Russell Mason



February 2002 – present

**Institution**: University of Surrey

# Unit of Assessment: 12 Engineering

**Title of case study**: Developing spatial audio experiences for global audiences

# Period when the underpinning research was undertaken: 2013 – 2020Details of staff conducting the underpinning research from the submitting unit:Name(s):Role(s) (e.g. job title):Period(s) employed by submitting HEI:Philip JacksonReaderNovember 2002 – presentAdrian HiltonProfessorNovember 1994 – presentPhilip ColemanLecturerApril 2014 – present

Period when the claimed impact occurred: 2013 – 2020

Is this case study continued from a case study submitted in 2014? N

Senior Lecturer

# 1. Summary of the impact (indicative maximum 100 words)

Surrey's spatial audio research has generated substantial economic, societal and cultural impacts via three main knowledge-transfer activities: (1) commercialisation of sound zone technology, (2) public engagement with spatial audio experiences, (3) next-generation audio technology for UK and global audiences. Surrey's sound zone technology was patented and commercialised by Bang & Olufsen (B&O), featuring in their BeoLab and BeoSound Balance products. The pioneering object-based audio research at Surrey has resulted in next-generation technology and vital new capabilities for audio production, which have engaged audiences of many thousands worldwide. As a direct result of Surrey's collaboration, the BBC have new spatial audio tools and European Broadcasting Union (EBU) standards for next-generation audio and are embedding this technology into UK-based media production infrastructure for global distribution.

# 2. Underpinning research (indicative maximum 500 words)

Spatial audio technology offers completely enveloping sound from all around the listener, as in real life, creating a far more immersive experience than conventional audio. It is the natural and necessary complement to high-resolution video. Sound zones deliver this without headphones by cancelling crosstalk, enabling users to experience personalised content in the same room.

The University of Surrey's Centre for Vision, Speech & Signal Processing (CVSSP), led by Hilton since 2012, is globally-recognised for its strengths in acoustical and visual signal processing. Complementing these strengths is Surrey's Institute of Sound Recording (IoSR), with particular expertise in auditory perception.

Working collaboratively within two research projects on multi-channel acoustical systems: Perceptually Optimised Sound Zones (POSZ) and Future Object-Based Spatial Audio (S3A), CVSSP developed sound zone technology **[R1, R2]**, audio device orchestration **[R3]** and object-based audio systems **[R4, R5, R6]**.

# Sound zone technology

Sound zone systems aim to control the sound such that multiple listeners can enjoy different audio programmes within the same room with minimal acoustic interference.

#### Impact case study (REF3)



The POSZ project (2010-14) was conceived by Jackson to combine sound zone engineering with psychoacoustics to predict and evaluate perceived sound quality. CVSSP conducted engineering research collaboratively with B&O engineers, whilst IoSR performed psychoacoustic research with B&O psychoacousticians resulting in 24 publications, 4 these, 3 patents and 2 datasets. This project established scientific foundations for the practical realisation of sound zones including the suitability of energy focusing, energy cancellation, and synthesis approaches for sound zone reproduction **[R1]**. Regularisation was also shown to control effects on the array effort and achieved acoustic contrast between zones alongside the trade-off between acoustic contrast and fidelity of the target audio programme in the zone **[R2]**.

#### Audio device orchestration and object-based audio systems

The S3A project (2013-19 EP/L000539/1) brought leading international experts in 3D audio and visual processing (including The University of Salford and University of Southampton) together with the BBC and UK creative industry. This Surrey-led, EPSRC funded five-year programme aimed to enable listeners at home to experience the sense of "being there" at live events (e.g., concerts, football matches) through immersive sound. The project has resulted in a large body of work including 8 datasets and over 170 publications, seminal outcomes include:

