Live Audiovisual Remote Assistance System (LARAS) for Person with Visual Impairments

Zachary Frey
Department of Physics
University of South Florida
Tampa, Florida, USA
zfrey@usf.edu

Nghia Vo
Department of Computer Science
University of South Florida
Tampa Florida, USA
nghiavo@usf.edu

Varaha Maithreya
Department of Mechanical Engineering
University of South Florida
Tampa,Florida, USA
doddi@usf.edu

Tais Mota
Department of Mechanical Engineering
University of South Florida
Tampa, Florida, USA
mota5@usf.edu

Urvish Trivedi
Department of Mechanical Engineering
University of South Florida
Tampa, Florida, USA
udt@usf.edu

Redwan Alqasemi
Department of Mechanical Engineering
University of South Florida
Tampa, Florida, USA
alqasemi@usf.edu

Rajiv Dubey
Department of Mechanical Engineering
University of South Florida
Tampa,Florida, USA
dubey@usf.edu

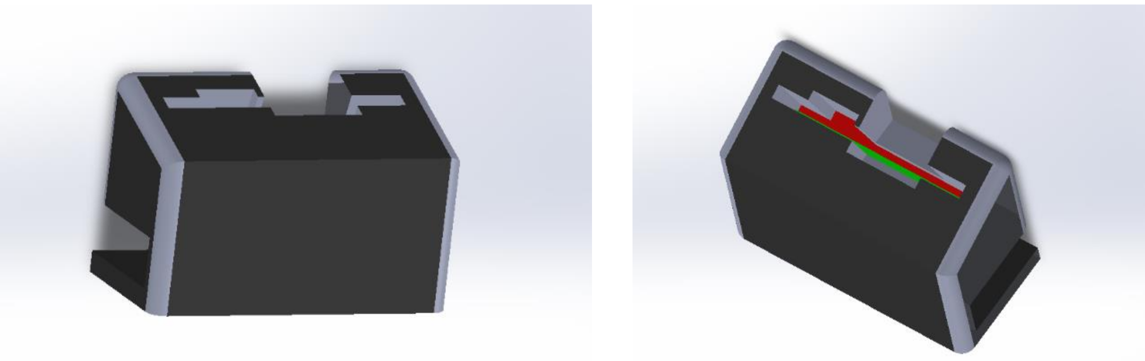
***Abstract*— According to "The World Report on Vision" by World Health Organization (WHO) [1], there are more than 2.2 billion people who have near or distant vision Impairments, out of which 36 million people are classified as entirely blind. This report also emphasizes the importance of social and communal support in enabling individuals with vision impairments to integrate into society and reach their full potential. While performing daily activities and navigating the environment, people with visual impairments (PVIs) often require direct or synchronous assistance [2]. Consequently, there is a growing need for automated solutions to assist in this regard. However, existing automated solutions can provide limited assistance with limited reliability [3]. Although assistants and caregivers can help PVIs when they have the time and they are able to be physically present onsite, other assistants may have the time to help, but no means to be physically present onsite to help (such as a friend with disabilities or an elderly family member who may not be physically able to help onsite). In this context, remote help can be utilized to provide PVIs with the help they need from persons who want to help but can't be physically there in person. In this paper, our team designed a wearable hat equipped with a camera system to facilitate remote audiovisual communication between a visually impaired user and a remote assistant. The hat features three 120-degree cameras positioned around its rim and houses electrical components discreetly within its structure. A mobile application installed on both the user's and the assistant's smartphones provides two-way audio communication. The user with visual impairments can request assistance from a pool of registered assistants who can control the camera feed on the user's hat and provide live guidance.**

This hat-based camera system enables real-time assistance for the visually impaired to perform daily activities such as job-related tasks, shopping tasks, meal preparation tasks, street navigation tasks, and other tasks. A hat-based design and remote assistance allow for a flexible and comprehensive solution to meet the needs of individuals with visual impairments and provides assistants the capability of remotely helping PVIs.

