Audio Engineering Society

Conference e-Brief

Presented at the International Conference on
Immersive and Interactive Audio
2019 March 27–29, York, UK

Applied Multichannel Recording of a Contemporary Symphony Orchestra for Virtual Reality.

Luke Reed[1], Martyn Harries[1], Alexandre Hurr[2] and Mathew Knight[3]

[1] *Creative Technologies Lab / Impulse Studio, University of the West of England, United Kingdom*

[2] *Student, Audio and Music Technology BSc., University of the West of England, United Kingdom*

[3] *Student, Broadcast Audio and Music Technology BSc., University of the West of England, United Kingdom*

Correspondence should be addressed to Luke Reed (luke.reed@uwe.ac.uk)

Fig. 1: A screenshot from the CineVR artefact.

ABSTRACT

Capturing musical performances for Virtual Reality (VR) is of growing interest to engineers, cultural organisations and the public. The application of ambisonic workflows in conjunction with binauralisation through head related transfer functions enables perception and localisation of sound sources within three dimensional space, crucially enabling height perception. While there are many excellent examples of orchestral recordings in VR, however, few make use of the height perception and favour 'on stage' horizontal positioning. This engineering brief presents a contemporary symphony orchestral performance captured with multichannel spot recordings and produced in second order ambisonics in which 51 performers were individually split and positioned across five levels of the performance space. The paper looks to critically discuss the methods employed addressing the workflow through pre-production, capture and post-production.

1 Introduction

Presented by The Paraorchestra and Friends [1] in April 2018, a live orchestra conducted by Charles Hazlewood performed *The Four Sections*, a rare work for symphony orchestra by contemporary composer Steve Reich, by 51 musicians individually split across five levels of Bristol's Colston Hall atrium. An audience would move in and amongst this large-scale sonic installation, experiencing their own unique balance of instruments and creating an unprecedented 'access all areas' orchestral invitation [2]. *The Anatomy of the Orchestra: A VR Experience* is a Cinematic VR artefact (360° film) that captured the dress rehearsal of this visceral, live performance as an immersive audio-visual experience. This engineering brief critically discusses the recording and post production strategies involved.

2 Recording Strategy

Colston Hall Atrium lacks the infrastructure commonly found in performance and recording environments (such as acoustic control, permanent cabling, patch bays and access points) requiring thorough planning of cabling logistics and microphone positioning. The orchestra itself was not arranged in traditional sections with individual performers dotted around the five levels and at the extremes of the performance space.

The final film artefact would produce a second order ambisonic (2OA) soundtrack achieved through the panning of sources with ambisonic synthesis plugins [3]. Due to the scale, positioning and concept of the performance and artefact it was decided initially to spot mic groups of performers in addition to the use of a first order ambisonic (1OA) microphone to provide an ambient fill [4]. This was intended to yield greater sound quality, isolation, control and spatial resolution in comparison to single point ambisonic or array recording strategy [5].

Initial planning lead to the inspection of building schematics and draft instrument positions (Fig. 2.) which informed cable runs and required networking. A Roland REAC [6] audio over Ethernet (AES50) network was used to reach the extremities of the network. A mobile Zoom F8 [7] field recorder was used for quick deployment and positioning between takes, coupled with a Sennheiser AMBEO [8] 1OA microphone to capture the sound field beneath the 360 camera. The Zoom F8's ambisonic recording mode ensures unity gain across the four capsules necessary for stable sound field capture in addition to offering real time decoded stereo monitoring for level checks.

