Biomedical Signal Processing and Control 42 (2018) 45–52

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Efficient wavelet-based artifact removal for electrodermal activity in real-world applications

Jainendra Shukla[a,c,*], Miguel Barreda-Ángeles[b], Joan Oliver[a], Domènec Puig[c]

[a] Instituto de Robótica para la Dependencia, Av. Artur Carbonell, 11, 08870 Sitges, Barcelona, Spain
[b] Eurecat, Technology Centre of Catalonia, Planta 9, Av. Diagonal, 177, 08018 Barcelona, Spain
[c] Intelligent Robotics and Computer Vision Group (IRCV), Rovira i Virgili University, Av. dels Paisos Catalans, 26, 43007 Tarragona, Spain

ARTICLE INFO

Article history:
Received 28 July 2017
Received in revised form
18 December 2017
Accepted 17 January 2018

Keywords:
Denoising
Stationary wavelet transform
Electrodermal activity
Wavelet analysis

ABSTRACT

Online monitoring of electrodermal activity (EDA) may serve as an economical and explicit source of information about actual emotional state and engagement level of users during their interaction with information and communications technologies (ICT) applications in *real-world* situations. In such contexts, however, EDA signal is affected by motion artifacts that introduce noise in the signal and can make it unusable. As the scope of movement minimization during EDA data acquisition is limited, this scenario demands online methods for detection and correction of artifacts with low computational cost. We propose an efficient wavelet-based method for artifacts attenuation while minimizing distortions, using a stationary wavelet transform (SWT) modeling the wavelet coefficients as a Laplace distribution. The proposed method was tested on EDA recordings from publicly available driver dataset collected during real-world driving, and containing a high number of motion artifacts, and the results were compared to those of three state-of-the-art methods for EDA signal filtering. In addition, the proposed method was tested for the online filtering of EDA signals collected while 12 volunteers conducted tasks designed to elicit various stress states. The results evidenced that the prediction of arousal states can be significantly improved after motion artifacts removal, and that the proposed method outperforms existing approaches and it has a lower computational cost. Taken together, these results evidence the effectiveness of the proposed method for online EDA filtering in real world scenarios.

1. Introduction

Electrodermal activity (EDA) refers to the changes in conductivity in the skin, which is affected by the activity of the sympathetic branch of the autonomous nervous system and is usually considered as a correlate of psychological processes such as emotional arousal, stress, or cognitive effort ([3], Chapter 3) [10]. It is commonly measured by placing two electrodes in palmar sites of the hand, although also other alternative locations, such as the wrist [28], have been proposed.

Due to its low-cost and easy-to-collect nature, EDA measurement has been commonly used in research in psychology [10] and as a tool for the assessment of user's experience in a variety of contexts such as recreational and serious games [12,26], driving [16], visual discomfort with 3D media [2], internet browsing [1], or patient–robot interaction [31].

EDA signal is a non-stationary signal, with mean levels usually ranging between 2 and 20 μS, and varying within a range between 1 and 3 μS for an individual. Its values typically show a slow decrease over time when the subject is at rest, and increase more rapidly when novel stimulation is introduced, and, once the stimulation is over, gradually decrease again [10]. EDA signal is considered to have two components: a tonic, or general, level, and a phasic component, characterized by more rapid and momentary changes in EDA levels. Such momentary changes, usually associated to the presence of arousing stimuli, are called Skin Conductance Responses (SCRs). SCRs can be seen as peaks over imposed to the tidal drifts in the general EDA levels.

The availability of wearable devices able to measure online EDA provide the opportunity for real world data collection in a comfortable manner in real-life scenarios outside the lab. This allows broadening the scope of EDA analysis from simply gathering information on user's psychological state to the online adaptation of the system according to the user's cognitive and emotional states.

* Corresponding author at: Instituto de Robótica para la Dependencia, Av. Artur Carbonell, 11, 08870 Sitges, Barcelona, Spain.
E-mail addresses: jshukla@institutorobotica.org (J. Shukla), miguel.barreda@eurecat.org (M. Barreda-Ángeles), joliver@institutorobotica.org (J. Oliver), domenec.puig@urv.cat (D. Puig).

https://doi.org/10.1016/j.bspc.2018.01.009
1746-8094/

The advantages of such online adaptation are especially remarkable in the cases of the users with acute problems to identify, report, and regulate their emotions, such as children or individuals with intellectual developmental disabilities (IDD). In this sense, novel cognitive stimulation systems based on the robots or virtual reality may greatly benefit from an online input of the user emotional state from EDA signal (e.g. [23]).

However, compared to laboratory studies, the online collection and analysis of EDA signals in real-life context involves new challenges related to signal processing. While in lab experiments the participants are usually asked to not to move the hand to which the electrodes are attached, hand's movement can be hardly controlled in real-life settings. As a consequence of such movement, partial detachment of the electrodes or pressure over them may occur, leading to the appearance of artifacts in the signal [3]. If these artifacts remain in the signal when it is analyzed they can easily be misinterpreted and skew the analysis [32]; for example, they may be mistaken for a SCR indicating an increased stress, especially when the analysis is conducted automatically. Even in a context in which not much movement should be expected, the presence of such motion artifacts can be critical in the case of users whose ability to control the level of movement is limited (e.g. individuals with IDD).

The remainder of the paper is structured as follows. Section 1.1 discusses the previous researches related with denoising of the EDA signal. Section 2 details the techniques used in this research for wavelet based artifact removal of EDA signals. Section 3 discusses the computational advantages of the proposed method and Section 4 presents the analysis of the method on the publicly available driver dataset. Section 5 discusses the online validation of the proposed method via an experiment involving stress eliciting tasks. Finally, Section 6 concludes the article by summarizing this work and highlighting the future steps.

1.1. Related work

Methods used in previous research to correct artifacts mainly consist of exponential smoothing [17] and low-pass filtering [24]. Over the last two decades, wavelets have already proven their significant value in signal processing and image. Wavelets have also been found suitable for EDA activity modeling, given its non-stationary behavior [22]. Wavelet based sophisticated de-noising of motion artifacts have been widely used in research [19,25,31,6] and has offered better results due to the good localization property of the wavelet transforms [25]. On the other hand, other sophisticated methods, not based in wavelets, have also been proposed. Recently, Greco et al. [14] proposed a deconvolution based approach (called cvxEDA) using Maximum a Posteriori (MAP) estimation and convex optimization which provided a decomposition of the EDA that is robust to noise.

