Design and Evaluation of a Remote Assistance Application for Persons With Visual Impairments

Varsha Nagaraj, Naveen Mukundan, Vinod Namboodiri
Department of Electrical Engineering and Computer Science
Wichita State University, Wichita, KS, USA
Email: vxnagaraj@shockers.wichita.edu, nxmukundan@shockers.wichita.edu,
vinod.namboodiri@wichita.edu

Seyed Ali Cheraghi
Smith Kettlewell Eye Research Institute
San Francisco, CA, USA
Email: ali.cheraghi@ski.org

***Abstract*—Blind or Visually Impaired (BVI) individuals often face many challenges while performing daily tasks or exploring new places. Getting assistance from strangers is not always desirable in such situations. With the advancement of technology, BVI individuals can utilize assistive technologies like screen readers or accessibility features on smartphones to access apps to solve some of these daily challenges, but there remain many tasks that still require some sort of human assistance. Some current approaches to provide remote assistance through video calls are either too expensive or do not use helpers whom a BVI individual can fully trust. This work develops a cross-platform mobile application called GuideCall that enables BVI individuals to draw assistance through a video call with a single volunteer helper selected from one of many pre-constructed situation-appropriate groups of trusted individuals. GuideCall provides many features not present in commodity video-calling applications and is specifically built to meet the needs of BVI individuals. Preliminary evaluations show GuideCall to be fairly effective in many daily tasks BVI individuals encounter, potentially proving to be an inexpensive option for receiving assistance while being more confident about the quality of assistance and safety.**

I. Introduction

Visual perception plays a central role in completing many tasks of human daily routine, such as indoor and outdoor wayfinding, locating items of interest at a store or office, comprehending visual signs and printed text, and getting a general sense of the current state of the surrounding. These tasks can pose great challenges for blind or visually impaired (BVI) individuals leading to the need to spend significant amount of additional time and effort (compared to sighted users) to complete these tasks (if they can be completed at all), potentially taking on undue physical risks in some cases.

There has, thus, been a lot of research to overcome these challenges. For example, there has been a lot of work in the area of indoor and outdoor wayfinding through the use of global-positioning systems (GPS), computer vision and artificial intelligence (AI), and wireless technologies to provide location and associated contextual information for BVI users (for example, [9], [14], [17]). Even with these emerging advancements, there will continue to be many instances when the limitations of these solutions (such as lack of infrastructure or conditions unsuitable for the technology to work) will result in a BVI user not fully being confident in relying on them. In such cases, it always helps to be able to rely on another human's assistance to bridge the gap and provide the necessary assistance. Unfortunately, increasing automation has led to a decrease in the number of human personnel used in various occupations; such personnel in the past would have been able to anticipate the special needs of BVI persons and provide assistance upon request.

There has thus been a growing trend of BVI individuals resorting to remote assistance from others by transmitting real-time images or videos [1], [2]. The sighted, remote, assistant comprehends the received images visually and passes along any information gleaned to the BVI user thereby "filling in" any of the latter's information gaps towards completion of the task. These systems, unfortunately, are either expensive to use due to high labor costs of the assistants (in the case of Aira), or untrustworthy due to the use of unknown and typically untrained volunteers (in the case of BeMyEyes) [18]. Additionally, contacting outside help may be restricted in situations that involve the workplace. Personal video calls through applications like FaceTime are commonly used as well, serving as an inexpensive, trustworthy option utilizing known helpers. Finding someone to help in a hurry may not be easy with such commodity applications, and they were never designed to serve BVI user needs [18].

This paper presents the design, implementation, and evaluation of the GuideCall remote video-based assistance system that allows a BVI user to seek and get assistance from a trusted set of known individuals through a free smartphone-based (usable also from other computing devices) application. GuideCall allows the user to populate and create trusted groups for specific life scenarios (such as work, personal) and reach out simultaneously to all members of a group when assistance is needed. The first person to accept the call takes on the assistant's role with all others notified that assistance is no longer needed. Beyond a simple video call interface that is designed to be BVI-friendly, GuideCall provides tools for a remote assistant to (i) control the BVI user's device to better assist them, and (ii) see real-time location information in embedded maps as a user moves around utilizing GPS or other indoor positioning information available. Such unique features are designed to make GuideCall to be more effective in completing the daily tasks where they need assistance.

