A Face Recognition Application for People with Visual Impairments: Understanding Use Beyond the Lab

Yuhang Zhao[1, 2], Shaomei Wu[2], Lindsay Reynolds[2], Shiri Azenkot[1]

[1]Information Science, Cornell Tech, Cornell University
{yz769, shiri.azenkot}@cornell.edu

[2]Facebook Inc. Menlo Park
{shaomei, lindsayreynolds}@fb.com

ABSTRACT
Recognizing others is a major challenge for people with visual impairments (VIPs) and can hinder their social engagement. We present *Accessibility Bot*, a research prototype bot on Facebook Messenger, which leverages state-of-the-art computer vision algorithms and the existing set of tagged photos of a user's friends on Facebook to help people with visually impairments recognize their friends. Accessibility Bot provides users information about the identity of friends in camera and their facial expressions and attributes. To guide our design, we interviewed eight VIPs to understand their challenges and needs in social activities. We then conducted a diary study with six VIPs to study the use of Accessibility Bot in everyday life. While most participants found the Bot helpful, their experience was undermined by perceived low recognition accuracy, difficulty aiming a camera, and lack of knowledge about the phone's status. We discuss these real-world challenges, identify suitable use cases for Accessibility Bot, and distill design implications for future face recognition applications.

Author Keywords
Visual impairment; face recognition; social activity

ACM Classification Keywords
H.5.1. Information interfaces and presentation: Multimedia Information Systems; K.4.2. Computers and Society: Social Issues.

INTRODUCTION
Recognizing people is a major challenge for people with visual impairments (VIPs) [15,42], preventing them from fully engaging in many social activities and undermining their sense of privacy and physical security [4]. For example, when a VIP enters a meeting room, classroom, or cafeteria, it is difficult for her to know who is present. As a result, VIPs can be reluctant to leave their homes, which may cause more anxiety and depression [17,42].

Face recognition technology presents an opportunity for VIPs to overcome this challenge. State-of-the-art computer vision algorithms can detect [31,39,57] and recognize faces [19,43,60] in the presence of blur, occlusion, noise, and with different poses and lighting conditions. Some face recognition algorithms provide recognition results in near real time with high accuracy (*e.g.*, [47,49]). In addition, computer vision algorithms can recognize people's facial expressions (*e.g.*, [10,44]) and facial attributes such as eyes and facial hair (*e.g.*, [11,29]), which are also important cues for effective social communication [23,33,37]. When integrated into mobile or wearable devices, these technologies can potentially enable VIPs to recognize their friends and better engage in social activities.

Researchers have designed different face recognition systems [14,24,26,27,46] to help VIPs recognize their friends; however, these systems had major limitations. First, most of them were not equipped with state-of-the-art face recognition models based on neural networks (*e.g.*, [58]) and thus were less accurate. Second, they did not include (or have access to) images of faces for training, requiring users to create such a collection from scratch. Third, they did not tackle potential privacy concerns associated with this technology. Given these limitations, prior systems are not ready for widespread, practical applications.

We present *Accessibility Bot* (Figure 1), the first application for people with visual impairments that leverages the state-of-the-art face detection and recognition algorithms [58] and existing face images on Facebook. Accessibility Bot is a bot[1] available on the Facebook Messenger platform. It was designed based on the feedback from our interview study with eight VIPs. When a user scans the environment with the Bot's camera, she gets information about the number of people in front of her in real time. With a double-click gesture, the Bot recognizes and announces the names of people in the current frame if they are among the user's Facebook friends and have tag suggestions turned on[2]. The Bot also describes people's facial expressions and attributes (*e.g.*, facial hair), as shown in Figure 1.

We evaluated Accessibility Bot in the wild with a seven-day diary study with six VIPs. Unlike prior work that evaluated the recognition algorithms themselves [27,54] or conducted a single-session study in a controlled lab environment [9,46], our participants used the Bot in their

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ACM ISBN 978-1-4503-5620-6/18/04…$15.00
https://doi.org/10.1145/3173574.3173789

[1] A Facebook Messenger Bot is a contact on Facebook Messenger, which parses and understands messages sent to it and responses or automates a task accordingly.

[2] At the time of this research, when a user turned on the tag suggestion feature (now under the Face Recognition setting), Facebook would recognize her face in photos uploaded by her friends, and suggest her name as a possible tag.

daily lives (*e.g.*, the workspace, parties, outdoors) without the researchers' supervision. During the seven days, we asked them to use Accessibility Bot in different scenarios, completing a daily survey to document their experiences.

In general, most participants found the Bot helpful and enjoyed using it in social activities. However, the main problem they encountered was incorrect recognition results, as some participants perceived the Bot's accuracy to be much lower than its performance on a standard test dataset [59]. We found that the low perceived accuracy was largely due to participants' difficulty in aiming the camera, which led to the photos' low quality, such as blurriness, low luminance, or cropping of faces in frame. The problem was further exacerbated by the lack of feedback on the photo quality and the smartphone's status. In terms of social acceptance, most participants found it appropriate to use the tool in public, and suggested use cases, such as a big party with lots of people and noise. We finally discuss the challenges of designing and evaluating a face recognition application for real-world use, and distill a set of design implications for future similar applications for VIPs.

RELATED WORK

People with visual impairments face challenges in social activities. Prior research has revealed a low involvement of VIPs in social activities [15–17,42,48]. For example, Desrosiers *et al*. [16] interviewed 64 older adults with visual impairments about their participation in 77 life habits and found that they had significantly lower participation in social roles, compared to those without disabilities. Prior research demonstrated the gap between VIPs and people without such impairments in social activities. While this gap is not entirely due to challenges in recognizing others, researchers have found this as one major issue. Cimariolli *et al*.'s study [15] with 365 older VIP adults showed that recognizing people was one of the most common challenges in their social life.

Face Recognition for People with Visual Impairments

Advanced face recognition technology has created opportunities to improve social activities for VIPs. Several face recognition systems have been designed [14,21,24,54] on smartphone platforms. For example, Kramer *et al*. [24] designed a face recognition application on a smartphone for VIPs. They evaluated their prototype in classrooms and meeting rooms and found that it recognized faces with an accuracy of 96%, even when the faces were looking away from the camera at a 40-degree angle. Researchers have also designed face recognition systems on wearable devices [26,35,37,38,46], such as glasses and smartwatches. Krishna *et al*. [26] mounted a camera on the nose bridge of a pair of sunglasses to detect faces and identify face locations. However, it required a user-constructed dataset of face images and was not formally evaluated with real users. GEAR [46] was a smartwatch prototype equipped with a camera in its wristband. After the system detected a face, the user needed to hold the camera still for several seconds to conduct the face recognition. When evaluated with five blindfolded sighted people and two people with low vision, the system showed an accuracy of 0.83 for sighted people and 0.63 for low vision people.

Unlike research described above that used audio feedback, McDaniel *et al*. [32] used haptic feedback to communicate to the user where nearby people were located. They designed a belt that has a camera resting on the front and seven vibration motors around the user's waist. When a face was detected, the belt generated vibrations from the motor that corresponded to the face's position. The duration of the vibrations indicated the distance of the face from the user. The researchers evaluated the effectiveness of the vibration feedback in informing direction and distance, but no evaluation was done for the whole system.

All prior systems were designed and evaluated in controlled lab environments. However, more challenges are involved in designing for everyday use in an uncontrolled environment. No prior work has considered the effect of recognition failures on the user experience, the difficulty of collecting a sufficiently large set of face images for model training, and potential privacy concerns of nearby people associated with real-time face recognition. OrCam [36] is a commercial product that recognizes people for VIPs. With a camera mounted on the frame of a user's eyeglasses, a user can take a photo of a friend and hear her name in real-time. Another mobile application, Seeing AI [34], works similarly; except that it uses a smartphone's camera instead of a mounted camera. Both OrCam and Seeing AI require users to construct their own training dataset by taking and labeling photos of their friends. There are no usability studies of OrCam or Seeing AI published in the literature.

