Composing and Producing Spatial Music for Virtual Reality and 360 Media

Dr. Enda Bates

Trinity College Dublin

ISSTA 2016

International Festival and Conference on Sound in the Arts, Science & Technology

Derry, September 7-9th, 2016

Abstract

Virtual Reality (VR) and associated 360 audio and video technologies, are in many ways ideally suited to the aesthetic of spatial music in which the location and/or movement of sounds in space is a composed aspect of the work. In traditional concert performances the musicians and performers are located on stage in front of the audience, however, this type of presentation cannot take full advantage of a 360 degree medium in which the viewer can shift perspectives at will. Spatial music, and in particular acoustic spatial music, is in contrast very well matched to this type of presentation as the viewer's ability to change perspective serves a real purpose in this type of aesthetic. Indeed, many composers of spatial music have in the past arranged the seating in the venue in spirals or circular arrangements so that no one direction is prioritized over another.

This paper describes some of the technical and artistic issues which arise when documenting spatial music performances using these newly emerging 360 recording and distribution technologies. A number of spatial compositions are discussed in terms of the compositional aesthetic, recording techniques and hardware, mixing strategies, 360 camera rigs, and online distribution channels for 360 media.

Keywords: Sound Diffusion and Spatialization, Soundscape Art, Recording Engineering, Music Software, Music Recording and Production, Electroacoustic Music

Composing and Producing Spatial Music for Virtual Reality and 360 Media

Dr. Enda Bates

Introduction

Figure 1 – Spatial Music Performances by NY Philharmonic (left) and Karlheinz Stockhausen (right)

Recent developments in Virtual Reality (VR) and 360 Video technologies represent the emergence of an entirely new medium for artistic expression with its own particular capabilities and requirements. This is particularly true for composers, sound designers and artists working in the broad area of spatial music, as this aesthetic is very well matched to this new medium in a variety of ways. Traditional concert performances in which all of the musicians are placed on stage can struggle to justify the extended field of view and different perspectives offered by this new medium. In addition, the arbitrary re-arrangement of musicians in the round may be counterproductive if the musical content and compositional aesthetic does not lend itself to this type of performance. Spatial music is in contrast very well suited to this new medium as the viewer's ability to change perspective serves a real purpose in this type of aesthetic. Indeed, many composers of spatial music have in the past arranged the seating in the venue in spirals or circular arrangements so that no one direction is prioritized over another (see Figure 2). This aesthetic also deals directly with the musical implications of spatially arranging musicians and sounds in this way and integrates the musical parameter of space fully within the compositional approach. While the medium of 360 video is therefore highly suitable for the documentation and recording of works of acoustic spatial music, the broader area of VR also offers many creative possibilities for composers of other types of spatial music such as soundscapes and acousmatic music. One of the themes of this festival and conference is the creation of new types of sensory engagement with space, something which this new medium can achieve in an unique way. The ability of VR to immerse the viewer within a scene, and the increased sense of empathy and engagement that results, is potentially an extremely powerful means of creative expression, which is just beginning to be developed.

The following paper will outline some of the technical and artistic issues which arise when documenting spatial music performances using these newly emerging 360 recording and distribution technologies, and outline some possible directions for future creative work in this area.

