Intelligent Environments 2016
P. Novais and S. Konomi (Eds.)

doi:10.3233/978-1-61499-690-3-416

Monitoring Electrodermal Activity for Stress Recognition Using a Wearable

Arturo MARTÍNEZ-RODRIGO [a,1], Antonio FERNÁNDEZ-CABALLERO [b], Fábio SILVA [c] and Paulo NOVAIS [c]

[a] *Universidad de Castilla-La Mancha, Instituto de Tecnologías Audiovisuales, Spain*
[b] *Universidad de Castilla-La Mancha, Instituto de Investigación en Informática, Spain*
[c] *Universidade do Minho, Departamento de Informática, Portugal*

Abstract. This paper introduces a novel and unobtrusive wearable monitoring device to be used for stress recognition. The paper offers details on the hardware design as well as on the software processes needed to classify between calm and stressed subjects. Only one physiological variable, the electrodermal activity, is used in this work. Moreover, the experimentation, training and testing of the mathematical classification model are also provided. A correctness of 82.5% is achieved by the combination of three skin conductivity response features. It is also interesting to highlight that the proposed approach reports a sensitivity of 89% when differentiating stressed from calm people.

Keywords. stress recognition, electrodermal activity, wearable

1. Introduction

Nowadays, emotion intelligence plays a key role in improving Human-Machine Interaction (HMI). One objective of HMI is to fill the gap between the emotional state of the human and the reaction of the computer in accordance with this behaviour. In this relationship, emotion recognition is the first step towards the development of Affective Computing (AC) systems. Indeed, emotion-aware computing builds emotional interactions between a human and a computer by developing computational models. The main objective consists in creating affective user interfaces to improve the user experience [1].

From a psychological point of view, emotions are defined by a discrete or dimensional model. In the discrete model, a finite set of emotions are defined. Some of these emotions are happiness, surprise, sadness, fear, disgust and anger [2]. Nevertheless, the number of defined emotions depends on the proper model. Thus, you can find a model with up to fifty different emotions [3] in the literature. Now, in dimensional models, emotions are defined by two (or more) variables. The most used dimensional model was proposed by Russell [4] in the eighties. Here, emotions are defined by the levels of arousal and valence. Arousal corresponds to the degree of excitement or calmness that a stimulus produces, while valence assesses how pleasant or unpleasant a stimulus is to the

[1] Corresponding Author: Universidad de Castilla-La Mancha, Instituto de Tecnologías Audiovisuales, 16071-Cuenca, Spain; E-mail: arturo.martinez@uclm.es.

subject [4]. Albeit a high number of emotions can be defined by combining different levels of arousal and valence, negative stress (high arousal, low valence) has gained a special attention in the last years [5]. The main reason relays in the fact that stress may interfere negatively in hazardous tasks. In this regard, stress has been measured during a real-world driving task to determine a driver's relative stress level [6]. Similarly, stress has been studied on aircrew to quantify the performance during military strategies [7], and in medicine for the duration of stressful surgical procedures [8]. Short stress episodes may not suppose any risk for the health of people, being even advantageous for facing some complex situations. Indeed, stress is considered as a reaction from a calm to an excited state for the sake of protecting ourselves [6]. Nevertheless, when people suffer from stress in a long term, it may suppose a high risk and trigger some kind of mental disease such as generalised anxiety disorder and depression [9]. Moreover, some physical disorders seem to be related with a continuous stress, like irritable bowel syndrome, gastro-esophageal reflux disease or back pain [10]. Furthermore, the relationship between stress and cardiac complications like hypertension or coronary artery disease [9] is well-known.

From a neuro-physiological point of view, there exists several aspects that cause physical or mental stress on people. Nevertheless, each individual feels a different stress degree when facing a same stimulus. In this sense, each personality type has a different impact in the response to stress [11]. Indeed, stress affects differently each person because humans are all different from each others. Since they have different personalities and come from several cultures, they show different personal histories [11]. However, regardless of the subjects' personalities, neuro-physiological researchers coincide that feeling stress leads to changes in physical and physiological activity [12]. Therefore, the development of new devices capable of detecting and discriminating stress are very useful in the improvement of HMI, as well as in the creation of intelligent environment where the control of stress is crucial.