Recognizing Facial Expressions Attributes for VIPs

Not being able to recognize facial expressions and physical attributes can prevent VIPs from engaging in social activities. Qiu *et al*. [41] interviewed 20 VIPs about nonverbal information in communications. They found that because participants didn't perceive visual signals, they found it difficult to perceive useful communication signals (*e.g.*, facial expressions) and others' feelings. Researchers in the iCare project [37] spoke with VIP students, their teachers, and experts in disability studies, and concluded that descriptions of people's appearances such as gender and facial hair can help VIPs identify their friends.

Researchers have designed systems to help VIPs perceive facial expressions and facial attributes [5,7,25,28,38,50,51]. VibroGlove [25] was a glove with 14 vibration motors on the phalanges of each finger. Seven facial expressions were mapped to seven vibration patterns that simulated the shape of the mouth in each facial expression. Eleven blindfolded sighted participants and one VIP used the glove to recognize facial expressions. The results demonstrated the potential of conveying facial expressions via haptic feedback. Anam *et al*. [5] designed Expression, a Google Glass application that recognized people's facial attributes and expressions on a remote server and informed a user of social signals via audio feedback. Expression was evaluated with six VIPs and four blindfolded sighted participants in two dyadic 10-minute conversations with and without Expression. Participants mostly liked the system but had minor concerns with the delay of the feedback.

Similar with prior face recognition systems, these facial attribute recognition systems were designed and evaluated in controlled lab environments. By evaluating our facial attribute recognition system outside of a lab environment, our study yielded deeper insights than prior lab studies.

Blind Photography Technology

Although face recognition can help VIPs recognize friends in social activities, it is challenging for them to capture good photos [3,22] for recognition. Researchers have developed applications to assist blind photography [1,2,6,8,12,22,52,55]. EasySnap [22,55] helped blind users take good portraits by verbally reporting the face location and size in the frame. PortraitFramer [22] further allowed users to explore a photo on a phone screen by touching, providing vibration feedback to inform the face size and position. While both systems received positive feedback, no evaluation reported how long it took a blind user to capture a good photo when using these systems. Balata *et al*. [8] improved the efficiency of camera aiming by designing BlindCamera, which supported two predefined target areas (central and golden-ration), specifying desired face locations to achieve a good portrait. The system provided audio and vibration feedback to guide the user to move the camera until the detected face matched the predefined target area. BlindCamera was evaluated with 4 VIPs and 12 blindfolded sighted people in an ideal sitation where the shooting target was a portrait on a wall. The results showed that blind users spent 34.69s in average to take a photo, and blindfolded sighted users spent 8.33s.

Although improved the photo quality captured by VIPs, prior design for camera aiming did not consider the real-life social situations. According to the result of prior study [8], adjusting a camera based on the audio guidance can take long, which may not fit for a dynamic social environement where people are moving and turning their heads all the time. Moreover, prior systems did not address the blurriness and luminance issues in the photo-taking process, which are also prevalent in a social environement. Our design focused on the real-life social situations, simplifing the audio guidance to balance the photo quality and interaction time. We discuss the impact of our design on the recognition accuracy and users' experience in the Results section.

EXPLORATORY INTERVIEW STUDY

We conducted an exploratory interview study to understand VIPs' need in social activities. Compared to prior work [26,37] that explored VIPs' needs of nonverbal information, we investigated more deeply into their real-life experiences, studying the challenges they face, their strategies to overcome these challenges, and their unmet needs.

ID	Age/Sex	Visual Condition
P1	39/F	Blind since she was 13 months old
P2	36/F	Blind since 10 years ago.
P3	28/F	Ultra low vision.
P4	40/F	Blind since she was 3 years old.
P5	64/F	Blind. She lose her vision 10 years ago.
P6	22/F	Ultra low vision.
P7	20/M	Blind since birth.
P8	48/F	Blind since she was 15 years old.

Table 1. Participants' demographics of the formative study.

Method

We recruited eight VIPs (1 male, 7 female, Table 1), whose ages ranged from 20 to 64 (mean=37). All participants were legally blind, meaning that either (1) their best-corrected visual acuity in their better eye (*i.e.*, the eye with higher visual acuity) was 20/200 or worse, or (2) their visual field was 20 degrees or narrower [62]. We conducted in-person semi-structured interviews, asking participants to describe their recent social activities. For each activity, we asked them about the difficulties they experienced when finding, recognizing, and interacting with people. We concluded by asking them to describe the information they need when navigating the social activities, and to assess the importance of each kind of information. The interview lasted one hour. We compensated each participant with a $100 gift card.

We coded the interview transcripts using Burnard's method [13]. Two researchers coded two samples separately and discussed the categories. One researcher then coded the remaining transcripts based on the agreed categories.

Findings

The Importance of Recognizing People

Being able to recognize friends was extremely important to the participants. They wanted to know who was around to engage socially and feel safe. In the study, all participants agreed that it was their right to know who was around because it was what sighted people always know. As P3 mentioned, "It's not that I want to be able to judge people, but it's because this is the visual information that I used to get that I'm not getting anymore."

Strategies for recognizing people in social activities

Besides asking for sighted assistance, our participants have developed strategies to identify and locate people nearby. All participants mentioned that they could identify their friends by voice. but they usually needed to walk around and waited until they heard their friends' voice. Another common practice is to call out names and ask about their locations: "I had to go outside and say, 'Who's here?' and wait for people to announce themselves" (P3). People also used their phones to locate others. Despite the development in location sharing services such as PeopleFinder Lite [63], most participants did not use these services but simply call or text their friends when they were nearby.

However, all these strategies have limitations. Calling out for names would not be appropriate in quiet environment such as libraries or classrooms. It is also not as effective in noisy settings such as train stations or busy restaurants. Calling or texting on the phone can suffer from the lack of location references or synchronization. As P5 described, "My [blind] husband and I were going to have dinner and decided to meet inside the front door of a mall. I went in one side and he went through the other. Finally, some woman came up and said, 'Are you meeting a blind gentleman? He is standing about ten feet away.'"

Information Needs in Social Interactions

Participants wanted to know the following kinds of information in order of priority: a) identity, b) relative location, c) physical attributes, and d) facial expressions. Krishna *et al*. [26] showed that facial expressions, identity, and body gestures were the top three kinds of information that VIPs need in social interactions. Our findings extended this prior finding and revealed a different set of priorities.

Identity. All participants believed that knowing the identities of people around them, especially their friends, was helpful in social interactions. In particular, participants were interested in finding a specific person: "I may not care about everybody, it'll be easier for me to be more specific, say looking for certain people" (P1). Some participants wanted to be able to recognize celebrities or some relatively famous people to build connections. "Sometimes you didn't know someone, but you know she's the president of Syrian Action Fund. It would be cool to be like, oh there she is, I want to meet her. That'd be helpful for networking" (P6).

Relative Location. Six participants wanted to know where others were relative to themselves. When noticing a friend is nearby, they wanted to know which direction and how far away he is, in order to find him and start a conversation.

Physical attributes. Five participants agreed that knowing people's appearance was important because it reflected their personalities and fashion trends. As P3 said, "People have their own style and it says things about their personality. I could also ask where they bought [their outfits] and get something similar." However, the other three participants didn't care about others' appearances. Most participants who became blind later in life (P2, P5) or had ultra low vision (P3, P6) were interested in people's appearance, while most congenitally blind participants (P1, P4) were not. There were only two exceptions: P7 who was born blind wanted to learn about fashion trends from descriptions of others' appearance, while P8, who became blind at 15, did not think this was necessary.

Facial Expressions. Three participants believed that it was important to know information about others' facial expressions. P3 also mentioned that knowing others' facial expressions would be useful in business-related interactions. However, others did not think they needed such information because they could perceive a person's affect through conversations and deduce their facial expression. "[Facial expressions] can kind of be sensed through the way people talk. I'm not certain if that would be important if I can grasp it from other cues" (P1).