1 VR as a New Medium

VR is fundamentally different to traditional video and this is equally true in terms of the role that sound plays in this new medium. In particular, the spatial aspect of sound is significantly more important here, as in this new medium, just as in real life, the spatial location and distance of sounds provide a vital role in informing the viewer as to the location, content and relative importance of elements in the scene which are not currently in view. Much has been written on VR's ability to create a powerful sense of immersion and presence within a virtual scene, however, the importance of sound in achieving this suspension of disbelief has been somewhat underestimated to date. In cinema, it has been said that "s*ound is half the picture"* [1] and this is equally if not more true in 360 video and VR, however, the implementation of spatial audio for VR has been slow to emerge across the many VR platforms currently available. With the increasing standardization of first order Ambisonics / B-format as the audio format for VR, it is now entirely possible to record matching 360 sound and video, which is particularly useful for the recording and documentation of works of instrumental spatial music. However, this medium is also potentially of benefit for other musical aesthetics too, particularly soundscape composition. The use of 360 video may seem counterintuitive for an aesthetic which is heavily focused on the sonic experience (which is sometimes referred to as acoustic ecology). However, this medium allows the composer to control the visual component of the presentation in a way that is generally not possible when a soundscape composition is presented in a concert performance, or simply reproduced in a domestic listening environment. While multichannel loudspeaker reproduction can provide an effective sense of envelopment within an audio scene, there is always and inevitably a conflict between the visual perception of the loudspeakers and reproduction environment compared to the recorded soundscape. In a 360 video, the composer has in contrast complete control over both the visual and audible scene, and this ability to present an immersive visual environment which is then perhaps removed, can potentially be used to focus the attention of the audience on the audible soundscape in a powerful way. An excellent demonstration of this approach may be found in the accompanying VR experience to the award winning 2016 documentary *Notes on Blindness*[2]. Based on the audio diaries of John Hull, the film describes the sensory and psychological experience of losing sight. The accompanying VR presentation utilises spatial audio and sparse, dimly lit 3D animations to represent this experience of blindness in a highly evocative manner. A similar process could be used in the context of soundscape composition by first presenting a visible and audible representation of a real world location using 360 video, and then selectively or perhaps entirely removing the visual component to deliberately focus the attention of the audience on the soundscape alone. If so desired, more elaborate processing or manipulation of the recorded soundscape could be supported visually through the augmentation or replacement of the recorded 360 video with computer generated or synthesized visual imagery. There are perhaps some parallels here with Iannis Xenakis' multimedia work and the concept of the *polytype* [1], and perhaps also Edgard Varèse' *Poème Electronique,* which was first presented at the Philips Pavilion designed by Xenakis, at the 1958 Brussels World's Fair. Interestingly the latter work had already been the subject of a VR production in the very early stages of modern VR I 2009 [2].

Figure 2 – Trinity 360 Orchestral Performance

[1] This line is usually attributed to George Lucas but is probably a paraphrasing of the following quote, *"The sound and music are 50% of the entertainment in a movie"*

[2] www.notesonblindness.co.uk

While the technical production of VR and 360 video content still presents many challenges, the default audio format for this medium, Ambisonics, has been widely used by many electroacoustic composers for decades, such as in Trevor Wishart's *VOX* cycle for example. In addition, the combination of spatial audio field recordings with synthesized or generated material has long a feature of much electroacoustic music, and the compositions and writings of Natasha Barrett on the use of soundfield microphones is particularly relevant [3]. As such, composers of spatial music are well placed to effectively exploit the creative possibilities of this medium, however, despite recent advances many technical challenges remain. The Trinity 360 Project [3]was developed by the author to investigate these aesthetic and technical issues through the composition and 360 audiovisual recording of a number of works of spatial music.

2 The Trinity 360 Project

The Trinity 360 Project was initiated by the author to explore the relationship between spatial music, VR and 360 video. In autumn 2015, the project secured funding from the Trinity Creative Initiative[4] to support the live performance and 360 recording of newly composed works of spatial music. Two performances for orchestra and acoustic quartet respectively have been recorded to date with the first of these released as a 360 video via YouTube[5] in May, 2016 (see Fig. 2). The orchestral work, entitled *From Within, From Without* is the first movement of a larger work of the same title, composed by the author and performed by Pedro López López, Trinity Orchestra, Cue Saxophone Quartet, and Miriam Ingram, in the Exam Hall of Trinity College Dublin on April 8th, 2016. The orchestra was arranged in two similar instrumental groupings positioned on stage with additional brass and percussion instruments along the sides and rear of the hall (see Figure 3).

