Novel approaches to production and post-production of immersive VR/360 audio-visual experiences.

Ronan Breslin, Dr. Jessica Argo and Poli Petrova

SCHOOL OF SIMULATION AND VISUALISATION THE GLASGOW SCHOOL OF ART

Why 360 or VR?

- immersion and presence
- place the viewer in an other world
- allow them to experience an other's perspective
- induce emotions and/or visceral sensations
- replicate reality? or **go BEYOND reality?**

Applications – Interactive (Games/Movies) or **Linear** (Experiential, Storytelling/Cinematic)

- sports (TV, tourism)
- healthcare (relaxation or exposure therapy)
- empathy, social change (UN VR)
- conceptual art and narrative filmmaking – needs simple workflows for a new generation of content creators.

Challenges: technical and aesthetic

- key objective is to create a meaningful, realistic (or hyperreal) soundscape that coincides with the viewer's head movements – crucial for immersion!

Technical

- the point of audition must adhere to the point of view – accurate head tracked spatial audio movements
- processing power for metadata individual object-based-audio sources
- processing power for higher order ambisonics (Lee 2016)
- reverberation zones (for walk throughs, 6 Degrees of Freedom interaction)

Aesthetic

- an authentic rendering of the real-world? Or additional foley, voice-over and music (according to audience expectations from fixed, static film and TV)
- is there too much going on? sound to focus the viewer? "real virtuality" and "salience" (Chalmers 2009 in Grimshaw 2015)
- spatial sounds to push the narrative, induce intense emotions.

Aesthetic opportunities

- VR and 360 film is an emerging medium – we will devise new techniques to elicit intense emotions and visceral sensations - sound will be the main driver of the narrative
- how can we get the audience to focus in on different elements of the 360 visual through audio? (conversely how can the viewers' gaze attenuate the levels of the sounds heard?)
- imagine you are a soldier in a field…

Accessibility for new content-creators

- Two stages of democratization
  - From high end military to technical computer science university departments
  - now the workflow must be accessible to diverse content creators such as artists and filmmakers.
- new workflows should diminish the elitist exclusivity (until recent years VR previously just in military or academic labs)
- artists and filmmakers can now create 360 and VR content, not just for computer scientists. Artists used to be dependent on a technology consultant – now the artist can work autonomously. (e.g. fish in the gallery Sidsel Meineche Hansen, *No Right Way 2 Cum* Transmission 2016 / *Siren Servers* (2017) The Butler Brothers, isodesign, Giles lamb & numbercult).
- "the sound of the smell of my shoes" Grimshaw 2015)

360° Video and Audio

- Not strictly VR but often referred to as such.
- Facebook 360 spatial audio workstation and Google VR spatial audio engine – binaural audio for headphones (Facebook supports 8channel 360 master and a separate stereo head-locked audio encoded with its video)
- Unity (Ffmod and Wwise – event-line similar to timeline in digital audio workstations

There are a range of 360° cameras currently available.

- Consumer/Prosumer: Samsung Gear 360 (£400)
- Mid to High: GoPro Omni (£5k)
- High: Nokia Ozo (£42k)
- FFS!!!: 360 Designs Eye ($250k to $500k)
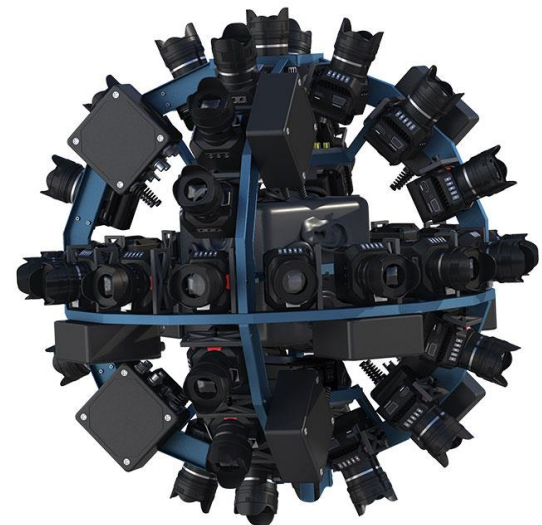
No longer channel based audio – Scene-based-audio and Object-based-audio

- Spatial audio simulates how people hear sound in the real world – sound originates from all around. simulate this by anchoring sounds to objects and positing them in relativity to the camera (Facebook 360 or Ambisonic Toolkit)

- Channel based audio – difficulty adapting to dynamic head tracking, sound coloration

- 5.1 only moves sound along a horizontal plane.

