Implementing a portable augmented/virtual reality auralisation tool on consumer-grade devices

Tim Beresford (1) and Jack Wong (1)

(1) Acoustics, Norman Disney & Young, Auckland, New Zealand

ABSTRACT

Auralisation is a powerful tool for presenting acoustic design options to stakeholders, enabling them to preview an acoustic environment using appropriately processed audio signals. Current state-of-the-art auralisations require a dedicated listening room with very low reverberation times, very low background noise, and a multi-channel loudspeaker setup. More recently, auralisations have been coupled with virtual reality (VR) visualisations to enhance the user experience, through use of VR headsets.

Recent computational power increases in portable consumer-grade devices, such as smart phones and tablets, have meant that rendering a 3D augmented reality (AR) or VR simulation is now a possibility on such devices. AiHear® is an AR/VR auralisation application which utilises the portable device platform in conjunction with an off-the-shelf (calibrated) headphone setup to playback accurate auralisation audio. This paper looks at some of the technical aspects of successfully implementing an AR/VR auralisation application for use on a low-cost consumer-grade platform.

1 INTRODUCTION

Auralisation is a powerful tool for presenting acoustic design options to stakeholders, enabling them to preview an acoustic environment using appropriately processed audio signals. Current state-of-the-art auralisations require a dedicated listening room with very low reverberation times, very low background noise, and a multi-channel loudspeaker setup. More recently, auralisations have been coupled with virtual reality (VR) visualisations to enhance the user experience, through use of VR headsets. However, such listening rooms setups are very expensive and are in no way readily accessible or portable.

Recent computational power increases in portable consumer-grade devices, such as smart phones and tablets, have meant that rendering 3D graphics in an augmented reality (AR) or VR simulation is now a possibility on such devices. Simultaneously rendering binaural audio in the 3D environment is also a possibility, whereby the audio source direction moves/rotates relative to the participant's head position/orientation in the 3D simulated world.

The use of headphones as the audio playback medium has many advantages for a portable AR/VR auralisation system, as discussed later in this paper. Through careful calibration of the AR/VR system playback levels and frequency responses, it is possible to reproduce audio signals with a high degree of spectral accuracy, thus making this platform suitable for engineering-precision auralisations.

Over the past several years, NDY Acoustics has incrementally developed a successful software implementation of AR/VR auralisation, called AiHear®, which utilises portable consumer-grade device technology as its platform. The features of the AiHear® application, along with some of the technical challenges, are discussed in this paper.

1.1 VR versus AR

VR technology has been in common use for several decades now, and many people are familiar with this form of simulation. In virtual reality, the user is fully immersed in a simulated visual environment, whereby they real world is completely replaced by the simulated world.

AR is similar in many ways to VR, in that the user experiences simulated elements which do not exist in reality, however, the simulated elements do not necessarily block out the real world. Instead, the real world is enhanced or augmented with the simulated elements; for example, a simulated chair can be placed within a real room and will remain stationary within the room, even when the user's viewpoint moves/rotates.

2 THE SYSTEM

2.1 AiHear® Hardware Setup

To make the AiHear® auralisation system as portable as possible, and to keep setup costs low, the application utilises only two key hardware components:

- An Apple portable device (iPad Air 3rd generation, iPhone X or better, running iOS 11 or newer*).
- Medium-to-high quality stereo headphones which have been calibrated through the AiHear® system. The current list of calibrated headphones includes: Audio-Technica ATH-M40x and JBL Live 660NC.

*For lower latency and better AR spatial tracking, an Apple device equipped with LiDAR sensors is recommended (e.g., iPad Pro 2nd generation, iPhone 12 Pro).