A major contribution of this work is the presentation of implementation details of GuideCall on a hybrid mobile development platform that enables it be used not just from Android or iOS enabled mobile devices, but also through web browsers and desktop applications. Such cross-platform capabilities allow both BVI users and helpers to connect across a diverse array of computing systems unlike prior systems that are typically limited to one form-factor device and/or operating system. Evaluation results from GuideCall provide confirmation that GuideCall is effective in helping accomplish tasks that ordinarily would have been challenging for a BVI user. For some tasks, GuideCall is shown to be almost essential to be able to complete the task; for other tasks, GuideCall proves to be a convenient and adaptable solution that can ensure that the task is completed efficiently.

II. Related Work

Numerous assistive technologies have been developed over the years to assist BVI persons with daily tasks. With the advent of GPS-based positioning, outdoor wayfinding has become easier using mapping software on smartphones. Even though indoor wayfinding has still remained a big challenge, there have been many recent efforts in the area utilizing wireless devices or computer vision to provide location information and context within indoor spaces [16], [13], [8], [14], [9], [10]. In addition to its use for wayfinding, computer vision promises to serve as the "artificial eye" allowing a BVI person to capture and analyze images using a smartphone and identify text and objects around them as captured within images [17] resulting in many applications emerging such as CamFind [4], TapTapSee [15], KNFB [5], BeSpecular [3] and SeeingAI [6], or something as simple as an app for magnification [7]. Other advances have been in the area of web accessibility and screen readers allowing BVI persons to participate in today's increasingly cyber-enabled society.

The various assistive technologies mentioned above have limitations, however. GPS location accuracy isn't always good enough to guide BVI users on walkways outdoors. For indoor wayfinding, the use of wireless communication devices presents challenges in terms of interference with other wireless signals and fading, and of course creating the necessary infrastructure with adequate precision. Computer vision based approaches have limitations in terms of how well (angle, lighting, etc.) a BVI person can independently capture images for analysis in addition to being able to pose and receive meaningful queries and responses. Web and computer screen accessibility technologies can fail periodically due to inadequate conformity to standards by content designers or device resource limitations. In general, assistive technologies can fail due to software or hardware malfunctions or incorrect usage by the human involved and keeping a human in the loop can be less stressful towards overcoming daily visual challenges and/or breaking down social accessibility barriers.

Given that on-call human involvement can be important to provide BVI users the confidence to meet many daily challenges, it is natural that there have been efforts to facilitate such involvement through the use of technology. It can be challenging to find human assistance in the geographic vicinity of a BVI user when the need arises. Thus, most approaches have focused on remote assistance from someone that is expecting such a request and prepared to assist. BVI users can make individual calls and share images or live video from smartphones through apps like Skype, FaceTime, etc., but it may take a while before someone accepts to assist given the unannounced nature of the request and the sequential nature of finding such a person. Group calls on these applications, the faster option, on the other hand do not easily allow selecting one user and continuing a call with them for further assistance. These commercial collaboration services were also never designed keeping in mind the needs of a BVI user and to serve as an assistive technology with a remote person helping another for visual tasks and wayfinding and in some cases having the ability to control smartphone functions of a remote user. VizWiz was introduced as a way for a BVI user to pose questions about their environment to remote crowdworkers who could look at shared images and answer questions [11]. Aira is a successful commercial venture that allows the BVI user to call in to a dedicated number and receive assistance from a trained orientation and mobility specialist [1]. The BVI person can use a smartphone or an extra-cost eyeglass device that provides images to the remote helper to look at and provide assistance. BeMyEyes is a similar service to Aira but is based on the use of a smartphone app and relies on unpaid volunteers to serve as assistants [2]. All three of these approaches have limitations. Both VizWiz and BeMyEyes provide assistance from people who have not established trust with the BVI user. This is a very important factor for a BVI user to feel confident about the assistance they are receiving. Previous work has also identified safety as a factor to consider due to incidents where a BeMyEyes volunteer attempted to come to the geographic vicinity of the BVI user to provide assistance [12]. Because Aira uses trained employees, there is likely to be greater trust established with those providing assistance. Because helpers in systems such as Aira, BeMyEyes, and VizWiz are employees of another organization or outside volunteers, sensitive corporate information from a BVI employee could be transmitted outside the premises. The biggest disadvantage with Aira is the high cost of the service, which runs at about $0.50/minute and can easily add upto hundreds of dollars a month for a user. In summary, the disadvantages with existing remote assistance applications are well known and were elicited out in our prior survey of BVI users [18].

