FINAL PROJECT

Massimiliano Borghesi - Ma Sound Design 2017-2018

_ _ _ _ _ _ _ _ _ _ _ _ _

PORTFOLIO

_ _ _ _ _ _ _ _ _ _ _ _ _

SPATIAL AUDIO FOR 360 VIDEO AND VR
The technical and creative aspects of the full chain of producing Spatial Audio for 360 video and VR

Abstract

This research focuses on the practice of Spatial Audio production for 360 film and VR. The complete Spatial Audio production chain is investigated, the pre-production, production, post-production and delivery processes. The research will focus on the Spatial Audio tools featured in AVID ProTools HD, comparing Ambisonics techniques with traditional techniques. The findings of the practice-based research suggest that the combination of traditional and Ambisonics techniques might be the most viable option for producing Spatial Audio content. Further findings suggest that the creative aspects of Spatial Audio sound design might benefit from adopting a hyper-realistic aesthetic. In the research microphone techniques and post-processing workflows are suggested. The output of the research comprise five 360 videos and cinematic VR experience. In the process of making these immersive contents workflow guidelines, issues and solutions are proposed. The author intends to suggest good practices that might help peers in creating Spatial Audio for 360 film and VR.

In the Appendix 1 to this research there is a broken-down analysis of the technical aspects and further considerations on the production of Spatial Audio content for 360 video and VR.

1. INTRODUCTION

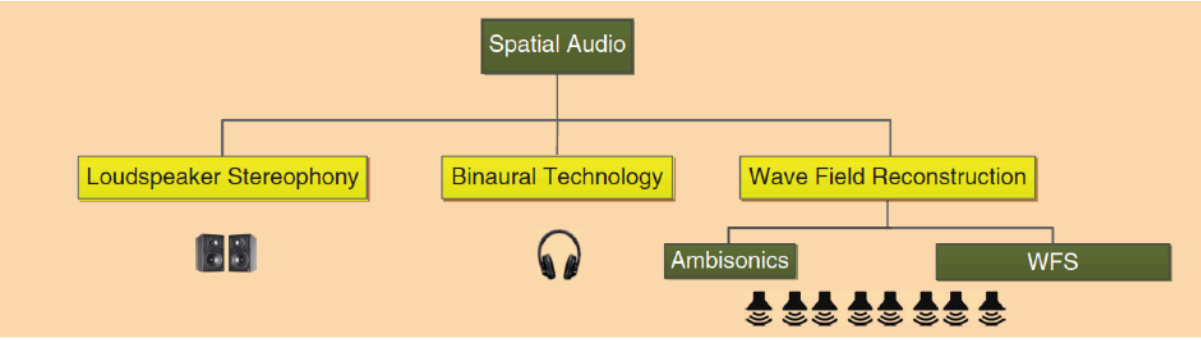
Undoubtedly we are witnessing an increasing interest and diffusion of immersive audio in all of its declinations.

We are continuously bombarded with the terms: immersive, 3D, 360. These terms are often related to visual contents that can be experienced on computers, smartphones, VR headsets, such as 360 films and cinematic VR. Similarly we are witnessing the coupling of this terms with the words ‘audio’ and ‘sound’.

The terms 360, 3D and immersive, when associated with the words audio and sound, all share the same meaning. For simplicity purposes this varied terminology will be included under the umbrella of Spatial Audio and the whole of these terms will be referred to as Spatial Audio in this paper.

In the book ‘Spatial Audio’ (2001), Rumsey defines the term in this way: “The goal of Spatial Audio is to capture the salient characteristics of a sound field into a recording and reproduce them at a possibly distant place and time such that a listener perceives a similar acoustic sensation as at the original.” (Rumsey, F., 2001)

Considered the state of the art of Spatial Audio techniques, this quote seems dated. This is because Spatial Audio not only refers to sounds captured in the real world, but thanks to the latest developments in technology that allow to synthesise virtual realities, the purpose of spatial audio is not anymore only to ‘capture the characteristics of a sound field into a recording’, but also to *synthesise* audio characteristics ‘such that a listener perceives’ an acoustic sensation undistinguishable from a real auditive experience. Thus, beside analysing the spatial audio technologies for recording and reproducing a sound source, this paper will investigate the potential of new technologies for synthesising a sound field, that have been developed over the last decades.



This research will focus on Ambisonics, binaural and traditional stereophony techniques, how these techniques can be implemented in the process of pre-producing, recording, post-producing and delivery Spatial Audio for 360 film and VR. The advantages and disadvantages of each these technologies will be comparatively discussed, highlighting their strengths and weaknesses at the current state of the art.

The diffusion of Spatial Audio techniques has induced AVID, developer and owner of the industry standard DAW ProTools, to introduce some tools to allow to create spatial mixes based on Ambisonics technology. This project will investigate the spatial audio tools available for ProTools HD, highlighting the technical strengths and limitations of the DAW and suggesting creative solutions for producing spatial audio for 360 videos and VR content.

This project will try to answer to the question: How can traditional filmic recording, editing, processing and mixing techniques be implemented in immersive media?
What deliverables need to be met?
Is Ambisonics a self-sufficient technique to produce Spatial Audio?
Are the tools offered by ProTools sufficient for a professional spatial mix?
What might be the areas of improvement?
Is it possible to draw some guidelines in the creation of Spatial Audio for VR and 360 films?

The project will feature an analysis of problems and solutions for recording Ambisonics and stereophony for VR and 360 film, overlook the technical positive and negative aspects of creating a spatial audio mix on Pro Tools HD and the creation of the right delivery formats for the main existing platforms. The research will be supported by practical tests and experiments on spatial audio recording, processing, mixing and delivering. The practical outcome of the project consists in a portfolio comprising of five 360 videos with spatial audio.

2. RATIONALE

VR and 360 videos are increasingly being implemented for any kind of media and purpose. Behind this trend lies the interest of companies to provide and sell new technologies. Whereas the technology has achieved a technical success, still contents that take advantage of the potential of the immersive media are lacking. Studies highlighted how spatial sound is perceived as the most important element in delivering a truly immersive experience (Reilly, E.,2017 'Headspace VR' Study Examines Audience and Empathy).
However, in the recent years spatial audio capabilities for VR were limited to game audio engines, and whether a dynamic spatial mix was achieved, the audio quality resulted poor. With the democratisation of the immersive technologies, spatial audio capabilities have been gradually implemented in more widely used DAW, first Reaper and now ProTools and Nuendo. These technological advancements give access to spatial audio to sound designers who were used to traditional mixing techniques. Now spatial audio with dynamic mixes, responding to head tracking, are not anymore programmer territory, but are accessible for a wider portion of the audio community. Nonetheless many new parameters are to be taken into account and the industry standard DAW presents possibilities as well as limitations. To comprehensively understand the process of creating Spatial Audio, this research will investigate the complete Spatial Audio production chain: pre-production, production, post-production and delivery. The purpose of this research is to investigate these processes and to gather the knowledge and the know-how to deal with 360 content, deliver professional spatial audio mixes and being able to articulate the possibilities of spatial audio for VR and 360 to fellow sound designers, post production houses and companies.

3. AIMS AND OBJECTIVES

To the moment this research is being carried out, there are no academic studies on how ProTools offers tools to implement traditional mixing in spatial audio. Nonetheless there are many studies on Ambisonics and Spatial Audio. The purpose of this research is to investigate these points and elaborate a workflow that mixes the

traditional channel based approach with the spatial audio capabilities now integrated in ProTools.
This final project overlooks the current spatial audio techniques for 360 Videos and VR.
It comprises three main parts:

1) a study phase were a research on Ambisonics and spatial audio plugins was carried out
2) A practical phase in which 5 VR/360 videos with Spatial Audio projects were pre-produced, produced, post-produced and delivered
3) An evaluation phase, where the experience gained on the field is summarised and analysed to point out the possibilities and limitations of Spatial Audio techniques and outline the best practises on the field.

The project's main body consist in the development of two experimental preparatory projects and three major projects involving Spatial Audio. These three practical projects offered the possibility of experimenting with the Ambisonics technology mixed with traditional mono and stereo techniques to output an audiovisual product featuring spatial audio.
The 3 projects were pre-produced, produced, post-produced and delivered, this meaning that the full chain of spatial audio production was tackled. The complexity of the project allowed the us to face unexpected challenges and work out solutions and possible good practices in the field of spatial audio for VR and 360 Videos. We hope that the experience gained through this process will be valuable for peers facing the possibilities and limitations of spatial audio related to immersive video contents.

4. METHODOLOGY

This research comprises three main sections: a literature review on Spatial audio history and techniques (in the CONTEXT section), a main body section where the field practice is carried out through the realisation of 360 videos / VR with Spatial Audio, a conclusive chapter where the findings are summarised. The research will refer to the document in Appendix 1, where a broke down of the practice on the field will delve into the technical aspects of Spatial Audio. This research, beside investing the technical aspects of creating a spatial audio mix in Pro Tools, will focus on the creative challenges and opportunities offered by sound for immersive media. The research culminates with the production of a portfolio of 360 videos featuring spatial audio, investigating different genre and audio solutions. The project focuses on the post production spatial capabilities of ProTools HD 12.8.3, and onto the creative solutions devised to work with tools that are still in their infancy and formats that are not yet standardised. Furthermore audio recording techniques for spatial audio will be tested, through the comparison of Ambisonics recording and multichannel recording. The recording thus gathered will be compared on the basis of audio quality and mixing possibilities in ProTools.
Thanks to a collaboration with students from the Northern Film School and lecturers from the Computing and Creative Technology school of Leeds Beckett University some new and original 360 content with spatial audio has been created. The project's final portfolio features 5 products in total: 2 experimental test videos and 3 VR/360 projects. The creation of these immersive contents involved other students and provided an invaluable experience in all the stages of creating spatial audio contents: ideation, pre-production, production, post-production and delivery.
The final pieces have been delivered both on Youtube and Facebook, to compare how these two publishing platforms work with spatial audio, their differences, strengths and issues.
While Pro Tools supports up to 3rd Order Ambisonics, Facebook supports 2nd Order Ambisonics and Youtube 1st Order Ambisonics. The ability to upload spatial audio to Facebook or YouTube is still in early stages as well, with both platforms providing limitations and differences from one another in what you can do (Forsythe, 2018). This

research will include a comparison between the two most popular publishing platforms.

4.1 THE 3 IMMERSIVE PROJECTS

All the content produced as output for this project is a combination of video and audio material, the video part being filmed or rendered in 360 and the audio part being rendered with spatial audio.
For the video part I collaborated with students from the Northern Film School and with 360 technology professionals Fluido TV. For the audio part I personally carried out all the tasks involving sound.

In a time window of 8 months the world of spatial audio for immersive video contents was penetrated, experimentation with it were carried out, three main projects to develop for this Final Individual Project were identified, the needed field research was carried out, two experimental tests 360 videos with spatial audio and three final projects were produced and released.
An initial experimentation was carried out with limited audio and video equipment, focused on the possibilities of spatial audio post production plugins. Links to this experimentations can be found in the OUTPUT section.
Once we established that we were able to produce 360 content with spatial audio, we identified three projects to develop in collaboration with other professionals inside and outside the University.

The two experimental test 360 videos with spatial audio:
1. “Spatial Audio Demonstration”
2. “Lazarus” - music video in 360

This two experimental videos were created to test the possibilities and issues of 360 filming and Spatial Audio. “Spatial Audio Demonstration” has an educational purpose, showcasing some artistic possibilities that can be achieved through Spatial Audio. “Lazarus” is an experimental 360 videos, recorded live. The same singer performed four times the same song, the four performances were then put together, side by side in a 360 video. The audio was captured on set and spatialised accordingly to the singer’s position.