However, the aforementioned techniques present several important limitations when used for online filtering of the EDA signals, such as the following:

1. Exponential smoothing [17] and low pass filtering based denosing [24] methods are not able to atone the unexpectedly occurring artifacts which have higher values than EDA and undiscriminated filtering of the whole signal also distorts the EDA signals without artifacts [6].
2. Denoising with the traditional DWT wavelet transform [31] can exhibit visual artifacts due to lack of translation invariance and "pseudo-Gibbs" oscillations are especially pronounced in the vicinity of discontinuities [9].
3. Estimation of the noise level σ as proposed by Swangnetr and Kaber [31] is based on the data collected during the rest period. This type of noise level estimation is an off-line and static measure and is an overhead cost on denosing. Moreover, as the actual nature of noise in the noisy signal is dynamic, hence the noise level estimation needs to be done online.
4. Gaussian mixture based modeling for the distribution of the wavelet coefficients [6] requires estimation of three model parameters, γ_j the mixture parameter, σ_j^2 and $c^2\sigma_j^2$ the variances of the two Gaussians (Eq. (1)) which demand employment of iterative algorithms such as Expectation Maximization (EM) algorithm. These algorithms have high computational complexity of $O(n+2n^2)$ for 2 mixture of Gaussians of one dimensional data as in the current case, where n is the number of data samples [7].

$$\tilde{d}_j \sim \gamma_j N(0, \sigma_j^2) + (1-\gamma_j) N(0, c^2\sigma_j^2) \quad (1)$$

5. The cvxEDA method (Eq. (2)) depends on convex Quadratic Programming (QP) [14] which demands polynomial time algorithms, such as interior point algorithms, for solution. The time complexity of convex QP solving algorithms is $O(n^3)$ [4], where n is the number of data samples. Denosing in online scenarios can be limited by the computational complexities of Chen et al. [6] and Greco et al. [14] methods.

$$y = Mq + Bl + Cd + \epsilon \quad (2)$$

Whereas in the offline analysis of EDA signals visual inspection allows identifying and removing the parts of the recording containing artifacts, online adaptation of systems based on EDA signal requires automated methods for such purpose. These methods not only need to be accurate enough to provide a good signal quality, but also need to be computationally affordable enough to work online.

The aim of the proposed work is to present a wavelet based method for filtering motion artifacts in EDA signal that fits these requirements. Experimental results described in Sections 3, 4.3 and 5.3, demonstrate the benefits of the algorithm through a comprehensive comparison with Swangnetr and Kaber [31], Chen et al. [6] and Greco et al. [14] methods.

2. Method

De-noising with the traditional wavelet transforms can exhibit visual artifacts due to lack of translation invariance [9]. Stationary Wavelet Transform (SWT) is redundant, linear and hence shift invariant in comparison to the Discrete Wavelet Transform (DWT) [27]. SWT also provides better sampling rates in the low frequency bands compared with a standard DWT [27].

The literature suggests that the selection of the mother wavelets can be done based upon their resemblance to either the shape of the signal [15] or the shape of the typical motion artifact [20]. Since EDA signals have asymmetric nature, asymmetrically shaped Daubechies (dbN) wavelets have been used to analyze them [21]. Swangnetr and Kaber [31] suggests that *db*3 is the most appropriate choice of the mother wavelet to represent the EDA signal. *Haar* wavelet has also been used for detecting edges and sharp changes, commonly seen in motion artifacts [6]. In addition, Chen et al. [6] suggests that *Coiflet*3 wavelet can have potential as the basis function, since it resembles the shape of the typical motion artifact. Hence, to analyze the effects of the different basis functions, we examine the *db*3, *haar* and the *coiflet*3 wavelets separately, as a mother wavelet during EDA signal denosing.

As explained in Section 1.1, EDA artifacts may result from the recording procedure or from the physiological responses, hence, artifact removal in EDA for above types is discussed separately.

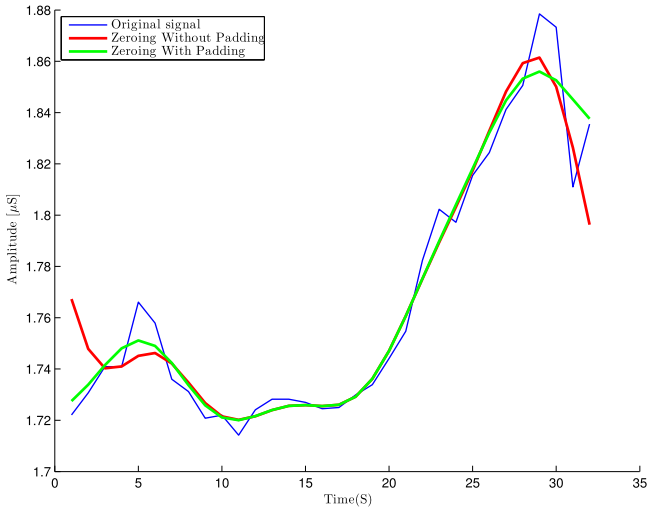

Fig. 1. Low pass filtering using SWT.

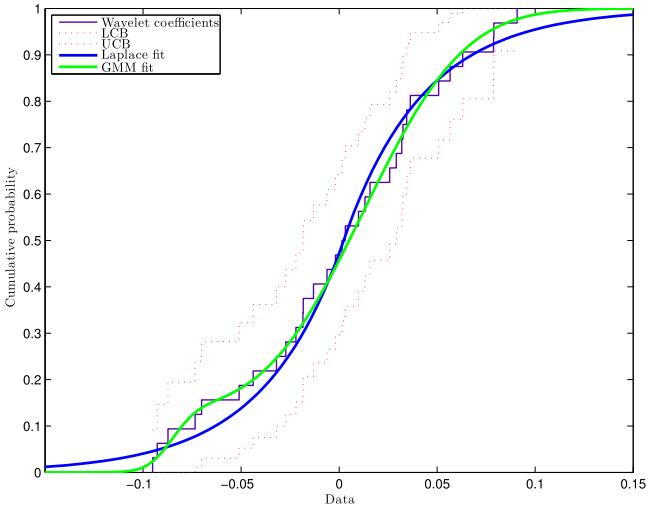

Fig. 2. CDF of wavelet coefficients with Laplace and GMM distribution.