GuideCall as a system attempts to provide the level of quality of assistance BVI users receive from Aira but at a much lower cost or for free. It is built as an application that a BVI person can use from a mobile phone, tablet, desktop application or browser to request assistance from a trusted group of potential helpers who, over time, may become adept at knowing how to help a BVI user. Because assistance is sought from a chosen group simultaneously (in parallel), the delay in getting someone to help is likely to be much shorter than sequential calls for assistance. Multiple groups can be configured

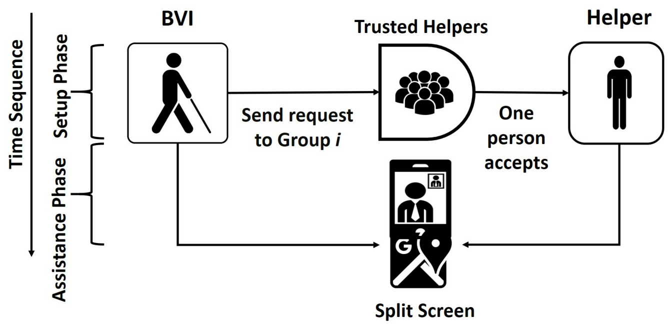


Fig. 1: GuideCall System Concept

for various assistance-requiring scenarios, including at work, home, or somewhere else. At work, where sensitive corporate information may not be shared with an outside person, a group could be created comprising only of co-workers. At home, trusted friends and relatives could provide assistance in finding documents or objects. Each potential helper can be labeled with trust levels. GuideCall provides real-time location updates for helpers using embedded maps allowing them to provide more knowledgeable assistance, not just in GPS-covered satellite locations, but also indoor locations that are so provisioned. A major advantage of GuideCall (by utilizing a hybrid mobile application development framework) is the ability to use it over multiple platforms; an BVI individual using an iOS smartphone can receive assistance from a helper on an Android device, web browser, or desktop application. This is a much more flexible paradigm improving the potential of connecting users and helpers in a mutually convenient manner.

III. GuideCall System

This section provides some details about the system architecture of the GuideCall application and its various components. Designed primarily as a single smartphone application, a scenario where it most useful, it operates in two modes: BVI user and Helper. The overview of system architecture is shown in Fig 1.

A. System Workflow

Whenever the BVI user needs assistance, the following steps will be followed through the application:

Step 1: The sign-in page is common for both BVI user as well as Helper. Once they create their account based on the user profile selection, they would be navigated to their own respective home page.

Step 2: For the BVI user , BVI mode screen activity with a help button is displayed. whereas, for the helper it directly navigates to the group chat page where list of users logged in to the application and chat messages would be displayed.

Step 3: To receive assistance, when the BVI user selects the help button, a notification "help needed" is sent to the devices of all potential helpers through a group chat application.

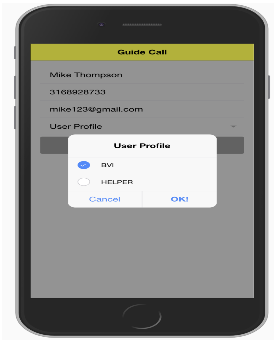
Step 4: One of the trusted helpers (who elects to provide assistance) will select the received message and click a call option that becomes available. This will result in video call to be connected between this user and the BVI user and a notification to be sent to all other users in the chat that a helper is connected successfully with the BVI user. This will assure other potential helpers that someone has accepted to be a helper for providing assistance.

Step 5: When the call is connected, the video of the BVI user will be shared with the helper. With the help of the video call, the helper can be the "eyes" of the BVI user in assisting with visual information. To assist effectively, a helper user can access features of the BVI user's phone such as cameras (front and back), flash light, and microphone (to switch the speaker on if necessary). The video call can be viewed in full screen by double tapping on the frame and allows pinching by the helper to zoom in and out.Along with the video call, Google Maps is integrated into the application to assist BVI users in outdoor environments. This allows the helper to study both the default view and the satellite view and understand the BVI user's location and orientation and guide them towards the destination. Incorporation of indoor maps is a unique feature for GuideCall. In locations provisioned with an accessible indoor wayfinding system such as GuideBeacon [14], real-time updates of the user walking can be shown on an image of the floor plan. This feature allows a helper to continuously learn about the context surrounding the BVI user and incorporate that in their instructions.

