Systematic Analysis of non-verbal cue requirements and design of the Social Interaction Assistant

From Research Question 1, it is certain that there is no analysis on the needs of the visually impaired and the blind community about their needs during social interactions. Though rsearch supports the need for social interactions, no efforts have been taken towards determining the specific necessities that this community has. In order to identify the unmet needs of the visually impaired community, two focus groups consisting primarily of people who are blind, as well as disability specialists and parents of students with visual impairment and blindness where conducted[[1]](#footnote-1). 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[[2]](#footnote-2).

As an example of the type of social disconnect that people who are visually impaired face, 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 to-wards 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).

While various other examples were cited by individuals during these focus group studies, the inability to access non-verbal cues were considered of highest priority. In this chapter, we introduce a self-report survey that was conducted based on the focus group study results that highlight the various non-verbal cues that are considered important from the perspective of the user population. Further, with the non-verbal cue priority list determine, the design of a platform that can be used for extracting and delivering these non-verbal cues is presented.

# Requirements for a Social Interaction Assistant

Based on the discussions conducted through the two focus groups, a list of needs was complied that characterized social needs often experienced by people with visual impairments. In doing so, two important aspects of social interaction were identified. These included

1. Access to the non-verbal cues of others during social interactions, and
2. How one is perceived by others during social interactions.

These needs correlated with the psychology studies conducted by Jindal-Snape with children who were visually impaired. She identifies these two needs under the *Social Learning* and *Social Feedback*. As discussed in Chapter XXX, Section XXX, these are the two important aspects of providing assistance and rehabilitation for people who are blind and visually impaired. While these two important categories were identified, for simplicity, the non-verbal cue needs were reduced to 8 aspects of social interactions that focused primarily on the physical characteristics of the interaction partner and the behaviors of the interaction partner. These questions were developed with the help of visually impaired professionals and students:

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.

While these 8 aspects of social interaction were important from the perspective of enriching social interactions of the people who are blind or visually impaired, it was not sufficient to just identify them, but it is important to determine the relative importance of these needs with respect to each other. To this end, an online survey was carried out to determine a self-report importance map of the various non-verbal cues. This list of questions included both the importance from the perspective of allowing access to the non-verbal cues of the interaction partner (for enabling Social Learning), while also focusing on the personal body mannerism (for enabling Social Feedback) of the individual.

# Online Survey

The online survey was anonymously completed by 28 people, of whom 16 were blind, 9 had low vision, and 3 were sighted specialists in the area of visual impairment and vocational training. 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, the metrics being

(1) Strongly disagree,

(2) Disagree,

(3) Neutral,

(4) Agree, and

(5) Strongly agree

The survey can be analyzed as having 3 groups (individuals who are blind, individuals with visual impairment and specialists with 20/20 vision) and 8 question groups each corresponding to the 8 aspects of social interactions that were identified from our focus group.

# Results:

## Mean Score Table:

Table 1 shows the eight aspects of social interactions, sorted by descending importance, as indicated by the survey respondents (the question numbers correspond to the need listed in the previous section). The mean score is the average of the respondents on the 5 point scale that was used to capture the opinions. A score closer to 5 implies that the respondents strongly agree with a certain question and that they consider inaccessibility to that particular non-verbal cue to be important deterrent to their social interactions. On the other hand, a score closer to 1 represents the respondent did not consider the access to a specific non-verbal cue to be important during their social interactions.

|  |  |  |
| --- | --- | --- |
| **Need** | **The Question** | **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

