Understanding How People with Visual Impairments Take Selfies: Experiences and Challenges

Ricardo E. Gonzalez Penuela
Information Science, Cornell Tech, Cornell University, New York, New York, USA
reg258@cornell.edu

Paul Vermette
Cornell University, Ithaca, New York, USA
prv25@cornell.edu

Zihan Yan
Information Science, Cornell University, Ithaca, New York, USA
zihanyan@zju.edu.cn.

Cheng Zhang
Information Science, Cornell University, Ithaca, New York, USA
chengzhang@cornell.edu

Keith Vertanen
Computer Science, Michigan Technological University, Houghton, Michigan, USA
vertanen@mtu.edu

Shiri Azenkot
Jacobs Technion-Cornell Institute, Cornell Tech, Cornell University, New York, New York, USA
shiri.azenkot@cornell.edu

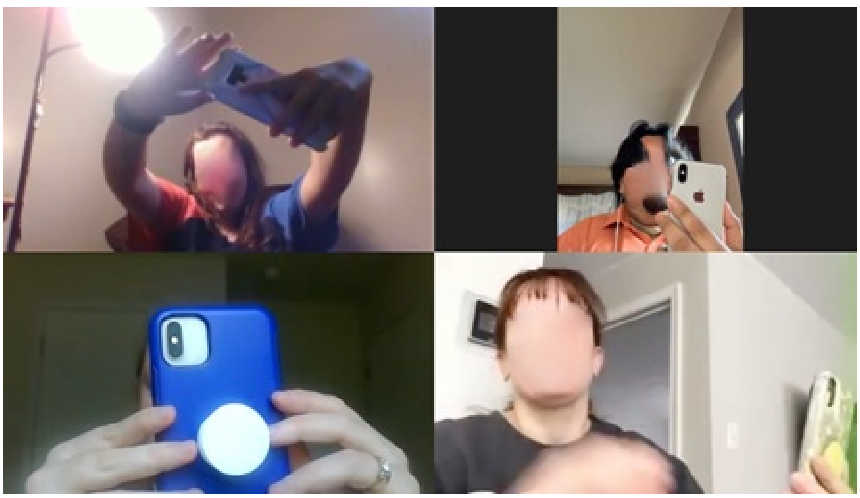

Figure 1: Participants demonstrating how they take a selfie with their smartphone (P3, P6, P5, and P9).

ABSTRACT

Selfies are a pervasive form of communication in social media. While there has been some work on systems that guide people with visual impairments (PVI) in taking photos, nearly all has focused on using the camera on the back of the device. We do not know whether and how PVI take selfies. The aim of our work is to understand (1) PVI selfie-taking experiences and challenges, (2) what information do PVI need when taking selfies, and (3) what modalities do PVI prefer (e.g., tactile, verbal, or non-verbal audio) to support selfie-taking. To address this gap, we conducted interviews with 10 PVI. Our findings show that current selfie-taking applications do not provide enough assistance to meet the needs of PVI. We contribute design guidelines that researchers and designers can implement for creating accessible selfie-taking applications.

ASSETS '22, October 23–26, 2022, Athens, Greece

ACM ISBN 978-1-4503-9258-7/22/10.
https://doi.org/10.1145/3517428.3550372

CCS CONCEPTS

• **Human-centered computing** → Accessibility; Empirical studies in accessibility; Human computer interaction (HCI); Empirical studies in HCI.

KEYWORDS

Selfies, Accessibility, Challenges Photo-Taking, Design Guidelines, Photography, PVI

ACM Reference Format:
Ricardo E. Gonzalez Penuela, Paul Vermette, Zihan Yan, Cheng Zhang, Keith Vertanen, and Shiri Azenkot. 2022. Understanding How People with Visual Impairments Take Selfies: Experiences and Challenges. In *The 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22), October 23–26, 2022, Athens, Greece.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3517428.3550372

1 INTRODUCTION & RELATED WORK

Like their sighted counterparts, people with visual impairments (PVI) take photos, share, and recall memories [2, 12, 19, 23, 26]. Moreover, for PVI, taking photos can be an important way to get information about their environment, whether it is from remote friends, the crowd [14-16, 20, 22] or computer vision services [18]. However, since they cannot see the photos, it is difficult for them to capture the target and ensure a photo is aesthetic and high quality (e.g., well-lit, and properly framed).

