A Systematic Requirements Analysis and Development of an Assistive Device to Enhance the Social Interaction of People Who are Blind or Visually Impaired

Sreekar Krishna*, Dirk Colbry, John Black, Vineeth Balasubramanian, and Sethuraman Panchanathan

Center for Cognitive Ubiquitous Computing (CUbiC)
School of Computing and Informatics (SCI)
Arizona State University, Tempe AZ 85281, USA

Abstract. People who are blind can be at a disadvantage in social situations because face-to-face communication relies heavily on non-verbal cues. Lack of access to cues such as eye contact can sometimes lead to awkward situations, such as answering a question that was directed to another person. This problem is compounded by the fact that sighted people are not always aware of their use of these nonverbal cues, and they do not take this into account when communicating with people who are blind. In the long term, this can limit employment opportunities for people who are blind, can adversely impact their professional advancement, and can even isolate them. To address this problem, we propose a methodology for determining and enumerating the most important needs of people who are blind, as they interact with others in social situations. Next, we establish design guidelines for building a Social Interaction Assistant to provide access to essential visual information during social encounters for the visually impaired. Finally, we use our design guidelines to develop a wearable Social Interaction Assistant prototype that is aimed at providing this essential information in real-time, during social interactions.

1 Introduction

In the United States 1.3 million people are legally blind and 11.4 million people have some sort of visual impairment. People with visual impairments use creative methods for coping with the lack of visual information in their daily lives. For example, specialized training can help people with visual impairments use their sense of hearing more effectively, training in the use of a white cane can help them navigate successfully, and training in the use of screen reading software can make computers more accessible.

Historically, the development of assistive devices has tended to be characterized by a technology-centric approach, which begins by asking "What can we do?" This approach is often inspired by a newly emerging technology, and it tends to produce *one-size-fits-all* technological solutions to the obvious problems that people with disabilities might

* Sreekar.Krishna@asu.edu; Ph: +1 (480) 326 6334; Fax: +1 (480) 965 1885

have already largely solved for themselves. One example of this type of technology-centric approach is the development of numerous *robotic guide dogs* [1] [2] [3]. Sadly, the technology-centric approach often fails to address some less obvious, but more important, unsolved problems.

Another problem with the technology-centered approach is that it often focuses only on the disabilities of the user, without taking into full account the user's abilities. For example, people who are blind are often able to perceive the presence of large objects in the environment around them. Ambient sound sources in the environment provide a form of *audio illumination* and the resulting sounds bouncing off of objects (or sounds shadowed by objects) allow a person who is blind to detect the presence of those objects [4]. Sometimes in attempting to overcome a disability, developers of assistive devices unintentionally interfere with the user's abilities. For example, navigational assistive devices that require the user to wear headphones or earphones deprive the user of sounds that are vital to the perception of the environment.

In an effort to avoid the problems of technology-centric design, this paper describes a human-centric approach that begins by asking "What is needed?" It answers this question by identifying unmet needs (as seen by people with disabilities) and guides the development of assistive devices to specifically address these needs. This human-centric approach often uncovers a wide variety of such needs, inspiring the development of a variety of different devices.

2 Motivation

In order to identify unmet needs of the visually impaired community, we established two focus groups consisting primarily of people who are blind, as well as disability specialists and parents of students with visual impairment and blindness. Members of these focus groups who were blind or visually impaired were encouraged to speak freely about their challenges in coping with daily living. During these focus groups, the participants agreed on many issues as being important problems. However, one particular problem - that of engaging freely with their sighted counterparts - was highlighted as a particularly important problem that was not being addressed by technology specialists. As an example of this type of social disconnect, consider a simple form of nonverbal communication: glancing at a watch to signal that it is time to wrap up a meeting. The sighted participants might respond to such a glance automatically, without consciously realizing that this visual information is not accessible to a participant who is blind. Similarly, a sighted person asking a question in a group will use gaze direction and eye contact to indicate to whom the question is directed. Without access to this visual cue, people who are blind might be left wondering whether the question was directed towards them. They can answer immediately (at the risk of feeling foolish if the question was not directed at them) or they can wait to see if anyone else answers (and risk being thought of as rather slow witted).

