BubbleCam: Engaging Privacy in Remote Sighted Assistance

Jingyi Xie*
Pennsylvania State University
University Park, PA, USA
jzx5099@psu.edu

Rui Yu*
University of Louisville
Louisville, KY, USA
rui.yu@louisville.edu

He Zhang
Pennsylvania State University
University Park, PA, USA
hpz5211@psu.edu

Sooyeon Lee
New Jersey Institute of Technology
Newark, NJ, USA
sooyeon.lee@njit.edu

Syed Masum Billah
Pennsylvania State University
University Park, PA, USA
sbillah@psu.edu

John M. Carroll
Pennsylvania State University
University Park, PA, USA
jmc56@psu.edu

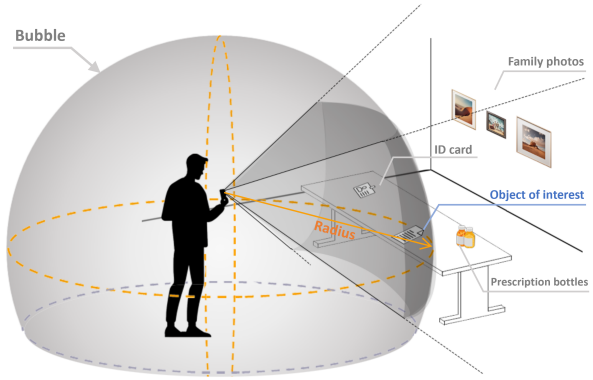


Figure 1: Illustration of Privacy-Preserving RSA with BubbleCam. BubbleCam creates a virtual "bubble" (depicted in light grey semisphere) with an adjustable radius, where only objects within are visible to remote-sighted volunteers (camera view depicted in dark grey). This figure demonstrates the user placing an object of interest inside the bubble for visibility, while sensitive items (ID cards, prescription bottles, and family photos) remain outside and hidden from the volunteer.

ABSTRACT

Remote sighted assistance (RSA) offers prosthetic support to people with visual impairments (PVI) through image- or video-based conversations with remote sighted assistants. While useful, RSA services introduce privacy concerns, as PVI may reveal private visual content inadvertently. Solutions have emerged to address these concerns on image-based asynchronous RSA, but exploration into solutions for video-based synchronous RSA remains limited. In this study, we developed *BubbleCam*, a high-fidelity prototype allowing PVI to conceal objects beyond a certain distance during RSA, granting them privacy control. Through an exploratory field study with 24 participants, we found that 22 appreciated the privacy enhancements offered by BubbleCam. The users gained autonomy, reducing embarrassment by concealing private items, messy areas, or bystanders, while assistants could avoid irrelevant content. Importantly, BubbleCam maintained RSA's primary function without compromising privacy. Our study highlighted a cooperative approach to privacy preservation, transitioning the traditionally individual task of maintaining privacy into an interactive, engaging privacy-preserving experience.

CCS CONCEPTS

• **Human-centered computing** → **Accessibility**; *Empirical studies in accessibility.*

KEYWORDS

People with visual impairments, privacy, remote sighted assistance, computer vision

*Both authors contributed equally to this research.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05…$15.00
https://doi.org/10.1145/3613904.3642030

ACM Reference Format:
Jingyi Xie, Rui Yu, He Zhang, Sooyeon Lee, Syed Masum Billah, and John M. Carroll. 2024. BubbleCam: Engaging Privacy in Remote Sighted Assistance. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3613904.3642030

1 INTRODUCTION

People with visual impairments (PVI) are confronted with various privacy challenges in their daily lives. These challenges span across physical spaces, such as navigating public transit [10, 12, 13], avoiding home intrusions [12, 13], and secure transactions [12, 13, 18, 28, 41], to the digital world where they may encounter inaccessible CAPTCHAs [12, 44, 69] or become targets of scams and malicious software or emails [41, 44, 63, 66, 69]. Furthermore, even the assistive technologies designed to aid them could pose privacy concerns, such as aural eavesdropping when using screen readers [10, 12, 62, 85], or visual eavesdropping when using screen magnifier or enlarge screens [12, 41, 48, 83].

One such assistive technology, remote sighted assistance (RSA), is a camera-based conversational aid for PVI. It operates in either asynchronous or synchronous mode by connecting a visually impaired user (*"user"* for short) with a remotely-located, sighted assistant (*"assistant"* for short). Asynchronous RSA systems accept photos and queries sent by users and return responses later [20, 21, 26, 36]. Conversely, synchronous RSA offers real-time interactions, where users share the live camera feed with the assistants, who interpret the video feed and converse with the users to provide assistance [17, 42, 46, 56]. Although RSA services provide extensive assistance from low-stake inquiries [17, 20, 26] to high-stake navigational tasks [57, 80], they introduce privacy concerns. Users might share private visual content on RSA platforms [11, 15, 35, 37, 45, 74], as they cannot review such content before sharing [12].

Computer vision-based solutions have emerged to address privacy concerns on asynchronous, imaged-based content sharing [38–40, 49, 58, 60]. This process involves the user capturing a picture of the scene of interest, followed by an AI model (e.g., a deep neural network), automatically detecting and obscuring any sensitive information in the image before sharing. These approaches are not developed specifically for RSA and rely on asynchronous, image-based communication. However, the privacy implications of live video-based RSA have yet to be explored. The significant challenges include the absence of highly precise models specifically crafted for the detection of private content and the limited computing speed on mobile devices for real-time applications.

To address this gap, we developed a high-fidelity prototype – *BubbleCam* that can conceal objects beyond a certain distance. It empowers users to control the degree to which they preserve their privacy. We utilize iOS devices featuring a LiDAR scanner to accurately measure the distance between the device and scene objects, then re-render the camera image based on the pixel-wise distance in real-time. During synchronous RSA sessions, BubbleCam facilitate users to establish a virtual "bubble" concealing everything behind it (Fig. 1). This introduces a simple, distance-based privacy-preserving strategy that is not reliant on individuals' subjective definitions of private objects.

In BubbleCam, we select "radius" as the dependent variable for its conceptual simplicity in representing the bubble's size. This allows for straightforward adjustments of the bubble's size, thereby controlling the protection strength by making items visible or not. Moreover, the radius is an intuitive metaphor for 3D space. It reduces the complexity of defining a 3D area with a single parameter, making it more perceptually accessible for users to comprehend and manipulate.

In this study, we explore whether participants are willing and able to appropriate the new interaction that BubbleCam is offering to work on the privacy goals, and how accurately BubbleCam can meet the goals that participants are achieving. We investigate the following *research questions*:

RQ$_1$. *How do visually impaired users and sighted assistants experience BubbleCam in protecting the privacy during synchronous RSA interactions?*

RQ$_2$. *What are the trade-offs between privacy and utility in privacy-preserving RSA interactions?*

To answer the questions, we conducted an exploratory field study with 24 participants in two non-profit agencies, where users either work or receive services. These settings are realistic and familiar to them, providing more ecological validity than lab environments. For the assistants, we recruited volunteers from Be My Eyes [4], a free RSA service. This platform adopts a relatively looser privacy policy than subscription-based RSA services [73], elevating the risks of potential privacy breaches. We mocked up three scenarios – office space, home environment, and shared social space, incorporating visual content deemed private or sensitive by PVI in RSA interactions [15, 35, 74]. In this study, we situated PVI's stated concerns [15, 74] within a field study and analyzed their actual behavior to supplement prior work.

We found that BubbleCam balances utility and privacy, effectively obscuring unintended objects without undermining RSA's core function, which is to assist users with visual cues from their cameras. Significantly, 22 out of 24 participants appreciated the privacy enhancements offered by BubbleCam, a contrast to regular RSA interactions where users often do not engage or have to compromise their privacy to obtain help.

We observed a cooperative approach to privacy preservation between users and volunteers, with both parties engaged in establishing and maintaining privacy. In this dynamics, volunteers actively alerted users to privacy breaches and suggested reducing the bubble radius, and users adjusted the radius with the help of volunteers. Unlike the single-actor approach to privacy protection (e.g., concealing one's identity on social media [75]), BubbleCam introduces a novel arrangement. Here, two actors co-productively configure privacy, transforming the traditionally individual task of maintaining privacy into an interactive and engaging experience. Different from prior cooperative privacy protection between PVI and their trusted allies [41], we delved into BubbleCam's potential in fostering such cooperation between PVI and strangers, specifically Be My Eyes volunteers.

Our findings align with the broader body of privacy literature, demonstrating their implications in the domains of contextual integrity, differential privacy, and user-centered privacy. This intersection presents notable contributions to ensuring more secure and respectful user experiences in RSA interactions while reflecting broader, foundational principles of privacy preservation across various contexts and technologies.

2 BACKGROUND AND RELATED WORK

Researchers have explored the nuanced perceptions of privacy among PVI. Hayes et al. [41] found that PVI view privacy as ownership and control over their personal information, or "the right to be let alone" [72]. This view reflects PVI's desire for people to respect their privacy by understanding that they don't want to share certain information. Similarly, Stangl et al. [74] reported PVI's definition of privacy as being a safeguard or maintaining a sense of control or ownership. Thus, privacy concerns are related to the loss of aforementioned factors like control, ownership, or the ability to manage [74]. Next, we outline PVI's privacy concerns and tools designed to enhance their privacy.

2.1 Asynchronous and Synchronous Remote Sighted Assistance

The remote sighted assistance (RSA) platform provides prosthetic support to PVI by connecting them with remotely located, sighted individuals. Utilizing various communication mediums, such as images [21, 52] and video [2, 4, 19, 25, 34, 42], RSA services incorporate both asynchronous and synchronous modalities.

Asynchronous RSA accepts photos and queries sent by PVI and returns responses after a period. One such system, Vizwiz [20], enables PVI to upload images with audio-recorded questions, and in response, they receive text-based answers via crowdsourced assistance. This has proven effective in tasks such as text reading [20], color identification [20], finding object [21], and fashion advice [26]. Building on this, Gurari et al. introduced Vizwiz Social [36], connecting PVI to their social networks like friends and family members. Given the constraints of single-photo and single-query input [20], asynchronous RSA is generally deemed unsuitable for complex, sequential contextual inquiries [54].

Conversely, synchronous RSA offers real-time, extended, and continuous interactions. PVI can thus receive instantaneous guidance tailored to their immediate environment and context. Synchronous RSA systems are equipped with video streaming capability. With the advent of technologies, synchronous RSA has transitioned from utilizing wearable digital video cameras [19, 34, 43] or webcams [25, 29, 70] to video apps on mobile devices [2, 4, 42, 82], and even enabling real-time screen sharing [53, 54, 81]. The immediate nature of synchronous RSA is particularly beneficial for high-stake or complex tasks that require timely feedback or decision-making like navigation [19, 25, 27, 29, 34, 42, 46, 55–57, 70, 80–82], shopping [57, 80, 81], and social engagement [27, 55, 56]. However, the effectiveness of synchronous assistance largely depends on the reliability of internet connectivity and the quality of video or audio transmission [29, 34, 42, 43, 46, 56]. In this study, we examine BubbleCam within the specific context of synchronous video-based RSA.