2.1. Artifacts originating from recording

A main source of artifacts is the power line noise resulting from the AC frequency input of 50 Hz (in Europe). Since the typical frequency of the EDA signal is 0.0167–0.25 Hz [10], this type of artifact can be treated as high-frequency noise and hence can be removed as low-pass filtering.

In the proposed approach, EDA signal is modeled via SWT and the level of decomposition J is set to $(log_2 freq + 2)$, where $freq$ is the signal frequency. Detail coefficients representing high frequency (>0.5 Hz) of the signal (i.e. noise) are set to zero, thereby low pass filtering the EDA signal [27]. Inverse SWT is then applied on the coefficients to recover the signal.

The low pass filtering via this method distorts the signal at the beginning and at the end (Fig. 1). To get rid of the distortion, a padding can be applied at the beginning and at the end of the signal before low pass filtering. After applying the extension, the low pass filtering preserves the signal.

2.2. Physiological response based artifacts

The most important source of physiologically based artifacts are motions, which includes skin movements beneath the electrodes and muscular activity occurring in non recording locations [3]. Swangnetr and Kaber [31] addressed the filtering of such artifacts by using a zero-mean Laplace distribution modeling (Eq. (3)).

$$\tilde{d}_j \sim Laplace(0, b) \tag{3}$$

The Laplace probability density function can be represented by Eq. (4), where μ is location parameter and b is a scale parameter. An estimator of μ is the sample median $\hat{\mu}$, which is equal to zero for the distribution of wavelet coefficients. Hence, the wavelet detail coefficients can be modeled using zero-mean Laplace distribution as indicated by Eq. (3).

$$f(x) = \frac{1}{2b} \exp\left(-\frac{|x-\mu|}{b}\right) \tag{4}$$

Chen et al. [6] used a Gaussian mixture based modeling of wavelet coefficients for removing such artifacts (Eq. (1)). As is evident from Eq. (1), wavelet coefficient model proposed by Chen et al. [6] requires estimation of three model parameters, γ_j the mixture parameter, σ_j^2 and $c^2\sigma_j^2$ the variances of the two Gaussians, whereas as suggested by Eq. (3) Laplace distribution modeling requires estimation of a single scale parameter b. Hence, computational complexity of using the Laplace distribution is significantly lower than the Gaussian Mixture Model (GMM) distribution. Fig. 2 displays a plot of the empirical cumulative distribution function (cdf) for the wavelet detail coefficients and the corresponding theoretical cdf for the Laplace and GMM distribution along with the lower and upper confidence bounds for 5% confidence interval.

As can be seen in Fig. 2, both Laplace and GMM distributions model well the wavelet coefficients within the 5% confidence intervals. Wavelet detail coefficient obtained from a sample baseline EDA signal of a publicly available dataset [30] was tested for fitting against Laplace and GMM distribution using the Kolmogorov–Smirnov test. Results confirmed that the detail coefficient distributions were not significantly different from neither the Laplace distribution, nor the GMM distribution ($p > 0.05$). These results reconfirm the propositions presented by Swangnetr and Kaber [31] about Laplace distribution fit and the propositions presented by Chen et al. [6] about the GMM distribution fit for modeling of wavelet detail coefficients.

Laplace fitting of wavelet coefficient requires an estimation of a single scale parameter b, which can be estimated for each level j from the wavelet coefficients of the original signal d, using Maximum-likelihood estimation as per Eq. (5) [29].

$$\begin{aligned} \hat{b} &= \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{\mu}| \\ &= \frac{1}{N}\sum_{i=1}^{N}|x_i|, (\because \hat{\mu} = 0) \end{aligned} \tag{5}$$

As $\hat{\mu}$ of wavelet distribution is zero, Eq. (5) reduces to later form.

Let δ be the proportion of motion artifacts in original signal d, then the thresholds of wavelet shrinkage to remove the noise coefficients from the signal can be obtained by Eq. (6).

$$\Phi(T_{low}) = 1 - \Phi(T_{high}) = \delta/2 \tag{6}$$

where Φ is the cumulative distribution function of the Laplace distribution, and T_{low} and T_{high} are the thresholds. It can be shown that for a zero mean Laplace distribution, T_{low} and T_{high} can be obtained using Eqs. (7) and (8).

$$T_{low} = -T_{high} \tag{7}$$

$$T_{high} = \hat{b} * \log_e(\delta) \tag{8}$$

where $\hat{b}$ is the estimated scale parameter of Laplace distribution.

Based on the points above, a SWT based denosing with *db3*/*haar*/*coiflet3* as the mother wavelet and using Laplace distri-

Table 1
Computational complexity for evaluated methods.

Method	Big O	Running time (mS)	Parameters required
Chen et al. [6]	$O(n+2n^2)$	14.931	3
Greco et al. [14]	$O(n^3)$	21.685	3
Proposed method	$O(n)$	0.017	1

bution for wavelet coefficients modeling is employed for removing motion artifacts from EDA. The procedure is described below:

1. The highest level wavelet detail noise coefficients of the low pass filtered EDA are modeled with the Laplace distribution and the scale parameter $\hat{b}$ of Laplace distribution is estimated using Eq. (5). The T_{low} and T_{high} thresholds of wavelet shrinkage are calculated using Eqs. (7) and (8).
2. Motion artifacts are removed from the EDA signal as per Eq. (9).

$$\hat{d}_j^k = \begin{cases} \hat{d}_j^k & T_{low} \le \hat{d}_j^k \le T_{high} \\ 0 & otherwise \end{cases} \quad (9)$$

3. Steps 1 and 2 are repeated on the wavelet coefficients of all higher levels and then inverse SWT wavelet transform is applied to the thresholded wavelet coefficients to obtain the fully denoised EDA signal.

3. Computational complexity

To assess the computational complexity of the proposed algorithm, it was compared with the computational complexity of Chen et al. [6] and Greco et al. [14] methods. A Matlab implementation was performed for the proposed method and Chen et al. [6] method, while Matlab implementation of the author's version was used[1] for Greco et al. [14] method. Repeated measurements for 10,000 times were performed for all the methods while processing a sample EDA signal from a publicly available dataset [30]. The Matlab implementations of all the three methods were executed on an Intel® Core™ i7-4790 CPU@3.60 GHz × 8 processor with 8 GB of RAM running 64-bit Ubuntu 14.04 LTS operating system. Since the proposed method uses the same Laplace distribution for modeling of wavelet coefficients, as demonstrated in Swangnetr and Kaber [31], computational complexities for both these methods are the same and hence are not compared.