## Histogram of Responses:

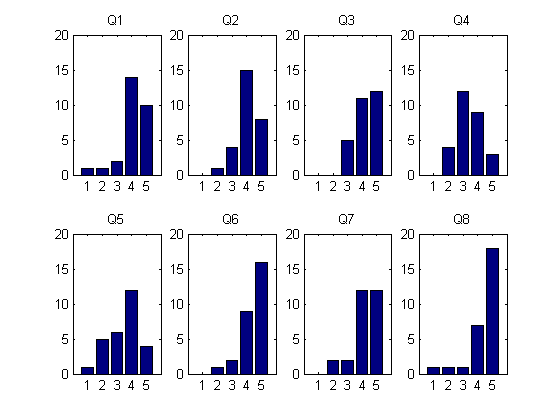


Figure 2: Histogram of Responses grouped by Questions

Figure 2 shows the histogram of responses for the 8 Questions that were asked as part of the survey. Each subplot refers to a single question and shows the number of times users responded to that particular question with answers from 1 to 5 on the Lickert Scale. Each histogram adds up to a total of 28 that corresponds to the 28 participants that took part in the online survey.

## Box Plot Analysis:

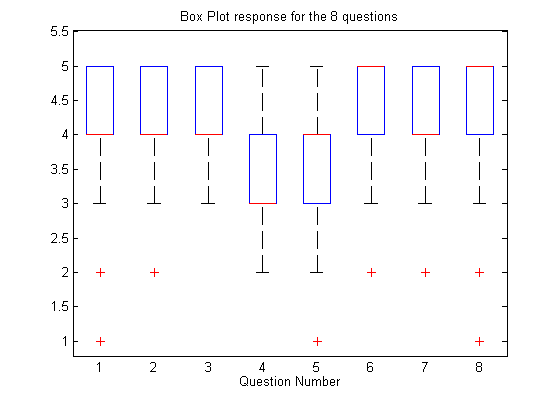


Figure 3: Box Plot of user responses for the 8 questions on the survey

The Box Plot of the 8 question responses is shown in Figure 3. The median values of the responses are shown as red lines for each of the 8 questions. While the blue box shows the enclosure for all responses between the 75 percentile and the 25 percentile points. Since the responses were on an integer scale of 1 through 5, the median coincides with the upper or lower 25 percentile. The whisker corresponds to the upper and lower limit of the values represented under that particular question. The plus marks represent any outliers under each question. Outliers are indentified based on whether they are outside the 3 sigma (variance) from the mean value. Note that the median value for questions 6 and 8 are at 5, median value of 4 for questions 1, 2, 3, 5 and 7, and median of 3 for question 4. Historically, Licket Scale data has been analyzed using Box Plot analysis as the plot captures all the descriptive statistics of minimum value, maximum value, median, variance, and the inter quartile range that encapsulates the 50 percentile of the data around the mean.

## Questionnaire Bias:

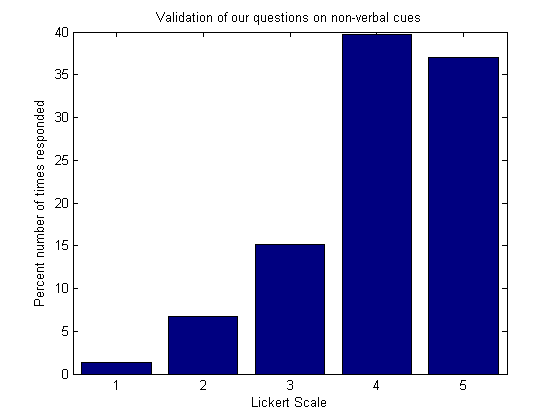


Figure 4: Questionnaire Bias

Figure 4 shows the number of times the users chose to answer the 8 questions with their agreement or disagreement. The y-axis has been normalized to 100 points. The graph shows that users chose to answer the most with agreeing with the 8 questions. Followed closely behind was the strong agreement with the questions asked in the survey. The respondents chose to answer the least through strong disagreement to what was asked in the survey.