Other Information. Participants also mentioned other kinds of information that they wanted to know in social activities. Some participants felt that knowing relationships between people was important and could help them develop conversations. "I used to go to parties as a sighted person. You didn't just observe age and all that, you observed the connections and who's coming with who, who's meeting who" (P4). Some participants were interested in who was available for a conversation during social activities. "What sighted people do is they walk up to somebody and start a conversation. So it would be nice to know if a person was already engaged in a conversation or looking around and available for a conversation" (P5).

ACCESSIBILITY BOT

After learning the needs for facial information by VIPs, as well as issues with their current strategies, we designed Accessibility Bot as a research prototype to provide low-cost, mobile, real-time support for such needs. Using Facebook's face recognition algorithm [58], Accessibility Bot provides facial information including identity (which is usually deduced from one's face), face locations, and facial expressions and attributes through screen reader software to a VIP user.

Accessibility Bot is a Facebook Messenger Bot (footnote 1). We integrated the Bot into Facebook Messenger since it supports easy and reliable access to the camera and the Facebook API, which provides face recognition services to recognize a user's Facebook friends. Facebook Messenger is also a widely-used platform with 1.2 billion monthly active users [64], so that people can access the bot easily with a much lower learning curve than a brand new application. Since Facebook is a unique platform that contains both connections among people with real names and a large number of tagged photos, the Bot on Facebook Messenger enables the face recognition to be trained on the existing photos of the users' Facebook friends, and relieves the users from having to construct a collection of face images to train the models by themselves.

We illustrate the interaction flow of Accessibility Bot in Figure 1. The Bot works with TalkBack [20], the built-in screen reader on Android. We chose Android because it is the most widely used mobile platform [65]. When a user opens the Messenger App and starts a conversation with the Bot, it replies with a short introduction, instructing the user to turn on the camera (Figure 1b) for face recognition.

When aiming the camera, we do not provide audio guidance as the framing methods from blind photography technology (*e.g.*, [8,22]). This is because adjusting the camera based on audio feedback may take long [8], which will slown down the interaction significantly. Instead, we simplify this aiming process by reporting face presence and face numbers in real time (Figure 1c). As long as the user hears

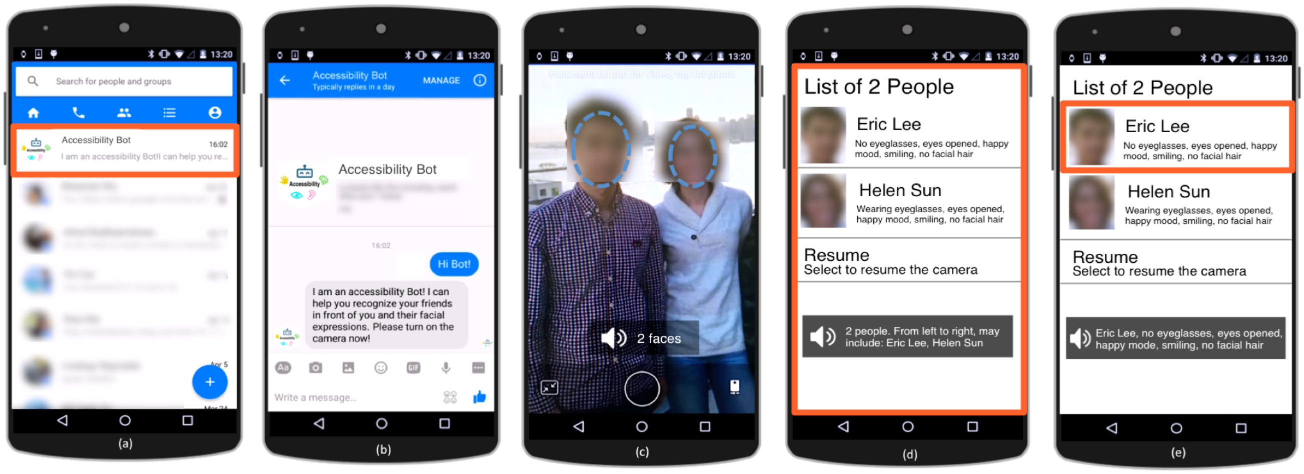

Figure 1. Accessibility Bot: a) the Bot appears as a contact in Facebook Messenger; b) it automatically replies to the user and tells her to use the camera; c) it detects faces in real time and verbally reports the number of faces; d) it lists the recognition results and reports people's identities and relative locations; e) a user can navigate the list and selectively listen to a specific person's detailed facial information. The text on the bottom is for demonstration purposes and does not appear in the app.

that there is a face in the frame, she can trigger the face recognition directly, without adjusting the camera further.

The user can use a double-tap gesture to trigger the face recognition. We chose this gesture since it's a standard phone screen reader gesture to trigger a tap event, which the users are already familiar. This gesture can also ensure that the face recognition is only performed with clear intent. The Bot then sends the current frame to a remote server to recognize the faces. The Bot only recognizes a person if he is the user's Facebook friend, and if the user allows tag suggestions (footnote 2). Otherwise, he will be referred to as "unknown person." Before the recognition, the Bot announces, "start recognition," to notify the user to wait for the recognition results (which usually takes a few seconds). The image is not stored locally or remotely on the server, but discarded after the detection and recognition process is completed.

In addition to recognizing faces, Accessibility Bot also provides people's relative location (left or right), facial expressions (happy, surprised, angry, sad, and neutral), and facial attributes (*e.g.*, whether the face is smiling, or has glasses or facial hair). Once the face recognition process is complete, the Bot organizes and returns all the facial information person by person in a list (Figure 1d).

Since VIPs rely heavily on audio signals to understand the environment, too much audio feedback from Accessibility Bot could be distracting. We minimized the distraction by prioritizing the recognized facial information. Based on participants' feedback on the information priority in the interview study, we grouped the facial information into two priority levels: (1) People's identity and their relative location, and (2) Facial expression and facial attributes. The Bot only automatically announces people's identity and relative face location after the recognition. For example, in Figure 1d, it reports, "two people, from left to right, may include Eric Lee, Helen Sun (fake names)." To reduce distraction from too much audio feedback, we simplified the relative location information by reporting the names in the order of left to right based on the face position, so that a user would know who was on the left, in front of, and to the right of her camera. If the user is interested in a specific person, she can navigate to him in the list with a swipe gesture and listen to his detailed facial information: "Eric Lee, no eye glasses, eyes opened, happy mood, smiling, no facial hair" (Figure 1e). The user can then navigate to the "Resume" button at the end of the list and double tap to resume the real-time face detection.

Accessibility Bot uses the proprietary Facebook API to conduct face recognition. The face recognition model was designed and trained in a similar way to Zhang *et al.*'s method [58]. Our algorithm reached above a 97% accuracy on the People In Photo Albums (PIPA) dataset [59], which consists of over 60000 instances of ~2000 individuals collected from public Flickr photo albums. For each face detected in a photo, we used SVM classifiers to predict different characteristics of the face, including both facial expressions (*e.g.*, happy, angry) and facial attributes (*e.g.*, glasses, facial hair). Those classifiers were trained separately and reached a precision of 0.9 or above on public Facebook photos. To increase the accuracy, the Bot only recognizes a user's top 200 friends on Facebook. We rank the friends by their tie strength to the user using a method similar to Gee *et al.*'s work [18].

DIARY STUDY

We evaluated Accessibility Bot with a seven-day diary study. Our goal was to evaluate its effectiveness and social acceptability in different real-life situations.