In this work, a novel mathematical algorithm capable of detecting stress is introduced. Moreover, the associated hardware design, as well as the training and test model to be embedded in the wearable, are included. Finally, the classification performance is described and some conclusions are commented. The main objective of this work is the description of a system, and its associated architecture, that is useful to take certain actions in order to control or decrease the stress level in multiple environments.

2. Stress Recognition Using Physiological Signals

In the last years, there has been a growing number of works that determine stress by means of physiological signals [13]. On the contrary that physical variables like facial and speech recognition, physiological features are more robust against cross-cultural discordances, since it is almost impossible to control the physiological reactions. Moreover, acquisition sensors can be easily embedded in the environment, this way improving their acceptance. Indeed, recent advances in micro-controllers of ultra low power consumption allow to insert the sensors on clothes. In addition, such hardware architecture permits an intelligent management of the consumption, increasing the operation time [14]. On the other hand, the emergence of devices with transceivers integrated in the own micro-controller, as well as efficient communication protocols, enables the development of sensor networks that communicate with each other.

All these novelties allow a continuous monitoring of different physiological signal simultaneously. Some of the most used variables in stress management are Heart Rate Variability (HRV), Electro-Dermal Activity (EDA), ElectroMyoGraphy (EMG), and ElectroEncephaloGraphy (EEG), among others. All of them are in some way connected with the Autonomous Nervous System (ANS) and, consequently, they are representative of physiological changes produced in the human body. Nevertheless, the suitability of each physiological variable has to be studied. Thus, it has been reported that EMG measures unconscious myo-neural responses, thus assessing exclusively partial aspects of emotions (e.g., arousal level) [15]. On the other hand, EEG seems to be the best option to measure stress. Indeed, EEG detects a wide range of the emotional state dynamics by directly accessing the fundamental structure in the brain [5]. It seems logical to measure within the source where emotions are originated. However, although new devices are emerging in the last years, EEG is still far from being considered wearable. Furthermore, HRV is linked with the sympathetic and para-sympathetic components of the ANS. Although it has been reported that HRV is able to access some subspaces of emotion [15], HRV is highly affected by confounding variables like exercise, heart disease, allergic reactions and inhalation of irritants, among others [16]. Finally, EDA variable has been validated as suitable to measure stress and calm states [6,17]. Indeed, EDA is linked with the sympathetic component and reflects the changes produced by an increase of alert [18].

Therefore, in this work the hardware development and the mathematical classification model are focused in conditioning the signal and maximising the performance through using exclusively the EDA variable.

3. Description of the Wearable

A deep study of the behaviour of EDA signals is required to carry out the electronic design. Indeed, the choice of electronic components thoroughly depends on the electric characteristics of the signal under study. EDA measures the conductivity changes produced in the skin due to the increase of the activity of the sweat glands. An EDA signal is formed by the superposition of two different components. On the one hand, the Skin Conductivity Response (SCR) corresponds to the activation of the sudomotor nerve, which is fired by the ANS. Given this relationship, the SCR component has been previously used to measure stress [19]. SCR is represented morphologically as a peak or burst of peaks with different durations, intensities and decays depending of the stimulus produced. On the other hand, the Skin Conductance Level (SCL) corresponds to the baseline, overlapped to the SCR component. SCL variates due to several aspects like skin type, genetic aspects, and autonomic regulation [13] and, consequently, it is different for each person. Therefore, SCL uses to act as a confounding variable when developing a model for classifying stress in different subjects.