People who are blind often find themselves unable to independently access essential visual information. This can be a particularly difficult problem in social situations, where as much as 65 % of two-person conversation is non-verbal [[5], p.30]. Compounding this problem is the fact that sighted people are often unaware of the nonverbal cues that they use in social situations, and do not make allowances for this when communicating with people who are blind. People who are blind are often unwilling, or unable, to constantly ask others for this non-verbal information, because it places a perceived burden on friends and family. Over time such situations can put a person who is blind at a disadvantage, potentially leaving them socially isolated [6].

This social isolation is an important concern. Humans are social beings who need meaningful communication and interaction with others to lead healthy and productive lives [7]. However, while many assistive devices have been developed to meet a wide range of the needs of people who are blind, there has not been enough attention given to the development of assistive devices that satisfy the need for access to nonverbal communication cues. In response, our goal has been to develop assistive devices that will allow people who are blind to access important visual information during social encounters.

Section 2 lists the needs that the members of our focus groups and the respondents to our online survey felt were most important in social situations. Section 3 discusses existing computer vision algorithms that might be used to meet those needs. Section 4 presents the three design alternatives that we considered for our Social Interaction Assistant, and Section 5 describes the hardware and software implementation that we developed. Section 6 provides a summary, and a discussion of our plans for future development.

3 Requirements for a Social Interaction Assistant

Members of our focus groups voiced many concerns about social interaction and isolation. Based on these conversations, we compiled a list of needs that are often experienced by people with visual impairments. In doing so, we identified two aspects of social interaction that are particularly important: *access to the non-verbal cues of others* during social interactions, and *how one is perceived by others* during social interactions.

Access to Non-Verbal Cues Access to the non-verbal cues of others during a social interaction is something that sighted people take for granted. Non-verbal cues such as eye contact, hand gestures, and body posture play very important roles in social communication [8]. As an important first step in the design of our Social Interaction Assistant, we decided to develop a better understanding of the most important nonverbal cues that we need to convey to a person who is blind or visually impaired.

How One is Perceived by Others Although people who are blind cannot visually perceive their own appearance and demeanor during social interactions, members of our focus groups indicated that understanding how others perceived them was important. For example, assistive technology that makes the user stand out (such as the Jordy [9] enhanced vision system) were rejected. "*Don't make me look like a Martian*" was a sentiment that was shared by most of the attendees. This sentiment makes it important that assistive technologies be extremely discreet. Ideally, assistive devices should allow a person who is blind or visually impaired to interact with sighted peers without those peers even being aware of their disability, or their assistive device.

Another cause of social isolation for people who are blind are distracting or unusual body mannerisms [5]. For example, a person who is blind might rock back and forth during a conversation without being fully aware of how this is being perceived by sighted people. Training can make people who are blind more aware of the social norms, expectations, and needs of sighted people - for example, turning their head to face a sighted person as they speak to her as a substitute for establishing eye contact. Assistive technologies could also play a role in helping individuals who are blind recognize, learn, and practice body mannerisms that are considered socially appropriate.

3.1 Essential Requirements

As a first step toward the development of a Social Interaction Assistant, we used our focus group results to identify and enumerate the following list of needs for information that is not always accessible by people who are blind, as they engage in social interactions:

1. Knowing how many people are standing in front you, and where each person is standing.
2. Knowing where a person is directing his/her attention.
3. Knowing the identities of the people standing in front of you.
4. Knowing something about the appearance of the people standing in front of you.
5. Knowing whether the physical appearance of a person who you know has changed since the last time you encountered him/her.
6. Knowing the facial expressions of the person standing in front of you.
7. Knowing the hand gestures and body motions of the person standing in front of you.
8. Knowing whether your personal mannerisms do not fit the behavioral norms and expectations of the sighted people with whom you will be interacting.

