A Machine Learning Approach for Emotion Detection Through low-cost Hardware*

Emilio López-Ales[1], Mara Trinidad Herrrero[2,3], and José Palma[1]

[1] Artificial Intelligence and Knowledge Engineering Group. University of Murcia
emilio.lopez4@um.es,jtpalma@um.es

[2] Clinical and Experimental Neuroscience (NiCE-IMIB). School of Medicine.
University of Murcia. mtherrer@um.es

[3] Institute for Aging Research. University of Murcia.

Abstract. In the field of Affective Computing, one of the most important issues is the identification of the emotional state of a subject. There are a plethora of research works in emotion identification, works that have their foundations in other fields such as philosophy, psychology, neuroscience, and cognitive sciences. Nowadays, with the emergence of wearable devices and DIY electronics kits, the interest in developing emotion identification systems with these low-cost devices has gained more attention. The use of low-cost devices came out with new challenges related to the low quality of the signals acquired due to less noise-tolerant sensors which are used in real-life environments. In this context, the main objective of this work is to present a methodology, based on machine learning techniques for time series forecasting, to build models able to identify emotional states, from signals acquired from low-cost devices, as accurately as a professional medical device can do. To this end, we proposed the use of two devices: *Nexus-10 MKII*, a biofeedback and neurofeedback system from MindMedia, used to obtain reference measure, and *BiTalino (r)evoltuion Boar Kit* (BiTalino hereinafter), a low-cost physiological signals acquisition device from PLUX Wireless Biosignals S.A. In this work,11 Machine Learning models have been developed to predict the emotional state, identified by *Nexus-10*, with the signals provided by *BiTalino*. Our experiments show that the best model was a **Random Forest** which can predict the emotional state in the test set with a $RMSE$ of 0.172 and a R^2 of 0.858.

Keywords: Emotion identification · Affective Computing · Time Series Forecasting · Machine Learning.

1 Introduction

Emotions are a fundamental part of human behaviour and, as pointed out in [7,8], play an important role in human decision tasks. However, not until Rosling

Picard coined the term "Affective Computing" (AfC) [19], do we realise the need of emotion aware computational systems. Currently, it is widely assumed that a system capable of identifying the affective state of the user and reacting to them can offer a better human-computer interaction experience and, in many cases, make less frustrating the use and adoption of new technology. From its beginnings, AfC has been a prolific research field, making possible the development of effective systems in a long range of applications domains such as, for example, medicine [15], assisted learning [22], arts [9], entertainment [14] and ambient intelligence [1]. In all of these areas, AfC aims to reduce the communicative difference between human emotions and computers, developing systems capable of recognising and reacting to the emotional states of users.

From its beginnings, AfC research has been focused on developing systems able of 1) human emotion identification, 2) expressing emotions and 3) "feeling" emotions [5]. Apart from the recent advances in 2) and 4), the topic that has received more attention from the AfC community is emotion identification. Without a reliable emotion recognition process, it is impossible to develop emotion-aware systems. It is in this context in which this research is conducted.

Emotion identification requires representational models in which identified emotional states could be measured. Multiple models have been proposed by researches of a wide range of fields, ranging from psychology and philosophy to neuroscience and cognitive science (see [10,16] for a review). Among all the available, the OCC model [18], based on the appraisal theory proposed by James Russel [21], is the most widely used in AfC. In the OCC model, emotions are represented in an orthogonal two-dimensional space. One of the dimension is the valence in which states ranging from pleasure to displeasure can be represented. The other dimension, arousal, is in charge to capture the intensity of the emotion (from excited to calm).

In the field of neuroscience, relevant studies have revealed a correlation between the response of the Autonomous Nervous Systems (ANS) to human emotions and the valence-arousal plane [4,11]. More specifically, a great number of research studies has pointed out that the Galvanic Skin Response (GSR) correlates with the arousal levels and Heart Rate (HR) with the emotional valence. However, although GSR and HR are widely used, there are a huge number of research focused on detecting emotions from other physiological signals (see [5] for a review). Among the most commonly used physiological signals, we can find Electromyogram (EMG), Electrocardiogram (ECG), Electroencephalogram (EEG), Electrooculogram (EOG) and Blood Volume Pressure (BVP).

