Physiological Signal Processing for Emotional Feature Extraction

Peng Wu[1], Dongmei Jiang[2] and Hichem Sahli[1,3]
VUB-NPU Joint AVSP Research Lab
[1]*Department of Electronics and Informatics(ETRO), Vrije Universiteit Brussel (VUB), Pleinlaan 2, Brussels, Belgium*
[2]*Shaanxi Provincial Key Lab on Speech and Image Information Processing, Northwestern Polytechnical University (NPU), Xi'an, China*
[3]*Interuniversity Microelectronics Centre (IMEC), Leuven, Belgium*

Keywords: Signal Denoising, QRS Detection, Electromyography Segmentation.

Abstract: This paper introduces new approaches of physiological signal processing prior to feature extraction from electrocardiogram (ECG) and electromyography (EMG). Firstly a new signal denoising approach based on the Empirical mode decomposition (EMD) is presented. The EMD can decompose the noisy signal into a number of Intrinsic Mode Functions (IMFs). The proposed algorithm estimates the noise level of each IMF. Experiments show that the proposed EMD-based method provides better denoising results compared to state-of-art. In addition, a real-time QRS detection approach is proposed to be directly applied on the noisy ECG signals. Moreover, an adaptive thresholding approach is employed for the EMG segmentation. Both approaches are validated using synthetic and real physiological data resulting in good performances.

1 INTRODUCTION

Affective Computing is a new area of computing research described as "computing which relates to, arises from, or deliberately influences emotions" (Picard, 2000). It emphasizes the importance of adding new emotional features to the human-computer interaction. The use of physiological sensors as a means of recognizing user's affective state has a number of advantages: (i) the size of such sensors is rapidly decreasing to the extent that it is nowadays incorporated into body wireless network (Pantelopoulos and Bourbakis, 2008), (ii) they are less disturbing than being "watched" by a camera as is the case with facial expression recognition, and (iii) they are less susceptible to social masking (Kim, 2007).

However, emotional state recognition by means of biosignals analysis is also problematic. This is due in part to the fact that biosignals are usually corrupted by white noise (Üstündağ et al., 2012) and other various types of noise, such as baseline wander, muscle contraction and electrode motion artifacts (Andrade et al., 2006; Blanco-Velasco et al., 2008). Furthermore, despite the evidence from psychophysiology suggesting a strong correlation between human emotional states and physiological responses, determining an appropriate mapping between them requires the estimation of reliable features.

In this paper, we propose advanced signal processing for the characterization of ECG and EMG physiological signals. Firstly, a new signal denoising approach is introduced. The focus is on signals with white Gaussian noise.

Wavelet-based denoising has been the dominant technique in the area of non-linear and non-stationary signals (e.g. physiological signals) denoising for many years (Donoho, 1995). The goal is to estimate the signal from the noisy observations such that the Mean Square Error (MSE) is minimum. To achieve this, the observed signal is transformed into wavelet domain, which decomposes it into many subbands. The small coefficients in the subbands are dominated by noise, while coefficients with large absolute value carry more signal information than noise. Replacing noisy coefficients (small coefficients below certain value) by zero and an inverse wavelet transform may lead to reconstruction signal that has lesser noise. Often, hard thresholding and soft thresholding techniques are used for such denoising process (Donoho, 1995).

The main drawback of the wavelet approach is that the basis functions are fixed, and do not necessarily match signals with large variabilities, such as the electromyography (EMG). Huang et al. pro-

Wu P., Jiang D. and Sahli H..
Physiological Signal Processing for Emotional Feature Extraction.
DOI: 10.5220/0004727500400047
In *Proceedings of the International Conference on Physiological Computing Systems* (PhyCS-2014), pages 40-47
ISBN: 978-989-758-006-2

posed the Empirical mode composition (EMD) as a tool to adaptively decompose a signal into a number of components, called Intrinsic Mode Functions (IMFs) (Huang et al., 1998). In contrast to the wavelet approach, the EMD relies on a fully data-driven mechanism that doesn't require any predefined basis. A general scheme of signal denoising using EMD is as follows: 1. Decompose the noisy signal into a number of IMFs; 2. Estimate the noise level of each IMF and threshold the estimated IMFs; 3. Reconstruct the denoised signal using the thresholded IMFs.