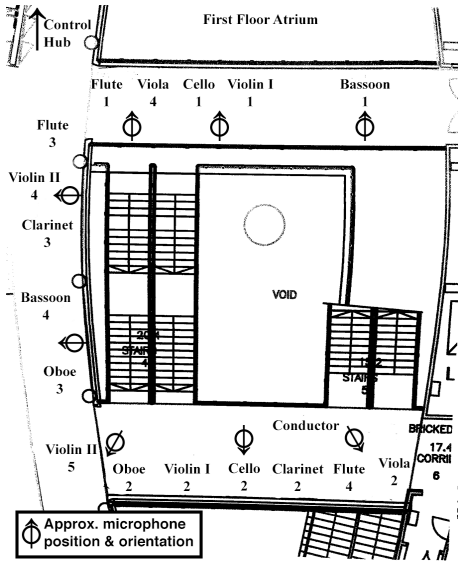

Fig. 2: An example of the first floor layout with drafted instrument and microphone positioning.

Two MOTU 8M, one MOTU 8PRE and one MOTU 896 [9] were used as USB interfaces connected to an iMac tracking direct to the digital audio workstation Reaper [10]. The four interfaces were split into two pairs of primary and secondary units connected in series via ADAT/optical. The two primary MOTU interfaces were connected in parallel to the iMac via USB and were combined as an Aggregate Device within the Mac OS Audio Midi Setup utility. This was necessary as the interfaces were limited to a maximum of three units when daisy-chained in series and thus enabled all 32 channels to be tracked simultaneously to REAPER.

The selection of microphones was limited to the university's stock and not enough channels were available to close microphone all 51 instruments. Where possible the spectral response of the group of local instruments informed microphone choice and positioned at 1-2 meters with an attempt to balance

the group. Directional polar patterns were favoured to isolate groups and a greater density of microphones were used for groups that would appear closer the camera as it was presumed that they would require more spatial precision and isolation of source.

Synchronisation between all devices was critical. Word Clock was passed over coaxial (BNC) between the two parallel primary MOTU interfaces, to help alleviate potential jitter or clocking issues. The Zoom F8 was synchronised with Linear Timecode (LTC) from an LTC generator VST in Reaper with a Tentacle Sync E [11] as an intermediary. An audio click track with verbal bar number prompts was recorded, this was transmitted to the musicians over FM in ear monitors and was essential in aligning different takes during post-production. See Figure 3 for a systems diagram.

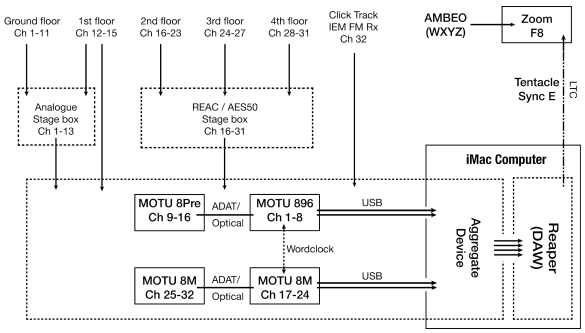

Fig. 3: A system diagram showing the use of multiple synchronised recording devices.

The 360 camera operation was carried out by Biome Productions [12] who utilised the Kandao Obsidian S [13], a 6K stereoscopic 360 camera. The camera has built in microphones but no audio input meaning use of a secondary Tentacle Sync E to burn LTC into the video audio track was not possible, therefore traditional clapperboard slates and log sheets were used to sync and organise rushes. Additional information was captured on log sheets including microphone polar coordinates (azimuth, elevation and distance values). As only one camera was available the performance was recorded as multiple takes so as to cut to different positions in the edit to create a more dynamic and interesting final artefact. A total of six takes were captured, two from each of the three camera positions.

3 Post-Production

The audio mixing was performed in Reaper [10] utilising the Facebook 360 Spatial Workstation (FB360) [14] suite of tools and plugins, which at the time of production only supported up to 2OA workflows. Biome Productions provided ready-to-use footage for the editing and mixing stages to be undertaken by UWE Bristol students. The footage was delivered as 6K DNxHR codec in a Quicktime container (.mov), it had been warped, stitched and colour balanced with some minor FX plates for large areas of overexposed footage. Lower resolution video proxies were produced for smooth video previews with the FB360 video client and while editing in Adobe Premiere [15] which supported basic 360 video editing functionality.

