Emotional State Detection Based on EMG and EOG Biosignals: a Short Survey

João Perdiz, Gabriel Pires and Urbano J. Nunes

***Abstract*—With its acquisition of muscular potentials linked to displays of affective state, facial Electromyography (EMG) is the most fitting physiological signal for performing emotional expressions detection. EMG and computer vision-based facial recognition have very different strengths and weaknesses, but their information may complement each other. Additionally, Electrooculographic (EOG) signals can enable gaze tracking and detection of eye movements, but blinks and saccades cannot directly indicate emotional state.**

In this paper, we survey and analyze the methods, strengths and challenges of using these biosignal sources for detecting emotional states, and then we propose a framework that combines simultaneously facial expression detection, using EMG, with saccade detection using EOGs, to classify four basic expressions: neutral, sad, happy, and angry.

I. Introduction

The range of potential applications for biosignals has widened exponentially in the last four decades. This was spurred by the development of psychophysiology in the 1970s, a field which focuses on the influence a person's psychological state can have in acquirable physiological signals, or biosignals. In the early 1980s a relation was found between biosignal variation and the recall of past emotional states in [1]. The furthering of this type of research lead researchers to focus on facial Electromyographic (EMG) signals to detect facial gestures, which were conclusively considered displayers of emotional state in [2]. This research progressed hand in hand with a more systematized organization of its physical setups [3]. While not directly linked to emotional state, Electrooculography (EOG) also benefited from these advancements, being proposed as a mean of controlling user interfaces, but also as an indicator of the affective state of a user, pointing to fatigue or stress [4]. Overshadowed by computer vision-based detection techniques, more recently biosignals' strengths relative to digital imaging have merited a resurgence of their applications for emotional recognition. Here we survey the field centering on EMG and EOG signals, while proposing a framework to combine both and use the latter signal's overlooked potential.

Authors are with the Institute of Systems and Robotics, Department of Electrical and Computer Engineering, University of Coimbra.

This work has been supported by FCT under projects "AMS-HMI2012-RECI/EEIAUT/0181/2012", "UID/EEA/00048/2013" with FEDER funding, programs QREN and COMPETE, and Research Grant no SFRH/BD/104985/2014. Emails: {joao.perdiz, gpires, urbano}@isr.uc.pt

II. Biosignals and emotional states

A. Emotional states, facial expressions, and the FACS

One of the first researchers explicitly linking an alteration on physiological signals to the utterance of facial expressions that are universally linked to emotional states was Ekman et al. [1]. Voluntary facial expressions caused measurable variations in biosignals related to involuntary Autonomic Nervous System activity. Skin conductance and temperature were monitored, showing that facial expressions cause variation in biosignals not acquired from the face. Since it was established in [5] that some facial expressions are closely tied to certain emotional states to a nearly universal degree, this indirect proof then served as evidence of the tie between facial expressions and emotions, namely in its dimensions of valence and arousal [6].

The universality of facial expressions lead the authors of [5] to propose both a group of six basic emotions – disgust, fear, joy, surprise, sadness, and anger – and the Facial Action Coding System (FACS). The FACS breaks down facial expressions into their muscular components, or Action Units (AU), while linking groups of AUs to displays of emotional states. The six basic expressions and examples of the AUs can be seen in [7].

Facial EMG research only became more widespread with the publication of a systematized process for acquiring and processing facial EMG signals. The electrode positioning procedures and how to prepare an unbiased experimental paradigm are clearly described in [3]. Together with the FACS, it made it possible to accurately pinpoint which signals – from individual facial muscle groups – to acquire and where to expect variations for each facial expression made by the user, as EMG can be used to classify AUs. The FACS allowed manual categorization of facial expressions and their comparison with self-assessment by experimental participants, opening the way for a faster development of automatic computer vision-based recognition of AUs.

B. Computer vision works

Although not all computer vision-based recognition works employ the FACS to detect facial features, several implementations are already capable of high accuracy rates in various paradigms of recognition: recognizing individual AUs or full-face emotional displays are the most common. However, obtaining good results recognizing spontaneous reactions is still difficult [8].

The main challenges of computer vision recognition are the correct classification of spontaneous expressions, real-time detection of expressions, and the ability to classify occluded, rotated or off-centered faces [9]. A method is proposed in [10] to classify the six emotional expressions through continuous tracking of the onset and offset of AUs, achieving simultaneously recognition of expressions and face tracking in real time. Finally, a head pose invariant system is tested in [11], where a regression-based approach is capable of extracting geometric features from rotated faces to detect expressions as a fixed-pose system would.