Keywords—Assistive Technologies, Android application, Visually Impaired

I. Introduction

Independence is a highly valued aspect of human life and is acknowledged globally as essential in fostering a good quality of life. However, individuals with disabilities often face limitations that impede their independence, including physical, mental, and behavioral constraints and negative societal attitudes and prejudices [4]. The field of disability studies is continuously advancing, and technology has played a significant role in providing viable solutions. Technological advancements have significantly impacted various aspects of



(a) Front iso metric view

(b) Top iso metric view

Fig. 1. Camera Mount

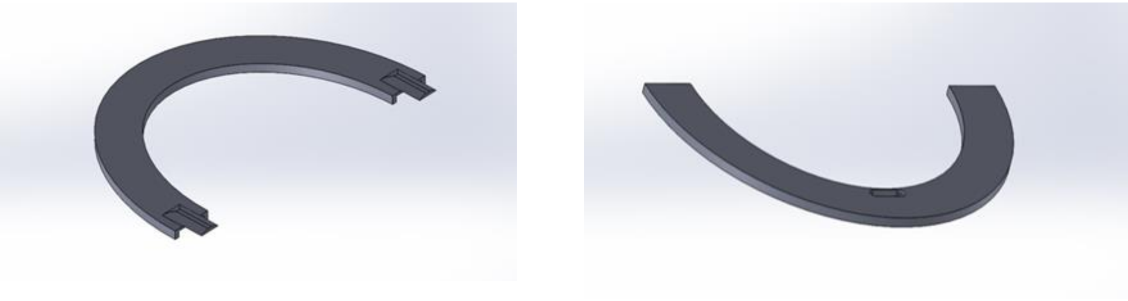


Fig. 2. Hat Rim

life for individuals with disabilities, including rehabilitation robotics and technologies [5], prosthetics [6], assistive software [7, 8], transportation [9], and ambulatory aids [10, 11], to name a few.

Despite technological advances in disability studies, accessibility to essential daily activities remains a significant challenge for people with visual impairments. Researchers have developed technologies based on computer vision and artificial intelligence for individuals with visual impairments. TapTapSee [12], CamFind [13] , and Talking Googles [14] are successful mobile applications that recognize environments and provide virtual assistants.

Despite their best efforts, these assistive devices need to improve their ability to assist solely by employing video feeds. In the case of not enough light, or objects with a similar texture or color, will result in delayed response, making it impossible to use them effectively on a daily basis. A person with visual impairments may also be unable to achieve as much independence as they would like when dependent on external assistive devices. This makes it even more critical to conduct practical interventions, such as Human In the Loop systems [15], to overcome these limitations and help individuals become more independent.

In light of the limitations identified in assistive devices, this work offers a novel method of addressing accessibility challenges faced by people with visual impairments. This research aims to develop a wearable hat with a camera system that allows the user to communicate with an assistant remotely through audiovisual communication. A mobile application facilitates easy and intuitive communication, with just two buttons that both parties can easily access, enabling easy and intuitive use.

Among the key features of this hat are that it does not require the user to occupy their hands during use. The assistant can view the user's live camera feed remotely and guide them through the desired tasks, making the process easier and increasing the user's independence. Three 120-degree cameras are integrated along the rim of this hat, and the electrical components are carefully concealed within the hat's structure. This allows for seamless and unobtrusive operation, making daily tasks and navigation easier for users.

