



# Designing an Interactive and Collaborative Experience in Audio Augmented Reality

Valentin Bauer<sup>1</sup>(✉), Anna Nagele<sup>1</sup>(✉), Chris Baume<sup>2</sup>, Tim Cowlshaw<sup>2</sup>,  
Henry Cooke<sup>2</sup>, Chris Pike<sup>2</sup>, and Patrick G. T. Healey<sup>1</sup>

<sup>1</sup> School of Electronic Engineering and Computer Science,  
Queen Mary University of London, London, UK  
{v.m.bauer, a.n.nagele}@qmul.ac.uk

<sup>2</sup> BBC Research and Development, London, UK

**Abstract.** Audio Augmented Reality (AAR) consists of adding spatial audio entities into the real environment. Existing mobile applications and technologies open questions around interactive and collaborative AAR. This paper proposes an experiment to examine how spatial audio can prompt and support actions in interactive AAR experiences; how distinct auditory information influence collaborative tasks and group dynamics; and how gamified AAR can enhance participatory storytelling. We are developing an interactive multiplayer experience in AAR using the Bose “Frames” audio sunglasses. Four participants at a time will go through a gamified story that attempts to interfere with group dynamics. In this paper we present our AAR platform and collaborative game in terms of experience design, and detail the testing methodology and analysis that we will conduct to answer our research questions.

**Keywords:** Audio Augmented Reality · Collaboration · 3D audio · Audio game · Storytelling · Experience design

## 1 Introduction

In the past few years, the technological development of 3D audio for headphones using binaural audio has facilitated the delivery of Audio Augmented Reality (AAR) experiences. AAR consists of adding spatial audio entities into the real environment [13]. The technology has been applied to a range of fields such as teleconferencing, accessible audio systems, location-based games or education. AAR Research has mainly been focusing on the perception of sound quality [18], realism, or discrimination between real and virtual sounds [13]. Yet, interaction and collaboration remain under-researched. One of the big challenges is acoustic

---

Supported by the EPSRC and AHRC Centre for Doctoral Training in Media and Arts Technology at Queen Mary University London and BBC R&D.

The first two authors have equally contributed to the paper.

© Springer Nature Switzerland AG 2019

P. Bourdot et al. (Eds.): EuroVR 2019, LNCS 11883, pp. 305–311, 2019.

[https://doi.org/10.1007/978-3-030-31908-3\\_20](https://doi.org/10.1007/978-3-030-31908-3_20)

transparency, so that the user can stay connected to his environment as if they had no headphones. Bose Frames (BF) audio sunglasses [5] are a newly available wearable AAR consumer technology that embed the speakers and technology in the frame of the sunglasses, and are therefore perfectly acoustically transparent.

We are developing an interactive AAR multiplayer experience, for four players at a time, that encourages human interactions. Our prototype, *Please Confirm you are not a Robot*, explores three research questions: How can the affordances of the technology and spatial sound prompt and support actions in interactive AAR? How can asymmetric information influence group dynamics and support or distract from collaborative tasks? How can a participatory performance create empathy and behaviour change through interactive storytelling? We will test our game in a user experience study, from which we want to derive design implications for interactive storytelling and multiplayer AAR game design.

## 2 Background

### 2.1 State of the Art

In AAR, spatial audio is often rendered over headphones. Issues have been reported regarding front-back inversions, sound timbre artifacts, or externalisation, due to the use of non-individual HRTFs [3]. These are more noticeable in static than in dynamic binaural, which consists of the addition of a headtracking system, and also increases the user immersion and localisation accuracy. A valuable asset of wearable devices is that they can offer headtracking possibilities and thus make dynamic binaural audio more widely available.

Representations of the AAR sound field can be either natural or pseudoacoustic. For the former, virtual audio entities are directly added to the auditory real environment. For the latter, binaural microphones are added into the listener's ears and routed to the earphones so that the listener perceives a synthesized version of his environment. This system, also called “hear-through” audio, is for instance common in hearing aids. In all cases, the aim is that the user should not be able to determine which sources are real and which are not. This requires using high-quality 3D audio rendering [13] and a careful mix between the virtual sources and the auditory environment.

### 2.2 Challenges

Previous AAR studies have mainly focussed on hear-through audio. Transparent earphones remain under-explored and questions arise about a seamless integration of audio entities onto the real auditory environment. Some open-ear systems, such as the BF system presented in Sect. 4, exist. The mixed reality Microsoft HoloLens glasses can render dynamic binaural audio and holographic 3D images, using small loudspeakers, a camera, eye tracking, and headtracking sensors integrated in the frame of the glasses [14]. Bone conduction headsets can also be used to render binaural audio. Despite some localization accuracy issues, good externalization and spatial discrimination can be achieved [2].

Designing sounds in AAR still requires more investigation, but binaural audio has been shown to increase the user immersion in comparison to stereo audio [7]. Mixing remains a challenge due to the dynamic nature of real sounds that change over time in both level and frequency, which can lead to audio masking.