Experimenting with the immersive technology laid the foundation for the three major 360 / VR projects:

1. “La Traviata” Opera in 360
2. “Spirit of Cornwall” Contemporary ballet with live ensemble in 360
3. “In the Cave” Cinematic VR

“La Traviata” is the live capture of some excerpts of the Italian Opera “La Traviata” by Giuseppe Verdi. The opera was performed by a 60 pieces orchestra and 40 pieces chorus with 8 lead singers. The performance was recorded live and the famous aria “Brindisi” was chosen to be featured in this project. In the Appendix 1 can be found further technical details and photo reportage of this project.
“Spirit of Cornwall” is the live capture of a contemporary ballet performance with live ensemble. The performance took place at the Chelsea Flower Exhibition in London. The music and choreography were commissioned to create a multi-disciplinary event, mixing architecture, garden-art, sculpture, music and dance.
More on this project can be found in the Appendix 1.
“In the Cave” is a cinematic VR commissioned by La Biennale - Venice Film Festival - College VR. This is a film in 360, that mixes filmic and videogame techniques. The process of making this project differs from the other two, utilising a game engine to build the film architecture and playback the experience. A technical report on the making of “In the Cave” can be found in the Appendix 1.

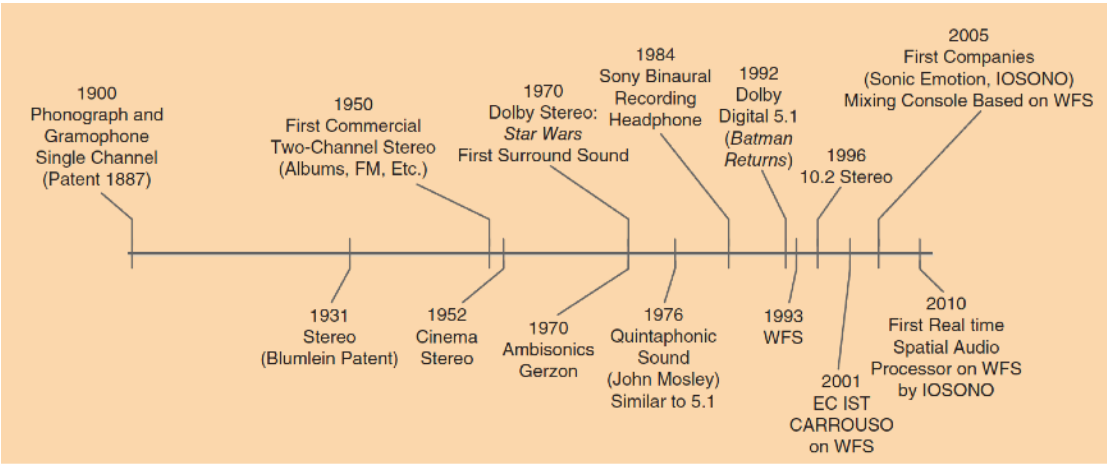
The process of making these immersive media encompassed the complete chain of Spatial Audio production for 360 and VR: pre-production, production, post-production and delivery. In this research there will be four major chapters, summarising the experiences, challenges and solutions devised for all the three main projects.

5. CONTEXT

5.1 HISTORY OF SPATIAL AUDIO

This research focuses on the use of Ambisonics and binaural techniques applied to 360 videos and VR, however a brief historical analysis outlines how the origin of Spatial Audio dates back to the very invention of sound recording and reproduction techniques. From the invention of the phonograph to the development of Ambisonics the history of spatial audio can be considered as a continuous evolution. In this context is important to highlight that what is commonly referred to as monophonic, stereophonic and surround sound, are all declinations of what can be broadly defined as Spatial Audio. Invention of **Stereophony** lays a milestone in the development of Spatial Audio techniques. Ethimologically the word stereophony is formed by the ancient Greek words:
στερεός (*stereós*, "firm, solid") + φωνή (*phōnḗ*, "sound, tone, voice") and it was coined in 1927 by Western Electric, by analogy with the word "stereoscopic" (Henry George Liddell, Robert Scott, *A Greek-English Lexicon*, on Perseus Digital Library).
Modern stereophonic technology was invented in the 1930s by British engineer Alan Blumlein at EMI, who patented stereo records, stereo films, and also surround sound. ("Early stereo recordings restored". *BBC*. 2008) The development of multichannel technology had begun and while “Two-channel stereophony is the oldest and simplest audio technology” it “has been progressively extended to multichannel stereophony systems, through 5.1, 7.1, 10.2, and 22.1 surround sound systems. (Rishabh and Woon-Seng, 2013)

Is important to stress that Ambisonics technology is based on Blumlein’s work on stereophony and mid-side techniques.



There have been some paradigmatic changes in the approach to spatial sound, but for long time the most popular declination of it was the so called surround-sound. Despite of being a technology that has been devised almost 80 years ago, of which a seminal idea can be found in Walt Disney 1940 Fantasound system, featuring three channels played across 54 speakers (WM. E. Garity and J. N. A. Hawkins, “Fantasound,” *Journal of the Society of Motion Picture Engineers* 37, no. 8 (1941): 127–46), surround sound is still being developed by the entertainment and cinema industries.
This interest in creating an immersive soundscape is based on the evidence brought by scientific research, that has demonstrated that the listener is capable of perceiving different sound sources as emanating from different positions and experiencing a sense of the environment in which the recording was made (Hirst, J.M., 2016).
The most popular developer of surround sound systems is Dolby Laboratories. While there were some surround sound for film experiments in 1970s, the standardised 5.1 format was developed for the film Apocalypse Now (1979) by Dolby and multi-Academy-Award-winner sound designer, sound and picture editor Walter Murch. The 5.1 was replaced by the 7.1 and 9.1 and 10.2 formats. In the past two decades have been developed sound systems capable of reproducing the height dimension, bringing the surround sound technology to a three-dimensional phase. Other immersive formats with added height have been developed by Dolby competitors, as the Auro3d, and despite being more efficient than Dolby systems, have not yet caught on (http://ce-pro.eu/2016/12/2017-auro-3d-dolby-atmos-dtsx/).
In 2012 a further change of paradigm was brought by the advent of Dolby Atmos and DTS X technologies, implementing object-based technology versus the traditional channel-based technology of surround sound.
While these formats have been implemented mainly in cinemas, the diffusion of 360 videos and VR on smartphones and headsets has witnessed the resurgence of the Ambisonics format. Ambisonics technologies support for immersive media represents yet another change of paradigm in the Spatial Audio universe.

5.2 AMBISONICS

Ambisonics techniques both for recording and post producing the audio material gathered are largely utilised in this project. Ambisonics has progressively grown to become a standard format for immersive audio nowadays, but its origin dates back in the 1970s.
Ambisonics was the brainchild of a small group of British academics, notably Michael Gerzon of the Mathematical Institute in Oxford, and Professor P. B. Fellgett of the University of Reading. They worked in the early Seventies to create an integrated high-resolution surround-sound system. The Ambisonics project was built on the astonishing work on stereo recording and reproduction performed by Britain's early audio genius, Alan Dower Blumlein. The term Ambisonics was chosen for its meaning: “surround sound”. (Elen, R. 2001)
The system was designed from the beginning to enable recordings to be made with a special surround microphone (the Soundfield Microphone).

“Rather than capturing a spatially defined “window” of sound from a fixed perspective, an ambisonic recording captures a full sphere of sound from which multiple perspectives—whether mono, stereo or surround—can be derived later, in mixing, or even by an end user.” (Pior, D. 2016)
Ambisonics technology seems to have find its place more than 40 years after being devised. The possibility to capture a three-dimensional image of a space might well apply to 360 content and VR. There are undoubtedly many positive aspects of the technology to deliver an immersive experience. Ambisonics technology is capable of providing a full surround, with added height, while requiring only four channels total. Not only the surround-sound image remains largely unaffected by listener position, in other therms there is no sweet spot, but also the audio images can appear anywhere, not solely in the speaker positions. (http://www.ambisonic.net/)
This capability of producing phantom sources might be a solution for delivering the proximity effect that multichannel stereophony is not able to recreate.
There are further advantages in Ambisonics, when compared to other 3D sound technologies “regarding CPU and hardware (number of

channels and loudspeakers) make it a very attractive approach for 3D audio reproduction" Hollerweger, F. (2005).
It is expected that Ambisonics may outperform multichannel stereophony "within its sweet spot if same number of microphones and loudspeakers are available." (Otani, M. 2015)
Nonetheless Ambisonics reproduction requires certain standards for a faithful auditive image reproduction and as Otani, M. (2015) observes "this leads to another difficulty in applying these approaches to consumer use, especially to domestic use at home, because such loudspeaker arrangement necessitates an audio room specifically designed to allow it."

More doubts are arisen by Rishabh, R. and Woon-Seng, G. (2013) on Ambisonics, when they point out that "it has yet to gain acceptance in the commercial field and extensive research is being carried out on various derivatives of higher-order Ambisonics to improve its commercial feasibility."
"Rather than capturing a spatially defined "window" of sound from a fixed perspective, an ambisonic recording captures a full sphere of sound from which multiple perspectives—whether mono, stereo or surround—can be derived later, in mixing, or even by an end user." (Pior, D. 2016)

5.3 BINAURAL

The practical outcome of this research is intended to be consumed on headphones. Hence, while the production and post production workflow was based on Ambisonics, the final delivery is intended through binaural.
Binaural literally means "having two ears"or "relating to two ears". Binaural hearing, along with frequency cues, lets humans and other animals determine the direction and origin of sounds. The term "Binaural hearing" is more than 200 years old, having been used in early research papers from the late 18th century. During this period this term was generally used to describe recordings or reproduction techniques made with two signals. However at that time 'binaural' wasn't described as we know it today, as two audio signals modified from the reflections of our body (Paul, S. 2009).
Recent research has shown that listeners are able to use these binaural cues to determine the distances of nearby sound sources. (Brungart,

Douglas S. 2002). In other words, binaural techniques mimics how we perceive sound and is the ideal way to deliver Spatial audio through headphones.

In a head tracked 360 system the viewer looks around, seeing only the portion of the 360 sphere he/she is looking at. Similarly with sound, he/she will hear mainly the sound coming from where he is looking at, while the other sounds existing in the sphere will be attenuated. In this system, whether Ambisonics is the full sphere of audio in which the user is immersed, the user will perceive mainly the sounds related to where he/she is pointing to. In other words, he/she will not perceive the full sphere of sounds existing in the 360 Ambisonics sphere at once, but mainly the portion of sphere he/she are pointing to, while the rest of the sounds will be “out of focus”. Thus, being the sphere still an Ambisonic-encoded sphere, what the final user hears is a binaural rendering of the sphere. This binaural rendering changes depending on where the user is pointing to.

To this extent all the Ambisonics techniques implemented in this research are decoded and rendered to binaural for an optimal headphones playback system.

6. PRE-PRODUCTION

Pre-production is a key phase for any media production process, particularly relating to 360 filming. The main challenge presented by 360 filming is that there is no blind spot to hide he crew. The lack of a behind-the-camera scenario presents specific challenges for the lighting and sound department. The main issue consists in not being able to place microphones on set that are visible to the camera. Moreover, being the camera recording a 360 field of view, new microphone placement techniques have to be adopted.

