Bilkent University
Department of Computer Engineering

Senior Design Project

Project short-name: EyeContact

High Level Design Report

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

Supervisor: Hamdi DİBEKLİOĞLU
Jury Members:
Selim AKSOY
Çiğdem GÜNDÜZ DEMİR
Innovation Expert: Çağla ÇIĞ KARAMAN
Website: eyecontact.in

Dec. 22, 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 Contents

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.

This report is the transformation of the analysis model into a system design model. It defines design goals of the project, and decompose the system into smaller subsystems. Strategies for building the system, such as the hardware/software platform on which the system will run, the persistent data management strategy, the global control flow, the access control policy, and the handling of boundary conditions are also mentioned in this report [1].

1.1. Purpose of the system

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 platform 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 JavaScript. Processing part and WebRTC(Real Time Communication) part will be handled on cloud. For computer vision, OpenCV will be used running on cloud. Cloud computing will be handled with REST service and WebRTC signaling methodology in order to be efficient.

1.2. Design goals

- **Compatibility:** Parts of the system must be properly working on all platforms that are supports WebRTC [2].

- **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 use computer vision on cloud, the system is designed to be able to give feedback simultaneously.

- **Usability:** The system is designed to 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 is about to be done to 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 people. Therefore, the interface should be inviting and easy-to-use for both sighted and visually impaired people.

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

1.3. Definitions, acronyms, and abbreviations

- **WebRTC:** WebRTC ("Web Real-Time Communication") is a collection of communications protocols and application programming interfaces that enable real-time communication over peer-to-peer connections.

1.4. Overview

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.

2. Proposed Software Architecture

2.1. Overview

Our project consists of 2 systems interacting with each other. One of them is responsible for data collection from camera and headphones; also recognition of the items. This system will run on a regular computer or phone.. This system will serve as a basic

requests system to our main application. Our second system will be responsible for its logic about the application part of the project and database connections. We will decompose the main system into smaller subsystems in order to reduce the complexity.

In the decomposition of the system to the manageable subsystems, 3-Tier architectural style will be used. 3-tiers of the system are composed of Presentation Tier, Application Logic Tier, and Data Storage Tier.

2.2. Subsystem Decomposition

For the main system of our project, we decided to have the client-server architecture style which is designed to be more general type of a 3-tier architectural style. Client side deals with only client side of the application whereas Server side is responsible for handling Database Management part of the system.