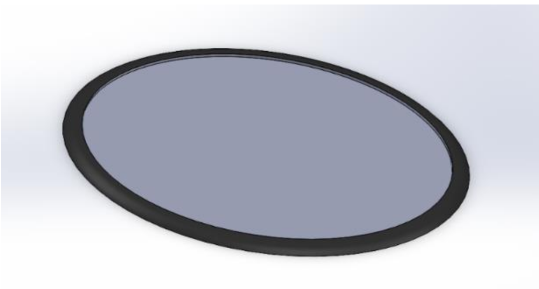


Fig. 3. Hat Insert

II. System design and implementation

A. Hardware Implementation

A series of 3D-printed modules were utilized to house the system components, including the camera mounts, hat rims, and inserts. This hat mount was designed to securely fasten three 120-degree wide cameras at a 120-degree angle to the hat's rim to create a human-like perspective. The camera mounts were designed to fit snugly in their slots, eliminating the need for adhesive material and allowing them to be easily removed for repair or replacement. In addition to providing stability to the camera mounts, the hat rim conceals the camera's wiring. Furthermore, the camera insert serves as a secure enclosure for the electrical components. A 3D-printed SolidWorks model of each of these components are shown in Fig. 1-3. The camera mount depicted in Fig. 1 is carefully designed to allow it to slide and fit onto the cutout, which can be easily observed in green color notation. As depicted in red color, the cutout is shallower to ensure increased tightness of the assembly. Its wide horizontal cut allows it to be securely mounted on the 3D-printed rim, which is proportionate to the camera mount. The additional cuts in Fig 1. are designed to accommodate peripheral components of the camera, such as its cable output.

This system features a 120-degree Arducam camera [16] that captures high-quality images and videos with wide-angle coverage. A raspberry pi three B [17], a PKCELL 6600mAh

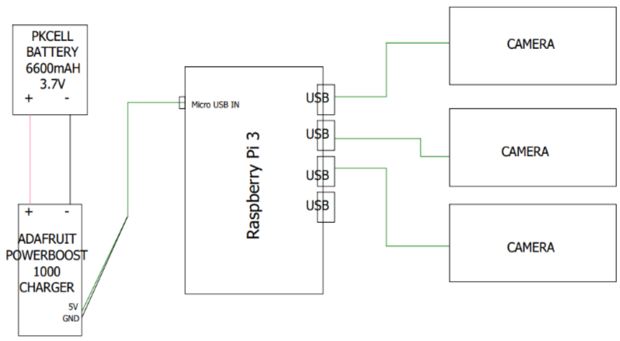


Fig. 4. Circuit Design

3.7V battery, and an Adafruit Power Boost 1000 charger [18] (voltage converter) are all used in this system to support the camera and its functionality. The hat insert depicted in figure 3 can accommodate the entire circuitry and fits securely inside the hat cavity, providing a stable base for placement while ensuring the user's comfort is not compromised. The electrical system's schematic is illustrated in Fig. 4.

This system uses a slanted hat on the edges to integrate an embedded 3D-printed rim without being noticeable. The hat has been chosen to not obstruct the user's view or cause discomfort. A visual representation of how the camera and the hat have been integrated into a functional device is shown in Fig. 5, which illustrates how the hat and the camera have been assembled.

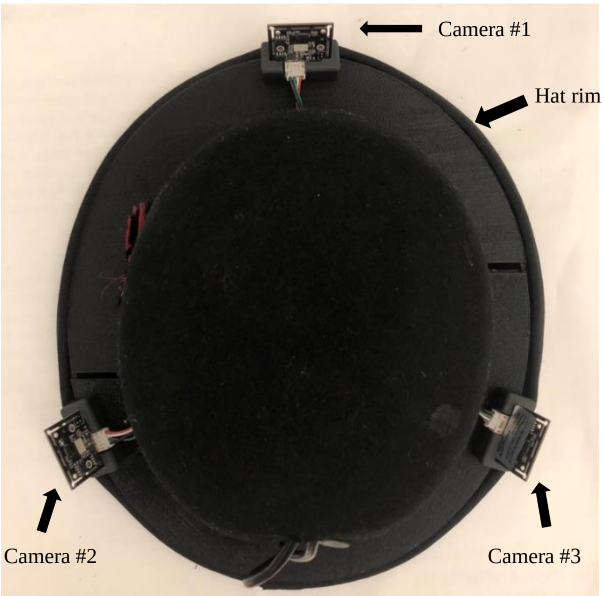


Fig. 5. Live Audiovisual Remote Assistance System (LARAS)

Due to the limitations of the 3D-Printer, the design of the 3D-printed rim had to be divided into two sections, as shown in Fig 2. In order for the rim to remain intact, the interlocking parts of the two sections were secured with Velcro. The wires within the hat area were accommodated within the rim by incorporating a small cut, shown in Fig. 2.