Researchers developed methods that provide audio and tactile guidance to assist PVI in taking photos [1-11]. For example, Jayant et al. developed Easy Snap [2] a mobile device application that helped PVI capture other people in a photo by providing verbal instructions, vibrations, and audio cues. Similarly, Cutter et al. [9] developed a method to help PVI capture a document. Their method provided the user precise verbal instructions specifying the direction and distance to move the device. This body of work focused on taking photos with the back camera (the camera on the back of the device) and ensuring that a certain target was captured. Furthermore, these systems have mainly focused on whether a PVI can take a "compliant" photo with the provided guidance (e.g., Whether they can capture a document, people, or a scene in an urban area). Only a small number of researchers have considered the task of taking selfies, which is typically more open ended and requires more detailed information to meet the user's needs.

Today, people frequently take selfies (i.e., photos of themselves). They share selfies to present themselves in a desired light (particularly in social media platforms). PVI want to participate in social media just like their sighted counterparts, taking and posting selfies as well as other types of photos [11, 13, 19, 21, 26], so it is important to ensure that selfie-taking is accessible.

Only one project has proposed an accessible selfie-taking application for people with visual impairments [24]. Yungjung et al. [24] designed SelfieHelper, an application that provided verbal feedback about the number of faces, sizes and their locations in the screen. Although this application addressed some challenges in the selfie-taking process (e.g., framing the user's head in the photo), participants mentioned that they desired more information (i.e., their physical appearance, or the background). Fang et al. [25] also designed an application that provides feedback throughout the selfie-taking process, but it was not intended for PVI. These applications did not provide holistic support for accessible selfie-taking—they did not provide enough information to meet PVI needs, or the feedback was not designed to be understood by PVI. Thus, at present, we still do not know what specific challenges PVI experience when taking selfies, what type of support can best assist them and what information they need to take a selfie.

To address this gap, we investigated the selfie-taking experiences of PVI. Our research questions were: (1) What are the current challenges that PVI face when taking selfies? (2) What information do PVI want when taking selfies? (3) What modalities do PVI prefer (e.g., tactile, verbal, audio non-verbal cues) in a hypothetical accessible selfie-taking application?

In this paper, we describe an interview study to address the research questions and contribute design guidelines for accessible selfie-taking applications. Building on prior work, our research takes one more step towards making photography fully accessible to PVI.

2 METHOD

We recruited 10 participants, all of whom had a visual impairment. Their ages ranged from 22 to 68 (mean=38, SD=14). All participants were based in the United States and were recruited through the National Federation of the Blind mailing list. Four identified as totally blind, five identified as legally blind, and one participant identified as low vision. Eight were female and two were male. All participants were iPhone users.

We conducted remote interviews that were approximately one-hour long via videoconferencing software. We first asked participants to recall instances where they needed to use their phone's front-facing camera. Next, we asked them about their selfie-taking behavior, e.g., "which applications do you use to take selfies?". Then, we asked participants to demonstrate how they take a selfie with their phone while they "thought out loud" [17]. Finally, we asked participants to provide feedback about a set of potential features and modes of interaction for an accessible selfie-taking application.

During the interviews, one of the researchers took notes of significant moments and relevant quotes. In addition, the interviews were videorecorded and transcribed. Two researchers coded two of the transcripts individually through an inductive coding process. Then, the two researchers went through the generated codes together until there was an agreement on a single set of codes. The resulting codebook was utilized by one of the researchers to code the rest of the transcripts.

3 FINDINGS

3.1 Selfie-Taking Experiences and Challenges

Participants described their weekly selfie-taking behavior. Five participants took selfies infrequently (1-4 every week), three frequently (5-9) and two very frequently (10+). Participants took selfies for various reasons: to communicate with friends (P3, P8), to take photos with their newborn (P6), for online dating (P1, P8, P9), to remember an occasion with their peers (P1, P5, P10), and to share how they looked (P4, P7).

All participants except P4 sent selfies to a sighted person for validation before sharing them. P4 had extensive experience taking selfies, and she felt that if "you do not care too much about how [the photo] looks" then she would send the photo without validation.

Many participants expressed desire but also frustration over taking selfies, and using frontal camera features to take augmented selfies, like face filters (P1, P4-10). Two participants mentioned using inaccessible camera applications in social media (e.g., Snapchat and Facebook Messenger) solely because they are forced by social circumstances (e.g., friends only use these platforms).

