SOCIAL SIGNAL PROCESSING AT LAIV

Giuseppe Boccignone, Paola Campadelli, Elena Casiraghi, Claudio Ceruti, Vittorio Cuculo and Raffaella Lanzarotti
LAIV Lab - Dipartimento di Informatica
Universita' degli Studi di Milano
Via Comelico 39/41, Milano, 20135
Italy *

Abstract Facial expressions, body postures, autonomic signals are all examples of signals that are of interest for the computational modelling and understanding of human behaviors. In this note, we summarize the research activities pursued within the LAIV lab aimed at sensing, analyzing and understanding such behavioural signals.

Keywords: Behavioural analysis, social signal processing, affective computing, machine learning

1. Introduction

Current research programs at LAIV (Laboratory for the Analysis of Images and Vision) involve the computational modelling and understanding of human behaviors and patterns of behavioral signals captured by a variety of sensors (Figure 1). This relates to a number of tasks: i) sensing and analyzing displayed behavioral signals including facial expressions and body gestures, in the context in which observed behavioral signals were displayed; iii) understanding human behavior by translating the sensed human behavioral signals and context descriptors into a description of the shown behavior.

Realizations of behavioural modelling are of interest for surveillance, behavioral biometrics, affect-sensitive systems in computer-aided learning environment, customer service, intelligent driver assistance, and entertainment industry.

In such framework, our largest on-going research project (see Figure 2) concerns the design of a computational tool to support facial expressions and

*Partial funding provided by grant FIRB - Future in Research RBFR12VHR7

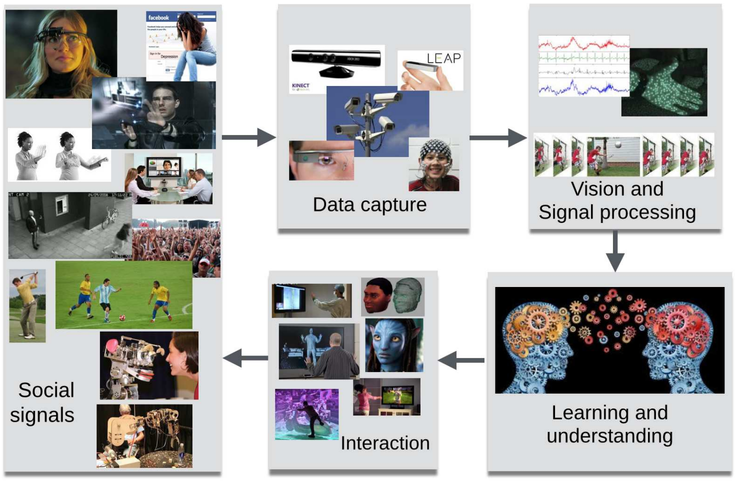

Figure 1. LAIV research activities at a glance

biosignals based shape analysis and Bayesian networks for interpreting emotions (Brombin et al., 2012).

This project has been motivated by the need to provide objective and quantitative tools to measure/identify emotions which will allow a better understanding, assessment, and treatment of neuropsychiatric patients. More precisely, the aim is to clarify theoretical knowledge in the context of anxiety and eating disorders, so to improve treatment techniques, rehabilitation methods, prevention and early diagnosis.

Specific research issues that have been recently addressed within and beyond the above mentioned project are summarized in the following Sections.

2. Processing of autonomic signals

Emotional experience is embodied in peripheral physiology. Hence, systems can detect emotions by analyzing the pattern of physiological changes associated with each emotion (assuming a prototypical physiological response for each emotion exists). Meanwhile, the amount of information that the physiological signals can provide is increasing, mainly due to major improvements in the accuracy of psychophysiology equipment and associated data analysis techniques. Sympathetic and parasympathetic functions of the autonomic nervous system have been analyzed through Hearth Rate Variability signals (Costa et al., 2012).