Although, huge number of medical devices are available for acquiring these signals from the medical community, currently there is a growing interest in developing emotion recognition systems using low-cost devices, such as wristbands and electronic DIY kits [13,12,17,20,23]. Some of the advantages of using low-cost devices, apart from their cost, are their portability which makes possible the design of experiments in real-life situations, outside of highly controlled laboratory environments. Apart from the portability capabilities, their autonomy, due to a low energy consumption hardware, makes possible to extend the period in which the signals are recorded. However, despite these advantages, one of the main problems that has to be faced when working with low-cost devices is the quality of sensors. In this sense, a mechanism to deal with poor noise-tolerant sensors, which introduce more artefacts than those obtained by medical devices, are needed to obtain reliable measures. It is in this context in which this work has been developed. The main objective of this work is to present a methodology, based on machine learning techniques for time series forecasting, to build models able to identify emotional states, from signals acquired from low-cost devices, as accurately as a medical device can do. To this end, we proposed the use of two devices: *Nexus-10 MKII*[1] (Nexus-10 hereinafter), a biofeedback and neurofeedback system from MindMedia, used to obtain reference measure, and *BiTalino (r)evoltuion Boar Kit* (BiTalino hereinafter), a low-cost physiological signals acquisition device from PLUX Wireless Biosignals S.A.

2 Data and Experimental procedure

To obtain the required data, an experiment was performed with the collaboration of students of the University of Murcia. Students between 18 and 28 years old have been studied. The volunteers have been contacted individually, following a methodology for the experiment, giving them an appointment with exact date and time. The experiment consists of the visualisation of a collection of 40 well known paintings, arranged randomly for each participant. While the subjects visualise the stimuli, the necessary physiological signals are acquired through the sensors of the mentioned devices. The data have been treated with the utmost confidentiality, in accordance with Spanish Law 3/2018 of 5 December on the Protection of Personal Data and Guarantee of Digital Rights.

After accommodating the subject as best as possible with the sensors in place, the phases of the experiment are remembered. It's also reminded that is crucial that, for the duration of the entire experiment, the subject must look at the screen.

Fig. 1. Representative diagram of the experiment's phases.

The figure 1 depicts the experimental protocol followed. The description of the phases of the experiment is as follows:

- *Eyes Closed Phase*. This phase lasts 60 seconds. During this phase, the subject must remain still with their eyes closed until they are told to open them

[1] *Nexus-1* is Medical CE certified and FDA registered

again. This phase allows us to know the rhythms and frequencies of the subject when they are in a state of minimal activity.

- *Basal Phase*. This phase lasts 60 seconds. During this phase, the subject has to look directly at a black screen. This phase makes possible to know what is the "normal" state of the subject when they are active without receiving any stimuli.
- *Stimulus Phase*. In this phase, the subject will observe a collection of forty well-known paintings randomly arranged on the screen. These paintings are the visual stimuli that are projected individually one after the other, with a duration of 8 seconds.
- *Basal Phase*. Another basal phase exactly the same as the first one.

Once the experiment is finished, all recordings are stopped and the corresponding files, with the collected data, are saved. Sensors are then removed from the subject and cleaned.

3 Data recording

During the experiment, two different devices have been used for physiological signals acquisition. To obtain a reliable emotional index for training the machine learning model, *NeXus-10* has been used. *Nexus-10* is capable of acquiring multiple physiological signals: EEG (2 derivations), EOG (electrooculography), GSR, BVP and temperature. For the objective of this work, GSR and BVP (Blood volume Pulse) signals have been considered. During the experiment, EEG (2 derivations), EOG and eye-tracking information have also been acquired for other purposes beyond the scope of this work.

The other device used is*BITalino*, from Plux Wireless Biosignals S.A. *BiTalino* is a physiological signal acquisition device which is based on similar projects, such as Arduino and Raspberry Pi. It is a low-cost, modular, multi-purpose, easily accessible and configurable acquisition device capable of capturing multiple physiological signals in real-time: Accelerometer, ECG, EDA (Electrodermal Activity), EEG (1 derivation), EGG (Electrogastrography), EMG (Electromyography), EOG, temperature and light. For acquiring the pulse signal, a pulse sensor, connected to one of the analogical channels has been used. Its cost and its open hardware and software philosophy make *BiTalino* a very interesting tool for developing projects.

In this work, the following physiological signals have been acquired:

- *NeXus-10 MKII*: GSR and BVP, both at a sample rate of 32Hz. GSR sensor is placed in the proximal phalanges II and III of the left hand. BVP sensor is placed in the distal phalanx II of the right hand.
- *Bitalino*: GSR and Pulse both at a sample rate of 1000Hz. The GSR sensor is placed in the middle phalanges II and III of the left hand. The BVP sensor is placed in the distal phalanx I of the left hand.