- Inventing 'audio orchestration' (AO); using ad-hoc consumer media devices delivering objectbased spatial audio to enhance immersion [R3]. In contrast to traditionally-conceived spatial audio, this system relied on semantic metadata describing the audio objects and a rule set for automatically adapting the reproduction across the loudspeakers on all connected devices. Listener feedback showed a positive experience due to the increased immersion compared with stereo reproduction, and that the orchestration approach worked particularly well for drama content.
- 2. An object-based approach for capturing, parameterising, editing, and rendering reverberation over a 3D spatial audio system **[R4]**. A Reverberant Spatial Audio Object (RSAO) embellishes a recorded sound object by synthesizing the desired room impression. In listener tests, RSAO was shown to convey room size and source distance correctly and modulate listener perception of these.
- 3. An end-to-end object-based spatial audio pipeline from sound recording to listening **[R5]**. The high-level system architecture introduced audio-visual interfaces to support object-based capture and listener-tracked rendering, and proposed objectification components, recording content directly into objects. Through experimentation, the research team demonstrated that the system's novel components extended scope and capabilities beyond state of the art.
- 4. Finally, a system architecture for semantically informed rendering (SIR) that combines objectbased audio rendering with high-level processing of object metadata [R6]. This processing uses novel, advanced metadata describing the objects to adjust the audio scene to the reproduction system and listener preferences. Examples demonstrated that SIR can significantly improve the media experience and provide advanced personalisation controls, for example, by maintaining smooth object trajectories on systems with few loudspeakers or providing personalised envelopment levels.

These outcomes from the POSZ and S3A projects represent the origin of an expansive research portfolio that culminated in the impact described.

**3. References to the research** (indicative maximum of six references), University of Surrey researchers highlighted in **bold**.

Seminal outputs from the POSZ and S3A projects include:

[R1] P.D. Coleman, P.J.B. Jackson, M. Olik, M. Møller, M. Olsen, and J.A. Pedersen, (2014) "Acoustic contrast, planarity and robustness of sound zone methods using a circular



loudspeaker array", Journal of the Acoustical Society of America, 135(4): 1929-1940, April 2014. DOI: 10.1121/1.4866442

- [R2] K. Baykaner, J. Francombe, R. Mason, P.D. Coleman, M. Olik, P.J.B. Jackson and S. Bech, (2015) "The relationship between target quality and interference in sound zones", Journal of the Audio Engineering Society, 63(1/2): 78-89, Jan.2015. DOI: <u>10.17743/jaes.2015.0007</u>
- [R3] J. Francombe, J. Woodcock, R.J. Hughes, R. Mason, A. Franck, C. Pike, T. Brookes, W.J. Davies, P.J.B. Jackson, T.J. Cox, (2018) "Qualitative evaluation of media device orchestration for immersive spatial audio reproduction", Journal of the Audio Engineering Society, 66(6): 414-429, June 2018. DOI: <u>10.17743/jaes.2018.0027</u>
- [R4] P.D. Coleman, A. Franck, P.J.B. Jackson, R.J. Hughes, L. Remaggi and F. Melchior, (2017) "Object-Based Reverberation for Spatial Audio", Journal of the Audio Engineering Society, 65(1/2): 66-77, Jan. 2017. DOI: <u>10.17743/jaes.2016.0059</u>
- [R5] P.D. Coleman, A. Franck, J. Francombe, Q. Liu, T. de Campos, R. Hughes, D. Menzies, M. Simón Gálvez, Y. Tang, J. Woodcock, P.J.B. Jackson, F. Melchior, C. Pike, F. Fazi, T.J. Cox and A. Hilton, (2018) "An audio-visual system for object-based audio: from recording to listening", IEEE Transactions on Multimedia, 20(8): 1919-1931, Aug. 2018. DOI: 10.1109/TMM.2018.2794780
- [R6] A. Franck, J. Francombe, Q. Liu, T. de Campos, R. Hughes, D. Menzies, M. Simon Galvez, Y. Tang, J. Woodcock, P.D. Coleman, P.J.B. Jackson, F. Melchior, C. Pike, F. Fazi, T.J. Cox and A. Hilton, (2019) "A system architecture for semantically informed rendering of object-based audio", Journal of the Audio Engineering Society, 67(7/8): 498-509, July 2019. DOI: <u>10.17743/jaes.2019.0025</u>
- 4. Details of the impact (indicative maximum 750 words)