EMD-based denoising has been applied to physiological signals in many studies (Boudraa et al., 2005; Andrade et al., 2006; Jing-tian et al., 2007; Blanco-Velasco et al., 2008; Karagiannis and Constantinou, 2009; Agrafioti et al., 2012). However, the estimation of the noise level, of each IMF, for a reliable threshold is still an open question. The strategy of Donoho (Donoho and Johnstone, 1994) is widely used (Boudraa et al., 2005; Jing-tian et al., 2007; Karagiannis and Constantinou, 2009), nevertheless, it is only suitable for noise dominant scales (i.e. IMFs), thus applying this threshold to high scales causes signal distortion. This is due to the fact that each scale of the EMD occupies lower frequencies than its preceding ones and most noise components lie in the first several scales. The high scales are signal dominated. Using a single threshold may give misleading results. State-of-art approaches either reconstruct denoised signals using all thresholded IMFs while ignoring the signal distortion (Boudraa et al., 2005), or by thresholding the first IMFs and keep the other IMFs unchanged (Jing-tian et al., 2007; Blanco-Velasco et al., 2008). For the latter, the number of IMFs to be thresholded is decided empirically (Jing-tian et al., 2007) or via statistical test (Blanco-Velasco et al., 2008). Alternatively, instead of estimating the noise level from the whole IMF, Andrade (Andrade et al., 2006) manually selected a window of noise from the original signal and then the boundaries of this window were used for the noise level estimation.

In this paper, we follow the idea of EMD-based denoising, however different from state-of-art methods, rather than estimating the noise level of all IMFs using Donoho's strategy, we propose a novel approach for noise levels estimation of dominant IMFs.

In this study, we focus on the electrocardiogram (ECG) and electromyography (EMG) signals which have been confirmed to be useful for emotion recognition (Kreibig et al., 2007). For the ECG, heart rate (*HR*) and heart rate variability (*HRV*) are the cardiovascular response features most often reported as indicators of emotion (Kreibig, 2010). The first step in extracting them starts from the exact detection of R peaks in the QRS complex (see Fig.1). Within the last decade many approaches to R peaks detection have been proposed (Kohler et al., 2002). However, most of these approaches are off-line and targeting the noiseless signal, which don't meet the requirements of many real-time applications. To overcome this problem, we make use of a Change Point Detection (CPD) algorithm proposed by Guralnik and Srivastava (Guralnik and Srivastava, 1999) for event detection in time series.

The EMG also plays an important role in expression recognition (Van Boxtel, 2010). A basic problem in EMG applications is the determination of muscular active periods within a given EMG signal. Within each muscle activity period the root mean square (*RMS*) and absolute mean value (*AMV*), are used as features (Farfán et al., 2010). Most state-of-art approaches uses thresholding for muscle activity segmentation. However, a fixed threshold value may cause misleading results (Özgünen et al., 2010) as muscular movement has different intensities. In this study, inspired by the energy-based speech voice activity detection (Van Gerven and Xie, 1997; Ghosh et al., 2011), we propose an algorithm for the automatic segmentation of EMG signals by dynamically calculating instantaneous values for the estimation of the segmentation threshold based on an adaptive scaling parameter.

2 EMD-BASED SIGNAL DENOISING

Signal denoising can be described as following: given a noisy signal $x(t)$:

$$x(t) = f(t) + n(t) \tag{1}$$

where $f(t)$ and $n(t)$ are signal and noise components, respectively. The objective is to estimate the noise level $\sigma(n)$ and then filter out the noise. Translating this idea to the case of EMD-based denoising, firstly the noisy signal is decomposed via EMD into a number of Intrinsic Mode Functions, c_i, i.e.:

$$x(t) = \sum_{i=1}^{N} c_i(t) + r_N(t) \tag{2}$$

where N is the number of IMFs and r_N is the residual. The noise component, $n(t)$, of the signal $x(t)$ is now decomposed and dispersed among the IMFs. Thus we have to estimate the noise level $\sigma_i(n)$, where $i = 1, \ldots, N$ of each IMF.

In this study, we propose an approach for estimating $\sigma_i(n)$, which avoids removing useful information from the dominant IMFs, hence causing signal distortion. The idea, is that only $\sigma_1(n)$ is computed using Donoho's approach (Donoho and Johnstone, 1994) since the first IMF basically consist of high frequency noise component, i.e.:

$$\sigma_1(n) = MAD_1/.6745 \tag{3}$$

$$MAD_1 = Median(\mid c_1(t) - Median(c_1) \mid) \tag{4}$$

Then the noise level of the ith IMF $\sigma_i(n)$ is computed as:

$$\sigma_i(n) = \delta_i \sigma_1(n) \tag{5}$$

The procedure for computing δ_i is as following: at first the Hilbert transform (HT) is applied to each IMF. The HT of the ith IMF $c_i(t)$ is defined as (Huang et al., 1999):

$$H[c_i(t)] = c_i * \frac{1}{\pi t} = y_i \tag{6}$$

or using the convolution definition,

$$y_i(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{c_i(\gamma)}{t-\gamma} d\gamma \tag{7}$$