These sample rates produces signals of 17440 samples for the *Nexus-10* and 545000 samples for signals for the *BiTalino*.

4 Singals proceessing

Despite the quality of signals acquired with *NeXus-10 MKII*, some processing is needed. As we are interested in the Skin Conductance Level (SCL), the tonic component of the GSR, a Continuous Decomposition Analysis using Non-negative Deconvolution have been applied to the GSR signal using Ledalab Software [2,3]. *NeXus-10 MKII* provides the HR values directly from the BVP signal, so no processing is required.

Signal acquired through *BiTalino* required some processing to remove both noise and artefacts. First, a *Butterworth* filter of order 3 and cutoff frequency of 2.5 Hz has been applied. The filtered signal is then processed by a *Savitzky-Golay* filter of order 1 and a frame length of 75. No prepossessing has been done on the EDA signal. The figure 2 shows a comparison between the signal BVP obtained by the *NeXus-10 MKII* and the *BITalino* after filtering.

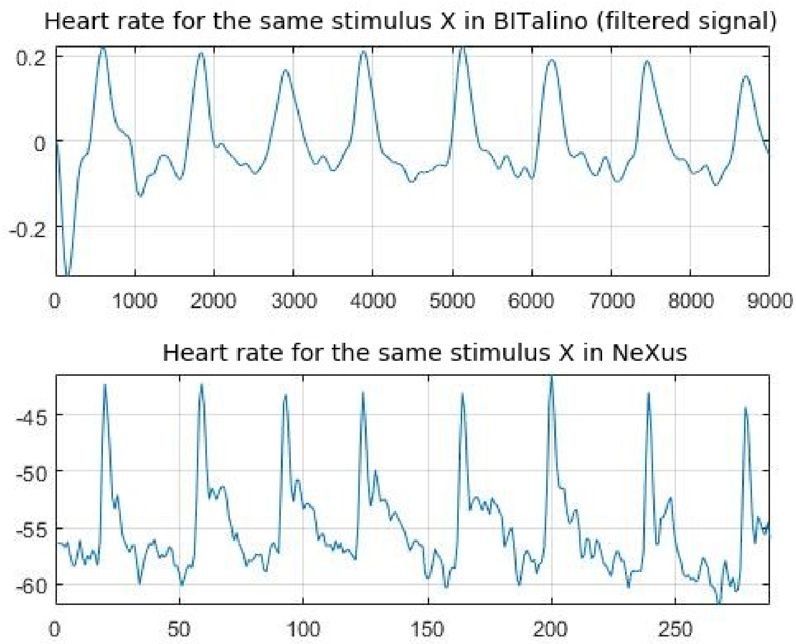

Fig. 2. BVP NeXus-10 MKII and BITalino sensor signals.

As *BiTalino* and *Nexus-10* acquired signals at different sample rates, a down-sampling process was applied to *BiTalino* signals to equal the number of samples and synchronise the timestamps. Then, the signals acquired were processed and segmented according to the stimuli presented. Finally, for each painting, four time series, each one composed of 364 samples, have been obtained.

5 Emotional Index

In this work, the Emotional Index (EI) is calculated as proposed in [6,24,25]. The idea under EI is to obtain a monodimensional variable from the two variables that define the effects plane [16]: HR, (horizontal axis) associated to the valence and SCL (tonic component of GDR), vertical axis associated to the arousal.

Using this approach, the emotional state of a subject can be define as:

$$EI = 1 - \frac{\beta}{\pi} \tag{1}$$

where

$$\beta = \begin{cases} \frac{3}{2}\pi + \pi - \vartheta \; if \; SCL_z \geq 0, HR_z \leq 0 \\ \frac{\pi}{2} - \vartheta \qquad\quad otherwise \end{cases} \tag{2}$$

and

$$\vartheta = \arctan(HR_z, SCL_z) \tag{3}$$

SCL_z and HR_z represent the Z-score variables of the SCL and HR, acquired from *Nexus-10*, respectively. The σ and μ required for the transformation are calculated from the corresponding signals acquired during the 2 baselines phases (at the beginning and the end of the experiment). The EI, obtain through t1, 2 and 3 equations, varies between $[-1, 1]$, where positives values are associated with positive emotions and negative values to negative emotions. Once EI has been calculated, all the signals are downsampled to produce one sample per second. At the end, a dataset with 13832 samples is obtained.