Method

Participants

We recruited six VIP participants (2 male, 4 female), whose ages ranged from 32 to 39 (mean=36), as shown in Table 2. All participants were legally blind. Two of them were totally blind, two had ultra low vision and relied on screen readers, while the other two had functional vision and used

Name	Sex/Age	Visual Condition	Technology Experience	Android Phone
Anne	F/39	Blind since she was 13 months.	She has an iPhone and an Android phone. She used Android device and TalkBack for her last job. She used a Bluetooth braille display.	Kyocera Hydro Icon c6730; Android 4.4.2
Kate	F/32	Ultra low vision; only has a little vision in the right eye; can see shadows and light.	She has a Samsung Mega Android smartphone. She has been using Android and TalkBack for five years.	Nexus 5; Android 4.4.3
Peter	M/38	Ultra low vision; totally blind on the left eye and a little vision on the right eye.	He normally uses an iPhone. He uses Android because he teaches adaptive technology. He had a braille display with Android phone.	Nexus 5; Android 4.4.3
Matt	M/32	Blind since two years ago.	He uses an iPhone. He has experience with Android and TalkBack, but he does not like using Android phone.	Nexus 5; Android 5.0.1
Marie	F/33	Low vision; cannot see stuff far away or details; cannot recognize faces.	She has an HTC Desire smartphone and an iPad. She uses both TalkBack and zoom on Android phone.	HTC Desire 700; Android 5.1
Susan	F/39	Low vision; has Coloboma and retinal detachment; no peripheral vision; has blurry and dark spots in the central vision.	She has a Samsung Galaxy S5. She thinks it has a bigger screen than an iPhone. She uses magnification on the smartphone. She does not use screen reader, such as TalkBack.	Samsung Galaxy S5; Android 5.0

Table 2. Participants' demographics in the diary study.

magnification on their smartphones. All participants had experience with Android smartphones. They were all Facebook and Facebook Messenger users. Anne (P1 in Table 1) took part in both the exploratory interview study and this diary study. Participants were compensated with a $100 Amazon gift card for each day of the study.

Procedure

The study consisted of three parts: a one-hour tutorial, a seven-day diary study, and a one-hour final interview.

The tutorial session was conducted in the lab. We started by asking participants' demographic information and their engagement in social activities. We gave participants who didn't have an Android smartphone a Nexus 5 and set up Accessibility Bot for them. We then demonstrated the Bot and asked them to test it on the researchers and themselves (*i.e.*, a selfie). Participants practiced using the Bot until they could confidently recognize all facial information and get the audio descriptions. Susan had little experience with TalkBack, so we also trained her to use TalkBack.

Participants used Accessibility Bot for one week in their daily life. During the course of the study, we asked them to use Accessibility Bot in at least four days. We sent each participant an eight-question survey (Table 3) once each day, asking them whether they used the Bot and their experience with the Bot for that day. We asked them to fill in the survey everyday, including the days in which they did not use the Bot. To avoid interrupting participants' social activities, we only asked them to fill in the daily survey by the end of each day. Participants could email or call us at anytime during the week for troubleshooting.

Lastly, we conducted a follow-up semi-structured interview, where we asked participants whether the Bot was helpful, in what situations they wanted to use it, and how it could be improved. We also asked them to estimate the accuracy of the Bot based on their daily experiences. Finally, we asked participants to demonstrate how they used the Bot both on themselves and on other people.

Q1	Did you use the Accessibility Bot today?
Q2	In what situations did you use Accessibility Bot?
Q3	Approximately how many people were present when you used Accessibility Bot?
Q4	How helpful was the Accessibility Bot today? (Extremely, very, somewhat, a little, not at all)
Q5	How accurate was the Accessibility Bot today? (Extremely, very, somewhat, a little, not at all)
Q6	What was the biggest difficulty that you encountered while using the Accessibility Bot today?
Q7	What can be done to improve Accessibility Bot?
Q8	If you did not use the Bot, why didn't you use it today?

Table 3. Eight-question survey for the diary study

Analysis

We video recorded the tutorial and final interview sessions. We transcribed the videos and coded the transcriptions in the same way as the exploratory interview study. We also aggregated participants' answers to the daily survey.

Results

Usage Patterns

According to the daily survey, all participants used Accessibility Bot in at least seven instances over the whole week, while three participants used it more than ten times. All participants used the Bot on at least four days: two participants used the Bot on four days, three used it on five days, and one used it on all seven days.

Each participant used Accessibility Bot in at least three different scenarios during the week. Participants identified the following scenarios:

- At home with family (6 participants)
- At a small gathering with friends (5 participants)
- At a work-related event (3 participants)
- For taking a selfie (3 participants)
- On the train (1 participants)
- While camping (1 participants)
- While browsing photos (1 participants)

Percieved Helpfulness of Accessibility Bot

Four participants found the Bot helpful and were impressed when the Bot provided accurate recognition results. Peter said, "It was useful because I could tell who was in the crowd [with the Bot]. When I was with people, it would tell me she's smiling, he's smiling, looks angry." However, Matt and Anne found it hard to use because it was difficult to aim the camera (see *Interaction Design* section).

We observed a relationship between participants' visual condition and the perceived helpfulness of the Bot. Two participants with ultra low vision (Kate and Peter) rated the Bot as "very helpful," two participants with medium low vision (Marie and Susan) thought it was "a little to somewhat helpful," while another two who were totally blind (Anne and Matt) had difficulty using the Bot and generally did not find it helpful at all. However, as opposed to Matt who gave negative feedback throughout the whole week, Anne's assessment changed the last two times she used the Bot, saying that it was more accurate and a little helpful. It suggests that users who are blind may experience a higher learning curve than people who still have functional vision when using the Bot.

Our results did not show a novelty effect. While Anne's rating increased from not helpful to a little helpful, Peter's rating reduced from very helpful to somewhat helpful in the last day. All others did not have noticeable trends.

The Scope of Face Recognition

Participants liked being able to recognize their Facebook friends. It enabled them to recognize more people than using other methods. As Kate mentioned, "It's helpful, especially if it's based on your friends on Facebook. Because I know I'm definitely Facebook friends with way more people than I know their voices just by hearing." Participants also noted that limiting the recognizable people to a user's Facebook friends could protect nearby people's privacy. "It did not pick up people who were not on my Facebook page. That was really, really neat. I think that'll probably take care of the privacy issue" (Peter).

However, some participants still found the scope of face recognition to Facebook friends to be limited. For example, Anne met some friends but could not use Accessibility Bot because they were not her Facebook friends. "If there were people in your life you regularly interact with, but you weren't friends with on Facebook, or they don't have a Facebook [account], or they're people who prefer not to have photos on Facebook, I hope the Bot could recognize them too." This finding suggests the need to supplement the collection of photos on Facebook with photos taken by users of their non-Facebook friends, as in OrCam [36].

Recognition Accuracy

Perceived Accuracy

Recognition accuracy heavily influenced participants' experience with Accessibility Bot. When using the Bot, participants often asked a sighted person to check whether the descriptions were accurate. We asked participants to estimate the Bot's accuracy and referred to their estimate as the *perceived accuracy* in this paper. We found that participants' perceived accuracy varied, ranging from 0.2 to 0.9 with a mean of 0.63 (SD=0.25). This perceived accuracy was much lower than the accuracy of our algorithm on the test dataset (see *Accessibility Bot* section). This finding highlighted the mismatch between the high performance of computer vision technologies on standard benchmark datasets and the poor experience of visually impaired users with those technologies in the wild. We identified the factors that contributed to such mismatch based on participants' feedback, as follows.

First, the recognition algorithm performed poorly with photos taken by our participants due to various image quality issues. Our study underlined the challenges of taking good photos for VIPs (see *Interaction Design* section). As a result, photos taken by our participants were more likely to have low luminance, blur, and partially obscured faces. Indeed, some participants noticed that the face recognition was only accurate when the photo was taken in well-lit conditions and captured clear full faces: "The accuracy for finding the friends was pretty decent as long as it was a full face… If it was dim, not accurate at all" (Matt). On the other hand, not being aware of photo quality issues also contributed to participants' disappointment with the Bot, leading to low accuracy ratings.

Second, some participants had high expectations on Accessibility Bot, or beyond its designed use case. For example, Susan described a situation where she used the Bot to recognize a co-worker who was moving away from the camera. "He has a crutch, so I thought maybe it will say something about the crutch. But it didn't. Actually, I don't think it said anything about him at all."

