Pervasive Assessment of Social Behavior

James M. Rehg

Professor, School of Interactive Computing
Georgia Institute of Technology

Background

I am the Lead PI on a recent NSF Expeditions award, which proposes to develop novel technologies for capturing and analyzing the social and communicative behaviors of children, in the context of the diagnosis and treatment of developmental disorders such as autism. This effort targets a largely-unexplored dimension in the creation of pervasive computing technologies. More generally, I have spent the past 20 years conducting research in computer vision, with the goal of creating new capabilities for interaction between people and the digital world. Some notable achievements include *DigitEyes*, the first real-time video-based hand tracking system; the DEC *Smart Kiosk*, an early experiment in creating digital avatars that could interact with people in public spaces; state-of-the-art methods for detecting skin, faces, and pedestrians in images and video; algorithms for tracking Fred Astaire's dance movements from movie footage; and a projector-guided painting system, which uses a novel "on-canvas" interface to guide novice oil painters in following classical painting techniques. Through these efforts, I have contributed to the steady erosion of the gap between the physical and digital manifestations of human existence, a process which is fundamental in making computing truly pervasive.

Vision

Beginning in infancy, individuals acquire the social and communication skills that are vital for a healthy and productive life. Our understanding of this developmental process is incomplete, due primarily to a paucity of data. This limitation impacts not only our scientific understanding, but also our ability to detect and treat a broad range of behavioral and developmental disorders. **I envision the creation of novel pervasive computing technologies to support the large-scale capture and analysis of social and communicative behaviors under natural conditions**, beginning in infancy and extending throughout childhood (and even adulthood). My goal is to be able to model and quantify the verbal and nonverbal interactions between children and their caregivers and peers that are instrumental in the development of socialization and language.

The need to capture multimodal behavioral data (e.g. vision, speech, and wearable signals) inobtrusively and under natural conditions poses substantial challenges in pervasive computing, such as the need for new forms of smart sensor networks and wearable sensors. In addition, basic research in human-centered computing is required to insure that behavioral sensing technologies are understood and accepted by a broad range of stakeholders. Furthermore, unambiguous interpretation of multimodal sensor data requires knowledge of context, which requires further integration with pervasive computing technologies (e.g. location- and context-aware services).

Impact

Children with developmental delays face great challenges in acquiring social and communicative skills, resulting in substantial lifetime risks. Autism, for example, affects 1 in 110 children in the US and can lead to substantial impairments, resulting in a lifetime cost of care of

$3.2M per person. At the opposite end of the age range, elderly individuals face challenges in sustaining social relations, for example due to decreased mobility or cognitive impairment. In addition to their vital role in interpersonal relations, social and communicative behaviors can also provide an important source of information about consumer preferences, the effectiveness of advertising, and the function of organizations. In spite of their importance and ubiquity, there has been very little previous work on sensing and modeling social and communicative behaviors in a comprehensive and integrative manner. Moreover, there have been very few computational studies of the behavior of children in general and of pre-linguistic children in particular. Previous areas of significant research investment, such as automatic speech recognition or location-based activity modeling, have tended to focus on adult users and goal-oriented tasks, usually in the context of a single sensing modality.

Human behavior is inherently multi-modal, and individuals use eye gaze, hand gestures, facial expressions, body posture, and tone of voice along with speech to convey engagement and regulate social interactions. The NSF Expedition that I am leading is developing multiple sensing technologies, including vision, speech, and wearable sensors, to obtain a comprehensive, integrated portrait of expressed behavior. Cameras and microphones provide an inexpensive, non-invasive means for measuring eye, face, and body movements along with speech and nonspeech utterances. Wearable sensors can measure physiological variables such as heart-rate and skin conductivity, which contain important cues about levels of internal stress and arousal that are linked to expressed behavior. This project is developing unique capabilities for synchronizing multiple sensor streams, correlating these streams to measure behavioral variables such as affect and attention, and modeling extended interactions between two or more individuals. In addition, novel behavior visualization methods are being developed to enable real-time decision support for interventions and the effective use of repositories of behavioral data. We refer to this technology collectively as **Behavior Imaging (BI)**, by way of analogy to the medical imaging technologies which have revolutionized medical science in the past century. Looking forward, we believe there is an opportunity to have a similar impact on the science of behavior.

Behavior imaging technologies can support a variety of applications. For example, they could enable novel objective screeners for behavioral disorders based on short interactions between a child and a clinician (e.g. in the context of a well-baby visit). In addition, BI could support real-time decision support systems for behavioral interventions and therapy based on automatic quantification of progress towards behavioral goals. Beyond the treatment of disorders, BI could be used to assess more general patterns of socialization and interaction in natural settings (e.g. monitoring the number of social bids children initiate and receive throughout the day in the context of a daycare). Related methods could be applied to older adults to assess the quality and quantity of their social interactions, as a measure of the stability of their social network.

If I am selected to attend this meeting, I will use the opportunity to identify points of contact and potential opportunities for collaboration between my Expeditions effort and related and interconnected research topics in the domains of pervasive computing, wearable and ubiquitous computing, and intelligent sensor networks, among others.