6 Model Building and results

Once signals have been processed and EI has been calculated, for each stimulus three temporal series, *BITalino* EDA and pulse signals together with the EI, are used to create a multivariate time series. Therefore, the problem for building a model for emotion identification from *BITalino* can be approached as a multivariate time series forecasting problem using Machine Learning Techniques.

To build the model, different data set configurations, with a different number of lagged variables, have been tested:

- **Forma_1**. No lagged variables considered, to predict EI_{t_i} only values of GSR_{t_i} and HR_{t_i} are taken into account as predictors.
- **Forma_2**. Two lagged version of predictors has been added to the previous data set producing two new datasets: **Forma_2_v2** and **Forma_2_v3** with one and two lagged versions of GSR and HR respectively.
- **Forma_3**. Two lagged version of the EI has been added to **Forma_2** datasets generating two new datasets: **Forma_3_v2** and **Forma_3_v3** with one and two lagged versions of EI respectively.
- **Forma_4**. Two new datasets have been created: **Forma_4_v2** and **Forma_4_v2** with with one and two lagged versions of EI added to **Forma_1** respectively.

From the original dataset, 20% of the samples have been reserved for testing. In this work, we have considered the following regression models: Linear, Knn, CART (Classification and Regression Trees), Random Forest, Bayesian Ridge, Lasso, Linear SVM (Support Vector Machines), ϵ-SVM, ν-SVM, SGD (Stochastic Gradient Descent) and Multilayer Perceptron. All the models have been trained over the seven datasets previously generated using 10 folds stratified cross-validation with a grid hyperparameter search. $RSME$ and R^2 have been chosen as performance measures. At the end of the process, 77 models were generating (11 regression models $\times$ 7 datasets).

First of all, in order to reduce the number of models to be analysed, for each model, the best pair (model, dataset), according to their evaluation in the test set, has been chosen (Table 1).

Model	**Dataset**	**RMSE (train)**	**RMSE (test)**	$\mathbf{R^2}$**(train)**	$\mathbf{R^2}$**(test)**
LR	*forma3_v3*	0.393 ± 0.011	0.387 ± 0.092	0.502 ± 0.024	0.410 ± 0.183
KNN	*forma4_v3*	0.367 ± 0.009	0.443 ± 0.078	0.570 ± 0.017	0.164 ± 0.485
CART	*forma3_v3*	0.000 ± 0.000	0.562 ± 0.084	1.000 ± 0.000	-0.259 ± 0.303
RF	*forma3_v2*	0.191 ± 0.014	0.172 ± 0.084	0.881 ± 0.016	0.858 ± 0.177
BRR	*forma3_v3*	0.393 ± 0.011	0.387 ± 0.092	0.502 ± 0.024	0.410 ± 0.183
Lasso	*forma4_v3*	0.395 ± 0.011	0.384 ± 0.100	0.498 ± 0.026	0.426 ± 0.196
Lin-SVM	*forma4_v3*	0.176 ± 0.020	0.179 ± 0.079	0.433 ± 0.070	0.305 ± 0.250
ϵ-SVM	*forma1*	0.563 ± 0.006	0.561 ± 0.055	-0.021 ± 0.013	-0.266 ± 0.318
ν-SVM	*forma3_v3*	0.155 ± 0.009	0.168 ± 0.076	0.503 ± 0.0243	0.350 ± 0.248
SGD	*forma4_v3*	0.396 ± 0.013	0.390 ± 0.090	0.495 ± 0.032	0.400 ± 0.184
MLP	*forma3_v3*	0.394 ± 0.014	0.412 ± 0.090	0.501 ± 0.033	0.385 ± 0.366

Table 1. Best results of each model through all datasets.

In order to determine if the observed differences in performance are statistically significant, statistical hypotheses tests have been applied. Due to the small number of sample in each group, if difficult to prove the parametric assumption (normality and sphericity), therefore the non-parametric *Friedman's test* has been conducted, rendering an χ^2 of 56.44 and 45.22 for $RSME$ in train and test data and 58.45 and 40.84 for R^2 in train and test data, which are considered significant ($p < 10^{-4}$).Additionally, *Nemenyi's Post-Hoc Test* tests were conducted and revealed that, in the case of $RSME$ in test data:

- $CART$ performs significantly different than *lasso*, RF, ν-SVM and *linear-SVM*, with *p-values* 0.029,0.001,0.001 and 0.029 respectively.
- *mlp* performs significantly different than RF and ν-SVM with *p-values* 0.007 and 0.017 respectively.
- ϵ-SVM performs significantly different than RF and ν-SVM with *p-values* 0.022 and 0.049 respectively.