C. Facial EMG applications survey

1) Expression detection and recognition: While the works discussed here focus on expression recognition methods, they show great variety in the target classes, and the six basic emotions model is not always used. Table I summarizes the most relevant works. "Detection mode" column in Table I informs of how feature extraction is implemented in real time: in "full-event" it is performed over all data in one event, whereas in "non-overlapping windows" features are extracted from sliding windows. In "continuous detection", features are continuously extracted from successive data points.

EMG electrodes are usually placed over the *Zygomaticus major* and *Corrugator supercilii* muscles. This stems from the fact that these are the most reliable electrodes for displaying EMG variation upon utterance of emotional expressions, as seen in [21] and [12]. In [12], it is reported a study where ten participants rated their emotional valence when reacting to a series of photos and videos; using these evaluations and features from EMG data gathered by two electrodes over the above mentioned muscles, the goal was to classify reactions in real time into three valence classes.

Other studies use just one electrode or electrodes unconventionally placed, not following the guidelines of [3]. In [13] only the *Zygomaticus major* source is used to classify expressions linked to six custom emotional states, attaining an accuracy of $69,5\%$. This result shows that performance is not always degraded by short acquisition setups if the feature extraction and processing steps are well developed, as verified by the comparison made of the results in [13] to those of studies showing multimodal classification from Electrodermal Activity, Blood Volume Pulse, EMG, Skin Temperature, and R-R distance signal features. In [14], ten different facial gestures over several groups of features are classified. Target gestures are posed and not directly related to emotional expressions, being closer to AUs. In [15], on the other hand, emotional expressions are classified. Facial electrodes are unconventionally placed, above the line of the eyes; this enables the setup to detect an "eyebrow pulling" gesture, a good indicator of surprise and therefore an extension to recognition of states outside the six basic emotions. Both [14] and [15] employ time-domain features, which are favored when processing EMG signals due to their computational simplicity. Another example of classification using single-channel data is reported in [17]. Facial EMG signals from a dataset described in [20] were employed, and used to classify the emotional reactions that participants in the trials of [20] showed to musical stimuli into four distinct classes.

In [16], a larger number of electrodes, eight pairs, was used, resulting in a higher than usual classification accuracy applying a probabilistic classification method.

2) Usability and applications: While good detection accuracies were obtained in the works discussed above, their physical setups would not translate well into practical applications. This is why published research studies aiming at introducing real-world applications of emotional recognition systems often do so in a binary classification paradigm. About this conflict between usability and performance, a counterexample to the papers above can be seen in [18] and [19]. Both present binary classification results, the former with "Smile" and "Neutral" classes and the latter with "Smile" and "Frown" classes. In both cases, distal electrodes are employed so that the user's face is not obstructed, and to allow long utilization periods. In [18] a headband houses the electrodes, and the method is used for enhancement of communication between autistic children and caregivers; in [19], the goal is to provide human feedback to a robotic platform's actions.

The studies discussed in this section show that EMG-derived features can be powerful tools to decode facial expressions on their own. The use of electrodes permit nearly instantaneous processing and classification, and may provide information invisible to computer vision systems. However, besides the cumbersome setups, electrodes still suffer from the risk of crosstalk between EMG sources and other biosignals when multimodal setups are considered. This is one of the challenges that we are trying to overcome in our proposed framework (Section III).

D. EOG detection and classification

Although EOG cannot be used for direct emotional state detection, it can provide multiple forms of complementary information in multimodal setups. A person's gaze and its shift in response to changes in a social setting have long been considered indicators of social engagement [22]. In [4], the rate of blinks and saccades is found to provide information regarding fatigue or anxiety, whereas [23] points to gaze as an indicator of a user's level of attention. On the other hand, saccades, frowns and blinks were studied as indirect indicators of emotional content by [24]. Measurement of stress and fatigue levels through eye movements can inform a Human-Computer Interface (HCI) about the user's tendency to display certain emotional states, while saccades and blinks may indicate both attention and gaze.

Approaches in EOG classification are very disperse. Some works only classify among four to eight types of saccades [25], [26] – some of them to implement HCI control paradigms. Others are that, alongside several saccadic directions, also classify blinks and frowns [24], [27], [28]. More rarely, some works aim at distinguishing saccades from other ocular movements, such as smooth pursuits and vestibulo-ocular reflexes [29].