Additionally, people's appearance could have differed from the way they looked in tagged photos, leading to mis-recognition. As Anne said, "Sometimes [the Bot] did not recognize a person just because the way they had on their profile was a lot different than who they were [in real life]."

Reactions to Inaccurate Recognition

Participants reacted differently to Accessibility Bot's recognition failures. Some participants felt confused or frustrated by an incorrect recognition result and would not trust the Bot afterwards. As a result, they always asked a sighted person to double-check the Bot's output. "[The inaccurate recognition] leaves me something I wasn't able to trust. I'd rather trust a human" (Anne).

On the other hand, some participants were more tolerant to the incorrect information. They believed that, although there was some inaccurate recognition, getting some additional information was always better than no information at all. As Peter mentioned, "Nothing is perfect. If you're visually impaired, here you are not going to find a perfect world. All it does is enhancing the information I have and what I can use. Whether it is accurate or not, at least it gives me something to work with."

Participants provided valuable suggestions to address recognition inaccuracy. For example, Susan suggested adding the confidence esitatimation to the results, "I'd rather it say something like 60 percent sure, rather than saying we think it's this, but we cannot guarantee."

Facial Expressions and Facial Attributes

Almost all participants found the information about facial expressions and facial attributes helpful. Anne was the one exception since she felt that she could get that inforamtion easily by talking to people. Other participants explained that knowing people's facial expressions before talking would help them start a conversation. As Peter emphasized, "[The Bot] told me when people were smiling or angry. If I was able to tell people's moods, I could tell someone was upset, and decide, is this a good moment to come over and strike up a conversation? I'm not going to know [people's mood] [without the Bot] until I really start talking."

Participants expressed some concerns about the facial expression and attribute information provided by the Bot. Some found the expressions too dynamic to be captured in a static photo. Anne described a situation where the Bot reported a "sad" facial expression: "The facial expression depends on when I snap it. I don't think you were particularly sad, but you just weren't smiling at that moment." Some participants found facial attributes useless at times. For example, when Matt used the Bot with his girlfriend, the Bot announced "no facial hair." He explained that "it was trying to find facial hair on my girlfriend. If there's no facial hair, just don't say it."

Interaction Design

As described in *Accessibility Bot* section, we required a user to perform a double tap to trigger the face recognition, making sure that the recognition was conducted with a clear intent to alleviate nearby people's concern for being unknowningly recorded. However, all participants preferred real-time face recognition over using a gesture to trigger it. The current interaction was complicated to perform and introduced latency, making it difficult to capture a desired scene. Matt explained, "That is adding an extra step. The photo you got was several seconds ago and the moment has gone already." Moreover, participants sometimes shook the phone when conducting the double tap gesture, leading to a blurry photo. This is especially obvious for Marie, who could only use her right hand because of her disability.

Participants encountered difficulties with aiming the camera, even when trying to capture a person who was standing still. This echoed results from prior studies that focalization was challenging for VIPs (*e.g.*, [22,52,53]).

Participants' ability to aim a camera was closely related to their functional vision. The four participants who had some functional vision (including those with ultra low vision) were able to decide when to trigger the recognition by listening to the face presence in real time. When the Bot indicated that there were faces in in view, they could aim the camera with their existing vision. As Kate described, "It did say one face detected. So I kind of knew, okay, I can take the photo. That was the guide I needed to know when was a good time. But I'm kind of glad I can see a little bit, enough to take a photo." For those with no functional vision, however, the current indicators were not sufficient. Although they knew where people were generally located based on the face presence information reported by the Bot, it was too hard for them to capture good (not blurry) photos that included people's entire faces. This explained our finding that low vision participants felt Accessibility Bot was more helpful than blind participants did, since people with low vision were more likely to take a good photo. It also supported the result from Adam *et al*.'s study [3] that people with low vision are more likely to take photos themselves than those who are totally blind.

Although doable, aiming the camera was still difficult for the two participants with ultra low vision. In a crowded environment, the participants had to get very close to other people, and even ask them to pose for the photo. Kate talked about her experience in a big party, "I still found it hard to take photos at a big gathering. People would be like turning their heads or walking away. I had to get close to them. Like, I'm sorry, can I take your photo?" Moreover, the participants needed to see the face in a large format on the screen so that they could check its quality with their remaining vision. This led them to capture only one face at a time. "I only took one person at a time. It would be a little harder for me to get both of [my friends] in there, especially with the small phone screen" (Kate).

Social Suitability

Social Acceptability of the Bot

We asked participants about the social acceptability of Accessibility Bot. Most participants thought it was acceptable for both private or public settings and claimed that they had an equal right to see and recognize their friends; they wanted to benefit from advances in face recognition technology. "Facial recognition is out there all the time. You walk through an airport, I guarantee you there's some facial recognition software running to pick people out. If computers can do it, why can't I take advantage of it? You're sighted, you can see and tell who I am, then why can't I? I'm not taking any information, but just want to see who you are" (Peter).

Compared with potential privacy issues, some participants cared more about possible safety benefits, and thought face recognition could help them be aware of the surrounding environment and improve their or their children's safety. Peter gave an example of a situation where he hoped Accessibility Bot could recognize his child: "People can scream about privacy, but to me it's a safety issue. I'm visually impaired. Have you ever seen those leashes for kids? I want the Bot to recognize where my kid is."

Unlike other participants, Matt did not think the Bot was socially acceptable. He worried that people would feel uncomfortable if he used a camera to scan around. He thought the Bot would emphasize his disability and further marginalize him: "You are videoing everybody and making people uncomfortable. This is not ideal especially in terms of we're already disabled. We're already being looked at in a certain way, and now, having our phones out and recording everybody can make us look like a creep." Although other participants did not mind using the camera on the smartphone, they preferred using a less noticeable device such as a pair of smart glasses or a camera on a cane.

The performance and usability of the Bot also affected its social acceptability [45]. Some system failures caused social stigma and lead to participants abandoning the tool quickly. For example, Anne was fine with the idea of using the phone to conduct face recognition, but the problems she encountered, such as inaccurate recognitions, discouraged her from using it in public. "I really wish I could use it at a social gathering, but I hesitated. Because when it works flaky on my phone, it's hard for people to understand what I'm doing. I don't want people see me struggling. If it works perfectly, then I can use it" (Anne).

We also asked participants about others' reaction to Accessibility Bot. Participants mostly used the Bot to recognize their friends or family members and said that they were all supportive of its use. This echoed the result from Profita *et al.*'s study [40] that the camera-related device was considered more socially acceptable if it was used to support a person with a disability.

Appropriate Use Cases

According to the daily survey, all participants used Accessibility Bot when gathering with familiar people, such as their families or friends (see *Usage Patterns* Section). Three participants used the Bot at work-related events. They indicated that the Bot could be helpful in some work situations, such as meetings where people wouldn't introduce themselves to the participants. "I want to know who is in the room when nobody else is talking. Or when we're at a big meeting I want to find out who is sitting there. So I would take the Bot out" (Kate). Participants said the most appropriate use cases for the Bot were activities with many people and a lot of noise, where it was difficult for them to recognize their friends by their voices alone.

Some participants felt more comfortable using the Bot in activities with many blind people because they were concerned about sighted people not understanding what they were doing. Anne explained: "I used the Bot mostly with a group of my blind friends because a lot of time blind people will understand better. With my sighted friends, I have to explain how I use my phone and what does this whole camera have to do with the phone. They don't really know about blindness and the accessibility stuff."

Reappropriation

Besides recognizing friends in social activities, participants also reappropriated Accessibility Bot for other purposes.