- scene based audio can be captured using a soundfield microphone – orientation rotates according to the viewer's head movements.

- object-based-audio (Dolby Atmos) – each sound object has metadata attached that dictates their position in the space, and their movement behavior and reverberation characteristic, should be mono sources, panned around the "virtual speaker array" (Google VR).

- interaural time, level and phase difference (ILD, ITD, IPD), head-related transfer function (HRTF). reflected and direct sound.

- "physically-based" (scene-based-audio) versus "artistic-based" (object-based-audio) (Altman, Krauss, Susal and Tsingos 2016)

Healthcare

Respite

- a virtual sensorial opposite offers a distracting analgesic for a burns victim who was immersed in a "snow-world scene" (Hoffman et al., 2011)
- *Atmosphaeres* 360 nature, travel and wellness experiences. (Fassbender 2014)

Simulation of PTSD/phobias.

- in exposure therapy a user confronts their fear in a controlled environment (such as augmented reality cockroaches crawling over the user's hand or traumatised war veterans re-enacting a combat situation with a virtual reality visualisation) (Breton-Lopez et. al 2010)
- *Immersive Soundscapes to Elicit Anxiety in Exposure Therapy: Physical Desensitization and Mental Catharsis*

Empathy

- UN VR
  - a refugee camp, guided by a young girl – inspire a virtual connection between policy makers and the displaced
  - ebola-stricken streets, schools, hospital and burial grounds in Liberia

- *Notes on Blindness: Into Darkness (2016),* conveying "a world beyond sight" the imagery produced by attentive listening when deprived of sight.

- *EXIT (2015)* panoramic data visualisation of cultural theorist Paul Virilio Scofidio et al 2015
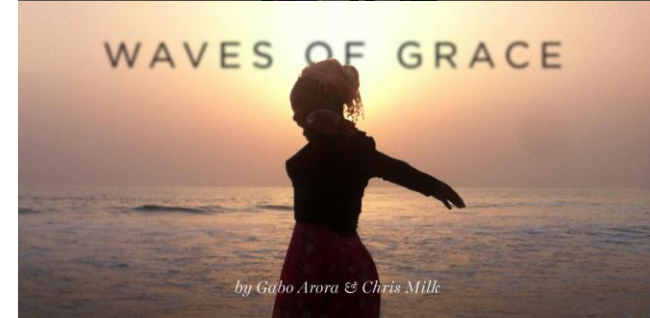
Conceptual Art & Filmmaking

- Char Davies' (1995) *Osmose* a seminal VR artwork - "a space for exploring the perceptual interplay between self and world"

- Avril Furness (2016) *The Last Moments*, Dignitas

- Jayisha Patel (2017) *Notes to My Father,* a story of a human trafficking survivor, places the viewer in a train carriage full of ogling men.
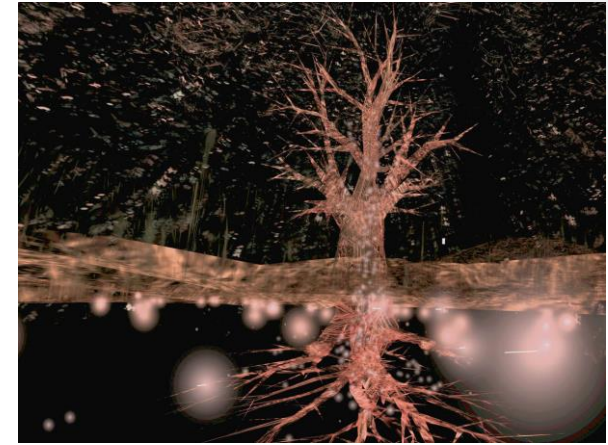
Sports

