



# Multi-channel reproduction system comparison dataset

*to accompany*

## Production and reproduction of programme material for a variety of spatial audio formats

Audio Engineering Society 138th Convention, Warsaw, Poland, 7–10 May 2015.

Jon Francombe<sup>1\*</sup>, Tim Brookes<sup>1</sup>, Russell Mason<sup>1</sup>, Rupert Flindt<sup>1</sup>,  
Philip Coleman<sup>2</sup>, Qingju Liu<sup>2</sup>, and Philip J.B. Jackson<sup>2</sup>

<sup>1</sup>*Institute of Sound Recording, University of Surrey, Guildford, UK.*

<sup>2</sup>*Centre for Vision, Speech, and Signal Processing, University of Surrey,  
Guildford, UK.*

\*[j.francombe@surrey.ac.uk](mailto:j.francombe@surrey.ac.uk)

May 15, 2015

## Abstract

This documentation describes the S3A multi-channel reproduction system comparison dataset, which is available to download (free for non-commercial research use) from <http://cvssp.org/S3A>.

## Contents

|          |                                 |          |
|----------|---------------------------------|----------|
| <b>1</b> | <b>Introduction</b>             | <b>3</b> |
| <b>2</b> | <b>The content</b>              | <b>4</b> |
| <b>3</b> | <b>The reproduction methods</b> | <b>5</b> |
| <b>4</b> | <b>The files</b>                | <b>8</b> |
| <b>5</b> | <b>Acknowledgements</b>         | <b>8</b> |

## 1 Introduction

A set of spatial audio programme material items in different formats was developed for the purpose of performing listening tests in spatial audio systems. The stimulus set was designed to cover a wide range of potential aspects of 3D audio content, including: various genres and musical characteristics; different numbers of sources; different types of source (small, large, electric, acoustic, etc.); different recording environments (studio, large hall, outside, etc.); different source positions (in front, above, behind) and movement; and various technical aspects (dynamic range, LFE content, etc.).

Some items in the stimulus set were captured in a specifically designed recording session in which multiple techniques were used simultaneously to capture various ensembles [1, 2]. The remaining items were created by remixing pre-existing multitrack content [2].



Figure 1: Brass quintet recording session, University of Surrey Studio 1

The stimulus set is available to download free for non-commercial research use from <http://cvssp.org/S3A>. If you use the stimuli in any work that is published, please include a reference to the following paper:

Francombe, J., Brookes, T., Mason, R., Flindt, R., Coleman, P., Liu, Q. and Jackson, P., 2015. ‘Production and reproduction of programme material for a variety of spatial audio formats.’ *Audio Engineering Society 138th Convention*, Warsaw, Poland, 7–10 May 2015.

## 2 The content

The programme material items described in the following sections are available in the stimulus set. Each excerpt is 20 seconds long and available for the reproduction methods described in Section 3. The items were loudness matched by a panel of listeners in a method-of-adjustment experiment.

### Brass quintet

A brass quintet (two trumpets, horn, trombone, and tuba) playing *Quintet* by Michael Kamen. Recorded by Jon Francombe in Studio 1 at the University of Surrey.

### Film

An excerpt of the soundtrack to *Skyfall* (2012) featuring music, speech, and sound effects. The soundtrack was augmented with a B-format recording of rain in order to add a height element to the surround sound format from the DVD.

### Pop track

*Just Another Frame* by Paul Hedges, Peter Walters, and Rupert Flindt. Mixed for various formats by Rupert Flindt.

### Experimental music

A first-order B-format rendering of *Rotating psychoacoustic tuning curves* by Florian Hecker, decoded to different speaker layouts.

### Jazz quintet

A jazz-rock quintet (two electric guitars, electric bass, piano, and drums) playing *Freddie Freeloader* by Miles Davis. Recorded by Jon Francombe in Studio 1 at the University of Surrey.

### Big band

A live big band performance by the RAF Squadronaires playing *Little Brown Jug* by Glen Miller, arr. Barrie Forgie. Recorded in the Royal Albert Hall by Rupert Flindt.

### 3 The reproduction methods

Each programme item is provided in seven formats:

- mono;
- stereo;
- 5.1;
- 9.1;
- 22.2;
- ambisonic cuboid; and
- headphones.

Table 1 contains details of the loudspeaker positions according to ITU-R BS.2051 [4], and also the channel order in the files. Table 2 contains more detail on the production of each stimulus for each reproduction method.

| Chan. | Mono  | Stereo | 5.1   | 9.1   | 22.2  | Cuboid | HP |
|-------|-------|--------|-------|-------|-------|--------|----|
| 1     | M+000 | M−030  | M−030 | M−030 | M−060 | B−045  | L  |
| 2     |       | M+030  | M+030 | M+030 | M+060 | B+045  | R  |
| 3     |       |        | M+000 | M+000 | M+000 | B+135  |    |
| 4     |       |        | LFE1  | LFE1  | LFE1  | B−135  |    |
| 5     |       |        | M−110 | M−110 | M−135 | U−045  |    |
| 6     |       |        | M+110 | M+110 | M+135 | U+045  |    |
| 7     |       |        |       | U−030 | M−030 | U+135  |    |
| 8     |       |        |       | U+030 | M+030 | U−135  |    |
| 9     |       |        |       | U−110 | M+180 |        |    |
| 10    |       |        |       | U+110 | LFE2  |        |    |
| 11    |       |        |       |       | M−090 |        |    |
| 12    |       |        |       |       | M+090 |        |    |
| 13    |       |        |       |       | U−045 |        |    |
| 14    |       |        |       |       | U+045 |        |    |
| 15    |       |        |       |       | U+000 |        |    |
| 16    |       |        |       |       | T+000 |        |    |
| 17    |       |        |       |       | U−135 |        |    |
| 18    |       |        |       |       | U+135 |        |    |
| 19    |       |        |       |       | U−090 |        |    |
| 20    |       |        |       |       | U+090 |        |    |
| 21    |       |        |       |       | U+180 |        |    |
| 22    |       |        |       |       | B+000 |        |    |
| 23    |       |        |       |       | B−045 |        |    |
| 24    |       |        |       |       | B+045 |        |    |