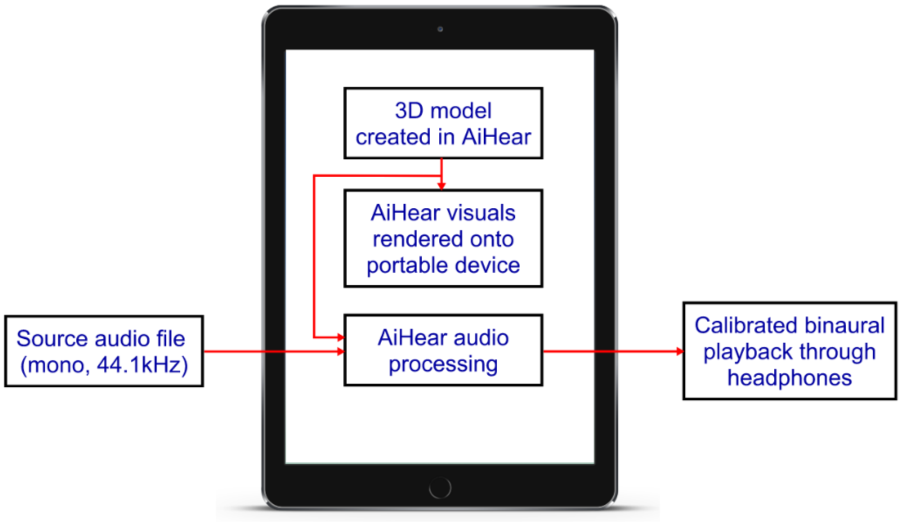
NDY (Tim Beresford, 2021)
Figure 1: Typical AiHear® hardware components

2.2 Simulation Operation

In operation, AiHear® combines pre-recorded audio signals with a 3D simulated model to produce an auralisation which is linked to AR/VR visualisations. Users can choose between AR or VR visuals, depending on their required application.

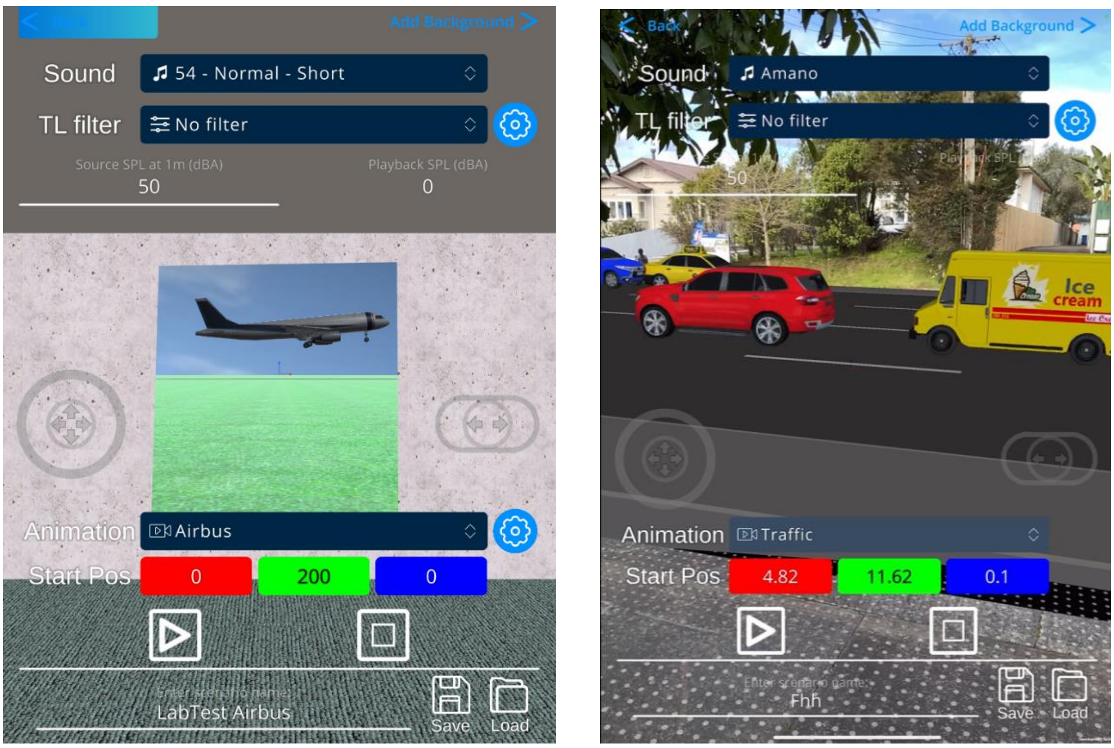
The environment to be auralised is created in 3D using the portable device running the AiHear® application, as follows:

- The simulation room is created by placing coordinates for each of the room's corners, or by using the simple box-shaped room creation tool.
- The room surfaces are assigned sound absorbing materials (with associated 1/3 octave absorption coefficients).
- The sound source is placed at the desired location, either inside or outside the simulated room. The source type can be selected as a stationary point source, moving point source or line source.
- 1/3 octave transmission loss filters can be applied to represent different sound insulation properties of the room's walls, for the situation where the sound source is located outside.



