A Data Driven Empirical Iterative Algorithm for GSR Signal Pre-Processing

Arvind Gautam[1,2], Neide Simões-Capela[3,1], Giuseppina Schiavone[4], Amit Acharyya[2], Walter de Raedt[1], Chris Van Hoof[1,3]

[1]IMEC, Kapeldreef 75, Heverlee, Belgium

[2]Electrical Engineering Department, Indian Institute of Technology Hyderabad, Sangareddy, India

[3]Electrical Engineering Department - ESAT, KU Leuven, Kasteelpark Arenberg 10, Heverlee, Belgium

[4]IMEC-NL, High Tech Campus 31, Eindhoven, The Netherlands

ee15resch01003@iith.ac.in, neide.simoescapela@kuleuven.be, giuseppina.schiavone@imec-nl.nl, amit_acharyya@iith.ac.in, walter.deraedt@imec.be, chris.vanhoof@imec.be

***Abstract*— In this paper, we introduce a data driven iterative low pass filtering technique, the Empirical Iterative Algorithm (EIA) for Galvanic Skin Response (GSR) signal pre-processing. This algorithm is inspired on Empirical Mode Decomposition (EMD), with performance enhancements provided by applying Midpoint-based Empirical Decomposition (MED), and removing the sifting process in order to make it computational inexpensive while maintaining effectiveness towards removal of high frequency artefacts. Based on GSR signals recorded at the wrist we present an algorithm benchmark, with results from EIA being compared with a smoothing technique based on moving average filter - commonly used to pre-process GSR signals. The comparison is established on data from 20 subjects, collected while performing 33 different randomized activities with right hand, left hand and both hands, respectively. In average, the proposed algorithm enhances the signal quality by 51%, while the traditional moving average filter reaches 16% enhancement. Also, it performs 136 times faster than the EMD in terms of average computational time. As a show case, using the GSR signal from one subject, we inspect the impact of applying our algorithm on GSR features with psychophysiological relevance. Comparison with no pre-processing and moving average filtering shows the ability of our algorithm to retain relevant low frequency information.**

Keywords*— EMD, Data driven, Iterative algorithm, GSR, EDA, Skin conductance.

I. INTRODUCTION

Galvanic Skin Response (GSR), also referred to as Skin Conductance (SC) or Electro-Dermal Activity (EDA), measures the skin conductance level and its change over time. Sweat secretion causes changes in skin conductance, thus, correlating with GSR value. Emotional stimulation and arousal triggers the release sweat, producing measurable variation in GSR [1][2]. The hand, wrist and foot are common anatomical locations to measure GSR, due to the high concentration of sweat glands. GSR signals can be characterized by two basic components: tonic component (0-0.16 Hz) and phasic component (0.16-2.1 Hz)[3][4]. The tonic component relates to slow changing baseline levels and individual background characteristics. Whereas, the phasic component includes the fast changing element which can be event related[1][2].

Traditionally, the measurement of GSR is used in psychophysiological research in laboratory environment. Measurements are conducted under stationary conditions and using experimental protocols that isolate other nefarious effects from the physiological aspect being investigated[5]. Recently, due to advances in hardware design, miniaturized and low power sensing technologies became broadly available, leading to crescent access to wearable sensors in the consumer market [6-9]. Wearable technology allows to shift from measurements of GSR in laboratory, to recordings in free-living conditions, with large amounts of data being collected without controlled stimulus or strict measuring standards. This shift has numerous advantages from the application point of view, as continuous and unobtrusive monitoring of physiological data is desired both in medical and wellness applications. At the same time, it comes at a cost: uncontrolled artefacts, induced by improper contact between the wearable and skin during free motions, reduce the signal quality of the GSR, making acquisitions in ambulatory conditions less reliable.

There is so far, no standardized way to deal with artefacts in GSR signals. While some will directly use raw GSR as provided by the sensors to extract physiologically relevant information [3], others will apply a moving average filter (smoothing filter) as pre-processing step [4][12]. In free-living, we found these methods to be insufficient to deal with the aforementioned artefacts[5].