3.2 Online Survey

We conducted a web-based survey in order to validate the list of needs that we identified from our focus groups, and to help establish the relative importance for each of these needs. This survey was anonymously completed by 27 people, of whom 16 were blind, 9 had low vision, and 2 were sighted specialists in the area of visual impairment.

The online survey consisted of eight questions that corresponded to the previously identified list of needs. Respondents answered each question using a five-point Likert scale: (1) Strongly disagree, (2) Disagree, (3) Neutral, (4) Agree, and (5) Strongly agree. Table 1 shows the eight questions, sorted by descending importance, as indicated by the survey respondents (the question numbers correspond to the need listed in the previous section).

Need	The Question	The Mean Score
8.	I would like to know if any of my personal mannerisms might interfere with my social interactions with others.	4.5
6.	I would like to know what facial expressions others are displaying while I am interacting with them.	4.4
3.	When I am standing in a group of people, I would like to know the names of the people around me.	4.3
7.	I would like to know what gestures or other body motions people are using while I am interacting with them.	4.2
1.	When I am standing in a group of people, I would like to know how many people there are, and where each person is.	4.1
2.	When I am standing in a group of people, I would like to know which way each person is facing, and which way they are looking.	4.0
5.	I would like to know if the appearance of others has changed (such as the addition of glasses or a new hair-do) since I last saw them.	3.5
4.	When I am communicating with other people, I would like to know what others look like.	3.4

Table 1: Results of the online survey

The histogram of responses (Fig. 1) from the online survey reveals the importance levels of the various needs. The responses to question 8 suggest that the respondents are highly concerned about how they are perceived by their sighted peers. The responses to questions 3, 6, and 7 suggest that respondents would like to know the identities, facial expressions and body gestures of the people with whom they are communicating. The responses to questions 4 and 5 indicate that there was a wide variation in respondents' interest in (4) knowing the physical appearance of people with whom they are communicating and (5) knowing about changes in the physical appearance of people with whom they are communicating. Many respondents indicated moderate, little, or no interest in either of these areas.

4 Relevance of Computer Vision in Building Solutions

Many of the needs for social interaction could be served by current and future computer vision research. For example, the problem of person recognition has been approached using algorithms for face recognition and gait-based identity recognition. Other research efforts in computer vision could be employed to meet other needs in the list. Table 2 shows the potential uses of existing computer vision algorithms for addressing the various needs in the list.

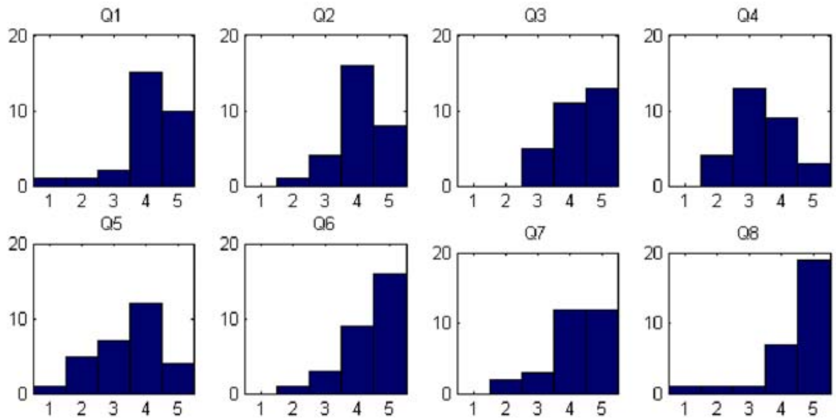

Fig. 1: Histograms of online survey responses

Need	Description	Existing Computer Vision Algorithms
1	People detection and localization	a. Face Detection [10] b. Face Tracking [11] c. Skin region extraction [12]
2	Assessing attention	a. Head pose estimation [13] b. Eye detection [14] c. Eye contact detection [14] d. Facial feature extraction [11]
3	Person identification	a. Face recognition [15] [11] b. Gait Recognition [16]
4 & 5	Extracting physical descriptors	a. Facial Feature extraction [11] b. Eye glass detection [11] c. Clothing detection [17]
6	Facial expression	a. Facial Feature extraction [11] b. Facial action analysis [11]
7	Gesture and mannerism recognition	a. Motion analysis [16] b. Skin region extraction [12] c. Activity recognition [16] d. Body part detection [17] e. Facial expression recognition [11]
8	Social training	a. Rocking detection (Motion estimation) [16] b. Social gaze training (Face detection and eye detection) [14]