B. *Software Implementation*

1) Software Overview: The development of this application incorporated the use of open-source software libraries, including the Python programming language [19], Flutter [20], Flask [21], OpenCV [22], and Firebase [23]. The implementation consists of two main elements: a mobile application and a video-streaming web server. This design aims to enable assistants to assist individuals with visual impairments through remote access to the hat-based cameras, which are integrated into the design via the user-friendly mobile application interface, allowing them to assist remotely. With the combination of these software components, individuals with visual impairments are expected to be more independent and be able to interact more effectively with the world around them. In Fig. 6, we can see an overview of the application architecture.

Users can initiate assistance requests using the mobile application based on its user-centered architecture. A one-to-many relationship allows multiple assistants to receive requests from a user's contact list. This enables multiple assistants to receive the request. In the event that multiple assistants receive the request, only the first one to accept will be able to continue the session, with the other instances being terminated. The selected assistant will be granted access to the video streaming server in response to the user's acceptance.

2) Video Streaming Raspberry Pi Server: A raspberry pi video streaming server is implemented to provide the mobile application with the necessary video feed. Through the OpenCV library, frames and images are retrieved from the three cameras connected through USB ports to the raspberry pi server. Python and Flask are used to implement the video feed as web servers, with a unique endpoint for each camera feed. Embedding these endpoints into HTML files allows the feeds to be displayed on a web browser. As a final step, the

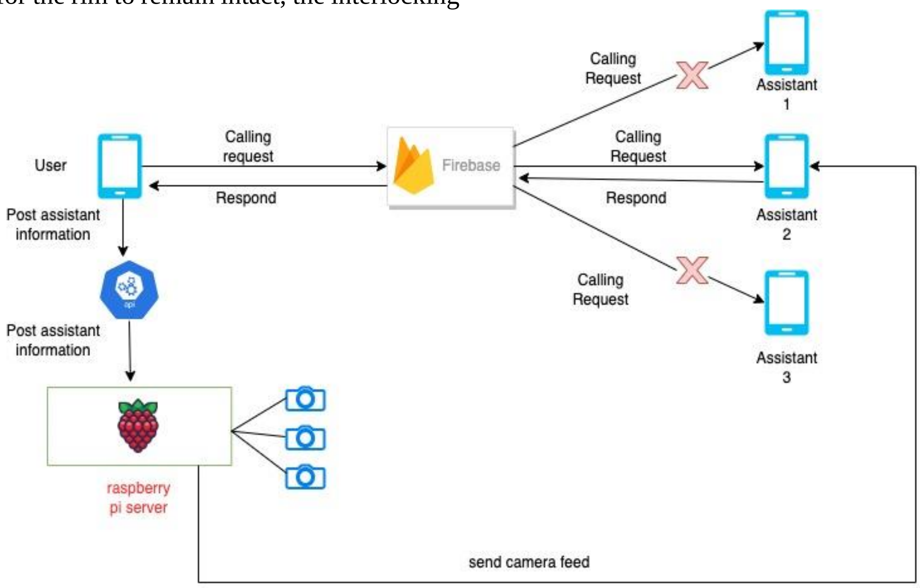
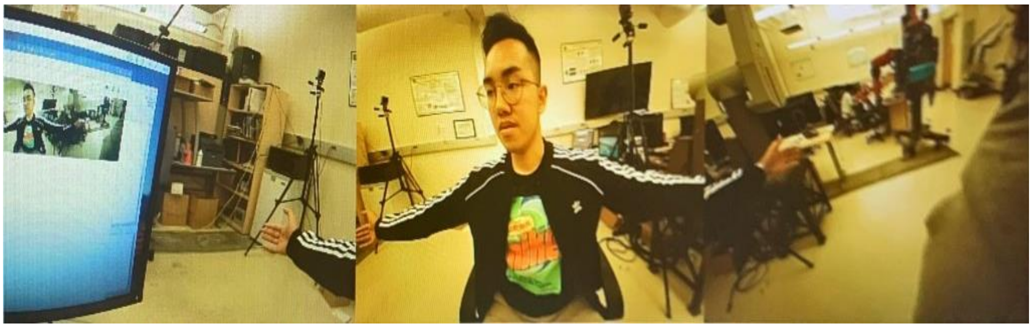


