Smart Talk Video Conference Assistance Device

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*—After the events of the covid-19 pandemic many day-to-day activities of people were transitioned into online platforms. Because of this transition the adoption of video conferencing software for the purposes of communication was done by a multitude of different fields. Because of this new reliance on conferencing software, it is vital that they provide adequate support for those who may be visually impaired when it comes to interfacing with their software. Unfortunately, most conferencing software while featuring different kinds of GUI's suffer from different issues that can cause problems. If the user is unfamiliar with the layout for instance it can cause confusion if they have used other kinds of conferencing software with a different layout in addition, the often lack of feedback can make it difficult for users to visualize their actions. many different solutions have been attempted to resolve the problems this software, but many are still lacking overall when it comes to considering the needs of both the end users as well as the other members who will be present while it is being utilized. This paper describes the process taken by the team thorough out the design process of the creation of a device for visually impaired users who may need to use conferencing software in their daily lives.**

I. Introduction

For our project we proposed the creation of a device that will allow users with disabilities particularly those with visual impairment a more accessible method of controlling their microphone as well as other elements of video conferencing software. Increasing the accessibility of video conferencing software is important as more and more industries adopt a work from home style where users are required to attend meetings online. Because of this, the ability of conference software to meet the needs of those with these impairments is important so that they are able to adequately interact with the software. If the needs of these users are not met then it can cause them to have an overall negative interaction and may even cause issues not just to the person with the disability but also those participating in the meeting. The Smart Talk is a device created for use with video conferencing software with the goal of improving the experience for users who may be visually impaired. The Smart Talk leverages the use of passive sensors in addition to physical controls in order to offer the user a global control system that will allow them to interact with any video conferencing software while also providing different forms of feedback, in order to provide a more standardized control layout and to help alleviate some of the common problems that can occur.

II. Methods

During our prototyping process we conducted multiple different forms of preliminary research to locate and identify common problems our device would need to solve. We also tested them in order to determine if our device met our requirements when it came to resolving issues from our research.

A. Research

For our initial steps we began by research the topics surrounding accessibility in video conferencing software in order to determine if our problem statement held any validity and then we began to search for papers that focused primarily on studying the accessibility of different conferencing software [1] [2] [3]. From these sources we did gather that there was indeed a gap in support for visually impaired users in conferencing software. After we began to look for journals that discussed common accessibility problems that were common with conferencing software and that also provided guidelines in how future developments could be implemented in order to provide options for visually impaired users [4] [5]. From these studies we utilized their information throughout the early development to make we prevented these problems as well as look to the journals to see if they provide and information on how we could resolve any issues should we encounter any.

B. Bodystorming

After our initial prototypes we then advanced on to the bodystorming process by which we then outlined the normal task a user would perform when interacting with the conferencing software and observing how they are normally done and created a use case diagram See Appendix B. After this we acquired a participant that met the characteristics similar to our ideal end user See Appendix A and had them attempt to perform our use case tasks while describing any issues they encountered See Appendix A. Once we were finished we presented the user with our prototype in order to see if it resolved any of the issue the users had See Appendix A. Once all these steps were completed we had a member of the project team repeat the same steps while still noting down the information See Appendix A. Once all results were captured we recorded which elements from our prototype functioned well and which ones would need to be improved See Appendix A.

C. User Feedback

Lastly, once we completed the previous steps in our design process we then moved on to evaluating our device to determine if their were still areas where we could improve our device or if there were any potential problem areas. To conduct these tests we acquired users to test our device and then presented them with a questionnaire. The information from the questionnaire would then be input into 3 different data visualization tools this being the System Usability Scale See Appendix E, Presence Questionnaire See Appendix C, and lastly the NASA Task Load Index D. We then analyzed all of these tools in order to evaluate and improve our current device.

III. Results

A. Research

From our preliminary research we were able to acquire multiple pieces of useful information that influenced the early design process while also providing us with guidelines to follow as we continued to prototype and develop our product. For example, early research indicated to us that multiple different kinds of conferencing software featured a lack of keyboard accessibility [2]. This indicated that the addition of a universal button layout would be useful for users in order to create a control scheme that was the same across all software and could be catered for users needs. On top of this, some sources noted that auditory feedback could be beneficial for visually impaired users on conferencing software [1] [4]. This information further reinforced the idea for buttons or switches which provide feedback and that pursuing a passive sensor device had solid support in improving this problem.

B. Bodystorming

During our bodystorming were able to test potential solutions to the problem as well as come to our own conclusions as to what we needed to design around when it came to generating new versions of our solution. Firstly, we discovered that the usage of screen readers causes issues during conferencing since it can be picked up by microphones causing other members to hear the screen reader which can cause distractions. Additionally we learned from our conducted scenarios with our participants that a big problem was the lack of consistency in controlling actions in the call with layouts being different for each piece of software which could cause issues if the user was unfamiliar with the layout. Because of this we opted to keep the design of an outside control box that could interact with the software in order to allow users to make a consistent control layout that could be customized by them.

C. User Feedback

Lastly, we have the data we gathered from the user feedback studies, firstly, from our NASA TLX See Appendix D we were shown that our device functioned properly in all 4 categories we did however note that the physical rating of the device was slightly higher. This information indicated to us that we need to improve our design slightly so that is was much easy to interface with physically we deemed these scores to be due to the stiffness of the buttons suggesting they need to be changed or loosened. From our SUS scale calculations we found that we scored 80.625 See Appendix E which indicated that our device was functioning well with only small and minimal problems. Our individual score analysis demonstrated to us that the device was easy to use but again that it might be too cumbersome See Appendix E again we estimated this was again because of the rigidity of the buttons. As for our presence questions the data demonstrated to us that the device was performing well in all categories See Appendix C even in the quality of the interface however again because of the buttons it was only slightly lower.