- Nowadays, sports broadcasts are a huge business (Scuda et al., 2016). These films inspire people to take up a new hobby, learn about the athlete's physical and mental limit, and take the audience to remote places of the world that only a limited few can go to.
- audio for sports videos is produced with regards to the audience's expectations (Scuda et al., 2016).
  - e.g. hear the swish of the net at a basketball game, despite the sound being inaudible to those present at the actual arena.
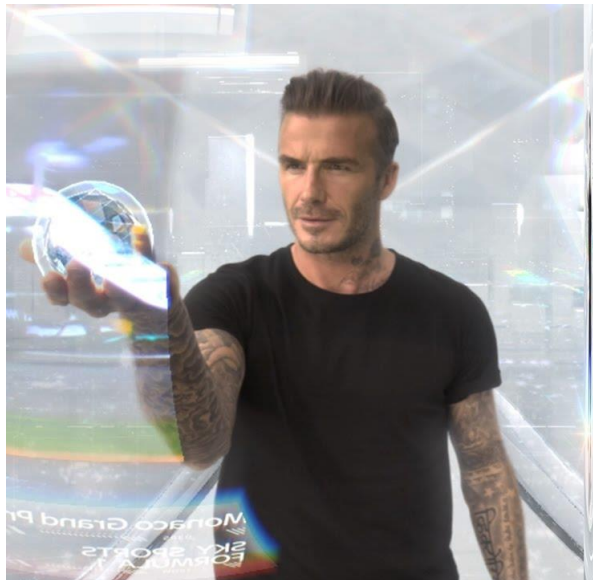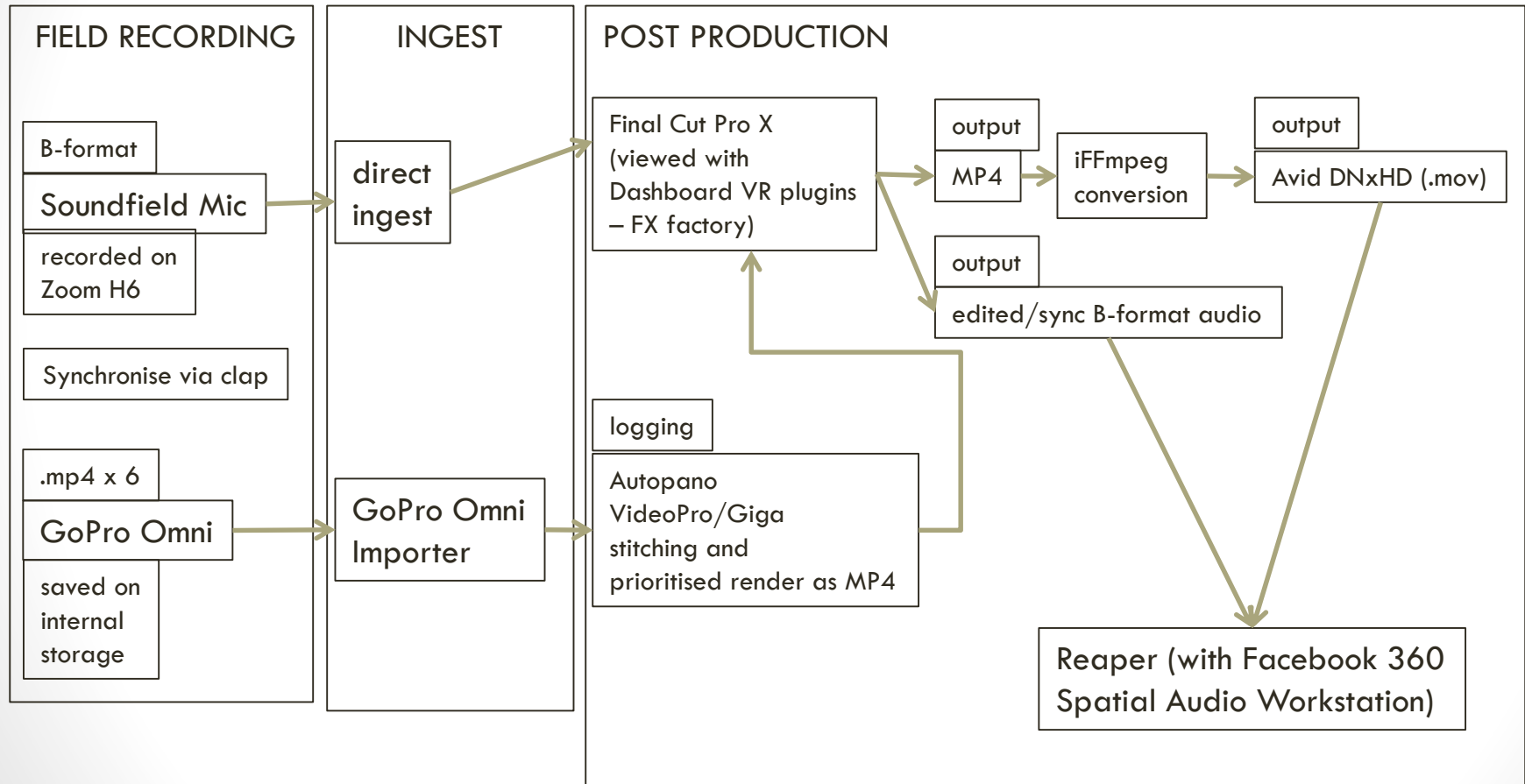
Glasgow 2018 European Championships