After evaluating the results, two models stood out from the others: $RandomForest$ and the ν-SVM. Although ν-SVM has a slightly higher $RMSE$ value, the Random Forest algorithm was chosen, as the $RMSE$ difference is approximately 0.005 while the $Random\ Forest\ R^2$ value is approximately 0.508 (out of 1) higher than ν-$SVM\ R^2$ value and also present less variability. Another conclusion is that $CART$ is the worse model and the unique model presenting overfitting. Summarising, $Random\ Forest$ the best model for predicting EI from GSR and Pulse signals provided by *BiTalino*, with an $RMSE$ of 0.172 and an R^2 of 0.858 on test set. The figure 3 shows an example of algorithm prediction.

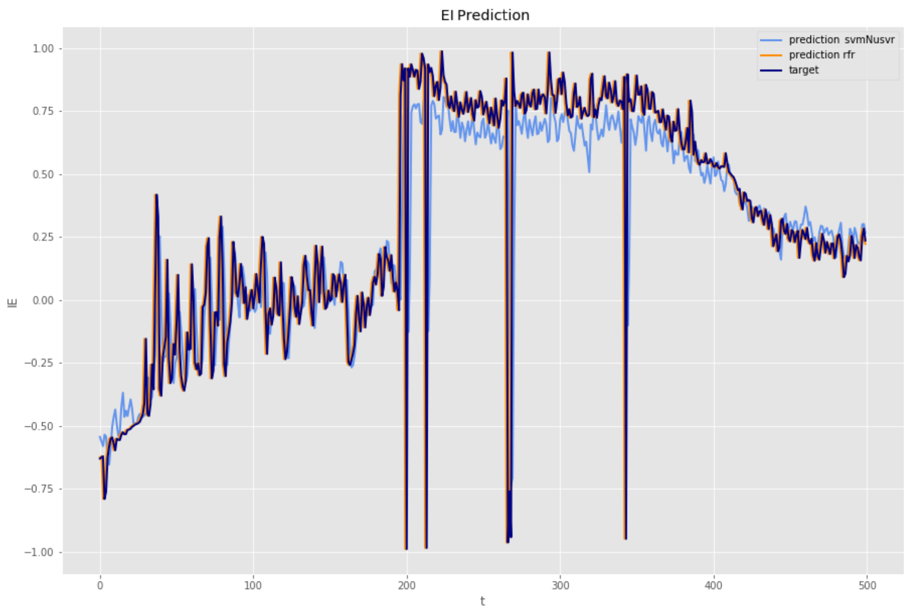

Fig. 3. Prediction of EI values with the two best models trained for low-cost device *BITalino*.

7 Conclusions

In this work, a methodology, based on Machine Learning techniques, for building models for emotions detection with low-cost hardware. As low-cost hardware, *BiTalino* from PLUXWireless Biosignals, S.A. has been chosen, and the results obtained show a good performance of the models obtained, producing reliable predictions of the Emotional Index EI very close to those obtained by medical certified equipment as the *Nesux-10-MRKII* of MindMedia. Another important advantage is that, through the process described here, a big part of the signal processing stack could be avoided.

Another important conclusion, based on model performance measures, is that the use of lagged variables, in our case 6 (two for each time series) is a good approach to overcome problems due to noise in signal acquisition.

Among future works, we are working in real-time implementation of the generated models. To this end, a real-time version of the two filters considered are being implemented. Apart from this, new experiments are being scheduled to increase the size of data sets. Another line is focused on the implantation of the filters on hardware or firmware.