Hereby we propose a data driven iterative low pass filter, the Empirical Iterative algorithm (EIA) inspired on Empirical Mode Decomposition (EMD), to pre-process GSR signals and remove motion artifacts. The algorithm can be applied both on signals affected by motion artefact, and on signals collected during stationary conditions, without interfering with the relevant physiological information. We compared our algorithm with standard moving average filtering, and determined its performance in removing both motion artefacts and quantization noise. We tested it on GSR signals collected from 20 subjects while performing 33 randomized activities using one or both hands. We compare execution times for EMD and EIA for these 20 subjects. Also, to show the validity of our approach for applications in psychophysiology, 4 parameters with physiological relevance commonly derived from the GSR signal were calculated, on

ISBN 978-90-827970-1-5 © EURASIP 2018

1 subject, and were compared when calculated from raw signals, moving average and EIA processed signals.

To the best of our knowledge, this is the first work of its kind, where a new data driven algorithm is proposed for GSR signal pre-processing.

The remainder of the paper is structured as follows: Section II provides the necessary background for this paper, Section III presents the proposed algorithm along with the performance comparison with the moving average, the computation time evaluation and the validation results, Section IV includes results and discussion and Section V is the conclusion.

II. BACKGROUND AND METHOD

A. *Empirical Mode Decomposition (EMD) Basics*

Empirical Mode Decomposition (EMD) was proposed by Huang et al. [10]. EMD is a data driven nonlinear technique for decomposing non-linear and non-stationary signals into a set of amplitude and frequency modulated components known as intrinsic mode functions (IMFs). The IMFs are meant to be zero-mean, mono component, oscillatory functions, which are orthogonal to each other. The IMFs are obtained from a time series by an iterative process known as sifting process [10][11]. In this process, upper and lower signal envelopes, respectively defined by the cubic spline interpolation of the local maxima and the local minima of the signal, are used to estimate a mean envelope. Iterations are repeated till the criteria that define an IMF are not satisfied, anymore. These criteria are [10]:

- Mean value of the input signal envelopes should be zero in order for the mean signal to constitute an IMF.
- IMF must have equal number of extrema and zero crossings, or the numbers can differ by one.

After extraction of IMFs from a time series signal the residue tends to become a monotonic function, such that no more IMFs can be extracted. Finally, after the iterative process, the input signal is decomposed into a sum of IMF functions (C1, C2, C3, C4, ...,Cn(t)) and a residue, r(t).

B. *Midpoint-based Empirical Decomposition (MED)*

A known problem of the EMD is the overshooting or undershooting of the signal related to the cubic spline interpolation. This is partially attenuated by doing the interpolation on the midpoints between consecutive minima and maxima, and not on the signal envelopes. An approach suggested by He et al. [11].

C. *Method and Data Acquisition System*

To investigate the effect of motion artefacts on GSR measured on the wrist, with a wearable device we collected data from 20 subjects while performing 11 different activities (detailed in Table I) with right hand, left hand (not analyzed here), and both hands, respectively. Altogether 33 randomized activities. A stationary phase was included at the beginning of the experiment to be taken as the individual's physiological baseline (i.e. reference at rest). All the activities are performed for 30 seconds, followed by 30 seconds of recovery. These activities include finger, hand and arm movements (from 2 to 7) and real-world-like activities (from 8 to 12). Using this protocol, around 38 minutes of data were collected for each subject. As for the protocol, data was collected synchronously from both wrists during the experiment, though for the purpose of this exploration only data collected from the right wrist will be used.

The data acquisition system used in this protocol was the Chillband (IMEC vzw, Belgium), a wrist worn sensing device show Fig. 1. It collects GSR data at a sampling frequency of 256 Hz, within a dynamic range of 0.05-20μS. Synchronously, it measures 3D acceleration at a sampling frequency of 32 Hz, in the range of ±2 g.

Figure 1. Chill band wearable sensor from IMEC, Belgium

Table. I: Activities

Serial Number	Activity performed	Time (s)
1	Rest	300
2	Individual finger tapping	60
3	Radial/ulnar deviation	60
4	Pronation/supination of hand	60
5	Wrist extension/flexion	60
6	Fist opening/closing	60
7	Whole arm lateral swing	60
8	Pen fidgeting	60
9	Random mouse movements and clicking	60
10	Random keyboard typing	60
11	Writing/drawing on paper	60
12	Cleaning table	60

III. PROPOSED ALGORITHM

A. *Problem statement*

An example of the raw GSR signal obtained from one subject with the protocol described in the previous section is shown in Fig. 2. Out of the 33 movements (30 seconds) and respective pauses (30 seconds), we highlight the signal corresponding to movements involving the right hand. It is visible that due to movement, the GSR is affected by motion artefacts of different amplitudes, on top of the overall quantization noise.