2.2 Privacy Concerns with Camera-based Assistive Technologies

PVI might knowingly or unknowingly share private visual content, like photos or videos, with asynchronous or synchronous RSA [11, 15, 35, 37, 45, 74], as they cannot review such content before sharing [12]. Their awareness of such inadvertent privacy breaches can sometimes be limited [13].

Various research efforts have sought to categorize the types of visual content considered private by PVI. Gurari et al. [35] classified 19 types of private visual content from the VizWiz dataset [21]. Expanding on this, Stangl et al. [74] gauged PVI's concern levels for 21 types across human-powered RSA and AI-powered visual description services, with direct input from PVI through interviews. Of particular relevance to our study is Akter et al.'s work [15], which examined the VizWiz dataset and observed 5 types of private visual content: address, prescription labels, credit card information, contents of digital screens, and the presence of faces or body parts of PVI and bystanders. They probed PVI's privacy concerns about disclosing background objects (irrelevant to the task at hand) to various RSA assistants, including Be My Eyes volunteers, across three scenarios – restaurant, office, and home.

Researchers noted that low-vision users are more concerned than totally-blind users about disclosing private information (e.g., personally identifiable information) [12, 15]. Prior work uncovered various users' needs and concerns for privacy-preserving camera-based assistive technologies. This study focused specifically on RSA prosthetics and leveraged previous findings on private visual content to guide task design. Our study situated PVI's stated concerns within a field study and analyzed their actual behavior to complement prior work. Moreover, our participants included both totally blind and low-vision users to explore the potential impact of different visual conditions on their actual behavior.

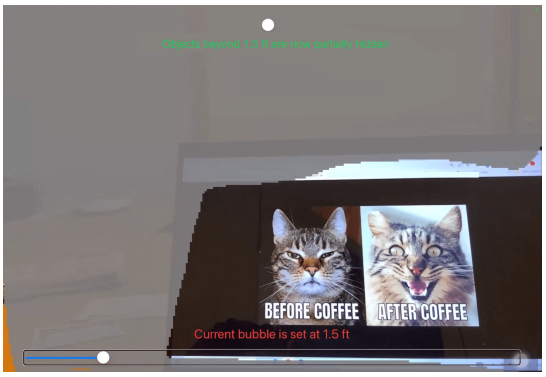
PVI's Concerns of Intruding Bystanders' Privacy. When using camera-based assistive technologies, PVI may capture and share bystanders in images or videos [13] on platforms like social media or RSA services [14, 15, 88]. Akter et al. [15] noted that PVI are more concerned about bystanders' privacy than their own in images. Researchers [14] explored both PVI's and bystanders' perspectives and revealed mutual privacy concerns about AI misrepresentations of bystanders' actions and attributes (e.g., gender). These insights call for future solutions to safeguard bystanders' privacy [88].

Previous studies investigated the perspectives of three stakeholder groups [14, 15, 41] (PVI, their trusted allies, and bystanders), but not the viewpoints of anonymous volunteers. This study focused on synchronous RSA and examined the privacy concerns of both stakeholders: PVI and volunteers. Building on prior work, we probed the privacy perceptions of Be My Eyes volunteers, identifying content that they find uncomfortable to view during RSA, and how our prototype could alleviate these concerns. Also, we examined the efficacy of BubbleCam in protecting bystanders' privacy.

2.3 Privacy-Enhancing Interventions for Camera-based Assistive Technologies

To protect privacy in camera-based assistive technologies, PVI often physically clear exposed areas and remove sensitive items before camera use [15]. Service providers have established privacy policies to regulate the collection, length of retention, use, and dissemination of visual content shared by PVI. Stangl et al. [73] found that RSA services like Aira and Be My Eyes collect data for AI training, and even share it with third parties. Although both platforms allow general data (not visual data) deletion, only Aira allows users to opt out of data collection.

(a) Fully hidden mode. This example, sourced from the field study (utilized by U6 for Task 3), obscured the W-2 form and medical record summary present in the background.



(b) Partially hidden mode. This example, sourced from the field study (utilized by U2 for Task 2), obscured the Social Security Card and boarding pass present in the background.

Figure 2: Fully and partially hidden modes in BubbleCam. Note: The names and tabs on the laptop were obscured by the authors.

Besides policy-based interventions, researchers suggested computer vision algorithms to detect sensitive or inappropriate content in camera-based assistive technologies. Current solutions are primarily designed for online photo sharing rather than camera-based remote assistance. These solutions typically use computer vision algorithms to detect specific sensitive elements within an image (e.g., faces, bystanders [38, 58]), then employ obfuscation techniques (e.g., blurring, pixelating, inpainting, avatar, DeepFake [39, 40, 49, 60]). Notably, Zhang et al. [87] introduce a human-AI hybrid method to help PVI detect and modify sensitive content in images. It begins with computer vision algorithms identifying potential sensitive information, which is then reviewed by trusted sighted allies for confirmation and processing. This enables PVI to safely share photos on social networks while maintaining privacy.

It's worth highlighting that the existing methods are primarily developed for sharing static images rather than live videos. To the best of our knowledge, the utilization of computer vision for privacy preservation in camera-based assistive technology, particularly in synchronous RSA, remains conspicuously absent from the literature. This gap presents two core challenges: firstly, the absence of a highly precise model specifically crafted for the detection of private objects intended for PVI; and secondly, the need for an efficient model capable of facilitating real-time detection and obfuscation on mobile devices, a crucial requirement for practical deployment.

In contrast to prior interventions, the novelty of BubbleCam is twofold. First, it introduces a simple distance-based privacy-preserving strategy that does not hinge on individuals' subjective definitions of private object categories. Second, it achieves real-time depth estimation and object occlusion directly on mobile devices, representing a notable step forward in the development of privacy-preserving solutions for synchronized RSA. Depth estimation of BubbleCam is achieved by adopting an off-the-shelf ARKit framework and LiDAR-equipped iOS devices. Using the same framework, Zea and Hanebeck [86] found that iPad Pro consistently obtained depth estimation error within 1% and 2% for different distances to a paper chessboard. Another independent study [76] reveals that the errors of about 90% of depth measurements by iPad Pro are less than 1 mm. Thanks to the accurate depth estimation facilitated by ARKit and LiDAR scanner, BubbleCam reliably achieves the distance-based occlusion effects.

3 SYSTEM IMPLEMENTATION

This section details the design and implementation of BubbleCam, and its application within the context of the RSA paradigm.

3.1 BubbleCam Mobile App

We introduce BubbleCam, an iOS app that can enhance the camera functionality of mobile devices. BubbleCam automatically obstructs objects positioned beyond a user-defined distance threshold from the camera. In Fig. 2(a), we illustrate an example that BubbleCam blocks all objects located more than 0.8 feet (ft) away from the camera. When applied to RSA, PVI can employ BubbleCam to establish a virtual "bubble" concealing everything behind it.

At the bottom of the BubbleCam interface is a slider that enables users to adjust the bubble's radius. Setting the radius to 0 ft blocks all camera content. Moreover, a top switch button provides control over the visibility of objects situated behind the bubble, with "fully hidden" mode completely obscuring the view and "partially hidden" mode rendering scenes behind the bubble translucent. The difference between the two modes is shown in Fig. 2. Additionally, BubbleCam has been designed for integration with the VoiceOver screen reader on iOS to ensure accessibility for PVI.

For implementation, we use iOS devices featuring a LiDAR Scanner, like the new iPad Pro or iPhone Pro, in conjunction with ARKit [3]. The LiDAR Scanner effectively measures the distance between the device and scene objects, while ARKit provides a buffer to restore the depth values. To render the camera image to BubbleCam in real-time, we've devised a small Metal [7] renderer that leverages the device's GPU to draw the graphics on the app. During the rendering process, the renderer examines the depth texture at each pixel and applies color according to the pixel's distance from the device. If this distance surpasses a specified threshold, we substitute the pixel's color with the occlusion color (grey in

Table 1: Blind and low-vision users' demographics.

ID	Gender	Age Group	Condition of Vision Impairment	Age of Onset	Occupation Type
U1	F	45-50	Totally blind, retinopathy of prematurity	At birth	IT consultant
U2	M	40-45	Totally blind, retinal detachment	since 1995	General assembler
U3	F	25-30	Low vision, Peter's anomaly, low vision in the left eye, no vision in the right eye	4 yrs old	Unemployed
U4	M	35-40	Low vision, nystagmus, high myopia, stem cell deficiency, shaky eyes	At birth	Material handler
U5	F	65-70	Totally blind, acute congenital glaucoma	At birth	Retiree
U6	M	30-35	Low vision, Red X-linked retinoschisis	At birth	Marketing manager
U7	M	40-45	Low vision, optic atrophy, pigmentosa	5 yrs old	Nursing home staff
U8	F	55-60	Totally blind, retinitis pigmentosa	5 yrs old	Stay-at-home mom
U9	M	35-40	Totally blind, retinopathy of prematurity	6 yrs old	Amazon employee
U10	M	55-60	Low vision, optic atrophy	At birth	Interpreter
U11	F	65-70	Totally blind, retinal blastoma	1 yr old	Retiree
U12	F	30-35	Low vision, albinism	At birth	Healthcare fundraiser

our app). The partially hidden mode is accomplished by applying a weighted blend of the original color and the occlusion color.

In detail, the processed data includes a stream of color images and depth images at 60 fps. Consider a pixel in the color image with original RGB value vector (r, g, b), where each component is a 16-bit floating-point value ranging from 0 to 1. Correspondingly, the depth image contains a 32-bit floating-point value d to represent the depth in meters at that pixel. To render occlusion, we designate the occlusion color as a grey shade with RGB values of $(0.5, 0.5, 0.5)$. The color displayed on the BubbleCam app at that pixel is computed as $(r', g', b') = \alpha(r, g, b) + (1 - \alpha)(0.5, 0.5, 0.5)$, where $\alpha \in [0, 1]$ denotes transparency. Assuming the current bubble radius is R, if the depth value $d < R$ (*i.e.*, the object on that pixel lies within the bubble), α is set to 1, signifying fully transparent. In the case of $d \geq R$ (*i.e.*, the object on that pixel is outside the bubble), it should be hidden. For fully hidden mode, α is set to 0. For partially hidden mode, α is set to 0.08 to balance occlusion and visibility. The low transparency ($\alpha = 0.08$) ensures an effective occlusion effect while allowing for a discernible representation of the background's geometric layout.

