Initial Report

Video to Audio Conversion for the Visually Impaired

Joseph Redfern
Supervised By: Dr. Kirill Sidorov
Moderated By: Dr. Andrew Jones

Module Code: CM3203
Module Title: One Semester Individual Project
Module Credits: 40

February 2015

1 Project Description

According to the World Health Orgnisation [4], as of 2012 there are 285 million visually impaired people, 39 million of whom are blind. The same source suggests that 90% of the visually impared live in developing countries. Official data from the Guide Dogs for the Blind Association [1] suggests that a the lifetime cost of a guide-dog is around £50,000. This paints a sad picture of the support available to a majority of the visually impaired — that kind of money is not available to those living in third world countries.

With that in mind, the goal of this project is to allow a visually impared or blind person to navigate around a room without the use of a guide dog, or other assistance. By coming up with a cheaper alternative to guide-dogs, it would be possible to make life-changing technology much more accessible to the people all arond the world. This project aims to develop a prototype system to so through the use of a depth sensing camera [2] and a laptop. By analysing the image and depth data from the camera, we believe it possible to produce an audio stream that can convey the importatant information that would normally by recieved by the human visual system.

Although some similar research and development has been undertaken [3], existing solutions appear to suffer from "information overload", and seem too low-level to be practical in navigating around a room. By transmitting only relevant information to the user of the system and relying on their intuition to mentally re-construct the scene, we hope to create a system that can be used in real-world scenarios. In this sense, the project is unique — rather than relying entirely on computer vision techniques, we can rely on the human brain to help understand and contextualise the scene.

2 Project Aims and Objectives

The main aim is to devise a method of encoding video (high-bandwidth in nature) as sound (of a lower bandwidth) — it is fundementally a compression problem. In order to solve this problem, the following objectives need to be completed:

- To investigate and gain an understanding of the main issues faced by the visually impared.
  - This will likley involve analysing past research and medical literature
  - It may become necessary to conduct interviews/survey of the visually impared
- To be able to interact with an Asus Xtion PRO LIVE [2] using Matlab
  - Investigate existing MATLAB modules — they may require some modification to allow Depth/RGB image registration

- To be able to filter irrelevant information from the video stream, so as to avoid "information overload"
  - Decide what constitutes an "important feature" — i.e. faces, near-by walls etc
    - This may involve some degree of experimentation/trial-and-error.
  - Develop ways of extracting important features
- To develop a functional prototype capable of conveying relevant image data in the form of audio:
  - The best way to "encode" the information needs to be established. This may involve using different MIDI instruments to transmit different information in the form of melodies.
  - The system should be able to use biaural playback techniques, helping the user of the system to infer the direction of an object.
  - Information extracted from the video should be identifiable in the transmitted sound.
  - The system should be able to process the video in real-time.

Given the amount of time available, it would be an un-realistic to attempt to develop a system to entirely replace current solutions. Instead, the aim is to develop a prototype that can prove that the system is feasible — a 1% improvement to the quality of life of a visually impaired person would still be a great acheivement.

3 Ethical Considerations

In its current form, the project does not meet the criteria defined by the University for an ethical review to be necessary. We do not currently anticipate needing to conduct interviews or surveys, and the initial prototype will not undergo any clinical trials.

Should this change, the appropriate approval and documentation will be acquired — this may take some time.

4 Work Plan

I intend to have weekly meetings with my project supervisor, where current progress can be evaulated and the system can be tried out.

4.1 Weeks -8-0

No real work has been completed - some background research into the subject has been done, depth sensing camera [2] has been ordered.

4.2 Week 1

- Complete Initial Report

4.3 Week 2

- Begin research into issues faced by the visually impaired
  - It may be necessary to conduct interviews or surveys. In such case, this task will be delayed by 1 week as the necessary ethical approval is acquired. A buffer over the Easter Recess allows for this.
- Investigate existing solutions used by the visually impaired
- Test and Configure MATLAB for use with the Xtion Camera
  - This assumes that the camera will be available in Week 2 - if the department has not yet taken delviery of the Xtion, the task may need to be delayed

4.4 Weeks 3-4

- Complete any research (as started in Week 2)
- Gather test data from a room with obsticles
- Experiment with filtering video, removing un-necessary features

4.5 Weeks 5-6

- Create basic template for final report, to be filled in throughout the project when appropriate
- If video filtering experimentation is incomplete, continue experimentation with filtering/compressing video footage
- Begin investigating how different components of the video footage can be encoded as audio

4.6 Weeks 7-8

- Implement biaural (directional) playback of the encoded audio
- Complete development of audio encoding techniques

4.7 Weeks 9-10

Easter Recess is between Weeks 9 and 10, which can be used as a buffer zone for any previous, un-expected delays.

- Begin Final Report
- Tie the various components together, create UI for the system

4.8 Week 10

- Final testing of system, implemenatation of any final bugfixes
- Continue work on Final Report

4.9 Week 11

- Submit Final Report

As the latter stages of the plan are dependant on earlier steps having been completed, and given that this is an exploratory project, there may be some slight in-accuracies with timing. These dates are not intended to be concrete. However, by using the Easter Recess as a buffer zone, this should not impact on the project deliverables.

References

[1] Guide Dogs for the Blind Association. *Fact File*. 2014. URL: http://www.guidedogs.org.uk/aboutus/guide-dogs-organisation/facts (visited on 01/28/2015).

[2] ASUSTek Computer Inc. *Xtion PRO LIVE*. URL: http://www.asus.com/uk/Multimedia/Xtion_PRO_LIVE/ (visited on 01/29/2015).

[3] New Scientist Joanna Carver. *Camera-to-sound app lets blind people 'see'*. 2012. URL: http://www.newscientist.com/blogs/nstv/2012/11/seeing-with-sound.html (visited on 01/29/2015).

[4] World Health Organization. *Visual imparement and blindness*. 2014. URL: http://www.who.int/mediacentre/factsheets/fs282/en/ (visited on 01/28/2015).