

An immersive journey preparation tool for people with vision impairment

Dung Lai, Chris McCarthy, David Sly, Harrison Bennett, Matt Shackleton, Stuart Favilla

Introduction

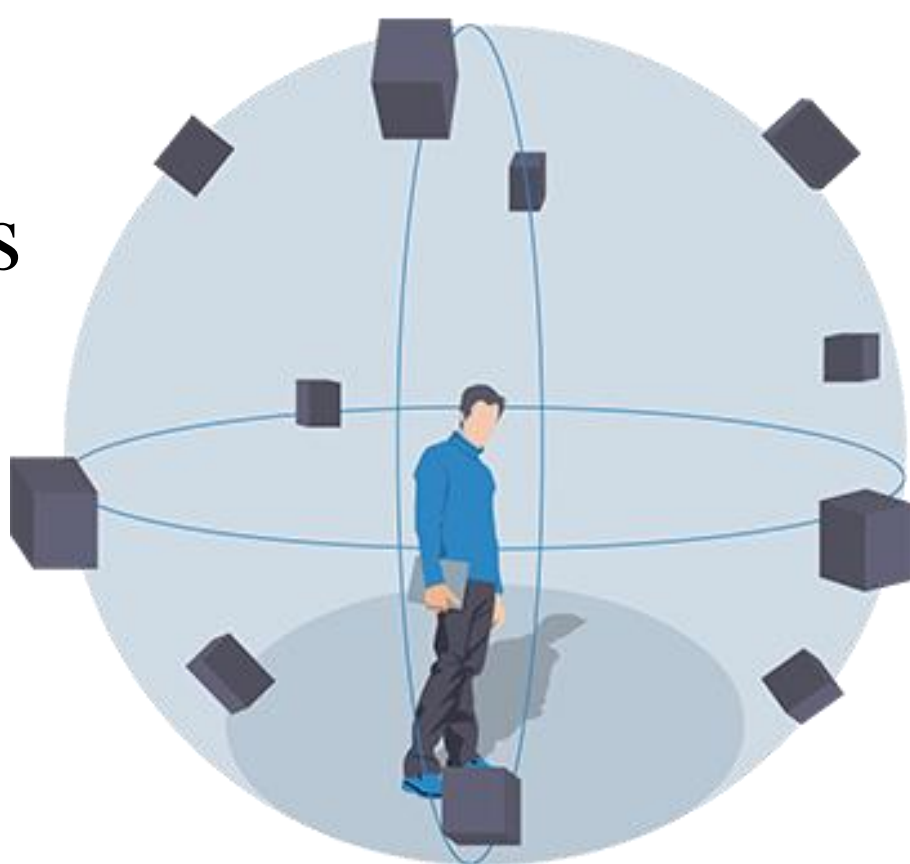
With almost one in five Australians experiencing some form of disability, a large proportion of the community face challenges to actively participate in city life.



This project demonstrates a long-term technology-enabled solution to issues experienced by people with vision impaired.

