Project Description

1. Introduction

Civilization and society has been deemed as the greatest evolutionary progress in human history, creating a social system where every individual is trained, exposed and required to operate on a complex social plane, the governing rules of which are as hazy as the mortar that holds societies together. This eventual metamorphosis happened due to the fact that humans are social animals. We need social connectivity with our fellow citizen to coexist and cooperate within the society. Such connectivity is made possible mainly through interpersonal communication, which in large parts includes verbal and non-verbal interactions. From a quantitative analysis, 35% of our interpersonal interactions are verbal speech, while nearly 65% of communication happens through non-verbal cues [1] [2]. In a typical bilateral interaction, while speech encodes most of the information, non-verbal cues facilitate the elegant delivery, interpretation and exchange of the verbal message. For instance, eye gaze, iconic body or hand gestures, and prosody (stress and intonation of speech) are essential for seamless, synchronized exchanges in social interpersonal interactions. On a day-to-day basis, people communicate so effortlessly through intertwined verbal and non-verbal cues that they do not realize the complex interplay of their voice, face and body. The complexity and importance become evident when one (or more) of the participants in a social interaction is unable to interpret their interaction counterparts due to sensory disability, such as blindness and low vision. 72% of non-verbal communication (Figure 1) is encoded through visual channels including the human face and body. Expressions, mannerisms and behaviors of interaction partners form a significant portion of non-verbal cueing which an individual with blindness or visual impairment is unable to access. This inhibition impedes with his/her ability to make appropriate judgments leading to a sense of social isolation or diminished returns in any social exchange. This, in turn, affects the way one conceptualizes himself/herself as part of a society that imposes social needs and requirements. From learning social skills to utilizing them, individuals who are blind face fundamental social challenges in coping with the society. This proposal addresses this fundamental challenge that is largely unexplored.

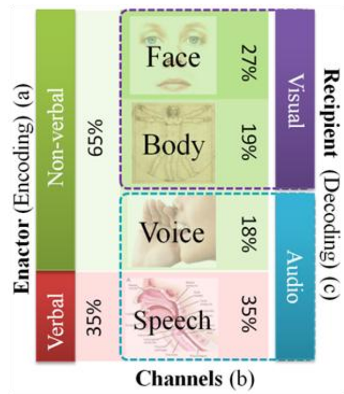


Figure 1: Components of human bilateral communication.

Human interpersonal communication involves perception and cognition of the actions and speech of the interacting individuals. The study and understanding of social interaction is, unfortunately, a very difficult process – since even simple everyday social interactions involve complex sensory, motor and cognitive tasks. For instance, what seems to be a simple act of shaking hands between two individuals is a sequential cognition of several sensorimotor events. Two individuals who engage in shaking hands first engage in eye contact; then exchange the desire to interact (usually through face and body gestures, such as smile and increased upper body movements); determine the distance between themselves; move towards each other maintaining an appropriate interpersonal distance; engage in shaking hands; and finally, move apart assuming a conversational distance which is usually wider than the handshake distance. Evidently, people who are blind and visually impaired are at a loss in most of these steps. Compensating for their deficit, they usually initiate a handshake by standing at one place and extending their arm in a handshake posture in the direction where they hear most people conversing, hoping to draw the attention of their sighted counterparts. Unfortunately, when there is a group of individuals who are all interacting among themselves, the individual who is blind may not have sufficient information to determine if any of the sighted counterparts has noticed his/her intent to make social contact. This example not only illustrates the inherent complexity of the problem, but also demonstrates the significance and need for (currently non-existing) assistive social aids in enriching the experienced quality of life for individuals with disabilities.

In order to address this basic human need, this proposal seeks to develop a novel next-generation assistive social situational awareness aid for individuals with visual impairments. Based on our past successes in designing, prototyping, developing, evaluating and deploying assistive technologies for individuals who are blind or visually impaired, as well as motivated by our recent explorations on addressing the social needs of these individuals in interpersonal interactions, we plan to develop an experiential technology solution that provides real-time access to non-verbal communication cues in dyadic (one-on-one) interactions. Understanding the nuances of social actions, behaviors and perceptions towards developing assistive technologies, requires cross-cutting specializations within computational science and beyond, including affective computing, human communication engineering, human behavioral modeling, human-machine interfaces, usability engineering, assistive technology design and development, and multimedia computing. We have brought together an interdisciplinary team which has a proven track record and prior successes in ubiquitous and embodied multimedia computing, human-computer interfaces, haptics, machine intelligence (Panchanathan); human communication studies, non-verbal communication (Ramirez); and usability engineering, disability studies (Hedgpeth). The team also includes a postdoctoral researcher (Balasubramanian) and a research engineer (Krishna), each of who has worked extensively in the developing and applying computational methods and technologies that are designed for individuals with disabilities. The team will also comprise of 2 graduate students and several STEM senior undergraduate students who will participate through Capstone projects. Our collaborator and community partner, Arizona Center for the Blind and Visually Impaired (ACBVI), have vouched their strong support and commitment to the success of this project (please see letter of support).

2. Motivation

Social interactions are an essential component of both personal and professional growth amongst individuals. From one's home to workplace, people require a socially conducive environment, the lack of which can lead to psychological stress [3] and finally, in an urge for the individual to disconnect from the society. Studies have shown that people who are socially connected with their work environment tend to show higher level of work satisfaction and increased productivity. Unfortunately, the social disability faced by individuals who are blind could isolate them from their social work environments. According to AFB statistics, 80% of individuals who are blind have the potential to be employed with appropriate technological assistance that promotes independence in their daily activities. While current assistive technologies focus on providing access to gadgets and similar inanimate entities, no technology focuses on the more fundamental problem of providing access to the social environment in which one lives and works. While existing assistive technologies are focused on problems such as navigation, text reading, computer/web access and everyday appliance/computer access, little or no work has been focused at the real-time accessibility to social and behavioral cues, or to promote access to relationships with fellow humans. *The proposed research intends to pioneer the development of social assistive aids that can create a level playing field for all individuals, by reducing the social disability associated with individuals with sensory/cognitive disabilities*. Reducing the need for formal care services using such assistive aids could save billions of dollars annually in the US and relieve caregivers of the "free care" they provide at the sacrifice of their own time and health.

From the studies conducted on the effects of visual impairment on social interactions, especially experiments done on children [4] [5], individuals with visual impairments face two distinct levels of difficulties: (i) the inability to decode visual non-verbal cues, and (ii) the missed opportunity to immerse oneself into new socio-cultural environments. While we discussed the former challenge earlier, the second aspect is equally important. Humans are in a constant state of learning through interpersonal communication. Cultures of a geographical location, a work place, or even a group of friends impose a requirement on individuals to constantly learn from their social interactions. As an example, people who migrate from the East Asian countries to the US are overwhelmed by the strong upper body movements that are demonstrated typically in the West, but learn over time by close observation. Unfortunately, individuals who are blind may lack this essential aspect of social feedback. Similarly, people with visual impairments are unable to visually inspect their own body movements (self-monitoring) and learn what body movements are appropriate in a particular cultural setting. For example, crossing legs, leaning back on a chair, moving ones hand within their own personal space are all well-suited body behaviors for a professional work environment in the US, while none of these would be considered appropriate in Japan. Hence, people with visual impairments tend not to display much body movement with the trepidation of displaying inappropriate body mannerisms. Unfortunately, these factors add to the general feeling of the target population as being asocial, and could lead to social isolation.