An individual ambisonic mix was created for each of the three camera positions before video editing was finalized [16]. This would allow for the musical performance to dictate the picture edit in respects of form and timing and reduce the need for excessive remixing. Two review sessions were held with Charles Hazlewood (the artistic director of the project) to ensure that the mixes met his expectations and to inform any creative decisions. From this, the decision was made for the mix to spatially contract into the omnidirectional centre as the music crescendos to its finale, this signification of climax was accompanied with a progressive shift in colour saturation as the video cut to different positions in the atrium.

Each source channel had EQ and light compression applied before being placed in the sound field using the FB360 Spatialiser plugin, initially based on the polar coordinate measurements noted down on log sheets. Due to the warping of the image from stitching the camera's six lenses, these real-world measurements often did not correspond accurately to the post-processed video. During playback the perceived location of sources were not congruent with the positioning within the user interface and at times appeared inverted 180 degrees. It is thought that the room modelling performed by the FB360 Spatialiser in combination with the abundant spill of microphones and large reverberant acoustic of the location created phasing issues within the

synthesised sound field. The room modeling was bypassed on all channels and positions aligned based on perceived location rather than as displayed in the user interface.

In an attempt to address further potential phasing issues, distant sound sources (greater than ~15 meters) were grouped together and summed to a mono bus before being spatialised at an average location. This resulted in a subtle clarity of these sources as they were too distant to perceive envelopment and allowed for phase coherence to be monitored before spatialisation.

Critical listening tests were made using varying amounts of the Sennheiser AMBEO microphone blended with the spatialised spot microphones. The AMBEO audio was encoded to B-format ambiX with the generic correction filter within the Sennheiser AMBEO A-B Converter VST plugin [17]. The AMBEO recordings were a useful reference but were not used due to the lacking spatial resolution of 1OA recordings, high frequency colouration [18] and the complications of phase alignment and sound field stability produced when mixed with the large array of spot microphones.

2OA bounces of each camera position were printed as nine channel poly-WAV files, using Reapers online render mode as using offline bouncing had produced timing errors in conjunction with FB360 in previous projects. Channel 1 (ACN 0, FuMa X, the omnidirectional centre) of the ambisonic mix and the musicians' click track were imported into Adobe Premiere and synchronised to the video allowing for a simple mix overview to dictate video edit and for the three camera positions and accompanying mixes to be accurately time aligned.

During editing and review it was found that jump cuts were more aesthetically pleasing than dissolves and fades, potentially as a reflection of the rhythmic and spatial complexity of the performance. However, an unpleasant disorientation was particularly strong when cuts occurred outside of the rhythmical or phrasal coherence of the music. Furthermore, having an identifiable motif and clear localisable source close to camera (for example an isolated lead violin line) helped to quickly anchor and orientate after a sudden cut and jump in position. These observations only revealed themselves when monitored in a head mounted display as opposed in the FB360 video client running as windowed on the desktop.

4 Summary

This project involved a challenging and complex recording strategy that had to mediate a number of resourcing, infrastructure and environmental compromises. Despite this the sound quality of the final artefact far surpasses what could be achieved using a single FOA microphone offering a more detailed envelopment and clarity of distant sources.

A sufficient sense of the atrium acoustic is preserved despite not using the AMBEO as an ambient fill, perhaps due to the approach not using close positioning as described by Nettingsmeier [4] and Riaz et al [5].

The non-standard positioning of instruments in combination with resourcing limitations was the fundamental cause of complexity and while a more conservative layout would have potentially led to a technically cleaner result it would have diminished the unique characteristic and interaction of spatial and rhythmic components of the performance.

The use of the Roland REAC helped alleviate some infrastructure issues however the unit used presented significant self noise that may be alleviated if using a more modern Audio over IP/Dante solution.

The mixing strategies utilised were sufficient and further investigations could be made on the optimum use of mono fold down of distant sources.

This project demonstrated a large scale orchestral performance that makes use of height in a way that is rarely undertaken. Considering that it is this height dimension that ambisonics and VR filming affords over conventional methods it is hoped that this example may aid and encourage further projects.