BubbleCam is built upon off-the-shelf ARKit framework and LiDAR-equipped iOS devices. Depth estimation is achieved by integrating depth and color streams through Apple's machine learning algorithms, as illustrated in Apple's patent [84]. Although the errors of most measurements are less than 1 mm [76], the accuracy of depth estimation is affected by various factors, like the distance between the sensor and the object, lighting conditions, and the surface texture. Notably, depth measurement accuracy tends to be lower on highly reflective or absorbent surfaces [86]. The LiDAR scanner in iOS devices has a maximum measurement range of 5 meters. To accommodate both measurement accuracy and common usage distances, we set the adjustable bubble radius range from 0 to 10 ft (around 3.048 meters).

3.2 Applying BubbleCam to RSA

We integrate BubbleCam into RSA to enhance the camera capabilities for privacy preservation. Since the current RSA services [2, 4] are limited to accommodating only two individuals, we have chosen to utilize the Zoom teleconferencing app [9]. Zoom closely aligns with the RSA applications in terms of audio connectivity and video transmission features, facilitating effective communication between the visually impaired and sighted participants. This resemblance has been exploited in previous studies [81] to replicate RSA interactions in research settings.

In synchronous RSA sessions, users share the live stream from BubbleCam with remote volunteers. By utilizing Zoom, remote researchers can access the same view of BubbleCam as the volunteers, which greatly improves their comprehension of the volunteer's BubbleCam experience. This setup also allows remote researchers to conduct focus-group and one-on-one interviews to gather more insights.

4 METHOD

We conducted an exploratory field study (IRB-approved) with 12 visually impaired users and 12 sighted volunteers to investigate the feasibility, desirability, and challenges of BubbleCam as well as to understand our research questions.

4.1 Participants

Recruiting blind or low-vision users. We recruited 12 visually impaired participants (6 males and 6 females, 6 blind and 6 low-vision) by collaborating with two non-profit agencies – Lighthouse Guild in New York and North Central Sight Services Inc. in Pennsylvania. Each visually impaired participant uses and is familiar with RSA services (e.g., Be My Eyes [4], Aira [2]). Their common age groups are 30-35, 35-40, 40-45. One of them is unemployed, two are retired, and the rest are full-time employees. Table 1 presents their demographics. Each visually impaired participant received a $45 gift card per session for their time and effort. We will hereafter refer to visually impaired participants as *users*.

Recruiting sighted volunteers. We recruited 12 sighted participants (6 males and 6 females), with the most common age group of 25-30. Table 2 presents their demographics. Most are students and have received fewer than 5 calls. V1 is a notable exception, he

Table 2: Sighted volunteers' demographics.

ID	Gender	Age Group	Number of Calls Received	Year Registered	Occupation Type
V1	M	40-45	≥50	2018	Digital accessibility programs coordinator
V2	M	25-30	<5	2023	Student
V3	F	30-35	≥15	2020	Teacher
V4	M	25-30	<5	2021	Student
V5	F	20-25	<5	2021	Student
V6	F	30-35	<5	2021	Student
V7	F	25-30	≥10	2020	In-between jobs
V8	M	20-25	<5	2021	Student
V9	F	20-25	<5	2021	Student
V10	F	25-30	<5	2022	Student
V11	M	25-30	<5	2023	Student
V12	M	25-30	<5	2022	Student

is a coordinator of a specialized help program on Be My Eyes and has received over 50 calls. Each sighted participant received a $45 gift card per session. We will hereafter refer to sighted participants as *volunteers*.

4.2 Apparatus

The user was provided an iPad Pro equipped with BubbleCam, along with a holder to keep the iPad Pro stationary on the table if needed. Given the limitation of current RSA services to support only two individuals simultaneously, we chose the Zoom teleconferencing app (Section 3.2), where volunteers and at least one researcher joined remotely.

4.3 Environment

We conducted 12 RSA sessions in two non-profit, social support agencies, where users are either employees or customers. These agencies provided us with their office spaces for the research sessions. Throughout the experiments, the users remained seated in a conference room to ensure their safety. Concurrently, volunteers participated via Zoom.

We mocked up three scenarios: office space, home environment, and shared social space. Those are similar choices to previous research [15], which reflect everyday, widespread use of RSA services [1, 6, 56] and carry risks of inadvertent disclosure of sensitive information [10, 15, 41, 88].

In these scenarios, users had the autonomy to select and position relevant objects on the table. Sensitive information on these items was pre-fabricated by our research team, as outlined in Table 3. To mimic real-life settings, researchers consistently reminded users to regard the environment as their own workspace or home when selecting and arranging the objects. Specifically, researchers handed objects to users one at a time, described each object, and encouraged them to exclude any objects that they typically wouldn't have on their own table or wouldn't seek help with in past Be My Eyes calls. Then, users were asked to organize the chosen objects on the table according to their usual practices.

This approach proved realistic [23, 59], as reflected in the users' strategies. They divided the table into sections based on the significance of items (U4, U7, U12), describing the arrangement as "in front", "on top", "to the left", and "to the right". Others (U8, U10, U11) chose a more "random" placement, like "all over the place". Self-arrangement was necessary for the subsequent task performance. Moreover, involving users in the task space enhanced the study's ecological validity.

4.4 Task Design

Based on prior work [17, 46, 56] and publicly available data from Be My Eyes websites [1, 6, 8], we identified five tasks that are commonly rendered in RSA and could be safely simulated in our setting through role-playing (Table 3): (1) reading a hand-written letter; (2) describing a meme; (3) finding a medicine bottle and a box; (4) reading a restaurant menu; and (5) parenting a tween to select storybooks. Tasks were performed in the same order as this list during each session. We contextualized the tasks within the scenarios of office, home, and shared social spaces (Section 4.3). Although some of the tasks entailed similar activities, such as reading text, they occurred in distinct scenarios featuring varied background objects and users' different interactions with the surroundings. Each scenario involved 1 or 2 tasks.

To assess the performance of BubbleCam, we chose a set of foreground and background objects, similar to prior work [15]. Users need to manipulate BubbleCam to perform tasks in the foreground while hiding sensitive information in the background. The foreground objects, which were non-sensitive, needed to be visible to volunteers for task completion. In contrast, the background objects, while relevant to each scenario, weren't essential for the task completion and contained potentially sensitive information that needed to remain obscured from volunteers. The foreground and background objects used in each scenario are listed in Table 3. To ensure the relevance and authenticity of our tasks and object selections (both foreground and background), we consulted and validated them in meetings with the local chapter of the National Federation of the Blind (NFB). The NFB members found these tasks to be consistent with real-life situations they often encounter on RSA platforms.

Users' self-arrangement of items was involved in each task, including Task 3, which was finding a Tylenol bottle and a COVID-19 test box amidst other bottles and boxes. Due to the items having similar contours and lacking tactile labels, even with self-arrangement,

Table 3: List of scenarios, tasks, and corresponding foreground and background objects. Foreground objects are essential for task completion and need to be shown to remote volunteers. Conversely, background objects are relevant to each scenario but unnecessary for task completion. Background objects contain *potentially sensitive information* that can either be positioned outside the camera feed or hidden with BubbleCam if within the camera feed.

Scenario	Task with Foreground Objects	Background Objects with *Potentially Sensitive Information*
Office Space	Reading an unfolded, hand-written letter from NFB	Flyers, *envelopes with name and address, passport, social security number, credit cards, W-2 form, insurance form, family photos*
	Describing a meme on a desktop screen	
Home	Finding a Tylenol bottle and a COVID-19 test box	General medicine bottles, boxes (similar in size to the COVID-19 test box), *prescription bottles, medical record summary, COVID-19 test result, passport, social security number, credit cards, W-2 form, insurance form, family photos*
	Parenting a tween to select storybooks*	*Face or body part of the tween, roommates, family photo on the wall, messy area*
Shared Social Space	Reading a restaurant menu*	Plate, napkins, water bottle, *face or body part of another customer and the customer's child (doll used for simulation)*

* Role-playing tasks. One researcher played the role of the tween and the customer. When there were two researchers, the other researcher played the role of a roommate and another customer, aligned with the respective tasks. The order of these two tasks was reversed in the study.

finding target objects was challenging for users. This setup realistically reflects the complexity users often face in distinguishing medicine bottles or boxes with similar shapes in their daily environments [1, 16, 22, 32].

In the tasks of parenting a tween and reading a restaurant menu, one researcher played the role of a tween in the home setting and the role of a customer in the shared social space. If there were two researchers, the other researcher played the role of a roommate or another customer, aligned with the corresponding scenarios. The integration of role-playing made the tasks more realistic and engaging, thereby broadening the scope for exploring BubbleCam's capabilities (e.g., hiding bystanders). Note that a doll was utilized to simulate the presence of a child. No individuals under the age of 18 were involved at any stage of the study.

4.5 Procedure

We conducted 12 RSA sessions in total, with users attending in-person at non-profit agencies and volunteers participating via Zoom. One or two researchers were present on-site for equipment setup, BubbleCam instruction, and role-play facilitation. Meanwhile, at least one researcher joined via Zoom, sharing the same view of BubbleCam as the volunteer, to monitor network connections, understand the volunteer's experience, and conduct a one-on-one interview with the volunteer. We recorded the sessions after consent. Each lasted for 75 to 90 minutes and was divided into four parts.

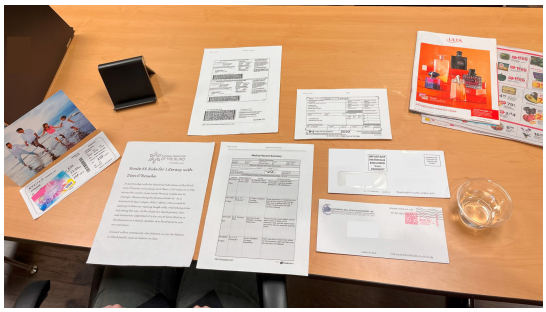
Part 1: training. Users and volunteers received an overview of the experiment and non-technical instructions on using BubbleCam. Participants were given sufficient time and guidance to familiarize themselves with BubbleCam and ask questions, including a practice trial to adjust the bubble's radius. During the trial, users tweaked the radius to ensure that flyers, whether on the table or in their hands, were visible to volunteers. Volunteers provided feedback on what they could see at various radii and read flyer content. Different information was placed on the table for the subsequent formal assessment. This part lasted for 10 to 15 minutes, concluding when participants were ready to proceed.