- Go Pro Omni camera and Soundfield ST350 microphone rig (and spot microphones)

- Velodrome racing, Scottish Universities Rowing Regatta, Ladies Scottish Open (golf), gymnastics, The British Diving Championships,

Capture Workflow

FIELD RECORDING

B-format

Soundfield Mic

recorded on Zoom H6

Synchronise via clap

.mp4 x 6

GoPro Omni

saved on internal storage

INGEST

direct ingest

GoPro Omni Importer

POST PRODUCTION

Final Cut Pro X (viewed with Dashboard VR plugins – FX factory)

output

MP4

iFFmpeg conversion

output

Avid DNxHD (.mov)

output

edited/sync B-format audio

logging

Autopano VideoPro/Giga stitching and prioritised render as MP4

Reaper (with Facebook 360 Spatial Audio Workstation)

Monitoring on Oculus Rift

Sound mixing process

SoundField audio was rendered to a 4 channel B-Format track

Location B-Format sound synchronized in FCPX/Premiere Pro, individual audio tracks were imported into Reaper.

The B-Format audio placed on a track with Facebook 360° Spatial Audio Station™ (FB360) Spatialiser plugin.

Other production sounds were track laid as mono or stereo tracks

the object-based-audio are positioned in different places in space

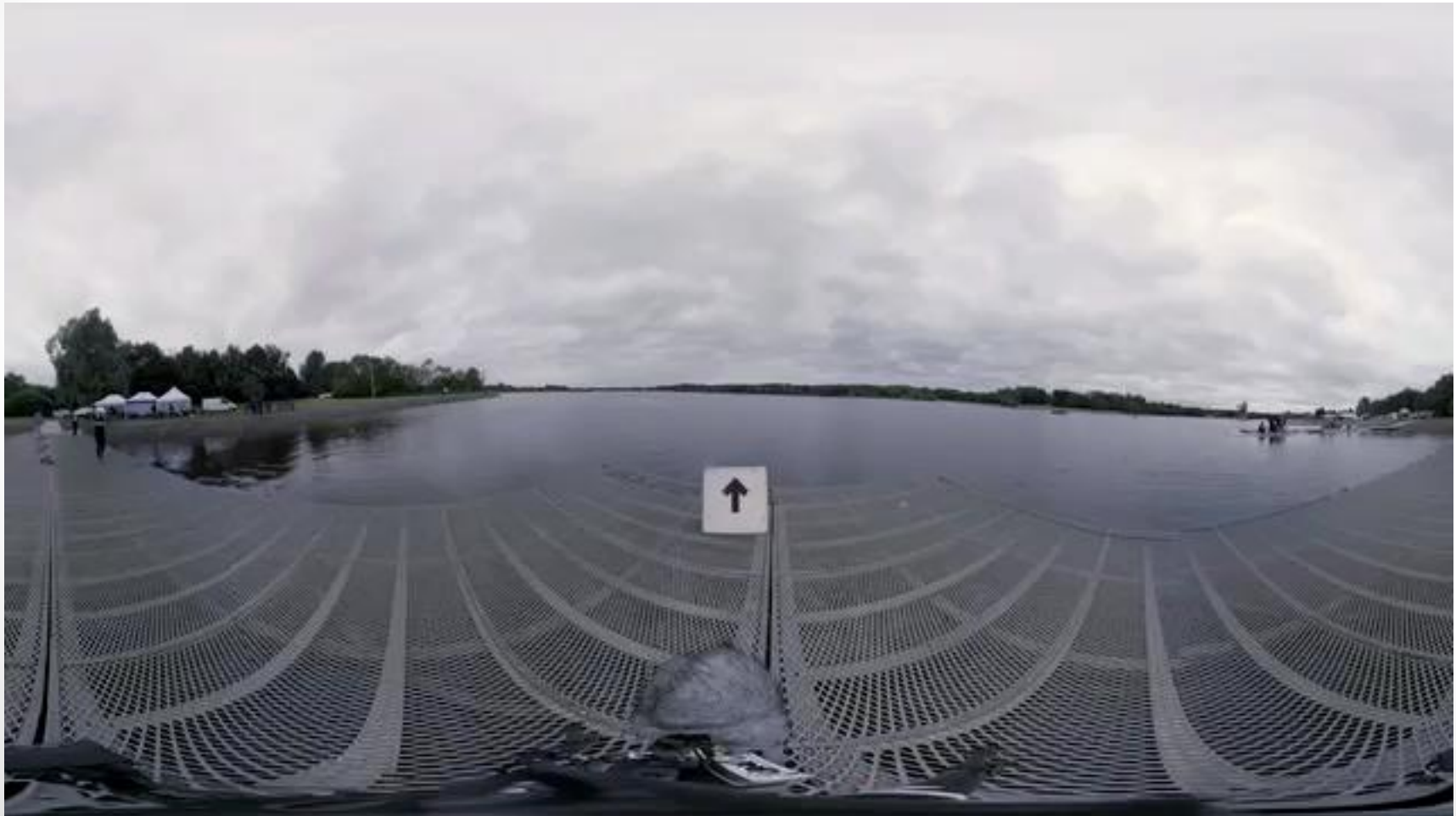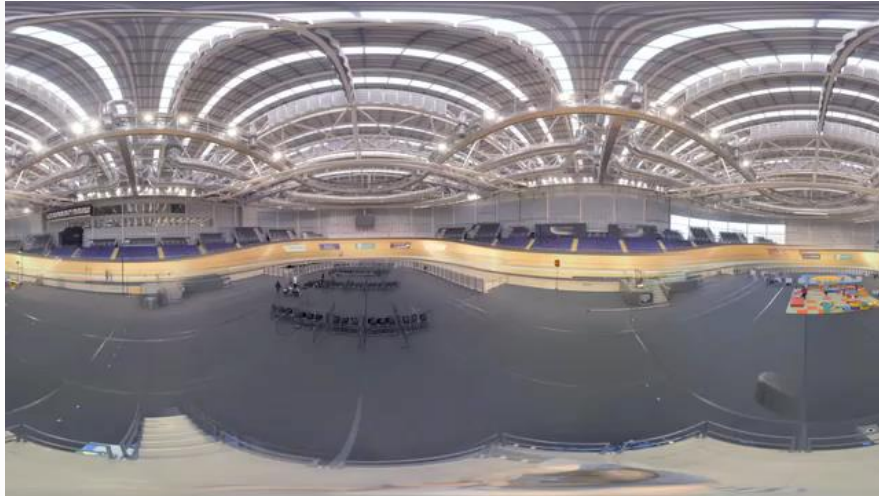The mix needs to be separated into SBA, OBA and Head-locked stereo.