Most AAR applications remain individual. Yet, some studies have focused on collaboration through location-based AAR games in stereo, using sounds triggered at specific locations [11,16]. Regarding spatial AAR, Mariette and Katz [17] developed *SoundDelta*, a mobile multi-user AAR architecture which uses mobile user devices and servers communicating over WiFi. They explore the potential of the *Ambisonic cell* approach to deliver personalized audio to a large number of users over a specific area.

In the following sections, we present the development of our AAR game and architecture, and introduce our research methodology and planned studies.

### 3 AAR Experience Design and Storytelling

Headphones and similar devices divide auditory spaces into private and public. BF, in contrast, do not create a sound barrier but allow individual augmentation of sonic experiences. Lyons et al. [16] suggest that AAR has potential to bring people together in the same location and enhance social interactions. We are designing an environment to foster face-to-face interaction, exploiting three features of AAR: Asymmetric information; Layering augmented sounds over “real life” sounds; Triggering sounds with head gestures and movement.

A limited amount of applications for Bose AR exist. Some apps allow users to explore a soundscape by selective listening [19], other ones use BF as a gaming device with taps and head movements as interactions [1], or make use of the technology’s mobility through soundwalks [8]. *Dead Drop Desperado* [10] is the only known game that requires two players.

Apart from those Bose AR applications, spatial audio is used in immersive theatre to create imaginary spaces and parallel realities [9]. AAR experiences often assign a role to the user, asking them to perform. Looking at this in a multiplayer context, this is reminiscent of choreography and theatre performance. The theatre practice developed by theatre maker Augusto Boal [4] blurs the boundaries between everyday activities and performance. It is used to rehearse for desired social change [15]. Inspired by this practice, our multiplayer game will result in a choreography prompting users to observe, reenact and subvert behaviours around digital devices.

#### 3.1 Game Overview

*Please Confirm you are not a Robot* is a speculative fiction, constructed of four individual games. At the start, each participant meets their guide who introduces the scenario and the gesture controls: tapping, nodding and shaking head. In the first game participants are prompted to simultaneously draw a circle with one arm, and a cross with the other arm, in the air. Spatialised sounds of drawing a

circle and cross will play for some players alongside the movement. We will look at whether sound cues have any effect on the participant's performance.

For the second game participants pair up and mirror each other's movements while being prompted to ask each other questions. We will look at whether this contributes to interpersonal closeness or affect between participants, or whether different layers of sound are distracting.

The third game uses the BF as a gaming interface. A variety of notification sounds will appear in the sonic sphere around each participant. To turn them off they have to look at the sound and double-tap the side of the frames. Participants collect points for each sound they turn off. We will test different feedback sounds for finding sounds in space.

The last game requires the participants to tap each other's frames, following prompts of what they like about each other, to collect points. We will look at the interaction between participants. At the end, one participant will be separated from the group with a separate story-line. They will become the agent to end the whole experience by taking the other player's frames off their face.

### 3.2 Game Design

Early designs of interactive audio-only experiences highlighted the importance of sound design [16]. Sounds have to be put in context with other sounds or narration to establish a cause and effect relationship between the actions and sounds, and match the player's mental model [16]. Since varying loudness levels are a challenge in AAR, we decided on a specific room where we will conduct the experiments. We created the sound design with sounds gathered from personal recordings or Freesound.org (under the Creative Commons Licence), and using the software SoundParticles in combination with Reaper.

## 4 Audio Augmented Reality Architecture

In our modular AAR platform, we use BF because of their acoustic transparency, headtracking system, user interaction and ergonomics. In addition, BF have a Bluetooth low energy system and offer three interactions: nodding, shaking the head, and tapping the glasses. The headtracking system has an accelerometer, gyroscope and magnetometer, and a latency of around 200 ms (higher than the 60 ms optimal latency [6]). This may affect audio localisation but the other BF aspects make it suited to achieve a good user engagement in the game. Since our system had to be modular and support 3D audio, GPS tracking, BF API, and multiplayer possibilities, we chose to work with Unity software (version 2018.3 for compatibility). We work with phones on iOS due to the compatibility with BF SDK, but future developments may also include Android phones.

For multiplayer collaboration we designed a Local Area Network (LAN) over WiFi using the Unity's *UNet* system. The first player who connects to the game is the host and starts broadcasting its IP address. Following players (clients) automatically detect it and join the game. Some objects are synchronised over

the network, such as player dependent objects, or objects that keep track of global variables. Asynchronised objects can also trigger events locally for each player. With this system events can be player specific, and different players can listen to different sounds synchronised over time. The BF API gives us access to the sensor data of BF. We used Google Resonance Audio SDK for 3D audio rendering because of its high-quality with 3rd-order ambisonics [12]. The architecture supports GPS tracking with Mapbox API, Audio Interactive Programming with Pure Data, and gives access to the phone's affordances (sensors, vibrator) (Fig. 1).