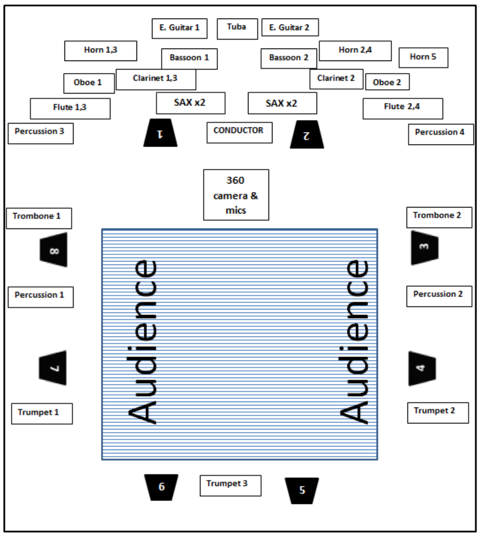

Figure 3 - *From Within, From Without* Stage Plan

The piece follows a spectralist approach in that all pitches are derived from the harmonic spectrum of the commencements bell of the universities campanile (see Figure 4), which is located just outside the concert venue and can be seen in the opening shot of the video. The piece begins with a call and response dialogue (which i in many respects the oldest form of spatial music) between the orchestra and a recording of the bell reproduced using loudspeakers at the rear of the hall. While the bell's spectrum informed the harmonic language of this work, the rhythmic structure was inspired by the synthesis technique of granulation, and a variation of the approach developed by composers such as Charles Ives and Henry Brant. Like Ives, Brant frequently used highly contrasting material with independent tempi, meters and harmonies to articulate and

3 www.endabates.wordpress.com

4 www.tcd.ie/trinity-creative

5 www.youtube.com/watch?v=WQvHH-u9W5w&index=1&list=PL00434BCD01707E05

highlight the different spatial locations of the performers. In this way, the difficulty of maintaining exact rhythmic coordination between the spatially distributed players is avoided, with only the entry of each player cued by the conductor. While Brant's approach is very effective at highlighting the spatial distribution of the instruments, it does inevitably result in very dissonant harmonies and textures. In contrast, in this work the melodic lines of the spatially distributed brass instruments are rhythmically independent, but much closer in terms of the harmonic language. As such, the independent lines overlap in ways that are sometimes consonant, and sometimes dissonant, resulting in a texture that is in some respects reminiscent of granulation; the electronic processing technique in which many fragments (or grains) of the original sound are layered over each other. This is particularly prominent in the middle section of *From Without, From Within* as the spatially distributed brass instruments perform very similar melodies that are deliberately desynchronized, resulting in an interestingly complex texture which avoids the issue of maintaining precise synchronization between the spatially distributed musicians.

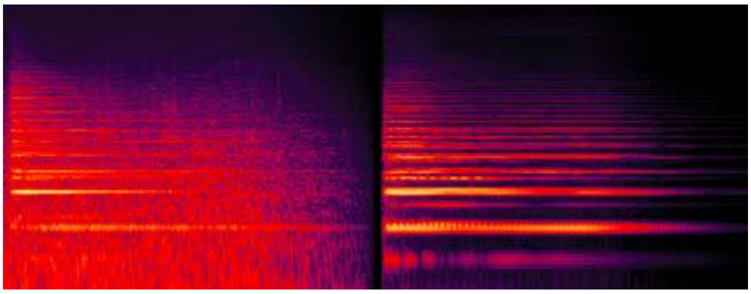

Figure 4 - Bell and Orchestra harmonic spectra (left) and Trinity Campanile (right)