Rowing
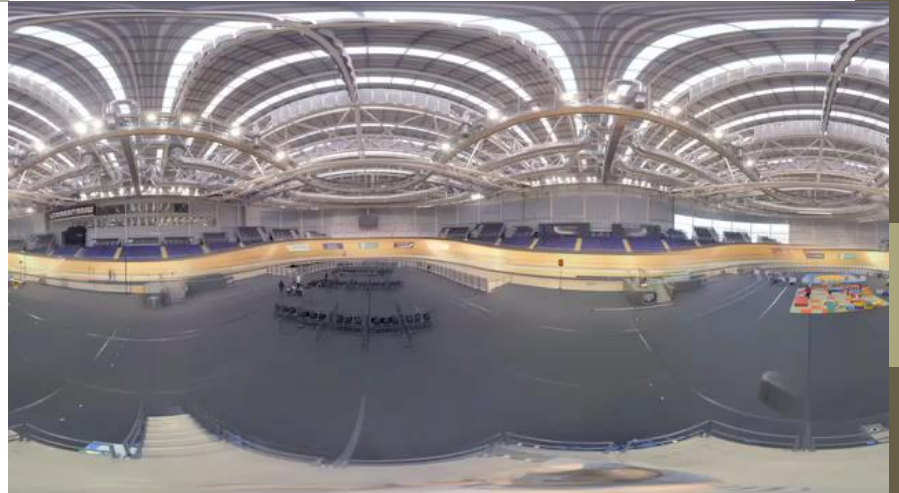
- GoPro Omni versus Samsung Gear VR.

Velodrome

the authentic, unedited soundfield recording

versus

enhanced audio

The 360 Degrees of Bouldering in Dumbarton (Petrova, 2017)

- Poli Petrova constructed the sound design to communicate inner-subjective sounds of the climber – heartbeats, low frequency droning thuds – over 100 individual sound sources.