ISBN 978-90-827970-1-5 © EURASIP 2018

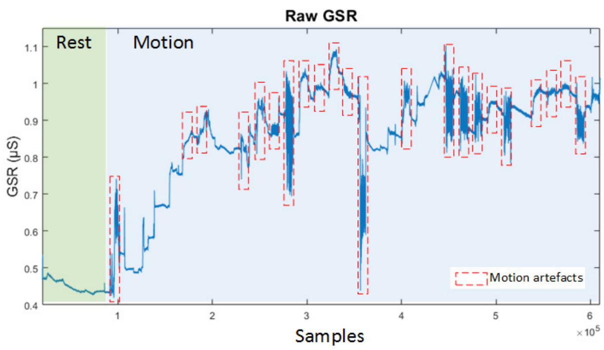

Figure 2. Raw GSR: rest and motion phases. Signals corresponding to the movements involving the right hand are delimited by red lines.

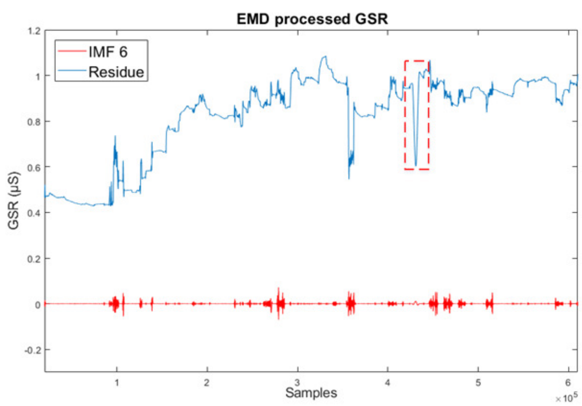

Figure 3. GSR decomposition based on EMD, IMF6 and its respective residue. Detail of overshooting shown inside the red rectangle.

Applying EMD on the raw signal on Fig. 2, results on 6 IMFs, and respective residues. The 6th IMF along with its residue is shown in Fig. 3. The filtered GSR signal can be obtained from the residue, not the IMF itself.

The EMD mostly removes the high frequency effect of motion from residue obtained in Fig. 3, while preserving the global trend of the signal. Though, for breaking down the raw GSR into these 6 IMFs (and corresponding residues) the sifting process takes around 50 to 60 iterations. Which would produce high system latency on real time applications. Therefore, there is a need for an algorithm that can be easily implement on real-time applications for GSR signal pre-processing, and also can intelligently remove the high frequency noise without affecting the signal.

B. Empirical Iterative Algorithm

Inspired by the EMD algorithm, next we propose a new iterative algorithm for denoising the GSR from motion artefacts and quantization noise. The algorithm does not rely on the sifting process of EMD, and provides the filtered signal directly as an output of each iteration. This algorithm also relies on MED to minimize the effects of overshooting or undershooting.

The pseudocode for the proposed method is presented in Fig. 4. The steps 1 and 2(a) are similar to EMD. From step 2(b) the EIA differs, in the sense that the resulting signal is obtained without subtracting the midpoint envelope from the input signal. Conveniently performing the iterations on the filtered signal, while obtaining the noisy content from the final residue.

1. Consider $x(t)$ as an input time series GSR signal.
2. For $i = 1:7$ where i is number of iterations.
    a. Identify all local maxima's as $p^i(n)$ and local minima's as $q^i(n)$ of input signal.
    b. Find mid-point $m^i(n) = (x_m, y_m)$ for consecutive maxima $p^i(n) = (x_p, y_p)$ and minima $q^i(n) = (x_q, y_q)$ and vice versa for all the maxima's and minima's as :

    $$(x_m, y_m) = \left(\frac{x_p+x_q}{2}, \frac{y_p+y_q}{2}\right)$$

    c. Interpolate all the midpoints $m^i(n)$ by cubic spline interpolation to get a signal $y^i(t)$.
    d. Consider $y^i(t)$ as input signal: $x(t) = y^i(t)$
3. End.
4. Subtract $y^7(t)$ from $x(t)$ to get $r(t)$.

    $$r(t) = x(t) - y^7(t)$$

5. Then $r(t)$ will represent residue of the GSR.

Figure 4. Pseudocode for the Empirical Iterative Algorithm (EIA)

The output of EIA on the raw signal from Fig. 2 is shown in Fig. 5.

