



Audio Engineering Society

Convention e-Brief 459

Presented at the 145th Convention
2018 October 17–20, New York, NY, USA

This Engineering Brief was selected on the basis of a submitted synopsis. The author is solely responsible for its presentation, and the AES takes no responsibility for its contents. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Audio Engineering Society.

Streamlined 3D sound design: the capture and composition of a sound field

Wieslaw Woszczyk¹ and Paul Geluso²

¹ Graduate Program in Sound Recording and Centre for Interdisciplinary Research in Music Media and Technology,
McGill University, Montreal, Québec, Canada H3A 1E3

² Dept. of Music & Performing Arts Professions
New York University, New York, NY 10012, USA

Correspondence should be addressed to Wieslaw Woszczyk (Wieslaw.Woszczyk@mcgill.ca)

ABSTRACT

A pragmatic approach to 3D sound design is described that employs a minimum number of sound fields captured with tetrahedral microphones. The captured sound fields, each extended in a horizontal and vertical dimension, are combined to provide the essential segments of the entire 360° sound design. Supplemental single-capsule microphones are used as needed for balancing of spaciousness and clarity. A compatible scaling of sound design from 3D to 2D can be easily accomplished without distortion of timbre or space.

1 Introduction

This paper is a follow up to the first part of the investigation, which was presented in Tokyo during August 7-9 at the AES International Conference on Spatial Reproduction - Aesthetics and Science [1]. The goal of both publications is to encourage experimentation in and to augment the understanding of 3D capture, mixing, and sound design by sharing experimental results and knowledge about relevant tools and techniques that may accelerate personal entry into immersive audio production. In particular, the focus is placed on pragmatic approaches, using affordable tools to enable a wider group of practitioners to successfully enter the world of 3D sound.

Music productions in 3D formats such as 22.2 or Dolby Atmos are known to require 40-60 microphones and more when close (spot) microphones supplement the 3D capture of the room acoustics and of the overall musical staging. Even if some of these microphones will not be used, the required infrastructure is huge and available only in

dedicated fixed installations. Placing so many individual microphones presents logistical and physical challenges, is time consuming and costly in location recordings. It is also difficult to assess the quality of sound capture in each of the many microphones, and to ascertain the degree of functional consonant interaction between them without a suitable 3D monitoring environment, which typically is not available in a location recording.

The technique we investigated discussed here, utilizes a small set of tetrahedral microphones, which allow for a strategic composition of captured 3D sound fields. Each component sound field provides a critical structural support of the entire 3D spatial image, and plays a primary role in forming the overall immersive sound design. Because the tetrahedral capsule configuration is well defined and calibrated, monitoring of sound in 3D is not necessary during the recording, and headphones are sufficient to judge the aspects of distance to the sources, and the spacing between the tetrahedral microphones. Subsequently, captured sound fields

become fully integrated during mixing within a suitable 3D monitoring environment.

2 The Concept

In a traditional approach to recording with multiple microphones, monophonic transducers are distributed within a space to bring focus to sources and to the acoustic environment. This is essentially a discrete spatial sampling approach, where spatial continuum is fragmented, lending a highlighted structure to the sound image. If the microphones are widely spaced, and the transducers are of a directional type, their signals are largely independent from each other. Even if at the lowest frequencies the flow of propagating waves is captured in a coherent manner, at mid and upper frequencies the signals are incoherent. To regain source presence, individual microphones are often moved closer to sources and their signals are panned to desired directions. This fragmented approach cumulatively lends a diffusing effect to the ambience and reduces the size of source image due to the lack of stereophonic continuity characterized by the overlapping capture of sources.

If the microphones are omnidirectional and closer to each other, their signals share a common representation of the sound field, with strong relative temporal cues describing the space. Typically, in such case, microphones' spatial relationship matches the layout of 3D playback loudspeakers. Morten Lindberg achieved beautiful results with this technique [2].