IV. Discussion and Conclusions

For our final device we added a buzzer which provides auditory feedback, which will notify user that their actions were recognized. This helps differentiate the different buttons. Overall, the device does a good job of controlling a user's microphone and video. One problem is that we have no way of actually knowing if the operating system is muted. Our device instructs users to unmute before they plug in the device so that it can keep track of the current state. The force sensor does not do a great job of distributing weight, and the device would do better at detecting users if we had a larger sensor sandwiched between pads. The device would also be easier to use if it were wireless, as users would not need to worry about getting stuck on the wires connecting the box to their PC. Some of the buttons could also be better as toggle switches. A different method of feedback could be useful, such as haptic feedback, as the auditory feedback may not be particularly easy to understand for some users. We were mostly constrained by funds and availability of hardware. Without cost free access to different options it's simply not feasible for us to do lots of prototyping with different components. All in all though, we believe that we've created a device that adequately addresses the problems that users with limited sight need to overcome to use video conferencing software.

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

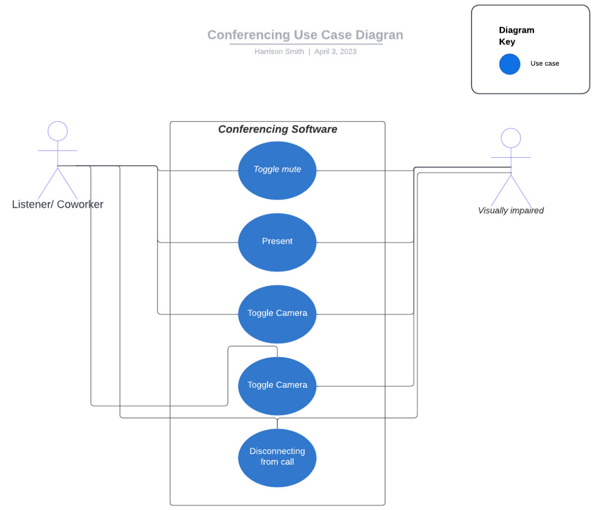
Scenario 2:

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

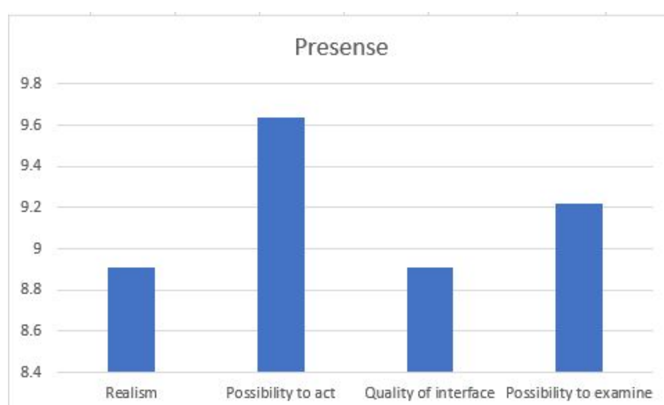
Post-VR Prototype

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.

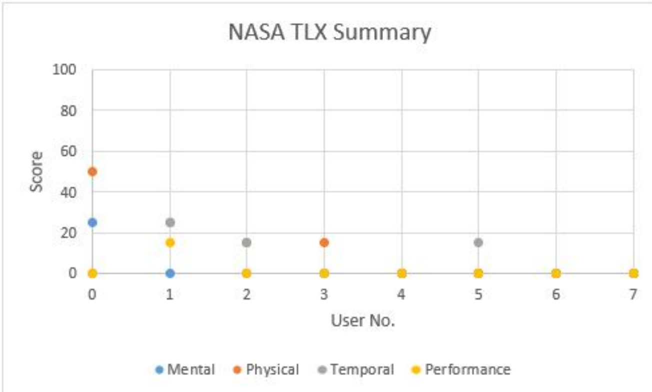
Appendix B
Use case diagram



Appendix C
Presence



Appendix D
NASA TLX



Appendix E
System Usability Scale

Paticipant	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	
	I think tha	I found the	I thought t	I think tha	I found the	I thought t	I would im	I found the	I felt very (	I needed t	Score
1	2	1	5	2	3	5	5	1	5	1	75
2	3	1	5	1	4	5	5	1	5	1	82.5
3	3	1	5	1	4	5	5	1	5	1	82.5
4	4	1	5	1	4	5	5	1	5	1	85
5	4	1	5	1	4	5	5	1	4	1	82.5
6	5	1	5	1	5	5	5	1	4	1	87.5
7	5	1	5	1	5	5	5	2	4	1	85
8	5	2	4	1	5	5	5	3	1	3	65
									Average Score:		80.625

References

[1] N. Anderson, "Accessibility challenges of video conferencing technology," in *Design, Operation and Evaluation of Mobile Communications*, vol. 12796 of *Lecture Notes in Computer Science*, pp. 185–194, Cham: Springer International Publishing, 2021.

[2] I. A. Doush, A. Al-Jarrah, N. Alajarmeh, and M. Alnfiai, "Learning features and accessibility limitations of video conferencing applications: are people with visual impairment left behind," *Universal access in the information society*, pp. 1–16, 2022.

[3] M. Hersh, B. Leporini, and M. Buzzi, "Accessibility evaluation of video conferencing tools to support disabled people in distance teaching, meetings and other activities," in *ICCHP*, p. 133, 2020.

[4] L. Moreno, P. Martínez, and R. Alarcon, "Requirements and design patterns for an accessible video conferencing tool," in *Proceedings of the XXII International Conference on Human Computer Interaction*, pp. 1–9, 2022.

[5] B. Maus, "Designing video call spaces with and for adults with learning disabilities: A remote participatory design approach," 2021.