- intimate sounds paired with distant visuals?

- A re-invention of cinematographic and editing concepts – leaving the audience to choose their own view, using sound as a guide, rather than a dense prescription of emotional clues.

- "The camera was positioned at the same height as my own head and treated as a person, rather than just a filming tool"

- calibration time at the start essential.

- "Bouldering consists of short, but hard routes and is usually done in a group of 2 or more people…While someone is climbing, the others "spot"… the viewers … will take part as spotters."

SCHOOL OF SIMULATION AND VISUALISATION THE GLASGOW SCHOOL OF ART

Sound recording process

- scene based audio recorded with Soundfield ST350 (+ windshield + zoom H6 recorder + sound devices box to split b-format into W, X, Y, Z)
- object based audio recorded with lavalier microphones, and shotgun microphones.
- Acquiring all these separate OBA sources allows for higher fidelity of the soundscape as well as more control and flexibility when their sources move.
- dialogue was captured as a separate take.
- sea breeze – return to get the right conditions for clean wild track, the sound of the water, wind in the trees and footsteps on rock.

Visual Experiments

- no means of live preview of the video
- if the filmmaker wants to capture long shots, ideally the entire film should only consist of them. (surprise or confusion by the sudden change of perspective, switch to close up)
- 15 minute takes = 530GB!
- camera attached to a rope above a sport climber.
  - parallax errors, due to the proximity of the rock on one side, and motion sickness inducing visuals, as the camera was not stable.
  - too close to the climber
  - camera must be a couple of meters away from the wall in order to capture the entire climber – drone?

Final approach

Camera was treated as a person and three types of angles were used.

1. camera is a passive viewer 'outside the action'.

2. 'a passive listener', where one of the protagonists addresses the camera directly and talks to it as if it is a person.

3. the viewer is 'part of the action', where the camera is in between the people watching the boulder

Post production issues

- Essential to have powerful computer - synchronization of OBA sources challenging, video player would go out of sync with the timecode (Re-exporting the video in a smaller size did not improve the situation as it was too pixelated to see important details)

- export a separate audio track for all sync sound (before mixing in Reaper, which does not support the OMF format) - create separate tracks containing "all W" or "all sync dialogue" tracks and separate them in Reaper

- FB360 software's "mix focus" control determines what the level attenuation of the source will be, when the viewer turns their head around.

- Had the "mix focus" control been present on individual channels, rather than on the control plug in, it would have been possible to allow the viewer to choose between whether or not to hear the climber or spotter's thoughts. This method would have been beneficial to the story as it would have shown some character development and verbally describe the situation.

- in order to watch the video within the session, all tracks need to be off-set by -19s. FFMPEG?

- According to the audiece reports - the authentic soundscape seems most natural and compelling compared to recordings made in controlled conditions (Such as no wind; no background noise)

The 360 Degrees of Bouldering in Dumbarton (Petrova, 2017)
subtle dropout of music, low frequency rumble to simulate fear, inner thoughts/reflections of the climber

The 360 Degrees of Bouldering in Dumbarton (Petrova, 2017) climber's breaths

References

Altman, M., Krauss, K., Susal, J., Tsingos, N., 2016. Immersive Audio for VR. [PDF] AES Conference on Audio for Virtual and Augmented Reality, 2016 September 30-October 1, Los Angeles, CA, USA. AES: Available [Accessed 10 August 2017]

Arte, 2016. *Notes on Blindness: A VR Journey into a World Beyond Sight*. Available at: http://www.notesonblindness.co.uk/vr/ [Accessed 31 August 2017]

Beaudoin, J-P., Nain, V.,(2016), "Audio for Cinmatic VR" [ONLINE VIDEO] Available at: http://www.gdcvault.com/play/1023654/Audio-For-Cinematic Accessed on: 14 August 2017

Breton-Lopez, J., Quero, S., Botella, C., Garcia-Palacios, A., Banos, R.M. & Alcaniz, M., 2010. An augmented reality system validation for the treatment of cockroach phobia. Cyberpsychology, Behavior and Social Networking, 1-16. [photograph] Available at < https://ispr.info/2010/10/20/treating-cockroach-phobia-with-augmented-reality/> [Accessed 16 January 2017]

Chalmers, A. G., Howard, D., and Moir, C. 2009. Real virtuality: A step change from virtual reality. In SCCG '09:Proceedings of the 25th Spring Conference on Computer Graphics, 9-16. New York: ACM Press.