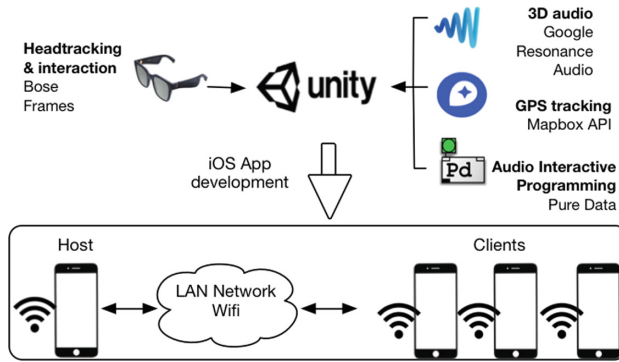


Fig. 1. Architecture of the AAR game

## 5 Testing Methodology

We conducted preliminary testing with a group of four participants from BBC R&D to detect technical and narrative flaws that helped us to refine our prototype. We will soon conduct a user study with five groups of four participants with different levels of expertise in 3D audio and augmented reality. A pre-study questionnaire will assess previous experience with 3D audio as well as interpersonal relationships of the group. One researcher will be with each user during the experience and take notes about their behaviours. After each game, users will be asked to answer specific questions about the game. A post-study questionnaire and guided group discussion will assess aspects of the game such as enjoyment, interactive ease, problems, storytelling and feelings about the group. Experiments and discussion will both be filmed and recorded. We will conduct qualitative analysis of the recordings and participant responses, in comparison with the performance measures we set out for each game. From this analysis we will attempt to answer our research questions, derive design recommendations for AAR multiplayer games and give an indication of areas for further research.

## 6 Conclusion

We reviewed previous AAR studies and discovered that newly available technologies such as BF, which do not cover the ears, offer new opportunities for collaboration and interaction in AAR. We were inspired by previous multiplayer experiences, methods of interactive storytelling, and theatre practices to develop *Please Confirm you are not a Robot*. This game immerses a group of four players into a scene where they play and act out several scenes, guided by asymmetric information and binaural sound cues. This paper details the development of the modular AAR architecture that supports our experimental game, and can be extended in the future to create other multiplayer games. This is one of the first studies to our knowledge that evaluates BF AAR experiences.

## References

1. Audiojack Homepage. <https://www.audiojack.com>. Accessed 3 Jul 2019
2. Barde, A., Helton, W.S., Lee, G., Billinghamurst, M.: Binaural spatialization over a bone conduction headset: minimum discernable angular difference. In: Proceedings of the 140th Convention of the Audio Engineering Society (2016)
3. Begault, D.-R.: 3-D Sound for Virtual Reality and Multimedia. Academic Press Professional, San Diego (1994)
4. Boal, A.: Games for Actors and Non-actors, 2nd edn. Taylor and Francis e-Library, Abingdon (2005)
5. BOSE Frames Homepage. <https://www.bose.co.uk/engb/products/frames.html>. Accessed 3 Jul 2019
6. Brungart, D.S., Kordik, A.J., Simpson, B.D.: Effects of headtracker latency in virtual audio displays. *J. Audio Eng. Soc.* **54**(1), 13 (2006)
7. Chatzidimitris, T., Gavalas, D., Michael, D.: SoundPacman: audio augmented reality in location-based games. In: 18th Mediterranean Electrotechnical Conference (MELECON), Lemesos, Cyprus, pp. 1–6 (2016)
8. Cooke, H.: Audio AR: Sound Walk Research The Missing Voice (2019). <https://www.bbc.co.uk/rd/blog/2019-03-audio-sound-walks-tours>. Accessed 3 Jul 2019
9. Darkfield Homepage. <https://www.darkfield.org>. Accessed 3 Jul 2019
10. Dead Drop Desperados App Store Preview. <https://apps.apple.com/us/app/dead-drop-desperado/id1454393037>. Accessed 3 Jul 2019
11. Ekman, I.: Sound-based gaming for sighted audiences- experiences from a mobile multiplayer location aware game. In: Proceedings on 2nd Audio Mostly, Germany (2007)
12. Gorzel, M., Allen, A., Kelly, I., Kammerl, J., Gungormusler, A., Yeh, H.: Efficient encoding and decoding of binaural sound with resonance audio. In: Audio Engineering Society Convention 12 (2019)
13. Härmä, A., et al.: Augmented reality audio for mobile and wearable appliances. *J. Audio Eng. Soc.* **52**(6), 618–639 (2004)
14. Hololens Homepage. <https://www.microsoft.com/en-us/hololens/>. Accessed 28 Jul 2019
15. Kochhar-Lindgren, K.: Towards a communal body of art: the exquisite corpse and Augusto Boal's theatre. *Angelaki: J. Theor. Hum.* **7**(1), 217–226 (2002). <https://doi.org/10.1080/09697250220142137>

16. Lyons, K., Gandy, M., Starner, T.: Guided by voices: an audio augmented reality system. In: Online Proceedings of the International Conference on Auditory Display, Georgia Institute of Technology Atlanta, Georgia, USA (2000)
17. Mariette, N., Katz, B.: SoundDelta - large scale, multiuser audio augmented reality. In: EAA Symposium on Auralization, Espoo, Finland (2009)
18. Tikander, M.: Usability issues in listening to natural sounds with an augmented reality headset. *J. Audio Eng. Soc.* **57**(6), 430–441 (2009)
19. Traverse Homepage. <https://www.traverse.fm>. Accessed 3 Jul 2019