The second movement, entitled *The Silent Sister*, was performed by Miriam Ingram and composed for solo voice and multichannel tape reproduced using a eight channel loudspeaker system (the title is an archaic slang term for Trinity College). This piece explores the relationship between Trinity College and the rest of the city, both from a sociohistorical perspective, and in terms of the sonic attributes of these two, quite different spaces. During its long history Trinity's relationship with Dublin, and indeed with the rest of the country, has often been difficult and even combative at times. While questions of religion and national identity were a significant factor, tensions between liberal and conservative voices, both from within and without, have also played their part in this complex relationship between *"town and gown".* It is perhaps not surprising therefore that the walls that surround Trinity were sometimes perceived as a barrier meant to keep people out! In Mary Muldowney's fascinating book *"Trinity and its Neighbours. An Oral History"* [4], many local residents employed by Trinity during the second half of the 20th century mention this perceived separateness of the university from the city (for many, their job interview was the first time they had been inside the college walls). As John McGahern puts it, *"when I was young in this small country Trinity College was so far removed from our lives and expectations that it seems a complete elsewhere".*

The third and final movement is a soundscape composition for eight loudspeakers entitled *Of Town & Gown*. This movement was constructed from field recordings made around the outskirts of the campus, which were then manipulated, processed and combined with the sound of the commencements and commons bells from the college campanile. In this way the piece again explores the relationship between the university and the rest of the city (between town and gown) through the blending and relating of these different sounds from both inside and outside the college walls.

The second work composed for the Trinity 360 project consists of an acoustic quartet (guitar, cello, flute and saxophone) with the musicians arranged symmetrically around the central recording position in the debating chamber of Trinity College Dublin (see Figure 5). The compositional aesthetic here follows a more traditional contrapuntal approach in the form of a modified round, a form of strict canon in which each part performs the same melody but starting at different times. The title of *A Round, Around* reflects this approach and the spatial arrangement of players. At various points in the piece, heavily accented notes and overlapping dynamic markings are used to create different types of rotational spatial effects. Footage and recordings from this performance are currently being used for an ongoing assessment of the performance of various soundfield microphones [5] but it is expected that the 360 video of this performance will be released early in 2017 once this study is complete.

Figure 5 - *A Round, Around* quartet recording

The next composition in the Trinity 360 project will focus on soundscape compositions and VR, as discussed earlier in this paper. This VR soundscape composition will function as a companion piece to the concert soundscape composition *Of Town & Gown* performed earlier this year and again here at ISSTA 2016.

3 Shooting 360 Video and Creating Visual Content for VR

Figure 6 – The Trinity 360 camera system (left) and GoPro Omni (right)

The last year has seen a rapid expansion in the number of 360 camera systems available, ranging from low cost spherical lens cameras, to highly expensive multi-camera rigs. While these low cost devices are quite simple to use and generally require little or no post-processing to create a 360 video, the picture quality is unsurprisingly generally quite poor. In contrast, much improved fidelity and picture quality can be achieved using a multi-camera rig, however, this improvement comes with a far higher price tag and more complex post-processing to stitch the individual camera feeds into a 360 video. It is worth noting that just as a distinction must be made between standard video and 3D video, there is also a distinction between 360 video and 3D (or stereoscopic) 360 video. While the use of stereoscopic imagery will add an increased sense of visual depth and can readily be delivered using a typical VR Head Mounted Display (HMD), designing camera systems which can shoot 3D 360 video is extremely challenging. Typical 3D cameras rely on a pair of lens to capture a different image for each eye. Arranging multiple video cameras to achieve this over a full 360 degree field of view is difficult and extensive post-processing of the individual camera feeds is required. In addition, any mismatches or drift in the synchronization of the individual camera feeds can potentially be much more of an issue in a stereoscopic system when compared to standard 360 video. Google's JUMP system[6] (based around 16 GoPro cameras arranged in a circular rig) can achieve this, however, this relies upon uploading all the camera footage to the online JUMP Assembler and the extensive computing power which Google can then apply to render the footage. It is worth noting that despite being launched as the Go Pro Odyssey[7] many months ago, this system is still not commercially available. The camera rig used to date in the Trinity 360 project (see Figure 6 - left) is an experimental system based around 12 Go Pro cameras in a circular array and intended to provide footage to aid in the development of our own stitching software for stereoscopic 360 video, while also providing acceptable monoscopic 360 video (using VideoStitch Studio 2). The lack of genlock and time-code synchronization remains an issue however, and for this reason we intend to reuse the GoPro cameras from this system with the newly released Omni rig (see Figure 6–right), released by GoPro just a few months ago. The Omni system contains full synchronization of all 6 cameras and preset stitching templates and software which potentially strikes a good balance between relative affordability, image quality and ease of use.

