CS 491
Senior Design Project

Project Specifications Report

Project Name: UThere

Students

Bilgehan Akcan - 2180290

İlke Kaş - 21803184

Kimya Ghasemlou - 21801412

Melike Demirci - 21702346

Mustafa Yuşa Bababemez - 21703083

Supervisor

Ayşegül Dündar

Jury Members

Erhan Dolak

Tağmaç Topal

Table Of Contents

1. Introduction

Developments in technology have been reshaping the way humans communicate with each other for a long time. Video conferencing is one of the most recent and widely used real-time communication types. There are different uses of video conferencing. It is widely used by companies that have remote workers or multiple offices. Another usage of video conferencing is for educational purposes such as workshops, training, university lectures, etc. It provides many benefits such as reducing travel expenses, increasing productivity and time efficiency, and promoting collaboration for individuals and businesses [1].

While providing these benefits, video conferencing has the drawback of limited personal interactions [2]. Due to this limitation, the attention span of the participants tends to decrease during online meetings. During face-to-face meetings, it is possible for presenters to observe their audience's facial expressions and respond accordingly. By scanning the room, narrators can use different techniques to grab others' attention when needed. However, it is not easy to examine all the faces during a video conference, which is the main problem we want to provide our innovative solution.

Other than tracking the audience's attention, there might be other aspects that presenters may want to track. One thing that presenters may want to know is which part of their presentation took more attention. With the help of eye gaze tracking techniques, presenters may be informed about which part of the screen is more eye-catching. This can be beneficial for them to improve their presentation content over time. For example, if images and graphs are more catchy than the textual content, they may prefer to use them more.

In addition to these, presenters might want to be aware of the audience's emotional state; such as when they reacted positively, negatively, or stayed neutral. By knowing that, they can adjust themselves to the audience's needs immediately. For instance, if the majority of people react negatively, there might be a need for further explanation. If they stayed neutral during the parts that should create excitement, presenters may want to work on their presentation skills.

Online meetings have various challenges and online presenters have several needs. These needs can be met by a video conferencing tool that adopts the latest machine learning techniques to track and analyze the audience.

1.1 Description

UThere is a video conferencing web application with various audience tracking and analysis features. The main aim of this web application is to improve the attention span of the audience by providing instant feedback to the presenters during online meetings. The application will monitor the audience's faces and analyze their attention rates. It will track if the people are looking at the screen and if they do which area they are exactly looking at. UThere will notify the presenters about the attention status of the audience in real time. So that, presenters will be able to arrange themselves as they would be able to do in face-to-face meetings. With the help of the latest techniques in machine learning, the application will be able to provide a more human interaction friendly online meeting experience. UThere aims to break the taboo in people's minds that online meetings are often boring and monotonous.

The application will also turn the audience's real-time videos into an advantage for presenters with eye gaze tracking and emotion tracking features. By analyzing the audience's reactions to the presenters, it will be possible to gain insight on things that would stay unnoticed during face-to-face meetings. Presenters will be notified about how the majority of people react and whether they are staying neutral, positive, or negative towards what is being presented.

Normally, it is less possible to keep track of who is paying attention and who is not and to find out the correlation between audience profiles and their attention spans. UThere will also explore these correlations by conducting data analysis on attention span information and user profiles.

When the distinguishing functionalities of UThere are considered, it can be suggested that the type of innovation that will be implemented is Product Performance [3]. The innovation that we want to bring is incremental, we want to improve existing video conferencing tools with new features. This improvement is crucial due to the transformation of communication platforms and practices.

1.2 Constraints

1.2.1 Implementation Constraints

- Version control will be conducted via Git on Github.
- Trello will be used as the project management tool.
- Project will be implemented by using Python, HTML, CSS, Javascript and React.
- Django will be used as the web framework.
- Agora SDK will be used for handling video, voice and other video conferencing features.