Fig. 6. Application Architecture

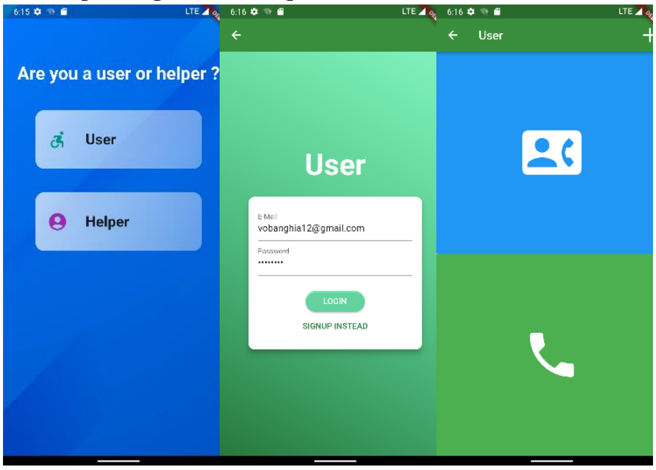


(a) Camera #1 feed (b) Camera #2 feed (c) Camera #3 feed

Fig. 7. The video feed consists of three cameras from the video raspberry streaming server.

IP address of the video streaming server is integrated into the mobile application's codebase to display the live stream of the video feed. The Raspberry streaming server shows the video feed from the three cameras, as shown in Fig. 7.

3) Android mobile application: The Dart programming language and the Flutter framework were used to develop the mobile application. The application's architecture has two main components: the front end and the back end. Front-end components handle user interfaces and facilitate interaction with applications. On the other hand, back-end components process data and communicate with other systems or devices. A division of responsibilities allows for an application implementation that is more efficient and streamlined while also improving the user experience.



(a) Application Startup Page (b) Login Page (c) User Session Interface

Fig. 8. User Interface for user and helper

a) Front-end: A front-end component of the application was developed using Dart and Flutter programming languages. The application has two distinct log-in sessions for users (individuals with visual impairments) and helpers. Both sessions have similar sign-in and sign-up pages, with fields for email, password, and name. After successfully logging in, the user interface differs between the two sessions. User Interface Designs are shown in Fig. 8.

As part of the helper session, users can be added, removed, or updated based on their availability. On the other hand, the user session features two large buttons, one for calling and one for the user's contact list. When a call is executed, all registered helpers for that user will be notified to assist. If a helper accepts the request, the corresponding camera feed will be returned to the user, while other calling sessions will be terminated simultaneously.

The contact button allows users to view the list of helpers and perform operations such as adding, removing, and updating them. Users are only able to add helpers who are available in the database to their contacts. Live video streaming can be stopped by clicking the red button when accessing the video feed.

b) Back-end: A cloud-based platform called Firebase provides a variety of functionalities, including real-time database management, authentication, hosting services, and real-time database management [23]. In the application, Firebase serves as the backbone, allowing CRUD methods to be implemented on the user and helper tables of the database, such as Create, Read, Update, and Delete.

In the database schema, two tables have been created: the users table and the helpers table, which are interconnected. Helpers and users share a foreign key, with the user field containing a helperID field and the helper field containing a userID field. The implementation of the calling feature needs to have this relationship, as it enables the user to send a request signal to all their contacts, who can accept or decline the call.

In response to a call, the helper sends a signal back to the database, which the user receives as authorization for the helper to access the video streaming server. Upon signing up to the app, the userID and helperID fields are generated by a random ID generation process within the database.

III. CONCLUSION & FUTURE WORK

This paper provides details about design and development of an audiovisual remote assistance system, consisting of an Android mobile application wirelessly integrated with three 120-degree cameras that provide a clear and comprehensive view of surroundings that assists people with vision impairments in performing activities of daily living with the help of a human assistant. The device enables remote audiovisual communication between individuals with visual impairments and their assistants by integrating a camera system into a hat, providing both the user and the assistant with the necessary guidance and assistance. As a result, individuals with visual impairments can increase their independence by enhancing their ability to perform daily tasks and navigate their surroundings. One of the unique features of this device is that it allows for hands-free assistance, which means that it does not require the user to have their hands occupied at any given time. This makes it a very convenient and accessible device.