## Rank Average and F-score:

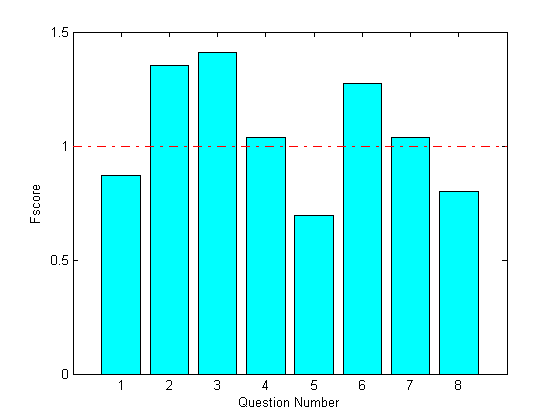
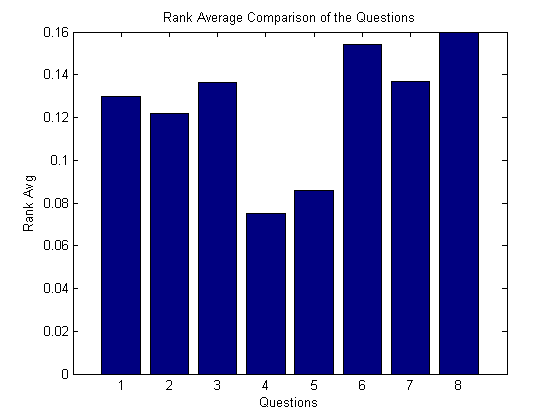
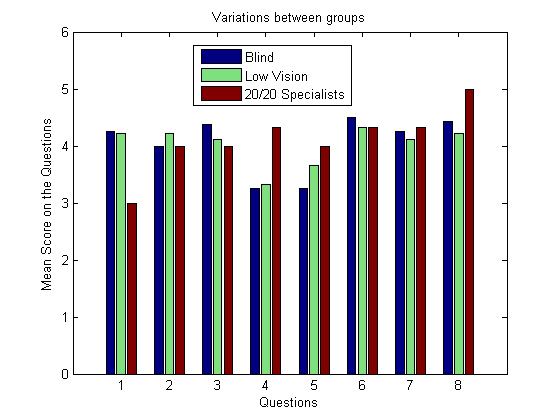


Figure 5: (a) Rank Average of responses for the 8 questions asked in the online survey. (b) F Score for the 8 questions with unity line marked with a dash-dot line.

As can be seen from Figure 4, the questionnaires were biased due to the focus group studies that lead to the questions themselves which were then answered in the online survey. Because of this bias, just the mean of the Lickert Scale will immediately show the same bias. This is due to the Gaussian iid assumption that is made while extracting the mean for the answers. In order to overcome this non-Gaussianity that is introduced due to the bias, we resort to non-parametric mean for the responses. Rank average of the responses is estimated instead of the typical mean of the responses for each of the questions. Procedure for estimation of the rank average is shown below:

1. Rank all data from all (question) groups together; i.e., rank the data from 1 to N ignoring group membership. Assign any tied values the average of the ranks they would have received had they not been tied.
2. Rank Average for each group is then given as

## Average Per Group:



# Analysis of the survey responses

## Histogram of the responses:

A histogram of responses is shown in Figure 1 below. From the average score of the participants who took the online survey, it can be seen that,

1. Respondents are highly concerned about how they are perceived by their sighted peers (based on the response to Question 8 on the survey).
2. Facial expressions form the most important visual non-verbal cue that individuals who are blind or visually impaired feel they do not have access to. This correlates with the studies into non-verbal communication that highlights the importance of facial cues that requires visual decoding (based on Question 6 on the survey).
3. Followed by facial expressions, body mannerisms seem to be of higher importance for individuals who are blind and visually impaired. This can be correlated to the table shown Chapter XXX, Section XXX, where body follows the face in terms of displaying non-verbal cues (based on Question 3 of the survey).
4. 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.