TABLE I
RELEVANT WORKS RECOGNIZING EXPRESSIONS USING EMG CHANNELS EXCLUSIVELY

Ref.	**Detected expressions**	**Channels**	**Expression nature**	**Detection mode**	**Classification method**	**Application**	**Accuracy (%)***
[12]	positive, neutral and negative valence	2 (Z+C)†	spontaneous	continuous detection	Regression analysis	not reported	39,2 (O)
[13]	happiness, sadness, fear, surprise, disgust, neutral	1 (Z)†	spontaneous	full-event window	k-Nearest Neighbor	not reported	69,5 (F)
[14]	10 non-emotional expressions	3	posed	non-overlapping windows	Neural Network and SVM (either)	not reported	87,1 (F)
[15]	happiness, anger, rage, neutral, frowning	4	posed	non-overlapping windows	Fuzzy C-means	not reported	90,8 (F)
[16]	surprise, happiness, anger, disgust, sadness, neutral	8*	posed	full-event window	Gaussian model probability	not reported	92,19 (O)
[17]	joy, pleasure, sadness, anger	1	spontaneous	full-event window	Support Vector Machine	reactions from listening to music‡	52 (F)
[18]	smiles, neutral	1	natural setting	non-overlapping windows	Support Vector Machine	Detect expressions in autistic children	>90 (O)
[19]	smiles, frowns	4	natural setting	non-overlapping windows	Neural Network	Provide feedback to HRI	87,52 (O)

$*$ – All canonical muscle positions according to [3]. † – Z – *Zygomaticus superior* channel. C – *Corrugator supercilii* channel.
‡ – From database in [20]. $\star$ – O – Online classification; F – Offline classification

Discrete classification of ocular movements may be insufficient to discern gaze direction; this could be addressed with a robust method for directly measuring horizontal saccadic angle, as signal variation is nearly linear to the resting saccadic angle [26], [30]. The inability to apply this approach with vertical saccades does not diminish its value, as social interactions and attention shifts are mostly performed with horizontal drifts in ocular position.

III. COMBINED BIOSIGNALS FRAMEWORK

We propose the inclusion of EOG in emotional detection systems to complement facial expression classification capabilities of EMG features (see our framework setup in Fig. 1). EMG and EOG decoding is applied concurrently. Their combined use is hampered by the risks of sensor crosstalk, and the related physical superposition of sensors in close positions across the face. In this setup, EMG signals are used to classify anger, sadness, happiness and neutral expressions using the time-domain features detailed in [31], [32]. EMG electrodes are placed over the *Zygomaticus major* and *Corrugator supercilii* muscles, the most representative for emotional expressions, while EOG electrodes follow the usual horizontal and vertical placements, detecting vertical saccades and doing horizontal tracking. Horizontal saccades are processed from signal windows so that an angular component is calculated, similarly to [30]. Vertical saccades are classified (as downward or upward), in a process triggered when vertical EOG potential subsample's maxima are the only EOG features crossing a predetermined threshold. This selection allows the EMG and EOG processes to work separately without mutual interference in the respective outputs. Figure 2 shows waveforms from three sources during various events of a laboratory trial, and the effect of crosstalk on both EOG sources when muscular contractions are elicited during the shaded band shown in Fig. 2. The threshold approach is effective for a certain amount of crosstalk, but prevents simultaneous classification of events. Different processing paradigms need to be researched for spontaneous detection. To complement gaze, user wears an Inertial Motion Unit on the back of the head. Head pose angles can then be used to provide a complete picture of the user's gaze target, which can be shown in a proof-of-concept avatar that mimics classified and detect user actions, as is done in [16] for facial expressions.