Part 2: task rendering. Each pair of the user and volunteer completed a total of 5 tasks (Section 4.4) in the same order. One exception is the U1-V1 pair. They only completed the first two tasks because BubbleCam was sensitive to U1's micro-movements, which caused it to constantly shift on the screen and distract V1. This exacerbated V1's neurodivergence, so we shortened the task-rendering part and focused on collecting their narrative feedback.

Before each task, users arranged objects on the table according to their habits (Fig. 3) or removed objects not typically present during prior Be My Eyes calls. For example, U2 removed prescription bottles, U4 removed a passport, and U5 removed credit cards. Remote volunteers could not see users placing or removing objects. During the tasks, users had complete control over the bubble's radius, with the option to partially or fully hide objects. Apart from role-playing, researchers only observed and avoided any judgments or assistance. The duration of task rendering varied from 30 to 45 minutes and concluded once all the tasks were accomplished.

Part 3: focus group. We organized a focus group via Zoom, involving both user and volunteer, to bring all perspectives and consider them together. We prepared prompts to encourage dialog, including the (in)feasibility of BubbleCam in different scenarios and their suggestions for improvement. Involving other participants might introduce a source of bias in opinion sharing. However, this potential bias originated before the focus group, as participants had conversed with each other during collaborative experiences in privacy preservation. The focus group reflected and extended the collaborative dynamic, enabling participants to evoke memories and vocalize more issues. Besides, researchers actively moderated to make sure everyone share [61, 68]. This part lasted for 10 to 15 minutes.

Part 4: one-on-one interviews. Finally, we conducted one-on-one interviews with each participant. The interviews concentrated on the individual level, with participants focusing on their own perspectives. We aimed to elicit feedback from distinct and individual perspectives in a more profound manner. This allowed us to capture insights that were potentially overlooked or left unshared during the focus group [61]. These interviews were semi-structured. For the user, a researcher conducted the interview face-to-face at the



(a) Arrangement of objects by U12 for Task 1 on the table.

(b) Arrangement of objects by U4 for Task 3 on the table.

Figure 3: Layout of the user's table during tasks. Note: The names and addresses on the letters and credit cards were obscured by the authors.

non-profit agency. Simultaneously, another researcher, who joined remotely, conducted the interview with the volunteer over Zoom. The interviews lasted between 10 to 15 minutes.

Part 5: data analysis. After the participants' consent, we recorded all sessions, including focus groups and interviews. The first author manually transcribed the recorded data and analyzed the transcripts. The analysis was conducted through an iterative coding process involving initial coding, identification of new concepts, and categorization into themes and sub-themes [24]. All authors reviewed the concepts, themes, and sub-themes in weekly research meetings to finalize the codebook. Next, we present our findings.

5 FINDINGS

Privacy is missing in regular RSA interactions, with users disengaged or even forced to compromise it. They angle the camera in accordance with the volunteers' instructions to get help. With BubbleCam, however, users regain privacy and engage in managing it. We also revealed that both blind users and sighted volunteers participated in and were actively involved in the privacy-preserving RSA. They collaboratively contributed to the creation and maintenance of privacy during these interactions. This section analyzes participants' engagement in privacy, technology, and coordination.

5.1 Engagement in Privacy

Overall, 22 participants (U2-U12, V2-V12) appreciated the privacy benefits introduced by BubbleCam, ensuring they were showing or seeing only the intended content. By examining both parties' perspectives, we revealed that BubbleCam reduced users' embarrassment, boosted their self-esteem in blocking content, and enabled volunteers to comfortably avoid seeing unnecessary or inappropriate objects beyond the task at hand. Meanwhile, we found that BubbleCam preserved the primary function of RSA without compromising privacy: assisting users through visual cues from their cameras.

5.1.1 Users' Privacy in Blocking Content. In current RSA services, users' entire camera field is visible to assistants. This leads to privacy concerns, as eight participants (U1, U4, U7, U9, U12, V1, V3, V7) concurred that users *"really have to give up [privacy] in order to accept help"* (V7). They need to trust volunteers and follow instructions to *"pan... and look through everything"* (V3), which could potentially expose sensitive information. This risk increases in free RSA services with anonymous, not background-checked volunteers, raising concerns about information misuse or exploitation.

Compared to their previous RSA experiences, twenty-two participants (U2-U12, V2-V12) believed that BubbleCam enhances users' feelings of being *"protected"*, *"secured"*, *"safer"*, *"more relaxed"*, *"comfortable"*, and *"confident"*. It allows users to only share specific areas that need help while effectively obscuring irrelevant background information. With this capability, twelve participants (U2, U4, U6-U12, V2, V3, V9) thought that BubbleCam could eliminate their feelings of embarrassment by blocking what they were hesitant to share, thereby bolstering their self-esteem.

U7 pointed out that blind people are not necessarily organized, *"They have things all over the place. They put certain things that they're not aware of, or they may have forgotten"*. Consequently, users (U2, U4, U6-U12) may feel embarrassed when they are exposed to messy or private spaces. An example is U2, who apologizes to sighted assistants for unintentional exposure of the mess. After experimenting with BubbleCam, participants praised its potential to mitigate such awkwardness, elevating their confidence during RSA interactions. Specifically, BubbleCam is practical to obscure areas or objects they're not comfortable sharing and *"it's very, very effective, you know, covers up the things in the background"* (U7). This includes a range of areas and items such as garbage (U2), dirty dishes in the sink (U6, U10), unorganized fridge (U6), intimate settings like bedroom, bathroom, or toilet (U4, U10, U12), personal clothes like pajamas, bra, pants, socks (U4, U7, U8, U11), and even naked body parts when getting dressed (U9).

> *"If you're in your house, you may have things around you don't want anyone to see, like clothes or, you know, private clothes... Maybe your house is not clean. And, you know, you don't want them looking at that. So you could just have the bubble to where you want it, and I think that's good. I like that."* (U11)

Beyond their own spaces, a majority of users (except U1, U2, U6) valued protecting the privacy of bystanders, viewing it integral to their own privacy. They particularly consider the anonymity of

close social connections like families, friends, or roommates (U5-U7, U9-U12). This concern extends to public spaces, being mindful of other customers' and passengers' privacy (U3, U4, U7-U12). The commitment to privacy was reinforced by BubbleCam, solidifying users' confidence in using RSA service without infringing upon others' personal spaces or moments. As U5 put it, *"It's the same thing with privacy again. I have the option to not let that camera see anybody other than the people that I want it to see."* even if the bystanders inadvertently wander into the camera frame.

5.1.2 Volunteers' Comfort in Avoiding Unnecessary or Inappropriate Views. Six volunteers (V1, V3, V5-V7, V9) denoted that the absence of privacy preservation in regular RSA platforms often puts them in *"an awkward position"* (V7). It occurs when they encounter irrelevant or private content without knowing if users intend for them to see it. For example, V1 saw users' messy area, address, and account numbers; V5 felt *"a little bit invasive"* when users revealed too much; V6 felt *"uncomfortable"* and *"intruding"* upon viewing users' photos; and V7 glimpsed a user's bare leg. In such situations, it depends on volunteers to act *"trustworthy"* (V1) and *"good-naturedly"* (V7), driven by ethical considerations to either ignore, refrain from judgment, or not misuse such information. Therefore, volunteers hope for enhanced protective measures in the RSA paradigm to ensure they only view content that users intend to reveal.

BubbleCam fulfilled this requirement by enabling users to blur content they deem sensitive. It reassured volunteers that they only saw user-selected content intended for sharing. As a result, seven volunteers (V2, V5, V7-V10, V12) mentioned that they, as providers of assistance, also felt protected and comfortable not to see unnecessary or inappropriate content in privacy-preserving RSA interactions.

> *"I think I would feel more confident than usual because it's something that comes up every single time – the issue of privacy. And I really wanna make sure that I'm only seeing what they [users] want me to. So that'd be great."* (V7)

5.1.3 Preserving Utility Amidst Privacy. In the pursuit of preserving privacy, the effective obscuration of unintended objects with BubbleCam did not compromise the core function of the RSA prosthetic: assisting users via visual cues from their cameras. This suggests that BubbleCam retains its utility while upholding privacy.

Users reacted positively to the capability of seeking help from volunteers while simultaneously maintaining control over what to share. It not only enhances users' privacy and security in RSA but also boosts their confidence in seeking assistance from volunteers. Furthermore, three volunteers (V5, V9, V12) confirmed that BubbleCam was unobtrusive and did not impede their ability to assist users. This highlights BubbleCam's efficacy: it offers protection for both parties and maintains the functionality of synchronous RSA at a level consistent with existing platforms.

> *"A lot of times, blind and visually impaired people are very nervous about sensitive information... it's very easy for someone who's blind or visually impaired to be taken advantage of or manipulated because they can't see and someone else can. So I think that a really great feature of the app is the security [of] knowing that I'm getting the help that I am asking for, while also not jeopardizing any of my information."* (U6)

> *"... We are only interested in helping them [users] to find whatever they need. Adding the bubble is a way to help them protect their private information, and it's also protecting us as well for something we don't want to see. It's just a grey bubble. It's not really anything that bothers me."* (V5)

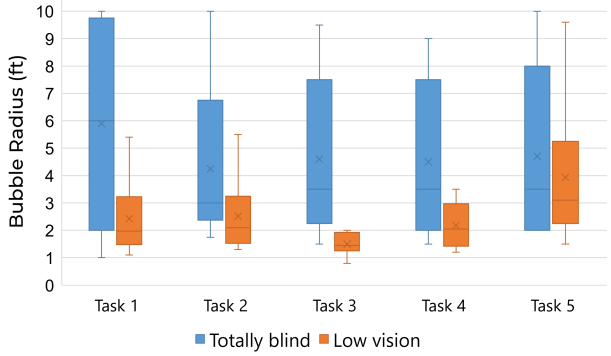


Figure 4: Comparison of bubble radius adopted by the totally blind and low vision users in the five tasks.

In summary, BubbleCam restores a sense of privacy that was notably absent in regular RSA interactions. Users are empowered to protect the privacy of their own spaces by not exposing messy areas and to respect the privacy of bystanders. For volunteers, BubbleCam enables them to avoid seeing unnecessary or inappropriate content, enhancing their comfort. Importantly, while upholding privacy, BubbleCam maintains the utility of synchronous RSA, allowing volunteers to assist users through visual cues.

5.2 Engagement with Technology

To investigate the choice of bubble radius by visually impaired users across five tasks, we recorded the shared BubbleCam screen during all sessions and meticulously documented the specific bubble radii employed by each user for each task by thoroughly reviewing the video footage. In instances where a user utilized multiple radii within a single task, we conducted a weighted average calculation based on their respective durations. Since U1 only completed Tasks 1 and 2, we decided to omit U1's data point in the quantitative analysis.