6.1 LOCATION SCOUTING

Scouting location is vital to know where the pitfalls are and what opportunities the location can offer to solve the problems.
Scouting locations multiple times is a good recommendation as there are many aspects that need to be taken into account: the light conditions, the acoustics of the space, the presence of natural blind spots to hide the microphones, the chance to run test recordings.
Two of the projects in this research were live recordings of live events, the third project was a film set production. These two different scenarios require different approaches.

6.2 LIVE EVENTS - MICROPHONES

In a live recording the presence of a microphone is usually accepted, so mic placement can be designed on capturing the best possible sound, with little compromise for intrusion in the camera frame.
Another element in a live situation is the presence of an audience. Primarily the microphone placement doesn't have to interfere with the audience field of view, secondarily the audience itself could accidentally move the microphones and produce unwanted sounds.
More time in preproduction allows for a better microphone placement and the choice of the right microphone for each sound source.
However one of the two live recordings projects presented further challenges, as the location could not be scouted due to strict safety policy of the event. The limited time given to set up the audio recording equipment, forced to quick decision making and precise microphone placement.

It has to be considered that most of the times the equation: more time spent in preproduction = less time spent in post production, is true. Another fundamental factor that needs to be studied during pre-production is what kind of instrument or ensemble are going to be recorded. In the live events it has been a mix of orchestra, chorus, opera singers. A complete list of microphones used and other equipment with photos of mic placement can be found in the Project 1 and Project 2 sections of Appendix 1.

6.3 360 FILM - MICROPHONES

For the project involving a filmic set location recording there were other factors that presented challenges and opportunities. The most important factor was the time allocated for placement and testing. There is no blind spot and contrary to the live events, the microphones have to be perfectly hidden to the camera. The traditional filmic audio recording set up relies on radio microphones and primarily on the boom microphone, manoeuvred by a boom operator. This set up is clearly unaffordable due to the lack of a out of frame space, where the microphone and the boom operator can hide.
In this scenario the production design department might be the best ally, providing props and furniture to hide plant mics.
The pre-production phase for this project was the most crucial and it took 5 days in total, compared to a two days shooting.
This precious time allowed for hiding the mics and running recording tests before the shootings.
A detailed list of microphones used and close ups of mic placement can be found in the Project 3 section of Appendix 1.

6.4 SOUND LIST

When woking on film is key to have a plan of all the sounds that need to be recorded on set. An important pre-production task of the sound recordist is to compile a sound-list, with all the sounds, Foleys, props movements, Ambiences, SFXs that might need to be recorded and later implemented in the mix.
The full sound list compiled for “In the Cave” can be found in the Appendix 1.

7. PRACTICE ANALYSIS

7.1 PRODUCTION PROCESS

The world of VR and 360 videos is constantly evolving at high speed. During the 8 months time spent working on this research, the tools have multiplied and ProTools has implemented TOA (third Order Ambisonics), the Facebook 360 Spatialiser plugin has up-stepped from SOA to TOA support and many other software companies are releasing their own spatial audio plugins and interfaces. Considering this constant evolution this research focused on the process of producing 360 content with Spatial Audio, from the ideation to delivery. There is no strict guideline on how to run a spatial audio production, and this research intended to fill in this gap, drawing some guidelines based on the experience gained in the making of the three main projects. Being there no framework specific to VR, this research implemented traditional stereo and mono techniques side by side with Ambisonics techniques both in the recording process as well as in the post production chain.

7.2 RECORDING PROCESS

The recording process implemented two main techniques: traditional mono/stereo recording and Ambisonics recording.
Whether the traditional mono/stereo recording techniques required a microphone placement as close to the sound source as possible to capture a pristine sound, with no background sounds, the Ambisonics microphone utilised was placed generally as close to the camera perspective as possible.
The placement of the Ambisonics mic can be appreciated in the pictures featured in the Appendix 1.

7.3 MONO/STEREO RECORDING

The microphones used varied depending on what sound source needed to be recorded: mono shotguns mics were used for voice of the opera singers, large diaphragm cardioids for the orchestra and

chorus, XY stereo pair for smaller acoustic ensembles, radio mics for actors, omnidirectional plant mics for cinematic 360 set.
All the microphones were recorded either on a mixing desk connected via Firewire to a laptop, running ProTools HD, or on a portable mixer/recorder, the Zoom F8.
More technical details can be found in the Appendix 1.

7.4 AMBISONICS MIC

This research tested two Ambisonics microphone, The SoundField ST450II and the SoundField SPS200. Even though these two microphones are from the same company, their characteristics are peculiar and they are ideal to different use.
The ST450II is a more sensitive microphone, coupled with its own preamplifier.
The SPS200 is less sensitive, but more agile.
Both Ambisonics microphones were recorded onto the Zoom F8.

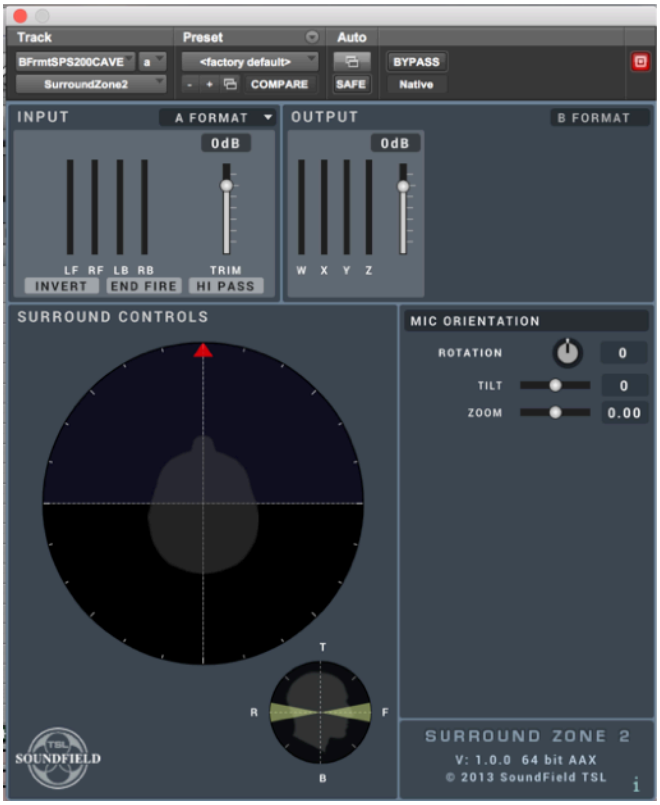
7.5 FINDINGS

The combination of the traditional mono/stereo and Ambisonics techniques can produce the best results. The Ambisonics recordings can be used as an ambiental recording that can be enhanced with the addition of mono/stereo recordings.
The SPS200 is preferable for location and outdoor recordings, being more robust, while the ST450II is the best choice for indoor and studio recordings.
There are some technical issues that can happen when mixing mono/stereo and Ambisonics, these are tackled in depth in the Appendix 1.
Further considerations on the findings can be found in the post production section.

8. POST-PRODUCTION

The post-production phase for VR and 360 videos is the core of the technology. Not only it renders the Ambisonics files to binaural, but also allows for the mono/stereo recordings to be placed in an Ambisonics sphere. These possibilities require technical passages that are specific to the Ambisonics workflow. However s it was demonstrated for preproduction and production phases, even for post-production the mixed use of traditional techniques with Ambisonics techniques can produce the best result.

8.1 AMBISONICS FORMATS

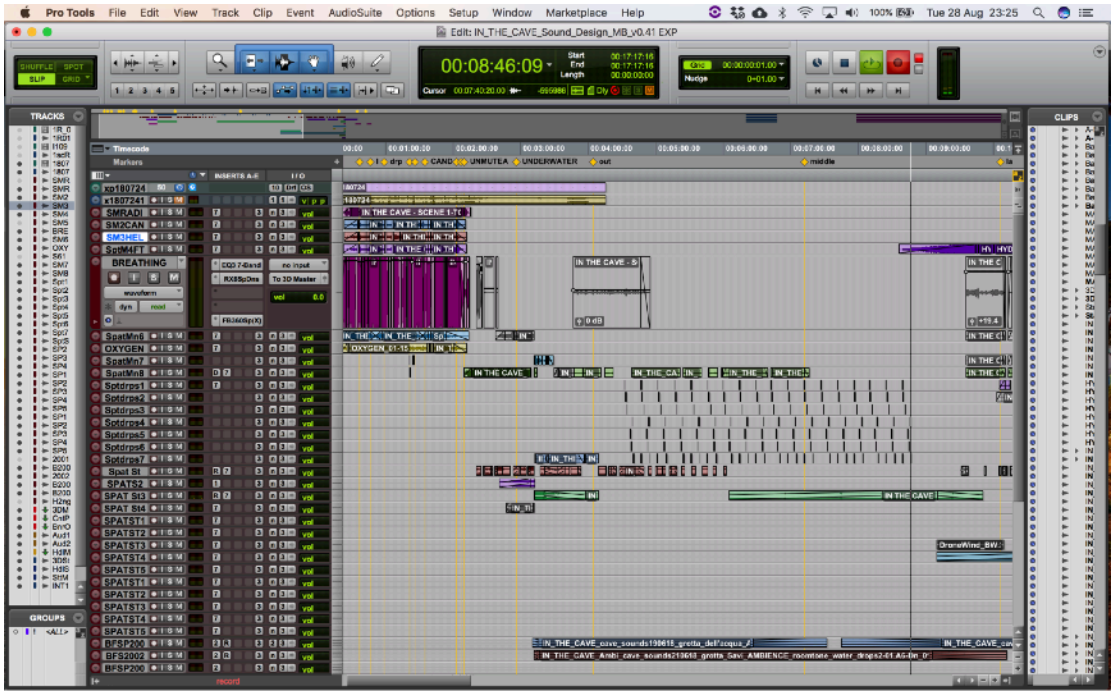


Ambisonics formats are diverse. Each format has its own peculiarity, but as the Ambisonics techniques are used more and more, there are some emerging standards. When working with Ambisonics, the files need to be converted to these standards.

8.2 A TO B FORMAT CONVERSION

A Format and B Format are two analog audio standards that are part of the ambisonics workflow.
A Format is the raw recording of the four individual cardioid capsules in Ambisonics microphones. Since each microphone has different capsules at slightly different distances, the A Format is somewhat specific to the microphone model. (postperspective, 2018)
B Format is the standardised format derived from the A Format. The first channel carries the amplitude information of the signal, while the other channels determine the directionality through phase relationships

between each other. Once you get your sound into B Format you can use a variety of ambisonic tools to mix and alter it. The SoundField ST450II that we used for the opera recordings outputs a B-Format file, which is the format that the Facebook 360 Spatialiser utilises for the processing. The SPS200 that we used for this production is an A-Format microphone, this means that for a binaural rendition it has to be firstly converted to B-Format and then rendered in binaural.