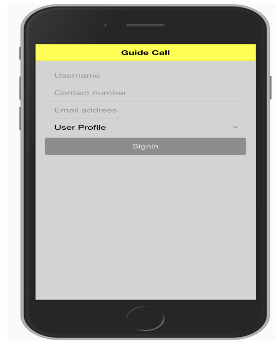
Step 6: Once a BVI user has got the assistance they need from a helper, they can end the call by clicking a done button. This button will send a message of "Thanks for helping me" to the group chat and enables others to understand that the BVI user was successfully assisted.

IV. Implementation

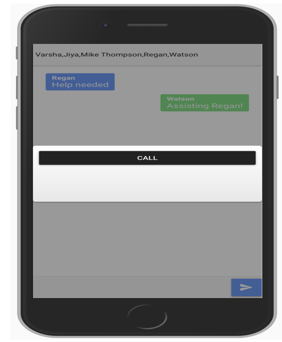
A major contribution of this work is how to design and implement an application such as GuideCall in a fashion that is not only functional and usable, but also affordable and available widely across various mobile platforms. A prior implementation of GuideCall [**?**] was only an Android application and used specific native constructs available only to that platform. Such native apps are developed specifically using a specific programming language and built using that platform's SDK and pre-defined functionalities. This has the major limitation that the mobile application needs to be developed separately for each platform. In the case of an application such as GuideCall where helpers or users could be on different platforms, it is very challenging to enable cross-platform interactions while preserving all features. Thus, in this section we present the design and implementation of GuideCall using a cross-platform development model. The rest of this section describes the use of such a paradigm in implementing GuideCall.



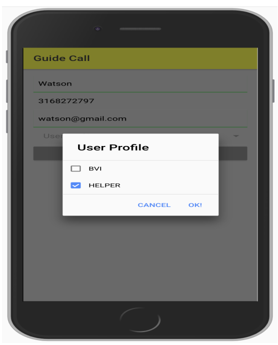
(a) BVI profile selection



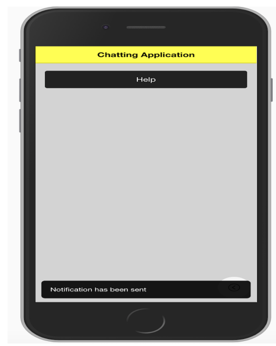
(b) BVI sign in



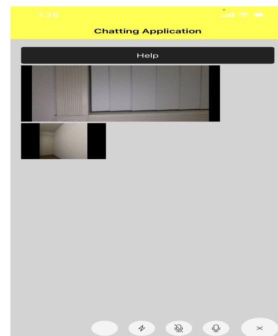
(c) Messaging and call features



(d) Helper profile selection



(e) Helper notification



(f) Assistance in progress

Fig. 2: Screenshots of various stages of GuideCall in action from both BVI and a helper's perspective.

A. *Hybrid application development framework*

In GuideCall, we have centered our design on the hybrid mobile application front-end framework "Ionic". This framework is an open source SDK (Software Development Kit) which allows the user to design and develop mobile applications with native-like impression and appearance. Ionic allows building apps that not only target iOS and Android, but also the web and native desktop, increasing the chances of interactions between helpers and BVI users. Using web standards through Ionic's components allows for a small code footprint that is also resilient against changes in platforms. For the current implementation, GuideCall was developed using Visual Studio and Cordova. Three elements used in GuideCall development were the Ionic command line interface (CLI), Ionic View, and the Ionic Model-View-Controller (MVC) architecture.user.