In our design, the EDA sensor measures the DC exosomatic EDA through a couple of Ag/AgCl disc electrodes with a contact diameter of 10 mm, as it can be appreciated in Figure 1. A small DC current is applied into the skin through the electrodes. The amount of current applied is limited to 10 mA/cm^2 to prevent damage in the sweat glands [20]. The SCL component is rejected in origin by applying an electronic third-order Butterworth low-pass filter in a Sallen-Key topology, and with a cut-off frequency of

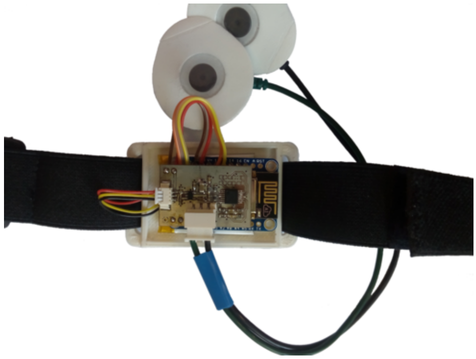

Figure 1. A photograph of the wearable for electrodermal activity acquisition.

0.05 Hz. Similarly, the SCR component is acquired by applying a similar low-pass filter with 1.5 Hz cut-off frequency. These frequencies correspond to the reported bandwidth of both EDA components [21]. Finally, the SCR component is amplified by means of a single-supply, rail-to-rail precision operational amplifier based on the AD860x (Analog Devices).

Once the signal has been acquired, filtered and amplified, a 12-bit Analog-to-Digital Converter (ADC) is used to register the signal with a sampling rate of 10 Hz. The micro-controller rules the global system and controls the sampling frequency for signal acquisition. In this work, an 80 Mhz low-power 32-bit ESP2866 micro-controller is used [22]. This micro-controller is specifically designed for mobile, wearable electronics and Internet of Things applications. The system is capable of managing the power consumption by acting on the power saving architecture and changing the operation mode. Moreover, the iBus RAM/ROM interface permits access to an external flash memory to store the data. The most interesting feature of the system relays in that a 2.4 GHz transmitter/receiver transept is completely embedded in the micro-controller. Therefore, the co-existence of wireless technologies such as Wi-Fi or Bluetooth is permitted. Moreover, the micro-controller features contemplate the use of network protocols for high connectivity like TCP/IP with DNS support over 802.11 b/g/n protocol. In this work, Wi-Fi technology is used to communicate the wearable with the control station, as well as with the rest of devices foreseen within the same network.

Having this context on mind, during the training step the wearable is connected with a base station that records the SRC signal and registers the timestamps for the elicited stimuli. This data is posteriorly processed and characterised to compute the classification model (see Section 5). Nevertheless, it is worth noting that the classification model is embedded into the wearable in a final version, where the device communicates with the rest of nodes that connected in the same network, in order to provoke changes in the environment.

4. Experimental Setup

In order to assess the functionality of the stress detection proposal, this section introduces the experimental design and the description of the study population, respectively.

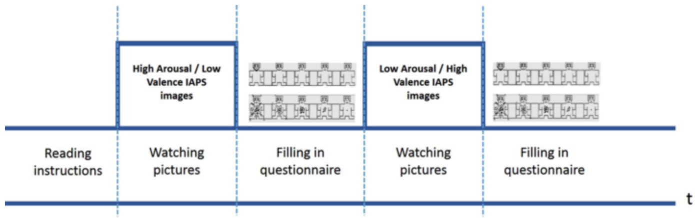

Figure 2. Phases of the experiment.

4.1. Experimental Design

An experiment has been designed with the objective of assessing the correct performance of the described system for stress detection. The core of the experimentation relies in the use of pictures from the International Affective Picture System (IAPS). These have been chosen in order to trigger predefined arousal and valence values [23]. The IAPS database is used in our case to show a series of images to some volunteer participants. All the pictures used in the experiment belong to one of the two following categories: high arousal-low valence and low arousal-high valence. These corresponding to stressed and calm, respectively, according to the circumplex affect model by Russell [4]. In addition, the images taken as representative for the calm category possess a given arousal value lower than 4 and a valence value in the range 4 to 6. In a similar manner, the stressed category of images consists in samples with an arousal value higher than 5 and a valence value lower than 3.

