Automatic, Dimensional and Continuous Emotion Recognition

Hatice Gunes, Imperial College London, UK

Maja Pantic, Imperial College London, UK and University of Twente, EEMCS, The Netherlands

ABSTRACT

Recognition and analysis of human emotions have attracted a lot of interest in the past two decades and have been researched extensively in neuroscience, psychology, cognitive sciences, and computer sciences. Most of the past research in machine analysis of human emotion has focused on recognition of prototypic expressions of six basic emotions based on data that has been posed on demand and acquired in laboratory settings. More recently, there has been a shift toward recognition of affective displays recorded in naturalistic settings as driven by real world applications. This shift in affective computing research is aimed toward subtle, continuous, and context-specific interpretations of affective displays recorded in real-world settings and toward combining multiple modalities for analysis and recognition of human emotion. Accordingly, this article explores recent advances in dimensional and continuous affect modelling, sensing, and automatic recognition from visual, audio, tactile, and brain-wave modalities.

Keywords: Bodily Expression, Continuous Emotion Recognition, Dimensional Emotion Modelling, Emotional Acoustic and Bio-signals, Facial Expression, Multimodal Fusion

INTRODUCTION

Human natural affective behaviour is multimodal, subtle and complex. In day-to-day interactions, people naturally communicate multimodally by means of language, vocal intonation, facial expression, hand gesture, head movement, body movement and posture, and possess a refined mechanism for understanding and interpreting information conveyed by these behavioural cues.

DOI: 10.4018/jse.2010101605

Despite the available range of cues and modalities in human-human interaction (HHI), the mainstream research on human emotion has mostly focused on facial and vocal expressions and their recognition in terms of seven discrete, basic emotion categories (neutral, happiness, sadness, surprise, fear, anger and disgust; Keltner & Ekman, 2000; Juslin & Scherer, 2005). In line with the aforementioned, most of the past research on automatic affect sensing and recognition has focused on recognition of facial and vocal expressions in terms of basic emotional states, and then based on data that
has been posed on demand or acquired in laboratory settings (Pantic & Rothkrantz, 2003; Gunes, Piccardi, & Pantic, 2008; Zeng, Pantic, Roisman, & Huang, 2009). Additionally, each modality—visual, auditory, and tactile—has been considered in isolation. However, a number of researchers have shown that in everyday interactions people exhibit non-basic, subtle and rather complex mental/affective states like thinking, embarrassment or depression (Baron-Cohen & Tead, 2003). Such subtle and complex affective states can be expressed via tens (or possibly hundreds) of anatomically possible facial expressions, bodily gestures or physiological signals. Accordingly, a single label (or any small number of discrete classes) may not reflect the complexity of the affective state conveyed by such rich sources of information (Russell, 1980). Hence, a number of researchers advocate the use of dimensional description of human affect, where an affective state is characterized in terms of a small number of latent dimensions (e.g., Russell, 1980; Scherer, 2000; Scherer, Schorr, & Johnstone, 2001).

It is not surprising, therefore, that automatic affect sensing and recognition researchers have recently started exploring how to model, analyze and interpret the subtlety, complexity and continuity of affective behaviour in terms of latent dimensions, rather than in terms of a small number of discrete emotion categories.

A number of recent survey papers exist on automatic affect sensing and recognition (e.g., Gunes & Piccardi, 2008; Gunes et al., 2008; Zeng et al., 2009). However, none of those focus on dimensional affect analysis. This article, therefore, sets out to explore recent advances in human affect modelling, sensing, and automatic recognition from visual (i.e., facial and bodily expression), audio, tactile (i.e., heart rate, skin conductivity, thermal signals etc.) and brainwave (i.e., brain and scalp signals) modalities by providing an overview of theories of emotion (in particular the dimensional theories), expression and perception of emotions, data acquisition and annotation, and the current state-of-the-art in automatic sensing and recognition of emotional displays using a dimensional (rather than categorical) approach.

BACKGROUND RESEARCH

Emotions are researched in various scientific disciplines such as neuroscience, psychology, and linguistics. Development of automated affective multimodal systems depends significantly on the progress in the aforementioned sciences. Accordingly, we start our analysis by exploring the background in emotion theory, and human perception and recognition.

THEORIES OF EMOTION

According to the research in psychology, three major approaches to emotion modelling can be distinguished (Grandjean, Sander, & Scherer, 2008): (1) categorical approach, (2) dimensional approach, and (3) appraisal-based approach.

The categorical approach is based on research on basic emotions, pioneered by Darwin (1998), interpreted by Tomkins (1962, 1963) and supported by findings of Ekman & his colleagues (1992, 1999). According to this approach there exist a small number of emotions that are basic, hard-wired in our brain, and recognized universally (e.g., Ekman & Friesen, 2003). Ekman and his colleagues conducted various experiments on human judgment of still photographs of deliberately displayed facial behaviour and concluded that six basic emotions can be recognized universally. These emotions are happiness, sadness, surprise, fear, anger and disgust (Ekman, 1982). Although psychologists have suggested a different number of such basic emotions, ranging from 2 to 18 categories (Ortony & Turner, 1990; Wierzbicka, 1992), there has been considerable agreement on the aforementioned six emotions. To date, Ekman’s theory on universality and interpretation of affective nonverbal expressions in terms of basic emotion categories has been the most commonly adopted approach in research on automatic affect recognition.
Related Content

Agile Wheeled Mobile Robots for Service in Natural Environment
[www.irma-international.org/chapter/agile-wheeled-mobile-robots-for-service-in-natural-environment/84920/](www.irma-international.org/chapter/agile-wheeled-mobile-robots-for-service-in-natural-environment/84920/)

Distributed Control of Robot Swarms: A Lyapunov-Like Barrier Functions Approach
[www.irma-international.org/chapter/distributed-control-of-robot-swarms/141997/](www.irma-international.org/chapter/distributed-control-of-robot-swarms/141997/)

Surgical Robots: System Development, Assessment, and Clearance
[www.irma-international.org/chapter/surgical-robots-system-development-assessment/63538/](www.irma-international.org/chapter/surgical-robots-system-development-assessment/63538/)

A Virtual Reality Study of Help Recognition and Metacognition with an Affective Agent

Novel Swimming Mechanism for a Robotic Fish
[www.irma-international.org/chapter/novel-swimming-mechanism-robotic-fish/78098/](www.irma-international.org/chapter/novel-swimming-mechanism-robotic-fish/78098/)