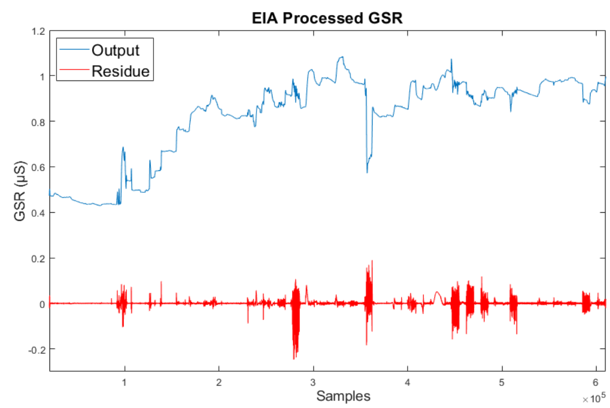

Figure 5. Output from Empirical Iterative Algorithm (EIA)

In this use case for EIA, the number of iterations is fixed on 7, as it was verified based on visual inspection that this is a reasonable number of iterations, after testing it on the full set of data collected from the 20 subjects. The number of iterations can be decreased or increased based on the application, and the characteristics of the raw data collected.

C. Benchmarking

We benchmark our algorithm with commonly use moving average technique. The moving average is applied with a window length of 64 samples. The comparison of raw GSR, GSR processed with moving average, and GSR processed with proposed EIA are shown in Fig. 6.

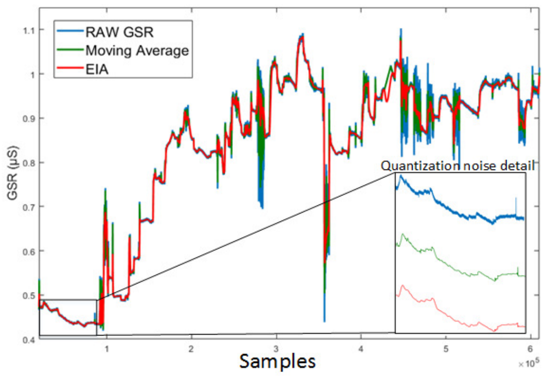

Figure 6. Comparison of raw GSR, smoothened GSR obtained with moving average filter and GSR processed with EIA. Emphasis on the detail of quantization noise removal.

The signal-to-noise ratio (SNR) for the raw GSR, the GSR processed with moving average and the GSR processed with EIA is calculated for all 20 subjects and the results are presented in Fig. 7. For the calculation of SNR, the signal content is extracted by low pass filtering the input signal using a Butterworth filter (3rd order) with a cut-off frequency of 1.99 Hz (to include phasic and tonic components). The noise content is extracted by band pass filtering in the band of 2-10 Hz (to include artefacts and quantization noise) using a Butterworth filter (3rd order). The average SNR value for all 20 subjects is 40, 47 and 61 dB in the raw GSR, moving average and EIA, respectively.

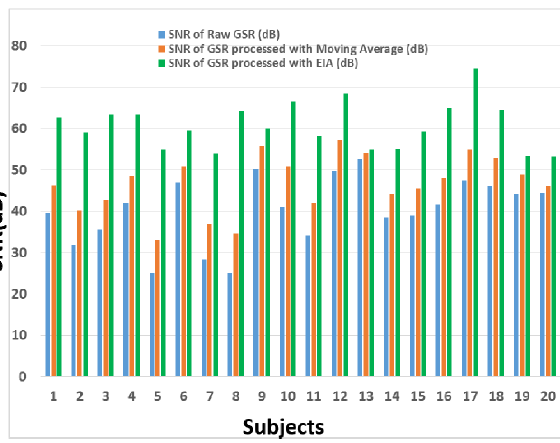

Figure 7. SNR comparison for 20 subjects across raw GSR, Moving Average and EIA.

