Audio Engineering Society

Convention e-Brief 430

Presented at the 144th Convention
2018 May 23–26, Milan, Italy

An efficient method for producing binaural mixes of classical music from a primary stereo mix

Tom Parnell[1] and Chris Pike[1]

[1] *BBC Research & Development, Salford, United Kingdom*

Correspondence should be addressed to Tom Parnell (tom.parnell@bbc.co.uk)

ABSTRACT

Radio audiences in the UK are increasingly listening using headphones, and binaural mixes are likely to offer more natural and immersive classical musical experiences than stereo broadcasts. However, the stereo mix is currently a priority for broadcasters, and producers have limited resources to create an additional, binaural mix. This engineering brief describes the semi-automated workflow used to produce binaural mixes of performances from the BBC Proms. Spatial audio mixes were created by repositioning the individual microphone signals from the stereo broadcast in three dimensions, and adding ambient signals captured using a 3D microphone array. A commercial mixing application was used for spatial panning and binaural rendering, and the resulting binaural audio was streamed live online. Comments on the production workflow were collected from the music balancers, and audience responses were surveyed.

1 Introduction

A third of radio audiences in the UK now uses headphones to listen [1], possibly prompted by the arrival of fast mobile data networks, content streaming services such as the BBC iPlayer, and behaviours around the consumption of digital music. At the same time, radio broadcasters are increasingly realising the benefits of using binaural technology to produce novel and enhanced audio content [2].

Binaural processing filters audio signals with head-related transfer functions (HRTFs) to simulate the hearing cues created by acoustic interaction between our bodies and the environment around us. This gives the impression that a sound source is located outside the head at a given location in space. Radio producers can benefit from applying this advanced audio technique to create a richer sense of space in the sound of their programmes. Previous research has found that the overall listening experience of stereo music is preferred to binaural [3]; consequently, further work needs to be done to understand what needs to be improved in binaural mixing systems to give better overall quality. Binaural mixes of classical music are likely to be better than stereo mixes at reproducing source positions and the spatial impression of the hall ambience [4]. However, audiences for recorded classical music ascribe high value to particular aesthetic qualities of commercial stereo recordings, and balances of classical music for radio broadcast reflect these preferences; for example, extra low-frequency content, a wider perspective, a closer perspective, more reverberation, and clarity in low-level details [5].

Binaural renders exist as two-channel audio files and can be modified with stereo audio processing, distributed via existing two-channel broadcast audio infrastructure, and listened to using stereo audio players and software, including the BBC iPlayer Radio application. Therefore, binaural sound is currently more convenient to deliver by radio broadcasters than loudspeaker-based spatial formats like 5.1 surround sound, or object-based audio transmissions for advanced sound systems.

Radio programmes are mixed in stereo and listened to using loudspeakers and headphones. The motivation for creating an additional, binaural mix is to offer an alternative version for headphone listeners which enhances source positioning and ambience. If some of the existing stereo production process can be shared with the binaural production process, this will be beneficial because the valued parts of the stereo mix will be retained and unnecessary duplication of effort will be avoided. The latter point is particularly important for producers who have limited budgets.

2 Context

The BBC Proms is a festival of classical music held annually at the Royal Albert Hall in London, with every concert broadcast live on BBC Radio 3. A variety of musical forces perform several genres of music, each of which requires different audio mixing skills and techniques. A range of microphones is permanently rigged in the hall (see Figure 1) for the duration of the season, and a collection of ad hoc 'spot' microphones is available for close-miking instruments and singers on stage.

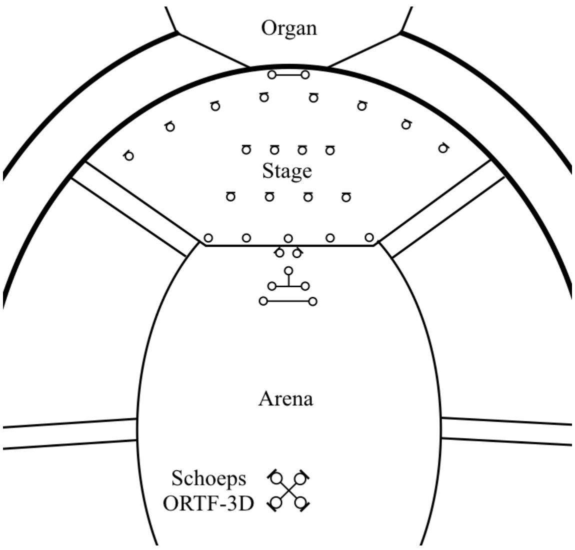

Figure 1. Diagram showing the Proms microphone rig at the Royal Albert Hall.