The present authors advocate capturing the continuum of sound field rather than its fragments by using tetrahedral microphone arrays, each made of four coincident directional transducers. In this approach, a total sound field is composed of regional sound fields that are captured within the ensemble, and within boundaries of the space, rather than derived from independent mono signals. This way, information captured in the spatial field has an inherent three-dimensional structure that lends strong 3D imaging to the sound design. Importantly, capturing continuous sound fields preserves the natural three-dimensional fabric of sound providing human listeners with a multitude of natural spatial

cues. This, compared to the assortment of fragments, produces a greater sense of realism and immersion.

The tetrasonic "soundfield" microphone is not new since Michael Gerzon [3],[4],[5] built its first prototype in 1970s for use in periphonic recording and reproduction [6]. In Gerzon's approach, capturing a single sound field was to deliver a complete authentic representation of a musical performance. However, with the single sound field captured, Gerzon could not solve the problem of balancing musical sources and perspectives. This was the same dilemma as encountered in early electrical recordings when only a single mono microphone was used. Gerzon acknowledged that "a profound philosophical problem with tetrahedral recording is where to put the microphones."

3 The Method

The method proposed here uses two or more sound fields that are captured and integrated into a single 3D sound design. So, instead of numerous monophonic fragments, the sound designer can integrate continuous sound fields, each containing precise three-dimensional spatial relationships detected by the tetrahedral microphone array. The sound designer may move these sound fields around in the listening space within a 3D loudspeaker array or in headphones, and adjust the horizontal and vertical extent of each to suit the intended sound design purpose. Because the tetrahedral capsules have a precise coincident relationship, sound fields can be repositioned in the playback space in a straightforward manner, and without negative side effects such as unstable or distorted imaging and sound coloration. This would not be the case if monophonic fragments of the sound field separated in distance were moved across the reproduced sound field triggering numerous interferences.

At the minimum, two tetrahedral periphonic sound fields need to be captured, one for the left side, and one for the right side of the sound design. Thus, the left-side continuum transitions into the right-side continuum through a common overlapping centre. The three-dimensional continuum, through the left, through the centre, and through the right, is extended in elevation and azimuth, like a sheet of canvas

wrapped around the listener. The overlapping sound fields give the impression of a complete immersive shell of sound. Each captured 3D region integrates the direct and the ambient information from a single acoustically coherent viewpoint. The direct and indirect sound cues reinforce the perceived unity of a cohesive spatial image.

The proposed left-to-right orientation of 3D sound fields is optimized for the human listener, whose auditory system compares spatial cues received in the left and the right ear. Perceptual cues presented by loudspeakers simultaneously along vertical and horizontal axes create a highly listenable, natural sounding spatial image. The rich redundant directional cues articulate perceptual dimensionality of sound design.

Tetrahedral 4-capsule microphone is the simplest transducer capable of extracting three-dimensional acoustical information from the sound field. It is, therefore, the only microphone able to capture spatial continuum, the most useful information communicating realism in recording. In contrast, a single-transducer monophonic microphone provides no spatial information.

While first-order directivity of tetrahedral capsules is insufficient to provide narrow angular selectivity of sources captured across the entire 360° sphere, it is quite sufficient to spatially characterize a zone of the overall sound field. At the minimum, two zones (zone LEFT and zone RIGHT) should be created, each having a vertical and a horizontal extent, in order to compose a 3D sound design giving highly immersive listening experience. Within each zone, sufficient spatial differentiation is possible using attentive microphone placement. Finally, once the zones are combined, spatial selectivity may be augmented as needed, even to the full sphere.

The balance of relationships captured between source and environment is adjusted with microphone placement, and a combination of the component balances gives the overall sound field impression. This is not unlike balancing individual single capsule microphones, except that instead of using zero-dimensional chunks, we now combine continuous three-dimensional volumes.

4 The Prague Sessions

We now present three examples of practical implementation of the proposed streamlined method of 3D sound design. Previous report [5] described experiments conducted at McGill University with graduate students using up to five tetrahedral microphones in 3D recording, and then mixing the captured components in a dedicated control room designed for 22.2, Dolby Atmos, and 9.1 sound monitoring.