Developing synthesized or computer generated imagery for VR, whether with or without recorded 360 video, requires the use of game engines such as Unity and involves a quite different workflow which is beyond the scope of this paper. However, it should be noted that Ambisonics, whether in the form of soundfield microphone recordings or synthetically panned sources, can still be used in either production environment.

The importance of a 360 video display system to monitor the 360 video when producing the accompanying spatial audio soundtrack cannot be overstated, however, support for this facility has also been slow to emerge. The Facebook Spatial Workstation[8] (consisting of a number of plugins and templates for the DAW Reaper) currently supports synchronized playback of audio and 360 video through a HMDs such as the Oculus Rift, albeit solely for OSX and a very restrictive number of video formats. The SpookSyncVR application provides similar functionality for the DAW Reaper and the Kolor Eyes desktop video player and similar functionality is now available for Google's Jump Inspector App for Android. It is hoped that increased support this type of video monitoring will be integrated into other audio production environments in the near future.

3 Recording and Producing Spatial Audio for 360 Video

While early 360 videos (such as Beck's 360 recording of David Bowie's Sound & Vision[9] for example) attempted to use modified binaural microphones to capture a 360 degree soundfield, this approach was problematic in a number of ways. These microphones essentially capture multiple, concurrent binaural recordings from different perspectives. During playback, the head-tracking system cross-fades between these different recordings as the listener's head moves, so that sounds hold their position rather than rotating with the listener. While this works to an extent, sounds which are located at positions in between the different angles optimally captured by the microphone may not be reproduced correctly, and smoothly rotating the sound to compensate for head movement is quite difficult to achieve. In addition, these microphones are highly idiosyncratic, non-standardized and somewhat bulky.

Ambisonics [6] is in contrast very well suited to 360 video as the entire audio scene can be readily and smoothly rotated for headtracking during playback, recordings can be directly produced using soundfield microphones, and the technique is generally well understood and widely available. First Order Ambisonics (FOA) or B-format describes a three dimensional soundfield using four channels of audio labelled W, X, Y & Z. These four channels of audio correspond to the overall non-directional sound pressure level [W], and the front-to-back [X], side-to-side [Y], and up-to-down [Z] directional information (see Fig. 7) which can then be decoded for a variety of loudspeaker arrangements. For VR and 360 video, the Ambisonics signal is generally

6 https://vr.google.com/jump

7 https://gopro.com/odyssey

8 www.facebook360.fb.com/spatial-workstation

9 www.wired.com/2013/02/beck-360-degree-online-video

decoded for a virtual loudspeaker array[10], in which each individual loudspeaker feeds are encoded into a binaural signal using HRTF processing, before being mixed together for the final output.

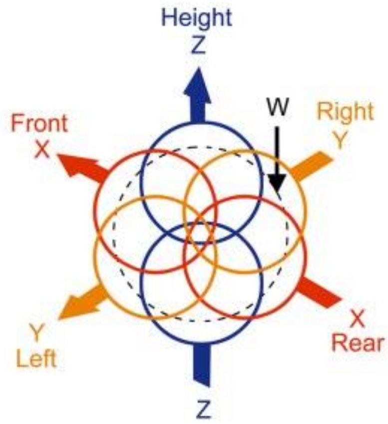

Figure 7 – First Order Ambisonics, or B-format

While B-format is certainly an efficient means of representing a three-dimensional sound scene which is well suited to VR and 360 video in many ways, this first order representation does present some issues. For example, the spatial resolution of FOA is relatively course [6], and this is particularly true for B-format recordings made with a soundfield microphone in a reverberant acoustic and/or when the microphone is placed quite distant from the musicians [7]. In addition, the use of computationally efficient HRTF sets for virtual loudspeaker rendering to head-tracked binaural can also introduce errors in localization. Finally, from an aesthetical perspective, this type of coincident microphone arrangement produces a very different sound to that of the traditional spaced microphone techniques which in the past have often preferred for orchestral recordings [8,9].