The BBC radio outside broadcast truck, Sound 3, uses the Stagetec Nexus audio system to route microphone signals and other audio between the hall and Sound 3. A dedicated team of sound engineers and producers creates the stereo music mixes, using Sound 3's Stagetec Aurus mixing console. Speech presentation is added at a second mix position inside the hall, from where the pulse code modulation audio is streamed to Broadcasting House, for broadcast on BBC Radio 3.

The BBC offered an automated 4.0 loudspeaker mix of Proms performances for several years until 2016. This was achieved by routing a sub-mix of the non-ambient elements in the stereo music mix to the two front loudspeakers, routing a sub-mix of the ambient elements to the two rear loudspeakers, and adding artificial surround reverberation. This approach required the listener to receive and route four channels (delivered using the web audio application processing interface) through their computer to four appropriately-positioned loudspeakers. The music balancer monitored the feed infrequently. It offered limited spatial benefit to the listener and obtained minimal uptake.

With increased headphone listening, and spatial audio tools advancing, an object-based binaural production seemed appropriate to complement BBC Radio 3's stereo coverage of the 2017 Proms. Twenty-one concerts were streamed live online in binaural sound to represent the range of classical music performed; they included orchestras, soloists, choirs, baroque ensembles, a contemporary music ensemble, solo piano, solo organ, and the world-famous *Last Night of the Proms* event, with its enthusiastic audience participation. The trial offered a unique opportunity to test binaural music production in an existing live broadcast environment, with experienced music balancers. The public broadcasts offered the chance to gather feedback from the audience.

3 Methodology

An efficient, semi-automated technical workflow was developed for creating binaural mixes based on the existing microphone rig, audio infrastructure, and mix process used to create the stereo music mixes. This involved rigging additional microphones in the Royal Albert Hall and integrating with the Stagetec Nexus audio network. For each binaural broadcast, the music balancer was supported in the operation of the binaural production system by one of the authors.

The trial was conducted for acoustic performances only; amplified concerts require a fundamentally different technical approach to achieving the stereo mix, and were therefore not considered here.

3.1 Microphone signal path

Each microphone signal used in the stereo mix was fed post-fader, post-EQ, post-dynamics processing, post-delay, from the Aurus console into IRCAM's Panoramix spatial audio mixing application, using two MADI connections. This ensured that the relative levels of microphone signals used in the stereo mix were maintained, and that the feeds included any dynamic level changes made by the balance engineer; this included extensive use of the Aurus console's control groups, which allow selections of individual channels (slaves) to be level-controlled together by master faders.

The Panoramix software was used to apply three-dimensional panning, EQ, delay and binaural rendering of each source. Since most of the microphones used were in static positions for the duration of each concert, their three-dimensional panning could be retained throughout; level changes for different musical pieces were automatically reflected in the binaural mix via the post-fader feeds 'inherited' from the Aurus console.

3.2 Binaural rendering

Microphone signals were grouped into three types: 'main' microphones, 'spot' microphones, and 3D 'ambient' microphones. Separate binaural busses were established in Panoramix for each specified microphone type and an additional bus was created for the post-fader feeds from the stereo reverberation devices used in the stereo mixes. See Table 1.

Panoramix performs object-based binaural rendering where each source is processed through a binaural filter corresponding to the intended source direction. Panoramix allows the use of user-loaded spatially oriented format for acoustics (SOFA) files [6]; in this case, anechoic head-related impulse responses (recorded using a Neumann KU-100 dummy head microphone) [7].

Bus	Source type	EQ applied	Delay
1	Main mics	None	28ms
2	Spot mics	None	28ms
3	Ambient mics	High-pass at 160Hz	0ms
4	Reverberation	None	28ms

Table 1. Binaural bus processing applied within Panoramix.

In stereo mixes, music peaks are compensated for by overall level control of the master stereo audio group; however, these alterations do not affect the individual post-fader sends from each microphone channel. Sufficient headroom exists in each Aurus channel that the MADI feeds did not peak; however, the consequent summing of these feeds in each binaural bus caused overloading to occur when attenuation was not applied. It was therefore necessary to attenuate each binaural bus by a default of 6dB each; this was a 'starting point' and the eventual level differed depending on the artistic preference for each music balance.

Overall level control of the binaural mixes was directly linked within the Aurus console to the overall level control of the stereo mixes in order to achieve the same dynamic control in both mixes.