The goal of this publication is to describe new experiments conducted in Prague by the authors with a group of NYU students during three 9.1 recording and mixing sessions in June 2018. In each of these productions, only two tetrahedral microphones were used, augmented by standard single-capsule transducers. The analysis of these sessions, with the explanation of practical approaches and results, illustrates the value of this pragmatic methodology for achieving successful 3D sound design in music recorded for the 9.1 format.

The three sessions included the recording of Woodwind Quintet at the Studios of the Czech National Symphonic Orchestra (scoring studio at CNSO), a mixed electronic/acoustic Didgeridoo ensemble recorded in Faust Studio, and a vocal sextet recorded in a church. The diverse musical styles and acoustical environments allowed for useful testing and verification of the proposed technique.

Woodwind Quintet Session

In a scoring stage spacious enough for a medium size orchestra, a woodwind ensemble was placed at a 2/3 distance from the control room window to the back wall. The musicians were seated in a semicircle from Left to Right: flute, oboe, French horn, bassoon, clarinet.

Two tetrahedral microphones were placed in A-B to capture a 3D image of the ensemble in its interior sound field. Further behind, a Decca Tree of 3 omnis (AKG 414) was set augmented with left and right surround wide-cardioids (also AKG 414), at 3 meters above the floor. Five DPA 4011 microphones

were placed 1.5 meters above each of the AKGs, pointed up to capture the acoustics of the elevation.

In listening to stereo on B&W801 monitors, the image heard was detached from the speakers, and open above and to the sides. There was liveness from the early reflections, and the image felt very 3D, spacious with only two loudspeakers. Listening in the control room felt as if there was no window and wall separating us from the musicians. The tetrahedral A-B pair gave strong sense of presence to the ensemble. The photos in Figures 1 and 2 try to show the relative placement of the microphones.



Figure 1. View from the Woodwind Quintet side towards the control room in the CNSA Studio. Left and Right tetrahedral microphones are circled in red.



Figure 2. View from the control room side. Decca Tree in the CNSA Studio employed in the Woodwind Quintet session.

Faust Studio Session – improvisation

In a pop studio two musicians used a variety of electronic and acoustic instruments to record long soundscapes in two improvisations. We had a grand piano and Nord Stage 2 keyboard synth performance, active mixing of dozens of guitar pedals in parallel treating the vocal microphone, a low drone of the didgeridoo, and percussive or shimmering sounds of numerous resonating metallic plates attached to two wooden boxes.

A tetrahedral microphone was installed inside each of the wooden boxes to capture internal sounds and reverberation from the ringing plates. All eight outputs of the tetrahedral mics were later routed in the 9.1 listening room to create tall Zone-A (on the left) and tall Zone-B (on the right), immersing the listener with full width and height. Two DPA 4011 were also used outside the boxes, 1 meter from the plates, to capture the high frequency transient detail.

The piano sound from an ORTF pair of Sennheiser Twins and from two room microphones was augmented with artificial reverberation. This blended well with complex lingering sounds and effects from the guitar pedals. Figure 3 photo shows some of the microphones used, with circles around the tetrahedral microphones, and the didgeridoo drone pipe with two microphones (a condenser and a dynamic on the red table).



Figure 3. Two tetrahedral microphones (in red circles) each set up inside a separate wooden box that amplifies the ringing of steel resonator-plates.

Vocal sextet session in the church

In the highly reverberant church, six singers performed acapella standing on the stairs near the altar. The musicians were in a line slightly curved and splayed vertically partly on the steps of the altar, and partly on the floor. See the photo in Figure 4.

Two tetrahedral microphones were set up in A-B configuration designating the Left and Right zones of the sound field near the ensemble. The microphones were capturing the inner sound field of the ensemble in vertical and horizontal dimensions. As musicians moved around, the motion was beautifully articulated in the captured sound. Two inner capsules of the A-B equally captured the center musician.

Further behind, more than twice the distance of the tetraphonic microphones to the vocalists was an M5 inline array of 5 Sennheiser MKH800 microphones set to wide-cardioid. It was capturing the entire ensemble, spread about the width of the 6 singers, 4 meters wide, precisely aligned to be a line array.