Four participants used Accessibility Bot to take selfies and used it as a mirror to check their appearance. As Peter described, "It's like a personalized mirror. It tells me no facial hair so that I know I got a clean shave today." However, some low vision participants found it harder to take a good selfie than to take a good photo of someone else because they had to hold the phone far away from themselves to capture their faces properly, but they couldn't see the image well at a distance. "I can kind of see a face in the screen. But when I hold it this way [holding the phone far to take a selfie], I can't see anything" (Kate).

Anne used the Bot as a photo-examination tool to check who's in a photo. As she described, "I was trying to take a picture of a photo on my other phone. There was one person in the photo and it worked. It was very accurate."

Usability Issues in the Wild

Lacking Knowledge of the Phone Status

Some participants reported that the recognition was slow. We observed how they used the Bot in the interview session and found that they had used it with a weak Internet connection but were not aware of it. For example, Peter first used the Bot under a cellphone network service that had a weak signal in the experiment room and the Bot responded very slowly. After we connected his phone to Wi-Fi, the speed improved a lot. "I never did connect it to my Wi-Fi at home. I didn't know that would make a difference. Maybe I did need to be on Wi-Fi" (Peter). Lacking knowledge of the system status may lead to confusion and system failures.

The Effect of TalkBack

TalkBack introduced additional challenges for our participants. Peter and Matt, who were mostly iPhone users, had difficulty using TalkBack. They felt TalkBack had a robotic voice and provided unnecessary information, such as describing the layout of the user interface. Matt was enthusiastic about the Bot at the beginning, but abandoned it because of TalkBack. "I think I was running into more barriers with TalkBack than anything else. It really diminished my ability to appropriate the Bot" (Matt).

Susan had difficulty with TalkBack since she normally used the screen magnifier on her phone rather than the screen reader. She only used TalkBack for the Bot, so she had to switch between TalkBack and magnification. "I don't really like [Talkback]. All of a sudden it's loud, so I have to wear headphones all the time. If I don't want to use it [for other apps], I have to turn it off." Susan wanted the Bot to provide its own audio feedback independent of TalkBack.

DISCUSSION

Our study showed that Accessibility Bot was helpful for VIPs with functional vision. People with no functional vision struggled using the Bot because they had difficulty capturing recognizable photos with the current interaction design. Accessibility Bot is suitable for various daily situations, including gathering with family or close friends, work-related events, loud parties, and activities with many VIPs. In this section, we discuss the challenges that arose when designing and evaluating a face recognition application for real-world use, and distill a set of design implications for future similar applications.

People with visual impairments were sometimes unaware of what information would benefit them. For example, half of our participants did not think it was necessary to receive people's facial expression information in the exploratory study. However, after using Accessibility Bot, most participants enjoyed having this information. P3 reflected on this in the exploratory study, "I'm not clear on how important that information is because I'm not getting it anymore. When people talk about how facial expressions are important, it does make me think that I'm missing something, but I don't know what I'm missing exactly." This demonstrates that while users' reported needs are critical for guiding design, researchers should also consider information that is important to sighted people. In our case, ample prior work has highlighted the importance of facial expressions in communication, so we sought to make that information accessible and understand its impact on VIPs' experience. This approach moves beyond classic user-centered design that focuses on target users' reported needs.

Potential privacy concerns posed constraints to the design of Accessibility Bot. Camera-based technologies have been fueling heated debates about privacy. Many people expressed the fear of being recorded and exposed in a live video stream in public without their knowledge [66]. To alleviate such concerns, our design required the user to perform an explicit gesture to trigger face recognition. However, this design decision made it more challenging for users to aim the camera and, in turn, reduced the recognition accuracy and speed, and the overall effectiveness of the Bot. Since recognition was only conducted on static photos, the recognition results could not be improved by analyzing consecutive frames. Moreover, facial expressions are dynamic, thus expression recognition would have been more robust if it were conducted over a time interval. While we understand the public debate about privacy and real-time face recognition, we note that the discourse neglects issues of access and equity. In our case, real-time face recognition would give VIPs access to information that sighted people already have. Concerns about real-time face recognition exemplify how the needs of VIPs are marginalized when technology is designed and analyzed solely from the perspective of people who do not have disabilities.

The weeklong diary study allowed us to examine the use of Accessibility Bot in real-world situations. As such, we discovered many challenges that researchers do not typically encounter in lab studies of research prototypes. For example, we found that a weak Internet connection had a major impact on the application performance, but participants were rarely aware of this and thought the application simply failed. Some participants also struggled with TalkBack because it had a "robotic" voice, spoke unnecessary information, and required extra navigation operations. Such banal challenges had a major negative impact on people's ability to benefit from the Bot. When designing applications, designers should consider these challenges and try to mitigate them with better feedback and instructions to the user.

The diary study also enabled us to reflect on the effectiveness of computer vision technology in real-world situations, particularly for accessibility [30,56]. Although current technology can provide VIPs information they could not previously access, its effectiveness is hindered by standard datasets and test procedures. Computer vision algorithms are trained on photos taken by sighted people, sometimes on photos shared on social media, resulting in lower-than-expected perceived accuracy in accessibility use cases [61]. For future applications that leverage computer vision as assistive technology, it is important to foster awareness of photo quality, as well as to train the algorithm with examples provided by target users.

The interaction design also had strong impact on the effectiveness of the computer vision technology. To balance the interaction speed and recognition accuracy to fit for the dynamic social environement, we simplifyied the audio feedback from pior blind photography research and only reported the face presence in real time. However, the diary study showed that VIPs had difficulty capturing recognizable faces (faces were cut-off, blurry, or dark). Thus sacrificing accuracy to achieve a fast interaction may not be a good design option. An improved version should adapt techniques from blind photography research to increase the recognition accurcay.

LIMITATIONS

Our evaluation had limitations. First, we had only six participants, so our findings are deep but not large scale. Second, we were unable to measure the recognition accuracy for photos taken during the diary study; we did not collect the photos taken by our participants because the photos included people who were not participants in the study. However, we did discuss the in-lab accuracy of the recognition algorithms and the perceived accuracy of our participants, generating a sharp contrast that emphasized the importance of designing suitable computer vision technology and a dataset for VIPs. We blieve that an accuracy test on a manually-generated dataset taken with our system by the researchers would serve as a better point of comparison. We will use this method in the future. Future research should also explore suitable approaches that would both safeguard non-users' privacy and effectively evaluate computer vision technology in real-world situations (*e.g.*, a diary study).

CONCLUSION

In this paper, we designed Accessibility Bot, a research prototype Facebook Messenger Bot that recognizes friends of people with visual impairments by leveraging Facebook's high performance face recognition algorithms and the large set of tagged photos on Facebook. We evaluated Accessibility Bot outside of a laboratory environment through a weeklong diary study to understand participants' daily experiences with the Bot in different social situations. We found that Accessibility Bot was helpful for most participants. However, people's experience was undermined by the low perceived accuracy, difficulty with aiming the camera, and other usability issues that do not typically arise in lab studies. We discussed these real-world challenges and provided design implications for future face recognition applications for users with visual impairments.

ACKNOWLEDGMENT

We thank our colleagues at Facebook for their insights and guidance on the design and implementation of the Bot. We thank the LightHouse for the Blind and Visually Impaired for sourcing participant and providing the space for the interview studies.

REFERENCES

1. Dustin Adams and Sri Kurniawan. 2014. A blind-friendly photography application for smartphones. *ACM SIGACCESS Accessibility and Computing*, 108: 12–15. https://doi.org/10.1145/2591357.2591358

2. Dustin Adams, Sri Kurniawan, Cynthia Herrera, Veronica Kang, and Natalie Friedman. 2016. Blind Photographers and VizSnap : A Long-Term Study. In *ASSETS '16 The 18th International ACM SIGACCESS Conference on Computers and Accessibility*, 201–208.

