Presence Detection for Video Conferencing Applications

Samuel John Mario Canonaco
Ontario Tech University
Oshawa, Canada
samuel.canonaco@ontariotechu.net

Harrison Bradley Smith
Ontario Tech University
Oshawa, Canada
harrison.smith1@ontariotechu.net

***Abstract*—With the Covid-19 pandemic the shift towards the usage of video conferencing software as a means of communication for different sectors has become more prominent in the last few years. However, this video conference software can sometimes be lacking in its support for those with disabilities who would be using them. Our project is centered around the creation of a device for users who might have vision related disabilities to allow for a much smoother and accessible experience. This is important because often times these programs can have limitations that make it difficult for these people to use them. In this document we seek to utilize leverage bodystorming in order to develop potential use cases as well as assess common problems that can often occur in normal day to day use of video conferencing applications. Then we create a prototype in order to present a potential solution to these issues and gain additional feedback from our actor in order to redetermine if any other additional use cases need to be considered as well as update and present our new findings.**

***Index Terms*—remote communication, telecommunications, control interaction**

I. Introduction

In this document we will outline our process we used in order to incorporate the user centered design process into account when designing, prototyping, and testing potential solutions to our problem. we leveraged the usage of body storming in order to determine potential use cases for idea in what our users would need to be able to accomplish using our device, in addition we also created a user persona in order to create and an example of who we hope to create the solution for and what they will look like. Alternatively, we utilized an alternative form of body storming in order to come to our conclusions where instead of utilizing role playing in order to simulate the interactions between different users with our potential solution we utilized VR simulations in order to create interactive prototypes in order to evaluate our use cases and tasks.

II. Methods

In conducting our bodystorming process we began by observing how certain tasks (raising hands, muting/unmuting, share their screen, turning on and off camera, and take questions.) are normally performed on video conferencing software we then generated use cases based off of what was observed during the evaluation See Appendix B and combined it with a persona See Appendix A in order to determine the types of people who would be interacting with our use cases. Once we acquired our use cases we identified potential problems with the traditional workflow when our persona would interact with the use cases See Appendix A. Once we were finished with our preliminary data gathering we then moved onto testing where we got an actor who exhibited qualities similar to our persona to interact with the video conferencing software. Specifically, we had a member of our team remove their glasses to simulate what these pieces of software would convey visually for someone with limited vision. In this scenario we had our actor go back and forth with the observer to accomplish the outlined scenarios both in the traditional method, and simulating our proposed solution See Appendix C. After the actor was done we had the actor and observer change places and redo the previous steps to collect additional findings. Finally, we created a VR simulation See Appendix D which represented how our final device should theoretically work and recorded notes on what elements from our solution functioned well as well as any potential changes we would need to make for our use cases See Appendix A.

III. Results and discussion

Through our research we identified many of the potential pain points likely experienced by people with limited vision when using video conferencing software. Having voice and video being transmitted makes it difficult for screen readers to be used, as they can be slow and they can be disruptive for other call participants. Additionally we discovered that much of the iconography used for audio and video indicators can be difficult to identify if they don't make heavy use of colour, which could also be a potential problem. Our actor identified that the only reason they were able to use the applications with limited vision is due to the fact that they had memorized the layout after using the apps for multiple years. These different applications are also not consistent in what they use as a mute or disable video hotkey, which means that additional third party software is needed to maintain consistency and ease of use, especially since hotkeys can be dependent on window focus. In conclusion, the current video conferencing applications are very close to unusable without additional tools for people with limited vision.

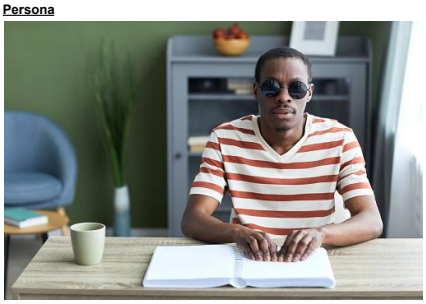
Appendix A
Bodystorming Notes

Bodystorming Information

Tasks

- Interact with a video call application
- Traditionally done by clicking a on screen button or entering a key shortcut

Persona



- Any user of any age who may experience impaired vision or blindness and be required to make frequent use of conferencing software in their day to day lives either for interaction or their profession.

Situation

Somebody who is visually impaired trying to interact with conferencing software (Google meet, Discord, Zoom, Skype, Teams) and needs to be able to interface with the software, via controls i.e raising hands, muting/unmuting, share their screen, turning on and off camera, and take questions.

Scenarios

1. Mute audio and disable video to talk to someone in your physical room
2. Giving a group presentation where you need to mute repeatedly
   (Scenarios are applied to both observer and actor in order to highlight differences)

Actions are traditionally done by clicking a on screen button or entering a key shortcut

Samuel (limited vision actor):

Potential Issues

Scenario 1:

1. User may not be able to easily identify buttons on screen for muting or unmuting
2. User may have multiple windows open causing the streaming software to not register keybinds
3. User may not be familiar enough with specific conferencing software layouts
4. Using screen reader software may result in other call participants hearing narration
5. Speed of muting is reduced

Scenario 2:

1. Can be difficult to detect when different slides change
2. Screen reader would be slow and could result in other participants hearing it
3. Could simply forget that you're muted due to subtle indicator

Harry (Observer):

Potential Issues

Scenario 1:

1. May have window covering conferencing software causing software not to register inputs or cause you to forget to press the mute button or disable the camera

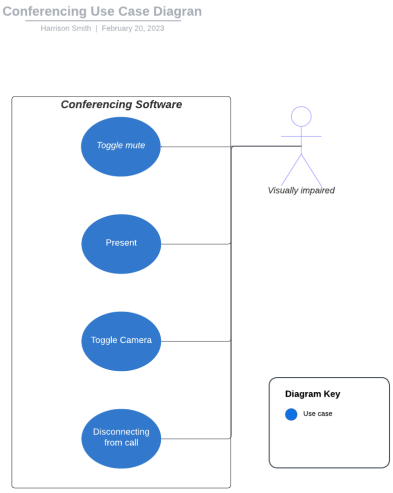
Scenario 2:

1. Displaying incorrect window / tab.
2. Delayed transmission

Post-VR Prototype

The prototyped interaction works properly in VR. It detects when the user is seated and when they stand up. Accordingly, the program shows a graphic of a muted or unmuted video call. This interaction feels useful and natural and demonstrates that we can progress with utilizing the same premise for other options. However, during our bodystorming we were able to identify that our initial use cases for our tool were still valid and instead we needed to focus on how exactly the user would reach the use cases which is the reason the use case diagram is not updated.

Appendix B
Use Case Diagram



Appendix C
Bodystorming breakdown video

Link To YouTube

Appendix D
Wizard of Oz Prototype

Link To YouTube