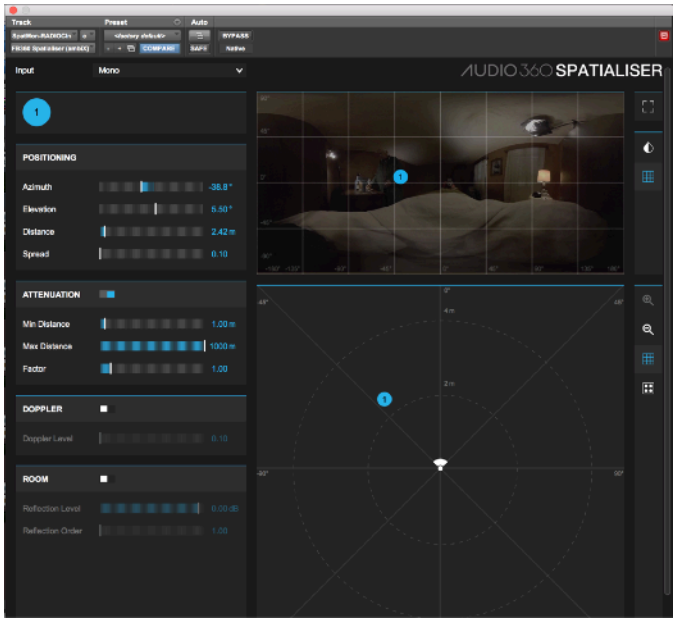
8.3 AMBIX AND FUMA FORMATS

A further difference in Ambisonics formats lies between Furse-Malham (FuMa) and Ambix formats. While Ambix is the most recognised as the new standard for channel order and normalisation protocols, still some plugins, like SurroundZone2, will operate the conversion from A to B format, but outputting a FuMa Ambisonic file. It is vital to the

Ambisonic workflow to keep track of these differences and convert the files accordingly to standards chosen to mix with. A good practice might be to always convert A Format to B format and FuMa to Ambix. In this way it will be obtained a B-Format Ambix standardised file, which is currently the most widely adopted standard.

There are several ways to do the conversion. In this researched were identified two methods that are explained in the Appendix 1.

8.4 AMBISONICS ENCODING



Mono and stereo recordings can be converted into an Ambisonics sphere. This can be performed even if there is no Ambisonics recording to start with. This process is called Ambisonics encoding. For one of the projects of this research, were used only a stereo pair recording and spatialised it in SOA through the Facebook 360 Spatialiser app. Here again working in an Ambisonic field with mono/stereo audio objects resulted the best way to get a more precise localisation, a more processable and mixable material and a simpler approach.

8.5 REDUNDANCY

Not being able to follow the action with a boom or to adjust the mics during the takes or the events that is being recorded, a viable idea is opting for a redundant mic set up. The microphone number was high for two main reasons: firstly to cover up the full range of action, secondly to have a back up recording in case of any mic failure. More

details on type of body microphones and recording techniques can be found in the Appendix 1 in the section Redundancy.

8.6 BODY MICS ARE THE NEW BOOM

Not being possible to use a boom microphone to follow the action a redundant set up of radio microphones was placed on the talents.Of the two body mics, one was transmitting to a radio receiver, the second recording on a small recorder unit placed on the talent. This redundant solution was adopted for back up reasons, in case of radio interferences, signal drops, or clothes rustling. The radio mic attached to the talent and can be used not only to pick up dialog, as there was no dialog in this production, but to record the Foleys live on set. Thus clothing sounds, footsteps, objects manipulation were recorded by the wireless mics placed on the talents' body.

8.7 PLANT MICS

Once the design of the room, the position of the furniture, the blocked actions of the actors, were studied during pre-production, a web of plant mics was designed, positioned in strategic places, hidden to the camera.
The plant mics were DPA 4060, being these microphones easy to hide and with an omnidirectional polar pattern. The 4 DPA 4060 were arranged in the space as can be seen in the picture below.

In the Appendix 1 can be found close up photos of the microphones.

8.8 Impulse Responses

A Convolution Reverb consists of a recorded sample (called an Impulse Response or “IR”) of an acoustic space to excitation from a signal such as a sweep tone, starter gun, or snare drum crack, and the effect on the space of that signal after it has been removed and usably transformed by the convolution processor. Convolution reverbs essentially record and process the reverberant behaviour unique to a real acoustic space. (bhphotovideo).
Impulse Responses of a room can be recorded and used in post-production to apply the same reverberant properties to sound that were recorded in a different location as the sound that were recorded on set.
To achieve a coherent reverberant space, Impulse responses were recorded for all the location of the film for project 3. Techniques to record and implement Ambisonics impulse responses were researched, and through the Ambisonics-related Facebook group was started a thread whom main contributor was Daniel Courville from Canada, an academic working with Ambisonics and developing Ambisonics tools. He drew the guidelines for the process of using Apple Impulse Response Utility. The paper P.A.I.R.S. by Andrew Rahman, in which the author tested some Portable Ambisonics Impulse Responses Systems (P.A.I.R.S.), was another major contribution to develop an effective Ambisonics Impulse Response method. Further research was carried out to find a portable self-powered speaker to play back a sine sweep (creative field recording website).
Multiple impulse responses were recorded both in Ambisonics and stereo. The obtained impulse responses were imported in the convolution reverbs by Altiverb and Logic X.

8.9 HYPERREALISM

This research project evolved from a realistic frame to pursue a hyper-realistic aesthetics. This was not a pre-determined decision, but the outcome of an initial research phase and an unexpected epiphany.

Here are two definitions of Hyper-realism from Wiktionary:

- **Hyperrealism***(Noun)*
  A style in art that attempts to reproduce highly realistic graphic representations

- **Hyperrealism***(Noun)*
  A compositional style defined by Noah Creshevsky as "an electroacoustic musical language constructed from sounds that are found in our shared environment (realism), handled in ways that are somehow exaggerated or excessive (hyper)."

Whether the first definition is related to graphic art, it is still pertinent to the sonic universe. The second definition summarise and explains the process of recording realistic sound in the real world, to subsequently rearrange them in the fictional world of the film. This definition describes exactly the process of field recording and audio editing for film. Kevin Bolen, Audio Designer for Immersive and Interactive Media at Skywalker Sound, points out in Soundworks Collection – VR Sound Design Panel at Pyramid Studios - how the spatial sound designer task is to craft and experience that feels authentic rather than realistic. (Kevin Bolen, 2018).

This leads to the most important consideration we can draw from this experience. In filming a live event we are not pursuing a documentarist, non-intrusive replica of the experience. We are creating a new experience, whose product is the output of the process of recording and post producing the live material we captured. We aim for an enhanced aesthetics. We created a hyperrealistic rendering of the event we recorded, pursuing an experience that feels authentic rather than realistic.
Becoming aware of this concept allowed us to freely reinterpret the material we had and despite not being unfaithful to it, we manipulated it pursuing an hyperrealistic aesthetic ideal.

8.10 CREATIVE SOUND DESIGN

The three projects were approached differently. The 360 Opera and the Ballet are live events recorded and post produced. The epiphany of

adopting a hyper-realistic aesthetics stroke during the post production of these two projects. The decision to use the mono/stereo recordings instead of the Ambisonics, resulted in a complete freedom to place the sound sources in the 360 field. There was an initial plan to use the natural reverb of the room recorded by a room mic set up during the performances. However this option was abandoned in favour of a digital reverb. This decision allowed to re-interpret the space, automating the level of the reverb and adjusting the distance and mix for every single microphone independently. More insights on the use of Reverb can be found in the Appendix 1.

The cinematic VR project, “In the Cave” offered up other challenges. Being a film, the sound was post produced as a traditional film, editing the recordings, polishing the files, recording and cutting the Foleys, adding FXs and eventually mixing. However, being a 360 film a whole new set of challenges were presented. The mono/stereo sources were spatialised in three dimensions and the microphones placed in the actors were automated to follow the action in a three dimensional sphere. Please consult the Appendix 1 for further considerations on the editing and polishing of the audio files.

The Ambisonics recordings were used only as a guide, there are technical reasons for this that are discussed in the Appendix 1. Thus the whole scenes were re-designed, in the pursuit of a hyper-realistic aesthetics. The first person perspective was rendered through the breathing sound, spatialised in the Zenith of the sphere. All the sound elements of the first scene were deliberately introduced one by one, gradually, to suggest an oneiric experience, lying halfway between reality and dream.

An extensive part of the sound design was based on the recordings made in real caves and on the underwater recordings made with hydrophones. There are some technical and artistic insights on these aspects in the Appendix 1.

The hyper-realism was abandoned in the second scene of the film, which is set inside a Cave. There is a strong symbology of rebirth driven by the water and underwater elements and the cave sounds blend with intra-uterine sounds, symbolising the existence of a foetus in the womb.

These sound elements, like the water drops, the underwater sounds and the medical foetus recordings were processed to craft the final piece of ambiental music that accompanies the credits of the film. A

version of this experimental piece of music can be listened in the 360 website that is set up as a summary of this portfolio. The link to he website can be found in the Output section.

8.11 LINEAR AND INTERACTIVE SOUND DESIGN

Video games are an interactive media and this is going to influence how the sound design work is faced. In traditional linear media, the sound designer has absolute control on what happens in the timeline. VR and 360 film are in between linear and interactive media. There is a degree of freedom that allows the viewer to rotate their view and this affects dynamically the sound mix that is created and rendered binaurally.
“In the Cave”, the cinematic VR that is the object of Project 3 of this research, is a film, but its architecture lies between film and videogame. It has a linear storytelling but with a degree of interactivity. The whole film has an interactive sound design, the interaction being the freedom of rotating the head and accordingly hear the binaural rendering of the Ambisonics sphere rotating as well. This contributes to create the illusion of realism and the suspension of disbelief. A further element to take into consideration in the sound design is: varying distance. To achieve this the game engine offers the possibility of setting up attenuation, eq curves and reverb responding to distance and orientation. So while layers of ambiences were kept constant throughout all the sequence, elements like the waterfall were treated as sound objects in the game engine.
In some parts of “In the Cave” there are playing simultaneously three layers of Ambisonics cave ambiences, plus individual water drops, underwater movements, waterfall, streams, and experimental sounds recorded with hydrophones. Some of the experimental sounds were captured placing the hydrophones on the body, recording blood flowing under the skin and breathing through the throat. Also some medical recordings of a foetus’ heart beat provided by a midwife were implemented. However all these elements were not played at once and were located in time and space, following an interactive sound design that was sitting on a linear storyline.
Further details on the double approach of linear and interactive sound design are explained in Appendix 1.

8.12 DOWN-STEPPING

Down-stepping is the process of converting a Higher Order Ambisonics file (HOA) to a lower order Ambisonics file. This might be needed due to some software or playback platforms not supporting HOA.
In this research the Ambisonics microphones recorded a First Order Ambisonics (FOA) in A-Format onto 4 channels. All the sounds were designed and mixed in Second Order Ambisonics (SOA). From a SOA file of 9 channels, the mixes were down-stepped to FOA for delivery.

8.13 FINDINGS

The whole post-production was carried out on ProTools HD 12.8.3. ProTools HD, supports Ambisonics up to third order. Whether this is sufficient to craft a spatial audio sound design for 360 video and VR, ProTools falls behind its competitors, especially to Reaper, a Digital Audio Workstation capable of supporting Ambisonics up to Seventh Order. Other disadvantages that have emerged in the use of ProTools and Ambisonics is the fewer number of plugins developed for the Ambisonics workflow in the industry standard DAW. A further comparison with the competitor Reaper shows that the majority of Ambisonics plugins are developed in the VST format, but not in AAX format, which is the format ProTools utilises for its plugins. Moreover, being the technology in its infancy, there is a wide community of audio software developers releasing Ambisonics open source plugins in every format but in AAX.
Furthermore, the bundled ProTools plugins, like the reverbs, compressors and limiters do not support Ambisonics. Hence the Ambisonics workflow in ProTools has to rely on third-party plugins which as stated before are in most cases available only for other DAWs.
The implementation of bundled ambisonics reverbs and limiters in the future updates of ProTools, would facilitate the spatial audio sound designers' Ambisonic workflow.