To further provide evidence on the effectiveness of the proposed EIA in adaptively removing motion artefacts, we included acceleration magnitude captured in the wrist as our reference for motion. In Fig. 8, we present the spectrogram of accelerometer magnitude and that of GSR, across 3 conditions: raw, processed with moving average and processed with EIA. The artefacts can be connected to the motion detected by the accelerometer and the filtering properties of the two methods can be compared.

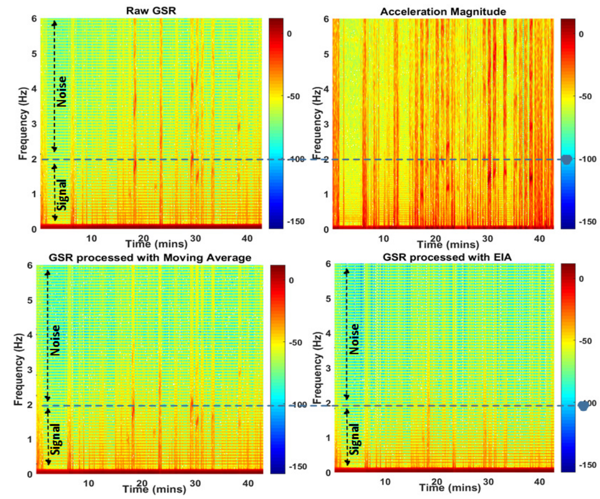

Figure 8. Spectrogram representation: (a) Raw GSR; (b) Magnitude of the acceleration; (c) GSR processed with moving average; (d) GSR processed with EIA.

D. Computation time

In order to provide data to support our claim that this algorithm represents a clear enhancement in terms of computation time, we provide the comparison between the traditional EMD and EIA in Table. II. EMD decomposition was limited to 7 IMFS and EIA to 7 iterations. For both algorithms simulations ran on MATLAB 2016a, on a PC with an Intel core i5 processor @ 3.50 GHz and 8 Gb RAM. Times were measured for each of the 20 subjects in the dataset previously described, and averaged.

Table. II: Computation time for EMD and EIA: mean ± standard deviation values for data of 20 subjects.

EIA (s)	EMD (s)
2.4 ± 0.2	326 ± 45

E. Proof-of-concept

As a proof of concept for the proposed EIA, we considered 4 features with psychophysiological relevance that are usually derived from GSR: tonic component, phasic component, absolute second difference and number of peaks [3][4]. These parameters were calculated on a sliding window of length 30 seconds, with 29 seconds overlap. The results are shown in Fig. 8 for the raw GSR, GSR processed with moving average and GSR processed with EIA.

In this sense, we use the phasic and tonic component as representing the signal physiological content whereas, the number of peaks and the absolute second difference, related to their inherent sensitivity to high frequency content, are being used to compare artefact effects.

IV. Results and Discussion

The results obtained by the SNR comparison on Fig. 7 for all the 20 subjects across raw GSR, moving average processed GSR and EIA processed GSR, show that the EIA can effectively reduce the impact of high frequency motion artefacts and quantization noise. The EIA based processing shows approximately 30 % increase in SNR when compared to moving average based processing.

From the spectrogram analysis shown in Fig. 8(a) the high frequency band (above 2 Hz) seems to contain principally motion artefacts that correlate to the motion reference, i.e. the corresponding acceleration magnitude, in Fig. 8(b). On the spectrogram of GSR processed with moving average, shown in Fig. 8(c), it is clear that there are still high frequency components related to motion. Whereas, on spectrogram of GSR processed with EIA, in Fig. 8(d), those components are softened and the result contains mostly low frequency signal information (bellow 2 Hz). This goes along with our claim that the EIA can adaptively soften the high frequency motion artefacts as well as the quantization noise (as visually depicted in Fig. 5 and 6).

Our EIA takes 136 times less computational time to provide an acceptable output, when compared to EMD, this represents a 99% enhancement. The high variance in the timing results for EMD is maybe related to high sensitivity of the sifting process to the signal content.

We start by showing in Fig. 2 that the GSR is indeed affected by motion artefacts, though these artefacts do not affect all the components used in psychophysiological research in the same way. From the comparison plot established in Fig. 9(a) and 9(b), we visualize that the key features of GSR such as tonic and phasic components are similar and maintain the same trend across the raw GSR, processed GSR from moving average and processed GSR from the EIA. Which provides a good indication that the EIA retains the relevant information of GSR, and can be used in general, whether there is movement or not. However, the results from number of peaks and absolute second difference show differences. Based on the results in Fig. 9 (c) and (d) the EIA decreases the values of both parameters, as the number of high frequency oscillations drops. We believe that left untreated artefacts will affect these values and provide a wrong physiological quantification.

