Immersive Audio Recording for Virtual and Augmented Reality

Lewis Thresh, *Research Technician*, **Hasham Riaz,** *Abbey Road Studios* **AND Gavin Kearney,** *Senior Lecturer*
(lewis.thresh@york.ac.uk) (Hasham.Riaz@abbeyroad.com) (gavin.kearney@york.ac.uk)

University of York, York, UK

Abbey Road Studios, London, UK

0 INTRODUCTION

This work follows on from [1]...

1 Recording at Abbey Road Studios

2 Test Material Preperation

2.1 Video

As two different 360°cameras were used during recording, two different methods of spherical video encoding were used. **[Look at has's AES/thesis for this?]**

2.1.1 Position A - 360°perspective

Writing

2.1.2 Position B - 180°perspective

2.2 Audio

3 Listening Tests

Two rounds of listening tests were conducted for viewing position A (test 1) and B (test 2). The procedure was identical however the data used, such as video and microphone configurations used were different for each. Both tests recruited participants both from the University of York and Abbey Road Studios. All participants were required to have some previous experience with mixing/producing and/or spatial audio. The number of participants recruited for each test were as follows:

Recruted From	Test 1	Test 2
UoY	15	29
Abbey Road	4	9
Total	**19**	**38**

3.1 Attributes Focus Group

The aim of the listening test was to assess the performance of each microphone array configuration for a VR environment in terms of its spatial and timbral quality. Due to the subjectivity of such a test, a focus group was assembled with the purpose of producing a list of mutually agreeable adjectives to use to describe certain spatial and timbral attributes. The attributes chosen to use within the listening tests are shown in table1.

3.2 Procedure

Participants was asked to wear and Oculus Rift DK2 headset and a pair of Audio Technica MH50x headphones and watch the first 80 seconds of the chosen take from their respective positions (A or B). Once the clip was finished participants would answer a questionnaire. This involved rating on a scale of 1 - 10 the level at which they experienced each of the spatial attributes listed in table 1 and to list any timbral attributes they felt best described the timbre of the recording.

To ensure that all participants understood each of the attributes in the same way a short training exercise was conducted before each test. This involved taking the participants through each of the attributes with audio examples.

4 Analysis

To best analyse the results, five analysis targets were defined:

Attribute	Description
Spatial	
Locatedness	How easily you can locate a sound source within the VR environment
Sense of Space	How well the space where the recording was made is perceived
Externalisation	Perception of sound coming from all around your head
Envelopment	Whether the sounds are perceived to originate inside of outside of the head
Timbral	
Full	Abundance of low frequencies present
Bright	Abundance of high frequencies present
Flat	Lack of high and low frequencies present
Rich	The mix sounds good with both high and low frequencies
Realistic	The sounds heard in the VR experience are realistic (sound like real instruments) and timbral characteristics have been preserved.
Loud	The perceived level sounds high

Table 1. Table of Spatial and Timbral Attributes

Analysis 1: Does viewing position effect Spatial Attribute rating?

Analysis 2: Does the choice of microphone array effect Spatial Attribute score?

Analysis 3: What is the effect of using Directional or Diffuse-Field Arrays?

Analysis 4: Is there a correlation between SA score and selected timral attributes?

Analysis 5: Is there a in perception of timbre with difference viewing positions?

4.1 Analysis 1: Does viewing position effect spatial attribute scores?

To assess the potential effect of viewing position on spatial attribute score, data was grouped into 8 sections: An average score for each spatial audio attribute (4 groups) each split into the average score for viewing position A and B (8 groups) illustrated in figure 1.

4.2 Analysis 2: Does the choice of microphone array effect Spatial Attribute score?

4.3 Analysis 3: What is the effect of using Directional or Diffuse-Field Arrays?

4.4 Analysis 4: Is there a correlation between SA score and selected timral attributes?

4.5 Analysis 5: Is there a in perception of timbre with difference viewing positions?

5 REFERENCES

[1] H. Riaz, M. Stiles, C. Armstrong, A. Chadwick, H. Lee, G. Kearney, "Multichannel Microphone Array Recording for Popular Music Production in Virtual Reality," presented at the *Audio Engineering Society Convention 143* (2017 Oct), URL http://www.aes.org/e-lib/browse.cfm?elib=19333.

THE AUTHORS

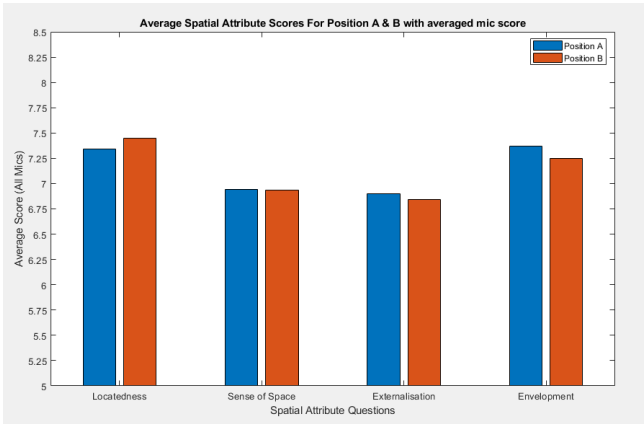

Fig. 1. Bar chart showing average spatial attribute score for each spatial attribute at viewing position A and B