References

1. Altieri, A., Ceccacci, S., Mengoni, M.: Emotion-aware ambient intelligence: Changing smart environment interaction paradigms through affective computing. In: International Conference on Human-Computer Interaction. pp. 258–270. Springer (2019)
2. Benedek, M., Kaernbach, C.: A continuous measure of phasic electrodermal activity. Journal of neuroscience methods **190**(1), 80–91 (2010)
3. Benedek, M., Kaernbach, C.: Decomposition of skin conductance data by means of nonnegative deconvolution. Psychophysiology **47**(4), 647–658 (2010)
4. Bradley, M.M., Lang, P.J.: Measuring emotion: behavior, feeling, and physiology. In: Lane, R.D., Nadel, L. (eds.) Cognitive Neuroscience of Emotion, chap. 11, pp. 242–276. Oxford university press (2000)
5. Calvo, R.A., D'Mello, S., Gratch, J.M., Kappas, A.: The Oxford handbook of affective computing. Oxford University Press, USA (2015)
6. Cartocci, G., Modica, E., Rossi, D., Maglione, A.G., Venuti, I., Rossi, G., Corsi, E., Babiloni, F.: A pilot study on the neurometric evaluation of effective and ineffective antismoking public service announcements. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 4597–4600. IEEE (2016)
7. Damasio, A.R.: Descartes' error: emotion, reason, and the human brain. G.P. Putnam (1994)
8. Damasio, A.R.: The Feeling of What Happens: Body and Emotion in the Making of Consciousness. G.P. Putnam (1994)
9. Gürkök, H., Nijholt, A.: Affective brain-computer interfaces for arts. In: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction. pp. 827–831. IEEE (2013)
10. Hamann, S.: Mapping discrete and dimensional emotions onto the brain: controversies and consensus. Trends in cognitive sciences **16**(9), 458–466 (2012)
11. Kop, W.J., Synowski, S.J., Newell, M.E., Schmidt, L.A., Waldstein, S.R., Fox, N.A.: Autonomic nervous system reactivity to positive and negative mood induction: The role of acute psychological responses and frontal electrocortical activity. Biological psychology **86**(3), 230–238 (2011)
12. Kosiński, J., Szklanny, K., Wieczorkowska, A., Wichrowski, M.: An analysis of game-related emotions using emotiv epoc. In: 2018 Federated Conference on Computer Science and Information Systems (FedCSIS). pp. 913–917. IEEE (2018)
13. Kutt K., Binek W., M.P.N.G.B.S.: Towards the development of sensor platform for processing physiological data from wearable sensors. In: Rutkowski L., Scherer R., K.M.P.W.T.R.Z.J. (ed.) Artificial Intelligence and Soft Computing. ICAISC 2018. Lecture Notes in Computer Science, vol. 10843, pp. 168–17. Springer (2018)

14. Lara-Cabrera, R., Camacho, D.: A taxonomy and state of the art revision on affective games. Future Generation Computer Systems **92**, 516–525 (2019)
15. Luneski, A., Konstantinidis, E., Bamidis, P.: Affective medicine. Methods of information in medicine **49**(03), 207–218 (2010)
16. Mauss, I.B., Robinson, M.D.: Measures of emotion: A review. Cognition and emotion **23**(2), 209–237 (2009)
17. Nalepa, G.J., Kutt, K., Giżycka, B., Jemioło, P., Bobek, S.: Analysis and use of the emotional context with wearable devices for games and intelligent assistants. Sensors **19**(11), 2509 (2019)
18. Ortony, A., Clore, G.L., Collins, A.: The cognitive structure of emotions. Cambridge university press (1990)
19. Picard, R.: Affective computing. cambridge, massachustes institure of technology. The MIT Press (1997)
20. Rouast, P.V., Adam, M.T., Chiong, R., Cornforth, D., Lux, E.: Remote heart rate measurement using low-cost rgb face video: a technical literature review. Frontiers of Computer Science **12**(5), 858–872 (2018)
21. Russel, J.A.: A circumplex model of affect. . Journal of Personality and Social Psychology (39), 11611178. (1980)
22. Santos, O.C., Saneiro, M., Boticario, J.G., Rodriguez-Sanchez, M.C.: Toward interactive context-aware affective educational recommendations in computer-assisted language learning. New Review of Hypermedia and Multimedia **22**(1-2), 27–57 (2016)
23. Singh, G., Bermúdez i Badia, S., Ventura, R., Silva, J.L.: Physiologically attentive user interface for robot teleoperation: real time emotional state estimation and interface modification using physiology, facial expressions and eye movements. In: 11th International Joint Conference on Biomedical Engineering Systems and Technologies. pp. 294–302. SCITEPRESS-Science and Technology Publications (2018)
24. Vecchiato, G., Cherubino, P., Maglione, A.G., Ezquierro, M.T.H., Marinozzi, F., Bini, F., Trettel, A., Babiloni, F.: How to measure cerebral correlates of emotions in marketing relevant tasks. Cognitive Computation **6**(4), 856–871 (2014)
25. Vecchiato, G., Maglione, A.G., Cherubino, P., Wasikowska, B., Wawrzyniak, A., Latuszynska, A., Latuszynska, M., Nermend, K., Graziani, I., Leucci, M.R., et al.: Neurophysiological tools to investigate consumers gender differences during the observation of tv commercials. Computational and mathematical methods in medicine **2014** (2014)