Figure 8 - Various Soundfield Microphones (from left to right): Core Sound TetraMic, Soundfield MKV, MH Acoustics Eigenmike, and Zoom H2n [4]

A number of new microphones have recently emerged which can be used to create a B-format recording and their relative performance is currently the subject of an ongoing study by the author. The first part of this study assessed monophonic timbral quality and localization accuracy using subjective listening tests and an objective directional analysis respectively [4]. Best results have been achieved overall by the Soundfield MKV system which despite being an older model in the Soundfield range, produced excellent results in terms of timbre, localization accuracy and levels of noise. While these microphones are to be recommended for music recordings, the accompanying mains powered rack unit is not suitable for field recordings. Core Sound's TetraMic is widely used for VR recordings, is significantly more affordable than the Soundfield range, and produced excellent results in our study in terms of timbre and localization accuracy. However, the output of this microphone is a very low level signal and so requires additional preamplifier gain compared to other

10 www.mee.tcd.ie/thrive

systems. As such, it is recommended that high quality preamps are used with this microphone to avoid excessive hiss in the recording (something which was notable in the results for speech signals in our original study [4]). Sennheiser's recently developed Ambeo microphone represents another alternative which is less demanding in terms of preamplification compared to the TetraMic, however, its timbral quality is yet to be formally assessed (this microphone will be included in the second part of our study). MH Acoustics Eigenmike is an interesting system in that it can produce both first and higher order Ambisonics recordings. However, achieving a high level of timbral quality using a system based on large numbers of microphone capsules is a significant challenge, particularly in its high frequency response [4].

The Zoom H2n is a low cost, portable digital audio recorder containing two opposite facing stereo microphone pairs. After a recently released Firmware update, the H2n can now natively record a horizontal only B-format signal, which can be directly attached a 360 video. For increased flexibility, this can also be achieved in post-production using the original 4-channel mode of two stereo pairs; the front facing mid-side recording, and the rear facing XY pair. Using a plugin for the DAW Reaper which will be freely released publicly[11] in the coming months these two recordings can be similarly converted into a horizontal only B-format signal, while still retaining access to the original stereo recordings. In our study, the H2n unsurprisingly performed worse than other microphones, however, its performance is reasonable given that it was not originally designed to produce B-format recordings, and its extremely low cost when compared to the other microphones. Surprisingly, the H2n has been used for audio capture in Google's JUMP system for immersive virtual reality content, and the associated commercial camera system, the GoPro Odyssey [1]. The portability and low cost of this device also make it quite useful for spatial field recordings, such as the technique adopted for the concert version of the soundscape composition *Of Town & Gown* discussed earlier. It was known in advance that the performance venue would present challenges in terms of the acoustic and large spacing between loudspeakers, and that little time would be available for tweaking the loudspeaker array on the day of the concert. For this reason, a conscious decision was made to avoid the use of Ambisonics and B-format recordings, and instead focus on a spaced microphone techniques suggested by Augustine Leudar [10]. All the field recordings for this octophonic piece were taken using four stereo recorders widely spaced at a distance that approximates the spacing between loudspeaker pairs in the performance space (see Figure 9). This resulted in an extremely spacious, diffuse recording which worked well in the context of this aesthetic and in the less than ideal acoustic of the performance venue. The pleasant, and in this case usefully diffuse sound does come at the cost of directional accuracy however, which is clear when this reproduction is compared to a B-format recording of the same scene constructed from a 4-ch recording from just one recorder. For the concert piece, certain sounds were therefore extracted to a mono file from the original field recordings and amplitude panned around the array to re-inject some of this missing directionality. This change in approach reflects the importance of matching the production technique to the particular medium and performance context in question. While spaced microphone techniques can be attractive and even beneficial for loudspeaker reproduction, and have long been the preferred technique for orchestral recordings, this must be balanced against the fact that VR is an entirely new medium, with its own demands and requirements. While much more practical experimentation is certainly required, a new medium will require a new recording aesthetic and it is noticeable that of the many 360 videos released by professional orchestras to date[12],[13] none have involved spatial music and all have used a traditional, static stereo mix.