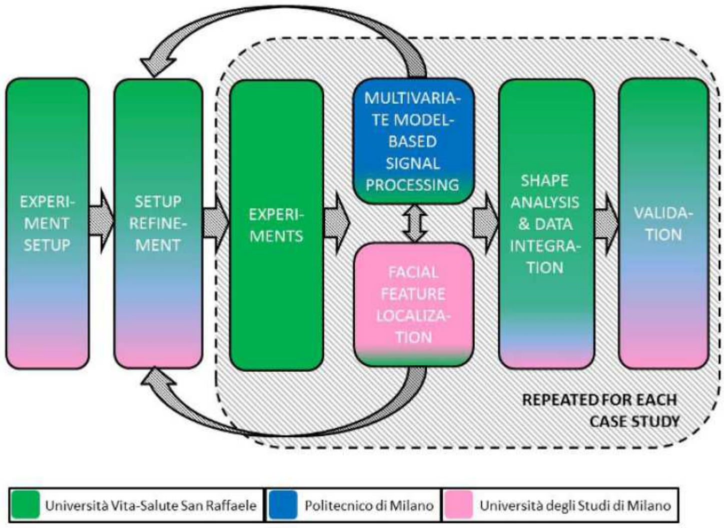

Figure 2. Interpreting emotions: integrating facial expressions, autonomic signals, shape analysis and Bayesian networks

3. Processing of facial signals

Facial expression, together with other communicative signals such as head gestures and gaze, is a rich source of information for nonverbal interaction. In such endeavour, a crucial initial step is the detection of certain facial feature points usually referred as *landmarks*. Sparse coding techniques are currently investigated for either feature learning for the extraction of reliable face landmarks (Boccignone et al., 2014a) and for face recognition (Adamo et al., 2012; Adamo et al., 2013). The rationale behind this approach is that unsupervised learning of sparse code dictionaries from face data can be an effective approach to cope with such challenging problems.

Also, in a different perspective, affective expression recognition has been addressed in the framework of the simulationist paradigm (Vitale et al., 2014).

4. Gaze behavior processing

The relationship between the social function of gaze and its primary cognitive function for the gazer - to foveate and hence analyse in depth a region of the visual world - has emerged in recent years. Different stochastic models for the analysis and the simulation of gaze behavior have been studied in terms of ecological foraging behavior (Boccignone and Ferraro, 2014; Boccignone and Ferraro, 2013b; Boccignone and Ferraro, 2013a; Clavelli et al., 2014; Clavelli et al., 2013). Interestingly, although mental states can be inferred from other modalities, the eyes, are the best and most immediate 'windows to the mind', and also the best indicators that we have 'connected' with another mind. Advanced machine learning techniques have been exploited for detecting the signature of expertise through eye-movements analysis (Boccignone et al., 2014b).

5. Human behavior understanding and Machine Learning

Social Signal Processing is a multidisciplinary venture, with topics such as computer vision, signal processing and machine learning. In particular, machine learning is of fundamental importance for both understanding behaviors and mining the hidden mechanisms that generated them. New general learning techniques for intrinsic dimensionality estimation and Fisher discriminant classifiers have been developed in our lab (Rozza et al., 2012; Campadelli et al., 2013; Ceruti et al., 2014).

Further, the epistemological issue of levels of explanation in behavioral sciences is under investigation (Boccignone and Cordeschi, 2012).

6. User Interaction

Face recognition systems aim to recognize the identity of a person depicted in a photograph by comparing it against a gallery of prerecorded images. Current systems perform quite well in controlled scenarios, but they allow for none or little interaction in case of mistakes due to the low quality of images or to algorithmic limitations. A guided user interface for improving automatic face recognition has been proposed, which allows to adjust from a fully automatic to a fully assisted modality of execution, according to the difficulty of the task and to amount of available information (gender, age, etc.) (Arca et al., 2012).

References

Adamo, Alessandro, Grossi, Giuliano, and Lanzarotti, Raffaella (2012). Sparse representation based classification for face recognition by k-limaps algorithm. In Elmoataz, Abderrahim, Mammass, Driss, Lezoray, Olivier, Nouboud, Fathallah, and Aboutajdine, Driss, editors, *Image and Signal Processing*, volume 7340 of *Lecture Notes in Computer Science*, pages 245–252. Springer Berlin Heidelberg.

Adamo, Alessandro, Grossi, Giuliano, and Lanzarotti, Raffaella (2013). Face recognition in uncontrolled conditions using sparse representation and local features. In Petrosino, Alfredo, editor, *Image Analysis and Processing, ICIAP 2013*, volume 8157 of *Lecture Notes in Computer Science*, pages 31–40. Springer Berlin Heidelberg.

