Faculty of Engineering
Summer Research Program 2019-2020

Project Title: Visual Attention in 360 degree videos
Supervisor(s): Prof. Tom Drummond, Dr. Akansel Cosgun
Department: Electrical and Computer Systems Engineering
Email: tom.drummond@monash.edu, akansel.cosgun@monash.edu
Website profile of project supervisor: https://www.monash.edu/engineering/tomdrummond, http://acosgun.com/

Objective
360 cameras has been around for only a short while and they have great potential to be used for robots. Compared to standard cameras, 360 cameras provide a lot more sensory information from the environment, making it much harder to process these images that capture a wide area. However, image context can be used to focus on visual attention. For example, the sky or ground isn’t interesting but the people are. Similarly, computer vision algorithms can focus on the labels, street names or texts in the image for a better understanding of the world. The objective is to develop algorithms for visual attention: which parts of the image is ‘interesting’ to have a closer look at?

Project Details
We will focus on a particular application: Given a 360 video recorded in a city center, can an algorithm infer the coordinates of the camera in Google Maps? This would require detecting street names, detecting the road and localizing with respect to prior images.

This project will have two parts.
1) Interfacing the camera images: This involves potentially writing drivers to access the camera images, and post-process it so that they are ‘unwrapped’.
2) Computer Vision: This is the main focus of the project. We will use state-of-the-art Deep Learning algorithms to detect text in the image, detect buildings as well as roads and the sky. The task would be to estimate the GPS coordinates (and direction), given the input image and previous ground truth data.

Prerequisites
- Strong programming skills are required (Python or C++ preferred). Exposure to Linux is a plus.
- Interest in publishing the research in a scientific conference