Table 1 presents the theoretical computational complexity in terms of *Big O*, the running time of the compared algorithms in milliseconds and the number of parameters estimation required for each methods. As can be seen from Eq. (5), the running time performance of the Laplace fitting method will grow linearly and in direct proportion to the size of the input data set, hence the computational complexity of Laplace fitting method is $O(n)$. On the other hand, parameters estimation for GMM distribution fitting of wavelet distribution coefficients employs the EM algorithm for fitting a 2 mixture of Gaussians which has the computational complexity of $O(n+2n^2)$[7]. Similarly cvxEDA method proposed by Greco et al. [14] depends on convex Quadratic Programming (QP) which demands polynomial time algorithms, such as interior point algorithms, for solution and has the time complexity of $O(n^3)$ [4]. Since $O(n) < O(n+2n^2) < O(n^3)$, the proposed method has the lowest computational complexity followed by Chen et al. [6] method and Greco et al. [14] method.

The reported running time is the median of the 10,000 repeated measurements. It can be seen that method by Chen et al. [6] requires estimation of three model parameters, γ_j the mixture parameter, σ_j^2 and $c^2\sigma_j^2$ the variances of the two Gaussians (Eq. (1)). cvxEDA approach by Greco et al. [14] also requires estimation of optimal values for three parameters, q an auxiliary variable to find sudomotor nerve activity (SMNA), l the vector of spline coefficients, d a 2×1 vector with the offset and slope coefficients (Eq. (2)). However, the proposed method requires estimation of only a single scale parameter b. Moreover, the computational cost of the proposed method is the lowest among all the three methods. In terms of execution time, it is around 863 times faster than Chen et al. [6] method and is around 1254 times faster than Greco et al. [14] method.

4. Off-line analysis: filtering quality

4.1. Driver dataset

In order to test the proposed method, we used a sample of EDA signals obtained from a publicly available dataset of physiological signals collected from participants while driving on routes in city streets and highways in the area of Boston, Massachusetts [16]. The authors of this dataset recorded, together with other physiological signals, EDA data from 24 drives with duration between 50 and 90 min, with a sampling rate of 16 Hz. However, since the recordings from 9 of those participants were missed, the actual sample contains about 1207 min of EDA recordings from 15 participants. The main reason for choosing this dataset for testing our filtering algorithm, beyond its availability, was that it was recorded during real driving. Therefore, it exemplifies well the type of motion artifacts that can be found when recording EDA in real-world scenarios.

Each signal was divided in non-overlapping 8 s segments, resulting in a total of 9056 segments. Eight seconds segment was used since it is adequate to capture the important features of the EDA signal [10]. A researcher with experience in the analysis of EDA signals visually inspected the signals and annotated which segments contained motion artifacts. Given the smooth waveform of EDA signal and its relatively slow change in magnitude (e.g. rise of a SCR takes 1–3 s in reach a peak), as well as its restricted range of variability (usually in a range of 1–3 mS for a given subject) [10], experts can easily detect artifacts in the signal due to motion. Indeed, visual inspection of the signal is a standard procedure in research using EDA [13]. The expert annotated as noisy those segments containing "sharp" changes in the EDA signal, which does not represent the usual behavior of EDA and are likely to be the product of electrodes motion or partial detachment. A total of 715 segments were considered as containing motion artifacts, while the remaining 8341 segments were considered as motion artifact-free.

4.2. Analysis procedure

The approach for testing the quality of the offline filtering was similar to the one described in Chen et al. [6]. It involves analyzing separately the performance of the filter for noisy and not-noisy segments. For the noisy segments, a metric of the algorithm's artifact power attenuation (APA) was calculated, while for the not-noisy segments, we calculated the normalized mean square error (NMSE) [25]. This way, an overall judgment of the algorithm's performance can be made based on quantitative measures of both artifact power reduction on noisy signals and distortion introduced by the filter in artifact-free signals (Fig. 3).

We adapted the equations from Molavi and Dumont [25] and Chen et al. [6] to better suit the signals of short duration, by eliminating the $\sum$ over artifacts as short duration signals do not carry multiple artifacts. Therefore, APA and NMSE were defined as

$$APA = 10\log_{10}\frac{Var[y]}{Var[\tilde{y}]} \quad (10)$$

[1] https://github.com/lciti/cvxEDA.

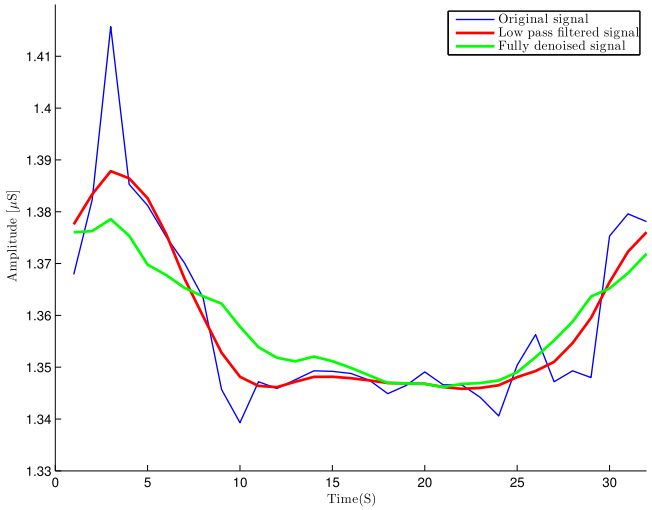

Fig. 3. Original EDA and denoised signals.

$$NMSE = 10\log_{10}\frac{[y-\tilde{y}]^2}{[y-\bar{y}]^2} \quad (11)$$

where y represents the original signal, $\tilde{y}$ represents the filtered signal, and $\bar{y}$ represents the mean value of the y signal. We calculated the APA values for each of the (715) segments with artifacts and the NMSE value for each of the (8341) artifact-free segments, for the output of each of the three versions of the proposed method (*db*3, *haar*, and *coiflet*3), as well as for the same segments filtered following the approaches by Chen et al. [6], the Swangnetr and Kaber [31] method, and the Greco et al. [14] approach. Concerning Swangnetr and Kaber [31] method, the σ of noise in the raw signal and hence the threshold of wavelet shrinkage was estimated based on the 15-min rest periods at the beginning of the drive [16]. For the cvxEDA parameters Greco et al. [14], the default values as suggested in the code provided by the authors, which are $\tau_0 = 0.7$, $\tau_1 = 10$, $\alpha = 0.008$ and $\gamma = 0.01$ were employed for all the subjects. Concerning the Chen et al. [6], and the proposed method, artifact proportion δ was set to 0.10 for all subjects as used by researchers in [6].