The partial perception of low-vision users might play a role in their bubble radius selection, whereas totally blind users rely relatively more on remote volunteers. Hence, we compared the bubble radius data between these two groups of individuals. As illustrated in Fig. 4, we have employed a box and whisker plot to visualize the bubble radius choices made by both totally blind and low-vision users across the five tasks. The key findings from Fig. 4 are as follows:

- Low-vision users tend to choose a smaller radius, with the medians being less than 4 ft in all tasks. Conversely, the radius mediums for totally blind users exceed 4 ft in all tasks.

- Totally blind users adopt a wider range of radii, while low-vision users' choices are more concentrated.
- The disparity in radius selection between totally blind and low-vision users is more pronounced in certain tasks (e.g., Task 3), but less significant in others (e.g., Task 5).

However, only 2 users (U2, U9) employed the partially hidden mode throughout the experiments. None of the users switched modes during the experiments, suggesting that this feature might have a limited impact on their overall experience or perceived privacy benefits.

Next, we delve into participants' feedback on their engagement with radius and blurring functions within BubbleCam.

5.2.1 Engagement with Radius. Twenty-one participants (U2-U12, V2-V8, V10-V12) emphasized that BubbleCam improved users' autonomy by allowing them to engage with radius and exert further control over the extent of what is shared with volunteers. This reshaped their RSA experiences because control is often absent in regular RSA interactions.

Varying on vision conditions, users employed different strategies to interact with the slider on BubbleCam, a feature that allowed users to dynamically adjust the radius of the bubble in real time. Low-vision users, often being more aware of privacy leakage, were more inclined to use the slider and maneuver it carefully to align the bubble with the contours of the objects they were showing. In contrast, blind users, often being less aware of the content they were showing and less adept at distance measurement, might unintentionally set the bubble too large and reveal more than necessary alongside the foreground objects. This observation elaborated on the quantitative findings that low-vision users opted for a smaller radius and a narrower range of radii, while totally blind users chose a larger radius and a broader range of radii. This difference was evident by U12: *"I think, because I have vision, I'm more aware of what I'm hiding and like, why I would want to hide it. Whereas, maybe someone who's totally blind, they may be more focused, actually trying to get the help that they need, as opposed to caring about everything else."*

As users maneuvered the slider, they had the immediacy to fine-tune the radius and the extent of visibility, affording them the flexibility to reveal or conceal parts of their surroundings in line with their comfort and discretion. This empowerment went beyond simple functionality. It offered users the opportunity to shape their own experience during RSA interactions. Participants acknowledged this autonomy, describing the slider feature as embodying *"customality"* (U4) and *"availability"* (U5). In contrast, regular RSA interactions often require users to relinquish this autonomy, as users have to follow volunteers' instructions and expose their surroundings to get assistance.

> *"Well, what I like most is the fact that I can block out what I don't want people to see. I have that availability to go from 0 to 10 [ft]. So, the wider the scope, the more people see; the smaller, the less they see... The person can see what you allowed him to see, no more, no less."* (U5)

Meanwhile, users appreciated the awareness that *"they [volunteers] can't adjust [the radius] to see more than what I want them to see"* (U2), highlighting the BubbleCam's commitment to upholding users' autonomy. The decision rested with the users: *"hide or not to hide"* (U10). Volunteers, on the other hand, could not manipulate or control BubbleCam and were specifically unable to swipe the slider. Knowing that they were the only ones to decide settings on BubbleCam, without external adjustments or interventions, strengthened users' sense of autonomy. This awareness motivated users to feel *"more confident and more comfortable"* when engaging with BubbleCam.

> *"I think it's a wonderful idea because more protection goes [with] more privacy. And it makes you more comfortable while using the app because you know that you're the one [who has] control. No one can change it. No one can move it. It makes you even more confident and more comfortable when using the app."* (U7)

5.2.2 Engagement with Blurring Feature. Volunteers were split on the blurring feature operating outside the bubble's coverage, where things inside the bubble remain visible and anything outside is obscured. Eight participants (U4, U7, U10, V2, V3, V5, V9, V12) indicated that the blurring feature did not hinder volunteers' performance. In fact, some noted it helped volunteers concentrate better on task-relevant objects. V3 highlighted that the blurring feature minimized distractions by focusing on what the user aimed to share, without the need to instruct them to adjust the camera angle.

> *"I think you're more focused on what you're doing. And there are not as many aspects to get you kind of distracted, like, you know, 'Oh, well, let's go check over there or something like that', and you have to think about all the directions, like let's move forward. So it's just something that's right in front of you."* (V3)

However, three volunteers (V1, V6, V9) were concerned that the blurring feature distracted them from focusing on the task at hand. V1 elaborated that the contour of the blurred area, defined by the bubble's edge, was highly sensitive to users' micro-movements. Although U1 attempted to hold the iPad steady, the unintended movements, even the tiny ones, led to continuous adjustments of the blurred area's boundaries. This constant change disrupted the clarity of the video and made it challenging for volunteers to accurately assess and focus on the task. This was particularly challenging for V1, as these disruptions exacerbated his neurodivergent condition.

> *"Your hand, as much as you're trying to keep it steady, is still moving just a little bit, right? And so it continues to make that bubble and the focus, everything kind of shifts and morphs and changes. And it's so difficult to get a nice, clear, you know, image of what it is supposed to assess."* (V1)

5.2.3 Engagement with Width. Participants pointed out that BubbleCam predominantly relies on distance, thus it is challenging to conceal *"side objects"* (V4) located closely and at the same distance as the foreground object that users intend to share. V5 illustrated this challenge with an example: *"When two documents are placed in parallel, it's difficult to hide the information for one of the documents because the distance is technically the same."*

To address this issue, nine participants (U3, U4, U6, U12, V3-V5, V7, V10) proposed the idea of shifting the bubble *"side to side"* (U6) or *"left to right"* (U6, U12, V3, V10) to adjust *"the wideness of the picture the volunteer can see"* (V5). V3 and V5 envisioned this as blocking side objects so that volunteers could *"only see exactly in the center"* (V3).

> *"Maybe the option to do left-to-right blocking. And so, maybe like setting up [a] parameter on how wide you can show. I think that would be helpful for left to right, especially [for] folks with even lower vision. Maybe they may not notice that they're showing documents on the left and right."* (U12)

V10 referred to this feature as *"image segmentation"*. Expanding on this, she suggested a feature of manually *"adjusting the location of the bubble on the screen"*. This would empower users to adjust the bubble directly with a tap on their device, effectively concealing side objects.

In summary, BubbleCam enabled users to engage with the bubble's radius via a slider, enhancing autonomy by granting them exclusive control over the extent to share with volunteers. Volunteers were divided over the blurring feature operating outside the bubble's coverage: some found it beneficial for focusing on task-relevant objects, whereas others were distracted by its constant adjustments. Participants also proposed a feature for adjusting the bubble's width to block side objects.

5.3 Engagement between Users and Volunteers

We observed a cooperative approach to privacy preservation between users and volunteers, where volunteers proactively intervened to remind users of privacy breaches and suggested radius reductions, and users adjusted accordingly. However, tension might arise when considering increasing the radius. While users prefer to safeguard their privacy by restricting what they share, they sometimes may seek visual-based assistance in scenarios that require a broader visual scope.

5.3.1 Assisting User-Controlled Privacy Preservation. Regardless of users' various vision conditions, eleven volunteers (V2-V12) understood and *"completely respect[ed]"* (V4) user's decision to obscure parts of their surroundings. V7 explained the reason as *"[users] might not know all that is on their table and what they're showing."* Six participants (U7, U12, V1, V7, V8, V10) resonated that users, especially blind users, are often less aware of what they show or hide. Thus, users need help from volunteers to collaboratively manage privacy. In this collaboration effort, volunteers actively intervened when users inadvertently overextended the radius and revealed sensitive information. They informed users about visible items, warned of potential privacy breaches, and recommended reducing the radius. In such situations, users generally accepted and acted upon these suggestions, adjusting to a suitable radius with the help of the volunteers.

Although Section 5.2.1 indicates that low-vision users often use smaller radii, instances of inadvertence still occur, underscoring the need for cooperative privacy management. An example involves U10 and V10 during the task of reading a handwritten letter. U10 initially set the radius at 7.5 ft, within which V10 noticed background objects like a passport and bank statement alongside the letter. Recognizing the potential privacy concern, V10 immediately suggested adjustments: *"Can you slide the slider to the left?"* Upon reducing the radius to 6 ft, U10 asked, *"Should I go down?"*, to which V10 directed further reduction, *"Yeah, a little more, more. Could you go more?"* until only the letter was visible.

Furthermore, participants (V5, V6, V12) envisioned more ideas to enhance this collaborative effort and protect users' privacy. They proposed granting volunteers limited control within BubbleCam, allowing them to increase blocking but not to decrease it. As V12 articulated, *"It would be great that I also have a certain amount of control to some extent, that I can add up more blocking. But from my side, I'm not able to reduce the blocking."* Similarly, V5 suggested permitting volunteers to *"black out the entire screen"* in case they detect any leakage of sensitive information.

5.3.2 Tension between Users' and Volunteers' Perspectives. Although both parties cooperatively preserve users' privacy by reducing the radius, opinions on increasing it vary. On one hand, users hesitate to expand the radius due to privacy concerns. On the other hand, volunteers are responsible for seeing the details users need and require a broader view to ensure users' safety in certain scenarios. This divergence could lead to communication tension.

Eight users (U2, U5, U7-U12) preferred not sharing control over enlarging the bubble. U5 expressed reluctance to volunteers' suggestions for revealing more: *"It will depend on what I want him or her to see. So it may not work in their favor, but it has to work in mine"* (U5). Conversely, volunteers conveyed the need to see more of the users' physical surroundings in certain scenarios. V1 and V8 noted the possibility that users might not always be conscious of their shared content, requiring assistance in pinpointing and identifying objects. From volunteers' perspectives, a comprehensive view is essential for *"giv[ing] feedback on where to find something to select the correct thing"* (V1). If given a restricted view, V8 questioned, *"How can I help, right? I will just say what I see."* This highlights a potential communication tension – users aim to maintain privacy by sharing less, yet sometimes they are seeking visual-based assistance that necessitates a broader view.

Building on this, three volunteers (V1, V10, V11) stressed the need for comprehensive environmental awareness in emergencies or high-stakes navigational tasks. They believed that BubbleCam's limited view, in such situations, could hinder prompt hazard detection, potentially delaying reactions to danger. Taking navigation as an example, V1 and V10 highlighted the risk of users colliding with obstacles or pedestrians due to the restricted view.