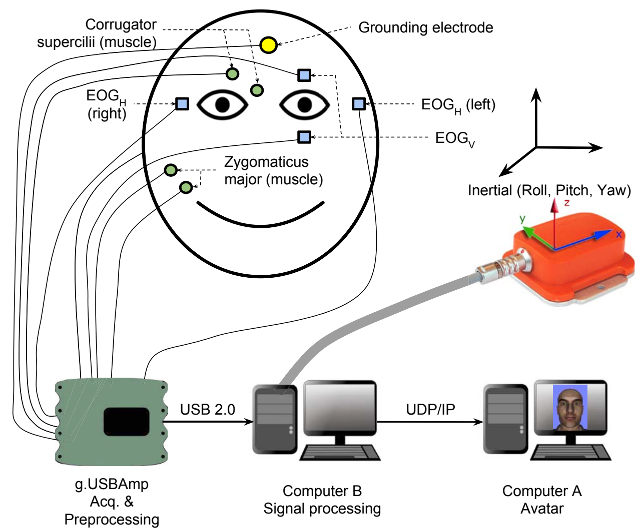

Fig. 1. Simplified block diagram of the proposed framework for a combined EMG and EOG emotional detection and attention state system and avatar reproduction of classified events. The yellow circle points to the reference electrode's position in the face, blue squares to the EOG electrodes, and green circles to the pairs of bipolar EMG electrodes, with target muscles indicated. The IMU system is worn by the user, at the occipital region of the head.

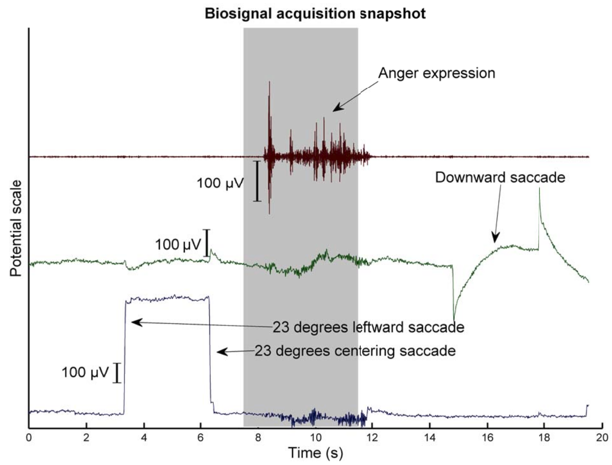

Fig. 2. A superimposed trace of simultaneous horizontal EOG (bottom trace), vertical EOG (middle trace) and *Corrugator supercilii* EMG (top trace) during three different, sequential actions: a horizontal saccade, followed by a facial expression, followed by a vertical downward saccade. Note how the EMG signal originates crosstalk on both EOG sources when an expression is made (shaded area).

REFERENCES

[1] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions," *Science*, vol. 221, no. 4616, pp. 1208–1210, 1983.

[2] J. T. Cacioppo, R. E. Petty, M. E. Losch, and H. S. Kim, "Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions." *Journal of personality and social psychology*, vol. 50, no. 2, p. 260, 1986.

[3] A. J. Fridlund and J. T. Cacioppo, "Guidelines for human electromyographic research," *Psychophysiology*, vol. 23, no. 5, pp. 567–589, 1986.

[4] M. W. Johns, A. Tucker, R. Chapman, K. Crowley, and N. Michael, "Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers," *Somnologie-Schlafforschung und Schlafmedizin*, vol. 11, no. 4, pp. 234–242, 2007.

[5] P. Ekman and W. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement.* Palo Alto: Consulting Psychologists Press, 1978.

[6] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *Journal of research in Personality*, vol. 11, no. 3, pp. 273–294, 1977.

[7] Y.-L. Tian, T. Kanade, and J. F. Cohn, "Facial expression analysis," in *Handbook of face recognition*. Springer, 2005, pp. 247–275.

[8] E. Sariyanidi, H. Gunes, and A. Cavallaro, "Automatic analysis of facial affect: A survey of registration, representation, and recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 37, no. 6, pp. 1113–1133, 2015.

[9] I. Kotsia, I. Buciu, and I. Pitas, "An analysis of facial expression recognition under partial facial image occlusion," *Image and Vision Computing*, vol. 26, no. 7, pp. 1052–1067, 2008.

[10] Y. Li, S. Wang, Y. Zhao, and Q. Ji, "Simultaneous facial feature tracking and facial expression recognition," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2559–2573, 2013.

[11] O. Rudovic, M. Pantic, and I. Patras, "Coupled gaussian processes for pose-invariant facial expression recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1357–1369, June 2013.

[12] T. Partala, V. Surakka, and T. Vanhala, "Real-time estimation of emotional experiences from facial expressions," *Interacting with Computers*, vol. 18, no. 2, pp. 208–226, 2006.

[13] S. Jerritta, M. Murugappan, K. Wan, and S. Yaacob, "Emotion recognition from facial emg signals using higher order statistics and principal component analysis," *Journal of the Chinese Institute of Engineers*, vol. 37, no. 3, pp. 385–394, 2014.