B. *Video calling*

Initially, Facetime was used for video calling features and this only supported iOS and not other platforms. To bridge that gap between iOS and other platforms we integrated our own custom solution by using ApiRTC from Apizee to the existing workflow. ApiRTC is an open source and cross-platform real-time communication service which supports audio and video calling features in GuideCall. Generally, the media flows are encrypted and uses peer to peer transmission. In the background it uses WebRTC technology, which is one of the communication standards developed by World Wide Web Consortium, further developed by the Internet Engineering Task Force (IETF). WebRTC technology facilitates the communication within browser-to browser which does not require registering or downloading other applications for audio and video calling features. GuideCall application includes a simple authentication sign-in procedure across all the platforms. In order to use various features of ApiRTC we must create an account in ApiRTC development center and obtain an API key. The obtained API key helps check for the user presence and user connection. The clients in WebRTC can directly communicate with each other bypassing the server which results in decrease of the latency. In GuideCall, if a user registers their details while creating an account, their information including their contact number would be stored in the database. When the BVI user seeks assistance, the helper will be notified on an application platform of their choice.

C. *Backend cloud architecture*

Firebase is a BaaS (Backend as a Service), which reduces configuration and setting up time. Firebase is an upgrade of APIs that allows extended functions such as messaging, data storage and service authentication. When the user creates their account and a profile, the data is synced to the Firebase real-time database including their location details as it can be used to locate them initially and guide them as per the requirement. The next time a user opens the application, it directly shows the user's homepage fetching the user details that were stored in the database. The registered details of the BVI user or helper is rendered in the group chat window to initiate WebRTC streaming with the appropriate individuals. Whenever the user downloads this application, a unique user ID would be created for every mobile device. When a BVI user seeks assistance and clicks on the help button, a notification would be sent to all the helpers present in the group. This unique user ID is helpful to send or receive notifications from other users in the application. For push notifications, we have made use of Firebase cloud messaging service which can be deployed on all the platforms. The back-end code in the application which responds to events triggered by real-time databases are executed using cloud functions for Firebase.

V. Evaluations

The primary objective in evaluating GuideCall was to measure its effectiveness in performing tasks that can be challenging for BVI persons to accomplish on their own. Four different tasks, that were identified through discussions with BVI individuals as independently challenging, were chosen.

A. Challenge Scenarios as Tasks

Outdoor Wayfinding: To locate and navigate to any unfamiliar outdoor location, BVI users typically employ a GPS-based mapping application. Unfortunately, the instructions from such mapping applications tend not to be fine-grained enough for BVI pedestrian navigation. Additionally, real-time information such as safety in crossing streets (especially those without any crosswalks and traffic lights) is lacking. For this task, we ask a BVI user to walk to a nearby (about 400 ft away using shortest path) building unfamilar to them. The paths require crossing a street without crosswalks or signals.

Indoor Wayfinding: Upon entry into any unfamiliar space, finding a specific destination within that space is always challenging. Without anyone to ask around, feeling around the entire space with hands is the only option currently if no other indoor wayfinding system for BVI users is present (as is typical). For this work, we ask a user to locate a laboratory room (about 100 ft away using shortest path) as they arrive at the entrance of a floor through an elevator.

Computer Screen Task: BVI individuals heavily depend on a screen reader to interpret textual information on a computer screen. In instances where a screen reader does not work (frozen computer screen, image on screen instead of text), help is necessary. In this task, a computer screen with an image mimicking a frozen screen with error message is displayed and BVI users are asked to comprehend the situation.

Locating Objects: A common challenge is that of locating an item within a space. Such situations arise in locating documents or objects within offices or homes especially if misplaced. In this task, BVI individuals are asked to find a misplaced handout document on an office desk kept second to last with seven other documents in the pile.

Both the BVI participant and helpers were using smartphones with GuideCall installed for these evaluations. For all tasks, participants were allowed to use any assistive technologies/techniques they would generally use to complete tasks. These include the use of smartphone mapping applications for outdoor wayfinding, apps like SeeingAI [6] or magnifiers like SuperVision [7].

B. Methodology

Eight BVI subjects with varying level of visual impairments (see Table I) were recruited to test GuideCall. Each participant was paid $50 to complete the four challenge tasks that took between 1-2 hours to complete. Results from only those participants who completed tasks without and then with GuideCall are shown in this paper. Some participants were unable to complete some of the tasks due to weather (outdoor wayfinding), or not having any mechanism to complete the task without GuideCall or due to lack of time. Effectiveness of GuideCall is measured using only the metric of tasks completion time in this paper; consideration of other quantitative and qualitative metrics will be explored in future work. For participants A,B, D, and F, the helper used to provide assistance was very with GuideCall and its features. For participants C, F, and H we used helpers who had no prior experience in using GuideCall.