9. DELIVERY

Delivery of VR and 360 videos with Spatial Audio represents the last stage of the VR production chain that was investigated in this research. To the current state of the art there is no format that can be accepted by all the platforms. Every platform supports and requires different formats. In the table below can be viewed some of the formats required.

Output platform	Input (Audio)	head-locked stereo	Focus	Container	Encoder option to use
Facebook News feed+ Facebook 360 app	9 channel 2nd order ambiX	YES	YES	.mp4	Facebook 360
	8 channel TBE	YES	YES	.mp4	Facebook 360
	4 channel 1st order ambiX	YES	YES	.mp4	Facebook 360
Oculus video (Mobile) [side-loaded videos]	9 channel 2nd order ambiX	YES	YES	.mkv	FB360 Matroska
	8 channel TBE	YES	YES	.mkv	FB360 Matroska
	4 channel 1st order ambiX	YES	YES	.mkv	FB360 Matroska
Oculus video (Mobile) [online streaming]	9 channel 2nd order ambiX	YES	YES	.mp4	Facebook 360
	8 channel TBE	YES	YES	.mp4	Facebook 360
	4 channel 1st order ambiX	YES	YES	.mp4	Facebook 360

Oculus video (Rift) [side-loaded video ONLY]	9 channel 2nd order ambiX	YES	YES	.mkv	Rift: Oculus Video
	8 channel TBE	YES	YES	.mkv	Rift: Oculus Video
	4 channel 1st order ambiX	YES	YES	.mkv	Rift: Oculus Video
YouTube	4channel 1st order ambiX	NO	NO	.mp4	YouTube video

The SOA 9 channel file was encoded through the Facebook 360 Encoder. The process of muxing (encoding video with audio and injecting metadata for the correct reproduction), represents a fundamental step. Every platform supporting 360 videos has specific technical requirements and at the moment there is no standard. Facebook can upload and play 360 video with spatial audio up to SOA, while Youtube accept only FOA files. The video requirements are specific as well, but both platforms require a .mov container and don't support h264 compression. It was needed to convert the video files multiple times to be able to play them in ProTools and then to create the best format to encode them in a 360 video.
FFworks, an open source application that allows to decide which video format to output and performs the conversion was used to convert the video files.
The video files thus obtained were encoded with the Spatial Audio and uploaded onto Facebook and and Youtube.
The video files that were uploaded have 8k resolution, however they can be played at lower resolutions depending on bandwidth and internet speed.
On the other hand, the cinematic VR was delivered in FOA for further implementation into Unreal engine. The cinematic VR will be played back on an Oculus headset. After the premiere at La Biennale - Venice Film Festival, it will be available on other platforms. At the moment it cannot be disclosed.

10. CONCLUSIONS

This Final Project focuses on the complete chain of pre-production/ production/post-production-delivery of Ambisonics files for 360 video and VR onto ProTools HD DAW. Encompassing full orchestral live recordings, light and quick set ups, 360 film productions and VR game engine-based creations, this project aims at pointing out some guidelines for the professionals approaching Ambisonics as a recording and delivering format.

While it has been stated that redundancy in microphones number and techniques might be of paramount importance, it has been also explained the complexity of the Ambisonics recordings set ups and processing

Whether Ambisonics is not a new technique, having been devised in the 1970s, the combination of traditional recording techniques, both for music and cinema, and Ambisonics techniques can offer the best solution in terms of flexibility, quality and mix-ability.

Ambisonics recordings bound the mixer to work with a full sphere of sound. While this in an ideal environment can simplify the mixing process, in the heavily post produced world of film, VR and 360 experiences it might lack of malleability.

Nonetheless Ambisonics offers the perfect format to encode mono/ stereo recordings into a full sphere.

The mixture of traditional recording, editing and mixing techniques can be implemented into an Ambisonics field, creating an object-based mix with mono and stereo sources, encoded into an Ambisonics sphere. This allows the sound designer to claim back its creativity, having the freedom to craft with traditional techniques, with the added dimension of a 360 space.

In the pursuit of a hyper-realistic aesthetics for the new immersive media like 360 videos and VR, Ambisonics plays the ideal tool when combined with traditional techniques.

11. OUTPUT

“Brindisi in 360” - La Traviata - Opera in 360 - Youtube link
https://www.youtube.com/watch?v=lCr2EAGFUBA&feature=youtu.be
Facebook link
https://www.facebook.com/allyouneedissound/videos/983675165143223/

“Spirit of Cornwall” - Ballet with live ensemble - Youtube link
https://www.youtube.com/watch?v=V-kNG8G3FGw&feature=youtu.be

“Guarea” - Violin solo live performance - Youtube link
https://www.youtube.com/watch?v=tvGa0k7dvuc&feature=youtu.be

“In the Cave” - Cinematic VR - not disclosable - private link to wip version
https://youtu.be/70R8LcnT_Ak

“Spatial Audio Demonstration” - Youtube link
https://www.youtube.com/watch?v=NOaZLo7ulJU

“Lazarus” - Captain Wilberforce - Youtube link
https://www.youtube.com/watch?v=4CJvU0LpzHc

SUMMARY - 360 WEBSITE WITH THE COMPLETE PORTFOLIO:

here's the 360 website with the links to the videos in the portfolio and the technical documentation.

https://go.wondavr.com/8dQrsqC3LP

There two ways of viewing it:
1. On a desktop/laptop. It works only on Mozilla Firefox browser. Firefox can be downloaded for free here: https://www.mozilla.org/en-US/firefox/new/
Just open Firefox, copy, paste and go to the link. Remember to put your headphones on for the best experience of Spatial Audio.

2. On a smartphone. Download the Wonda VR app from Google Play store or App store. Now put your headphones on. Launch the Wonda VR App. In the menu select Launch Experience > Via QR code > scan the QR code below.

Now you will open the interface. You can select to use Smartphone View or Cardboard View. Both are supported, but

the cardboard doesn't open the external links to the video, so select the Smartphone "magic window" View.
Now you can look around and choose which video to watch. Just click on the link, and the video will be opened in Youtube. Once in Youtube you can select the cardboard/headset view. I recommend this as the experience is really enhanced through the visor. So after playing the video, slip your phone into the cardboard, all the videos have opening credits that allow for few seconds of adjustments. Remember to plug your headphones in to get the best result for the spatial audio.
Enjoy the 360 videos with Spatial Audio!

12.BIBLIOGRAPHY

Akeroyd, Michael A., “Introduction to the Issue on Spatial Audio.” International Journal of Audiology, 14992027, Jul2006 Supplement, Vol. 45

Algazi, V., Duda, R. “Headphone-Based Spatial Sound.” IEEE Signal Processing Magazine, 10535888, 01/01/2011, Vol. 28, Issue 1

Elen, R.,(2001) “Ambisonics: The Surround Alternative—Surround 2001”

Garity, WM. E. and Hawkins, J. N. A. “Fantasound,” Journal of the Society of Motion Picture Engineers 37, no. 8 (1941): 127–46)

Hilton, K. “Old-school surround – the return of binaural” PSN EUROPE, May, 1 , 2014

Hilton, K. ”Sound Now and Next 2015: The future starts here” PSN EUROPE, Jul 13 2015,

Hollerweger, F. - Florian Hollerweger's Website, 2005 - flo.mur.at

Rahman, A. “P.A.I.R.S. - Portable Ambisonic Impulse Response System, An Introduction to Ambisonic Impulse Response Recording”, 2017.

Reilly, E.,2017 ’Headspace VR’ Study Examines Audience and Empathy

Rumsey, F. "Spatial Audio" in , Focal Press, 2001.

Ward, P. “The rise of 3D Audio” - PSN EUROPE APR 19, 2017.

Watson, P. “3D audio: reinventing the wheel” PSN EUROPE, Aug 23 2011

Zhang, W.; Samarasinghe, P.N.; Chen, H.; Abhayapala, T.D. “Surround by Sound: A Review of Spatial Audio Recording and Reproduction.” Appl. Sci. 2017, 7, 532.

13. WEBOGRAPHY

http://www.avidblogs.com/ambisonics-vr360-audio-pro-tools-hd/

http://www.avidblogs.com/ambisonics-vr360-audio-pro-tools-hd/

https://www.pro-tools-expert.com/home-page/2017/3/16/developing-a-3d-audio-workflow-for-virtual-reality-in-pro-tools

https://sonicscoop.com/2018/02/05/audio-mixing-for-vr-the-beginners-guide-to-spatial-audio-3d-sound-and-ambisonics/

https://nofilmschool.com/2016/10/detailed-guide-to-spatial-audio-for-360-video-using-mostly-free-tools

https://endabates.wordpress.com/2017/06/19/comparing-ambisonic-microphones/

http://www.creativefieldrecording.com/2017/03/30/ambisonic-field-recordingresources/

https://www.waves.com/ambisonics-explained-guide-for-sound-engineers

https://evangelriclapore.wordpress.com/tag/dolby-atmos

http://ce-pro.eu/2016/12/2017-auro-3d-dolby-atmos-dtsx/

http://ce-pro.eu/2016/12/2017-auro-3d-dolby-atmos-dtsx/

https://www.asoundeffect.com/ambisonics-primer/

http://www2.grammy.com/pdfs/recording_academy/producers_and_engineers/5_1_rec.pdf

https://www.asoundeffect.com/podcasts-vr-audio-selling-sfx-bass-fear-elephant-sounds-much-more/

https://postperspective.com/vr-audio-differences-format-b-format/

accessed on 21 Aug 2018

http://javierzumer.com/blog/2017/8/9/an-introduction-to-game-audio
Accessed on 28 Aug 2018

Glenn Forsythe, Audio Mixing for VR: The Beginners Guide to Spatial Audio, 3D Sound and Ambisonics, February 5th, 2018 by

14. USEFUL LINKS

www.visualise.com
www.bamsound.com
www.1618digital.com
www.bbc.co.uk/rd
www.sennheiser.com
https://endabates.wordpress.com/2017/06/19/comparing-ambisonic-microphones/
http://www.creativefieldrecording.com/2017/03/30/ambisonic-field-recordingresources/
https://www.waves.com/ambisonics-explained-guide-for-sound-engineers
https://evangelriclapore.wordpress.com/tag/dolby-atmos/
http://www.syntheticwave.de/Two_Ears-Two_Loudspeakers.htm
http://www.holoplot.com/
http://www.yarra3dx.com/technology/
https://en.wikipedia.org/wiki/Stereophonic_sound

15. APPENDIX 1

This Appendix to the research focuses on further technical details and reports issues and solutions encountered during the recording and post processing of the audio captured for the project.

THE 3 IMMERSIVE PROJECTS

All the content produced as output for this project is a combination of video and audio material, the video part being filmed or rendered in 360 and the audio part being rendered with spatial audio.
For the video part I collaborated with students from the Northern Film School and with 360 technology professionals Fluido TV. For the audio part I carried out all the tasks involving sound.
I will refer as “we” when talking about the team that created the output of this research, being this a filmmaking process and being aware that filmmaking is a collaborative effort.
In a time window of 8 months we penetrated the world of spatial audio for immersive video contents, experimented with it, found out three main project to develop for this Final Individual Project, carry out the field research, produced and released the three projects.
An initial experimentation was carried out with limited audio and video equipment, focused on the possibilities of spatial audio post production plugins. Links to this experimentations can be found in the OUTPUT section.
Once we established that we were able to produce 360 content with spatial audio, we identified three projects to develop in collaboration with other professionals inside and outside the University.