DAVIS, M. F. (2003), "History of spatial coding"; Journal of the Audio Engineering Society, 51, 554-569 Accessed on: 01 August 2017

DOLBY (2011), "Dolby History" [WEBPAGE] Available: http://www.dolby.com/us/en/about/history.html Accessed on: 01 August 2017

Dolby, 2015. Personalized Audio for Sports Broadcast: The Opportunity to Engage. Available at: < https://www.dolby.com/uploadedFiles/wwwdolbycom/Content/Technologies/Dolby_Personalized_Audio/Personalized-Audio-for-Sports-Broadcast_white-paper.pdf > [Accessed 31 August 2017]

EleVr, 2014, "Audio for VR film (binaural, Ambisonic, 3D,etc" [WEBSITE] Available at: http://elevr.com/audio-for-vr-film/ Accessed on: 12 August 2017

Ellis-Petersen, H., 2015. Global panic: art show Exit brings climate change to shocking life. *The Guardian* [online] 27 November. Available at: < https://www.theguardian.com/artanddesign/2015/nov/27/art-show-exit-climate-change-cop21-un-conference-paris > [Accessed 31 August 2017]

Facebook, 2017. *Facebook 360*. Available at: https://facebook360.fb.com Date Accessed: 31 August 2017.

References

Facebook Audio 360 documentation, (2017)," Video Format Guidelines" [ONLINE] Available at: https://facebookincubator.github.io/facebook-360-spatial-workstation/KB/VideoFormatGuidelines.html Accessed on: 14 August 2017

Faramarzi, S., 2017. How women are gaining ground in virtual reality. The Guardian, [online] 14 August. Available at: < https://www.theguardian.com/lifeandstyle/2017/aug/14/virtual-reality-women-reclaim-technology#img-2 > [Accessed 31 August 2017]

Fassbender, E., Heiden, W., 2014. Atmosphaeres – 360° Video Environments for Stress and Pain Management. *Serious Games Development and Applications 5th International Conference*. Berlin, Germany, October 9-10 2014. Switzerland: Springer International Publishing.

GERZON, M. A. 1973. Periphony: "With-height sound reproduction." Journal of the Audio Engineering Society, 21, 2-10 Accessed on: 01 August 2017

Google VR, 2017. *Spatial Audio*. Available at: < https://developers.google.com/vr/concepts/spatial-audio > [Accessed 31 August 2017]

Grimshaw, M., Walther-Hansen, M., 2015. The Sound of the Smell of my Shoes. *Proceedings of the Audio Mostly 2015 on Interaction With Sound*. Thessaloniki, Greece, October 07 - 09, 2015. New York: Association for Computer Machinery.

Hoffman, H. G., Chambers, G. T., Meyer III, W. J., Arceneaux, L. L., Russell, W. J., Seibel, E. J., Richards, T. L., Sharar, S. R. and Patterson, D. R. 2011. Virtual Reality as an Adjunctive Non-pharmacologic Analgesic for Acute Burn Pain During Medical Procedures. The Society of Behavioral Medicine, [e-journal] 41, pp. 183-191. Available at: http://dx.doi.org/10.1007/s12160-010-9248-7 [Accessed 31 August 2017]

Hughes, M., (2016), "VR Headset Prices are going to Crash and Here's Why" [ONLINE] Available at: http://www.makeuseof.com/tag/virtual-reality-price-crash/ Accessed on: 12 August 2017

Immerscence Inc., 2017. *Osmose*. Available at: < http://www.immersence.com/osmose/ > [Accessed 31 August 2017]

J. Riedmiller, N. Tsingos, and S. Mehta, P. Boon, (2014), "Immersive & Personalized Audio: A Practical System for Enabling Interchange, Distribution & Delivery of Next Generation Audio Experiences"[PDF] Accessed on: 21 July 2017

Lee,H.,(2016), "Capturing and Rendering 360° VR Audio using Cardioid Microphones "[PDF] Paper presented at: AES Conference of Audio for Virtual and Augmented Reality, 2016, 2016 September 30- October 1, Lost Angeles, CA, USA. Accessed on: 10 August 2017

References

MANTEL, J., (1975), "Optimum Multiphony" 50th Audio Engineering Society Convention Accessed on: 01 August 2017

Manthorpe, R., 2017. I 'died' in virtual reality at the assisted suicide clinic Dignitas. *Wired*, [online] 24 March. Available at: < http://www.wired.co.uk/article/virtual-reality-assisted-suicide-dignitas > [Accessed 31 August 2017]