4.3. Results

Table 2 shows the median values of the APA and NMSE metrics for each of the compared methods. In terms of artifact power reduction, as expected, the median APA evidences that all the proposed methods performed better than the approach by Swangnetr and Kaber [31]. However, the APA for *db*3 and *coiflet*3 approach is lower than those of the approaches by Chen et al. [6] and Greco et al. [14]. The *haar* version of our method outperforms Chen et al. [6] method and is close to the performance of the Greco et al. [14] method. In order to check the consistency of this difference, we conducted two paired samples *t*-test for comparing the performance of the *haar* version with the performance of the Chen et al. [6] approach and Greco et al. [14] approaches in terms of APA. The results of the *t*-tests showed that the performance of our *haar* method is significantly better than the Chen et al. [6] approach ($t(712) = 7.38$, $p < .001$), while the performance of the *haar* version and the Greco et al. [14] approach was not significantly different at the $p < .05$ level ($t(712) = 1.95$, $p = .05$).

Table 2
Median of NMSE and APA for evaluated methods.

Method	APA (dB)	NMSE (dB)
Chen et al. [6]	5.284	−3.085
Greco et al. [14]	5.887	−1.885
Proposed with *db*3	4.808	−3.515
Proposed with *haar*	5.712	−3.678
Proposed with *coiflet*3	4.766	−3.665
Swangnetr and Kaber [31]	1.624	−11.625

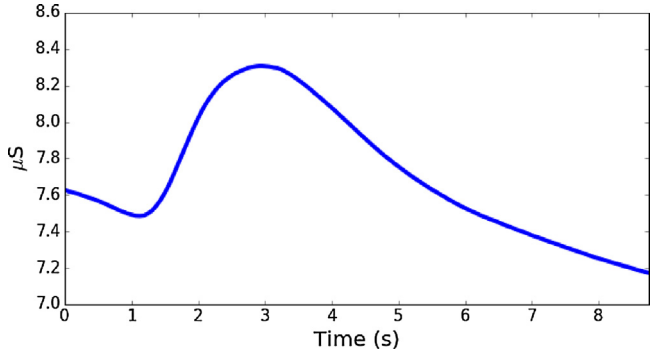

Fig. 4. Example SCR.

Regarding the NMSE values, Table 2 shows that all the three versions of our method have lower NMSE values, indicative of a smaller distortion introduced in the artifact-free signals, than the approaches Chen et al. [6] and the Greco et al. [14], being the *haar* solution the one obtaining the lowest NMSE value among them. This difference is statistically significant at the $p < .05$ level in the case of the comparison between the *haar* version and the Greco et al. [14] method ($t(8340) = 13.67$; $p < .001$), but not in the comparison between the *haar* version and the Chen et al. (2015) method ($t(8340) = 1.82$; $p = .07$). By contrast, the approach by Swangnetr and Kaber [31] is significantly better than our *haar* version in terms of NMSE ($t(8340) = 95.46$; $p < .001$).

Taken together, the present results indicate that the *haar* version of the proposed method is the one offering the highest advantages: it reduces artifact power similarly to the most powerful of the three state-of-the-art methods, while introducing much less distortion in the artifact-free segments. Compared to the approach by Chen et al. [6], our algorithm outperforms it in terms of APA while no significant difference is found in terms of NMSE. Finally, while the approach by Swangnetr and Kaber [31] is much more respectful for the artifact-free signals, it's potential to reduce artifact power is also much lower than our *haar* version. Thus, these results suggest that the *haar* version provides a good balance between a high artifact power reduction and a low distortion introduced in the artifact-free signals, which together with its lower computational cost, make it the optimal solution among the existing ones.

5. Online validation: MIST experiments

In order to test the performance of the algorithm in actual online tasks, we asked 12 volunteers to participate in an experiment aimed to elicit different levels of physiological arousal on them. The goal of this experiment was to analyze whether the online filtering of EDA signal done with our *haar* method improves the identification of SCRs, and, ultimately, allows a better detection of cognitive and emotional states. SCRs are one of the most well-known features of EDA activity, consisting of momentary increases in the EDA level occurring when the individual encounters a stimulus eliciting arousal in the sympathetic nervous system. SCRs commonly start between 1 and 3 seconds after stimulus presentation, and are characterized by a clearly identifiable waveform (see Fig. 4), which includes a sudden rise in EDA levels, with an amplitude typically between 0.1 and 1.0 μS, and a rise time between the onset of the response and the peak between 1 and 3 s [10].

The experiment involved three tasks aimed, respectively, to induce states of relax, moderate arousal, and stress, in the partici-

pants. The logic behind this experimental design is that, since SCRs are much more frequent when the individual is aroused [3], a higher number of SCR should be found for the more stressful tasks compared to the more relaxing tasks. We implemented the tasks using the Psychopy software.[2]

5.1. Tasks description

In the first task (*relax* task), the participant was instructed to watch a series of pictures of landscapes displayed in a PC monitor, while listening to a segment of calm classical music. Each picture was shown during two seconds, and the total duration of the task was two minutes.

In turn, the second and third tasks consisted of an adapted version of the Montreal Imaging Stress Task (MIST) [11]. It has two parts: in the first one (*training* task), the participant was asked to solve a series of arithmetic operations with one or two digits in the shortest possible time, but without any time limit. This task was expected to elicit moderate states of arousal and served to gather an approximation of the time that the participant needed for solving such operations. The second part of the test (*stress* task) presented similar operations but with a time limit, which was shown in a time bar on the screen. The software calculated the mean time required by the participant for each operation from the first part, and assigned a 10% less time for each operation in the second part, making it very hard to solve the operations correctly. The program provided feedback on whether the responses were correct or not, and also adjusted the time assigned to each operation, in a way such, after three correct responses, the time assigned was reduced another 10%, so making the participant unable to solve correctly the operations. In order to avoid excessive frustration in the participant, after three wrong responses the time assigned to the next operation was slightly increased, so the participant had more chances to give a correct answer. The *stress* task also included negative feedback by displaying on the screen a bar showing a (fake) comparison between the participant's performance and the mean population performance in the task, which indicated that the participant's performance was below the mean. After several operations with incorrect responses, a warning was presented on the screen indicating that the poor performance of the participant was compromising the results, and requested to invest more effort in the task. Since temporal pressure and negative feedback are two of the most common causes of stress, this task is expected to elicit higher stress compared to the training task. The training and stress tasks required the participants to manually type the responses on the keyboard, which involved moving the right hand to which the electrodes were attached, thus favoring the occurrence of motion artifacts in the EDA signal.