[14] M. Hamedi, S.-H. Salleh, M. Astaraki, and A. M. Noor, "EMG-based facial gesture recognition through versatile elliptic basis function neural network," *Biomed Eng Online*, vol. 12, no. 1, p. 73, 2013.

[15] M. Hamedi, I. M. Rezazadeh, and M. Firoozabadi, "Facial gesture recognition using two-channel bio-sensors configuration and fuzzy classifier: A pilot study," in *Electrical, Control and Computer Engineering (INECCE), 2011 International Conference on*. IEEE, 2011, pp. 338–343.

[16] G. Gibert, M. Pruzinec, T. Schultz, and C. Stevens, "Enhancement of Human Computer Interaction with facial Electromyographic sensors," in *Proceedings of the 21st Annual Conf. of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*. ACM, 2009, pp. 421–424.

[17] C. Zong and M. Chetouani, "Hilbert-huang transform based physiological signals analysis for emotion recognition," in *Signal processing and information technology (isspit), 2009 IEEE Int. Symposium on*. IEEE, 2009, pp. 334–339.

[18] Y. Takano and K. Suzuki, "Affective communication aid using wearable devices based on biosignals," in *Proceedings of the 2014 Conf. on Interaction design and children*. ACM, 2014, pp. 213–216.

[19] A. Gruebler, V. Berenz, and K. Suzuki, "Emotionally assisted human–robot interaction using a wearable device for reading facial expressions," *Advanced Robotics*, vol. 26, no. 10, pp. 1143–1159, 2012.

[20] J. Wagner, J. Kim, and E. André, "From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification," in *Multimedia and Expo, 2005. ICME 2005. IEEE Int. Conf. on*. IEEE, 2005, pp. 940–943.

[21] L. B. P. Ang, E. F. Belen, R. A. Bernardo, E. R. Boongaling, G. H. Briones, and J. B. Coronel, "Facial expression recognition through pattern analysis of facial muscle movements utilizing electromyogram sensors," in *TENCON 2004. 2004 IEEE Region 10 Conference*, vol. 100. IEEE, 2004, pp. 600–603.

[22] L. Schulze, B. Renneberg, and J. S. Lobmaier, "Gaze perception in social anxiety and social anxiety disorder," 2013.

[23] E. Gu and N. I. Badler, "Visual attention and eye gaze during multiparty conversations with distractions," in *Int. Workshop on Intelligent Virtual Agents*. Springer, 2006, pp. 193–204.

[24] A. Cruz, D. Garcia, G. Pires, and U. Nunes, "Facial Expression Recognition based on EOG toward Emotion Detection for Human-Robot Interaction," in *Bio-inspired Systems and Signal Processing, 8th Int. Conf. on (Biosignals 2015)*, 2015.

[25] S. Aungsakul, A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Evaluating feature extraction methods of Electrooculography (EOG) signal for human-computer interface," *Procedia Engineering*, vol. 32, pp. 246–252, 2012.

[26] R. Barea, L. Boquete, M. Mazo, and E. López, "Wheelchair guidance strategies using EOG," *Journal of intelligent and robotic systems*, vol. 34, no. 3, pp. 279–299, 2002.

[27] A. Banerjee, S. Datta, M. Pal, A. Konar, D. Tibarewala, and R. Janarthanan, "Classifying electrooculogram to detect directional eye movements," *Procedia Technology*, vol. 10, pp. 67–75, 2013.

[28] K. Pettersson, S. Jagadeesan, K. Lukander, A. Henelius, E. Hæggström, and K. Müller, "Algorithm for automatic analysis of Electro-oculographic data," *Biomedical engineering online*, vol. 12, no. 1, p. 110, 2013.

[29] M. Vidal, A. Bulling, and H. Gellersen, "Analysing EOG signal features for the discrimination of eye movements with wearable devices," in *Proceedings of the 1st Int. workshop on pervasive eye tracking & mobile eye-based interaction*. ACM, 2011, pp. 15–20.

[30] F. Behrens, M. MacKeben, and W. Schröder-Preikschat, "An improved algorithm for automatic detection of saccades in eye movement data and for calculating saccade parameters," *Behavior research methods*, vol. 42, no. 3, pp. 701–708, 2010.

[31] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7420–7431, 2012.

[32] D. Garcia, "Modelação de expressões faciais num avatar usando classificação de biossinais." Master's thesis, University of Coimbra, 2014.