NDY (Tim Beresford, 2021)
Figure 2: AiHear® functional schematic

The portable device's screen acts as the lens though which the simulation is viewed. By moving the device around, the user can look around the simulated room which will appear fixed in space relative to the real world. The relative motion/rotation of the portable device to the simulated sound source alters the directional arrival of the audio signals to the user's ears to create a 3D audio environment. For example, a simulated point source moving from left to right across the user's point of view will audibly shift from being louder in the left ear to the right ear to match the visuals. Similarly, if the point source is stationary, but the portable device's point of view is rotated away from the source to the right (clockwise), the audio will become louder in the left and quieter in the right ear.

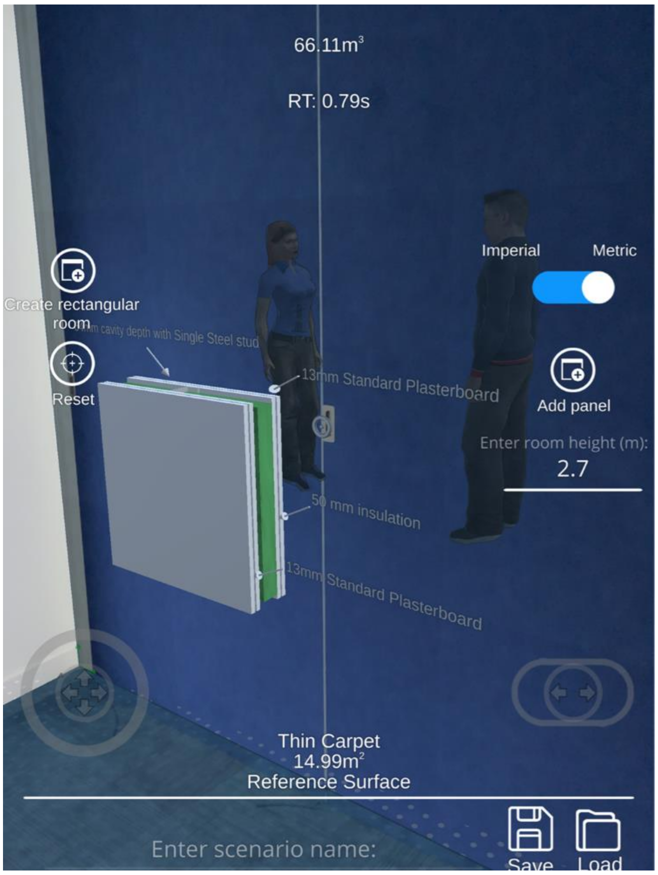


NDY (Jack Wong, 2021)
Figure 3: Examples of a simulated moving point source in VR (left) and line source in AR (right)

2.3 AiHear® Features

2.3.1 User-defined absorption coefficient and transmission loss data

Users can edit and save their own absorption coefficient and transmission loss data as either 1/3 octave or 1/1 octave band values. Absorption coefficients are saved as “materials” which also have user-defined visual textures and colours. An additional feature attached to the transmission loss tool is the ability to visualise wall constructions using the in-app wall build-up tool. Users can move closer to the wall cross section to inspect it, whilst hearing the transmission losses the wall construction has to offer.



NDY (Jack Wong, 2021)

Figure 4: An example of the wall build-up visualisation tool. In this scenario, a conversation is being simulated in the next-door room whilst the sound insulation performance of the separating wall is demonstrated.