When recordings music for 360 video, there is potentially a conflict between the optimal audio recording position (ideally relatively close to the musicians) versus the optimal position for the 360 camera. While it is natural to assume that the main 360 microphone should be placed in exactly the same position as the camera, this may not be the best approach, or indeed necessary. It is worth noting that standard video recordings of music performances generally combine relatively close audio recordings (above the conductors head for example) with video captured from further away. In the context of 360 video music recordings, it may therefore be beneficial to position the main microphone array closer to the musicians than the camera. The results of one study support this view and suggest that the perceptual binding of audible and visual components was maintained in the presence of incongruities in the respective distance of each stimulus, within a certain margin of error [11]. This indicates that microphones may be positioned closer than the camera while still being perceived as consistent with the visual position of the musicians in the video recording. However, this only holds for a certain margin of error, beyond which the binding of stimuli breaks down. The

[11] www.endabates.wordpress.com

[12] www.all360media.com/work/la-philharmonic

[13] www.youtube.com/watch?v=YVmyAPufGaE

degree to which this holds is naturally highly context dependent and as has been noted earlier this increases the importance of an appropriate 360 video monitoring system when producing such recordings.

Figure 9 – Spatial Field Recording for *Of Town & Gown*

In the specific context of spatial music, it is frequently the case that some or all of the musicians will be positioned at a greater distance from the main microphone than is ideal. This is particularly true if an audience is also present, such as at a concert performance for example. This was noticeable when recording the two performances conducted to date for the Trinity 360 project. The orchestral recording was taken during a live performance and as such was placed well beyond the critical distance of all musicians. As the hall was also highly reverberant, this resulted in highly diffuse main microphone recording with very little directionality and the addition of spot microphones was absolutely required. As no audience was present for the quartet recording, the musicians could be placed much closer to the camera and main microphone array resulting in much higher spatial fidelity in the soundfield microphone recordings. In this scenario, perhaps the more important question is how close the musicians may be placed to the camera in order to avoid stitching artefacts (most 360 camera systems struggle to accurately capture close sources within have a certain minimum distance), and to ensure a comfortable experience for the eventual viewer. In this regard, it is worth approaching such a recording as a performance for one, and arranging the musicians at a distance that feels natural and comfortable in the room.

Traditional surround recordings often use two microphone arrays for the respective capture of the direct and reverberant soundfields, as this allows for further adjustments in post-production [12]. This also allows the main microphone array to positioned quite close to the musicians to primarily capture the direct sound, which is simply not possible for spatial music due to the distribution of musicians around the hall. The relative balance of sources within a B-format recording can be manipulated somewhat using warping and dominance effects [5] but in general, additional spot microphones are needed to ensure the correct balance of instruments, to counteract for the limitations of first order Ambisonics in terms of localization accuracy, and to address the potentially increased distance of the microphone from the musicians. Günther Theile's technique of Room-Related time Balancing (RRB) is potentially useful here as a means of introducing these close, spot microphone recordings while maintaining a plausible, although perhaps not necessarily accurate, sense of distance [12]. The RRB technique involves converting the mono spot mic recordings into a number of artificial early reflections which are then delayed and panned according to the natural pattern of reflections in the hall. While this can be implemented manually using an Ambisonic encoder and appropriate time delays and reflection angles, various tools have emerged from the world of VR gaming which can be used for this purpose. The Facebook 360 Spatial Workstation supports various distance effects and room modelling using a computationally efficient approach while more elaborate systems are available as commercial plugins [7]. One additional benefit of this approach is that the use of this type of reflections based distance processing can improve the externalization of binaurally reproduced recordings, which is obviously important in the context of VR [13]. It should be noted however that Theile does not recommend the RRB technique if the main microphone recording does not contain a satisfactory directional image [12]. That was certainly the case for the highly diffuse soundfield microphone recording of the orchestra presented here, so for this production, spot microphones were used to re-introduce a certain amount of direct sound, and hence directionality into the final mix. In contrast for the quartet, the less reverberant acoustic and closer position of the musicians produced a much higher degree of spatial fidelity in the main microphone recording, and the standard RRB approach was far more useful here.