3. Dustin Adams, Lourdes Morales, and Sri Kurniawan. 2013. A Qualitative Study to Support a Blind Photography Mobile Application. *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments - PETRA '13*: 1–8. https://doi.org/10.1145/2504335.2504360

4. Tousif Ahmed. 2015. Privacy Concerns and Behaviors of People with Visual Impairments. *Proceedings of CHI 2015*: 3523–3532. https://doi.org/10.1145/2702123.2702334

5. ASM Iftekhar Anam, Shahinur Alam, and Mohammed Yeasin. 2014. Expression: A Dyadic Conversation Aid using Google Glass for People who are Blind or Visually Impaired. *Proceedings of the 6th International Conference on Mobile Computing, Applications and Services*: 57–64. https://doi.org/10.4108/icst.mobicase.2014.257780

6. Marynel V Azquez and Aaron Steinfeld. 2014. An Assisted Photography Framework to Help Visually Impaired Users Properly Aim a Camera. *ACM Transactions on Computer-Human Interaction* 21, 5: 1–29.

7. Shantanu Bala, Troy McDaniel, and Sethuraman Panchanathan. 2014. Visual-to-tactile mapping of facial movements for enriched social interactions. *2014 IEEE International Symposium on Haptic, Audio and Visual Environments and Games, HAVE 2014 - Proceedings*: 82–87. https://doi.org/10.1109/HAVE.2014.6954336

8. Jan Balata, Zdenek Mikovec, and Lukas Neoproud. 2015. BlindCamera: Central and Golden-ratio Composition for Blind Photographers. *Proceedings of the Mulitimedia, Interaction, Design and Innnovation*: 8:1--8:8. https://doi.org/10.1145/2814464.2814472

9. L. Balduzzi, G. Fusco, F. Odone, S. Dini, M. Mesiti, A. Destrero, and A. Lovato. 2010. Low-cost face biometry for visually impaired users. *BioMS 2010 - 2010 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications, Proceedings*: 45–52. https://doi.org/10.1109/BIOMS.2010.5610444

10. Marian Stewart Bartlett, Gwen Littlewort, Ian Fasel, and Javier R. Movellan. 2003. Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction. In *2003 Conference on Computer Vision and Pattern Recognition Workshop*, 53–53. https://doi.org/10.1109/CVPRW.2003.10057

11. N. Bellustin, Y. Kalafati, Kovalchuck, A. Telnykh, O. Shemagina, V. Yakhno, Abhishek Vaish, Pinki Sharma, and Shirshu Verma. 2011. Instant Human Face Attributes Recognition System. *IJACSA*, Special Issue Artif. Intel.: 112–120.

12. P Bigham, Chandrika Jayant, Andrew Miller, Brandyn White, Tom Yeh, and College Park. 2010. VizWiz :: LocateIt - Enabling Blind People to Locate Objects in Their Environment. 65–72.

13. P Burnard. 1991. A method of analysing interview transcripts in qualitative research. *Nurse education today* 11, 6: 461–466. https://doi.org/10.1016/0260-6917(91)90009-Y

14. Shonal Chaudhry and Rohitash Chandra. 2015. Design of a mobile face recognition system for visually impaired persons. *arXiv preprint arXiv:1502.00756*. Retrieved from http://arxiv.org/abs/1502.00756

15. V. R. Cimarolli, K. Boerner, M. Brennan-Ing, J. P. Reinhardt, and a. Horowitz. 2012. Challenges faced by older adults with vision loss: a qualitative study with implications for rehabilitation. *Clinical Rehabilitation* 26: 748–757. https://doi.org/10.1177/0269215511429162

16. Johanne Desrosiers, Marie-Chantal Wanet-Defalque, Khatoune Témisjian, Jacques Gresset, Marie-France Dubois, Judith Renaud, Claude Vincent, Jacqueline Rousseau, Mathieu Carignan, and Olga Overbury. 2009. Participation in daily activities and social roles of older adults with visual impairment. *Disability and rehabilitation* 31, 15: 1227–1234. https://doi.org/10.1080/09638280802532456

17. Laura E. Dreer, Timothy R. Elliott, Donald C. Fletcher, and Marsha Swanson. 2005. Social Problem-Solving Abilities and Psychological Adjustment of Persons in Low Vision Rehabilitation. *Rehabilitation Psychology* 50, 3: 232–238. https://doi.org/10.1037/0090-5550.50.3.232

18. Laura K. Gee, Jason Jones, and Moira Burke. 2017. Social Networks and Labor Markets: How Strong Ties Relate to Job Finding on Facebook's Social Network.

Journal of Labor Economics 35, 2: 485–518. https://doi.org/10.1086/686225

19. A.S. Georghiades, D.J. Kriegman, and P.N. Belhurneur. Illumination cones for recognition under variable lighting: faces. In *Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.98CB36231)*, 52–58. https://doi.org/10.1109/CVPR.1998.698587
20. Google. Get started on Android with TalkBack - Android Accessibility Help. Retrieved March 6, 2017 from https://support.google.com/accessibility/android/answer/6283677?hl=en
21. Rabia Jafri, Syed Abid Ali, Hamid R Arabnia, Saudi Arabia, and Ism-tec Llc. 2013. Face Recognition for the Visually Impaired. *Proceedings of the International Conference on Information and Knowledge Engineering (IKE)*.
22. Chandrika Jayant, Hanjie Ji, Samuel White, and Jeffrey P. Bigham. 2011. Supporting blind photography. *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '11*: 203–210. https://doi.org/10.1145/2049536.2049573
23. T. G. Knapp, M. L., Hall, J. A., Horgan. 2013. *Nonverbal Communication in Human Interaction*. Cengage Learning. https://doi.org/10.1080/17404620601014724
24. K. M. Kramer, D. S. Hedin, and D. J. Rolkosky. 2010. Smartphone based face recognition tool for the blind. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*: 4538–4541. https://doi.org/10.1109/IEMBS.2010.5626038
25. Sreekar Krishna, Shantanu Bala, Troy McDaniel, Stephen McGuire, and Sethuraman Panchanathan. 2010. VibroGlove: an assistive technology aid for conveying facial expressions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI) Extended Abstracts*: 3637–3642. https://doi.org/10.1145/1753846.1754031
26. Sreekar Krishna, Dirk Colbry, John Black, Vineeth Balasubramanian, and Sethuraman Panchanathan. 2008. A Systematic Requirements Analysis and Development of an Assistive Device to Enhance the Social Interaction of People Who are Blind or Visually Impaired. *Workshop on Computer Vision Applications for the Visually Impaired*. Retrieved from http://hal.inria.fr/inria-00325432
27. Sreekar Krishna, Greg Little, John Black, and Sethuraman Panchanathan. 2005. A wearable face recognition system for individuals with visual impairments. *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility - Assets '05*: 106. https://doi.org/10.1145/1090785.1090806
28. Sreekar Krishna and Sethuraman Panchanathan. 2010. Assistive technologies as effective mediators in interpersonal social interactions for persons with visual disability. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6180 LNCS, PART 2: 316–323. https://doi.org/10.1007/978-3-642-14100-3_47
29. Michael J. Lyons, Julien Budynek, Andre Plantey, and Shigeru Akamatsu. 2000. Classifying facial attributes using a 2-D Gabor wavelet representation and discriminant analysis. In *Proceedings - 4th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2000*, 202–207. https://doi.org/10.1109/AFGR.2000.840635
30. Haley Macleod, Cynthia L Bennett, Meredith Ringel Morris, and Edward Cutrell. 2017. Understanding Blind People ' s Experiences with Computer-Generated Captions of Social Media Images. https://doi.org/10.1145/3025453.3025814
31. M Mathias and R Benenson. 2014. Face detection without bells and whistles. *Computer Vision–ECCV* ….
32. Troy McDaniel, Sreekar Krishna, Vineeth Balasubramanian, Dirk Colbry, and Sethuraman Panchanathan. 2008. Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. *HAVE 2008 - IEEE International Workshop on Haptic Audio Visual Environments and Games Proceedings*, October: 13–18. https://doi.org/10.1109/HAVE.2008.4685291
33. Albert. Mehrabian. 1981. Silent messages : implicit communication of emotions and attitudes. 196. https://doi.org/10.1037/020475
34. Microsoft. Seeing AI | Talking camera app for those with a visual impairment. Retrieved August 25, 2017 from https://www.microsoft.com/en-us/seeing-ai/
35. Laurindo Britto Neto, Felipe Grijalva, Vanessa Regina Margareth Lima Maike, Luiz Cesar Martini, Dinei Florencio, Maria Cecilia Calani Baranauskas, Anderson Rocha, and Siome Goldenstein. 2016. A Kinect-Based Wearable Face Recognition System to Aid Visually Impaired Users. *IEEE Transactions on Human-Machine Systems* 47, 1: 52–64. https://doi.org/10.1109/THMS.2016.2604367
36. OrCam. OrCam - See for Yourself. Retrieved July 7, 2015 from http://www.orcam.com/
37. Sethuraman Panchanathan, John Black, Mike Rush,

and Vivek Iyer. 2003. ICare - A User Centric Approach to the Development of Assistive Devices for the Blind and Visually Impaired. *Proceedings of the International Conference on Tools with Artificial Intelligence*: 641–648. https://doi.org/10.1109/TAI.2003.1250252