In our past work, two focus group studies were carried out with people who are blind and visually impaired in Phoenix, AZ and Tucson, AZ in 2001, which revealed a basic need for access to social cues in everyday interactions. In [6], we identified 8 important social needs for individuals who are blind and visually impaired and rank ordered them through an online survey (Figure 2). This list shows that participants' most important need is the feedback on their own body mannerisms and how this affects their social interactions. Following this was the need to access facial expressions, identity, body mannerisms, eye gaze, proxemics and appearance of their social interaction partners, in the presented order. In a $43.8 billion market flooded with choices on various assistive technologies (projected 2013 market share for disability and elderly assistive technologies as reported by independent market analysis firm BCC Research, report #HLC047B, Published: October 2008), ***no*** device or technology delivers any form of social assistance with one's everyday interactions. Such a trend in assistive technology development can be traced to the common practice of adopting the latest state-of-the-art sensor technologies in addressing classic problems of assistance making the problem technology centric and not focused on the users. An example is the laser range finder on the white cane that translates distances in front of the cane to auditory tones (without considering the fact that the target population senses their surroundings primarily through audio). This proposal seeks to address this market gap by developing next-generation social assistive aids for users with visual impairments – and broadly speaking, for users from a larger society that will benefit from these advancements in human communication mediation and assistance.

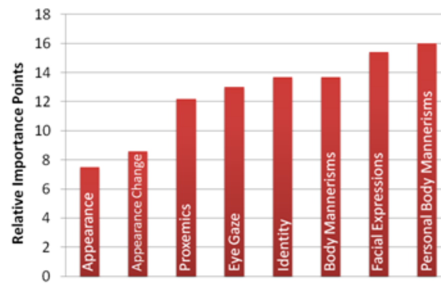


Figure 2: Rank-ordered list of social needs of individuals with visual impairments [6]

3. Approach and Significance

The focused research of developing social assistive technologies for people who are blind and visually impaired poses a broad and interesting spectrum of research questions including: (a) What components of the social interactions are diminished due to the sensory visual loss? (b) What coping mechanisms have evolved in the target population to deal with the vision loss? (c) What can be understood by studying the coping mechanisms in light of the difficulties faced by the target population? (d) What are the person-specific dimensions of user needs in everyday social interaction? (e) How can the implicit and explicit needs of an individual be identified towards enriching one's social interactions? (f) What actuators will offer the best sensory substitution for vision loss (within the scope of social interactions)? (g) What form factors of assistive technologies will effectively sense and deliver social cues in a subtle and unobtrusive manner? Answering some or all of these research questions promises impact on the broad area of enriching human interpersonal communications, whether the individuals involved are disabled or not.

Enriching Dyadic Interactions: Literature on human communication models classifies interpersonal interactions broadly into: *dyadic interactions*, and *group interactions*. Theories of leadership, conflict and trust base their findings on dyadic interaction primitives where the importance of the various non-verbal cues is heightened due to the one-on-one nature of dyadic interactions. Eye contact, head gestures (nod and shake), body posture (conveying dominance or submissiveness), social touch (hand shake, shoulder pat, hug, etc.), facial expressions and mannerisms (smile, surprise, inquiry, etc.), eye gestures (threatened gaze, inquisitive gaze, etc.) are some of the most important parameters that are studied closely in dyadic understanding of human bilateral communication [7]. Literature on group dynamics, in turn, builds on top of these bilateral dyadic communication models towards deriving each individual's contribution to the group. Hence, in this proposal, *we focus our efforts on developing an assistive technology aid that will enrich dyadic (one-on-one) social interactions of a user who is blind with their sighted counterpart. The proposed device will integrate computer vision and machine intelligence components that process data streamed from a miniature pan-tilt-zoom camera to extract the interaction partner's face and head-based communicative gestures. These gestures will then be delivered to the user via a high bandwidth visuo-haptic sensory substitution system.*

Sense-Enrich-Deliver Paradigm: Existing affective computing research largely follows a sense-decide-react model, i.e. sense the non-verbal cues, decide on the affective state of the user, and react through robot/avatar manipulation where necessary. In comparison, *we intend to propose a sense-enrich-deliver model, i.e. sense the non-verbal cues, enrich them based on the user's disability and specific choice, and deliver the non-verbal cue data in auxiliary sensory modalities that the user has access to.* Thus, while affective computing focuses on socio-emotional decision making and visual reproduction of affective cues, this research focuses on effective, low cognitive load sensory substitution of visual affective cues.

Person-Centered Design: Another essential component of assistive technology solutions is to cater to the varying needs of an individual user. Vision disability, like any other disability, is not a black-or-white issue, and users lie on a spectrum from normalcy (20/20 sight) to complete blindness with varying levels of optical, neurological and perceptive capabilities in between. Hence, it becomes imperative that assistive technologies are designed to accommodate the adaptation of features based on an individual user's specific abilities, needs and preferences. Also, from our past studies, we have found that users have remarkable learning capacities in using assistive devices, as they evolve to accommodate the device in their day-to-day life. This generates varying requirements even in a single user's experience of the technology. Target users also express a strong requirement to be involved in decision-making, and not let technology make decisions on their behalf. Hence, it becomes essential that assistive devices only provide access to otherwise inaccessible information, and let users decide the course of action. In order to address these issues, we intend to develop an assistive aid that operates in four distinct levels of sensory substitution on the social communicative cues, allowing the user to choose a suitable level based on context and convenience. These four levels of information vary from literal translation (where movements of the face and head are translated as is into a haptic representation) to purely symbolic representations (where iconic spatio-temporal patterns are delivered to represent higher level semantics – such as emotions - of face and head movements). More details of this approach are presented in Section 5. As the user adapts to the device, he/she could choose to dynamically change the level of detail of the social information they receive. For example, in an interview setting, a user may choose to receive all details of facial movements to interpret carefully; while in a casual setting, a user may choose to receive high-level semantic labels such as happy, sad, surprised or angry.

Broader Significance: The proposed research focuses on the development of social assistive technologies that will have widespread impact on a large portion of society, including individuals with other sensory and cognitive disabilities. The increasing numbers of the elderly population have led to a significant increase in the incidence of disability, thus making this research timely, important and widely relevant. Apart from the social disability faced by individuals with visual impairments, studies [8] have shown remarkable reduction in the symptoms of Alzheimer's and dementia patients when their social activity level is increased. Further, such social assistive aids could be useful for patients with Autism Spectrum Disorders, both in facilitating a richer social experience, as well as in serving as a social learning instrument. Given the increasing nature of virtuality in everyday interpersonal interactions and a peaking reliance on remote interactions for personal and work communications, this technology solution could prove to be a pivotal starting point for a new generation of enabling social technologies.

4. Results from Prior NSF Support

In the past, the PI has been successful in the design, development and deployment of innovative assistive ubiquitous technologies and systems, including assistive technologies and smart environments, through support from various NSF programs, as listed below. A detailed report of the results and achievements, along with the publications that resulted from these awards, are provided in Appendix A. A brief description of relevant NSF-funded projects follows.

Title: ITR: iLearn: IT-enabled Ubiquitous Access to Educational Opportunities for Blind Students (Award Id: IIS 0326544), **Amount:** $1,112,212, **PIs:** S. Panchanathan, K. S. Candan, F. Golshani, T.Hedgpeth, P. Green, **Duration:** Sep 2003–Aug 2008. **Description:** The iLearn project was focused on the development of a suite of assistive systems to help students with visual impairments succeed in their educational endeavors. These assistive systems included the *iLearn Reader*, the *iLearn NoteTaking Device,* and the *iLearn Information Assistant*. The success of the *iLearn* systems led to the expansion of its scope to support other activities of everyday living under the umbrella of the broader iCARE (**i**nformation Technology **C**entric **A**ssistive and **R**ehabilitative **E**nvironments) project, wherein a pilot prototype of the *iCARE Interaction Assistant* was designed and developed. The Interaction Assistant was intended to enhance the social interaction experiences of individuals with sensory/perceptual and cognitive disabilities, such as blindness and autism. More details of the iLearn projects are available in the NSF ITR final report at http://cubic.asu.edu/finalNSF-ITRReport.pdf. Two post-doctoral researchers, 5 PhD, 8 Masters and more than 10 undergraduate students were trained in human-centered ubiquitous technologies as part of this program. More importantly, 6 undergraduate students, who are blind or visually impaired, actively participated in the design, development and evaluation of the devices. These students were trained in the design of assistive technologies, and have authored/co-authored several of our publications in this project. This program also resulted in newer NSF awards such as a HRD award

(PPD-FRI: Ubiquitous Environment to Facilitate Engineering Education for Blind Persons, Award Id: 0333452), a Division of Undergraduate Education (DUE) award (MAISON: Middleware for Accessible Information Spaces on NSDL, Award Id: 0735014) and an Arizona Rehabilitation Services Administration (RSA) award (Ubiquitous Environment to Facilitate Access to Textbooks and Related Materials for Adults and School Age children who are Blind or Visually Impaired).