Conclusion

The preceding discussion has outlined some of the creative possibilities of 360 video and VR for the recording and production of spatial music, and indeed the perception of spaces in the broadest possible sense. While a number of music productions and performances have to date been presented using 360 video, these have often failed to fully utilise the spatial possibilities of this entirely new medium. The effective use of space in music requires the full integration of this parameter within the compositional aesthetic and much can be learned through the study of the long history of spatial choral, orchestral, electroacoustic, and soundscape composition. While many practical and technical challenges remain, the emergence of an entirely new medium such as this represents a significant and exciting new development for composers and sound designers and it is hoped that the works and experiences described here will inspire further, creative work in this area.

Acknowledgements

This work would not have been possible without the technical assistance of Prof François Pitié, John Squires and all the members of the Spatial Audio Research Group. The research and performance was supported by Trinity College Dublin's Creative Initiative, and Science Foundation Ireland.

References

[1] Harley, M. A., "*Music of Sound and Light: Xenakis's Polytopes*", Leonardo, The MIT Press, Vol. 31(1), pp. 55-65, 1998.

[2] Lombardo, V., *et al*, *"A Virtual-Reality Reconstruction of Poème Electronique Based on Philological Research"*, Computer Music Journal, Vol. 33(2), pp. 24-47 (2009).

[3] Barrett, N., *"The Perception, Evaluation And Creative Application Of High Order Ambisonics In Contemporary Music Practice*", IRCAM Composer in Research Report, www.natashabarrett.org/Barrett_IRCAM_Report_v2.pdf, 2012.

[4] Muldowney, M. "*Trinity and its Neighbours. An Oral History",* Dublin: School of Histories and Humanities, Trinity College Dublin, 2009.

[5] Bates, E., *et al "Comparing Ambisonic Microphones – Part 1",* Proceedings of the Audio Engineering Society International Conference on Sound Field Control, 2016.

[6] Frank, M., *et al*, *"Producing 3D Audio in Ambisonics",* Proceedings of the 57th Audio Engineering Society International Conference, Guildford, UK, 2015.

[7] Bates, E., *Boland, M. "Spatial Music, Virtual Reality, and 360 Media",* Proceedings of the Audio Engineering Society International Conference on Audio for Virtual and Augmented Reality, Los Angelus, California, September 2016.

[8] Jacques, R., *et al*, *"Empirical Comparison of Microphone Setups for 5.0 Surround",* in 22. Tonmeistertagung VDT International Audio Convention, 2002.

[9] Camerer, F. and Sodl, C. *"Classical Music in Radio and TV–a multichannel challenge; The IRT/ORF Surround Listening Test"*, http://www.hauptmikrofon.de/orf.htm, 2001.

[10] Leudar, A., *“An alternative approach to 3D audio recording and reproduction”,* Divergence Press, Journal, Issue 3, December 2014.

[11] Finnegan, D., *et al*, *“Compensating for Distance Compression in Audiovisual Virtual Environments Using Incongruence”,* SIGCHI Conference in Human Factors in Computing Systems, 2016.

[12] Theile, G., *“Multichannel Natural Music Recording Based on Psychoacoustic Principles”*, Proceedings of the 19th Audio Engineering Society International Conference, 2001.

[13] Y. Yuan *et al*, *“Externalization Improvement in a Real-time Binaural Sound Image Rendering System”,* International Conference on Orange Technologies, 2015.