> *"If I can't see it and I can't tell them about it, they're going to get themselves into trouble. Case in point, something like a low-hanging tree ranch. 'Hey, it was blurred out.' And then, all of a sudden, it smacks them in the face because... within a foot of them, that doesn't give me enough time to communicate it to them, nor does it give them enough time to react to it."* (V1)

Three volunteers (V3, V4, V11) emphasized the importance of negotiation when a broader view is required. They typically explained the situation, described the current view, and suggested possibly expanding the bubble. V3 specifically prioritized users' comfort and autonomy, advising *"you might have to increase the bubble for me if you feel comfortable or move your camera a different way."* This

again reflects volunteers' respect for users' decisions, as analyzed in Section 5.3.1.

5.3.3 Towards Automatic and Selective Privacy-Preservation. To better coordinate and ease tension between users and volunteers, eight participants (U4, U7, U12, V2, V4, V5, V7, V8) recommended enhancing the privacy-preserving nature of BubbleCam by automatically and selectively detecting and obscuring sensitive objects.

As for "automatic" detection, U12 and V5 elucidated that important documents, such as debit cards, birth certificates, or social security cards, typically have distinct shapes and sizes. They proposed integrating an automatic object detection system that recognizes and describes these items, and alerts users with warnings like *"This might be the document you don't want someone to see"* (U12), or asks *"Are you sure to proceed?"* (V5). This feature would partly take over volunteers' roles in identifying privacy risks and suggesting radius adjustments. It would thereby enhance users' awareness of what they show and streamline the coordination process.

Coordination could be further enhanced through "selective" object detection. As V12 remarked, *"different people have different levels of sense of privacy"*. This is evident in users' varied comfort with showing messy areas or bystanders (Section 5.1.1). Thus, V7 and V8 suggested an intelligent object detection mechanism in BubbleCam that adheres to user-defined preferences regarding which objects to show and which not.

In summary, both parties engaged in collaborative privacy protection, with volunteers alerting users to potential breaches and advising radius reduction, and users adjusting accordingly. However, tension arose over increasing the radius: users preferred to share less for privacy, while volunteers needed a broader view, especially in high-stakes tasks. Automatic and selective object detection systems proposed by participants hold promise in easing this coordination.

6 DISCUSSION AND DESIGN IMPLICATIONS

In this section, we will discuss the collaborative efforts of users and volunteers in maintaining privacy, explore the distinctions in interactions among users with different visual conditions when using BubbleCam, and implications regarding privacy and security associated with BubbleCam.

6.1 Co-producing Privacy on BubbleCam

In this study, we developed BubbleCam, a high-fidelity prototype for enhancing privacy within synchronous, video-based RSA. BubbleCam allows users to hide objects beyond a set distance from sighted volunteers, granting them increased control over their privacy. Of our participants, 22 users and volunteers appreciated the privacy enhancements that BubbleCam offered, ensuring they were showing or seeing only intended content. On the one hand, it alleviated users' embarrassment and boosted their autonomy in blocking private documents, messy areas, and bystanders. This is in contrast to users' previous RSA experiences, where privacy was often compromised in order to get help. On the other hand, BubbleCam helped volunteers avoid seeing unnecessary or inappropriate objects beyond the task at hand.

We observed a cooperative approach in managing privacy with both users and volunteers involved. Unlike traditional privacy measures that rely on individual actions, such as hiding one's identity on social media [75], BubbleCam introduces an arrangement where privacy settings are dynamically configured through interactive engagement between two parties. Volunteers actively alerted users to potential privacy concerns and suggested reducing the BubbleCam radius. Meanwhile, users took these recommendations into account and made the appropriate adjustments to the radius, often with the aid of the volunteers. This shared responsibility for privacy preservation created an interactive and engaging experience that went beyond the traditional, individualized approach to maintaining privacy.

This collaborative paradigm is also distinct from prior cooperative privacy strategies employed by PVI with their trusted allies. Before, they used to collaborate with friends, family members, or professional helpers (e.g., paid mobility trainers) to protect their privacy and security, while excluding strangers (e.g., Aira agents or Be My Eyes volunteers) [41]. Hayes et al. [41] highlighted that seeking help from strangers could in turn introduce new privacy risks. In this study, we explored how BubbleCam can cultivate this cooperative dynamic between PVI and strangers, specifically Be My Eyes volunteers. Our findings underscored that a co-production in managing privacy between PVI and Be My Eyes volunteers is indeed achievable, with both parties working together to fine-tune privacy settings to ensure PVI's privacy.

Although both parties co-productively preserve users' privacy by reducing the radius, they may have communication tension due to different perspectives. Users hesitate to increase visibility due to privacy concerns. Conversely, volunteers are responsible for seeing details that users need and providing detailed guidance. To fulfill this role, they need a broader view, especially when ensuring users' physical safety. This divergence in needs and goals can pose challenges, especially during high-stake tasks. Volunteers were concerned that BubbleCam might impede their ability to detect potential hazards, thus reducing their reaction time in the face of danger, like avoiding obstacles or pedestrians in navigation. Though our findings suggest that co-production in privacy management is feasible, it may not necessarily align with considerations of users' physical safety. This intersection between privacy and safety thus requires further exploration.

6.2 Variation in Engagement Based on Visual Conditions

Users with varying visual conditions engaged distinctively with BubbleCam. Low-vision users often meticulously aligned the bubble with objects they intended to show, opting for a smaller radius and a more narrow range of radii. In contrast, blind users occasionally set the bubble too large, unintentionally revealing more than intended. This resonates with research indicating low-vision users tend to be more concerned than totally blind users about disclosing sensitive information [12, 15]. The distinctions can be attributed to users' respective visual conditions. Low-vision users, having some residual sight, are generally more conscious of distance estimation and the content they display, making them more inclined to use the privacy feature. Conversely, blind users, with limited awareness of

the content they present and distance measurements, are thus less likely to utilize the privacy feature independently without volunteer intervention.

There remains a gap between the availability of privacy-preservation tools (e.g., BubbleCam) and their underutilization by blind users. In this study, the engagement from volunteers bridged this gap through cooperative privacy preservation. Volunteers respected users' decisions to hide certain parts of their environment. They proactively intervened when users accidentally expanded the radius more than necessary and assisted them in maintaining their privacy.

Moreover, this gap can be addressed through computer vision technologies, such as object detection [51, 67] and scene text reading [33, 47, 71]. Important documents like credit cards, birth certificates, or social security cards usually have identifiable features like unique shapes, sizes, or titles. These attributes make them recognizable to computer vision systems. If the system detects these objects or specific texts in the video frame that haven't been obscured by BubbleCam, it can instantly notify the blind users [80]. Real-time notifications enable users to quickly employ voice commands, ensuring that sensitive content remains hidden. This integration of technology not only safeguards privacy but also empowers users, particularly those who are blind, with enhanced control over their visual data.

Future privacy-preserving mechanisms for blind users may face the same challenge: they are less aware of and less inclined to use privacy features. Therefore, ongoing research is crucial to increase blind users' awareness of privacy breaches and engagement in privacy management, ensuring the mechanisms are accessible and effective for their needs.

6.3 Design Implications in Privacy and Security

We now discuss the implications of our findings in the broader field of privacy and security research, particularly focusing on contextual integrity [65], differential privacy [30, 31], and user-centered privacy [50, 78].

6.3.1 Contextual Integrity. The Contextual Integrity theory [65] provides a practical framework for evaluating privacy across different social contexts. Rather than solely relying on the principle of informed consent or data protection, the theory argues that societal norms, known as "context-relative informational norms," govern privacy expectations by regulating information flow within specific contexts.

In any given context, informational norms involve three parameters. First, the *actors*, which represent the sender, recipient, and subject of the information. Second, the *attributes*, which define the type of information shared. Third, the *transmission principles*, which set the conditions under which sharing occurs. The theory posits that a *breach* of privacy occurs when these established norms are violated, even if consent is secured or the data is safeguarded. Such disruptions often trigger discomfort or concern, as they deviate from expected norms of information flow within a particular social context.

Examining our findings through the lens of Contextual Integrity Theory offers valuable insights. For example, in the RSA paradigm, two kinds of actors emerge: blind users as the *senders* and remote sighted assistants as the *recipients*. The *subject of the information* can be the senders themselves, disclosing their personal, intimate details (i.e., attributes) such as IDs, prescriptions, or body parts; or others like friends, family, or bystanders, sharing faces, body parts, locations, and social situations (attributes).

Upon closer inspection, another actor arises—the RSA service providers (e.g., Aira or Be My Eyes)—who store live camera feeds along with conversations between users and assistants. The terms of these service providers differ [73]. Aira discloses its data retention period for general personal data, does not share personal visual data with third parties, and allows opting out of general personal data collection. In contrast, Be My Eyes does not specify its data retention length, explicitly mentions sharing personal visual data with third parties for potential monetization, and does not offer opt-out options for general personal data collection. In fact, Be My Eyes has partnered with OpenAI to allow access to stored data for AI model training [5]. Using stored data for training may align with acceptable norms, but selling data to third parties raises significant privacy concerns for users, assistants, and bystanders alike.

Our bubble-based camera intervention can safeguard users' visual context and bystanders to a certain extent from unexpected actions of assistants (e.g., disclosing information without consent) or service providers (e.g., selling data to third parties). However, our method provides no defense against natural audio conversations during a session, underscoring a direction for future research to enhance privacy and contextual integrity in RSA services.

6.3.2 Differential Privacy. Our intervention draws parallels with the differential privacy framework [30, 31]. This framework provides formal privacy guarantees by adding calibrated noise to the input data. Three primary parameters stand out. First is the *privacy budget*, denoted by ϵ, which sets an upper limit on permissible privacy loss. Second is the *noise function*, commonly Gaussian or Laplacian, responsible for injecting randomness into the data. The third is the *sensitivity* of a query, measuring the maximum impact a single data point could have on the output. A *privacy violation* arises when the privacy budget ϵ is surpassed or when the noise function inadequately masks individual data points. Such violations can evoke concern, signaling a deviation from the framework's formal privacy assurances.

The privacy parameter ϵ is pivotal in differential privacy. Lower ϵ values yield stronger privacy safeguards but introduce more noise, making the data less useful. Conversely, higher ϵ values compromise privacy by keeping more original information as-is.

In our method, the user-selected distance band serves a role akin to ϵ and the noise function in differential privacy—the shorter the distance, the greater the noise, rendering the raw input video less useful to adversaries who gain unauthorized access to the video feed during or after a session. Therefore, our technique could provide a human-centered approach for tuning privacy parameters in established frameworks like differential privacy.

6.3.3 User-Centered Privacy. We observed distinct preferences among users regarding the type of content they were reluctant to share with volunteers (Section 5.1.1). For example, nine users expressed concerns about revealing messy areas, whereas others (U1, U3, U5) didn't share the same concern. Similarly, nine users valued protecting the privacy of bystanders, in contrast to others (U1, U2, U6).

