Bilkent University

Department of Computer Engineering

Senior Design Project

Project short-name: EyeContact

Project Specification Report

Melisa ONARAN – 21301232
Nazlı ABAZ – 21400231
Sarp SAATÇIOĞLU – 21400375
Yunus ÖLEZ – 21401539

Supervisor: Hamdi DİBEKLİOĞLU
Jury Members:
Selim AKSOY
Çiğdem Gündüz DEMİR
Innovation Expert: Çağla Çığ KARAMAN
Website: cs.sarpsaatci.com

Oct 9, 2017

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfillment of the requirements of the Senior Design Project course CS491/2.

Table of Content

1. Introduction

Improvement of every society depends on to the contribution of individuals in it. More people contributing to it with higher rates mean faster and bigger development. Governments endeavor to create better standards to increase production of individuals. However, individuals with disabilities do not have the same living conditions since they have special needs and thus, they cannot be as productive as others.

The aim of this project is to increase the living conditions of people who are visually impaired, and thus increase the development of societies. These individuals cannot get proper visually feedback from the environment which causes them to have a limited communication with societies. This project offers a solution for visually handicapped individuals that verbalizes the environment by using image processing using computer vision. Therefore, these people will be able to communicate, create and produce more.

1.1. Description

The purpose of this project is to solve the human-interaction problem of visually impaired people in terms of non-verbal communication during video chatting. This problem is encountered by almost all visually impaired people. Hence they are visually impaired, they do not have a chance to use video chatting applications such as Skype, Google Hangouts, etc. We propose a system that assists visually impaired people during video chats.

The system is able to recognize the person whose face is detected by the camera of the user's computer or external camera while video chatting and tells the identity of the person verbally. Furthermore, it also helps during the conversation by notifying whether the person is looking directly to the user while he/she is talking or not. This will provide user to be more interactive in the conversation in terms of non-verbal communication. Moreover, the system recognizes some certain emotions expressed by whom the user is making video conversation with.

Camera that the system uses will be user's platform's camera. Verbal notifications will be output from the platform with respect to user's choice of output method (e.g. headphone, speaker, bluetooth devices etc.). This system is based on computer vision technology in order to recognize faces that user

sees. The visual input taken from the camera is sent to the cloud platform so as to be processed and the verbal output is sent to the user's smartphone to be presented.

This system is available for cross-platform as a web application to be used by the users. Development of the user application will be done using React. Processing part will be handled with a web service on cloud. For computer vision, Python and OpenCV will be used running on cloud. Cloud computing will be handled with REST service methodology in order to be efficient.

1.2. Closely Related Technologies

During our research to find a similar product and technologies, we have discovered a few already in-use features, products and studies. Even though these mentioned technologies are similar, their aim differs from our product. While there some applications that help visually impaired people, the major applications that are similar to ours are Orcam, Horus, Be My Eyes and Emotion as a Service. Orcam is an application that helps visually impaired people through a camera that can be attached to the glasses. Through this camera, Orcam reads texts, recognizes street signs and faces, then through a small speaker gives feedback to the user [1]. Horus is a wearable assistant that recognizes objects and faces that has not yet been released [2]. Be My Eyes uses volunteers to help visually impaired through video chat. Volunteers answer the questions and describe the environment for the visually impaired [3]. Emotion as a Service is an application that analyzes emotional response through images and videos [4]. There are several more emotion sensing technologies such as Affectiva and Emovu [5, 6].

While there are many technologies as described above to help visually impaired people visualize their environment through audial feedback and detect emotions through images and videos, there is not a specific product that helps the communication and the socialization between 2 or more people by aural feedback through video chat that combines emotion detection, facial recognition and aural feedback.

1.3. Constraints

1.3.1. Implementation Constraints

- The application will run on cross-platform.
- Application part of the system will be developed based on React.
- Computer Vision part of the application will be done using OpenCV (Open Source Computer Vision).
- For the version control, GitHub and Git will be used in order to collaborate within our group.
- Cloud computing will be handled by use of third-party cloud platforms and node.js.

1.3.2. Technical Constraints

- Internet connection is required for complete functionality.

1.3.3. Economic Constraints

- The application will be a web application which enables the user to sign in with their accounts.
- The libraries and technologies may require additional purchases.
- The application will require a computer with a high resolution or a computer with an external camera.

1.3.3. Ethical Constraints

- We will not share information with any third parties.
- Users will be able to access their account by signing in.

1.3.4. Sustainability Constraints

- System will work as a web application for video chatting. Therefore, it will be a web application. Hence, in the future, system may be approved to be able to work as a cross platform application.

1.3.5. Social Constraints

- Application will be released in English language.
- The system will be designed on the basis of making an opportunity to visually impaired people to live video chat.

1.4. Professional and Ethical Issues

The application will use and collect visual data along with contact information. The main issue is privacy of both the user and the contacts. The application should maintain the personal information that is collected from the user confidentially. Other users should not be able to access others' private information. Since cloud computing will be used, data will be stored in third-party platforms.

2. Requirements

2.1. Functional Requirements

- User will be able to communicate with others through application's own video chatting feature.
- The application will detect human faces.
- User will be notified when other participants are looking towards them.
- When an unknown face is detected, the user will be able to add this face as a contact.
- The program will be able to recognize contacts by their faces and give a verbal feedback to the user.
- The program will determine emotions and also, give verbal feedback about it to the user.
- User will be able communicate with multiple people through the conference feature of the application.
- User will able communicate via text chatting, visually impaired people will use the build-in text-to-speech feature of the platform.

2.2. Non-Functional Requirements

- **Compatibility:** For the facial detection, the environment must be well lit for the camera to detect humans.

- **Robustness:** The camera that is used must be high resolution and sensitive to detect facial changes in order to determine emotions. Unrelated objects should be overlooked.

- **Performance:** System should run efficiently. As the application will use computer vision on cloud, the system should be fast enough to give feedback simultaneously.

- **Usability:** The system will be implemented in such a way that visually impaired people will have a chance to video chat. In other words, the visually impaired people will gain equal footing while live chatting.

- **User-friendliness:** Interface should be user-friendly. It should be easy to use for a visually impaired person, even with ones who are not good at using computers. Sighted-people will also use this application to communicate with visually impaired. Therefore, the interface should be inviting and easy-to-use for both sighted and visually impaired.

- **Adaptability:** The system should be used cross-platform. Different computers have different cameras, so the application should be adaptable to be used with different cameras.

3. References

[1] “Making the World Accessible: The Latest Tech for the Blind and Visually Impaired” https://www.mytherapyapp.com/blog/tech-for-the-blind-and-visually-impaired. [Accessed: Oct 09, 2017].

[2] “Horus Tech” https://horus.tech/?l=en_us. [Accessed: Oct 09, 2017].

[3]”The App That Lets Users Lend Their Eyes, And Blind People See Things In A New Way” https://www.fastcompany.com/3041054/the-app-that-lets-users-lend-their-eyes-and-blind-people-see-things-in-a-new-way. [Accessed: Oct 09, 2017].

[4] “Emotion as a Service” https://www.affectiva.com/product/emotion-as-a-service/. [Accessed: Oct 09, 2017].

[5] “Affectiva: How It Works” https://www.affectiva.com/#how-it-works. [Accessed: Oct 09, 2017].

[6] “Emovu” http://emovu.com/e/technology/. [Accessed: Oct 09, 2017].