As described earlier, the 8 questions corresponding to the social needs of the individuals were identified from the focus group survey that was conducted. Thus, the questions presented in the online survey questions were biased towards the needs of everyday social interactions of individuals who are blind and visually impaired. In order to under

# Alternative Sensing 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).

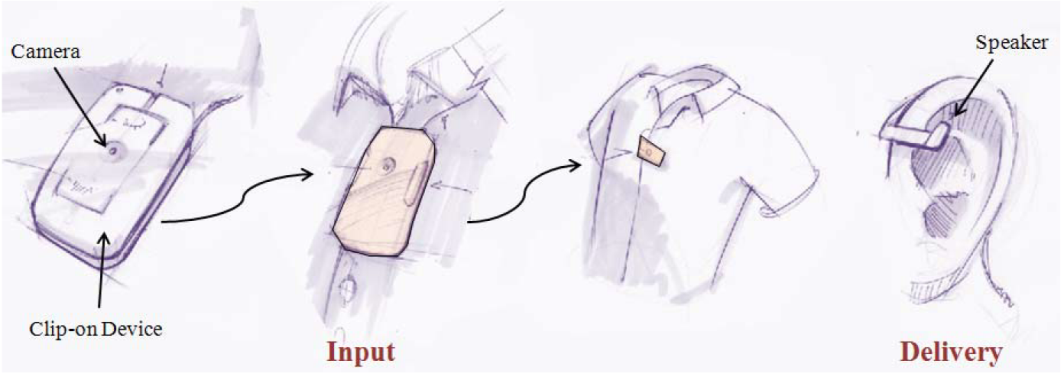


Figure XXX: A clip on camera and small speaker

## 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).

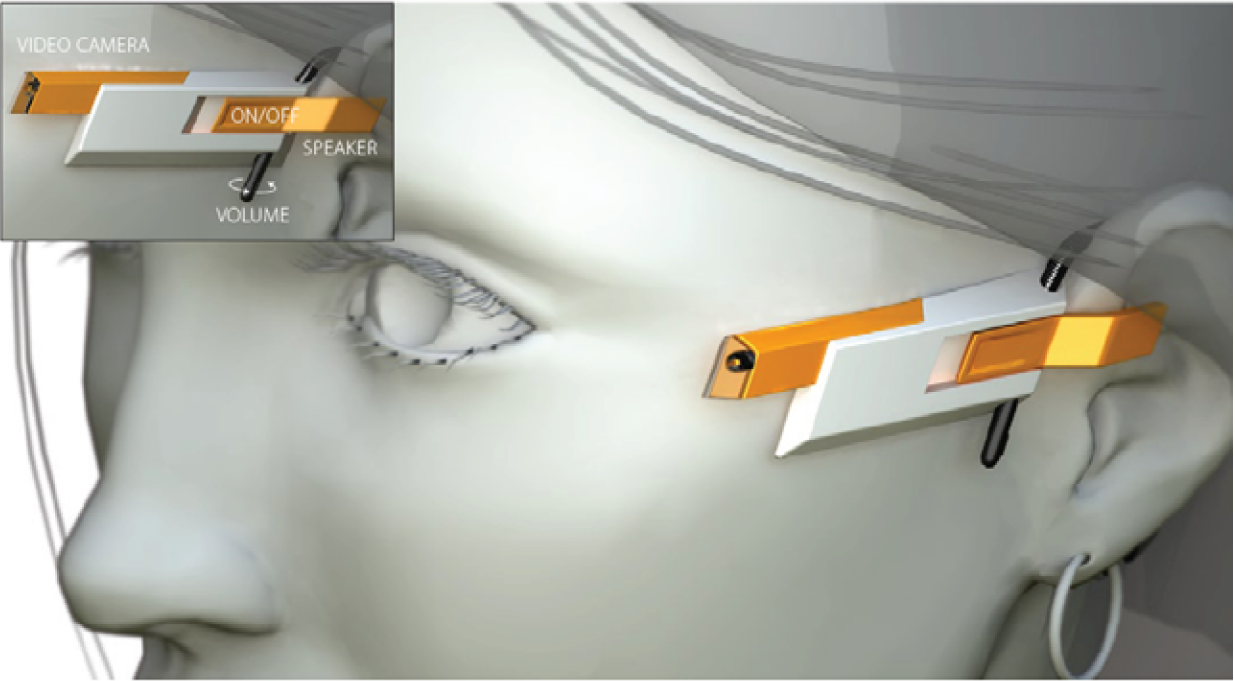


Figure XXX: A ear-mounted video camera and speaker

## 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).

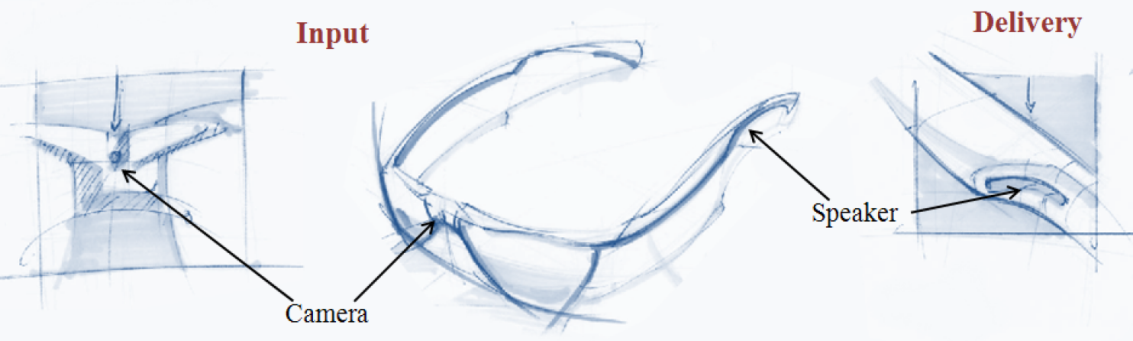


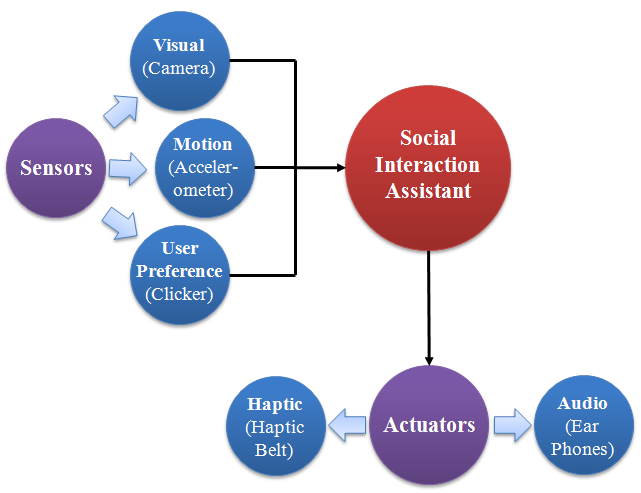
Figure XXX: A tiny video camera and speaker on a pair of glasses.

# Social Interaction Assistant:

Having analyzed the requirements and considering the various components of the sensing and delivery technology, we resorted to Concept 3 and incorporated the important aspects of egocentric and exocentric sensing into the prototype system.

## System Architecture

The system level architecture of the proposed social interaction assistant is shown in the Figure 1. The sensor suite consists of: (1) a visual sensor (1.4 Megapixel camera), (2) a motion sensor (± 12g accelerometer), and (3) a 5-button clicker, which serves as a user interface. The social interaction assistant software (implemented on a Windows Operating System PDA) uses these sensors to collect information about the various social and behavioral mannerisms of the user and participants in the vicinity of the user. Interpretations of the social interactions generated by custom algorithms are communicated to the user through an actuator suite, consisting of: (1) a haptic belt and (2) a set of ear phones. The haptic belt encodes information in the form of vibrotactile cues, while the ear phones provide short audio cues.



System level architecture of the Social Interaction Assistant

## Prototype System:

Figure 2 shows the implementation of the proposed Social Interaction Assistant. A tiny video camera is placed unobtrusively on a pair of glasses, and a tiny state-of-the-art accelerometer is placed unobtrusively in a hat, and is used to monitor the user’s body mannerisms – particularly those related to head movement. (Most communicative gestures are encoded in movements of head and the most widely occurring and problematic stereotypic body mannerisms are done with the head.) The accelerometer operates on a coin battery that allows for uninterrupted operation for over 4 hours. The user uses the 5-button clicker to control what types of information are delivered by the system. The haptic belt can be worn under the clothing, and the earphones are worn discretely under their hat. Thus, the proposed design of the assistive technology is (1) wearable, (2) portable, (3) unobtrusive, (4) self and other sensing, and (5) can be worn by the user for extended periods of time.