P5, who had a successful business with her online blog, described how the inability to take good selfies independently could affect her business:

> "I see when people post those types of things (selfies), you really get a feel for who they are and you get a feel for who they are as a person, in addition to who

Table 1: Participant demographics.

Participant ID	Gender	Age	Visual Impairment	Onset
1	Male	29	Totally blind	Birth
2	Female	44	Totally blind	1
3	Female	22	Legally blind	Birth
4	Female	24	Legally blind	Birth
5	Female	32	Legally Blind	Birth
6	Male	31	Legally Blind	11
7	Female	52	Totally blind	Birth
8	Female	38	Totally blind	Birth
9	Female	41	Low Vision.	37
10	Female	68	Legally blind	Birth

> they are as a business owner. And I feel like I'm kind of missing out on that because I can't express myself in that way".

P9 explained how she felt when she was creating a dating profile:

> "[Taking a selfie for the dating app] was stressful. I kept thinking **it is probably going to be a really weird shot** where I'm either cut off, where **it is going to look like I have a physical disability**."

3.2 Information Needs and Modality

Many participants pointed out how the lack of information in the guidance provided by the camera they typically use to take selfies (e.g., the iPhone camera) demotivated them entirely from taking a selfie (P1, P2, P5, P7, P8, P10).

P8 described how the feedback given by Apple's guidance system was not helpful: "It will say if a photo is bright or dark, but even photos that people send me when it's a good photo or they say it's a good photo, it will describe it as dark. So, I don't really have a good sense of what that means".

Participants desired to have more detailed and descriptive information about their physical appearance to share selfies confidently (P1, P2, P6, P7, P8). P1 provided an example of the kind of information he would like to receive: *"Bob is sitting at his desk in his new vinyl desk chair, and he is wearing a brown sweater with three buttons, blue jeans, and black socks."*

Six participants had a strong preference for human-like conversational guidance (P1, P4, P5, P7, P9, P10). Six participants cautioned that verbal information should strive to be natural and intuitive (P1, P2, P4, P5, P7, P9). They were especially concerned with verbal cues that may lead to cognitive overload (e.g., "turn the phone 45 degrees to the right") because (1) they would need to process that information and (2) such instructions would be harder to reliably follow.

Half of the participants also liked the idea of having auditory non-verbal guidance cues (P1, P4, P7, P8, P9). Many participants recalled some form of "hot-cold" tone guidance. This meant that the guidance tone would change its rhythm and volume to guide their movements. P8 provided an example: *"This is [in an application] for taking pictures of documents with the back camera to read them. It uses tonal guidance, it is louder and steadier when the image is clear, and softer when you are too far away or wavery when you are not steady. It is one of my favorite ways to use a camera."*

None of the participants were excited about using tactile cues to convey information. On the contrary, some participants discouraged the use of vibrations because they thought that it would be confusing or uncomfortable to have the phone vibrating while taking a photo (P1, P4). P4 elaborated on their reasoning: *"The only thing I can think of that your phone can do [with tactile cues] is buzz, and that would shake the camera."*

4 DESIGN GUIDELINES & FUTURE WORK

In prior studies, involving the back camera, researchers considered only one criterion for taking successful photos: the target object was captured in the photo [1–3, 9, 11]. In contrast, our findings showed that taking selfies was more complex, and involved three success criteria: PVI had to capture themselves, make sure they looked as desired, and not capture anything unwanted in the background. Based on this observation and our findings, we present design guidelines for accessible selfie-taking applications:

•**Tailor the system guidance to the goal of the user and support all success criteria**. For example, P4 mentioned that sometimes she wanted to take a quick photo and they did not care how she looked, but other times when she shared the photo on social media, she cared about how she looked and the photo quality. This is consistent with prior findings [24]. Most participants wanted to know how they looked (e.g., their facial expressions and clothing). In the first scenario the iOS camera guidance provides enough information since it helps PVI center their face and snap a quick photo; but in the second scenario, PVI need more information about their physical appearance (e.g., their facial expression, whether their eyes are closed or open, clothing detected), about background information (e.g., disorganized table, floor detected) and the quality of the photo (e.g., whether the photo is blurry, whether the faces in the photo can be seen clearly).

•**Guide the user through human-like conversational** prompts and descriptions. As many participants mentioned, PVI are accustomed to being assisted by sighted peers when they take photos. Thus, conversational prompts and descriptions of photo content and quality would be natural and easy to follow.