Label	Participant characteristics
A	Blind, cane user
B	Light perception (LP) only, cane user
C	20/500 one eye, LP other, cane user
D	20/150 both eyes, ¡20° visual field, cane user
E	Blind one eye, 20/800 other, dog user
F	Blind, cane user, speech impairment too
G	LP only, mobility impairment too, used walker
H	LP, cane user

TABLE I: Participant labels and characteristics.

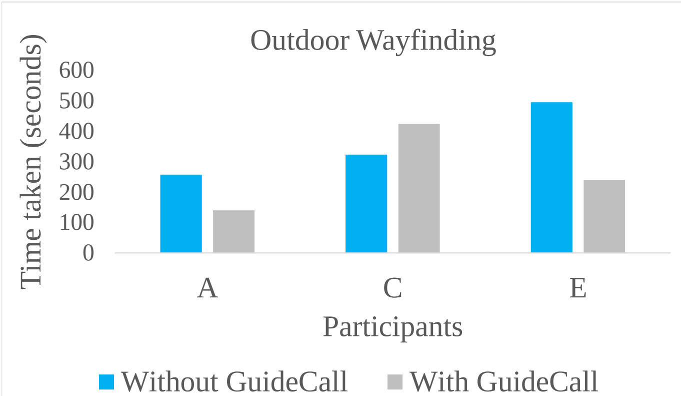
C. Results

Task completion times without and then with GuideCall are shown in Figures 3.

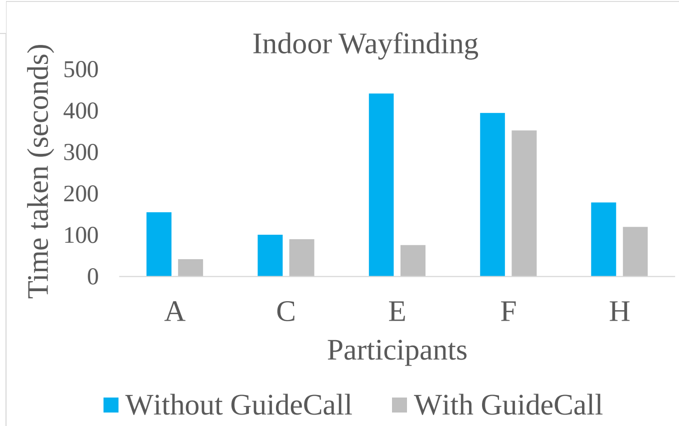
For the outdoor wayfinding task, results for participants A and E show notable benefits of using a helper for guidance over Google Maps. The result for participant C showed the participant taking slightly more time with GuideCall than without; this was due to the use of a helper that was new to GuideCall and its features.

The biggest benefits were seen for the indoor wayfinding task as the BVI subjects typically had to struggle to find the specified room in an unfamiliar environment, with no other technology options to complete the tasks. There were significant benefits in using GuideCall, except for participants C and f. We believe this was due to the fact that participant C had some useful vision that was utilized in conjunction with a smartphone magnifier app to wayfind indoors. Participant F's speech impairment in addition to the unfamiliarity of the helper used for that person may have contributed to the result.

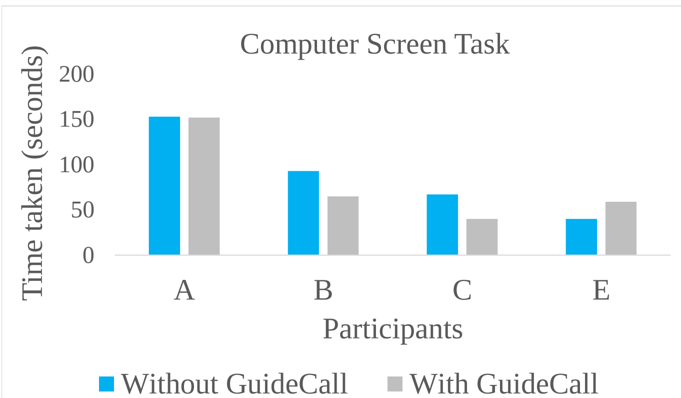
For the comprehending the computer screen task, those who had no way of identifying what is on screen were excluded from these results; GuideCall would have provided more benefits for such users. GuideCall still was typically better at interpreting screen information; the results for participant E was the only anomaly,. This was due to the user being very adept at using the KNFB app [5] for completing such tasks.