Two pairs of DPA 4011 were pointing up (front and rear slightly) on each side of the main aisle, near the ensemble, each about 3 - 4 meters away from the center line, and about 4 meters high, separated by a bar 1.2 meters in length, the distance between front and back microphones. The pairs were used as the 4 elevation channels of 9.1, containing mostly early reflections while avoiding the capture of direct sound and of rear wall reflections in the long dimension of the church. The sound was excellent, diffused, clear, not heavy, providing a very satisfying enveloping sensation.

For the capture of reverberation in a rear section of the church, two systems were set up near the walls. One system in each side of the church provided a strong intense acoustic image of the room, with some resonances, and repeats, characterizing the church. Two pairs of DPA 4011 pointed to the sidewalls, and to front-back walls, while up-down aiming Schoeps MK8 bi-directionals were used to matrix the signals into upper front and lower front, and rear up and rear down directions. The level of this component in the mix was minimal.

The live location recording was monitored on headphones. The final 9.1 sound design was based on a strong inner image of the ensemble from the tetraphonic A-B pair, enhanced by early reflections from two pairs of DPA near the ensemble, plus the frontal presentation of M5 array. The ensemble image was vivid and compelling within rich ambience they commanded as singers.



Figure 4. Vocal sextet ensemble in a church with two tetrahedral microphones located within the blue circles.

5 Routing of capsule signals to the speaker channels

In the presented examples of sound field capture using A-B configuration, the routing of tetrahedral microphone capsules was the same. While variations of course should be explored, the routing we used worked very well accomplishing a strong immersive 9.1 design.

Microphones' A-format output signals were used **directly**, without converting them into B-format first-order ambisonic signals, and without matrixing into subsequent channel-based signals. Each tetrahedral microphone has 4 cardioid capsules that have maximum on-axis sensitivities in four directions:

Capsule 1- **Front Left up**, Capsule 2- **Front Right down**; Capsule 3- **Back Left down**, Capsule 4- **Back Right up**).

The idea is to use the Capsules 2 and 3 as the left and the right side of the ensemble section (either the left or right section of the A-B), Capsule 1 aimed above the section, and Capsule 4 aimed away and across from the singers into the ceiling. The following routing was used:

Left Side - Tetrahedral Microphone - A

Capsule 3 (extreme left – panned to in-between surround left and front left), Capsule 2 (inside left – panned in between center and left), Capsule 1 facing above the singers on the left (left-front top), Capsule 4 pointing away and across from the three singers and up – panned to right-rear top surround.

Right Side - Tetrahedral Microphone - B

Capsule 2 (extreme right – panned to in-between surround right and front right, Capsule 2 (inside right – panned to in-between center and front right), Capsule 1 (facing above the singers on the right – panned to front right top), and Capsule 4 (facing away and up from the singers - to left top surround).

Figure 5 shows the capsules of the tetrahedral microphone with three capsules (C1, C2, C3) aiming at the source and one capsule away from the source (C4). The recommended routing to loudspeakers is indicated above. Figure 6 shows alternative aiming of tetrahedral capsules, with two of the three front capsules aiming higher (routed to elevation channels), and one capsule aiming lower (routed to one of the main horizontal channels).

We recommend trying the recommended routing and exploring alternatives in the search for what can best realize your 3D sound design concept.



Figure 5. The capsules of the tetrahedral microphone with three capsules (C1, C2, C3) aiming at the source and one capsule away from the source (C4).



Figure 6. Aiming of the Left tetrahedral microphone's capsules of an AB pair of tetrahedral microphones. Two of the three front capsules are aiming higher (routed to elevation channels), and one capsule is aiming lower (routed to one of the main horizontal channels).

The Right tetrahedral microphone will follow a similar plan but mirror-inverted.

6 Discussion

The useful and remarkable feature of the proposed method is its ability to magnify the presence, size, and perceived closeness of sources. Also noticeable is the corresponding fullness of timbres and ease of listening, no doubt due to the high variety and density of spatial and temporal cues reaching the listeners.

The best strategy for microphone placement is in the interior of the space immediately surrounding the sound sources. This is where directions of sound arrivals are diverse and well defined, and where spatially distributed sources counteract the limited directional selectivity of tetrahedral microphone. By capturing sound close to the source in 3D, extraordinary perceptual separation between direct and reflected sounds can be achieved.