•**Communicate low-level guidance with real-time feedback using non-verbal audio cues:** As mentioned by participants, PVI are familiar with systems that leverage the pitch, tone, or rhythm of sounds to convey a change of state in the system or progress of the task. This, combined with human-like conversational prompts, would provide PVI with the necessary information to make small adjustments to the camera position.

•**Include settings to filter information:** Consistent with the literature in other domains about accessible technology for PVI [27], we recommend providing an easy way to turn on or off information provided by the guidance system (e.g., turn on/off guidance about the photo's background or people's facial expressions). More generally, information should be modifiable with varying degrees of specificity to adapt to the needs and preferences of the user.

We hope that these guidelines will be used by researchers and designers to create accessible selfie-taking applications. In a future study, we will develop our own accessible selfie-taking application that PVI could use to share photos about their daily lives.

ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under awards IIS-1909248 and IIS-1909930. We thank all participants for their time and feedback.

REFERENCES

[1] Samuel White, Hanjie Ji, and Jeffrey P. Bigham. 2010. EasySnap: real-time audio feedback for blind photography. In Adjunct proceedings of the 23nd annual ACM symposium on User interface software and technology (UIST '10). Association for Computing Machinery, New York, NY, USA, 409–410. DOI:https://doi.org/10.1145/1866218.1866244

[2] 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 (ASSETS '11). Association for Computing Machinery, New York, NY, USA, 203–210. DOI:https://doi.org/10.1145/2049536.2049573

[3] Marynel Vázquez and Aaron Steinfeld. 2012. Helping visually impaired users properly aim a camera. In Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '12). Association for Computing Machinery, New York, NY, USA, 95–102. DOI:https://doi.org/10.1145/2384916.2384934

[4] Dustin Adams, Tory Gallagher, Alexander Ambard, and Sri Kurniawan. 2013. Interviewing blind photographers: design insights for a smartphone application. In Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13). Association for Computing Machinery, New York, NY, USA, Article 54, 1–2. DOI:https://doi.org/10.1145/2513383.2513418

[5] Marynel Vázquez and Aaron Steinfeld. 2014. An Assisted Photography Framework to Help Visually Impaired Users Properly Aim a Camera. ACM Trans. Comput.-Hum. Interact. 21, 5, Article 25 (November 2014), 29 pages. DOI:https://doi.org/10.1145/2651380

[6] Roberto Manduchi and James M. Coughlan. 2014. The last meter: blind visual guidance to a target. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). Association for Computing Machinery, New York, NY, USA, 3113–3122. DOI:https://doi.org/10.1145/2556288.2557328

[7] Jan Balata, Zdenek Mikovec, and Lukas Neoproud. 2015. BlindCamera: Central and Golden-ratio Composition for Blind Photographers. In Proceedings of the Mulitimedia, Interaction, Design and Innnovation (MIDI '15). Association for Computing Machinery, New York, NY, USA, Article 8, 1–8. DOI:https://doi.org/10.1145/2814464.2814472

[8] Minju Kim and Jungjin Lee. 2019. PicMe: Interactive Visual Guidance for Taking Requested Photo Composition. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, NY, USA, Paper 395, 1–12. DOI:https://doi.org/10.1145/3290605.3300625

[9] Michael P. Cutter and Roberto Manduchi. 2015. Towards Mobile OCR: How to Take a Good Picture of a Document Without Sight. In Proceedings of the 2015 ACM Symposium on Document Engineering (DocEng '15). Association for Computing Machinery, New York, NY, USA, 75–84. DOI:https://doi.org/10.1145/2682571.2797066

[10] Jongho Lim, Yongjae Yoo, Hanseul Cho, and Seungmoon Choi. 2019. TouchPhoto: Enabling Independent Picture Taking and Understanding for Visually-Impaired Users. In 2019 International Conference on Multimodal Interaction (ICMI '19). Association for Computing Machinery, New York, NY, USA, 124–134. DOI:https://doi.org/10.1145/3340555.3353728

[11] Dragan Ahmetovic, Daisuke Sato, Uran Oh, Tatsuya Ishihara, Kris Kitani, and Chieko Asakawa. 2020. ReCog: Supporting Blind People in Recognizing Personal Objects. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–12. DOI:https://doi.org/10.1145/3313831.3376143

