An immersive journey preparation tool for people with vision impairment

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

SWINBURNE UNIVERSITY OF TECHNOLOGY

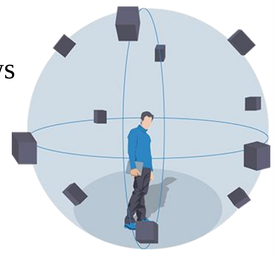
Introduction

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 is developing a long-term technology-enabled solution to assist journey preparation for members of vision impaired community.

Methods

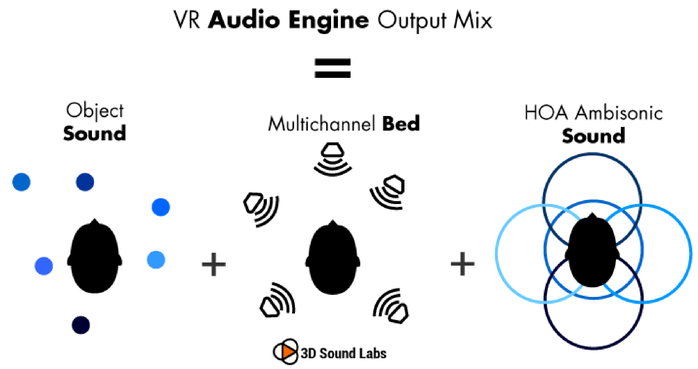
We develop an auditory-based simulator to simulate the sensory experience of a specific location in Melbourne's CBD. The idea is to provide people who have vision impairment and blindness an immersive tool that allows them to experience the sounds of environments they plan to walk through. The simulator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



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

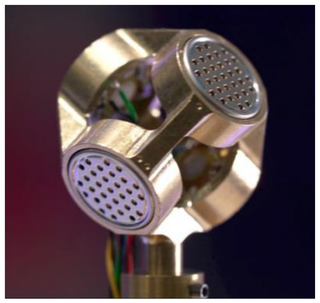
Surround Sound Technology

In order to achieve an "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.



$$W = \frac{1}{k}\sum_{i=1}^{k} s_i\left[\frac{1}{\sqrt{(2)}}\right]$$

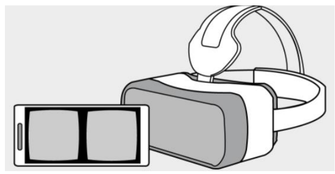
$$X = \frac{1}{k}\sum_{i=1}^{k} s_i[\cos\phi_i \cos\theta_i]$$

$$Y = \frac{1}{k}\sum_{i=1}^{k} s_i[\sin\phi_i \cos\theta_i]$$

$$Z = \frac{1}{k}\sum_{i=1}^{k} s_i[\sin\theta_i]$$

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.



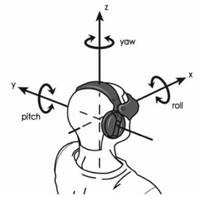
An auditory-based simulation of Flinders Street Station:

Voice Recognition and Speech Synthesis

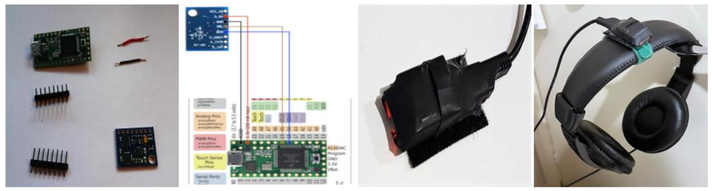
Users can interact with the interface using voice commands and then narration will be played back upon user's request.

Desktop Application

A Desktop application is also being developed to explore the use of headtracking, allowing the tracking of yaw-pitch-roll value of a user's head. This value is used to change sound field and generate directional narration.

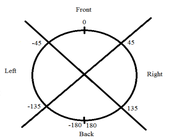


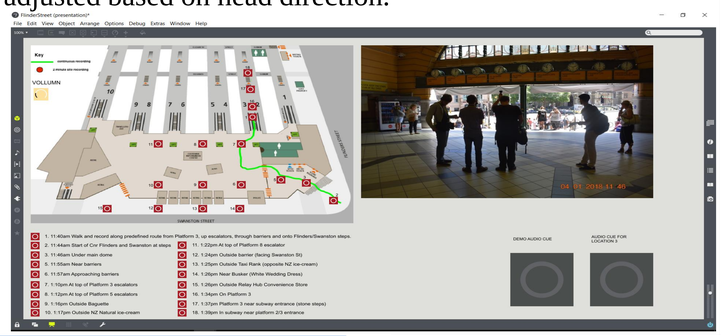
Thanks to Stuart Favilla, the Hedrot head-tracker is successfully implemented and the integration is shown below



Narration based on head direction

By separate the Yaw value from headtracker into 4 quarters representing front, left, right and back, 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 Mic
- Zoom H6 recoreder
- Teensy 3.2 board, Hedrot headtracker

Project website

https:\\dunglai.github.io\COM\index.html

References

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