5 Acknowledgements

The recording and post-production was carried out by students of UWE Bristol's Music Technology BSc programmes as part of the Impulse Studio, a staff and student powered audio technologies

enterprise and innovation initiative. They worked directly with the event organiser (Hannah Walton), and lead engineer (Andy Rose). The project was directed by Luke Reed and Charles Hazlewood.

The authors would like to thank the Paraorchestra and Friends, Bristol Plays Music and the Colston Hall for their time, support and trust.

This project was funded by Arts Council England through Bristol New Music and Research England through the Higher Education Innovation Fund and Impulse Studio.

References

[1] Paraorchestra.com, "The Paraorchestra and Friends", Available: http://paraorchestra.com/, 2019, online; Accessed on 20/02/19.

[2] Colston Hall, youtube.com, "Introducing The Anatomy of the Orchestra", Available at: https://youtu.be/lUVNXAMMRmM, 2018, online; Accessed on 20/02/19.

[3] Frank, M., Zotter, F., and Sontacchi, A., "Producing 3D Audio in Ambisonics", in *Audio Engineering Society International Conference 57*, 2015.

[4] Nettingsmeier, J., "Hybrid Higher-order Ambisonic Recording". In *(S)PACE Symposium*, De Montfort University Leicester, May 4, 2011.

[5] Riaz, H., Stiles, M., Armstrong, C., Chadwick, A., Lee, H., and Kearney, G., "Multichannel Microphone Array Recording for Popular Music Production in Virtual Reality", in *Audio Engineering Society Convention 143*, 2017.

[6] proav.roland.com, "Digital Snakes", Available at: https://proav.roland.com/global/categories/digital_snakes/, 2019, online; Accessed on 20/02/19.

[7] zoom.co.jp, "F8 Multitrack Recorder", Available at: https://www.zoom.co.jp/products/handy-recorder/zoom-f8-multitrack-field-recorder, 2019, online; Accessed on 20/02/19.

[8] sennheiser.com, "Sennheiser AMBEO VR MIC", Available at: https://en-uk.sennheiser.com/microphone-3d-audio-ambeo-vr-mic, 2019, online; Accessed on 20/02/19.

[9] motu.com, "Products", Available at: http://motu.com/products, 2019, online; Accessed on 20/02/19.

[10] Reaper version 5.80, Available at: https://www.reaper.fm/download-old.php, 2019, online; Accessed on 20/02/19.

[11] TentacleSync.com, "Tentacle Sync E", Available at: https://www.tentaclesync.com/sync-e, 2019, online; Accessed on 20/02/19.

[12] Biomeproductions.co.uk, "360 Video Production", Available at: https://biomeproductions.co.uk/, 2019, online; Accessed on 20/02/19.

[13] Kandaovr.com, "Kandao Obsidian S/R" https://www.kandaovr.com/obsidian-s-r/, 2019, online; Accessed on 20/02/19.

[14] Facebook 360 Spatial Workstation version 2.2.1, Available at: https://facebook360.fb.com/spatial-workstation/, 2019, online; Accessed on 20/02/19.

[15] Adobe Premiere CC 2015, Available at: https://www.adobe.com/uk/products/premiere.html, 2019, online; Accessed on 20/02/19.

[16] Holm, J., and Malyshev, M., "Spatial Audio Production for 360-Degree Live Music Videos: Multi-Camera Case Studies", in *International Conference on Audio for Virtual and Augmented Reality*, 2018.

[17] Sennheiser AMBEO A-B Converter VST plugin version 1.2.1, Available at: https://en-uk.sennheiser.com/ambeo-blueprints-downloads, 2019, online; Accessed on 20/02/19.

[18] Bates, E., Dooney, S., Gorzel, M., O'Dwyer, H., Ferguson, L. and Boland, F. M., "Comparing Ambisonic Microphones – Part 2", in *Audio Engineering Society Convention 142*, 2017.