Other relevant NSF awards (details in Appendix A):

Title: HCC: The iCare Ambient Interactive Shopping Environment (Award Id: 0739744), **Amount:** $156,927, **PI:** T. Hedgpeth, S. Panchanathan, **Duration:** Oct 2007 - Sep 2008.
Title: SGER: Incorporation of a psychological basis of haptics in the design of assistive haptic user interfaces (Award Id: 0554698), **Amount:** $196,574, **PIs:** S. Panchanathan, D Homa, D. Hansford, T. Hedgpeth, **Duration:** Nov 2005-Oct 2006.
Title: NSF-PPD-FRI: Ubiquitous Environment to Facilitate Engineering Education for Blind Persons (Award Id: HRD 0333452), **Amount:** $172,538, **PIs:** S. Panchanathan, K. S. Candan, M. E. Donderler, T.Hedgpeth, and P. Green, **Duration:** Oct 2003–Sep 2005.

5. Research Plan

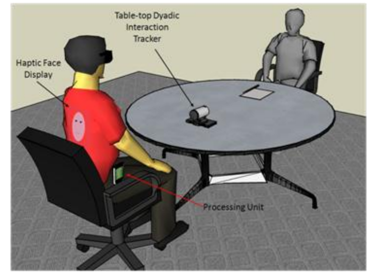
Our research plan focuses on four important components: (1) the dyadic mediation interface; (2) the suite of algorithms for extracting and understanding non-verbal communicative cues; (3) the haptic delivery system that translates visual non-verbal social cues to effective haptic cues; and (4) the real-world evaluation with the target user population. Unlike typical everyday assistive technologies, like web accessibility technologies or access to daily use appliances, concept of social assistance through technologies is rather new and unexplored, posing complex design challenges. This is owing to the fact that the technology has to be designed as a mediator between two individuals and thus requires that it does not interfere with the natural social communication of either of the participants. Unlike technologies (like Self-Cam [9], where a camera is supported on the user's shoulders to self-monitor facial expressions in autistic individuals) which can be considered socially disruptive, the proposed solution takes into account the social artifact of assistive technologies into design consideration. Our human-in-the-loop approach to translate research outcomes to technology solutions will be implemented through modular iterative and incremental deliverables that are regularly evaluated by the target user population.

Disability is an individual-specific concept. While the impacting factors causing disabilities may be common across groups of people, the experience of disability and its debilitating effects on daily living are unique to each individual. As mentioned earlier, to address this issue, we plan to develop an assistive aid that operates in four distinct levels of sensory substitution on the social communicative cues, providing the user with options to choose a suitable level based on context and convenience. The adjoining table shows the four levels of representation with corresponding examples of the level of information that is delivered through haptic cues. Moving from *Literal* to *Symbolic*, the haptic cueing bandwidth reduces, while the decisions made by the system increases. The device will also allow the users to move from symbolic to literal (or vice versa) in real-time based on the need. A user can thus choose desired levels of mediation considering the cognitive and computational loads in a particular space-time context. We believe that this research can serve as a pathway to the next generation of assistive technologies that provide access to complex social signals, and also serve as a platform for fundamental advancements in related application domains such as human computer interaction, remote interactions, robotics, haptic interfaces, and embodied computing. Current trends in assistive technologies require users to carry ever more technologies on person, and users who are blind have expressed their concern in our focus group studies on carrying too many technology artifacts with them all the time, be it assistive or not. Hence, at every stage of the implementation, care will be taken to ensure that the proposed dyadic communication aid, or essential components of it, can be built on top of existing wearable computing technologies that are in everyday use, like cellular phones. This will address user concerns about carrying too many devices, as newer technology solutions are developed.

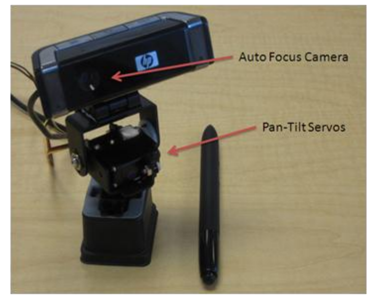
Literal	Facial features and fiducial points as is
Semi-Literal	Lip curved upwards, eye brow curve
Semi-Symbolic	Smile, eyes wide open
Symbolic	Happy

5.1 Research Objective 1: The Dyadic Interpersonal Mediation interface

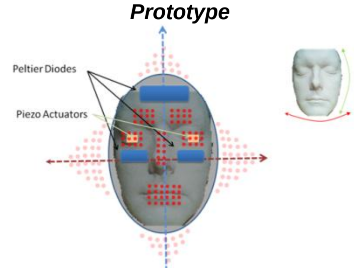
Camera-based affective interfaces have been explored in the past towards developing new human computer interaction paradigms. Though critical advances have been made in this front, mobile affect solutions towards effective real-time visuo-haptic sensory substitution are far from reality. The core of this proposal focuses on adopting a practical solution-centric approach to push the envelope in mobile social communication technologies. The proposed dyadic interpersonal mediation device will be investigated by PI Panchanathan, Co-PI Hedgpeth, Research Engineer Krishna, and the graduate students. Suitable modules of engineering and software will be conceived as development projects for senior undergraduate hardware/software Capstone projects. In the past, we have found this engagement of undergraduate students in research activities to be of significant impact on undergraduate students (in being exposed to cutting-edge STEM research). An important outcome of this coordinated research objective would be the development of a dyadic interpersonal communication mediator platform that is acceptable for both the users and the interaction partners. Important aspects of the prototype include: (a) form factor; (b) use case in everyday settings; and (c) available computational resources.



(a) Concept Prototype



(b) Tabletop Face Tracker: Pilot Prototype



(c) Haptic Face Display

Figure 3: Dyadic Interpersonal Mediation Interface

As illustrated in Figure 3, we propose to develop a table-top camera based upper-body non-verbal communication (NVC) sensor combined with a wearable haptic delivery system, united through a mobile computing platform. Recently, researchers at CUbiC@ASU have demonstrated an embedded real-time face tracking system based on a miniature servo pan-tilt mechanism incorporating a CMOS auto-focus digital camera [10]. The system makes use of open source software libraries, OpenCV, towards real-time face detection implemented on the open source hardware platform, FriendlyARM. The detected faces are then communicated to a mobile computing platform (like a PDA or smart phone) for further NVC cue processing. This form factor was developed after extensive feedback from the blind and visually impaired community about the typical dyadic social interaction scenarios where such an assistive technology would be of greatest help. While our work has demonstrated the feasibility of our approach, several computational as well as design challenges lie ahead in realizing an effective dyadic social mediation technology solution, which is the primary focus of this proposal. Our pilot prototype (Figure 3b) used a pan-tilt mechanism to capture the interaction partner even under lateral movements during the communication. As part of this proposal, we will explore zoom controls in similar configurations to capture finer spatiotemporal features of the face and the head. In keeping with our objective to build technologies that have a broad appeal, we will also investigate engineering designs where the servo-based pan-tilt-zoom mechanism is available as a cradle, on which a smart phone with a camera can be placed. In the project’s later phases, we will explore possibilities to export the sensing framework of the dyadic mediator to popularly available cell phone platforms. Considering the recent orientation of technologies towards mobile applications (such as iPhone Apps, Android Apps, and more recently, the Chrome Web Store), we plan to develop modular deliverables, which in future, may be ported as mobile applications.