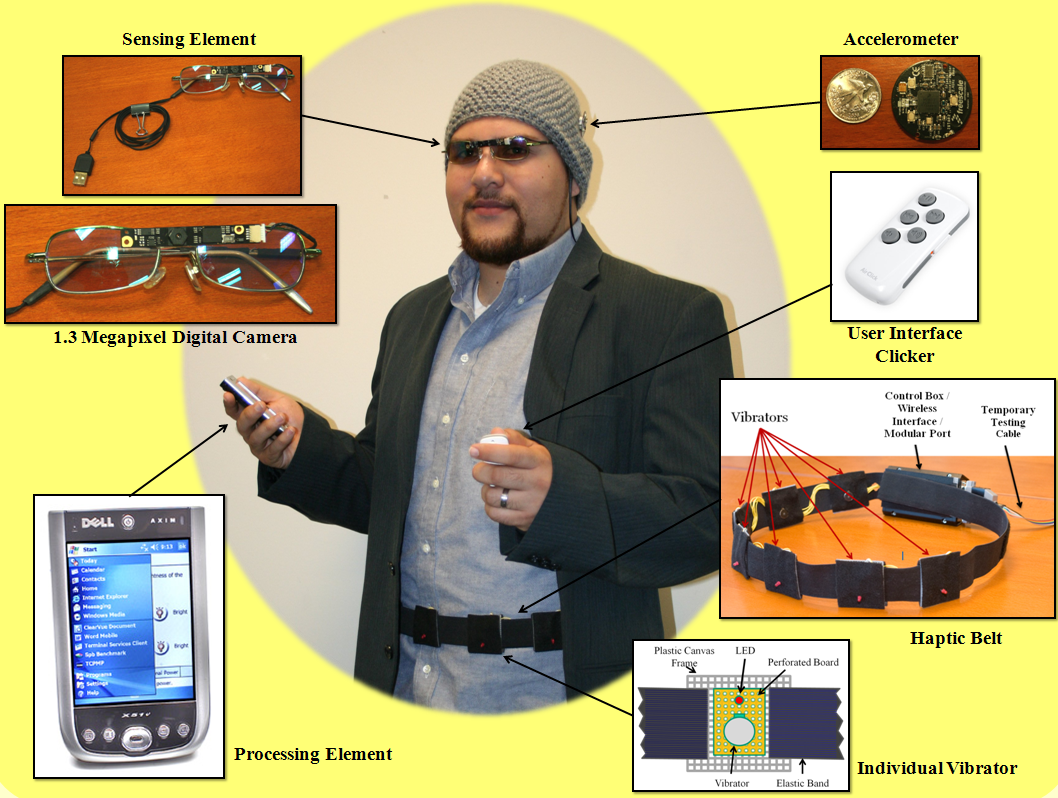


Figure 2: the implementation of the Social Interaction Assistant

1. In order to understand the assistive technology requirements of people who are blind, we conducted two focus group studies (one in Tempe, Arizona USA - 9 participants, and another in Tucson, Arizona USA - 11 participants) which included:

   1. students and adult professionals who are blind,

   2. parents of individuals who are blind

   3. professionals who work in the area of blindness and visual impairments.

   There was unanimous agreement among participants that a technology that would help people with visual impairment to recognize people or hear them described would signiﬁcantly enhance their social life. [↑](#footnote-ref-1)
2. To quote some candidates opinion about face recognition technology in a social setting:

   • “It would be nice to walk into a room and immediately get to know who are all in front of me before they start a conversation”.

   One young man said, “It would be great to walk into a bar and identify beautiful women”. [↑](#footnote-ref-2)