The Media Production Marketplace, (2017) "6 Reasons Why You should Focus on 360 ° Audio [WEBSITE] Available at: https://videostrategist.com/6-reasons-why-you-should-focus-on-360-audio-65810824a62b
Accessed on: 01 August 2017
Oculus, (2016), "Bring Your 360 Videos to Life with Spatial Audio" [ONLINE VIDEO], 12 October 2016 Available at: https://www.youtube.com/watch?v=MgI85YQch10&list=PLVKkUlcjSeb9H-GEWMuy5KfZZWLrJ18ux&index=5 Accessed on: 12 August 2017

Shivappa S, Morrell et al.(2016), "Efficient, compelling and immersive VR audio experience using Scene Based Audio / Higher Order Ambisonics" [PDF]; Paper presented at: AES Conference on Audio for Virtual and Augmented Reality, 2016 September 30- October 1, Lost Angeles, CA, USA Accessed on: 10 August 2017

Sarah Hill, (2016), "3D Audio for 360 Video Creators" [ONLINE VIDEO], 16 February 2016 Available at: https://youtu.be/uNJpPYo5Qp4 Accessed on: 12 August 2017

Scott, C., (2016), "How spatial audio can make your 360-degree videos more immersive" [ONLINE] Available at: https://www.journalism.co.uk/news/how-spatial-audio-can-make-your-360-degree-videos-more-immersive/s2/a695372/ Accessed on: 14 August 2017

Scuda, U.; Stenzel, H.; Baxter, D.,(2016), "Using Audio Objects and Spatial Audio in Sports Broadcasting" [PDF] , 57th Audio Engineering Society Convention

Sound On Sound, (2001), "You are Surrounded" [ONLINE] Available at: http://www.nada.kth.se/kurser/kth/2D2031/04_05/dokument/YOU%20ARE%20SURROUNDED%20VIII.pdf Accessed on: 12 August 2017

TORICK, E.,(1998),"Highlights in the history of multichannel sound" Journal of the Audio Engineering Society, 46, 27-31. Accessed on: 01 August 2017

University of Washington, (2015), "Sound Localization and the Auditory Scene"[ONLINE]
Available at: http://courses.washington.edu/psy333/lecture_pdfs/Week9_Day2.pdf Accessed on: 11 August 2017

United Nations Virtual Reality, 2015. *Syrian Refugee Crisis*. Available at: http://unvr.sdgactioncampaign.org/cloudsoversidra/#.WahZPI7KZFw [Accessed 31 August 2017]

United Nations Virtual Reality, 2015. *Surviving Ebola*. Available at: < http://unvr.sdgactioncampaign.org/wavesofgrace/#.WahZyI7KZFw > [Accessed 31 August 2017]

Vanian, J.,(2017), "Facebook Just Cut the Price of Its Virtual Reality Headset" [ONLINE] Available at: http://fortune.com/2017/03/01/facebook-oculus-rift-virtual-reality-headset-price/ Accessed on: 11 August 2017
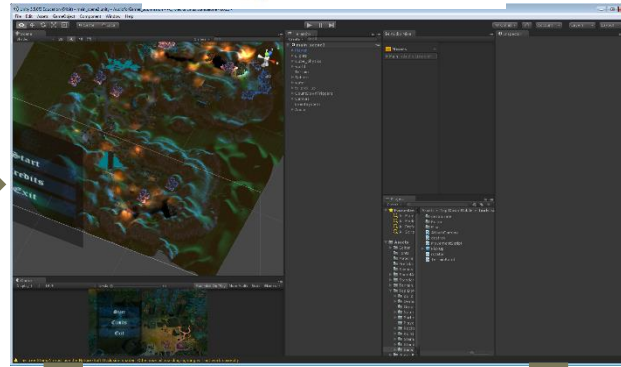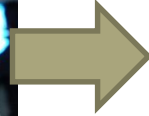
Whittington, W.,(2007)"Sound Design & Science Fiction", (P.223-p.240) [BOOK]

Potential VR Audio Workflow
(University of York)

OVR

Viewer Position (x,y,z)

Yaw, Pitch, Roll

Scene Description

$S1(x,y,z)$
$S2(x,y,z)$

$sN(x,y,z)$

Reverb parameters

HOA Sound Object Encoder / Mixer

HOA Reverb Engine

HOA Rotation Matrix

Virtual loudspeaker decoder

Ambisonic Engine

Left

Right