Arca, Stefano, Campadelli, Paola, Lanzarotti, Raffaella, Lipori, Giuseppe, Cervelli, Federico, and Mattei, Aldo (2012). Improving automatic face recognition with user interaction. *Journal of Forensic Sciences*, 57(3):765–771.

Boccignone, G. and Ferraro, M. (2013a). Gaze shift behavior on video as composite information foraging. *Signal Processing: Image Communication*, 28(8):949 – 966.

Boccignone, Giuseppe and Cordeschi, Roberto (2012). Predictive brains: forethought and the levels of explanation. *Frontiers in Psychology*, 3(511).

Boccignone, Giuseppe, Cuculo, Vittorio, and Lanzarotti, Raffaella (2014a). Using sparse coding for landmark localization in facial expressions. In *International Conference on Image Processing, ICIP 2014*. IEEE. Submitted.

Boccignone, Giuseppe and Ferraro, Mario (2013b). Feed and fly control of visual scanpaths for foveation image processing. *annals of telecommunications-annales des télécommunications*, 68(3-4):201–217.

Boccignone, Giuseppe and Ferraro, Mario (2014). Ecological sampling of gaze shifts. *IEEE Trans. on Cybernetics*, 44(2):266–279.

Boccignone, Giuseppe, Ferraro, Mario, Crespi, Sofia, Robino, Carlo, and de'Sperati, Claudio (2014b). The expert's eye: Detecting a signature of individual billiard expertise in eye movements using a multiple-kernel relevance vector machine classification technique. *Journal of Eye Movement Research*, 7(2):1–15.

Brombin, Chiara, Ferrario, Emanuela, and Lanzarotti, Raffaella (2012). Interpreting emotions: a computational tool integrating facial expressions and biosignals based shape analysis and bayesian networks. FIRB - Future in Research, Ministero dell'Istruzione, dell'Università e della Ricerca. Grant: RBFR12VHR7.

Campadelli, Paola, Casiraghi, Elena, Ceruti, Claudio, Lombardi, Gabriele, and Rozza, Alessandro (2013). Local intrinsic dimensionality based features for clustering. In Petrosino, Alfredo, editor, *Image Analysis and Processing ICIAP 2013*, volume 8156 of *Lecture Notes in Computer Science*, pages 41–50. Springer Berlin Heidelberg.

Ceruti, Claudio, Rozza, Alessandro, Bassis, Simone, Lombardi, Gabriele, Casiraghi, Elena, and Campadelli, Paola (2014). Danco: An intrinsic dimensionality estimator exploiting angle and norm concentration. *Pattern Recognition*. To appear.

Clavelli, Antonio, Karatzas, Dimosthenis, Llados, Josep, Ferraro, Mario, and Boccignone, Giuseppe (2013). Towards modelling an attention-based text localization process. In Sanches, Joao, Micó, Luisa, and Cardoso, Jaime S., editors, *Pattern Recognition and Image Analysis*, volume 7887 of *Lecture Notes in Computer Science*, pages 296–303. Springer Berlin Heidelberg.

Clavelli, Antonio, Karatzas, Dimosthenis, Llados, Josep, Ferraro, Mario, and Boccignone, Giuseppe (2014). Modelling task-dependent eye guidance to objects in pictures. *Cognitive Computation*, pages 1–27.

Costa, Tommaso, Boccignone, Giuseppe, and Ferraro, Mario (2012). Gaussian mixture model of heart rate variability. *PloS one*, 7(5):e37731.

Rozza, Alessandro, Lombardi, Gabriele, Casiraghi, Elena, and Campadelli, Paola (2012). Novel fisher discriminant classifiers. *Pattern Recognition*, 45(10):3725 – 3737.

Vitale, Jonathan, Johnston, Benjamin, Williams, Mary-Anne, and Boccignone, Giuseppe (2014). Reading emotions in facial expressions: a computational framework inspired by simulation theories of empathy. In *Annual International Conference on Biologically Inspired Cognitive Architectures*, BICA. Elsevier. Submitted.