The three projects are:
1. “La Traviata” Opera in 360
2. “Spirit of Cornwall” Contemporary ballet with live classical ensemble in 360
3. “In the Cave” Cinematic VR

PROJECT 1
“La Traviata” - Opera in 360

We collaborated with the Conductor David Levi who organised a performance of the Italian Opera “La Traviata” by Giuseppe Verdi. The cast featured professional singers, a full orchestra ensemble of 60 pieces and a chorus of 40 singers.
The identified location was unconventional: Sunny Bank Mills, an ex industrial mill located in Farnsley. The uniqueness of the location offered up the possibility of playing with the positions of the orchestra, chorus, lead singers and director. This resulted in an opportunity for a 360 capture of the event.
The performance took place on the 24th March 2018 at the Sunny Bank Mill, in the room managed by Trouble at Mill.

Pre-production
We scouted the location three times; the first two times to check the light conditions and the acoustics of the space, the third time we went to a concert to capture some test videos.

We agreed on the placement of the stage with the conductor and started working on a non traditional staging. Whether there was still the audience facing the stage, the number of instrumentalists, chorus and lead singers was exceeding the figure of 120 and forced us to spread the stage line across the whole length of the room.
This resulted in a interesting placement of the 360 camera rig, but represented a challenge for the sound department. Luckily we had three days of setting up and prepping the stage, allowing us to experiment with mic typology and placement.

Kit List

Video
360 cameras:
GoPro Omni rig (6 GoPros in a hexagonal array)
Ricoh Theta S

Flat cameras:
2x Sony HXR-MC50
1x Canon 700D DSLR

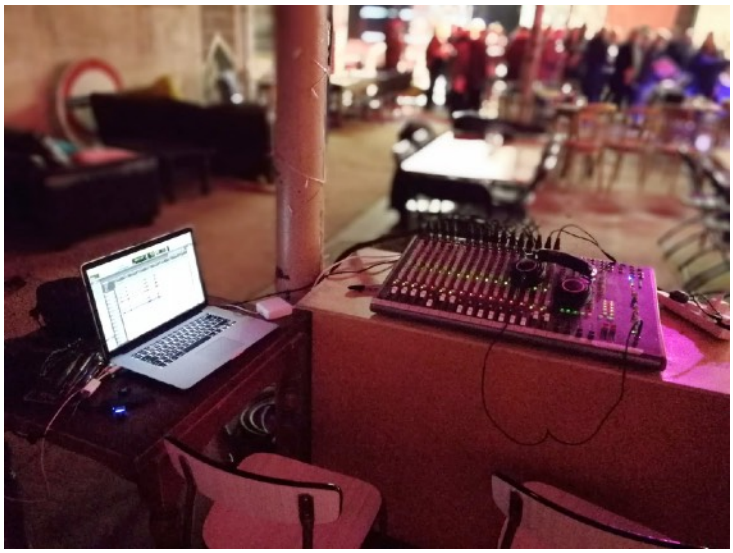
Audio
1x Allen & Heath Zed-r16 - 16 Channel FireWire Mixer
2x Audio Technica AT2020 Microphone
2x Audio Technica AT2035 Microphone
2x Audio Technica AT825 Stereo Microphone
2x Audio Technica AT835B Shotgun Microphone
6x Beyerdynamic MCE86 Shotgun Microphone
2x Audio Technica AT2050 Microphone
3x DPA 4060 Omni Microphone
1x Sandisk 32GB SD Card
1x Soundfield ST450 Microphone Kit
1x Zoom F8 Recorder
1x Sennheiser HD439 Headphones
1x Audio Technica ATH -m50x

Mic stands, stage boxes, multicore cables, XLR cables, cable shields, power extensions.

Recording on ProTools HD 12.8.3 on a MacBook Pro Retina late 2015

The audio set up was a mix of directional mono and stereo mics with an Ambisonic mic, the SoundField ST450 II.
The mixing desk was limited to 16 channels, recorded straight onto ProTools HD via FireWire from the Zed-r16 mixing desk. All these channels were implemented for mono and stereo microphones. To work around the limited number of channels we recorded the Ambisonic on a separate Zoom F8 audio recorder. Whether this represented a challenge for time alignment with the other 16 channels recorded straight onto ProTools, it allowed us to properly set the gains of the 4 capsules of the Ambisonic mic. To get a signal equal in gain from each capsule is a key factor when recording with Ambisonics set ups. If the levels of the 4 channels are not matched up, the spherical aural image we capture will be distorted. The latest firmware of the Zoom F8 provide an Ambisonic mode, that links up and locks the first 4 channels of the recorder, allowing to control all of them by moving just the knob on channel 1. The F8 wasn't updated so we downloaded and installed firmware 4.0. This firmware supports Ambisonics recordings allowing the user to

select if the recording is made in A or B format and if it is in FuMa or Ambix channel order.



MICROPHONES

Mono/Stereo Mics

The main goal of recording the full performance of the opera was to capture the lead singers in the best possible way, in the most possible isolated way from the orchestra. This would have given a higher degree of freedom and precision when coming to the mixing stage.

The use of individual radio mics would have provided the recording most isolated from the background and would have also been an easy solution to avoid directionality problems, when the singer would have sung off-axis. However the singers didn't want to be closely mic'd and radio mics are not reliable as frequency interferences and signal drops might occur.

Thus we consulted Stephen Hawkins, the sound engineer from Opera North, whose expertise in recording opera singer and orchestral ensemble was pivotal in designing our final microphone set up.

Due to the space configuration it was eventually decided to act the opera in a semi-dramatised version, with the singers having music stands to fix their positions and acting just the main actions of the drama. This resulted being really helpful to the sound department, facilitating the mic placement.

We placed a row of 6 highly directional shotgun mics to have the entire front stage covered, one mic for each music stand. These mics were on short mic stands, pointing upward towards the singers' mouth. In this way we wanted to intercept the singers' voice with the least possible background sound.

We added two shotgun mics on the sides, facing inward to capture any performance off-axis. Additionally we hanged an omni mic in the centre of the stage to capture a mix of the orchestra and singers singing inwards.

The orchestra played right behind the singer, we placed two wide cardioids on tall stands, facing downwards to capture the high strings and the low

strings, and a large diaphragm omnidirectional mic to capture the woodwinds, right in the middle of the section.
The percussions were loud enough to bleed into all the microphones. In general the whole orchestra was bleeding in every mic, due to frequency range of the mics and reverberations of the room.
The chorus was captured using just one large diaphragm wide cardioid on a tall stand.
A stereo XY microphone on a stand was placed next to mixer to capture the natural room reverb.

Ambisonics mic
The SoundField ST450II was placed right in front of the conductor.

It was as close as possible to the 360 camera. We found out this is a good practice, the 360 video and ambisonics audio perspective should be as coincident as possible, to record a faithful correspondence of image and sound. Nonetheless the acoustics of the place have to be taken into consideration, so whether is convenient to hide the ambisonics mic in the blind spot underneath the camera, this solution was not viable, as the camera was too low to have enough space for the mic. For this we tried to place them at the same height.
Is important to remember that the ST450II requires its own preamp, that coupled with the F8 recorder can result quite bulky. In this occasion we had the chance to hide it underneath the conductor's stage, but in an open field situation it might be more agile to use the SoundField SPS200, as we have done for the other two projects.

Production
We recorded the full opera with the Ambisonic mic, the mono/stereo mics and the flat cameras. We recorded some excerpts with the 360 GoPro Omni rig.

Post production
The video was post produced on Premiere Pro 2018, DaVinci Resolve and Autopano.
The audio was post produced on ProTools HD 12.8.3.

Phase issues
The main issue of working with multiple formats and microphones was phasing. The 16 channel need to be time aligned, we used Melda Productions AutoAlign, but this needed to be tweaked by hand. We have to consider that the full opera ran for 2 and a half hours, and that the singer were singing on different mics at different times. Specifically the task to cancel the phase issues between the mono/stereo mics and the Ambisonic mic was the most time consuming. A basic delay of the tracks to match the transients didn't work, as in the session there were times were the Ambisonics were earlier and other when were later.
We tried to apply the rule of delay compensation based on the simple formula that: delay=distance(metres)/c(speed of sound), but this didn't work consistently for the whole length of the recording. (Angelo Farina's post on Facebook Spatial Audio for VR Group)
We guess that the formula would work in a best case scenario, were the sound sources (singers) sing towards the microphones, so that we have to align just the direct sound. But in this case, many times they were singing facing away from the ambisonic mic, resulting in the direct sound being lost in the early reflections. If this would have been a phenomenon happening consistently it could have been fixed, however, due to the complexity of the set up and the randomness and unpredictability of the singer's movements on and off axis, it was impossible to have a perfect match of the transients.
The solution was to edit the Ambisonic file and nudge by hand the files to match the transients.

Polishing
Being a live performance, along with the music we captured all the unwanted noises occurred during the evening. Among these the most unpleasant were the creaking chairs of the chorus and orchestra. We edited them out when possible, attenuated some frequencies and applied multiple passes of de-clicking from the Izotope RX6 Advanced suite, to smooth them as possible.

Reverb
The room microphone resulted useless as it captured too much of the audience movements compared to the room natural reverberation. Instead of using it we opted for a digital reverb.

Hyper-realism
This leads to the most important consideration we can draw from this experience. In filming a live event we are not pursuing a documentarist, non-intrusive replica of the experience. We are creating a new experience, whose product is the output of the process of recording and post producing the live material we captured. We aim for an enhanced aesthetics. We created a hyper-realistic rendering of the event we recorded.
Becoming aware of this concept allowed us to freely reinterpret the material we had and despite not being unfaithful to it, we manipulated it pursuing an aesthetic ideal.

Following this conceptual milestone, we went back to the recordings and listened. We realised that many times the ambisonics and mono/stereo recordings were conflicting. We decided that it sounded brighter and clearer without the ambisonics files.
Once we muted the ambisonics tracks we weren't constrained to the spatial coordinates dictated by the ambisonic mic. Thus we spatialised the mono and stereo recordings as sound objects to match the position of the sound sources in the video, mixing it with the amount of digital reverb we preferred. We used the Facebook 360 Spatialiser for the 360 panning, RevibeII for the reverbs, and ProTools Eqs to sculpt the frequencies.
The outcome was a Second Order Ambisonics mix, but only created with mono and stereo recordings, encoded in an Ambisonic sphere.

Delivery
The SOA 9 channel file was encoded through the Facebook 360 Encoder. The process of muxing, encoding video with audio and injecting metadata for the correct reproduction, represents a fundamental step. Every platform supporting 360 videos has specific technical requirements and at the moment there is no standard. Facebook can upload and play 360 video with spatial audio up to SOA, while Youtube accept only FOA files. The video requirements are specific as well, but both platforms require a .mov container and don't support h264 compression. We have converted the video files multiple times to be able to play them in ProTools and then to create the best format to encode them in a 360 video. We used FFworks, an open source application that allows to decide which video format to output and performs the conversion.
We uploaded a SOA version on Facebook and a FOA version on Youtube.
The video file we uploaded is 8k resolution, however it can be played at lower resolutions depending on bandwidth and internet speed.