These findings align with previous research suggesting that privacy is a nuanced and normative concept, with individual preferences varying widely [79].

Privacy preferences are also influenced by context [64]. While users might share concerns about certain content, the degree of concern can differ based on the situation. Among the nine users concerned about the privacy of bystanders who unintentionally wander into camera feeds, seven users (U5-U7, U9-U12) were particularly attentive to those in their close social connections, such as their families, friends, or roommates. However, they may be less concerned about the privacy of bystanders outside these intimate spaces, such as other customers or passengers.

Therefore, recognizing and accommodating individual privacy preferences is essential [77]. It underscores the importance of integrating user-centered principles in privacy research and design, aiming to help users achieve a privacy level relative to their own desires [50, 78]. Participants recognized these differences and suggested the integration of a "selective" object detection feature. This would allow users to input their specific privacy preferences as text within the app, thereby ensuring appropriate content is blocked (Section 5.3.3). This caters to diverse privacy preferences within a single system, e.g., BubbleCam. An advanced approach is "user-tailored privacy" [50, 78]. This method offers automated default settings personalized to users' unique preferences. Users could preemptively specify their privacy concerns within their profiles, including the items, areas, or contexts. An algorithm, connected to this profile, can then use these individual choices as a basis for selective blocking.

6.4 Limitations and Future Work

This study has several limitations. First, although participants engaged with the prototype in a realistic situation, their interaction time was brief. This short duration resulted in a limited experience with BubbleCam, which in turn led to the second limitation: none of the users changed the mode during the session. The option to fully or partially hide content in BubbleCam is a design feature that requires more systematic exploration. Third, we examined a small set of scenarios, tasks, and sensitive or private information. Future investigations should delve into broader contexts, such as outdoor environments, and explore a more varied range of private information. This could span from the names of books to street signs, personal attire, or tattoos.

7 CONCLUSION

We present BubbleCam, a tool designed to enhance privacy for PVI during synchronous RSA. Utilizing depth estimation from iOS devices' LiDAR scanner, BubbleCam envelops PVI within a virtual "bubble" while concealing objects outside of it. This innovation, a first for real-time privacy preservation in synchronous RSA, led to success in an exploratory field study. Twenty-two out of twenty-four participants were satisfied with the privacy protection afforded by BubbleCam. Users reported reduced embarrassment and increased autonomy, controlling the visibility of private documents, messy areas, and bystanders. Volunteers also felt at ease avoiding unnecessary or inappropriate objects while maintaining the utility of RSA. Besides, BubbleCam fostered a collaborative effort in privacy protection, transforming the typically individual task of maintaining privacy into an interactive, engaging experience.

ACKNOWLEDGMENTS

We thank *Lighthouse Guild, North Central Sight Services Inc., Massachusetts Association for the Blind and Visually Impaired* for their cooperation in this work. We also thank the anonymous reviewers for their insightful comments. This research was supported by the US National Institutes of Health, and the National Library of Medicine (R01 LM013330).

REFERENCES

[1] 2023. 100 Ways to Use Be My Eyes. https://support.bemyeyes.com/hc/en-us/articles/360005528317-100-Ways-to-Use-Be-My-Eyes.
[2] 2023. Aira. https://aira.io/.
[3] 2023. ARKit 6 - Augmented Reality - Apple Developer. Retrieved September 12, 2023 from https://developer.apple.com/augmented-reality/arkit/
[4] 2023. Be My Eyes - See the world together. https://www.bemyeyes.com/.
[5] 2023. Be My Eyes uses GPT-4 to transform visual accessibility. Retrieved September 12, 2023 from https://openai.com/customer-stories/be-my-eyes
[6] 2023. Getting Started with Be My Eyes. https://support.bemyeyes.com/hc/en-us/articles/360005536558-Getting-Started-with-Be-My-Eyes.
[7] 2023. Metal Overview - Apple Developer. Retrieved September 12, 2023 from https://developer.apple.com/metal/
[8] 2023. What to Expect from Be My Eyes Calls. https://support.bemyeyes.com/hc/en-us/articles/360005542717-What-to-Expect-from-Be-My-Eyes-Calls.
[9] 2023. Zoom: Video Conferencing, Cloud Phone, Webinars, Chat, Virtual Events. https://zoom.us/.
[10] Ali Abdolrahmani, Ravi Kuber, and Amy Hurst. 2016. An empirical investigation of the situationally-induced impairments experienced by blind mobile device users. In *Proceedings of the 13th International Web for All Conference*. 1–8.
[11] Dustin Adams and Sri Kurniawan. 2014. A blind-friendly photography application for smartphones. *ACM SIGACCESS Accessibility and Computing* 108 (2014), 12–15.
[12] Tousif Ahmed, Roberto Hoyle, Kay Connelly, David Crandall, and Apu Kapadia. 2015. Privacy Concerns and Behaviors of People with Visual Impairments. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI'15)*. 3523–3532.
[13] Tousif Ahmed, Patrick Shaffer, Kay Connelly, David Crandall, and Apu Kapadia. 2016. Addressing physical safety, security, and privacy for people with visual impairments. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*. 341–354.
[14] Taslima Akter, Tousif Ahmed, Apu Kapadia, and Swami Manohar Swaminathan. 2020. Privacy considerations of the visually impaired with camera based assistive technologies: Misrepresentation, impropriety, and fairness. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–14.
[15] Taslima Akter, Bryan Dosono, Tousif Ahmed, Apu Kapadia, and Bryan Semaan. 2020. " I am uncomfortable sharing what I can't see": Privacy Concerns of the Visually Impaired with Camera Based Assistive Applications. In *29th USENIX Security Symposium (USENIX Security 20)*. 1929–1948.
[16] Moudhi Ali Almuzaini and M Abdullah-Al-Wadud. 2018. A review on medication identification techniques for visually impaired patients. In *2018 21st Saudi Computer Society National Computer Conference (NCC)*. IEEE, 1–6.
[17] Mauro Avila, Katrin Wolf, Anke Brock, and Niels Henze. 2016. Remote assistance for blind users in daily life: A survey about Be My Eyes. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. 1–2.
[18] Shiri Azenkot, Sanjana Prasain, Alan Borning, Emily Fortuna, Richard E Ladner, and Jacob O Wobbrock. 2011. Enhancing independence and safety for blind and deaf-blind public transit riders. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*. 3247–3256.
[19] Przemyslaw Baranski and Pawel Strumillo. 2015. Field trials of a teleassistance system for the visually impaired. In *2015 8th International Conference on Human System Interaction (HSI)*. IEEE, 173–179.
[20] Jeffrey P Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, et al. 2010. Vizwiz: nearly real-time answers to visual questions. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. 333–342.
[21] Jeffrey P Bigham, Chandrika Jayant, Andrew Miller, Brandyn White, and Tom Yeh. 2010. VizWiz:: LocateIt-enabling blind people to locate objects in their

environment. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*. IEEE, 65–72.

[22] Erin Brady, Meredith Ringel Morris, Yu Zhong, Samuel White, and Jeffrey P Bigham. 2013. Visual challenges in the everyday lives of blind people. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2117–2126.

[23] Stacy M Branham and Shaun K Kane. 2015. Collaborative accessibility: How blind and sighted companions co-create accessible home spaces. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2373–2382.

[24] A. Bryman and R.G. Burgess. 1994. *Analyzing Qualitative Data*. Routledge. https://books.google.com/books?id=KQkotSd9YWkC

[25] M Bujacz, P Baranski, M Moranski, P Strumillo, and A Materka. 2008. Remote guidance for the blind—A proposed teleassistance system and navigation trials. In *2008 Conference on Human System Interactions*. IEEE, 888–892.

[26] Michele A Burton, Erin Brady, Robin Brewer, Callie Neylan, Jeffrey P Bigham, and Amy Hurst. 2012. Crowdsourcing subjective fashion advice using VizWiz: challenges and opportunities. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 135–142.

[27] John M. Carroll, Sooyeon Lee, Madison Reddie, Jordan Beck, and Mary Beth Rosson. 2020. Human-Computer Synergies in Prosthetic Interactions. *IxD&A* 44 (2020), 29–52. http://www.mifav.uniroma2.it/inevent/events/idea2010/doc/44_2.pdf

[28] Brendan Cassidy, Gilbert Cockton, and Lynne Coventry. 2013. A haptic ATM interface to assist visually impaired users. In *Proceedings of the 15th international ACM SIGACCESS conference on computers and accessibility*. 1–8.

[29] Babar Chaudary, Iikka Paajala, Eliud Keino, and Petri Pulli. 2017. Tele-guidance based navigation system for the visually impaired and blind persons. In *eHealth 360*. Springer, 9–16.

[30] Cynthia Dwork. 2006. Differential privacy. In *International colloquium on automata, languages, and programming*. Springer, 1–12.

[31] Cynthia Dwork. 2008. Differential privacy: A survey of results. In *International conference on theory and applications of models of computation*. Springer, 1–19.

[32] William Easley, Ravi Kuber, and A Ant Ozok. 2017. An empirical study examining medication management among individuals with visual impairments. *Universal Access in the Information Society* 16 (2017), 483–495.

[33] Nobuo Ezaki, Kimiyasu Kiyota, Bui Truong Minh, Marius Bulacu, and Lambert Schomaker. 2005. Improved text-detection methods for a camera-based text reading system for blind persons. In *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*. IEEE, 257–261.

[34] Vanja Garaj, Rommanee Jirawimut, Piotr Ptasinski, Franjo Cecelja, and Wamadeva Balachandran. 2003. A system for remote sighted guidance of visually impaired pedestrians. *British Journal of Visual Impairment* 21, 2 (2003), 55–63.

[35] Danna Gurari, Qing Li, Chi Lin, Yinan Zhao, Anhong Guo, Abigale Stangl, and Jeffrey P Bigham. 2019. Vizwiz-priv: A dataset for recognizing the presence and purpose of private visual information in images taken by blind people. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 939–948.

[36] Danna Gurari, Qing Li, Abigale J Stangl, Anhong Guo, Chi Lin, Kristen Grauman, Jiebo Luo, and Jeffrey P Bigham. 2018. Vizwiz grand challenge: Answering visual questions from blind people. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3608–3617.

[37] Danna Gurari, Yinan Zhao, Meng Zhang, and Nilavra Bhattacharya. 2020. Captioning images taken by people who are blind. In *European Conference on Computer Vision*. 417–434.

[38] Rakibul Hasan, David Crandall, Mario Fritz, and Apu Kapadia. 2020. Automatically detecting bystanders in photos to reduce privacy risks. In *IEEE Symposium on Security and Privacy (SP)*. 318–335.

[39] Rakibul Hasan, Eman Hassan, Yifang Li, Kelly Caine, David J Crandall, Roberto Hoyle, and Apu Kapadia. 2018. Viewer experience of obscuring scene elements in photos to enhance privacy. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–13.