- MVC will be used as a software architecture pattern.
- OOP will be used as a programming paradigm.
- Only attendees having an open camera will be evaluated.
- English and Turkish will be supported.
- Cloud servers will be used for real time processing.
- Google Cloud will be used as a cloud system.
- The system will use third-party libraries like OpenCV, Pytorch and TensorFlow.

1.2.2 Economic Constraints

- github.io domain service is free of charge.
- Open source libraries are free of charge.
- Development and testing tools are free of charge.
- Cloud servers will be free of charge.
- Google provides a 12-month free trial with $300 credit to use with any Google Cloud services.[4]
- Github repositories are free of charge for student accounts.
- Agora provides a 10.000 minutes free trial for each month and further usage will be charged. 10.000 minutes for a month is satisfactory for the development stage, therefore there will be no cost during the development stage.
- 3.99$ will be charged after deployment for 1000 participant minutes, the payment method will be subscription based.[5]

1.2.3 Safety Constraints

- User's information will not be shared with any third party and will only be used to provide feedback.
- Password strength will be checked during registration.

1.2.4 Ethical constraints

- Participants should need to give permission to access cameras.

1.2.5 Environmental Constraints

- Since our project has no connection with the environment, there will be no environmental constraints.

1.2.6 Manufacturability Constraints

- Since our project has no connection with manufacturing, there will be no manufacturability constraint.

1.2.7 Technological constraints

- Demo will only support 4 participants (excluding the presenter) per meeting due to the limitation of processing power at the current stage.
- Device needs to be connected to the internet.
- Device should have an available camera.
- The program will be suitable for different platforms like Chrome, Firefox, etc.

1.3 Professional and Ethical Issues

In UThere, since the process is based on users' facial expressions, it is crucial that their privacy will not be violated for us to comply with ethical rules. In the database system, photos or videos of the users will not be held to respect the privacy of the users. Only data that are related to the emotions of the users will be stored for analyzing and offering analyzed information to the presenter. Collected data will not be shared with any third party companies for business purposes.

The protection of data becomes more significant over time. To process the data of the users, it is important to conform to the purpose limitation principle which means that data of the users can be collected or used for specific necessary purposes. These purposes also have to be transferred to the data owner in a well-defined way. Hence, UThere will collect the necessary data of the users when it is approved by users.

During implementation, when an open source software is used, it is important to take into account licensing it. These packages may have some disadvantages in terms of privacy. Some of the licenses ask to share all of the enhancements [6]. To avoid violation of privacy, licenses that do not require sacrifices will be chosen when it is needed.

2. Requirements

2.1 Functional Requirements

2.1.1 Presenter

- The presenter can register to the video-conferencing app and create a profile.
- The presenter can start a meeting and share that link with others.
- The presenter can share a screen.
- The presenter can enable and disable screen sharing with others.
- The presenter can access the real-time report indicating the attention level and emotional status of the participants.
- The presenter can send an alert to the selected participants.
- The presenter can end a meeting.
- The presenter can create a poll and share it with the participants in order to increase the attention rate.
- The presenter can view all the participants.
- The presenter is notified about who left the meeting.
- Before the presentation, the presenter can set an attention limit which the presenter is warned of when the attention rate of the participants goes down below.
- At the end of the presentation, a report displaying the overall attention rate and emotional status of the participants is formed and shared with the presenter.
- The presenter is notified about the percentage of participants not looking at the screen.

2.1.2 Participant

- The participant can register to the video-conferencing app and create a profile.
- The participant can join a meeting by using the link that is provided by the presenter.
- The participant can allow the system to use the camera and microphone.
- The participant can mute himself/herself.
- The participant can leave the meeting.
- The participant can raise a hand or send some emojis.
- The participant can view all participants.

2.2 Non-functional Requirements

2.2.1 Usability

- The user interface will be simple and provide quick access to essential features of the application. The design of the website will require minimal explanation to understand how to use it.
- Any user who knows how to use any video-conferencing application will be able to use UThere.
- The user can contact us to get help by filling the online contact form.