3.3 Artificial reverberation

Artificial reverberation was added to most stereo mixes using a Lexicon 960L and a Bricasti M7. The music balancers used different algorithms depending on the application and for different purposes; for example, to augment the natural reverberation of the hall, to imitate the natural reverberation of the hall, and to 'soften' the use of spot microphones in the stereo mixes. The stereo outputs from the two reverberation devices were mixed into the frontal sound stage of the binaural mixes to perform the same functions.

3.4 Ambient capture for use in binaural mix

The general ambience (early reflections, reverberation, and diffuse audience reaction) of the hall was captured spatially using a Schoeps ORTF-3D microphone array, rigged approximately 10m above the Arena, and approximately 12m into the auditorium from the front of the stage (see Figure 2).

Figure 2. Schoeps ORTF-3D microphone array (circled, top left), in situ at the Royal Albert Hall.

Mic	Mic label	Layer	Physical and panned azimuth	Default panned elevation
MK41	L	Lower	–45˚	0˚
MK41	R	Lower	45˚	0˚
MK41	Ls	Lower	–135˚	0˚
MK41	Rs	Lower	135˚	0˚
MK41V	Lh	Upper	–45˚	40˚
MK41V	Rh	Upper	45˚	40˚
MK41V	Lsh	Upper	–135˚	40˚
MK41V	Rsh	Upper	135˚	40˚

Table 2. Assignment and 3D panning of microphones in Schoeps ORTF-3D array.

The array uses two coincident ‘layers’ of four microphones, each positioned at the corner of a 17cm-sided square. The lower layer is orientated at –30˚ to the horizontal; the upper layer is elevated by +60˚. The eight microphones of the array were spatially panned as in Table 2.

3.5 Miscellaneous considerations

Individual channel delays are often applied to the microphone signals used in stereo mixes; this helps to compensate for the difference in time of arrival of the sound of an instrument at its ‘close’ microphone and the ‘main pair’ of microphones. It was therefore decided to further delay (as detailed in Table 1) all the signals used in the stereo mixes relative to the more ambient 3D microphone array.

A sense of height in the spatial sound field was achieved by applying elevation to the panned positions of orchestral wind, brass, and percussion microphones, and was perceived especially when chorus, organ, and offstage musicians were present in the mix.

The current settings of all the parameters in the Panoramix console can be saved as a template “session” file, making the recall of parameters for different musical pieces much quicker. The first author created a default template file as a ‘starting point’ for each binaural mix, from which further, musically-informed, creative decisions were made by the music balancers, in collaboration with the producers and the authors. Decisions made included 3D positions of sources, levels of each binaural bus, and levels of some individual sources, such as solo instruments and singers performing with orchestras.

By including all of the microphones and artificial reverberation used in the stereo mixes, it was hoped that the binaural mixes would be as musically and aesthetically similar as possible to the stereo mixes. Any changes to the overall tonal quality of the binaural mixes could be attributed to ad hoc balancing of the four binaural busses and/or the HRTF convolution. These differences were informally monitored on headphones by direct A-B comparison of the stereo and binaural mixes.

3.6 Distribution

A standalone IP stream was created for the live binaural transmissions; this used a MPEG-DASH stream to transport AAC-LC-encoded audio at 320kbps, the same maximum quality available on the BBC's standard live radio streams. A webpage (http://www.bbc.co.uk/taster/pilots/proms-binaural-2017) including background information, the audio player, and an audience survey was hosted on BBC Taster, an experimental online platform to showcase and gain audience insight in to new technological developments. The player automatically activated when required, according to the pre-programmed schedule of concerts.

Binaural audio clips of each musical piece were uploaded to the BBC Radio 3 programme webpages and made available on a dedicated Proms webpage (http://www.bbc.co.uk/programmes/articles/2913JxRtQl3ZTvw0wz5C4D1/bbc-proms-in-binaural-sound).

4 Outcomes

The method was informally evaluated to find out: (a) if the binaural process retained the tonal and musical qualities of the crafted stereo mixes; (b) what the music balancers thought of using it; and (c) whether it offered an enhanced experience to the audience.

4.1 Observations by the authors

The method provided a semi-automated workflow for creating binaural mixes from the pre-mixed elements used in the stereo mix. By grouping these individual elements into four main binaural busses (Table 1) the creation of binaural mixes was simplified and could be completed quickly. However, some aspects of the binaural mixes required further mixing to obtain satisfactory mixes; for example, in concerto performances or solo performers with an orchestra, solo microphones arranged as stereo pairs did not always transfer well to the binaural mix, and levels of solo microphones often needed altering from the levels in the stereo mix.