Next, we describe the way the experiment has been run (see Figure 2). A given participant sits in front of a computer monitor and the devices are put in the right wrist. The monitor is a high resolution 28 inches screen. Once the correct functioning of the devices and their communication with the software has been verified, the experiment can start. First of all, the participant carefully reads the instructions of the experiment. Then, the participant is shown ten pictures that are labeled with high arousal and low valence consecutively during 6 seconds. Blank images, which are used as silences, are inserted between each pair of consecutive images with a fixed duration of 1 second. The pictures are randomly selected from those images fulfilling the condition. Next, the participant is offered a distracting task in order to modify his/her emotional state to neutral. This task consists in asking the participant to express his/her feelings in terms of arousal and valence by using the SAM pictographs. Afterwards, the experiment randomly shows another set of ten low arousal/high valence valued images from IAPS. Again, silences are used between each pair of images. Lastly, the distracting task is offered a second time.

The duration of the experiment for each participant longs 140 seconds from the appearance of the first image. The pauses are designed to allow the participant to recover from the previously provoked affect. Notice that the pictures are randomly shuffled, such that the visualisation order is different for each participant. It is also important to highlight that the experiments are carried out in the safest possible way. This is why the participants are informed that they can stop visualising the sequence whenever necessary.

4.2. Study Population

Fifty participants (twenty-eight men and twenty-two women, aged 20 to 28) have enrolled as volunteers in the described experiment. All the participants were informed about the high emotional contents of some pictures and they agreed to take part in the study. All of them were students from the Cuenca Technical School (Spain). Before being accepted in the study, the students had to pass the PHQ-9 Depression Test Questionnaire. Four students were not accepted, and one experiment was not valid due to technical problems. So, the number of valid experiments was forty-five (twenty-five males and twenty females).

This study was carried out in accordance with the ethical standards of the responsible institutional committee on human experimentation. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

5. Feature Extraction

During the model's training phase data from the experiment are recorded and stored in the base station. Then, the data are post-processed in order to extract the most representative features for the prediction model. Firstly, the SCR signal is filtered again by using a 1.5 Hz cut-off low-pass FIR filter in order to decrease noise produced in the different electronic stages. Afterwards, different temporal features are extracted from the SCR signal for the two groups under study, that is calm and stressed. A total of ten time-related features are calculated over the SCR signal. More concretely, the mean (MSC), standard deviation (SDSC), maximum (MASC), minimum (MISC) and dynamic range (DRSC) are estimated. In order to highlight the sudden changes in skin conductivity, also the means of the first (FDSC) and second derivative (SDSC) of SCR are computed. Finally, also the integral (INSC), normalised average power (APSC) and normalised root mean square (RMSC) of SCR are calculated with the aim of identifying the morphological alterations produced when SCRs are present in the EDA signal.

6. Results

Shapiro-Wilk test was used to assess distribution normality for each single parameter, whereas homocedasticity was verified with Levene's test. Consequently, Table 1 shows the statistical differences between both calm and stressed groups throughout a one-way ANOVA test. Only those parameters that show some statistical significance are presented ($\rho \leq 0.05$). In this sense, all parameters, except MISC and SDSC, show some discriminatory power. Furthermore, all relevant parameters report comparable differences among the groups' means.

In addition, each feature's individual ability to discriminate between groups is evaluated by using a Receiver Operating Characteristic (ROC) curve. It is important to highlight that a ten-fold stratified cross-validation is used in this analysis. The reason for stressing the data is to avoid the possibility of classification results to be highly dependent on the choice of a given training-test segmentation. More precisely, the entire database is separated into ten equally-sized folds, such that each fold is a good representative of the data. Then, ten learning and test iterations are performed. At each iteration, nine folds are randomly used to train the system, while the remaining fold is used to test the performance. Furthermore, at each iteration a threshold is obtained for maximising the ROC performance. Such threshold is then used to assess the performance on the test set.