Where P indicates the Cauchy principal value. From $y_i(t)$ it is possible to define the analytical version of the ith IMF:

$$z_i(t) = c_i(t) + \mathbf{i} y(t) \tag{8}$$

where **i** is the imaginary unit, or in polar form:

$$z_i(t) = a_i(t) e^{\mathbf{i}\theta_i(t)} \tag{9}$$

in which:

$$a_i(t) = \sqrt{c_i^2(t) + y_i^2(t)} \tag{10}$$

$$\theta_i(t) = arctan(\frac{y_i(t)}{c_i(t)}) \tag{11}$$

The instantaneous frequency of the ith IMF ω_i is then obtained using the instantaneous variation of phase (Huang et al., 1999).

$$\omega_i(t) = \frac{d\theta(t)}{dt} \tag{12}$$

the mean period $\bar{P}_i$ is determined using the mean value of the instantaneous frequencies of the ith IMF:

$$\bar{P}_i = \frac{1}{mean(\omega_i(t))} \tag{13}$$

Following, the noise energy density of the ith IMF, E_i, is needed. Since the first IMF is always noise dominated, E_1 can be easily obtained by:

$$E_1 = \frac{1}{N} \sum_{t=1}^{N} c_1^2(t) \tag{14}$$

The product of energy density of IMF and its corresponding mean period must be a constant (Wu and Huang, 2004), i.e.:

$$E_i \bar{P}_i = C \tag{15}$$

Based on the mean period of the ith IMF $\bar{P}_i$ and the const C, the energy density of the ith IMF is obtained:

$$E_i = C/\bar{P}_i \tag{16}$$

and then:

$$\delta_i = E_i/E_1 \tag{17}$$

Having the δ_i, the noise level of the ith IMF $\sigma_i(n)$ can be obtained using Eq.5. The soft-thresholding approach (Donoho, 1995) is the applied to the IMFs:

$$c_i^{'}(t) = \begin{cases} c_i(t) - T_i, & if \;\; c_i(t) \geq T_i \\ 0, & if \;\; \mid c_i(t) \mid < T_i \\ c_i(t) + T_i, & if \;\; c_i(t) \leq -T_i \end{cases} \tag{18}$$

in which $c_i^{'}(t)$ is the thresholded version of the ith IMF, where:

$$T_i = \sigma_i(n)\sqrt{2log(L)} \tag{19}$$

Where L is the length of the signal $x(t)$. Finally, the denoised signal $x^{'}(t)$ is obtained via reconstruction using the thresholded IMFs:

$$x^{'}(t) = \sum_{i=1}^{N} c_i^{'}(t) + r_N(t) \tag{20}$$

3 Physiological Signal Analysis

3.1 Real-time QRS detection in ECG

The QRS complex is the most important segment in ECG signal, which reflects the electrical activity within the heart during the ventricular contraction. The moment of its occurrence gives us much information related to the current emotional state (Kreibig, 2010). Thus the detection of QRS complex, in particular R peak detection is the basis for ECG-based emotion recognition.

The real-time QRS detection approach proposed in this paper is based on the algorithm of change point detection of (Guralnik and Srivastava, 1999). Change point detection approaches apply data mining techniques to identify the time points at which the changes, i.e. events, occur. In (Guralnik and Srivastava, 1999) a method has been proposed for the detection of the appropriate set of number of points that minimizes the error in fitting a pre-defined function using maximum likelihood. There is no fixed number of change-points to be detected. Moreover, no constraints are imposed on the class of functions that will be fitted to the subsequences between successive change-points.

Following the notation in (Guralnik and Srivastava, 1999), let $y(t), (t = 1, ..., n)$ be the time series to be segmented. It is assumed that the time series can be modeled mathematically, where each model is characterized by a set of parameters. The problem of change-points detection, is formulated as finding a piecewise segmented model, given by

$$\begin{aligned} Y &= f_1(t, v_1) + e_1(t), (1 < t \leq \theta_1), \\ &= f_2(t, v_2) + e_2(t), (\theta_1 < t \leq \theta_2), \\ &= \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &= f_l(t, v_l) + e_l(t), (\theta_{l-1} < t \leq n). \end{aligned} \quad (21)$$

Where $f_i(t, v_i)$ is the function (with its vector of parameters v_i) that is fitted to the segment i. The θ_i's are the change-points between successive segments, and $e_i(t)$'s are error terms. The change-point in the first stage is the j minimizing $\mathcal{L}(1, j) + \mathcal{L}(j+1, n)$, say j^*.

