Computational Creativity Project Proposal

Project Overview

The aim of this project is to build a system that will be able to detect and recognise images. The idea is that the system will be able to take an input of an image, analyse it and describe the image to the user. Essentially, it will comprise of an image recognition program which then in turn will generate appropriate text to go alongside the image. This text will then be converted into an audio clip so it can be played back to the user. The purpose of this is to enable visually impaired users to be able to understand what an image is representing.

Project Plan

My goal is to tackle the project in stages working on both the code and the main report simultaneously. This can be shown in the breakdown below:

Stage 0	• Construct initial project proposal • Research my idea • Identify possible libraries to assist with implementation
Stage 1	• Begin initial development of project • Write abstract and background of the report
Stage 2	• Continue development until the system has finished the development stage • Perform testing of the system • Write introduction to the report
Stage 3	• Evaluate the system with unit tests • Write introduction, methodology and design sections of the report
Stage 4	• Perform evaluation of the system • Write up results and conclusion for the report • Record video of the system working

Technical

After presenting my idea to Colin and Caio, both have informed me that idea is on the complex side. As a result, I have researched my idea online and have found several libraries that will be able to assist me in the development stage of the project. After conducting some research, I have found a Python library called *ImageAI* which I believe I can use for this project. I will require large datasets in order to 'train' the system to identify images which I will find online in order to save time as hundreds to thousands of images will be required. As a result of this, one limitation I will have is that my system will not be able to identify any image you provide. It will only work with images that are relevant to the dataset it contains. If I had a large enough dataset, the system would in theory be able to identify a much larger variety of images.

Evaluation

The best approach to evaluate the results in my opinion is a variation of the Turing Test. If the images are going to described/recognised by a piece of software, it is important to determine whether humans think of the same words/phrases that the software produces. The main area of evaluation will be the text output that is generated for each image.