Table 1. Sensitivity (Se), specificity (Sp) and accuracy (Ac) for all parameters under study, as well as for training and test subsets, are shown. Moreover, the statistical significance through one-way ANOVA test is depicted.

Feature Acronym	**One-way ANOVA**	**Learning** Se (%)	Sp (%)	Ac (%)	**Test** Se (%)	Sp (%)	Ac (%)
MSC	1.03×10^{-5}	75.95	83.10	79.52	69.00	76.78	**72.95**
SDSC	1.33×10^{-5}	85.06	78.95	82.43	78.57	76.07	**77.38**
MASC	2.68×10^{-5}	81.45	76.35	78.91	74.07	72.28	**73.34**
DRSC	2.67×10^{-5}	82.92	75.42	79.18	76.64	71.00	**73.62**
FDSC	1.50×10^{-4}	74.58	80.43	77.50	69.21	74.85	**71.81**
INSC	1.11×10^{-5}	79.40	79.03	79.25	67.57	74.78	**71.00**
APSC	0.0026	84.98	79.88	82.43	79.92	78.71	**79.28**
RMSC	1.20×10^{-5}	85.92	78.92	82.42	81.57	78.85	**79.94**

All the significant parameters achieve a considerable performance in the classification. Statistical parameters show a comparable classification power, where the accuracies range from 71% to 77%. SDSC reports the highest accuracy among the statistical parameters. Nevertheless, the markers based on signal power show better performance than other statistical parameters. Indeed, RMSC achieves the highest performance, classifying correctly 79.94% of the subjects. However, INSC shows a distinct trend, being the worst parameter in the classification with 71% of correctness.

In order to increase the classification performance and to study the possible relationship among all the features, the parameters are combined by means of a tree-based classification model. These models are characterised by their simplicity and suitability to be embedded into small devices where computational capacity is limited. Indeed, in such models the decisions are taken by simple binary conditions, where different parameters are involved. In this work, the criteria to build up the tree is that any node only contains samples from the same group or less than 20% of the samples belonging the same group. Furthermore, the splitting criteria, which defines the number of tree levels, is performed by following the Gini index. Figure 3(a) shows the decision tree structure, calculated by using a ten-fold cross-validation technique to avoid biased results dependent of the training and test subsets. It is important to remark that all the significant parameters were included in this analysis. However, only tree parameters were selected, when the aforementioned building criteria were accomplished. Thus, SDSC is chosen as the most significant parameter to maximise the classification. Then, MSC and RMSC are also chosen to discriminate among groups. It is worth noting that these parameters show a relevant performance when analysed individually. In this regard, a global accuracy of 82.5% is achieved by the combination of the three temporal parameters, as shown in Figure 3(b). Moreover, the global performance increases more than 2.5% regarding the accuracy achieved by the best single parameter. It is also important to remark that the ratio of correctness within stressed subjects increases up to 89%, augmenting the sensitivity more than 9%. On the contrary, the specificity of the model decreases less than 3%.

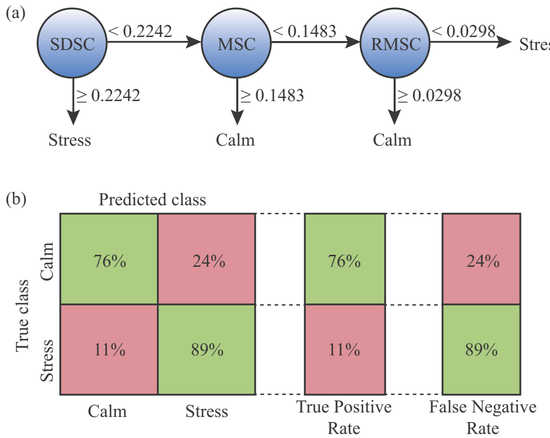

Figure 3. Combination of temporal parameters. (a) Decision tree structure. (b) Confusion matrix.