2.2.2 Privacy

- The collected personal data from the users such as their camera records will be only used for attention and emotion tracking. These personal data will not be processed for any other purposes which are incompatible with the main purposes of UThere.
- UThere will process the personal data of the user on the condition that users consent to such processing. The provided consent form will be specific, informed and explicit.
- The user's face will not be kept in the database.

2.2.3 Scalability

- UThere will be scalable in terms of the ability to handle increased workloads such as adding users more than 5 to a meeting if required hardware and software equipment are supplied.

2.2.4 Performance

- UThere will be a real-time system which processes the faces of the users and presents the current attention level to the presenter.
- The system can process at most 5 users' faces within a minimized time lag between the time when data is produced and the time when data is processed.
- The attention level of the presentation should have high accuracy.

2.2.5 Robustness

- The system will not crash under disturbances such as invalid or erroneous input by dealing with these stressful environmental conditions.

2.2.6 Extensibility

- The design and implementation of the system will not hinder future needs and updates of the system.
- The functionality of the project can be expanded in the future such as sleep detection, and specific emotion recognitions such as aggression, excitement, etc.
- The system will be implemented as object-oriented. Therefore, it will be easy to extend the system for future needs and requirements.

3. Risks

3.1 Techincal Risks

Although all of our group members have had experience in web development and machine learning, none of us has experience in the field of video communication or Agora SDK. Due to a lack of experience, there might be unexpected challenges. We plan to avoid this situation by starting early to project development to have enough time to deal with any unexpected challenges.

3.2 External Risks

After the covid-19 outbreak, video conferencing has been very popular and there are already advanced systems to fulfill the need. Our distinction from other products is the integration of machine learning inside the product with the aim of providing useful feedback to the users. It is possible that our competitors in the market adopt our technology and reduce our market share. We plan to prevent it by using high-quality and high-accuracy machine learning algorithms.

3.3 Social Risks

One might think it is possible that being "watched" can have negative effects on the attention and concentration span of the users. However, according to John Hopkins University in 2018, being watched can trigger "activity in the ventral striatum, an area of the brain that motivates action and motor skills".[7] Moreover, we plan to design our product such that the feedback will not be a distraction either to the presenter or the attendees.

4. References

[1] “10 benefits of use video conferencing” Microsoft.com, 29 Aug 2022. [Online]. Available: https://www.microsoft.com/en-us/microsoft-365/business-insights-ideas/resources/10-reasons-to-use-video-conferencing. [Accessed: 8 Oct 2022].

[2] “12 advantages and disadvantages of video conferencing” Techtarget.com, 04 Aug 2021. [Online]. Available: https://www.techtarget.com/searchcontentmanagement/tip/8-business-benefits-and-challenges-of-video-conferencing. [Accessed: 8 Oct 2022].

[3] Yabaş, S. E. “Innovation Management in Financial Services” cs.bilkent.edu.tr, 7 Oct 2022.[Online]. Available: https://www.cs.bilkent.edu.tr/~cs4912/current/files/intranet/07.10.2022_Innovation_Management_in_Financial_Services.pdf. [Accessed: 15 Oct 2022].

[4] “Free trial and free tier | google cloud,” Google. [Online]. Available: https://cloud.google.com/free. [Accessed: 17 Oct 2022].

[5] “Pricing,” Agora, 10 Oct 2022. [Online]. Available: https://www.agora.io/en/pricing/. [Accessed: 17 Oct 2022].

[6] P. Wayner, “12 ethical dilemmas gnawing at developers today,” InfoWorld, 21-Apr-2014. [Online]. Available: https://www.infoworld.com/article/2607452/12-ethical-dilemmas-gnawing-at-developers-today.html. [Accessed: 17-Oct-2022].

[7] Johns Hopkins University."When there's an audience, people's performance improves." ScienceDaily. [Online]. Available: www.sciencedaily.com/releases/2018/04/180420122913.htm .[Accessed: 17-Oct-2022].