PROJECT 2
"Spirit of Cornwall" - Contemporary ballet with live ensemble

We collaborated with Constella OperaBallet, a charity enterprise aimed at exploring new territories in ground-breaking collaborations which engage with and celebrate the cultures of diverse communities. Constella comprises some of the UK's finest young artists, who also work with companies including the Royal Opera House, English National Ballet, New York City Ballet and Vienna State Theatre.
Constella was involved in the multidisciplinary collaboration, culminated in the presentation of the garden installation called "Spirit of Cornwall" at the Chelsea Flower Show 2018.
The garden won a Silver-Gilt medal at the RHS Chelsea Flower Show 2018.
Embracing garden design, music, sculpture and architecture, a central feature of the garden were the performances of a composition inspired by Barbara Hepworth's work and commissioned by The Tate St Ives.
Six musicians and a dancer from Constella Opera Ballet gave three performances of the piece at the Chelsea Flower show, conducted by Leo Geyer, with choreography by Sarah Louise Kristiansen.
Over a million people engaged with Constella' s work at the Chelsea Flower Show 2018. Through broadcasts on TV and radio, its on-line audiences comprise 2 million people.
We recorded in 360 with Ambisonics and stereo microphone techniques the performance.

Preproduction

There was no opportunity to scout the location before the performance day, being the garden itself been constructed for the show. We had a 3d rendering of the space that we studied beforehand.
On the day of the performance due to the strict safety policy, we weren't able to set up our equipment until the 15 minutes before the show had begun. Considered this condition we decided to reduce the kit to the minimum. We knew that the six musicians were: singer, violin, viola, cello, harp and conductor.

Kit List
Video
360 cameras:
GoPro Omni rig (6 GoPros in a hexagonal array)

Flat cameras:
2x Sony HXR-MC50

Audio
SoundField SPS200
Samson CO2 stereo pair
Zoom F8

Production
We had 15 minutes of time to set up the cameras and the mics and there was no time for a sound check. The garden had shape reminding of an 8, in one circle there was the dancer, in the other there was the ensemble. The audience was in front of the dancer. We placed the 360 camera in the middle of the 8, to be equidistant from the dancer and the ensemble. The 2 flat cameras at the front, one wide, one zoomed in. The important audio sources were coming only from the ensemble, so we placed our mics close to them. We placed the ambisonics SPS200 near the strings and voice rather than close to the 360 camera. We used the stereo pair of cardioid microphones in a XY array, close to the sound sources. All the cameras and the microphones and recorded were running on battery power, not being a power main socket at disposal.
There was an artificial waterfall right next to the strings, this was part of the environment but we tried to get the recordings as isolated from the waterfall as we could. The performance lasted for approximately 15 minutes.

Post Production
The video was post produced on Premiere Pro 2018, DaVinci Resolve and Autopano.
The audio was post produced on ProTools HD 12.8.3.

A to B format conversion
A Format and B Format are two analog audio standards that are part of the ambisonics workflow.
A Format is the raw recording of the four individual cardioid capsules in ambisonics microphones. Since each microphone has different capsules at slightly different distances, the A Format is somewhat specific to the microphone model.
B Format is the standardised format derived from the A Format. The first channel carries the amplitude information of the signal, while the other channels determine the directionality through phase relationships between each other. Once you get your sound into B Format you can use a variety of ambisonic tools to mix and alter it.
The SoundField ST450II that we used for the opera recordings outputs a B-Format file, which is the format that the Facebook 360 Spatialiser utilises for the processing. The SPS200 that we used for this production is an A-Format microphone, this

means that for a binaural rendition it has to be firstly converted to B-Format and then rendered in binaural.

A further difference between Ambisonics formats is between Furse-Malham (FuMa) and Ambix. While Ambix is the most recognised as the new standard for channel order and normalisation protocols, still some plugins, like SurroundZone2, will operate the conversion from A to B format, but outputting a FuMa Ambisonic file. It is vital to the Ambisonic workflow to keep track of these differences and convert the files accordingly to standards chosen to mix with. A good practice that we worked out is to always convert A Format to B format and FuMa to Ambix. In this way we will operate with a B-Format Ambix standardised file, which is currently the most widely accepted standard.

There are several ways to do the conversion. We tested two: one is through the plugin SurroundZone2 by SoundField, which allows to convert the format and to virtually adjust the orientation of the microphone.
The other one is through an instance of the Facebook SPatialiser plugin opened on a quad-track with the A-Format file. There's a drop down list when can be selected the input format and then the output format will automatically be a B-Format.

Contrary to the phase issues occurred in the Opera recordings, the ambisonic mic and the xy stereo pair were in synch.

Voltage inconsistency
We started polishing the audio files and noticed how in the Ambisonic quad-channel track there were some interferences noises. We cleaned them out as best as possible with eq, de-crackle and de-rustle Izotope Rx plugins. We subsequently tried to diagnosed the origin of the noise. Talking with researcher and lecturer Bob Birch we realised it might have depended on the fact that running on internal AA battery power there might have been some incidental voltage drops, resulting in a crackling noise.
Another possible cause is the multicore cable, and the 10 pin connector that might be damaged. Using the same microphone for Project 3, I came a similar issue, checked the pin in the connector and found that some were bended. After fixing it, the noise disappeared.
However in this case the noise was recorded on file, so the only option we had was trying to reduce and mask it. Playing in the mixing stage we managed to bring it to an acceptable level, not noticeable unless you are aware there's a noise there.
Another good practice we learnt is to use powerful external batteries, like NP lithium-Ion via hi-rose pin connector or tracer batteries Lithium-polymer via a custom made cable.

Floor noise
Another unwanted sound element we had to work around was the noise of the shoes of the singer, moving on the concrete base. Being the floor of the garden made of concrete with some dirt on it, every time that the singer moved, slipped or dragged her feet, it resulted in a high frequency noise, cutting through all the frequencies and being really audible. Even though this sound was part of the performance, for the quest for hyper-realism we decided to attenuate trough some eq. This problem was the direct consequence of the lack of a sound check, and of the short time to set up the kit. The microphones had to be low to no be intrusive in the video, but the compromise that has to e accepted in this case is that a microphone close to the floor will pick up some floor noise as well. A microphone with a narrower polar pattern would have avoided in part the problem, but it would have been too narrow to capture the ensemble widely.

The waterfall
The waterfall was an element present physically in the space, so it captured on video and on audio as well. Even though we pointed to the organisers that the water was there and had a sound, so it would have ended up in the recordings, they asked if we could remove it.
We reduced the sound presence of the waterfall through eq and de-noising techniques, nonetheless it was impossible to completely cancel it from the sound wave without compromising the music. This was particularly evident in the Ambisonics recordings, as the SPS200 is an omni microphone, so is not directional and even if it was close to the instruments, on the other side it was facing the waterfall.
This condition led us to the conclusion that the best option we had was to mute the Ambisonics tracks and work only with the stereo XY pair.
The XY pair has a cardioid polar pattern, so it didn't record the direct sound of the waterfall that was flowing at its back.

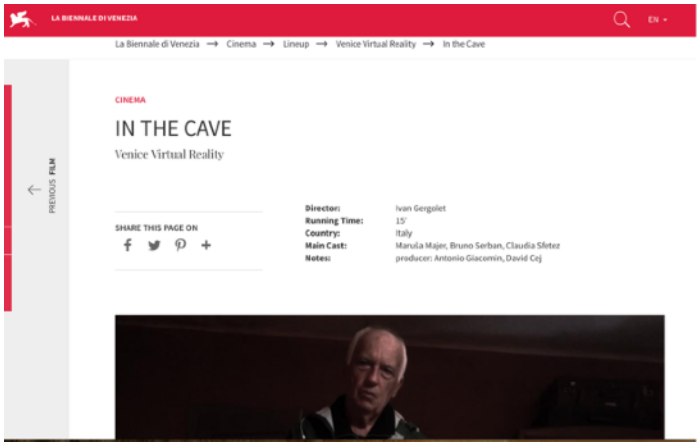
Stereo to Ambisonics
We used only the stereo pair recording and spatialised it in SOA through the Facebook 360 Spatialiser app. The two mics were facing one the singer and the other the ensemble, so split the stereo and treated them as two mono sources, changing the levels and plying with the width of the stereo image to better match the video position of the elements. Here again working in an Ambisonic field with mono/stereo audio objects resulted the best way to get a more precise localisation, a more processable and mixable material and a simpler approach.

Delivery
The video was delivered as an 8k FOA video on Youtube, encoded through the Facebook 360 encoder.

PROJECT 3
In the Cave - Cinematic VR

In the Cave is a cinematic VR, mixing filmed and computer generated video contents, by director Ivan Gergolet and produced by Antonio Giacomin.

In a cave full of beauty and mystery, nothing is what it seems. Exploring its wonders and looking for a way out of the dark, the viewer will be at the centre of an allegorical journey in search of life and rebirth.



The third project we took part to is the cinematic VR “In the Cave”. The film will premiere at the Biennale - Venice Film Festival and is one of the three project selected worldwide in the Biennale College VR. The cinematic VR features both 360 video and CGI. However the computer graphic generated section is completely rendered with the photogrammetry technique, heightening the realism of the visuals. From the audio point of view this project was really complex and was carried out in different phases.
Whether the project hasn’t premiered yet and we are bound to a non disclosure agreement, we can explain the technical challenges and the techniques we implemented for the realisation of the project.

Pre-production
The film is divided in three parts. Part 1 and 3 are set in the same room, Part 1 taking place 70 years later. Part 2 is set into a cave recreated through the use of photogrammetry.
Part 1 and 3 were actually filmed with a 360 stereoscopic camera, while Part 2 was recreated in post, compositing thousands of photographs shot in many caves.
I took on the roles of sound producer, sound recordist, sound editor and sound designer.
We read the script and compiled a sound list with all the possible sounds we needed to capture either on set or in a separate recording session.
There were three days of setting up and testing before shooting Part 1 and two days of preparation and testing before shooting Part 3. Even thought Part 1 is 1 minute and 30 seconds long and Part 3 is just about 30 seconds, this long preparation was necessary due to the testing that needed to be carried out and to adjusting all the details. The nature of filming in 360 is peculiar as there is no blind spot, no behind the camera, everything is on screen at the same time. Moreover there is no editing to help focus the attention and cover up mistakes, so the nature of the takes is more performative than cinematic, this meaning that is more like rehearsing and acting a theatre play rather than traditional cinema.
For these reasons the lighting and audio department need to study the location and plan beforehand the lights and mic placement. Being everything in frame, or better, being there no frame, but a total vision, there is no room for a boom pole, nor for a mic stand. In this scenario the production design department might be the best ally, providing props and furniture to hide plant mics.
The Preproduction phase for this project was the most crucial phase and it took 5 days in total, compared to a two days shooting.

This precious time allowed us to hide the mics and test the results before the shootings.

Kit List
Video
Kandao Obsidian 360 stereoscopic camera

Audio
SoundField SPS200 Ambisonics mic
Zoom F8 recorder
NP batteries with shoe adapter and h-rose connector
Sennheiser MKH 60 with blimp
4x DPA 4060 omni directional microphone
3x Sennheiser G3 transmitter-receiver with lav mics
3x Zoom F1 body recorders with lav mics
2x H2 Aquarian Hydrophone mics
JRF Hydrophone
Audio-Technica AT825 stereo XY microphone
Zoom H2n recorder
MM BSM-9 Binaural mics
Roland CS10em Binaural mics
Ursa chest straps, Rycote Overcovers, Undercovers, Batteries, XLR cables, Mic stands

Redundancy
Not being able to follow the action with a boom or to adjust the mics on the fly, we opted for a redundant mic set up. The microphone were redundant for two main reasons: firstly to cover up the full range of action, secondly to have a back up recording in case of any mic failure.