7. Conclusions

This work has introduced a novel and unobtrusive wearable monitoring device to be used in stress control environments. The details about the hardware, experimentation, training and testing of the mathematical classification model have been provided. Although only one physiological variable (EDA) has been used in this work, a correctness of 82.5% has been achieved by the combination of three SCR features. It is also interesting to highlight that the proposed model reports a sensitivity of 89% when differentiating stressed from calm people. In this sense, it is preferable to reach a high sensitivity, even at the expense of decreasing the specificity. Indeed, in terms of applicability, it is better to treat calm people with relaxing stimuli than not treating stressed subjects.

A number of works can be found in the literature covering the calm-stress differentiation. Nevertheless, only a few studies have focused their efforts on portable devices [17]. The main reason is that most of the works use complex classification approaches to raise the performance, using support vector machines or neuronal networks. In this regard, wearables need a special attention to keep the computational burden as low as possible. In this work, much efforts have been put to reduce the amount of information to be processed. Thus, only one physiological variable is acquired and processed. Moreover, all the features are calculated in the time domain, where the operations are reduced to simple sums and divisions. These specifications permit not only to embed the whole analysis into the wearable, but also analyse the data in real-time. Nevertheless, in order to decrease the consumption and to increase the wearable's autonomy, the device is in sleep mode most of the time. Only in concrete periods, the device wakes up, runs its routine and goes back to sleep. This sequence allows a drastic reduction in power consumption, specially in wireless systems where radio communication is by far the most expensive element within the whole system.

The applicability of this kind of systems are enormous, given the wide range of situations where stress control plays a key-role. Moreover, the continuous advances in chip miniaturisation and connectivity between devices makes ambient intelligence systems more unobtrusive and ubiquitous,offering a better acceptance by the user. Figure 4 shows a simple system where this application can be used. It consists of a home where the de-

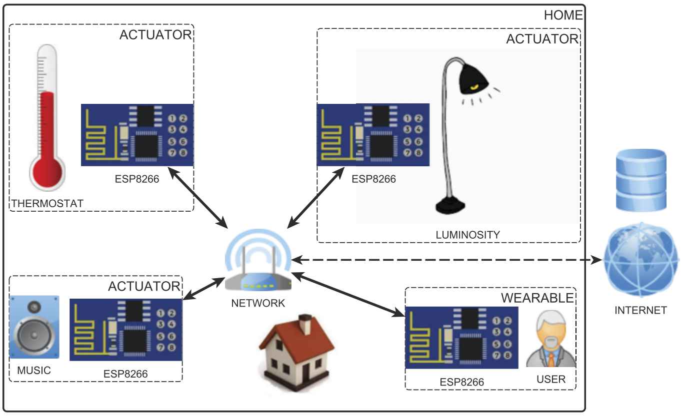

Figure 4. A possible application of the proposal.

vice is worn by the user. It evaluates the stress level each certain time and communicates with the rest of the nodes operating as actuators within the system. Thus, background music, lighting and temperature are regulated depending on the stress level of the user. Indeed, these variables have reported certain degree of influence in stress control [11].

Acknowledgements

This work was partially supported by Spanish Ministerio de Economía y Competitividad / FEDER under TIN2013-47074-C2-1-R and TIN2015-72931-EXP grants and PPII-2014-026-P from Junta de Comunidades de Castilla-La Mancha.