Table 2: Potential uses of computer vision algorithms

5 Alternative Platforms for a Social Interaction Assistant

Having determined the requirements for a Social Interaction Assistant, we next concentrated on a potential platform for the device. We observed that four important criteria had to be taken into account: (1) the device must be *unobtrusive* and *socially acceptable*; (2) the device must be *lightweight* and *compact* for easy everyday use; (3) the device should be *wearable*, so users have their hands free; and (4) the device should allow the user to control the direction of the wearable device unobtrusively. Based on these considerations, we investigated three different conceptual approaches, including:

Concept 1: A wearable video camera in a clip-on device, and a small audio emitter device that could be worn on the ear without obstructing normal hearing. Both of these devices would be connected to a compact computing element such as an Ultra-mobile PC (UMPC) (Fig. 2).

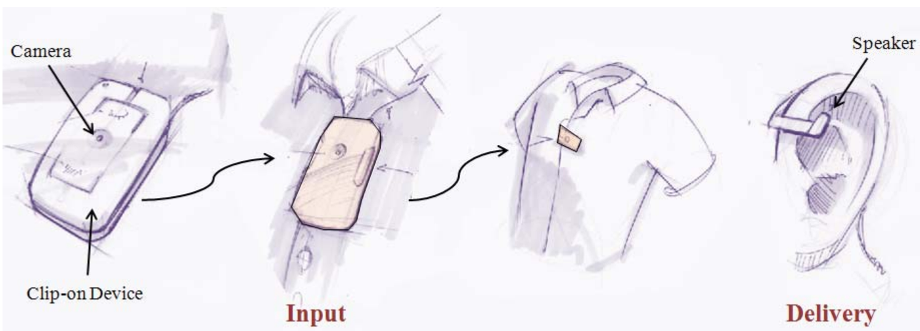

Fig. 2: A clip on camera and small audio emitter

Concept 2: A tiny, ear-mounted video camera and sound emitter (inspired by Bluetooth headsets) mounted on a small device that communicates with a UMPC (Fig. 3).

Concept 3: A tiny video camera and a sound emitter mounted unobtrusively in a pair of glasses - both of which are attached to a UMPC (Fig. 4).

Based on the previously mentioned criteria, we concluded that Concept 3 best met the requirements for unobtrusiveness, social acceptability, light weight, compactness, wearability, and ease of user control.

6 The Social Interaction Assistant Prototype

The following sections detail the hardware and software developed for our prototype Social Interaction Assistant.

