

Chapter 8.8

Facial Expression and Gesture Analysis for Emotionally–Rich Man–Machine Interaction

Kostas Karpouzis

*National Technical University of Athens,
Greece*

Athanasios Drosopoulos

*National Technical University of Athens,
Greece*

Themis Balomenos

*National Technical University of Athens,
Greece*

Stefanos Kallios

*National Technical University of Athens,
Greece*

Amaryllis Raouzaïou

*National Technical University of Athens,
Greece*

Spiros Ioannou

*National Technical University of Athens,
Greece*

Nicolas Tsapatsoulis

*National Technical University of Athens,
Greece*

ABSTRACT

This chapter presents a holistic approach to emotion modeling and analysis and their applications in Man-Machine Interaction applications. Beginning from a symbolic representation of human emotions found in this context, based on their expression via facial expressions and hand gestures, we show that it is possible to transform quantitative feature information from video sequences to an estimation of a user's emotional

state. While these features can be used for simple representation purposes, in our approach they are utilized to provide feedback on the users' emotional state, hoping to provide next-generation interfaces that are able to recognize the emotional states of their users.

INTRODUCTION

Current information processing and visualization systems are capable of offering advanced

and intuitive means of receiving input from and communicating output to their users. As a result, Man-Machine Interaction (MMI) systems that utilize multimodal information about their users' current emotional state are presently at the forefront of interest of the computer vision and artificial intelligence communities. Such interfaces give the opportunity to less technology-aware individuals, as well as handicapped people, to use computers more efficiently and, thus, overcome related fears and preconceptions. Besides this, most emotion-related facial and body gestures are considered universal, in the sense that they are recognized among different cultures. Therefore, the introduction of an "emotional dictionary" that includes descriptions and perceived meanings of facial expressions and body gestures, so as to help infer the likely emotional state of a specific user, can enhance the affective nature of MMI applications (Picard, 2000).

Despite the progress in related research, our intuition of what a human expression or emotion actually represents is still based on trying to mimic the way the human mind works while making an effort to recognize such an emotion. This means that even though image or video input are necessary to this task, this process cannot come to robust results without taking into account features like speech, hand gestures or body pose. These features provide the means to convey messages in a much more expressive and definite manner than wording, which can be misleading or ambiguous. While a lot of effort has been invested in individually examining these aspects of human expression, recent research (Cowie et al., 2001) has shown that even this approach can benefit from taking into account multimodal information. Consider a situation where the user sits in front of a camera-equipped computer and responds verbally to written or spoken messages from the computer: speech analysis can indicate periods of silence from the part of the user, thus informing the visual analysis module that it can use related data from the mouth region, which is essentially

ineffective when the user speaks. Hand gestures and body pose provide another powerful means of communication. Sometimes, a simple hand action, such as placing one's hands over their ears, can pass on the message that they've had enough of what they are hearing more expressively than any spoken phrase.

In this chapter, we present a systematic approach to analyzing emotional cues from user facial expressions and hand gestures. In the Section "Affective analysis in MMI," we provide an overview of affective analysis of facial expressions and gestures, supported by psychological studies describing emotions as discrete points or areas of an "emotional space." The sections "Facial expression analysis" and "Gesture analysis" provide algorithms and experimental results from the analysis of facial expressions and hand gestures in video sequences. In the case of facial expressions, the motion of tracked feature points is translated to MPEG-4 FAPs, which describe their observed motion in a high-level manner. Regarding hand gestures, hand segments are located in a video sequence via color segmentation and motion estimation algorithms. The position of these segments is tracked to provide the hand's position over time and fed into a HMM architecture to provide affective gesture estimation.

In most cases, a single expression or gesture cannot help the system deduce a positive decision about the users' observed emotion. As a result, a fuzzy architecture is employed that uses the symbolic representation of the tracked features as input. This concept is described in the section "Multimodal affective analysis." The decision of the fuzzy system is based on rules obtained from the extracted features of actual video sequences showing emotional human discourse, as well as feature-based description of common knowledge of what everyday expressions and gestures mean. Results of the multimodal affective analysis system are provided here, while conclusions and future work concepts are included in the final section "Conclusions – Future work."

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/facial-expression-gesture-analysis-emotionally/24398

Related Content

EA Multi-Model Selection for SVM

Gilles Lebrun, Olivier Lezoray, Christopher Charrierand Hubert Cardot (2009). *Encyclopedia of Artificial Intelligence* (pp. 520-525).

www.irma-international.org/chapter/multi-model-selection-svm/10296

Generative AI Methodology for Producing Assisted Art: Representation of the Historical-Cultural Identity of Southern Brazil

William Alberto Cruz Castañeda, Marcellus Amadeus, André Felipe Zanellaand Felipe Rodrigues Perche Mahlow (2024). *Making Art With Generative AI Tools* (pp. 179-196).

www.irma-international.org/chapter/generative-ai-methodology-for-producing-assisted-art/343427

Virtual Organisational Trust Requirements: Can Semiotics Help Fill The Trust Gap?

Tim French (2009). *International Journal of Intelligent Information Technologies* (pp. 1-16).

www.irma-international.org/article/virtual-organisational-trust-requirements/2448

Analysis of Older Users' Perceived Requests and Opportunities with Technologies: A Scenario-Based Assessment

Mari Feli Gonzalez, David Facal, Ana Belen Navarro, Arjan Geven, Manfred Tscheligi, Elena Urdanetaand Javier Yanguas (2011). *International Journal of Ambient Computing and Intelligence* (pp. 42-52).

www.irma-international.org/article/analysis-older-users-perceived-requests/52040

Adapting Technical Theatre Principles and Practices to Immersive Computing and Mixed Reality Environments

Tim Boucher (2010). *International Journal of Ambient Computing and Intelligence* (pp. 65-67).

www.irma-international.org/article/adapting-technical-theatre-principles-practices/43864