Body mics are the new Boom
Not being possible to use a boom to follow the action we managed to place on the talents two body mics, one transmitting to a radio receiver, the second one recording on a small recorder unit placed on the talent. This redundant solution was adopted for back up reasons, in case of radio interferences, signal drops, or clothes rustling. The radio mic is attached to the talent and can be used not only to pick up dialog, as there is no dialog in this production, but to record the Foleys right on set. Clothing sounds, footsteps, objects manipulation were recorded by the wireless mics placed on the talents' body.

Plant mics
Once we studied the design of the room, the position of the furniture, the blocked actions of the actors, we devised a web of plant mics, hidden from the camera, positioned in strategic places.
As plant mics e used DPA 4060, being these microphones easy to hide and with an omnidirectional polar pattern. We had 4 DPA 4060 arranged in the space as can be seen in the picture below.

One DPA was dedicated to capture footsteps and the creaking of the wood floor, one placed on the piece of furniture where the actor picks up a helmet, one was placed on the bedside table where the actor lights up a candle with a lighter and the last one was placed on the oxygen concentrator.

The SPS200 Ambisonic mic was placed behind the bed frame. This was the closest possible point to the camera, taking advantage of the furniture to hide the mic.

The Kandao Obsidian has a slot on top of the camera to allocate a Zoom H2n that can be remotely controlled by the camera remote controller. We placed the Zoom H2n on and recorded that as well. Even though it only records FOA only on the horizontal plane, we used it in post to synch the other mics to the take. In general the H2n recordings were usable, but sounded thinner when compared to the SPS200. This is probably due to the larger size of the capsules that the SoundField mic sports.

Production

All the cables were hidden behind the furniture.
Part 1 and Part 3 were shot on two separate days. While for Part 1 the sound recordist was hiding under the bed, for Part 3 there was no place to hide, and the sound recordist would start recording and then leave the room.

Foley sessions
We captured as many Foley sounds on set, this included footsteps, creaking floor, door, oxygen concentrator, binaural ADR breathing sounds. We had to wait for the crew to leave the set, so we worked in the night.

Ambience sounds
We recorded Ambisonics ambiences in the room were the film was shot, in the surroundings and in other rooms. We captured church bells and a full storm, that made it to the final edit. We will explain the processing in the post production chapter.
Water sounds
A prominent role in the sound design of “In the Cave” is played by water sounds. All the transitions between Part 1, Part 2 and Part 3 are based on water sounds and water sounds manipulation. In Part 2 the viewer is underwater and then explores the meanders of a cave.

Water sounds
We recorded a library of liquid and underwater sounds with the MKH 60 and with the two H2 hydrophones. We performed water movements, splashes, bubbles, water drops. We found that underwater sounds are generated by the movement of air bubbles coming to the surface. To obtain more bubbles we sprayed gas underwater and record the result.

Cave sounds
Caves have peculiar sounds, with reverberant dripping water and underground streams. We organised two expeditions in two different caves to collect a variety of cave sounds.
The two caves sounded differently, due to their unique conformation and a different presence of water. We recorded cave ambiences with the SPS200, the MKH60, the AT825 stereo mic, and the binaural mics placed on human heads.

Impulse Responses
We recorded impulse responses for all the location of the film. We researched techniques to record and implement Ambisonics impulse responses, and through

the Ambisonics-related Facebook group got in contact with Daniel Courville from Canada, an academic working with Ambisonics and developing Ambisonics tools. He guided us through the process of using Apple Impulse Response Utility. We studied the paper P.A.I.R.S. by Andrew Rahman in which the author tested some Portable Ambisonics Impulse Responses Systems (P.A.I.R.S.). Further research was carried out to find a portable self-powered speaker to play back a sine sweep (creative field recording website).
We decided to recorded multiple impulse responses both in Ambisonics and stereo for the room where Part 1 and 2 take place and for the caves we explored. We played back from an Anker Soundcore2 bluetooth speaker two sine sweeps developed by AudioEase for Altiverb reverb. Using the Zoom F8 soundcard capabilities, we plugged the F8 via usb to a laptop and record the sine sweeps reverbs straight into the computer. This allowed us to have a portable self powered impulse response system supporting Ambisonics.
The impulse responses obtained were imported in Altiverb and Logic X reverbs.

Post Production
“In the Cave” was post produced in collaboration with Berlin base VR post production company Another World VR. The director and producer came in contact with Another World VR after watching their work at Berlin Film Festival.
“In the Cave” mixes 360 stereoscopic video with CGI. To implement the two formats the whole VR runs on Unreal game engine.
The same approach is valid for both video and audio department. The video in Part 1 and Part 3 are played back by the game engine, while Part 2 is structured and rendered by the game engine. This is not a 6 DOF (6 degrees of freedom) experience where the user can walk in a space, but is an interactive 3 DOF, allowing head rotation on the three axis. Whether in the video filmed sections the sound mixed in the Ambisonics sphere provides the audio material that is then rendered to binaural when the user rotates his head, in the game engine part the approach is slightly more complex. Unreal support FOA, but works with sound objects as well placed on specific coordinates in the space. So when the mix in Part 1 and Part 3 is baked into the Ambisonics file, in Part 2 it is generated by the interaction of the viewer with the world. In this project the interaction is limited only to head rotation. We edited, designed, mixed and encoded part 1 and 3, while for Part 2 we edited and designed the individual sounds for the cave.The sounds were then implemented in the game engine by fellow sound designer Takuro Sakamoto from Another World VR.

Ambisonics
Once the take were selected we used the Zoom H2n recordings to synch all the tracks recorded on the F8, and onto the portable recorders. One issues we had to solve was rotating the Ambisonics field recorded by the SPS200 to match the position of the other microphones. Unfortunately there is no visual reference to do this, and we relied only on our ears. We used the Facebook 360 Spatialiser to rotate the Ambisonics sphere on the Yaw axis, until we get a satisfying result. Prominent, high-pitched sounds can be very important to help achieving the correct rotation. The A-format recording was then converted to B-format and from FuMA to Ambix standard with an instance of SurroundZone2 plugin and Facebook 360 Spatialiser. Once all the microphones were in synch and spatialised we started polishing the recordings and automating the position of the radio mics, accordingly to the movements of the talents. We used Harper X to extract a stereo file from the Ambisonics recording of the storm and then positioned the stereo recording in Ambisonics sphere.
Redundancy was helpful to find the right spatial proportions between microphones and have all the areas covered. Nonetheless having 2 Ambisonics recordings (H2n and SPS200) plus all the plant mics and more Foley and FXs, resulted in a mix too dense. Furthermore, when listening to both the SPS200 and the plant mics, the localisation resulted more blurred. We decided to mute the Ambisonics track of the take, while keeping some layers of Ambisonics room tones, and to mix only with the

mono plant mics and the radio mics, plus mono/stereo sound objects. The mixed resulted clearer and more spatialised.

Linear and Interactive sound design
Video games are an interactive media and this is going to influence how you face the sound design work. In traditional linear media, you absolutely control what happens in your timeline. “In the Cave” is a film, but its architecture lies between film and videogame. Has a linear storytelling but with a degree of interactivity. The whole film has an interactive sound design, the interaction being the freedom of rotating the head and accordingly hear the binaural rendering of the Ambisonics sphere rotating as well. This contributes to create the illusion of realism and the suspension of disbelief. In part 2 the viewer not only can rotate his head, but is also being carried on an invisible path, developing inside the cave. This adds a further element to take into consideration in the sound design: varying distance. To achieve this the game engine offers the possibility of setting up attenuation, eq curves and reverb responding to distance and orientation. So while we kept layers of ambiences constant throughout all the sequence, elements like the waterfall were treated as sound objects in the game engine.
For Part 2 we exported three layers of Ambisonics cave ambiences, plus individual water drops, underwater movements, waterfall, streams, and experimental sounds recorded with the hydrophones. Some of the experimental sounds were captured placing the hydrophones on the body, recording blood flowing under the skin and breathing through the throat. We also used some medical recordings of a foetus’ heart beat provided by a midwife.
We crafted an Ambisonics mix of all these elements, based only on a timeline, as the visuals weren’t ready yet. This mix served as a guideline for the implementation into Unreal.

For the credits we composed an experimental music using the water drops recorded in the caves.

Down-stepping
We recorded with a FOA in A-Format, designed and mixed in SOA and then down-stepped to FOA as Unreal currently supports only First Order Ambisonics.

Delivery
The FOA mixes and sound objects were implemented into Unreal. The cinematic VR will be played back on an Oculus headset. After the premiere at La Biennale - Venice Film Festival, it will be available on other platforms. At the moment it cannot be disclosed.

CONCLUSIONS
This Final Project focuses on the complete chain of preproduction-production-postproduction-delivery of Ambisonics files for 360 video and VR. Encompassing full orchestral live recordings, light and quick set ups, 360 film productions and VR game engine-based creations, this project aims at pointing out some guidelines for the professionals approaching Ambisonics as a recording and delivering format.
While it has been stated that redundancy in microphones techniques and number might be of paramount importance, it has been also explained the complexity of the Ambisonics recordings.
Whether Ambisonics is not a new technique, having been devised in the 1970s, the combination of traditional recording techniques, both for music and cinema, and Ambisonics techniques can offer the best solution in terms of flexibility, quality and mix-ability.
Ambisonics recordings bound the mixer to work with a full sphere of sound. While this in an ideal environment can simplify the mixing process, in the heavily post produced world of film, VR and 360 experiences it might lack of malleability.

Nonetheless Ambisonics offers the perfect format to encode mono/stereo recordings into a full sphere.
The mixture of traditional recording, editing and mixing techniques can be implemented into an Ambisonics field, creating an object-based mix with mono and stereo sources, encoded into an Ambisonics sphere. This allows the sound designer to claim back its creativity, disposing of the freedom to craft with traditional techniques, with the added dimension of a 360 space.
In the pursue of a hyper-realistic aesthetics for the new immersive media lie 360 videos and VR, Ambisonics plays the ideal tool when combined with traditional techniques.

IN THE CAVE

a Cinematic VR by I. Gergolet, A. Giacomin, D. Cej

SOUNDLIST

- 1. INT. NIGHT - BEDROOM

Roomtone

Footsteps

Wood creaking

Clothes when the man walks

Rope and helmet foley

Light buzz from abat jour

Oxygen concentrator / drip feed

Candle

Breathing - old and tired, slowly fading, try different rhytms and actors

Undistinguishabl voices from outside the room

Door?

Clock?

Rain outside?

Dripping coat of the man?

- 2. EXT. DARKNESS

Uderwater sounds - Hydrophone

Underwater bubbles

Underwater light switch on

Underwater rope - underwater cables strikes

Underwater breathing - bubbles

Water lapping, movements, emerging

Cave ambience

Dripping water - multiple heights and intensity

Waterfall in the cave

Underground river/ stream

Water springing out of the cave walls, droplets

Headlamp - switch on/off, buzz?

Headlamp - small flame - acetylene torch

Labour sounds

Breathing, panting, grunting

Lullaby

Worldize lullaby and labour sounds in the cave and record back with natural reverb

IR - Impulse responses of the cave diefferent perspectives

Walls getting tighter - skin movements, foley, leather scrunches

Organic sounds, bodily sounds,

Organic tissues movements

Fluid movements, squishes, slime, rip apart

- 3. INT. DAY – BEDROOM

Baby cry

Baby sounds

Blanket/bedsheet sound

Post partum mother: Heavy Breathing

Giggles, joyful

Lullaby

Breathing calming down

Nurse Cloths foleys

Scissors pick up, handling and cutting

Umbelical chord prop suonds

Roomtone

Same room but 70 years before, how does the sound change?

Sunny ext ambiences.

Different abat jour? - sound