Methods

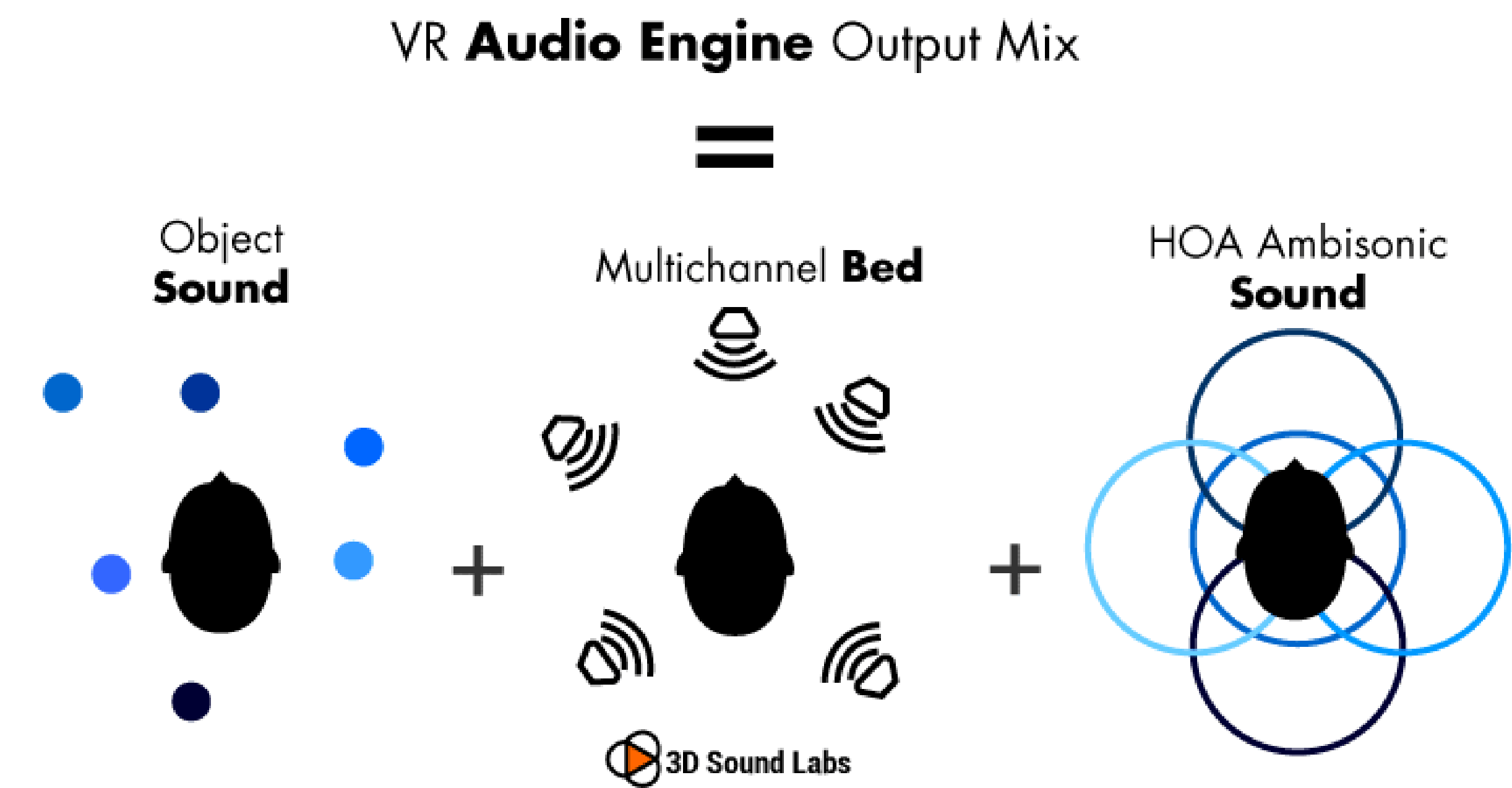
We develop an auditory-based simulator to simulate the sensory experience of a specific location in Melbourne’s CBD. The idea is to provide people who have vision impairment and blindness an immersive tool that allows them to experience the sounds of environments they plan to walk through. The simulator would allow them to rotate their body and hear the sounds change as they rotate.