From a processing perspective, while initial evaluations will be conducted using computing elements such as laptops, we plan to exploit the computational power on existing technologies carried on-body by the user population to achieve our objectives. To this end, we intend to lead an undergraduate Capstone project in Fall 2011 and Spring 2012 to develop communication interfaces between the table-top head tracker and popularly available smart phone platforms like Android, Apple iOS and Windows Mobile. We will also utilize the services provided by the ASU iPhone and Android Open Source Club, which has weekly meetings for interested developers, to suitably augment our efforts in this regard. In ongoing work, we are studying cloud computing options to complement the hardware capabilities of a mobile device to offer a flexible, platform-independent and scalable solution. We are developing a cloud-based computer vision framework with a web service front end to allow access to the algorithms from any device with an Internet connection. These efforts will focus on providing real-time capabilities to the mediation interface.

While affect sensing has been extensively explored in computing research, technologies for delivering social communication cues is still a new field of research (please see [11] for a detailed review of existing interpersonal communication technologies). To this end, we propose to develop novel wearable interfaces that will support intuitive visuo-haptic sensory substitution (Section 5.3) to deliver the high bandwidth communicative cues. Recently, we have developed and demonstrated a visuo-haptic sensory substitution system for delivering facial expression data in the form of discrete haptic icons [12]. Vibrotactile stimulators located on the back of the user's hands (See Figure 6 in Section 5.3) were used for delivering spatio-temporal vibration patterns that were specifically designed for communicating interaction partner's facial expressions. While this study demonstrated the effectiveness of our cueing system for delivering social cues, our evaluations with users who are blind have shown that they prefer interfaces that are less obtrusive and less dense in their haptic actuator distributions. In fact, the users expressed the acceptance of such technologies, if they can be integrated seamlessly with their everyday clothing [13]. To this end, we plan to explore haptic form factors that can be worn on the back of the user in the form of a vest. As shown in Figure 3a, we will explore the region above the waist to develop vibrotactile displays that can communicate intricate motion patterns of the interaction partner's face and head. One manifestation of the proposed project is shown in Figure 3c as the Haptic Face Display (HFD). It is envisioned that the HFD will replicate the entire human face through haptic actuators. The first HFD prototype will be built with vibrotactile actuators, for reasons of past experience and feasibility of development/evaluation. However, we plan to build a second HFD prototype which will include vibratory motors, Peltier diodes and piezo-actuators, as shown in Figure 3c. The vibrators will be used to deliver real-time facial movements mapped as spatio-temporal vibrations; the Peltier diodes will be used for communicating complex socio-emotional concepts like blushing, anger etc.; and the piezo-actuators will mimic eye blinks. The actuators on the HFD will be organized in a configuration that essentially allows for delivering social cues at all four levels identified in this research – literal, semi-literal, semi-symbolic, and symbolic. While we plan to evaluate the back of the user, the HFD can easily be translated to any other somatosensory location on the body, and we will seek user inputs in this aspect. Details on the visuo-haptic substitution methods are discussed in Section 5.3, where we present our most recent research on haptic interfaces, and demonstrate how the proposed solution will leverage expertise at CUbiC in developing novel interpersonal interaction interfaces. The mediation system, consisting of sensing, processing and delivery elements, will be evaluated with students and older individuals who are blind and/or visually impaired. Our evaluation methods are discussed in detail in Section 5.4.

5.2 Research Objective 2: Extraction and Understanding of Non-verbal Communication Cues

Non-verbal communication (NVC) is a complex process that involves several channels of communication including body language, facial expressions, postures, gestures, and vocal tones [2]. Historically, affect recognition systems have tackled the problem of extracting facial non-verbal cues, especially expressions, from videos and still images [14]. However, NVC represents a larger problem scope, with challenges from an *understanding* perspective (e.g., what NVC cues convey consent/agreement? What cues convey happiness?) and an *extraction* perspective (e.g., how can an computing element recognize the shape of eyebrow, decode the movement of upper arm, or extract prosodic features). Very recently, researchers have started studying this largely uncharted domain of modeling human communicative behavior. For example, Vinciarelli [15], Pentland [16] [17] and other researchers in the emerging fields of social signal processing and social intelligence [18] have attempted to model human social signals to decode communicative intent, to synthesize social signals and behaviors, and to react appropriately to the user's communicative intent. Morency addressed problems such as gesture recognition [19], head movement/pose recognition [20] [21], and listener backchannel prediction [22] using context-driven sequential probabilistic models such as Hidden Markov Models and Conditional Random Fields. Given this recent interest, we consider the proposed research timely for three reasons: (i) Firstly, the enhanced abilities of advanced computing and communication technologies allow the development of real-time algorithms that can understand and elicit NVC cues; (ii) Secondly, the growing interest of research in socio-behavioral computing has gathered sufficient momentum in the last 2-3 years for this project to be widely disseminated and generate impactful solutions; and (iii) Thirdly, the proposed work provides new perspectives to modeling human communication dynamics, by including individuals with visual impairments who have little or no access to non-verbal cues, and thereby providing pathways to obtain deeper insights into the needs of sighted individuals [23].

The necessary components to study and model non-verbal communication is fragmented across several scientific communities - including those in behavioral psychology, computer vision, speech, linguistics and signal processing - thus making this effort challenging. From a computational perspective, analysis of non-verbal social behavior in real-world settings is a complex problem that includes analysis of physical appearance, gestures and postures, face and eyes behavior, vocal behavior, and the surrounding space and environment [15]. Each of these problems, along with how they relate to one another, bring up fundamental research issues such as: (i) how to extract non-verbal communication cues in real-time and recognize them; (ii) how to combine NVC cues from different modalities and time frames to represent a particular emotion or stance; and (iii) how to model the temporal dynamics of NVC cues (e.g. speed, co-occurrence, etc) [15]. Integrating the user-in-the-loop to these research questions adds additional dimensions of sensory substitution (visual NVC cues to haptics). Further, it is also necessary to model interaction dynamics such as reciprocity [24], by combining cues of all participants of an interaction. Over the last decade, extensive amounts of research have been carried out on related problems such as body part detection [25][26][27]; silhouette/profile detection and body pose estimation [28]; automatic gait recognition [29]; gesture recognition [30]; body and hand posture recognition [31][32]; face detection [33]; face recognition [34]; head pose estimation [35][36]; expression/affect recognition [14]; gaze detection [37]; and vocal behavior/emotion recognition [38]. Despite the varied efforts, core problems of social interaction/NVC analysis – such as interpreting head and body cues in terms of communication semantics; sensing the context of the interaction; or classifying cues into target socio-behavioral categories in a context-sensitive manner - are still nascent. Several challenges at various levels of processing - robustness against errors in low-level processing, view and rate-invariant representations at mid-level processing and semantic representation of interaction dynamics at higher level processing - make this problem hard to solve.