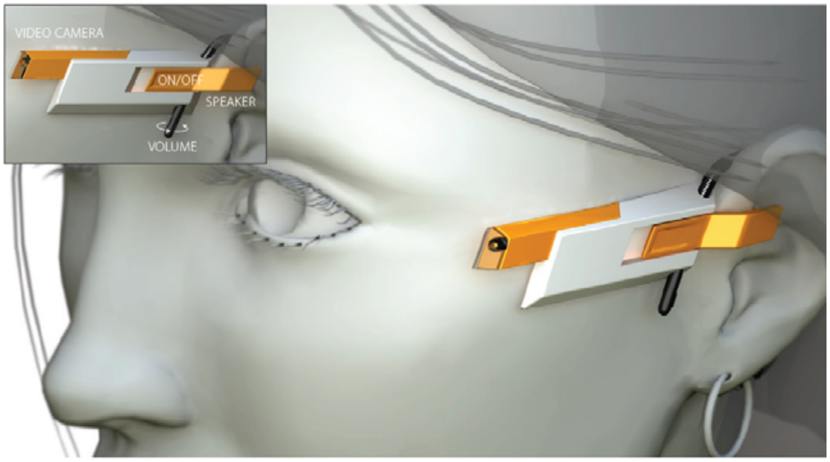

Fig. 3: An ear-mounted video camera and sound emitter

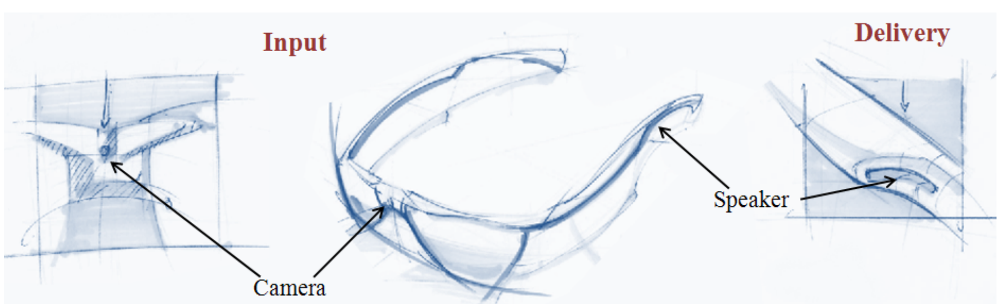

Fig. 4: A tiny video camera and sound emitter in a pair of glasses

6.1 Prototype Hardware

We mounted an analog camera ($1/3$“ CCD, 0.2 Lux light sensitivity, 92° field of view and 320 line scan NTSC output) on the nose bridge of a pair of sunglasses. The analog data were converted to digital using a video digitizer that connected to the computing element (a UMPC) through a USB port. The audio output (conveying processed information) was delivered to the user through the earpiece of the glasses. The prototype device is shown in Fig. 5.

Fig. 5: The prototype Social Interaction Assistant: Hardware

6.2 Prototype Software

The prototype Social Interaction Assistant processes the video stream from the camera, and uses existing computer vision algorithms to analyze the video frames (Fig. 6). The frames are processed by a face detection algorithm [10] that identifies the location of faces within the camera's field of view, sending a covert audio signal to the user when a face is detected. This is shown in Fig. 7, which is a screen shot of the demonstration software used to illustrate the functioning of the system.

The audio signal generated from the face detection result helps the user to know when someone approaches, allows the user to choose whether to initiate a conversation, and helps the user determine when to make eye contact. At the user's option, the Social Interaction Assistant can send detected faces to a face recognition component [15] that analyzes the face, compares the resulting feature vector to a user-specific database, and speaks the identity to the user if there is a successful match. The system also provides a quantitative measure of its belief on the top matches in the database (as shown in Fig. 8), which may help the user assess the confidence of the recognition result. If the face in the scene is a new acquaintance (i.e. no high confidence match), the system captures

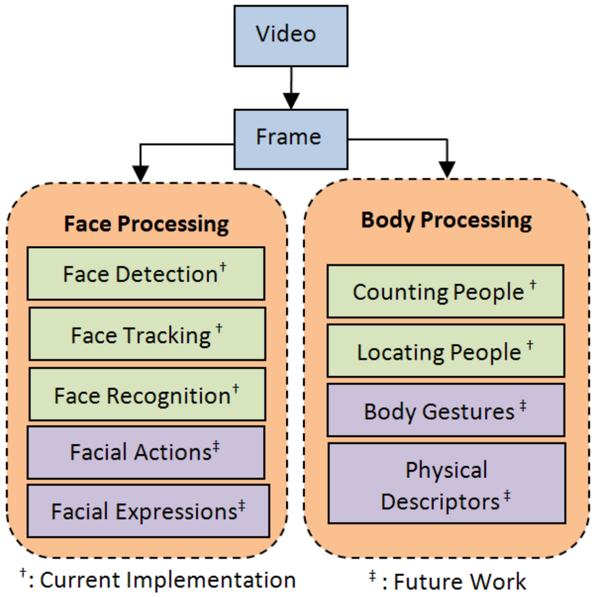

Fig. 6: The prototype Social Interaction Assistant: Software

Fig. 7: Screen shot of the demonstration software performing face detection; image captured from the camera on the prototype hardware.