In our experiment, we choose Flinders Street Station as the testing environment

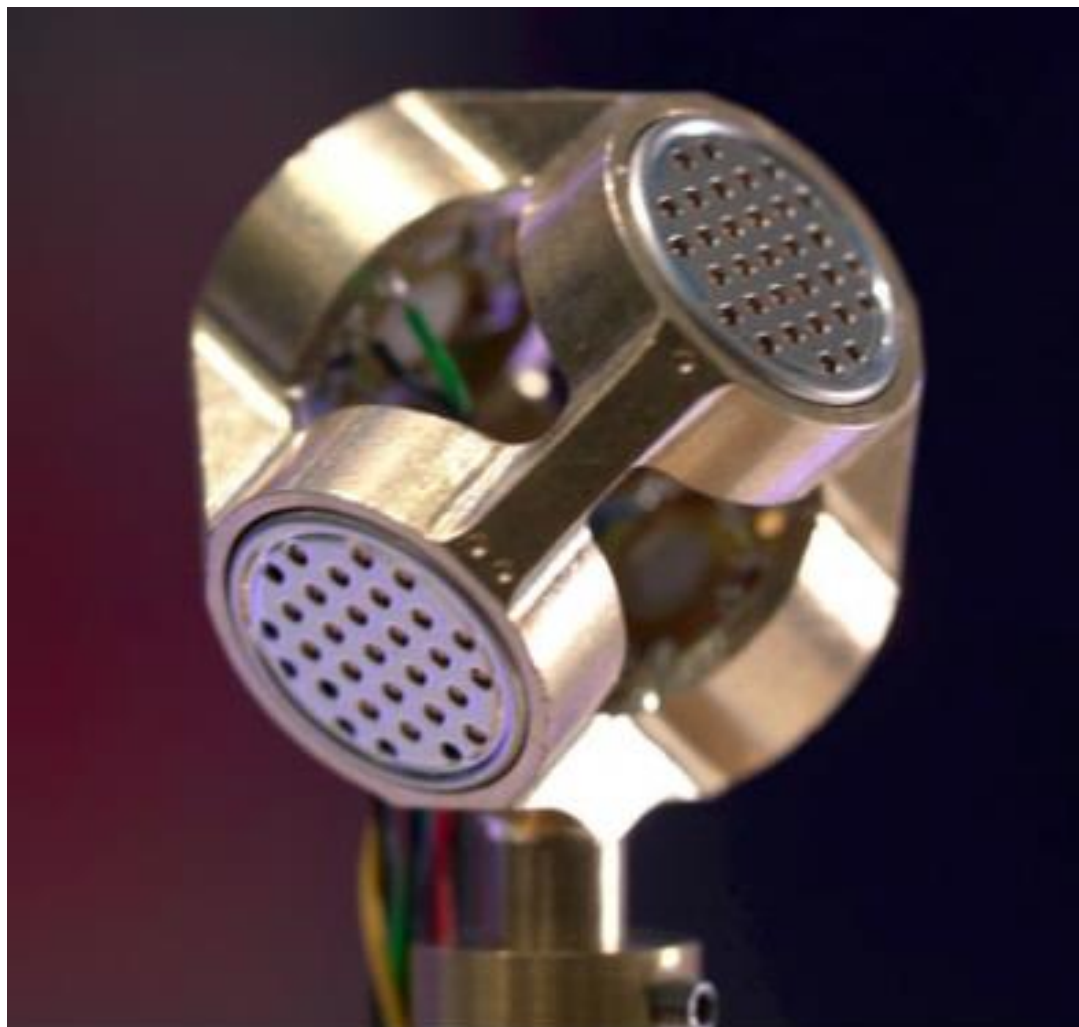
Surround Sound Technology

In order to archive what we call “immersive” experience, we use Ambisonic Technology [1]: a full-sphere surround sound technique, in addition to the horizontal plane, it covers sound sources above and below the listener



Recording Technique

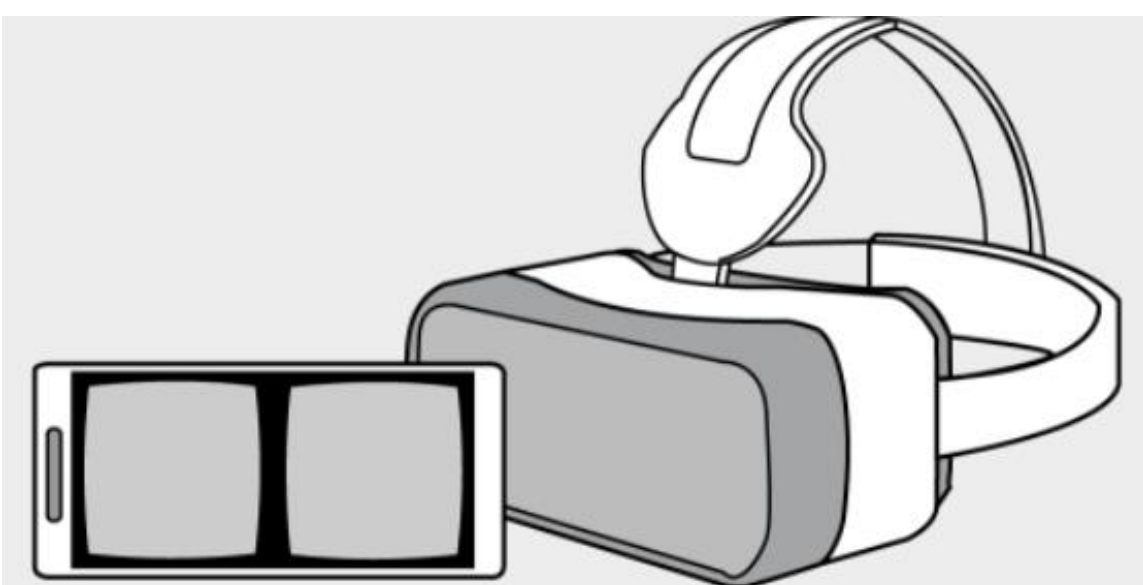
Sounds are recorded by a tetrahedron microphone (left image) to get 4-channel monophonic A-format. This is converted into 4-channel B-format using mathematical formula shown below [1]. B-format file contains XYZ directions which covers all 3 dimensions. W channel is called omnidirectional.