In order to maintain the viability of the development of an assistive prototype for dyadic interactions, our research in this proposal will remain focused on extracting non-verbal cues related only to the face and head. Face and head movements have been understood to express NVC with the highest effectiveness. Psychological experiments that study dyadic communication using a single behavioral cue have shown that facial expressions provide the best results [1]. While face/head-based movements cannot convey all possible NVC cues, prior research on predicting effective collaborative relationship development (such as trust building, accurate personality assessment, and interaction enjoyment) have offered evidence of specific face-based cues that help structure social interaction in dyads. Smiles are universally recognized as signs of pleasure and welcome [39], or to communicate social deference [40]. The gaze direction indicates where one's attention is focused: looking at a speaker indicates full attention, looking away indicates wandering attention, and a sudden change in eye gaze direction prompts partners to follow [41]. Head nods communicate a listener's ability to follow the speaker's message over time [42]. Facial expressions, defined using Ekman-Friesen's Facial Affect Coding System [43], express cognitive states like awe, amusement, pride, nurturant love, anticipation interest, anger, sadness, fear, embarrassment, confusion, surprise or puzzlement; psychological states like suicidal depression; social behaviors like accord and rapport; personality traits like extraversion and temperament; and social signals like status, trustworthiness, emblems (i.e., culture-specific interactive signals like wink); regulators (i.e., conversational mediators like nod and gaze exchange); and illustrators (i.e., cues accompanying speech like raised eyebrows) [15]. We believe that the lessons learnt from the study of head/face-based NVC cues will help expand this work into whole-body analysis at later stages of this project.

In the past we have developed algorithms targeted at mobile wearable platforms and have successfully demonstrated the use of computer vision, pattern recognition and machine learning technologies for enriching social situational presence, which we will build upon. To the best of our knowledge, while there have been related efforts in social signal processing, this will be one of the first efforts to create a technology in a field which is still in its infancy. Commercially available software (such as Seeing Machines FaceAPI) can identify and track facial features reasonably well from stable video frames (without much noise) allowing for significant real-world rotations of the head and face positions, which are typical of human interpersonal communication. Hence, in this proposal, we will focus on building knowledge layers on top of existing state-of-the-art methods (like FaceAPI). We will focus our efforts on addressing the further challenges in studying semantic mappings of facial features to interaction categories (such as emotions and other temporal dynamics). In particular, our preliminary work in this regard has shown us the necessity to address three significant challenges: (i) Facial expression recognition using hierarchical models of permanent and transient facial features; (ii)

Figure 4: FaceAPI® facial feature tracking software output

Reliable multi-view face detection; and (iii) Novel machine learning models of capturing the complex relationships between facial features/fiducial and interaction/emotion semantics.

Facial expression recognition research has been actively exploring detect-localize-track framework for over a decade and has reached a certain level of maturity in recognition reliability [14]. Most algorithms rely on either: (i) tracking permanent facial features to determine the expression, or (ii) treating the entire face expression image as a single input and extracting generic image features (like Gabor, Fourier and other Wavelet coefficients, Image Gradient Analysis, Local textural variations etc), followed by a learning algorithm. Generally, a Bayesian framework is adapted towards modeling the extracted features through Hidden Markov Models [44], Dynamic Bayesian Networks [45] or other temporally evolving probabilistic maps. However, human face reading experts (FACS experts) tend to use the combination of permanent facial feature movement with temporally evolving transient facial features (like crow's feet, forehead wrinkle, etc.) to determine the exact Facial Action [43]. For example, when analyzing a face for the incidence of AU 2 (eyebrow lower; which suggests sadness), face readers focus on the outer edges of the eyebrows to detect downward movement. Subsequently, they are known to analyze the edges of the temple to detect wrinkles formed above the eyebrow's outer edge. Motivated by human face readers, we plan to explore spatio-temporal Bayesian graphical models that encode both permanent and transient facial features. We believe that this approach can lead to fundamental and impactful innovations. In order to ensure that the dyadic interpersonal mediation system can operate at the four levels described earlier (literal, semi-literal, semi-symbolic and symbolic), we plan to ensure that the computational models that are developed are hierarchical by design, to be tightly coupled with the delivery possibilities. We have recently been studying the applicability of hierarchical Bayesian networks [46], decision networks and aggregation methods such as Dempster-Shafer theory [47] to provide for catering to levels of detail, each of which can be quickly combined to obtain information at the next higher level (e.g., upwards-curved lip and wide narrow eyes can be combined to obtain the smile concept).

Industry has successfully developed near-frontal face detection technology, and we will tap existing resources (such as the OpenCV implementation of face detection) in initial stages of this project to be able to develop a complete system that can be evaluated. However, going beyond, real-world multi-view face detection is still a significant research challenge [48]. In the proposed dyadic interpersonal communication mediator, typical face detection algorithms fail when there is a non-frontal head and face projection on the face tracker (Figure 3b). Further, atypical artifacts on the face like hair covering the forehead, person wearing glasses or a hat, or changes in environmental lighting conditions can mislead typical pattern classification algorithms. In recent years, we have explored various methods to overcome these challenges including an Evidence Aggregated Random Field Model for validating face detection results [47], real-time stochastic tracking to localize the head region from frame-to-frame using Structured Mode Seeking Particle Filter [49], and ensemble classifiers for person tracking on mobile machine vision platform using statistics of image features (like Histogram of Oriented Gradients, Edge Orientation Histograms, Shape Contexts and Active Shape Models) [50]. While these techniques have provided incremental improvements in detecting and tracking the interaction partner's face and head, we plan to study fundamentally different approaches to achieve reliable multi-view face detection. Specifically, we plan to exploit mutual context of nearby objects or body parts (motivated by [51]), and also explore multimodal approaches to face detection (such as [52]).

From perspectives of image-based object detection, the human face is a highly structured object across all possible morphologies when compared to any generic object, like a chair. Exploiting this structural persistence can be very useful in increasing the computational efficiency of any face-processing algorithm. In the past, human anthropometry has been used in validating face detection by comparing the ratios of distances between identified facial features. In recent work, we have built a human face anthropometric model derived from the 1988 Anthropometric survey conducted in the US Army [53]. The resulting model has offered rich insights into variations seen in human face and head measurements. These models contain 65 different measurements of the various facial fiducials with separate sex-based deviation models. Using these models, we have developed a Multivariate Gibbs Sampler for predicting the locations of facial features. The Gibbs Sampler is fast and numerically very stable, which can be implemented on embedded platforms easily. We believe that this approach will aid in achieving better performance in simultaneous face detection and facial feature localization, and in establishing a corresponding mutual contextual relationship. We are also currently exploring multivariate models with repeated bivariate filtering towards exploiting multithreaded embedded cores.

We will conduct a focus group study in the first year of the project, where NVC cues and interaction semantics of highest value during dyadic interactions are identified by individuals with visual impairments. The cues identified through such studies will be correlated through observational study of interactions of people who are visually impaired with their sighted counterparts. Videos (from various natural settings, including movies) will be collaboratively encoded with the representative NVC cue. In order to learn the relationship between spatio-temporal facial features and target NVC cues, we plan to adapt novel threads of machine learning – in particular, multi-label learning (where a facial artifact such as a raised eyebrow could correspond to different cues such as surprise or fear) and multi-task learning (which learns a problem together with another related problem at the same time using shared data/knowledge representations – such as expression recognition and interest level recognition, for example). From our past work on manifold learning for head pose estimation [36], we are also aware that face and head data can be understood to lie on an n-dimensional manifold, where n is the number of degrees of freedom in the movement. We intend to extend our prior work to identify suitable low-dimensional representations of face/head data whose geometry has a meaningful interpretation. Representing data in low dimensions also reduces the computational complexity of the learning problem. Further, as mentioned earlier, we believe that the novel goals of the proposed research in developing assistive social aids will provide us with newer perspectives of developing more effective and efficient computational models.

The tasks in this Research Objective will be coordinated by PI Panchanathan, Co-PI Ramirez, Postdoctoral Researcher Balasubramanian, and the graduate students. As stated before, components of the research will be conceptualized as software Capstone projects that can be provided to senior undergraduate students for porting algorithms onto the device. To maintain feasibility, the challenges in this objective will be addressed in four stages, each of which will result in adding features to the integrated device at one of the four levels. The stages of development will progress from the *Literal* level (where the facial features are presented as is) to the *Symbolic* level (which requires significant research effort to capture the semantics underlying the facial features and fiducials). This will also ensure that the integrated system can be evaluated periodically (at each stage) by users who are blind.