Table 1: Loudspeaker positions and file channels. The loudspeaker labels are from ITU-R BS.2051 [4]. The first letter indicates the layer (M = 0 degrees elevation, U = 30 degrees elevation, B = −30 degrees elevation, T = 90 degrees elevation; 0 degrees is at ear height and 90 degrees is directly above the listener. This is followed by an angle in degrees, where 000 is directly in front of the listener and +090 is directly to the right of the listener.

| Prog.<br>item       | Mono,<br>low quality               | mono   | Stereo  | 5-channel   | 9-channel   | 22-channel   | Cuboid  | Headphones                  |
|---------------------|------------------------------------|--|---|---|---|--|---|-----------------------------|
| <b>Brass</b>        | Single cardioid microphone         |  | Crossed stereo pair   | fig-8<br>Fukada tree and<br>Hamasaki square                               | Fukada tree,<br>Hamasaki square,<br>and upwards-<br>facing Hamasaki<br>square | Hamasaki <i>et al.</i> [3]<br>method 2                                     | Soundfield microphone (1st order B-format)  | Neumann binaural dummy head |
| <b>Jazz</b>         | Single cardioid microphone         |  | Crossed stereo pair   | fig-8<br>Fukada tree and<br>Hamasaki square                               | Fukada tree,<br>Hamasaki square,<br>and upwards-<br>facing Hamasaki<br>square | Hamasaki <i>et al.</i> [3]<br>method 2                                     | Soundfield microphone (1st order B-format)  | Neumann binaural dummy head |
| <b>Pop</b>          | Sum of stereo mix                  |  | Stereo mix by engineer  | 5-channel mix by engineer (no LFE)  | 9-channel mix by engineer (no LFE)  | 22-channel mix by engineer (no LFE)  | First-order ambisonic rendering of 22-channel mix   | Stereo mix                  |
| <b>Big band</b>     | Sum of stereo mix                  |  | Stereo mix by engineer  | 5-channel mix by engineer (no LFE)  | 9-channel mix by engineer (no LFE)  | 22-channel mix by engineer (no LFE)  | First-order ambisonic rendering of 22-channel mix   | Stereo mix                  |
| <b>Experimental</b> | First-order format decoded to mono | B-   | First-order format decoded to stereo  | First-order format decoded to 5 channels (no LFE)                         | First-order format decoded to 9 channels (no LFE)                             | First-order format decoded to 22 channels (no LFE)                         | First-order B-format decoded to cuboid  | Stereo version              |
| <b>Film</b>         | Mono sum of stereo version         | sum of L, C, R from DVD with added first-order B-format rain decoded to stereo | Stereo downmix of L, C, R from DVD with added first-order B-format rain decoded to stereo | 5.1 soundtrack with added first-order B-format rain decoded to 5 channels | 5.1 soundtrack with added first-order B-format rain decoded to 9 channels     | 5.1 soundtrack with added first-order B-format rain decoded to 22 channels | First-order ambisonic rendering of 5-channel mix with added first-order B-format rain decoded to cuboid | Stereo version              |

Table 2: Details of the creation of each stimulus. For the multichannel programme items, there was no dedicated LFE channel in any content with the exception of the film soundtrack.

## 4 The files

Each stimulus is a .wav file with one to twenty-four channels, as described above. The sampling frequency is 48 kHz, with a resolution of 24 bits per sample. The files are named as follows:

`[ProgrammeName]_[ReproductionMethod]_LoudnessMatchedMean.wav`

## 5 Acknowledgements

This work was supported by the EPSRC Programme Grant S3A: Future Spatial Audio for an Immersive Listener Experience at Home (EP/L000539/1) and the BBC as part of the BBC Audio Research Partnership.

We would like to thank the individuals and groups who assisted in performing and/or providing the programme material items: James Lackovic, Andy Macfarlane, Derryck Nasib, Chris Atchley, and Mark Prentice-Whitney; Paul Hedges, Peter Walters, and Rupert Flindt; Florian Hecker and Tony Myatt; and Tommy Ashby, James Wyatt, James Newman, Jake Willson, and Oliver Waton.

We would also like to thank the researchers and technicians that contributed and assisted during the recording session: Teo de Campos, Mariana Nunes de Oliveira, James Woodcock, Toby Stokes, George Oulton, Alan Haigh, Daisuke Koya, Kirsiten Hermes, Chris Pike, and Frank Melchior.

## References

- [1] Francombe, J., 2014. ‘S3A recording session (IoSR blog).’ <http://iosr.surrey.ac.uk/blog/2014-09-26.php>.
- [2] Francombe, J., Brookes, T., Mason, R., Flindt, R., Coleman, P., Liu, Q. and Jackson, P., 2015. ‘Production and reproduction of programme material for a variety of spatial audio formats.’ In: *AES 138th Convention, Warsaw, Poland, 7–10 May 2015*. eBrief No. 199.



- [3] Hamasaki, K., Hiyama, K. and Okumura, R., 2005. ‘The 22.2 multi-channel sound system and its application.’ In: *AES 118th Convention, Barcelona, Spain, 28–31 May 2005*. Paper No. 6406.
- [4] ITU-R, 2014. ‘Recommendation BS.2051-0: Advanced sound system for programme production.’