$$W = \frac{1}{k} \sum_{i=1}^k s_i \left[\frac{1}{\sqrt{2}} \right]$$
$$X = \frac{1}{k} \sum_{i=1}^k s_i [\cos \phi_i \cos \theta_i]$$
$$Y = \frac{1}{k} \sum_{i=1}^k s_i [\sin \phi_i \cos \theta_i]$$
$$Z = \frac{1}{k} \sum_{i=1}^k s_i [\sin \theta_i]$$

Web Application

The Ambisonic sound capsule has been encoded in 360 video using Facebook 360 Encoder and deployed in Youtube. This will make the video compatible with VR mode. Users can use Google cardboard or Oculus Rift to navigate and change direction. The video can also be rotated by sliding the video or using controller on the top left of the video.



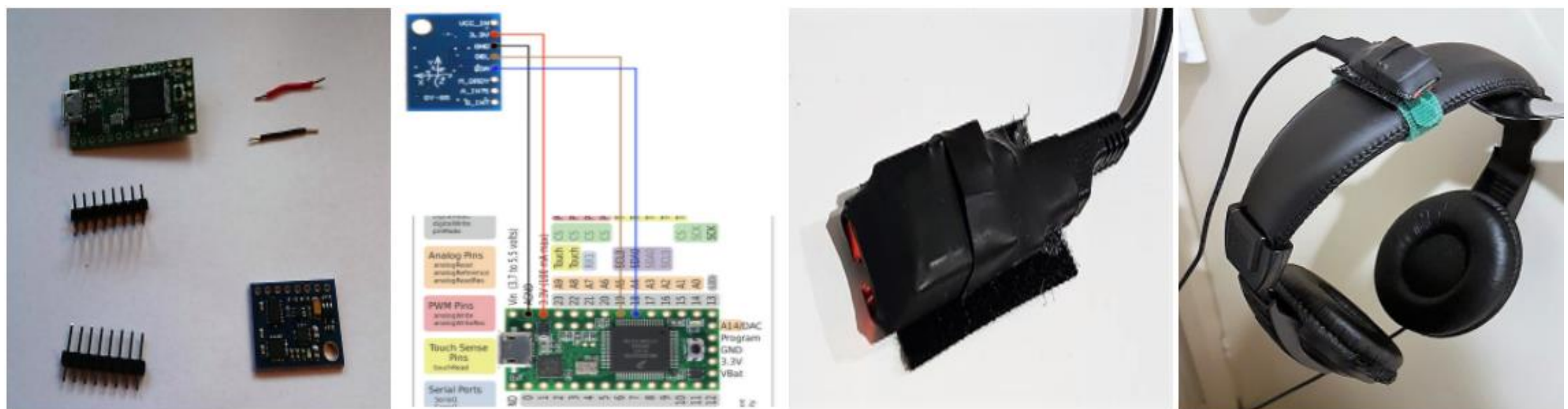
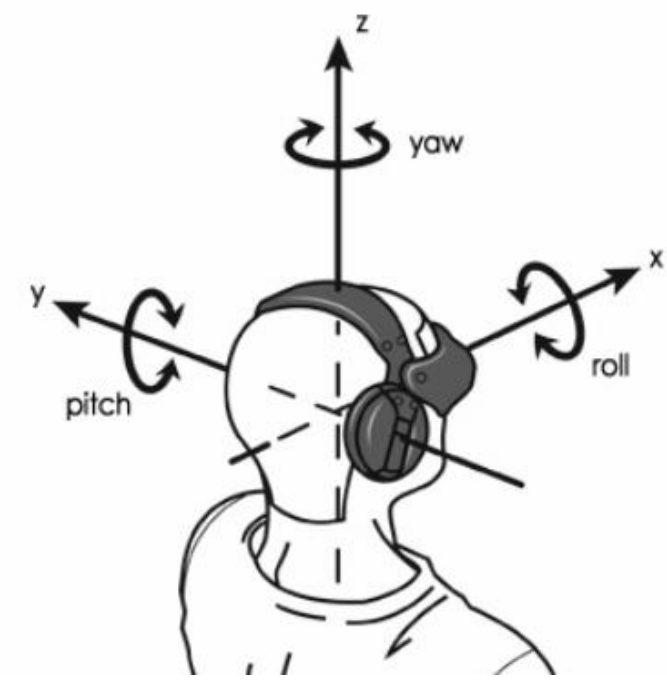
For a mobile user, rotating mobile device will also move the direction of the camera.



