Initial report - Text detection and text recognition in images for Text2Speech modules

TEAM: SA06_01

Eugen Leonid Cocalea Creoleanu (ELCC)
1408A

1 Project overview

The application we're going to develop is a Proof of Concept application that is attempting to identify text in images, recognize that text and output that in a format ready to be picked up by a 3rd party application.

Text detection and recognition applications can be used for a multitude of purposes like translating instructions (recipes, street directions, attraction descriptions, newspaper articles) written in a foreign language, helping visually impaired or illiterate people get around tasks that involve reading instructions, creating Augmented Reality applications that take information from the real world and present it to the user. The eventual end goal is to have the application run on portable / wearable computers that will accompany the user in the real world and continuously find information to work on in the surroundings.

The objective of the project is to achieve a reasonable rate of text detection and recognition in images that are already acquired and fed into the application while leaving its integration with 3rd party applications (like portable / wearable computers, translation engines, text2speech engines) for a later project.

Expected challenges are mostly related to the lack of field knowledge by the team and the scarcity of time to be allocated to the project.

Potential consumers: tourists, visually impaired people, illiterate people, kids learning to read, neglected kids that needs a story read to them one evening.

Competitors: Meta (1), Apple (3), Google (4)

2 Development approach

Task allocation

Task ID	Task description	Team member	Estimated time
Research	Research on existing solutions and technologies	ELCC	8h
Design	Application design Choosing relevant technologies	ELCC	8h
Implementation	Application development Application testing	ELCC	48h
Documentation	Writing application documentation	ELCC	8h

Git repository: https://github.com/Pucster/sa06t2s

References

[1] Meta AR/VR

[2] Google OCR

[3] Apple Glasses

[4] Google Glasses