[12] Cynthia L. Bennett, Jane E, Martez E. Mott, Edward Cutrell, and Meredith Ringel Morris. 2018. How Teens with Visual Impairments Take, Edit, and Share Photos on Social Media. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). Association for Computing Machinery, New York, NY, USA, Paper 76, 1–12. DOI:https://doi.org/10.1145/3173574.3173650

[13] Dustin Adams, Sri Kurniawan, Cynthia Herrera, Veronica Kang, and Natalie Friedman. 2016. Blind Photographers and VizSnap: A Long-Term Study. In Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '16). Association for Computing Machinery, New York, NY, USA, 201–208. DOI:https://doi.org/10.1145/2982142.2982169

[14] Jeffrey P. Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C. Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samual White, and Tom Yeh. 2010. VizWiz: nearly real-time answers to visual questions. In Proceedings of the 23nd annual ACM symposium on User interface software and technology (UIST '10). Association for Computing Machinery, New York, NY, USA, 333–342. DOI:https://doi.org/10.1145/1866029.1866080

[15] J. P. Bigham, C. Jayant, A. Miller, B. White and T. Yeh, "VizWiz::LocateIt - enabling blind people to locate objects in their environment," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2010, pp. 65-72, doi: 10.1109/CVPRW.2010.5543821.

[16] 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 (ASSETS '12). Association for Computing Machinery, New York, NY, USA, 135–142. DOI:https://doi.org/10.1145/2384916.2384941

[17] Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. the MIT Press.

[18] 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 (CHI '18). Association for Computing Machinery, New York, NY, USA, Paper 215, 1–14. DOI:https://doi.org/10.1145/3173574.3173789

[19] Violeta Voykinska, Shiri Azenkot, Shaomei Wu, and Gilly Leshed. 2016. How Blind People Interact with Visual Content on Social Networking Services. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16). Association for Computing Machinery, New York, NY, USA, 1584–1595. DOI:https://doi.org/10.1145/2818048.2820013

[20] Erin L. Brady, Yu Zhong, Meredith Ringel Morris, and Jeffrey P. Bigham. 2013. Investigating the appropriateness of social network question asking as a resource for blind users. In Proceedings of the 2013 conference on Computer supported cooperative work (CSCW '13). Association for Computing Machinery, New York, NY, USA, 1225–1236. DOI:https://doi.org/10.1145/2441776.2441915

[21] Shaomei Wu and Lada A. Adamic. 2014. Visually impaired users on an online social network. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). Association for Computing Machinery, New York, NY, USA, 3133–3142. DOI:https://doi.org/10.1145/2556288.2557415

[22] D. Gurari *et al.*, "VizWiz Grand Challenge: Answering Visual Questions from Blind People," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 3608-3617, doi: 10.1109/CVPR.2018.00380.

[23] Susumu Harada, Daisuke Sato, Dustin W. Adams, Sri Kurniawan, Hironobu Takagi, and Chieko Asakawa. 2013. Accessible photo album: enhancing the photo sharing experience for people with visual impairment. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). Association for Computing Machinery, New York, NY, USA, 2127–2136. DOI:https://doi.org/10.1145/2470654.2481292

[24] Lee, Y., & Oh, U. (2020). SelfieHelper: Improving Selfie Experiences for People with Visual Impairments. 한국HCI 학회논문지, 15(3), 23-30.

[25] Fang, N., Xie, H., & Igarashi, T. (2018). Selfie Guidance System in Good Head Postures. IUI Workshops.

[26] Yuhang Zhao, Shaomei Wu, Lindsay Reynolds, and Shiri Azenkot. 2017. The Effect of Computer-Generated Descriptions on Photo-Sharing Experiences of People with Visual Impairments. Proc. ACM Hum.-Comput. Interact. 1, CSCW, Article 121 (November 2017), 22 pages. DOI:https://doi.org/10.1145/3134756

[27] Hugo Nicolau, André Rodrigues, André Santos, Tiago Guerreiro, Kyle Montague, and João Guerreiro. 2019. The Design Space of Nonvisual Word Completion. In The 21st International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '19) Association for Computing Machinery, New York, NY, USA, 249–261. DOI: https://doi-org.proxy.library.cornell.edu/10.1145/3308561.3353786

[28] Pradhan, A., & Daniels, G. (2021). Inclusive beauty: how buying and using cosmetics can be made more accessible for the visually impaired (VI) and blind consumer. Cosmetics and Toiletries, 136(4), DM4-DM15. Association, USA, Article 109, 1929–1948.