38. Sethuraman Panchanathan, Shayok Chakraborty, and Troy McDaniel. 2016. Social Interaction Assistant: A Person-Centered Approach to Enrich Social Interactions for Individuals with Visual Impairments. *IEEE Journal on Selected Topics in Signal Processing* 10, 5: 942–951. https://doi.org/10.1109/JSTSP.2016.2543681
39. J Parris, M Wilber, and B Heflin. 2011. Face and eye detection on hard datasets. *... Joint Conference on*.
40. Halley P. Profita, Reem Albaghli, Leah Findlater, Paul Jaeger, and Shaun K. Kane. 2016. The AT Effect: How Disability Affects the Perceived Social Acceptability of Head-Mounted Display Use. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*: 4884–4895. https://doi.org/10.1145/2858036.2858130
41. Shi Qiu, Jun Hu, and Matthias Rauterberg. 2015. Nonverbal Signals for Face-to-Face Communication between the Blind and the Sighted. In *Proceedings of International Conference on Enabling Access for Persons with Visual Impairment*, 157–165.
42. Sharon Zell Sacks, Karen E. Wolffe, and Deborah Tierney. 1998. Lifestyles of students with visual impairments: Preliminary studies of social networks. *Exceptional Children* 64, 4: 463–478.
43. Ashok Samal and Prasana A. Iyengar. 1992. Automatic recognition and analysis of human faces and facial expressions: a survey. *Pattern Recognition* 25, 1: 65–77. https://doi.org/10.1016/0031-3203(92)90007-6
44. Caifeng Shan, Shaogang Gong, and Peter W. McOwan. 2009. Facial expression recognition based on Local Binary Patterns: A comprehensive study. *Image and Vision Computing* 27, 6: 803–816. https://doi.org/10.1016/j.imavis.2008.08.005
45. Kristen Shinohara and Jacob O. Wobbrock. 2016. Self-Conscious or Self-Confident? A Diary Study Conceptualizing the Social Accessibility of Assistive Technology. *ACM Transactions on Accessible Computing* 8, 2: 1–31. https://doi.org/10.1145/2827857
46. Laurindo de Sousa Britto Neto, Vanessa Regina Margareth Lima Maike, Fernando Luiz Koch, Maria Cecília Calani Baranauskas, Anderson De Rezende Rocha, and Siome Klein Goldenstein. 2015. A wearable face recognition system built into a smartwatch and the blind and low vision users. In *International Conference on Enterprise Information Systems*, 515–528. https://doi.org/10.1007/978-3-319-29133-8_25
47. Tolga Soyata, Rajani Muraleedharan, Colin Funai, Minseok Kwon, and Wendi Heinzelman. 2012. Cloud-Vision: Real-time face recognition using a mobile-cloudlet-cloud acceleration architecture. In *2012 IEEE Symposium on Computers and Communications (ISCC)*, 000059–000066. https://doi.org/10.1109/ISCC.2012.6249269
48. JA Stelmack and AA Rosenbloom. 2003. Patients' perceptions of the need for low vision devices. *Journal of Visual ...*.
49. Yaniv Taigman, Ming Yang, Marc Aurelio Ranzato, and Lior Wolf. 2014. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *CVPR 2014*. https://doi.org/10.1109/CVPR.2014.220
50. Md. Iftekhar Tanveer, A S M Iftekhar Anam, A K M Mahbubur Rahman, Sreya Ghosh, and Mohammed Yeasin. 2012. FEPS: a Sensory Substitution System for the Blind to Perceive Facial Expressions. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*, 207–208. https://doi.org/10.1145/2384916.2384956
51. Mohammad Iftekhar Tanveer and Mohammed Ehsan Hoque. 2014. A Google Glass App to Help the Blind in Small Talk. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility*, 297–298. https://doi.org/10.1145/2661334.2661338
52. Marynel Vázquez and Aaron Steinfeld. 2012. Helping Visually Impaired Users Properly Aim a Camera. *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*: 95–102. https://doi.org/10.1145/2384916.2384934
53. Violeta Voykinska, Cornell Tech, and Gilly Leshed. 2016. How Blind People Interact with Visual Content on Social Networking Services. FEBRUARY. https://doi.org/10.1145/2818048.2820013
54. Zixuan Wang, Jinyun Yan, Cong Pang, David Chu, and Hamid Aghajan. 2012. Who is Here: Location Aware Face Recognition. *Proceedings of the Third International Workshop on Sensing Applications on Mobile Phones*: 8:1--8:5. https://doi.org/10.1145/2389148.2389156
55. Samuel White, Hanjie Ji, and Jeffrey P. Bigham. 2010. EasySnap: Real-time Audio Feedback for Blind Photography. *Adjunct proceedings of the 23nd annual ACM symposium on User interface software and technology - UIST '10*: 409. https://doi.org/10.1145/1866218.1866244

56. Shaomei Wu, Jeffrey Wieland, Omid Farivar, and Julie Schiller. 2017. Automatic Alt-text: Computer-generated Image Descriptions for Blind Users on a Social Network Service. *Cscw*. https://doi.org/10.1145/2998181.2998364

57. C Zhang and Z Zhang. 2010. A survey of recent advances in face detection.

58. Ning Zhang, Manohar Paluri, Yaniv Taigman, Rob Fergus, and Lubomir Bourdev. 2015. Beyond Frontal Faces: Improving Person Recognition Using Multiple Cues. 4804–4813. Retrieved July 2, 2017 from http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Zhang_Beyond_Frontal_Faces_2015_CVPR_paper.html

59. Ning Zhang, Manohar Paluri, Yaniv Taigman, Rob Fergus, and Lubomir Bourdev. 2015. PIPA Dataset.

60. W Zhao, R Chellappa, P.J. Phillips, and a Rosenfeld. 2003. Face recognition: A literature survey. *Acm Computing Surveys* 35, 4: 399–458. https://doi.org/10.1145/954339.954342

61. 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, 2: 1–22. https://doi.org/10.1145/3134756

62. Legally Blind: What Does It Mean? Retrieved April 26, 2017 from http://www.allaboutvision.com/lowvision/legally-blind.htm

63. Sendero Group: PeopleFinder Lite App. Retrieved August 25, 2017 from http://www.senderogroup.com/products/PeopleFinder/index.html

64. Facebook Messenger hits 1.2 billion monthly users, up from 1B in July | TechCrunch. Retrieved August 25, 2017 from https://techcrunch.com/2017/04/12/messenger/

65. Android smartphone users in the US 2014-2016 | Statistic. Retrieved March 13, 2017 from https://www.statista.com/statistics/232786/forecast-of-andrioid-users-in-the-us/

66. Google Glass users fight privacy fears - CNN.com. Retrieved May 1, 2017 from http://www.cnn.com/2013/12/10/tech/mobile/negative-google-glass-reactions/