5.2. Procedure

Twelve employees (six males and six females) from first author's institution, aged between 21 and 54 ($M = 41.92$, $SD = 8.43$), and unaware of the experiment's purposes, voluntarily participated in the trial. Each one carried out the experimental tasks while his or her EDA signals were recorded and filtered online following the method described above using the *haar* mother wavelet. Fig. 6 shows the experimental setup.

During the experiment, participants were comfortably seated while wearing a noise canceling headphone in an experiment room. The experiment was presented on a computer screen while participants EDA was collected using a wireless wristband sensor Shimmer GSR+.[3] A laptop was placed behind the intervention table, hidden from the participant. It was used to present the MIST on the monitor, to receive the data from the Shimmer GSR+ sensor and to denoise and save the denoised EDA data.

Table 3
Mean (and standard deviation) of mean number of SCRs per segment across participants by task, as calculated from the raw and the filtered signal.

Task	Relax	Training	Stress
Raw signal	1.75 (1.36)	1.72 (1.19)	1.82 (0.97)
Filtered signal	0.22 (0.20)	0.34 (0.17)	0.39 (0.22)

The whole session lasted for around 15 minutes. To achieve an efficient time synchronization, the experiment execution and the EDA collection and filtering were done on the same system and were time stamped with the system time. For online filtering, the proportion of motion artifacts in original signal, $\delta = 0.10$ was employed for all participants. The data collection window was of 8 seconds duration after which the filtering was applied and the data was saved in a new file. A sliding window protocol of 1 second was applied. The original and the denoised data corresponding to each single window was saved in a separate file for the ease of further processing in validation of the denoising results. After finishing the tasks, participants were asked to rate in a 7-points Likert-type scale how stressful each task was, in order to have a measurement of the subjectively perceived stress. Finally, participants were debriefed about the goals of the experiment and the manipulation check in the stressful task.

For the validation of the denoised results, a SCR detector was implemented in Matlab. SCR detection was a two stage process. First, a valley is located on the signal. Then the highest peak after the valley is found. The SCR amplitude is the highest peak amplitude minus the lowest valley amplitude. SCR occurrences that were not large enough (<0.1 μS) were rejected.

5.3. Results

According to the self-reported estimates provided by the participants, the three tasks achieved their goals of eliciting different states of subjective stress. The mean rating for the relax task in the stress scale was 1.42 ($SD = 1.16$), while the mean rating for the training task was 3.00 ($SD = 1.76$), and the mean rating for the stress task was 5.83 ($SD = 1.34$), which, in a scale ranging from 1 to 7, can be considered as low, moderate, and high values, respectively.

Since the duration of the tasks were different, we consider the mean number of SCRs by 8 seconds segment in each task instead of the absolute number of SCRs by task. The first noticeable aspect of the comparison between the raw and the filtered signal is that the mean number of SCRs is much higher in the raw signals than in the filtered ones (see Table 3). The values for the raw signals indicate that, in some cases, more than one (even up to two or three) SCRs were detected by each 8-seconds segment, even for the relaxing task (Fig. 5). Since these values deviate largely from the typical number of SCRs [10], they are not realistic and probably a consequence of counting motion artifacts as SCRs. By contrast, the filtered signal shows much more realistic values (less than one SCR by segment in all cases). Particularly for the relax task, where few or none SCRs were expected, the raw signal shows a considerably high number of them, while they are kept to a minimum in most cases when using the filtered signal (see Fig. 5). Meanwhile, the training and stress task show a higher number of SCRs, which is in agreement of what can be expected from this test. Taken together, these results suggest that a number of artifacts were counted as SCRs when using

[2] http://www.psychopy.org/.

[3] http://www.shimmersensing.com/products/shimmer3-wireless-gsr-sensor.

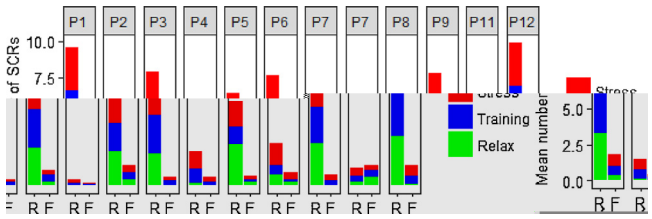

Fig. 5. Mean number of SCRs across participants by task for raw and filtered signal, Pi: participant *i*, R: raw signal, F: filtered signal.

Fig. 6. Experimental set-up.

the non-filtered signal, and that the filtering process helped to solve this problem.

In order to check whether the filtering improves the detection of SCRs, we analyzed the correlation between the number of SCRs in each task and the subjective stress scores given by the participants. Since the presence of SCRs shows a high individual variability, for analysis we used within-subject normalized scores [18] of mean number of SCRs by segment, for both the measurements with the raw and the filtered signal. In order to correct for possible different baselines in subjective perception of stress, the stress ratings were also normalized within-subjects. Then a linear regression was fitted to explore how the normalized mean number of SCRs by task predicted the normalized stress ratings, for both the raw and the filtered signal (Fig. 7). The results for the raw signal show no significant relationship between the number of SCRs and the stress ratings ($\beta = 0.23$, $p = 0.17$), while, in the case of the filtered signal, the number of SCRs is a significant predictor of subjective stress ($\beta = 0.51$, $p = 0.001$). As shown in Fig. 7, while there is not clear pattern of relationship between SCRs count and stress ratings when using the raw signal, in the data obtained with the filtered signal, even tough the presence of some outliers, the relationship is more linear. Moreover, while in the case of the raw signal the variance in the dependent variable explained by the model is quite low ($R^2 = 0.05$, $F(1, 34) = 1.96$, $p = 0.17$), in the regression considering the filtered signal it explains about a quarter of the variance in the dependent variable ($R^2 = 0.26$, $F(1, 34) = 12.12$, $p = .001$), that is, five times more variance in subjective ratings is explained by SCRs count when using the filtered signal. Noticeably, the data obtained with the raw signal show a relationship between SCRs count and which has a effect size in the range small to medium ($b = 0.23$) according to [8], and which is not statistically significant. It would lead a researcher to interpret that there is no statistical evidence of a relationship between both variables. By contrast, the results obtained with the filter signal shows a large to very large effect size ($\beta = 0.51$) [8], whose statistical significance is very high ($p = .001$), so there is strong evidence of a clear relationship between subjective stress ratings and number of SCRs.