Audio busses (of individual input channels), rather than control groups, were occasionally used in stereo mixes; therefore, any level changes made by the balancer to the audio busses did not affect the level of the bus members' channel post-fader outputs. This caused minor, but noticeable, differences between the stereo and binaural music balances.

Musical repertoire involving spatially-positioned performers benefitted most from binaural mixes. For example, Prom 44 featured a piece for onstage ensemble and five groups of instruments dispersed around the hall. This spatial effect was more successfully reproduced in the binaural mix than the stereo mix.

A few concerts included speech presentation from the stage; this was captured using a wireless microphone. Achieving a satisfactory balance between intelligibility and spatial perspective was challenging.

4.2 Responses from music balancers

Music balancers were impressed by the ease with which binaural mixes could be produced in this way. The level of engagement by balancers with the Panoramix software varied: some were keen to claim ownership of the binaural mixes while others preferred the authors to take control of the binaural mixes; reasons for this variation included lack of knowledge in binaural technology and time constraints hindering involvement in creating the additional mixes.

Balancers generally liked the binaural process and enjoyed the freedom to position microphone sources in three dimensions. There was a range of preferences in the approach to panning, and in the relative levels of 3D ambience and artificial reverberation included in the binaural mixes. Positioning was approached with different aims: some emulated the large width of the orchestral sound stage achieved when listening to their stereo mixes using headphones; others preferred to recreate the frontal sound stage as heard in loudspeaker stereo, which resulted in a narrower orchestral sound stage in the binaural mix, sometimes with a more distant perspective, and sometimes retaining a close focus on the individual musical elements.

The binaural mixes were perceived to provide extra spaciousness. However, minor differences between the timbral quality of the stereo and binaural mixes were noticed; these can probably be attributed to HRTF convolution and the introduction to the binaural mixes of the Schoeps ORTF-3D array.

4.3 BBC Taster

The live streaming webpage was visited by 8,962 unique browsers during the eight-week season, with over 20,000 page views. Listeners were encouraged to respond by answering several simple questions regarding their experience; 57% rated the experience 5 out of 5. The binaural mixes were rated as "more enjoyable" by 79% of respondents, whilst 75% said the experience was "somewhat" or "absolutely" like being there in person. Finally, 88% agreed that BBC Radio 3 should produce more binaural content.

5 Conclusions and future considerations

This paper has described how the individual microphone and artificial reverberation elements of a stereo mix of classical music can be further processed in order to produce an effective spatial audio mix, binaurally-rendered, that is suitably consistent with the original stereo balance. It has described how such binaural renders can be experienced with the same equipment used for stereo headphone listening, and it was shown that the audience enjoyed the listening experience. Balance engineers are enthusiastic about the potential for future binaural broadcasts of classical music, and have high expectations regarding consistency with stereo mixes. Some stereo balancing practices make direct transfer to binaural mixes difficult and a more consistent approach is required in future.

Further research into this method and associated applications will include:

- simplifying the binaural production interface so that 3D panning and level control is more closely aligned with the Aurus operation, and can be operated solely by the music balancer;
- formal subjective testing of the overall listening experience of the binaural and stereo mixes;
- making recommendations for the treatment of onstage speech in binaural mixes;
- the use of tracked headphone monitoring;
- the suitability of the static binaural mix to virtual reality and 360˚ video applications; and
- the use of 3D reverb and room simulation as additional creative tools.

Acknowledgements

The authors thank the BBC Radio & Music Operations team for assisting in designing the method, rigging additional microphones, and contributing to the binaural mixes and research outcomes of the trial.

References

[1] BBC Marketing & Audiences. Internal research document, 2017.

[2] BBC News, "3D sound experiments carried out by BBC R&D engineers", Online: https://www.bbc.co.uk/news/technology-26958946 (15 April 2014).

[3] Walton, T. "The Overall Listening Experience of Binaural Audio". VDT International Conference on Spatial Audio, 2017.

[4] Plenge, G. "On the differences between localization and lateralization." *The Journal of the Acoustical Society of America*, *56*(3), pp. 944–951, 1974.

[5] King, Richard. "Sound Board: Food for Thought, Aesthetics in Orchestra Recording." JAES Volume 63, Issue 4, pp. 303-304; April 2015.

[6] AES69-2015: AES standard for file exchange - Spatial acoustic data file format.

[7] Bernschütz, B. (2013). A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU 100. In *Proceedings of DAGA*.