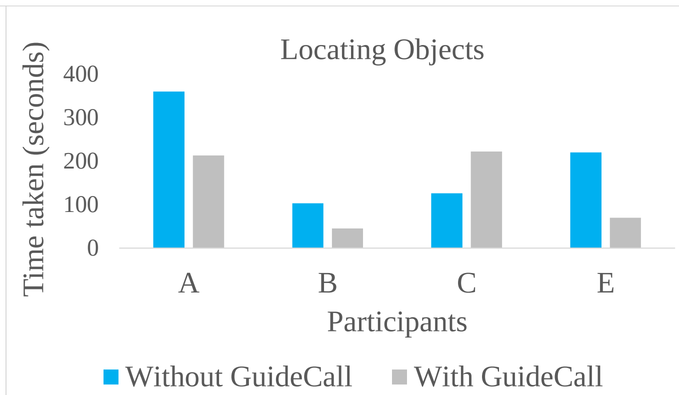
(a)



(b)



(c)



(d)

Fig. 3: Completion time for the four tasks under consideration: outdoor wayfinding, indoor wayfinding, computer screen comprehension, and locating a document of interest.

Finally, for the locating an object (document) task, only participant C was able to accomplish the tasks faster without GuideCall. This was primarily due to their residual vision which just required a magnifier to complete the task.

References

[1] Aira. https://aira.io/.

[2] Be My Eyes. http://bemyeyes.com/.

[3] BeSpecular. https://www.bespecular.com/.

[4] CamFind. https://camfindapp.com/.

[5] KNFB Reader. https://knfbreader.com/.

[6] SeeingAI. https://www.microsoft.com/en-us/seeing-ai.

[7] Super Vision. http://www.schepens.harvard.edu/superVision.

[8] Wayfindr open standard. https://www.wayfindr.net/.

[9] D. Ahmetovic, C. Gleason, C. Ruan, K. Kitani, H. Takagi, and C. Asakawa. Navcog: A navigational cognitive assistant for the blind. In *International Conference on Human Computer Interaction with Mobile Devices and Services*. ACM, 2016.

[10] D. Ahmetovic, M. Murata, C. Gleason, E. Brady, H. Takagi, K. Kitani, and C. Asakawa. Achieving practical and accurate indoor navigation for people with visual impairments. 2017.

[11] J. P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, A. Tatarowicz, B. White, S. White, and T. Yeh. Vizwiz: Nearly real-time answers to visual questions. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*, W4A '10, pages 24:1–24:2, 2010.

[12] S. M. Branham, A. Abdolrahmani, W. Easley, M. Scheuerman, E. Ronquillo, and A. Hurst. "is someone there? do they have a gun": How visual information about others can improve personal safety management for blind individuals. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '17, pages 260–269, 2017.

[13] Y.-J. Chang, S.-M. Peng, T.-Y. Wang, S.-F. Chen, Y.-R. Chen, and H.-C. Chen. Autonomous indoor wayfinding for individuals with cognitive impairments. *Journal of NeuroEngineering and Rehabilitation*, 7(1):1–13, 2010.

[14] S. A. Cheraghi, V. Namboodiri, and L. Walker. GuideBeacon: beacon-based indoor wayfinding for the blind, visually impaired, and disoriented. In *IEEE International Conference on Pervasive Computing and Communications*, March 2017.

[15] H. Kacorri, K. M. Kitani, J. P. Bigham, and C. Asakawa. People with visual impairment training personal object recognizers: Feasibility and challenges. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 5839–5849, 2017.

[16] J.-E. Kim, M. Bessho, S. Kobayashi, N. Koshizuka, and K. Sakamura. Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, SAC '16, pages 604–611, 2016.

[17] R. Manduchi and J. Coughlan. (Computer) vision without sight. *Communications of the ACM*, 55, 2012.

[18] V. Namboodiri, N. M. Ravindran, S. A. Cheraghi, and R. Bab. Guidecall: Affordable and trustworthy video call-based remote assistance. *Journal on Technology & Persons with Disabilities*, 8:53–68, March 2020.