These results suggest that the unfiltered signal contained a significant number of motion artifacts that were mistaken as SCRs by the detector. On the other hand, the online filtering allowed cleaning such artifacts, thereby improving the detection task and making the signal more descriptive of the actual subjective stress experienced by the participants.

6. Discussion and conclusion

The results demonstrate the effectiveness of the proposed method for online EDA filtering in real world scenarios. Specifically, *haar* variant of the proposed method performs an artifact power reduction similar to the currently available methods, with a higher respect for the original artifact-free signals. Furthermore, this is achieved with much lower computational cost than the state-of-the-art methods. In addition, the proposed method does not have any dependency on external libraries such as QP Solver required by [14] or Expectation-Maximization (EM) algorithm required by Chen et al. [6] approach. It is worthwhile noting that the cvxEDA approach by Greco et al. [14] requires tuning of at-least four parameters (τ_0 slow time constant, τ_1 fast time constant, α penalization for the sparse SMNA driver and γ penalization for the tonic spline coefficients) in pre-processing step. On the other hand, the proposed method requires knowledge of a single parameter (the proportion of motion artifacts in original signal δ) only in the pre-processing step.

We envision that future affective computing applications will be based on wearable devices, which are characterized by limited

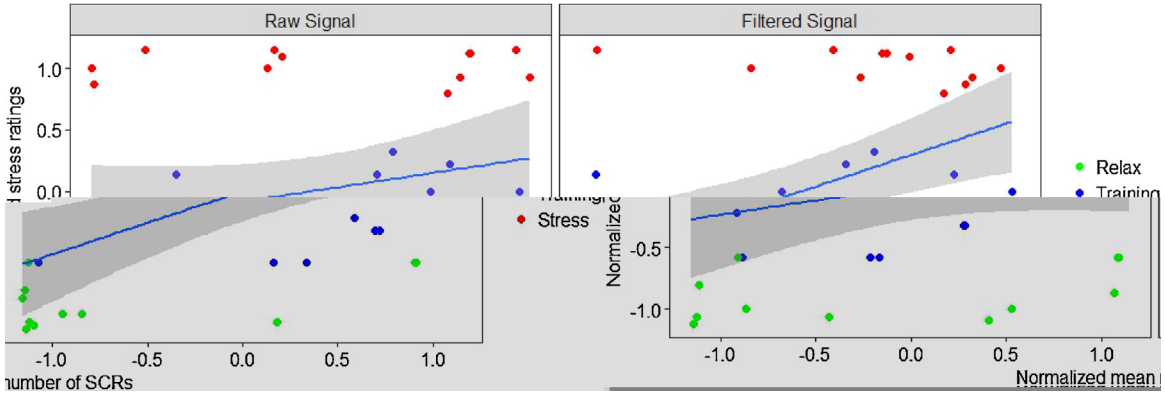

Fig. 7. Linear regression of SCRs by task predicted the normalized stress ratings.

computational power and storage. Hence, the lower computational requirements of the proposed method along with no external dependencies can facilitate the embedding of denoising algorithms into such devices. It can therefore boost the creation of applications able to monitor user's cognitive and emotional states. In this sense, stress prediction from EDA signal is still an open research issue, which falls far beyond our objectives in this work, and which probably will require the use of more significant features than the one used here. Even so, our results confirm that the online filtering significantly improves the quality of the signal in such a way that the correlation between the analyzed signal feature (SCRs) and subjectively perceived stress is substantially increased.

One limitation of the present study is that the performance of our filtering algorithm was tested in two contexts in which the users carried out actions (driving and typing in the keyboard) involving hand and arm movements while seated. However, wearable devices allow EDA recording in more dynamic situations, like walking or practicing sports, which can involve larger and faster movements in the subject's body and arms, and that may therefore produce more frequent or larger motion artifacts. Whereas testing the algorithm in all possible contexts of use of EDA monitoring is virtually impossible, further research should validate the effectiveness of the proposed method in some representative situations (e.g. customers walking in a grocery store, for neuromarketing studies). However, since many of the possible applications of EDA monitoring (including virtually in any ICT system addressed to recognize emotions and engagement level from physiological signals, such as tactile gaming consoles, robots during human-robot interaction, neuro-marketing, etc.) involve users in more static situations like the ones analyzed in the present study, the results of the present research mean a valuable contribution even without further validation in more dynamic situations.

Finally, there are reasons to think that the presence of motion in EDA signal cannot only been seen as a source of noise. Since it has been found that humans move or gesture in response to affective changes [5], the temporal concurrence of EDA changes and movements could mean a meaningful source of information on user's cognitive and emotional processes. For instance, a research question could be whether the levels of movement associated to EDA changes are informative of other processes related to physiological arousal (e.g. strategies for emotion regulation or coping with high cognitive load or stress). Since wavelet based motion artifact removal for EDA may also have advantages in this sense, the present research paves the way for future research addressing this issue.

Acknowledgements

This research work is supported by the Industrial Doctorate program (Ref. ID.: 2014-DI-022) of AGAUR, Govt. of Catalonia, is partially funded by the La Caixa via project AutonoMe (Ref. ID.: AD16-00877) and is partially supported by the Agència per a la Competivitat de lÊmpresa, ACCIÓ. Author J.S. gratefully acknowledges the cooperation of Dr. Manida Swangnetr from NCSU and Dr. Alberto Greco from University of Pisa in discussion of their work on related research.