2.3.2 Advanced room geometry setup

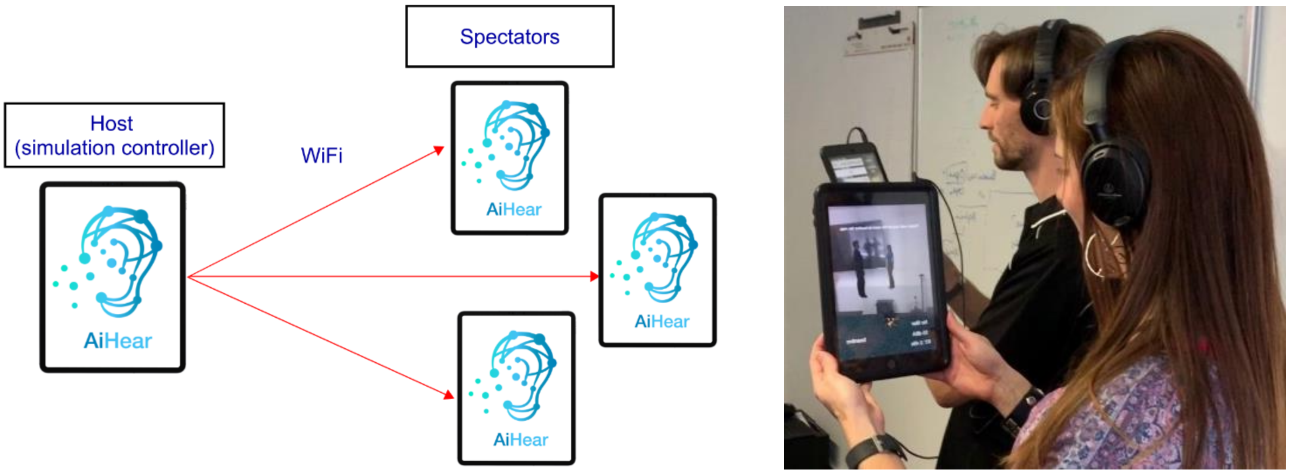
In addition to the in-app method of simulated room creation, advanced users can create rooms by writing the corner cartesian coordinates (x, y) to a .CSV file which is loaded into AiHear®. There is no limit on the number of corners (or sides) a room can have.

2.3.3 Saving and loading simulation scenarios

All simulation parameters can be saved as scenarios, which can then be loaded by the host during a presentation to rapidly switch between different scenarios. This enables direct A/B comparison of different acoustic treatment options.

2.3.4 Multi-user mode

AiHear® allows multiple users to join the same simulation, each with their own portable device. One user’s device hosts and controls the simulation parameters while other users’ devices can join and spectate. Each user views the simulation from their own point of view within the simulated environment and can move closer or further from the sound source, look towards or away from it, at their own free will. The theoretical maximum number of spectators is 100, although this has not yet been tested.



NDY (Tim Beresford, 2021)
Figure 5: Multi-user mode

The host user has control over the source audio file, sound source type and visual model, room surface finishes for reverberation control, transmission losses and background noise.

3 TECHNICAL DISCUSSION

This section provides technical discussion on a few key details which apply to the AiHear® method of auralisation.

3.1 Headphones for auralisation purposes

The use of headphones as the audio playback medium for auralisations has many advantages over the more typical “listening room” setup utilising multi-channel loudspeakers in an acoustically dead and quiet room. However, implementing headphones for AR/VR auralisations does present some additional technical challenges.

Table 1: Key advantages and disadvantages of using headphones for auralisation purposes

Advantages	Disadvantages
Highly portable	Head related transfer functions must be applied
Low cost	Headphone fit is important for good/repeatable frequency response
Good frequency response across the entire audible frequency range	Calibration of the headphones must be completed using an expensive binaural dummy head setup
Playback sound levels are independent of the environment in which the presentation is physically being held	The user’s head orientation relative to the sound source must be fed into the auralisation engine for correct binaural playback (panning)
Sound insulating headphones render the ambient noise levels in the physical presentation environment largely irrelevant. Auralisation of very low sound levels can be accurately presented	
Reverberation within the physical presentation environment is not a limiting factor, i.e., anechoic and free field scenarios can be auralised	

3.2 Modelling of reverberant and direct sound components

Although flexible enough to use in range of audio applications, AiHear® has been designed primarily as a tool for acoustic engineers. As such, the calculation algorithms within the application are based on the widely used classical room theory for predicting sound pressure levels inside rooms (Bies and Hansen, 2003):

$$L_p = L_w + 10\log_{10}\left(\frac{1}{4\pi r^2} + \frac{4}{R}\right) \quad \text{(for an omnidirectional point source)} \tag{1}$$

Where L_p is the total sound pressure level at the receiving (user) position, in dB

L_w is the source sound power level, in dB

r is the distance from the sound source to the receiving (user) position, in metres

R is the room constant, related to the mean absorption coefficient for the room.