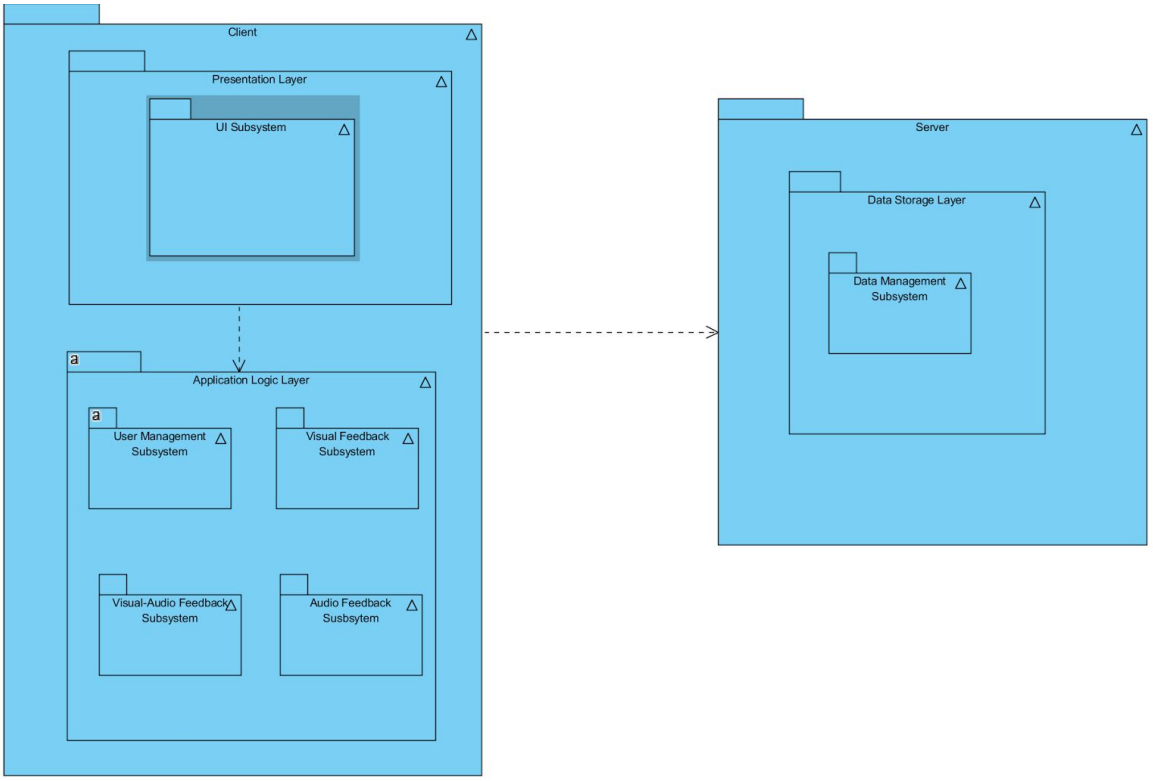


Figure 1: Subsystem Decomposition Diagram

2.2.1. Presentation Layer

Presentation Layer is designed to include only User Interface Subsystem. This subsystem is designed to serve as interface to the user while interacting with the layer down below to make sure that there is maintainability.

2.2.2. Application Layer

Application Layer is designed to be the layer where application logic is established. In this layer, main functionalities of the system will be performed such as visual-audial feedback, image processing, etc. This layer is composed of User Management Subsystem, Image Processing Subsystem, Visual-Audio Feedback Subsystem and Audio Feedback Subsystem. User Management Subsystem will be responsible from actions of users and interpret its commands. Image Processing subsystem will handle the image processing, afterwards Visual-Audio subsystem will be responsible for visual feedback to audio feedback transition and the Audio Feedback Subsystem will translate the detection output to the audio output for it to be translated through the headphones.

2.2.3. Data Storage Layer

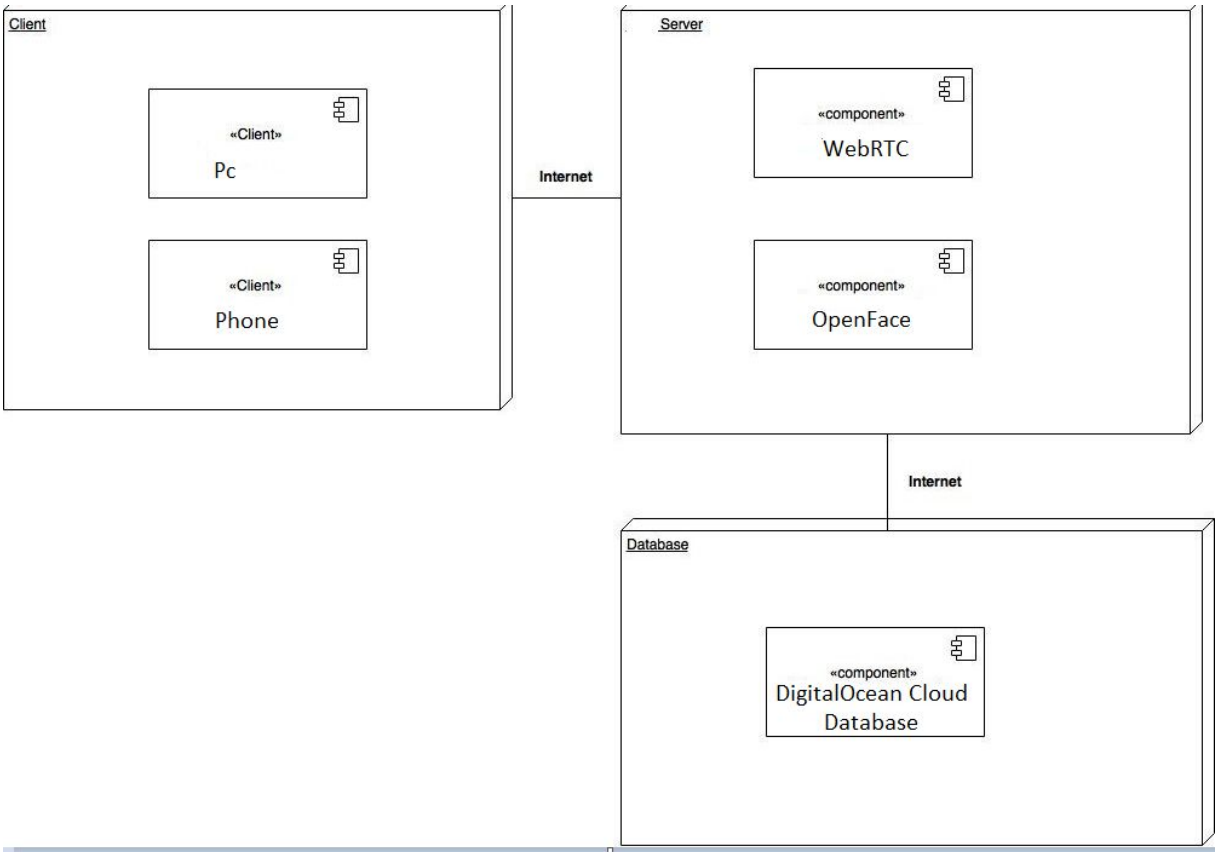
Data Storage Layer consists of Data Management Subsystem which stores the user data as well as contents of the user. This subsystem provides services to upper layer subsystem which are Application Layer and Presentation Layer.