Voice Recognition and Speech Synthesis: Users can interact with the interface using voice command and the audio cues will be played back upon user’s request

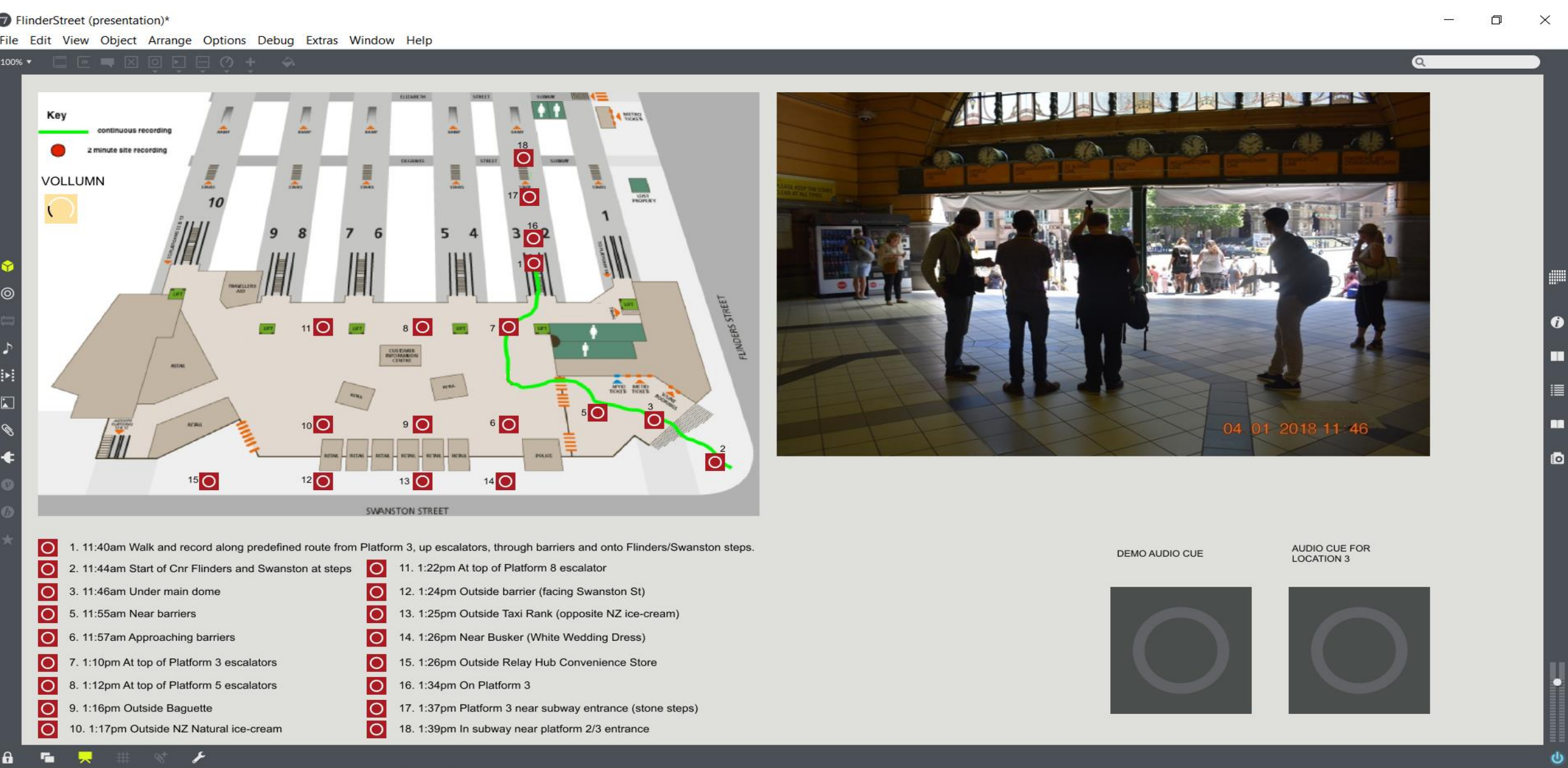
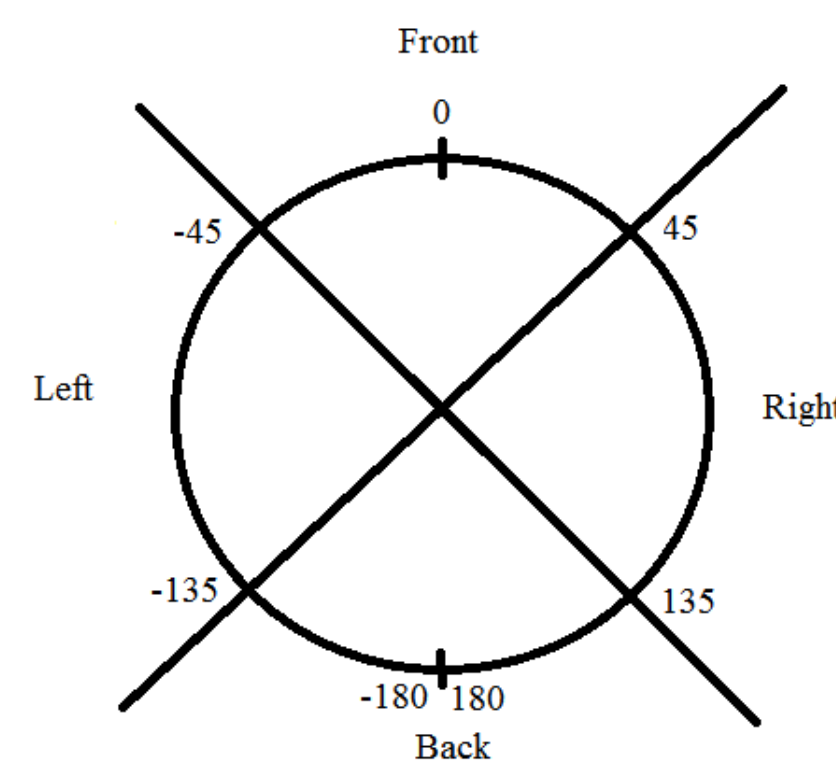
Desktop Application

For this platform, thanks to Stuart Favilla for the idea of using and integrating a head-tracker which can track yaw-pitch-roll value of a user’s head. This value is used to change sound field and generate directional narration. The image below shows the open-source head-tracker we used, called Hedrot and its integration on headphone.



Narration and Audio Description based on user’s head direction

The surround area will be described, description is on-request and dynamically adjusted based on head direction.



Main Equipment

- Garmin Virb 360 Video
- Sennheiser Ambeo VR Microphone
- Zoom H6 recorder
- Teensy 3.2 board, Hedrot headtracker

References

1. An Introduction to Higher Order Ambisonic
(<http://flo.mur.at/writings/HOA-intro.pdf>)

Project website

<https://dunlai.github.io/COM/index.html>