5.3 Research Objective 3: Visuo-haptic sensory substitution for delivering high bandwidth social data

Vision is the primary sense organ for most mammals; and for a few primates, including humans, trichromatic vision is so highly evolved [54] that a major portion of the neuronal pathway in the brain is dedicated to sensing, perceiving and cognizing visual stimuli. This allows human vision to process large amounts of visual data. Koch et al. estimated that human eyes, with 10^6 ganglion cells, could transmit up to 10 Mbits/s [55] to the brain. Hence substituting vision with any other sensory channel is a challenging task that requires careful design. This task becomes even more challenging when the visual data has high temporal bandwidth like facial expressions and head based communicative mannerisms. The task of sensory substitution for individuals who are visually impaired is further complicated by the fact that the users are already relying on their ears to listen to their interaction partners speak. Our studies in this domain have shown that a vibrotactile haptic communication language has 4 desired properties: easy-to-learn, easy-to-recall, intuitive, and hard-to-overlook (Figure 5).

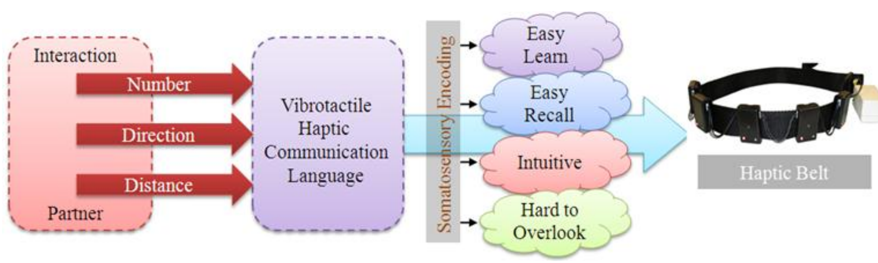


Figure 5: Haptic cueing towards delivering social situational awareness

Recently, at CUbiC, we have developed technologies that take advantage of the human somatosensory channel (the human skin) which is sensitive to vibration, temperature change and tap as the primary substitution channel and have developed technologies that have proved effective in sensory substitution for individuals who are blind and for telesurgery. Significant results for these projects include: (a) Sensing color through haptic actuations [56], (b) Distal object perception through standardized definition of object properties [57], and (c) Simulated orthopedic surgery training tool [58]. Our recent work on delivering social distances for people who are blind has resulted in wearable vibrotactile spatial awareness devices including the first [59] [60] and second [61] prototypes of the haptic belt for social situational awareness [10] (Figure 5). The location of vibrators, duration of vibration and number of active vibrators were adapted to encode direction, distance and number of interaction participants involved in a social interaction. Further, as part of our pilot studies towards this proposal, we recently developed the Vibroglove for communicating 6 classes of Ekman's basic facial expressions of the interaction partners to the dorsal surface of the palm of the user [12]. Figure 6 shows the Haptic Glove technologies along with the haptic cueing patterns that were used for communication. While these devices were able to provide visuo-haptic substitution to a certain level, the work proposed here requires more complex cueing patterns, given the increased complexity of the haptic interface and the necessity to deliver facial cues at user chosen levels of complexity. Unfortunately, there exists no established procedure or protocol for developing haptic cueing patterns. The science (both psychology and physiology) of haptics is still maturing towards understanding nuances of haptic communication.

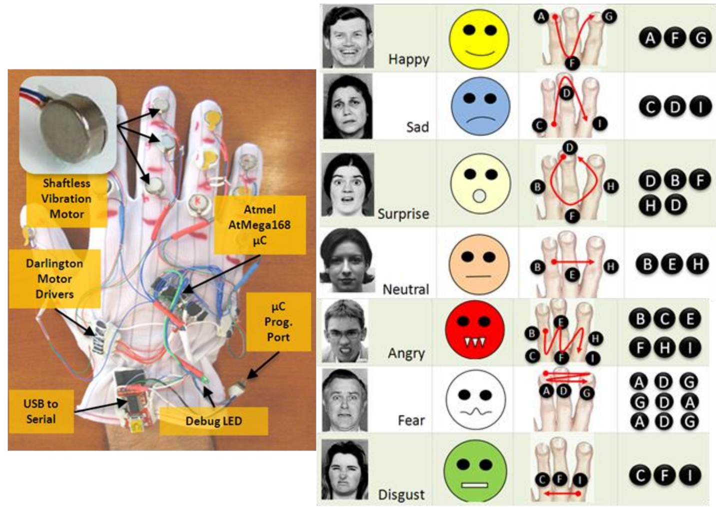


Figure 6: Haptic cueing towards delivering 6 basic expression concepts

As part of this proposal, we will develop new cueing patterns focused on delivering spatio-temporal facial expression data on the proposed Haptic Face Display (See Section 5.1), at four distinct levels of literal, semi-literal, semi-symbolic and symbolic cues, as described earlier. This poses important research challenges including: (a) identifying the various dimensions of haptic actuations on the HFD, (b) determining what discrete levels of haptic cueing are possible on the HFD, (c) determining natural haptic actuations that can easily correlate to human facial/head movements, (d) identifying basic metaphors of haptic cueing on the HFD, (e) ensuring that the haptic cues meet objective measures of usability, including easy to learn (learnability), easy to recognize (recognition accuracy), fast recall (response time), and easy to recall (retention), and (f) ensuring that the haptic cues meet subjective needs of the users.

Previously, complex vibratory patterns have been designed using combinations of vibration dimensions [62] [63], such as vibration frequency, amplitude, duration, rhythm and location, and by using human psychophysical perception effect (like sensory saltation [64]). There are infinite ways to map meanings to vibration dimensions, but conceptually, there are two extremes. On one end of the spectrum, tactons [65], or tactile icons, use a symbolic mapping to arbitrarily assigned meaning to vibration dimensions. On the other end, a literal mapping assigns vibrotactile cues to intuitive somatosensory signals that humans are already acquainted with, such as a shoulder tap to obtain one's attention. Encoding schemes may also fall somewhere in between such that the vibrotactile cues may be intuitive, but still require training. Similar results have been found with other haptic actuation mechanisms including solenoids, peltiers, and piezo-actuators. We will follow a structured evaluation methodology to identify the best set of cueing patterns that map facial movements at the four distinct levels spanning from literal to symbolic. In addition to the generic evaluation metrics (detailed in Section 5.4), the cueing patterns needed to develop an effective HFD will be studied using the specific metrics listed below.

<u>Objective Evaluation:</u> *Learning rate:* the number of training trials required towards memorizing all the cueing patterns; *Accuracy*: the number of correctly identified patterns presented during the training; *Response time:* the time to respond to vibrotactile patterns will be recorded for each participant. The individually recorded response times will be used to learn about the overall response time for each metaphor; *Retention:* participants will be divided into groups who are trained at once, but brought back for testing after 1, 2, 3 and 4 weeks of training towards understanding the retention time for the haptic cues.

<u>Subjective Evaluation:</u> *Questionnaire:* participants will be required to answer, on a scale from 1 (very difficult) to 10 (very easy), a series of structured questions on their experience of the cues as well as the interaction itself. *Feedback Discussions*: the participants, observers, caregivers and developers will participate in a joint feedback session after each user evaluation meeting, to discuss suggestions.

