Advances in Affective Computing for Psychological Applications

From the Fundamentals to the Future of Emotional Cognizant Artificial Intelligence Entities

Thomas B Talbot MD1, Albert S Rizzo PhD1, and Mohammed Soleymani PhD1
1University of Southern California Institute for Creative Technologies

The advance of artificial intelligence technologies such as visual, auditory and linguistic processing systems coupled with interactive virtual humans has advanced the potential of affective computing beyond an academic curiosity or exploratory research to the level of a practical, yet still evolving technology which has shown to be an effective tool for clinical patient-facing applications, psychological assessments, and clinical training. This symposium provides an introduction to the technology and techniques of today’s affective computing from a variety of perspectives. These perspectives start with the viewpoint of a computer scientist who introduces the conceptual framework for recognition of expressions and emotions. The discussion continues with a clinical psychologist’s experiences employing user state sensing in novel patient-facing efforts with provocative virtual human interviewers. The third perspective is that of a physician who employs a variety of sensory modalities for clinical training simulations. We start with an overview of the technologies and techniques used to identify emotional cues. Psychological evaluations and interventions have been employed and the results from these efforts as well as lessons learned are reviewed. We also employ experiences with sensory assisted virtual patients to share insights into the practicality of each major sensing technology that feeds intelligent, judging virtual agents with dynamic emotionally informed behaviors. This variety of applications along with the combination of the three distinct perspectives will serve to provide a well-rounded introductory survey of the techniques, results and potential of today’s state of the art affective computing applications.

Keywords: Affective Computing, Emotion Perception, Virtual Humans, Conversational Simulations

Affective computing technologies for automatic recognition of expressions and emotions
Mohammed Soleymani PhD

Affective Computing strives to enable machines with emotional intelligence, i.e., to recognize, understand and express emotions. Advancing automatic understanding of human behavior and emotions involves technical challenges in multiple domains, such as computer vision and natural language understanding. Computational understanding of emotions relies on psychological models of emotions, e.g., basic emotion theory and core affect, for quantifying and characterizing emotions either as distinct categories or multi-dimensional scales. Emotion sensing technologies have found their way to a multitude of applications including healthcare, virtual assistants, education and marketing.

Emotions are subjective experiences involving interaction between perceptual and contextual factors. Emotions are measured through their proxies, i.e., recognizing emotional expressions or physiological changes. The most common approaches of emotion recognition include recognizing facial or vocal expressions of emotions. Recognition of emotional expressions requires large amounts of labeled data, expensive to produce. Human behavior also has person-specific characteristics which poses a challenge for generalization of machine learning models. In this talk, we review the common approaches in automatic emotion recognition, including multimodal methods that can use different channels of information simultaneously, for example, facial and vocal expressions. To alleviate the scarcity of labeled data, we show how unlabeled videos can be leveraged for machine learning for expression recognition. We also show that using adversarial learning, a method that can make a machine learning model less susceptible to individual characteristics, such as gender, we can make emotion recognition more robust to between-person or between-group variations. Advancements in machine learning for
emotion recognition will enable future applications for socially and emotionally intelligent machines.

**Clinical Use of Virtual Human Interaction with Enhanced Capability to Infer User State via Automatic Behavior Sensing of Face, Gesture, Speech, and Physiological Responses**

Albert S Rizzo PhD

There has been growing interest in the use of artificially intelligent (AI) or semi-autonomous virtual human (VH) characters that allow users to engage in credible therapeutic interactions within Clinical VR simulations. Recent advances in this technology has driven application development across a number of fields, from education to military training to providing clinical assessment, treatment, and healthcare support. VH applications have been successfully used in support of exposure therapy for public speaking, social phobia, and role-play training with people on the autism spectrum. More complex VHs infused with varying levels of voice recognition, natural language processing, and AI have shown documented effectiveness in the role of virtual patients for training novice clinicians and as automated healthcare mentors. Moreover, we have developed VHs that can serve as clinical interviewers leveraging automated sensing of facial, gestural, and vocal behaviors useful for inferring the state of the user interacting with these VH entities and for assessing empathetic behavior with clinical trainees. VH’s can also conduct clinically oriented interviews within a safe, non-judgmental context that may encourage learning or honest disclosure of important information, while at the same time quantifying behavior that serves to inform pre/post clinical changes in clients undergoing treatment for sexual trauma. This talk will summarize our work and discuss our recent development of the Omnisense system focused on creating an adaptive system leveraging computer vision, vocal analytic, and physiological sensing of user behaviors to quantify user’s state and levels of psychological distress during interaction with a conversational VH agent.

**Acting on Perception: Practical emotion sensing techniques and considerations in conversational human-computer clinical simulations and training games.**

Thomas B Talbot MD

Through the creation of practical training applications which use affective computing coupled with conversational virtual patients, a great deal has been learned about practical considerations for the employment of emotion perception as a technology and practical tool for evaluators. Those considerations are both technological and approach based. Essential technologies for affective applications include natural language understanding (NLU), computer vision, auditory processing of speech tonal characteristics and contextual assessments. Important approaches to effective sensing include methods to identify context of the conversation in progress, a model of the virtual patient’s emotional state and use of evocative stimuli to provoke an emotional reaction when conducting measurements. We have successfully employed affective computing to create virtual psychiatric patients that take an emotional read from their human interviewers during the course of a clinical encounter. These virtual patients employ a judgment model which continuously updates and alters the patient behavior and degree of honest disclosure to the human interview accordingly. Our results indicate that opportunities for computer judgment happen at key moments and that identification of those moments allows for periods of critical sensing which enables great ‘reads’ from the human interviewer and allows our models to declutter the noise of continuous impressions taken during a prolonged encounter. There are opportunities to employ gaze detection, facial action unit decoding of emotional expressions, language usage and non-visual metrics to obtain a quality read of affective and entrainment aspects of clinical interviews.