25 images of the new face, and the user is later prompted to provide the name of the person. This combined information is then added to the database, for future use.

Fig. 8: Confidence levels in face recognition output.

7 Summary and Future Work

The Social Interaction Assistant described here is the result of a systematic exploration of the needs of individuals who are blind or visually impaired. As one member of our focus groups said, “*I appreciate the fact that we have a hand in designing it now, instead of it just being given to us, without a chance to see what we can do to make it better.*”

We believe that current and future computer vision techniques have the potential to meet many of the needs that we have enumerated. Advances in computing technology have allowed us to package our prototype system in a laptop that can be carried in a backpack. The current prototype (shown in Fig. 5) is intended as a proof-of-concept. It works well in a controlled laboratory setting, but it cannot handle drastic changes in pose angle or lighting conditions. More computer vision work will be needed to better handle these challenges. However, the needs that we have enumerated above provide a clearly defined target for such research. With respect to longer range goals, we are working on methods to recognize features such as hair color, facial expression, age, and gender.

We believe that a Social Interaction Assistant, such as the one described here, has the potential to transform the way that people with visual impairments conduct their daily lives, by improving their confidence and ability to engage in social interactions. In addition, we believe that ongoing work in this area will focus and extend computer vision research that will have important applications in many other domains.

References

1. Tachi, S., Tanie, K., Komoriya, K., Abe, M.: Electrocutaneous communication in a guide dog robot (meldog). IEEE Transactions on Biomedical Engineering **BME-32**(7) (1985) 461–469
2. Kulyukin, V., Gharpure, C., De Graw, N., Nicholson, J., Pavithran, S.: A robotic guide for the visually impaired in indoor environments. Proceedings of Conference on Rehabilitation Engineering and Assistive Technology Society of North America **21** (2004) 29–41
3. Kulyukin, V., Gharpure, C., Nicholson, J.: Robocart: toward robot-assisted navigation of grocery stores by the visually impaired. Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (Aug, 2005) 2845–2850
4. Shinohara, K.: Designing assistive technology for blind users. In: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility, ACM (2006) 293–294
5. Knapp, M.: Nonverbal communication in human interaction. Holt Rinehart and Winston Fort Worth (1992)
6. Wiener, W., Lawson, G.: Audition for the traveler who is visually impaired. Foundations of orientation and mobility **2** (1997) 104–169
7. Raver, S.A., Drash, P.W.: Increasing social skills training for visually impaired children. Education of the Visually Handicapped **19** (1988) 147–155
8. Burgoon, J.K., Buller, D.B., Guerrero, L.K.: Interpersonal deception: Ix. effects of social skill and nonverbal communication on deception success and detection accuracy. Journal of Language and Social Psychology **14** (September 1995) 289–311
9. : Low vision aids, low vision products, assistive technology from enhanced vision. `http://www.enhancedvision.com`
10. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition **1** (2001) 511–518
11. Zhao, W., Chellappa, R.: Face Processing: Advanced Modeling and Methods. Academic Press (October 2005)
12. Kakumanu, P., Makrogiannis, S., Bourbakis, N.: A survey of skin-color modeling and detection methods. Pattern Recognition **40**(3) (2007) 1106–1122
13. Murphy-Chutorian, E., Trivedi, M.: Head pose estimation in computer vision: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence (July, 2008)
14. Park, K., Lee, J., Kim, J. In: Facial and Eye Gaze Detection. (2008) 368–376
15. Turk, M., Pentland, A.: Face recognition using eigenfaces. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (1991) 586–591
16. Moeslund, T., Hilton, A., Krüger, V.: A survey of advances in vision-based human motion capture and analysis. Computer Vision and Image Understanding **104**(2-3) (2006) 90–126
17. Mori, G., Ren, X., Efros, A., Malik, J.: Recovering human body configurations: combining segmentation and recognition. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition **2** (2004) 326–333