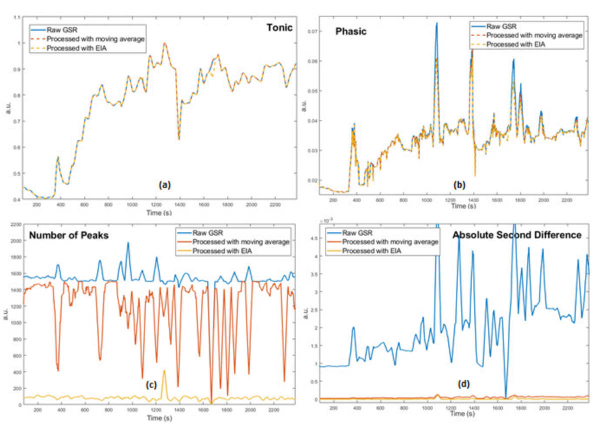

Figure 9. GSR parametrization results based on raw signal, moving average and EIA: (a) Tonic component; (b) Phasic component; (c) Number of peaks; (d) Absolute second difference.

V. Conclusion

In this paper, we present an empirical iterative low pass filter algorithm, which is able to reduce motion artefacts and quantization noise on GSR signals, that may in the future act as a pre-processing step on the estimation of parameters with psychophysiological relevance. The algorithm achieves an average of 51 % increase in signal quality, while moving average archives only 16 %. Also, the EIA performs 136 times faster than the EMD in terms of the computational time. The results from our exploration in terms of performance can be considered favorable for future integration into a sensor platform for real-time pre-processing. Future work should focus on expanding the validation of the algorithm, to include psychophysiological stressors in the protocol, in order to further evaluate the impact of the algorithm on the physiological content of the GSR signal. Developing robust stopping criteria for the number of iterations in EIA should also be considered.

References

[1] Boucsein, Wolfram. Electrodermal activity. Springer Science & Business Media, 2012.

[2] Venables, Peter H., and Margaret J. Christie. "Electrodermal activity." Techniques in psychophysiology 54.3 (1980).

[3] Smets, Elena, et al. "Comparison of machine learning techniques for psychophysiological stress detection." International Symposium on Pervasive Computing Paradigms for Mental Health. Springer, Cham, 2015.

[4] Singh, Rajiv Ranjan, Sailesh Conjeti, and Rahul Banerjee. "A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals." Biomedical Signal Processing and Control 8.6 (2013): 740-754.

[5] Tronstad, Christian, et al. "Model-based filtering for artifact and noise suppression with state estimation for electrodermal activity measurements in real time." Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE. IEEE, 2015.

[6] Poh, Ming-Zher, Nicholas C. Swenson, and Rosalind W. Picard. "A wearable sensor for unobtrusive, long-term assessment of electrodermal activity." IEEE transactions on Biomedical engineering 57.5 (2010): 1243-1252.

[7] Setz, Cornelia, et al. "Discriminating stress from cognitive load using a wearable EDA device." IEEE Transactions on information technology in biomedicine 14.2 (2010): 410-417.

[8] Fletcher, Richard Ribón, et al. "Wearable sensor platform and mobile application for use in cognitive behavioral therapy for drug addiction and PTSD." Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE. IEEE, 2011.

[9] Picard, Rosalind W., Szymon Fedor, and Yadid Ayzenberg. "Multiple arousal theory and daily-life electrodermal activity asymmetry." Emotion Review 8.1 (2016): 62-75.

[10] Huang, Norden E., et al. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis." Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences. Vol. 454. No. 1971. The Royal Society, 1998.

[11] He, Qingbo, Robert X. Gao, and Patty Freedson. "Midpoint-based empirical decomposition for nonlinear trend estimation." Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual Internal Conference of the IEEE. IEEE, 2009.

[12] Bakker, Jorn, Mykola Pechenizkiy, and Natalia Sidorova. "What's your current stress level? Detection of stress patterns from GSR sensor data." Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on. IEEE, 2011

ISBN 978-90-827970-1-5 © EURASIP 2018