listening tests.

2. REFERENCES

(1) 1999 - Multichannel Microphone Array Design by Michael Williams and Guillaume Le Dû
108th AES Convention in Paris - Preprint 5157

(2) 2007 - "Magic Arrays - Multichannel Microphone Array Design applied to Multi-format Compatibility" by Michael Williams
122nd AES Convention in Vienna - Preprint 7057

(3) 2012 - "Microphone Array Design for Localisation with Elevation Cues"
by Michael Williams, 132nd AES Convention in Budapest - Preprint 8601

(4) 2013 – "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 Convention 2013 May 4–7 Rome, Italy
Preprint 8839

(4) 2014 - Downward Compatibility Configurations when using a univalent 12 Channel 3D Microphone Array Design as a Master Recording Array" by Michael Williams
137th AES Convention in Los Angeles, USA - Preprint 9186

(5) 2013 – "Microphone Arrays for Stereo and Multichannel Sound Recording Vol 2" Pages 243 to 246 - The GOART Project.