# The commercialisation of sound zone technology

Bang & Olufsen (B&O) are world-renowned for their high-end luxury audio products; their innovative and progressive products are sold worldwide. B&O employ over 1,000 people and operate in more than 70 markets. Surrey's sound zone technology research **[R1, R2]** resulted in a number of new B&O products and commercial success, B&O's Research Director confirms:

"The fundamental knowledge developed in POSZ has made us capable of developing and using advanced directivity control where BeoLab 90 was the first product utilising this. Then followed BeoLab 50. BeoSound Balance and two new products will be released shortly." **[S1]** 

BeoLab 90 launched in November 2015, retailing at £61,000, and is described in the B&O 2015/2016 Annual report as their "<...> new flagship loudspeaker offering unprecedented level of control of acoustic behaviour and performance, making BeoLab 90 one of the most innovative speakers on the market". **[S2]** The technologies and know-how gained from the development of BeoLab 90 have been applied to subsequent Bang & Olufsen loudspeaker products.

B&O's Annual Report 2015/16 also confirmed the commercial success of BeoLab 90: "Revenue from the Speaker category increased 26 per cent compared to last year. Thereby, the share of revenue from the Speaker category was unchanged at 18 per cent compared to last year. Speaker sales were positively impacted by the launch of Bang & Olufsen's new flagship speaker, BeoLab 90 which has been very well received in the market." **[S2]** 

As a result of their research collaboration (2014-2016) with Surrey, B&O filed several patents relating to sound zone technology, each naming Surrey researchers as inventors **[S3]**. In May 2015, HARMAN International Industries (which later became a subsidiary of Samsung) announced its acquisition of Bang & Olufsen Automotive, B&O's car audio business unit. The sound zone patents **[S3]** were licensed as part of Harman's €145M purchase of B&O's automotive business (2015) **[S1]**.



HARMAN subsequently launched 'individual sound zones' (ISZ). In 2016, HARMAN was recognised by *Fast Company as one of the Top Ten Most Innovative Companies in Automotive* and reported sales of \$6.5B (2015) and \$6.91B (2016) **[S4]**. Automotive backlog of awarded business also hit a record high of \$24.1B in 2016 following the launch of their innovative audio systems, including ISZ **[S5]**.

# Public engagement with spatial audio

Showcasing the potential of, and engaging audiences in, next-generation audio is a critical step in the BBC and other providers investing in new tools and ultimately content. As such, the S3A project provided the BBC with such an opportunity:

"Three distinct pieces were produced to test out concepts relating to different strands of the research. All three of these productions were made possible through the cooperation and collaboration of the S3A project partners, the Universities of Surrey, Salford and Southampton, with BBC R&D." [S6]

**The Turning Forest** is the BBC's first publicly released VR experience built on spatial audio content designed, commissioned, recorded and produced with S3A and its object-based production system to render sounds from any direction **[R5, S7]**. It premiered at Tribeca Film Festival (attendance c150,000), then at festivals in England, Scotland, France, Germany, USA, Canada and Australia, plus many community, industry and scientific events. It won the TVB Europe Award for Best Achievement in Sound (2016) **[S8]** and a nomination for Best VR Experience at the Google Play Awards (2017).

The BBC publicised *The Turning Forest* via their Taster website. Its YouTube trailer has >24,000 views, and overall audience reach c100k since distributed globally on Google Play (58.5k downloads to Sept 2020) and for Oculus Go and Samsung VR (24.1k downloads to Feb 2020).

*The Vostok-K Incident* (2018) was commissioned and produced by S3A using ground-breaking research in audio orchestration **[R3]** of spatial sound across multiple consumer devices to create an immersive experience. **[S9]** 

Vostok-K was launched on BBC Taster, where it received >5.9k unique views, with 72% of viewers saying they "would use the technology again" and scoring it 4/5 overall **[S6]**. The launch was widely covered in the media, including the *Daily Telegraph, Financial Times,* BBC Radio South, BBC Radio Surrey and BBC Click. The BBC has subsequently formed a team to develop the Audio Orchestrator technology resulting in multiple productions, including BBC Philharmonic 2020 (www.bbc.co.uk/taster/pilots/pick-a-part), Monster 2020 (www.bbc.co.uk/rd/blog/2020-10-audio-drama-monster-interactive-sound), and public release of the Audio Orchestrator as a community tool for sound engineers (https://www.bbc.co.uk/rd/publications/whitepaper375), resulting in >300 requests to use the tool. Audience response research consistently indicates >70% of users say they would like to use this technology again.