In future work, a Live Audiovisual Remote Assistance System (LARAS) will be evaluated based on the communication protocol, user-friendliness, and other qualitative and quantitative metrics. This project has opened new avenues of research for future assistive technology research, especially concerning individuals with visual impairments. Considering that technology is continuing to advance, it is possible for the project's scope to be expanded to include features such as voice-activated navigation, augmented reality, and real-time object recognition in the future. Wearable technology, such as smart glasses or head-mounted displays, can be incorporated into the device to make it more engaging and interactive for helpers. Through this device, individuals with visual impairments can improve their quality of life and independence while laying a foundation for further assistive technology development and refinement.

ACKNOWLEDGMENT

The authors would like to thank the Florida Department of Education - Division of Vocational Rehabilitation for their support.

REFERENCES

[1] WHO. World Report on Vision. 2019; Available from: https://www.who.int/publications/i/item/9789241516570.

[2] Kommu, S.S., Rehabilitation Robotics. 2007: BoD–Books on Demand.

[3] Kelly, S.M.J.J.o.V.I. and Blindness, Use of assistive technology by students with visual impairments: Findings from a national survey. 2009. 103(8): p. 470-480.

[4] Aldridge, J.J.L.D.P., Promoting the independence of people with intellectual disabilities. 2010. 13(9).

[5] Trivedi, U., et al. A wearable device for assisting persons with vision impairment. in ASME International Mechanical Engineering Congress and Exposition. 2017. American Society of Mechanical Engineers.

[6] Toledo, C., et al., Upper limb prostheses for amputations above elbow: A review. 2009: p. 104-108.

[7] Menychtas, D., S.L. Carey, and R. Dubey. Simulation algorithm for the upper limb for better training and prosthesis prescription for amputees. in Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments. 2015.

[8] Trivedi, U., R. Alqasemi, and R. Dubey. Robot Learning From Human Demonstration of Activities of Daily Living (ADL) Tasks. in ASME International Mechanical Engineering Congress and Exposition. 2021. American Society of Mechanical Engineers.

[9] Habha, L., et al. Autonomous Wheelchair Indoor-Outdoor Navigation System through Accessible Routes. in The 14th PErvasive Technologies Related to Assistive Environments Conference. 2021.

[10] Ramakrishnan, T., M. Schlafly, and K.B. Reed. Evaluation of 3D printed anatomically scalable transfemoral prosthetic knee. in 2017 International Conference on Rehabilitation Robotics (ICORR). 2017. IEEE.

[11] Albraheem, L., et al. Third eye: An eye for the blind to identify objects using human-powered technology. in 2015 International Conference on Cloud Computing (ICCC). 2015. IEEE.

[12] Rajendran, P.S., P. Krishnan, and D.J. Aravindhar. Design and Implementation of Voice Assisted Smart Glasses for Visually Impaired People Using Google Vision API. in 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA). 2020. IEEE.

[13] Sahasrabudhe, S., R. Singh, and D.J.J.T.P.D. Heath, Innovative affordances for blind smartphone users: a qualitative study. 2016. 4: p. 145-55.

[14] Holton, B.J.A.A.M., A review of the taptapsee, camfind, and talking goggles object identification apps for the iphone. 2013. 14.

[15] Bringes, C., et al., Determining the benefit of human input in human-in-the-loop robotic systems. 2013: IEEE.

[16] Arducam. Available from: http://www.arducam.com/arducam-mini-released/.

[17] Upton, E. and G. Halfacree, Raspberry Pi user guide. 2014: John Wiley & Sons.

[18] PowerBoost 1000 Charger - Rechargeable 5V Lipo USB Boost @ 1A - 1000C. Available from: https://www.adafruit.com/product/2465.

[19] Van Rossum, G. and J.J.C.q. De Boer, Interactively testing remote servers using the Python programming language. 1991. 4(4): p. 283-303.

[20] Windmill, E., Flutter in action. 2020: Simon and Schuster.

[21] Ronacher, A.J.W.D., One Drop at a Time, Flask. 2010. 2017.

[22] Bradski, G. and A.J.D.D.s.j.o.s.t. Kaehler, OpenCV. 2000. 3(2).

[23] Tamplin, J. and A. Lee, Firebase expands to become a unified app platform. 1 September. 2011, Volume.