2.2.4. Data Processing Layer

This layer will swap in data from Data Storage Layer and the data will be prepared and the computed data will be passed to the Data Storage Layer.

2.3. Hardware/software mapping

User will interact with the application but to do that an internet connection is needed. Our server and database will be up and running constantly so anytime users will be able to use the application without any information loss or anything likewise. However as we stated above to application to work and show the relevant information or enabling its features users must have an internet connection as long as they use the application. Also WebRTC and OpenFace services must be on to be able to use the application with all of its features.



2.4. Persistent Data Management

What we aimed is to keep track of user information and the user calls constantly in our database. But since, we are already dealing with the video call and the face detection situation on a cloud server to have the best performance as possible, we thought it would be better if we ask the user to sign up with their google+ accounts, or accounts that are already memorized in their computer that they use constantly.

2.5. Access Control and Security

Initially, our system is designed to be Web Application that will be compatible with both phones and computers. All of the users need to make registration or log-in in order to use the system. Since, the personal information taken during these process is not shared or published, the user informations will be safe. If there is any kind of a feedback or problem, users are going to be able to contact the developers via their email.

2.6. Global Software Control

This application is a client-server application. The client side shows that this project is applicable for computers and phones. On the server side there are DigitalOcean Cloud services, OpenFace and WebRTC. There is a constant data flow between client and server sides. Since the data flow starts with the user inputs the flow of the program depends on user inputs. To be able to make the video call as aimed user needs to sign in or register make the call take the output as audio or as visual output printed on the screen.

2.7. Boundary conditions

2.7.1. Initialization

- To be able to use the system properly the users should have a proper computer or phone with a camera/ front camera.
- The users must have wireless connection.
- The users must register to the system.
- The users must log in to the system if s/he has already registered.

2.7.2. Termination

- The user should press either home button or back button while the current window is the main window. If the user quits the application, the process will be terminated.

2.7.3. Failure

- Assuming the electric went off, if the user was in e desktop then the system will be shut down mandatorily but if the system was running in a laptop or in a phone then the user still can use the system, if the internet connection was not gone, till their own battery runs out.
- If the computer or phone crushes for some reason then the problem may stop working efficiently.

3. Subsystem Services

3.1. User Interface Subsystem

User Interface subsystem is designed to provide the interface to the user as being a transition layer to the other subsystems in Application Logic Layer. Its basic operating logic is to display the data which is supposed to be available to the user. The idea of having a user-interface subsystem is to deal with complexities by handling them between other subsystems.

3.2. User Management Subsystem

In this subsystem, the user information is processed and managed. Its login and registration of the system is done at this subsystem. Also, user commands are processed and transferred to the other subsystems at the same layer and other layers.

User management subsystem is designed to manage the user information. The login and registration processes are handled and controlled in this subsystem. And the user commands in this subsystem are transferred to the related subsystems.

3.3. Image Processing Subsystem

Image Processing subsystem is designed to make image processing within the video-call and after having the process completed this subsystem transfers the output to the visual-audio feedback subsystem for further processing on audial feedback.

3.4. Visual-Audial Feedback Subsystem

Visual-Audio Feedback subsystem is a transition subsystem which is designed to take the output coming from the image processing subsystem, then by taking the output and transferring it properly to the audio feedback subsystem.

3.5. Audio Feedback Subsystem

Audio Feedback Subsystem is designed to manage the process of visually impaired user having audial feedback as an output about the emotions, attributes, etc.

3.6. Data Management Subsystem

Data Management Subsystem is the lowest layer of the system which is Data Storage Layer. This subsystem stores the data collected from upper layers and the outer system.

4. References

[1] "CS491 Senior Design Project I", *Ccs.bilkent.edu.tr*, 2017. [Online]. Available: http://www.cs.bilkent.edu.tr/CS491-2/CS491.html.

[2] "WebRTC", *Wikipedia*, 2017. [Online]. Available: http://www.wikizero.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvV2ViUlRD.