Most work in the area of haptic interpersonal communication has focused on delivering physical sensation of social touch like hugging [66], hand shake [67], tap on the shoulder [68], etc. Very few researchers [69] have explored the communication of very complex information, like facial expressions. Similarly in terms of sensory substitution very little has been done towards mapping complex visual or auditory information into haptics. The proposed work will thus provide novel perspectives in advancing both these research fronts. The tasks in this objective will be coordinated by PI Panchanathan, Co-PI Hedgpeth, Co-PI Ramirez, Research Engineer Krishna, and the graduate students.

5.4 Research Objective 4: Evaluation in Real-World Dyadic Interactions

Evaluations in the project will be coordinated by Co-PIs Ramirez and Hedgpeth. Co-PI Hedgpeth (Director, Disability Resource Center at ASU – ASU DRC) will actively seek the enrollment of students with visual impairments in our studies, and our community partner, Arizona Center for the Blind and Visually Impaired, has agreed to assist us with enrolling visually impaired individuals in the age group of 15-40, who will benefit maximally from these technologies. These evaluations will begin from the start of the project as some of the proposed prototypes are already in place. We will conduct quarterly evaluations and assessments in order to understand the experiences of the users and the interaction partners. The observations made from these assessments will be used to make design changes, thus completing the feedback loop in the proposed research.

Our past experience has revealed that there exists a knowledge gap between technology users and technology designers, and this gap can result in complete denial of a certain technology by the user population. Users often are unable to articulate their needs and requirements in terms of the available state-of-the-art technologies; and designers are now realizing that the design of inclusive technologies have their own guiding principles. To address this basic issue, we will take the following steps:
(1) As recently demonstrated in our NoteTaker project (NSF Award 0931278), involving students and participants who are blind in the design and development team offers a first-hand evaluation strategy where developers themselves become the users of the system. CUbiC already works closely with the disability resource center on ASU campus and annually recruits students with disability to: (a) increase exposure of the students to design and development of assistive technology, and (b) participate in experiments and ethnographic studies towards technology design and development. Similar efforts will be continued to ensure significant presence of students who are visually impaired, especially as part of undergraduate theses/projects. (2) From a different perspective, many assistive technologies seem easy to adapt for a younger, technology savvy population, while the elderly and technology-averse population tend to disregard them. In order to avoid this imbalance, we will engage with our community partner, ACBVI, who specialize in providing social interaction opportunities for people who have developed visual impairment and other disabilities due to their age. (3) Recent trends in social media and networking have shown that the successful adoption of a technology (be it assistive or not) depends on a set of highly skilled early-adopters and their opinions on the product's capabilities, learning curve and return on investment. These reviews are critical to allow mass appeal of assistive technologies. To this end, we propose to promote active blogging and social media presence while the technology is being developed. We intend to identify technology savvy users who can spawn assistive technology discussions related to social assistance. Through these early adopters, we propose to promote, evaluate and improve the design of the proposed social aid.

Learner Assessments: Assistive technologies may be considered effective only when they provide efficient means to users to help accomplish their tasks, or enhance their quality of life substantially. Learner assessments will be conducted to measure the impact of the dyadic social mediation tool on the quality of life of individuals who are blind and visually impaired in their ability to engage coherently in social interactions. Because one of the primary interests is assessing the extent to which the proposed technology assists users in managing interpersonal interactions, we will conduct laboratory experiments intended to document both perceptual and behavioral outcomes associated with the technologies. Following a period of training for learning how to use the assistive aids, participants will engage in two mock employment interviews. The technology will be utilized during only one of the interviews with the other counterbalanced in order to avoid bias; interviews will be recorded for behavioral analysis. Following each interview, participants will complete self-report measures evaluating perceptions of the interview including assessments of their own performance (communicator competence: appropriateness, effectiveness), comfort in and ease of using the technologies during the interaction, confidence during interaction, and enjoyment of the experience.

Nonverbal cues are instrumental in the management of interpersonal interactions, with participants displaying not isolated behaviors (e.g., a smile) but rather patterns of mutual influence, behavioral adaptation, or accommodation *in relation to their partners* in order to insure interactions run smoothly [70]. As such, the proposed technologies require the evaluation of the extent to which they help blind and visually impaired individuals coordinate their interaction behaviors in a similar manner. We would expect the patterns displayed when the assistive aids are used to differ significantly from when they are not (e.g., less behavioral adaptation to partner). Rather than focusing only on interviewee behavior, coding of the two interactions described above will focus on the behavior displayed by both participants (interviewer, interviewee) and analyzed statistically using existing protocols [70]; this allows for an assessment of how interviewer behavior, particularly facial expressions, influences that of the interviewee and vice versa. Other behaviors such as gestures and vocalic cues (e.g., backchanneling) that have been shown to be instrumental in interaction management will be also coded for use as controls. Our experiments will be motivated by the understanding that perception of another's socioemotional display is not a logical, deliberative process. These perceptions are instinctive, requiring little conscious cognitive effort, and we are often unaware of the nonverbal cues driving our perception of others' states [71]. Appropriate responses to a sender's emotions are best facilitated not by *knowing* that the sender is feeling an emotion, but by *feeling* an appropriate, corresponding emotion in response [72].

Technology evaluations: To facilitate acceptance of the developed social assistive technology solutions, surveys will be conducted online and at our collaboration centers (ASU DRC, ACBVI) to determine form factors that will best suit the user population in their daily lives. Designs of sensing and delivery interfaces will be evaluated using a protocol of observation, supported with measures of self-report through questionnaires and interviews to evaluate: (1) usability of the technology; (2) effectiveness of the technology, (3) user (the person who is blind and their interaction partner) perception of the technology, and (4) shortcomings. Stringent performance metrics that are motivated by the requirement for low computational complexity and real-time performance will be used to evaluate the algorithms used in the proposed framework at every step of development. Metrics that capture the naturalness of delivery and the cognitive overload due to training will be used to evaluate multimodal actuator technologies and delivery strategies. The results of the formative evaluation process will be continuously integrated into design decisions of subsequent stages of the project.

Summative Evaluation: In addition to the formative evaluations, a comprehensive summative evaluation will be carried out midway during the project, and in the 3rd year of the project. This evaluation will be performed to study the (1) usability of the developed technologies; (2) productivity in activities or interactions, as the case may be; (3) efficacy of non-verbal cues, and (4) overall impact. External evaluators from the caregiver organization (ACBVI) will be invited to conduct learner assessments and supervise the summative evaluations to quantitatively measure the usability of the social assistance technologies. Learner assessments of the user's skills and dispositions will be performed against the system's objectives and quality standards to determine the benefits of the proposed system. The system will also be evaluated by the training overload on the user using measures such as duration of training and efficiency of training.

6. Intellectual Merit and Broader Impacts

Intellectual Merit: The proposed research will result in the advancement of computational thinking in socio-behavioral computing (through the introduction of novel methodologies for computational analysis and evaluation of social behavior in dyadic interactions), human-computer interfaces (through the design of novel interfaces that deliver high-bandwidth social cues), machine intelligence (through the study of algorithms that elicit various levels of interaction semantics) and assistive technology/usability (through the development and evaluation of social assistive prototypes for individuals with visual impairments). The intellectual merit of this project is in the potential transformative impact of designing next-generation assistive social awareness aids, which not only answers fundamental research questions in the component disciplines, but more importantly, opens up new research directions at the intersections of these disciplines, and in healthcare technologies, disability studies, and behavioral psychology. Further, the technology evaluations by the user population and caregivers will provide us with valuable insights to develop better computational frameworks for use-inspired technologies in the most general sense.