In this equation, the direct and reverberant sound components are handled separately and then combined, and AiHear® replicates this approach. Using this theory enables users to easily match the in-app calculated sound levels and reverberation times to their desktop design calculations.

3.3 Air absorption

During development and testing of AiHear®, inclusion of sound absorption due to air (α_{air}) was found to be necessary to improve the realism of the auralisations. This was most important at frequencies from around 4kHz through to 20kHz, where the air absorption becomes rapidly greater with frequency (Harris, 1967).

3.4 Realtime 1/3 octave band filters

One key technical challenge of developing AiHear® was creating 1/3 octave band filters which sounded natural (free of audible artifacts), were acoustically accurate with good roll-off characteristics, and which could be implemented in real-time on consumer-grade portable devices across the full audible spectrum (30 bands simultaneously). Through trialling many different filter configurations, an acceptable balance between accuracy and processing speed was struck and implemented in the application.

3.5 Pre-recorded audio signals

Users can upload their own audio files to AiHear®, however, the quality of these audio files is of critical importance to the accuracy of the auralisation:

- The application only accepts audio files which are mono, 44.1kHz .WAV format
- The spectral balance of the audio files from 20Hz to 20kHz must match the intended sound source spectrum.
- Ideally the recordings should be anechoic/free-field or as close to this as possible. Reflections recorded in the source audio file will interfere with and muddy the reverberated sound playback.
- Time varying components of the audio signal must be carefully considered in the context of the auralised scenario. Generally, it is preferable to have source signals which do not vary in level over time. For example, if simulating a single motorbike pass-by as a moving point source in AiHear®, the source audio file should consist of a constant level recording of a motorbike. The movement of the motorbike, as well its sound level, will be varied accordingly within AiHear® based on its distance from the receiver location.

3.6 Background noise

One highly critical aspect in most auralisations is the background noise which is present during the simulation. This background noise provides audio masking which can either enhance the simulation's accuracy (if properly controlled) or significantly degrade it (if too high or too low). It may be obvious that if the background noise is too high, it will mask important sounds during the auralisation. Similarly, if the background noise is too low, auralised sounds which would normally be masked can appear more prominent than they should be. Therefore, it important that the background noise is controlled and played back at the correct level.

In the case of auralisations using headphones (assuming the headphones provide a good degree of sound insulation), it is most likely that when the user puts on the headphones, the background noise will drop below where it should be for the purposes of the auralisation. However, having background noise that is too low is preferable to it being too high because it is easier to add in background noise than it is to take it away.

AiHear® has a feature which allows users to load in background audio files and have them played into the auralisation at user-defined sound pressure levels. The control of this background noise is independent of the primary sound source audio controls so the two sources can be overlaid in the auralisation to provide the required masking effect.

3.7 System calibration

The AiHear® application was successively calibrated using a binaural head (and torso) simulator for different portable devices and headphone models. The calibration values were saved as 1/3-octave filters within the application allowing users to select the appropriate filters depending on their specific hardware setup. This calibration process meant the audio playback accuracy of the system could be verified. Calibration was completed with the listener facing directly towards the sound source, i.e., with a neutral-position head transfer function applied.

NDY (Tim Beresford, 2021)
Figure 6: AiHear® system calibration being undertaken on a head (and torso) simulator

4 FINAL REMARKS

Auralisation is a valuable tool for previewing how an acoustic environment will sound before it is constructed. Coupled with AR/VR visuals, auralisation can be a very immersive experience, and the power of this type of multi-sensory simulation holds a strong place in the future of acoustics.

AiHear® is an innovative front-runner in AR/VR auralisation technology, providing a platform which is immersive, low-cost, highly portable, and acoustically accurate. We look forward to sharing this technology with anyone who has an interest in acoustics.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the global Tetra Tech team for their ongoing support in the development of AiHear®, and in particular, Arif Zaher, for his tireless enthusiasm for innovation and the development.

We would also like to thank the University of Auckland's Acoustic Testing Service for allowing us to complete the system calibration testing in their laboratory, and Autex Insulation for their interest in, and valuable feedback on the application.

REFERENCES

Bies, D. A. and Hansen, C. H. (2003). *Engineering Noise Control*, 3rd ed. Spon Press, London.
Harris, C. M. (1967). *Absorption of sound in air versus humidity and temperature*, National Aeronautics and Space Administration, Washington, D. C.