References

[1] A. Luneski, P. D. Bamidis, M. Hitoglou-Antoniadou, Affective computing and medical informatics: state of the art in emotion-aware medical applications, Studies in Health Technology and Informatics 136 (2008) 517–522.
[2] P. Ekman, An argument for basic emotions, Cognition & Emotion 6 (3-4) (1992) 169–200.
[3] M. Schröder, R. Cowie, Towards emotion-sensitive multimodal interfaces, humaine, réseau d'excellence (2004).
[4] J. A. Russell, A circumplex model of affect, Journal of Personality and Social Psychology 39 (6) (1980) 11–61.
[5] M.-K. Kim, M. Kim, E. Oh, S.-P. Kim, A review on the computational methods for emotional state estimation from the human eeg, Computational and Mathematical Methods in Medicine 2013 (2013) article ID 573734.
[6] J. A. Healey, R. W. Picard, Detecting stress during real-world driving tasks using physiological sensors, IEEE Transactions on Intelligent Transportation Systems 6 (2) (2005) 156–166.
[7] M. J. Skinner, P. A. Simpson, Workload issues in military tactical airlift, The International Journal of Aviation Psychology 12 (1) (2002) 79–93.
[8] D. G. Duru, A. Deniz Duru, D. E. Barkana, O. Sanli, M. Ozkan, Assessment of surgeon's stress level and alertness using eeg during laparoscopic simple nephrectomy, in: 6th International IEEE/EMBS Conference on Neural Engineering, IEEE, 2013, pp. 452–455.
[9] C. Tucker-Ladd, Stress, trauma, anxiety, fears and psychosomatic disorders, in: Psychological Self-Help, 2008, p. chapter 5.

[10] H. Mönnikes, J. Tebbe, M. Hildebrandt, P. Arck, E. Osmanoglou, M. Rose, B. Klapp, B. Wiedenmann, I. Heymann-Mönnikes, Role of stress in functional gastrointestinal disorders, Digestive Diseases 19 (3) (2001) 201–211.
[11] M. Gomes, T. Oliveira, F. Silva, D. Carneiro, P. Novais, Establishing the relationship between personality traits and stress in an intelligent environment, in: Modern Advances in Applied Intelligence, Springer, 2014, pp. 378–387.
[12] R. A. Calvo, S. D'Mello, Affect detection: An interdisciplinary review of models, methods, and their applications, Affective Computing, IEEE Transactions on 1 (1) (2010) 18–37.
[13] G. Valenza, A. Lanata, E. P. Scilingo, The role of nonlinear dynamics in affective valence and arousal recognition, IEEE Transactions on Affective Computing 3 (2) (2012) 237–249.
[14] J. Li, X. Nie, Z. Jiang, Development specifications for arm application systems (2003).
[15] I. B. Mauss, M. D. Robinson, Measures of emotion: a review, Cognition & Emotion 23 (2) (2009) 209–237.
[16] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, J. S. Suri, Heart rate variability: a review, Medical and Biological Engineering and Computing 44 (12) (2006) 1031–1051.
[17] A. de Santos Sierra, C. S. Ávila, J. G. Casanova, G. B. D. Pozo, A stress-detection system based on physiological signals and fuzzy logic, Industrial Electronics, IEEE Transactions on 58 (10) (2011) 4857–4865.
[18] A. Martínez-Rodrigo, R. Zangróniz, J. M. Pastor, A. Fernández-Caballero, Arousal level classification in the ageing adult by measuring electrodermal skin conductivity, in: Ambient Intelligence for Health, Springer, 2015, pp. 213–223.
[19] A. Martínez-Rodrigo, R. Zangróniz, J. M. Pastor, M. V. Sokolova, Arousal level classification of the aging adult from electro-dermal activity: From hardware development to software architecture, Pervasive and Mobile Computing (2016) in press.
[20] R. Edelberg, T. Greiner, N. R. Burch, Some membrane properties of the effector in the galvanic skin response, Journal of Applied Physiology 15 (4) (1960) 691–696.
[21] W. Boucsein, Electrodermal Activity, Springer Science & Business Media, 2012.
[22] Q. M. Ashraf, M. Yusoff, I. Mohd, A. A. Azman, N. M. Nor, N. A. A. Fuzi, M. Saharedan, N. A. Omar, et al., Energy monitoring prototype for internet of things: Preliminary results, in: Internet of Things (WF-IoT), 2015 IEEE 2nd World Forum on, IEEE, 2015, pp. 1–5.
[23] P. J. Lang, M. M. Bradley, B. N. Cuthbert, International affective picture system (iaps): Affective ratings of pictures and instruction manual, Technical report A-8.