Introduction to the Special Section on Multimedia Computing and Applications of Socio-Affective Behaviors in the Wild

Affective Computing and Social Signal Processing are two developing fields of research that promise to revolutionize the way humans interact with multimedia and technology in general. Affective Computing is the science of creating emotionally aware systems, including the automatic analysis of affect and expressive behaviors. Social signal processing addresses all verbal and nonverbal communicative signaling during social interactions, be they of an affective nature or not. Despite the scientific and technological revolutions of the last decades in these fields, there is still an urgent need to advance emotion and social behaviors recognition for multimodal interface interaction to a level at which systems can deal with large volumes of nonprototypical naturalistic behavior collected in ecologically valid environments, as this is exactly the type of data that applications would have to face in the real world.

The first article of this special section addresses the analysis of malicious forms of communication in online social networks. Securing multimedia sharing and phishing attempts in social networks has become a critical issue with the rapidly increasing number of multimedia data exchanged by users. In the article entitled "Detection of Human, Legitimate Bot, and Malicious Bot in Online Social Networks Based on Wavelets," Sylvio Barbon Jr., Gabriel F. C. Campos, Rodrigo A. Igawa, Rordigo Capobianco Guido, and Mario L. Proença Jr. exploited two Twitter datasets to classify users as being human, a legitimate robot, or a malicious robot. In the proposed method, occurrences of the terms present in the tweets are weighted by lexicon-based coefficients and features are decomposed by a Discrete Wavelet Transform to feed a model based on a Random Forest classifier. An accuracy above 93% is reported for the detection of the three types of users (human, bot, malicious bot), with better results when the tweets are related to a specific hashtag (Super Bowl event) in comparison with miscellaneous themes.

The second article focuses on pose-invariant representation learning of facial expressions for emotion recognition. In the article "Spatially Coherent Feature Learning for Pose-Invariant Facial Expression Recognition," Feifei Zhang, Qirong Mao, Xiangjun Shen, Yongzhao Zhan, and Ming Dong exploit a 3D pose normalization technique that helps in preserving identity information while avoiding the use of multiple models for pose-specific training and parameter tuning. In the proposed method, a sequence of key regions in a synthetic frontal face image is selected for unsupervised representation learning. Dependencies between regions of the face are specifically encoded by introducing a linkage structure over the learning-based features and the corresponding geometry information of each key region. Experimental evaluations were conducted on the BU-3DFE and SFEW datasets, showing that the proposed model achieved superior results over other existing methods on those two datasets.

We hope that this special section will bring inspirational ideas to researchers working in the domains of affective computing and/or social signal processing. We would like to thank all the authors who submitted their work to the special section and the reviewers who contributed with their comments and observations to the high quality of the selected papers. We would also like to

https://doi.org/10.1145/3181711

especially thank the editorial members of ACM TOMM, Prof. Alberto Del Bimbo and Prof. Stefano Berretti, for their support and kind help in the realization of this special section.

Fabien Ringeval
Université Grenoble Alpes
Björn Schuller
Imperial College London/University of Augsburg
Michel Valstar
University of Nottingham
Jonathan Gratch
University Southern California
Roddy Cowie
Queen's University Belfast
Maja Pantic
Imperial College London/Twente University

Guest Editors