[40] Rakibul Hasan, Patrick Shaffer, David Crandall, Eman T Apu Kapadia, et al. 2017. Cartooning for enhanced privacy in lifelogging and streaming videos. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 29–38.

[41] Jordan Hayes, Smirity Kaushik, Charlotte Emily Price, and Yang Wang. 2019. Cooperative privacy and security: Learning from people with visual impairments and their allies. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*. 1–20.

[42] Nicole Holmes and Kelly Prentice. 2015. iPhone video link facetime as an orientation tool: remote O&M for people with vision impairment. *International Journal of Orientation & Mobility* 7, 1 (2015), 60–68.

[43] Ziad Hunaiti, Vanja Garaj, and Wamadeva Balachandran. 2006. A remote vision guidance system for visually impaired pedestrians. *The Journal of Navigation* 59, 3 (2006), 497–504.

[44] Fethi A Inan, Akbar S Namin, Rona L Pogrund, and Keith S Jones. 2016. Internet use and cybersecurity concerns of individuals with visual impairments. *Journal of Educational Technology & Society* 19, 1 (2016), 28–40.

[45] Chandrika Jayant, Hanjie Ji, Samuel White, and Jeffrey P Bigham. 2011. Supporting blind photography. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*. 203–210.

[46] Rie Kamikubo, Naoya Kato, Keita Higuchi, Ryo Yonetani, and Yoichi Sato. 2020. Support strategies for remote guides in assisting people with visual impairments for effective indoor navigation. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[47] Shaun K Kane, Brian Frey, and Jacob O Wobbrock. 2013. Access lens: a gesture-based screen reader for real-world documents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 347–350.

[48] Shaun K Kane, Chandrika Jayant, Jacob O Wobbrock, and Richard E Ladner. 2009. Freedom to roam: a study of mobile device adoption and accessibility for people with visual and motor disabilities. In *Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. 115–122.

[49] Mohamed Khamis, Habiba Farzand, Marija Mumm, and Karola Marky. 2022. DeepFakes for Privacy: Investigating the Effectiveness of State-of-the-Art Privacy-Enhancing Face Obfuscation Methods. In *Proceedings of the International Conference on Advanced Visual Interfaces*. 1–5.

[50] Bart Piet Knijnenburg. 2015. *A user-tailored approach to privacy decision support*. University of California, Irvine.

[51] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2017. Imagenet classification with deep convolutional neural networks. *Commun. ACM* 60, 6 (2017), 84–90.

[52] Aliasgar Kutiyanawala, Vladimir Kulyukin, and John Nicholson. 2011. Teleassistance in accessible shopping for the blind. In *Proceedings on the International Conference on Internet Computing (ICOMP)*. The Steering Committee of The World Congress in Computer Science, Computer …, 1.

[53] Walter S Lasecki, Kyle I Murray, Samuel White, Robert C Miller, and Jeffrey P Bigham. 2011. Real-time crowd control of existing interfaces. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 23–32.

[54] Walter S Lasecki, Phyo Thiha, Yu Zhong, Erin Brady, and Jeffrey P Bigham. 2013. Answering visual questions with conversational crowd assistants. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 18.

[55] Sooyeon Lee, Madison Reddie, Krish Gurdasani, Xiying Wang, Jordan Beck, Mary Beth Rosson, and John M. Carroll. 2018. Conversations for Vision: Remote Sighted Assistants Helping People with Visual Impairments. arXiv:1812.00148 [cs.HC]

[56] Sooyeon Lee, Madison Reddie, Chun-Hua Tsai, Jordan Beck, Mary Beth Rosson, and John M Carroll. 2020. The emerging professional practice of remote sighted assistance for people with visual impairments. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[57] Sooyeon Lee, Rui Yu, Jingyi Xie, Syed Masum Billah, and John M Carroll. 2022. Opportunities for human-AI collaboration in remote sighted assistance. In *27th International Conference on Intelligent User Interfaces*. 63–78.

[58] Fenghua Li, Zhe Sun, Ang Li, Ben Niu, Hui Li, and Guohong Cao. 2019. Hideme: Privacy-preserving photo sharing on social networks. In *IEEE Conference on Computer Communications (INFOCOM)*. 154–162.

[59] Franklin Mingzhe Li, Jamie Dorst, Peter Cederberg, and Patrick Carrington. 2021. Non-visual cooking: exploring practices and challenges of meal preparation by people with visual impairments. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–11.

[60] Yifang Li, Nishant Vishwamitra, Bart P Knijnenburg, Hongxin Hu, and Kelly Caine. 2017. Effectiveness and users' experience of obfuscation as a privacy-enhancing technology for sharing photos. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–24.

[61] David L Morgan. 1996. Focus groups. *Annual review of sociology* 22, 1 (1996), 129–152.

[62] Maia Naftali and Leah Findlater. 2014. Accessibility in context: understanding the truly mobile experience of smartphone users with motor impairments. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*. 209–216.

[63] Daniela Napoli, Khadija Baig, Sana Maqsood, and Sonia Chiasson. 2021. "I'm Literally Just Hoping This Will Work:" Obstacles Blocking the Online Security and Privacy of Users with Visual Disabilities. In *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*. 263–280.

[64] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.

[65] Helen Nissenbaum. 2020. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.

[66] Georg Regal, Elke Mattheiss, Marc Busch, and Manfred Tscheligi. 2016. Insights into internet privacy for visually impaired and blind people. In *Computers Helping People with Special Needs: 15th International Conference, ICCHP 2016, Linz, Austria, July 13-15, 2016, Proceedings, Part I 15*. Springer, 231–238.

[67] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2016. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence* 39, 6 (2016), 1137–1149.

[68] Stephanie Rosenbaum, Gilbert Cockton, Kara Coyne, Michael Muller, and Thyra Rauch. 2002. Focus groups in HCI: wealth of information or waste of resources?. In *CHI'02 extended abstracts on human factors in computing systems*. 702–703.

[69] Graig Sauer, Jonathan Holman, Jonathan Lazar, Harry Hochheiser, and Jinjuan Feng. 2010. Accessible privacy and security: a universally usable human-interaction proof tool. *Universal Access in the Information Society* 9 (2010), 239–248.

[70] Stefano Scheggi, A Talarico, and Domenico Prattichizzo. 2014. A remote guidance system for blind and visually impaired people via vibrotactile haptic feedback. In *22nd Mediterranean Conference on Control and Automation*. IEEE, 20–23.

[71] Roy Shilkrot, Jochen Huber, Connie Liu, Pattie Maes, and Suranga Chandima Nanayakkara. 2014. FingerReader: a wearable device to support text reading on the go. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 2359–2364.

[72] Daniel J Solove. 2005. A taxonomy of privacy. *U. Pa. l. Rev.* 154 (2005), 477.

[73] Abigale Stangl, Kristina Shiroma, Nathan Davis, Bo Xie, Kenneth R Fleischmann, Leah Findlater, and Danna Gurari. 2022. Privacy concerns for visual assistance technologies. *ACM Transactions on Accessible Computing (TACCESS)* 15, 2 (2022), 1–43.

[74] Abigale Stangl, Kristina Shiroma, Bo Xie, Kenneth R Fleischmann, and Danna Gurari. 2020. Visual content considered private by people who are blind. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–12.

[75] Frederic Stutzman and Woodrow Hartzog. 2012. Boundary regulation in social media. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 769–778.

[76] Tee-Ann Teo and Chen-Chia Yang. 2023. Evaluating the accuracy and quality of an iPad Pro's built-in lidar for 3D indoor mapping. *Developments in the Built Environment* 14 (2023).

[77] Daricia Wilkinson, Moses Namara, Karla Badillo-Urquiola, Pamela J Wisniewski, Bart P Knijnenburg, Xinru Page, Eran Toch, and Jen Romano-Bergstrom. 2018. Moving Beyond a" one-size fits all" Exploring Individual Differences in Privacy. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–8.

[78] Daricia Wilkinson, Saadhika Sivakumar, David Cherry, Bart P Knijnenburg, Elaine M Raybourn, Pamela Wisniewski, and Henry Sloan. 2017. User-tailored privacy by design. In *Proceedings of the Usable Security Mini Conference*.

[79] Pamela J Wisniewski and Xinru Page. 2022. Privacy theories and frameworks. In *Modern Socio-Technical Perspectives on Privacy*. Springer International Publishing Cham, 15–41.

[80] Jingyi Xie, Madison Reddie, Sooyeon Lee, Syed Masum Billah, Zihan Zhou, Chunhua Tsai, and John M Carroll. 2022. Iterative Design and Prototyping of Computer Vision Mediated Remote Sighted Assistance. *ACM Transactions on Computer-Human Interaction (TOCHI)* 29, 4 (2022), 1–40.

[81] Jingyi Xie, Rui Yu, Kaiming Cui, Sooyeon Lee, John M. Carroll, and Syed Masum Billah. 2023. Are Two Heads Better than One? Investigating Remote Sighted Assistance with Paired Volunteers. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS'23)*. 1810–1825.

[82] Jingyi Xie, Rui Yu, Sooyeon Lee, Yao Lyu, Syed Masum Billah, and John M Carroll. 2022. Helping Helpers: Supporting Volunteers in Remote Sighted Assistance with Augmented Reality Maps. In *Designing Interactive Systems Conference*. 881–897.

[83] Jian Xu, Syed Masum Billah, Roy Shilkrot, and Aruna Balasubramanian. 2019. DarkReader: bridging the gap between perception and reality of power consumption in smartphones for blind users. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*. 96–104.

[84] Xinyu Xu, Ahmad Al-Dahle, and Kshitiz Garg. 2020. Shared sensor data across sensor processing pipelines. US Patent 10,671,068.

[85] Hanlu Ye, Meethu Malu, Uran Oh, and Leah Findlater. 2014. Current and future mobile and wearable device use by people with visual impairments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3123–3132.

[86] Antonio Zea and Uwed Hanebeck. 2022. Modeling Spatial Uncertainty for the iPad Pro Depth Sensor. *Journal of Advances in Information Fusion* 17, 2 (2022).

[87] Zhuohao Jerry Zhang, Smirity Kaushik, JooYoung Seo, Haolin Yuan, Sauvik Das, Leah Findlater, Danna Gurari, Abigale Stangl, and Yang Wang. 2023. ImageAlly: A Human-AI Hybrid Approach to Support Blind People in Detecting and Redacting Private Image Content. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS)*. 417–436.

[88] Yuhang Zhao, Shaomei Wu, Lindsay Reynolds, and Shiri Azenkot. 2018. A face recognition application for people with visual impairments: Understanding use beyond the lab. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.