The panning of the four capsule signals allows one to successfully widen and elevate the image. The engineer should route the three outputs of left-side tetrahedral microphone to the left side, and route the fourth output to the opposite side. There must always be one signal component crossing the median plane of the listener, and it should be the capsule that aims away from the nearest intended source. The three capsules on the Left side should define the direct source in width and height as a triangle. The same principle applies to the tetrahedral microphone on the opposite, Right side of the ensemble.

The aiming of microphones should be done without the protective grid so that the capsules are clearly visible and their aim at the source can be confirmed. The removal of the protection grid improves the sound, as the grid is not acoustically transparent.

All tetraphonic microphone signals are used directly without matrixing, in the A-format. The four capsules aim at appropriate angles of the continuous sound field to record a particular 360° sound balance at that location using a minimum of resources (capsules). At each location, tetrahedral microphone records a specific spatial, spectral and temporal condition of that sound field as fully as possible. A unique robust 3D image of each sound field allows it to retain perceived identity even when a number of

different sound fields are presented simultaneously, mixed together in a 3D reproduction space. Balancing is straightforward because sound field imaging is not compromised by masking or interference from level or panning changes. But, it is essential to group and adjust all 4 tetrahedral capsule signals together and consider them as one. In the old approach boosting a mono source was immediately noticeable because its spatial image narrowed. The more it was amplified, the smaller it became.

7 Conclusions

The technique described here is grounded in spatial auditory perception and utilizes basic 3D tools (tetrahedral microphones) combined with standard stereophonic microphone techniques, such as A-B. This allows any engineer to make a comfortable transition from stereo or surround recording into 3D surround sound with height, while ensuring full compatibility across these formats (between 2D and 3D recording). The A-format signals of two captured 3D sound fields (A and B) are inserted into a 3D sound design without matrixing. When used in 2D sound design, the captured sound fields provide an added adjustability of direct to indirect sound ratio, without distortions of timbre and dynamics, thanks to the near-coincident arrangement of tetrahedral capsules aiming in multiple directions. Students found mixing of the session tracks in 9.1 listening environment to be straightforward, as was monitoring in stereo with headphones and loudspeakers during the recordings. Just two tetraphonic microphones provided the essential building blocks of a 3D sound design.

The Prague 3D recordings will be presented during the 145th AES Convention either in a 9.1 listening room at the convention site, or at the Dolan Studio in New York University.

8 Acknowledgements

The authors would like to express gratitude to Mr. Juergen Breitlow from Georg Neumann GmbH for arranging a generous loan of the Sennheiser Ambeo VR microphones for this study. Special thanks are also due to all students and musicians who

participated in the Prague 3D recordings, and to our hosts in Prague: Richard Mader and Sarah Coffey.

References

- [1] Woszczyk, W., Nulsen, J., Hahn, E., Degtyareva, K., Castillo, J., and Nagata, H., “Tetrahedral Microphone: A Versatile “Spot” and Ambience Receiver for 3D Mixing and Sound Design”. Audio Engineering Society Conference e-Brief 57, International Conference on Spatial Reproduction – Aesthetics and Science, 2018 August 7–9, Tokyo, Japan (2018)
- [2] Morten Lindberg, 2L (<http://www.2l.no>)
- [3] M. Gerzon, ‘Experimental Tetrahedral Recording: part one ’, *Studio Sound*, Vol. 13, pp. 396-398 (August 1971)
- [4] M. Gerzon, ‘Experimental Tetrahedral Recording: part two ’, *Studio Sound*, Vol. 13, pp. 472, 473 and 475 (September 1971)
- [5] M. Gerzon, ‘Experimental Tetrahedral Recording: part three ’, *Studio Sound*, Vol. 13, pp. 510, 511, 513 and 515 (October 1971)
- [6] M. Gerzon, “Periphony: With Height Sound Reproduction”, *Journal of the Audio Engineering Society*, January/February 1973, Vol. 21, No 1, pp. 2-10.