In 2019, the BBC and S3A co-produced an episode of *Casualty* to evaluate S3A's research on the use of object-based narrative importance to improve speech accessibility for broadcast integrated within a version of iPlayer (>8k views on BBC Taster and positive audience feedback).

Together, these trials were pivotal in informing the BBC's future strategy, and as a result of their success, the BBC established its Virtual Reality Hub, producing 14 VR experiences over two years **[S6]**. The BBC Lead R&D Engineer confirms, *"The S3A project has had a major influence on BBC R&D's work on object-based media in recent years, demonstrating the technical capabilities and audience value clearly. These ideas have fed into our thinking on projects such as the StoryKit and our plans for the New BBC Ecosystem."* **[S6]** 

# Next-generation audio technology

Since stereo's first broadcast 59 years ago, the systems and infrastructure have evolved worldwide to support channel-based audio. Next-generation audio (NGA) reflects a major global



transformation in broadcast and internet-delivered media services for improved listening experiences, using object-based techniques to make audio more immersive and personalisable. The BBC's Lead R&D Engineer confirms Surrey's critical role (via S3A) in enabling next-generation audio:

"The BBC's role has been in creating interoperable standards for NGA production and distribution. The S3A research work contributed directly to the BBC's work on the Rec. ITU-R BS.2127 standard that specifies algorithms for rendering the Audio Definition Model. This has recently been implemented in the EAR Production Suite (<u>https://ear-productionsuite.ebu.io/</u>), a project by the European Broadcasting Union." [S6]

The BBC and EBU have also adopted S3A object-based audio datasets as reference material for the NGA standards, including the Audio Definition Model Renderer and the Binaural NGA Renderer. This underpinned a binaural production of *Doctor Who in 2017*.

{*Text removed for publication*} **[S10]**.

5. Sources to corroborate the impact (indicative maximum of 10 references)

- [S1] Testimonial from B&O Research Director. (PDF)
- [S2] B&O Annual Report 2015/16 <u>https://investor.bang-olufsen.com/static-files/abfeafda-db36-4637-b29c-e9380a2b5574</u>
- [S3] July 2014 (filed January 2014), A Sound-field control method using a planarity measure (Coleman and Jackson (Surrey) <u>https://patentimages.storage.googleapis.com/2e/1d/e5/2d6de5f3b46b4b/WO2014108365</u> <u>A1.pdf</u>

March 2016, (filed March 2013) System for Optimizing the perceived sound quality in virtual sound zones Mason (Surrey) https://patentimages.storage.googleapis.com/18/41/fc/412bdc0da38682/US9277322.pdf

- [S4] February 2016 *Businesswire* article detailing HARMAN and Fast Company recognition <u>https://www.businesswire.com/news/home/20160216006480/en/HARMAN-Recognized-</u> <u>by-Fast-Company-as-One-of-the-Most-Innovative-Companies-in-Automotive</u>
- [S5] Harman 2016 Annual Sales Report https://materials.proxyvote.com/default.aspx?docHostID=300193
- [S6] Testimonial from the BBC's Lead R&D Engineer. (PDF)
- [S7] Coleman (Surrey) listed amongst S3A researchers on Turning Forest; https://www.bbc.co.uk/rd/blog/2016-05-virtual-reality-sound-in-the-turning-forest
- [S8] Turning Forest winner of 2016 TVB Europe Award (Best Achievement in Sound) https://www.tvbeurope.com/business/tvbawards-2016-winners-announced
- **[S9]** BBC article outlining Vostok-K incident with S3A project team listed <u>https://www.bbc.co.uk/taster/pilots/vostok</u>
- **[S10]** Confidential testimonial from {*Text removed for publication*}. (PDF available from the HEI upon request)