References

[1] M. Barreda-Ángeles, I. Arapakis, X. Bai, B.B. Cambazoglu, A. Pereda-Baños, Unconscious physiological effects of search latency on users and their click behaviour, in: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, New York, NY, USA, 2015, pp. 203–212.
[2] M. Barreda-Ángeles, R. Pépion, E. Bosc, P. Le Callet, A. Pereda-Baños, Exploring the effects of 3D visual discomfort on viewers' emotions, IEEE International Conference on Image Processing (ICIP) (2014) 753–757.
[3] W. Boucsein, Electrodermal Activity, Springer US, Boston, MA, 2012.
[4] S. Boyd, L. Vandenberghe, Convex Optimization Problems, Cambridge University Press, 2004, pp. 127–214.
[5] R. Chellali, S. Hennig, Is it time to rethink motion artifacts? Temporal relationships between electrodermal activity and body movements in real-life conditions, 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII) (2013) 330–335.
[6] W. Chen, N. Jaques, S. Taylor, A. Sano, S. Fedor, R. Picard, Wavelet-based motion artifact removal for electrodermal activity, Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2015 (2015) 6223–6226.
[7] Z. Chen, S. Haykin, J.J. Eggermont, S. Becker, Appendix E: Expectation-Maximization Algorithm, John Wiley & Sons, Inc., 2007, pp. 384–386.
[8] J. Cohen, in: J. Cohen (Ed.), Statistical Power Analysis for the Behavioral Sciences, Academic Press, 1988, Revised edition.
[9] R.R. Coifman, D.L. Donoho, Translation-Invariant De-Noising, Springer New York, New York, NY, 1995, pp. 125–150.
[10] M.E. Dawson, A.M. Schell, D.L. Filion, The electrodermal system, in: J.T. Cacioppo, L.G. Tassinary, G.G. Berntson (Eds.), Handbook of Psychophysiology, 3rd ed., Cambridge University Press, New York, NY, US, 2007, pp. 159–181.
[11] K. Dedovic, R. Renwick, N.K. Mahani, V. Engert, S.J. Lupien, J.C. Pruessner, The Montreal imaging stress task: using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain, J. Psychiatry Neurosci. 30 (2005) 319.
[12] A. Drachen, L.E. Nacke, G. Yannakakis, A.L. Pedersen, Correlation between heart rate, electrodermal activity and player experience in first-person shooter games, in: Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games, ACM, New York, NY, USA, 2010, pp. 49–54.
[13] S. For Psychophysiological Research Ad Hoc Committee on Electrodermal Measures, Publication recommendations for electrodermal measurements, Psychophysiology 49 (2012) 1017–1034, http://dx.doi.org/10.1111/j.1469-8986.2012.01384.x.
[14] A. Greco, G. Valenza, A. Lanata, E.P. Scilingo, L. Citi, cvxEDA: a convex optimization approach to electrodermal activity processing, IEEE Trans. Biomed. Eng. 63 (2016) 797–804.
[15] A. Gupta, S.D. Joshi, S. Prasad, A new approach for estimation of statistically matched wavelet, IEEE Trans. Signal Process. 53 (2005) 1778–1793.
[16] J.A. Healey, R.W. Picard, Detecting stress during real-world driving tasks using physiological sensors, IEEE Trans. Intell. Transp. Syst. 6 (2005) 156–166.
[17] J. Hernández, R.R. Morris, R.W. Picard, Call Center Stress Recognition With Person-Specific Models, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 125–134.
[18] J.R. Jennings, P.J. Gianaros, Methodology, in: J.T. Cacioppo, L.G. Tassinary, G. Berntson (Eds.), Handbook of Psychophysiology, 3rd ed., Cambridge University Press, 2007, pp. 812–833, Cambridge Books Online.
[19] L. Junli, H. Yanqin, W. Ping, C. Gang, A novel method for the determination of the wavelet denoising threshold, 2007 1st International Conference on Bioinformatics and Biomedical Engineering (2007) 713–716.
[20] V. Krishnaveni, S. Jayaraman, L. Anitha, K. Ramadoss, Removal of ocular artifacts from EEG using adaptive thresholding of wavelet coefficients, J. Neural Eng. 3 (2006) 338–346.
[21] J. Laparra-Hernández, J. Belda-Lois, E. Medina, N. Campos, R. Poveda, EMG and GSR signals for evaluating user's perception of different types of ceramic flooring, Int. J. Ind. Ergon. 39 (2009) 326–332.
[22] C.S. Lima, A. Tavares, J.H. Correia, M.J. Cardoso, D. Barbosa, Non-stationary biosignal modelling, in: D. Campolo (Ed.), New Developments in Biomedical Engineering, InTech, 2010.
[23] C. Liu, K. Conn, N. Sarkar, W. Stone, Online affect detection and robot behavior adaptation for intervention of children with autism, IEEE Trans. Robot. 24 (2008) 883–896.
[24] M.-Z. Poh, T. Loddenkemper, N.C. Swenson, S. Goyal, J.R. Madsen, R.W. Picard, Continuous monitoring of electrodermal activity during epileptic seizures using a wearable sensor, IEEE (2010) 4415–4418.
[25] B. Molavi, G.A. Dumont, Wavelet-based motion artifact removal for functional near-infrared spectroscopy, Physiol. Meas. 33 (2012) 259–270.
[26] L. Nacke, A. Drachen, S. Göbel, Methods for evaluating gameplay experience in a serious gaming context, Int. J. Comput. Sci. Sport 9 (2010) 1–12.
[27] G.P. Nason, B.W. Silverman, The Stationary Wavelet Transform and Some Statistical Applications, Springer-Verlag, 1995, pp. 281–300.
[28] M.Z. Poh, N.C. Swenson, R.W. Picard, A wearable sensor for unobtrusive, long-term assessment of electrodermal activity, IEEE Trans. Biomed. Eng. 57 (2010) 1243–1252.
[29] A. Sahli, P. Trecourt, S. Robin, One sided acceptance sampling by variables: the case of the Laplace distribution, Commun. Stat. Theory Methods 26 (1997) 2817–2834.
[30] J. Shukla, M. Barreda-Ángeles, J. Oliver, D. Puig, MuDERI: Multimodal Database for Emotion Recognition Among Intellectually Disabled Individuals, Springer International Publishing, Cham, 2016, pp. 264–273.
[31] M. Swangnetr, D.B. Kaber, Emotional state classification in patient–robot interaction using wavelet analysis and statistics-based feature selection, IEEE Trans. Hum. Mach. Syst. 43 (2013) 63–75.
[32] S.A. Taylor, N.M. Jaques, W. Chen, S. Fedor, A. Sano, R.W. Picard, Automatic identification of artifacts in electrodermal activity data, Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2015) 1934–1937.