Broader impacts: This project will have a significant impact on the lives of individuals with all kinds of visual impairments. More importantly, the concepts and technologies developed will provide pathways to technologies for individuals with other disabilities, such as autism, prosopagnosia and frontotemporal dementia, and - in the most general sense - a very large portion of society. The methodologies developed will catalyze interdisciplinary research in the component disciplines, and provide a wealth of data and information that will be made publicly accessible to promote further research. The diversity of the faculty team (by disability, gender and race), as well as the strong orientation of current and past projects of the PI to engage individuals with disabilities, will promote the active recruitment, retention and mentoring of students with disabilities, women and underrepresented/minority students across the disciplines. This project will also lead to invention of commercializable scientific methods and innovative technologies. Workshops and exhibit sessions will be proactively conducted, and portions of the research will be implemented as undergraduate Capstone projects, to promote training and learning in STEM education.

7. Education and Curriculum Development Activities

<u>Community and student training:</u> The recent socio-economic situation within US and around the world has triggered the larger community to question the importance of investments in research. It is therefore critical to demonstrate the benefits of research to society and to raise awareness of the importance of research to the general population. Since the proposed project naturally demonstrates the benefit to humanity, we plan to embark on specific efforts to educate the surrounding communities in the Phoenix area. In the past, we have invited high school children (including children who are blind and visually impaired) to visit our labs and offer demonstrations of the latest in computing technologies. We have also actively engaged high school students in summer research camps and projects. Our efforts have revealed that children assess opportunities based on the rewards that they receive and we plan to promote a reward-based STEM outreach for K-12 students through technology exhibits. Our team also has a long track record of effectively engaging undergraduates in senior undergraduate capstone projects, with a specific focus on including students with disabilities. The PIs are also jointly committed to improve diversity at both the undergraduate and graduate level. In particular, co-PI Hedgpeth (who is blind herself) has been a role model for the blindness and visual impairment community through extraordinary academic and educational accomplishments. Together, we will positively influence the lives of visually impaired students and community participants.

<u>Curriculum Development Activities:</u> This project has the potential to lead to new educational programs, as well as courses, in science and engineering, as well as in healthcare, assistive and wellness systems. Our past experience has shown us the importance of students learning by doing, and we will incorporate this approach in the possible course offerings that we have identified so far (listed below):

Socio-behavioral Computing: This interdisciplinary course will train students on the study of social behavior and context based on computational systems. Unlike existing social computing courses that study the virtual/web social behavior, this course will be based on the physicality of human social behavior and communication, and provide insights into computational solutions to address challenges in this field. PI Panchanathan and co-PI Ramirez will offer this course with the postdoctoral fellow/engineer.

Person-centered Computing: This course will be a seminar-driven graduate-level course that will seek to incite students to think about potential challenges in real-world case studies for person-centered technologies from perspectives of technology, usability and functionality. The students will be encouraged to design conceptual solutions to this futuristic problem, and will be co-offered by the PI team, the postdoctoral fellow and research engineer.

8. Diversity and Outreach

Dissemination Plan: The PIs have been very successful in technical knowledge diffusion by publishing in premier conferences and journals, while also organizing international workshops and conferences that have proven impactful towards broad knowledge discoveries in related areas. Examples include ACM ASSETS held in Tempe, AZ in 2007, BodyNets conference in 2008, International Workshop on Haptic Audio-Visual Environments and Games (HAVE) at Tempe, AZ in 2010, and upcoming ACM Multimedia Conference in 2011. We will continue to organize joint focused workshops/tutorial sessions based on the results of this project in premier conferences, including the California State University at Northridge (CSUN) Annual Disabilities Conference, Annual Rehabilitation Engineering and Assistive Technology Society of North America (RESNA) Conference, ACM Special Interest Group on Accessible Computing (SIG ASSETS) Conference, and the International Conference on Computers Helping People with Special Needs (ICCHP). We also plan to disseminate this work in the emerging field of socio-behavioral computational modeling in conferences/workshops such as Social Computing, Behavioral Modeling and Prediction (SBP) conference. We will publish our research in premier IEEE/ACM journals (e.g., IEEE

TPAMI, ACM SIGCHI), as well as in journals in disability studies (e.g., IJRR) and interpersonal communication (e.g., JSPR, JoC). We will publish our research in maximum-impact venues of each field and propose special issues.

We will harness the potential of the latest innovations in social web and media to generate widespread impact in this project. In a dedicated web portal for the project, we will host content about the progress of the project, and also moderate a wiki to form a relevant body of knowledge. Latest web technologies including social video sites (such as YouTube), webcasting, podcasting, twitter and social networking sites will be interlinked in our virtual portal for highest impact. When possible, relevant expert interviews and lectures will be recorded and shared to generate widespread interest. Efforts will be made to mix research cultures with the goal of training graduate students who feel less restricted in pursuing cross-departmental research. The proposed research will result in the following measurable concrete outcomes: *Datasets:* While there has been a recent surge in ideas for development of socio-computational systems, translation of these ideas into real-world applications has been limited. This is understandable due to the large amount of overhead involved in establishing real-world physical environments for data collection. We will maintain a well-documented webpage that serves as a community repository of datasets, and related computational tools for data exploration for free public access. *Textbook:* We anticipate that a tangible outcome of this project will be a new jointly-authored textbook on Socio-behavioral Computing. Although the limited timeframe of this project may make this effort challenging, we plan to begin the effort over the course of the project implementation, and intend to publish the gathered knowledge with our research outcomes in the coming years. In addition, it is also our belief that each of the research objectives in this project will lead to several edited books on topics such as mediation interfaces, socio-computational systems, haptic interfaces, etc. *School Exhibits:* In order to extend the impact of this partnership beyond the research community to the general population, we plan to demonstrate the developed technology as a science exhibit in high schools and local events. We have already interacted with students at the Basha Elementary School in Chandler, Arizona, and will initially conduct our event at this venue. We will subsequently explore further opportunities to excite young minds towards use-inspired and socially relevant science and technology.

9. Project Management

PROJECT SCHEDULE	***Year 1:*** Implementation of basic functionality (face detection/feature extraction/haptic cueing) on existing prototype; Conduct focus group studies to refine understanding of user needs/design requirements (obtain feedback on existing prototype); Collect data of dyadic interactions involving sighted-sighted and sighted-blind; Encode interaction videos with non-verbal communication cues; Generate datasets for further research; Explore computer vision/machine intelligence approaches for classifying non-verbal communication cues in real-time (initially on existing datasets, later on generated datasets): focus on modular deliverables; Explore designs for dyadic mediation interface (based on feedback from focus group studies); and design new haptic systems for social data delivery
	Year 2: Continue exploration of computational approaches for non-verbal communication cue recognition; as well as mediation interfaces and haptic delivery to assist in dyadic interactions: focus on iterative modular deliverables; Integrate computer vision/machine intelligence components with hardware interfaces in a modular fashion; Intermediate evaluations of developed designs/technologies/functionalities; Refine technologies based on user population feedback (engage in undergraduate Capstone projects for quick turnarounds)
	Year 3: Integration of developed software and hardware modules for assisting in dyadic interactions; Rigorous evaluation of developed technologies with user populations from different ages (including students with visual impairments); Explore future pathways, including mobile-based solutions using smart phone cameras
	Project-round Tasks: Dissemination and Education (Conference/journal papers, tutorials, special sessions, curriculum development, Capstone student projects); Outreach (website management, social media, webcasts, podcasts); Community Impact (high school student workshops, ACBVI visit, AZ School for Deaf and Blind visit)

Our team members have extensive experience in working as part of successful funded projects. One of the cornerstones of our coordination mechanisms will be the strong leadership provided by the PI. Panchanathan has proven coordination and leadership experience in current and past positions as the Director of the Center for Cognitive Ubiquitous Computing (CUbiC@ASU), Deputy Vice-President for Research and Economic Affairs at ASU, founding director of the School of Computing and Informatics and the Biomedical Informatics Department at ASU. We strongly believe that his leadership skills along with the co-PIs strong commitment will help this project in generating widespread impact. The geographical proximity of the institutions and the past experience of the team members in executing interdisciplinary projects will serve as a valuable factor in the success of this project.