The approach of (Guralnik and Srivastava, 1999) can be implemented in a batch or incremental. The batch algorithm is useful only when data collection precedes analysis. As in our case we are targeting on-line processing, we use the incremental version of the algorithm. The key idea is that if the next data point collected by the sensor reflects a significant change in phenomenon, then the likelihood of being a change-point is going to be smaller then the likelihood that it is not. However, if the difference in likelihoods is small, we cannot definitively conclude that a change did occur, since it may be the artifact of a large amount of noise in the data. Therefore a user-defined likelihood increase threshold is introduced.

$$(\mathcal{L}_{no\_change} - \mathcal{L}_{change}) / \mathcal{L}_{no\_change} > \delta, \quad (22)$$

where δ is a user-defined likelihood increase threshold.

Suppose that the last change-point was detected at time t_{l-1}. At time t_l the algorithm starts by collecting enough data to fit the regression model. Suppose at time t_j a new data point is collected. The candidate change-point is found by determining t_i, with likelihood criterion $\mathcal{L}_{min}(l, j)$, such that

$$\mathcal{L}_{min}(l, j) = \min_{l<i\leq j} \mathcal{L}(l, i) + \mathcal{L}(i+1, j). \quad (23)$$

If this minimum is significantly smaller than $\mathcal{L}(l, j)$, i.e. the likelihood criteria of no change-points from t_l to t_j, then t_i is a change-point. Otherwise, the process should continue with the next point, i.e. t_{j+1}.

In the incremental algorithm, execution time is a significant factor. If enough information is stored, some of the calculations can be avoided. Thus, at time t_{j+1} to find likelihood criteria

$$\mathcal{L}_{min}(l, j+1) = \min_{l<i\leq j} \mathcal{L}(l, i) + \mathcal{L}(i+1, j+1) \quad (24)$$

it is only necessary to calculate $\mathcal{L}(i+1, j+1)$, since (l, i) was calculated in the previous iteration.

Several types of basis functions, $f_i(t, v_i)$, can be considered, e.g. algebraic polynomials, wavelet, Fourier, etc.. In our current implementation polynomial fitting functions of degree 1 have been selected empirically. The user-defined likelihood threshold δ, has been set 0.97 in this study to avoid finding a lot of irrelevant change points. The remaining irrelevant change points (see Fig.1) can be filtered out by a predetermined threshold. The threshold is selected as 15% of the maximum value of the detected change points (Behbahani and Dabanloo, 2011).

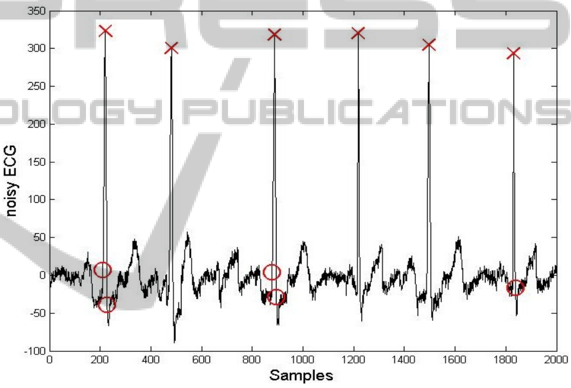

Figure 1: Detected change points including R peaks (cross) and irrelevant ones (circle)

3.2 EMG Segmentation

EMG signals are characterized by a discontinuous signal since information is carried only when muscles are active, such segments are the regions where activity information exists and are referred to as 'active segments'. The pauses between them are called 'inactive segments'. The features extracted from active segments are considered more suitable for expression recognition (Hamedi et al., 2011). For example, the active features of facial EMG is helpful for not only recognizing facial expressions, but also learning the actual mechanism of how facial expressions are formed (Aoi et al., 2011), which is less susceptible to social masking and may better reveal the emotional state behind the facial expression. The decision of determining to what class an EMG segment belongs reassembles the voice activity detection (VAD) in speech. Inspired by the energy-based speech VAD (Van Gerven and Xie, 1997; Ghosh et al.,

2011) approaches, in this paper we propose an *RMS*-based EMG segmentation. *RMS*s are extracted from each $L = 100$ ms of the EMG signal as an indicator of the muscle activation:

$$RMS(i) = \sqrt{\frac{1}{N}\sum_{t=1}^{N} x(t)^2} \tag{25}$$

where i is the index of the windows within EMG signal $x(t)$. The duration of the optimal analysis window depends on the purpose of the study. The window size of 100 ms is decided in this study to track the fast dynamic changes in facial expression (Van Boxtel, 2010).

The adaptive thresholding method for EMG segmentation is based on *RMS* levels, i.e., RMS_{min} and RMS_{max}, representing the minimum and maximum *RMS* value of the incoming windows respectively. The threshold T is calculated as (Ghosh et al., 2011):

$$T = \alpha RMS_{min} + (1-\alpha) RMS_{max} \tag{26}$$

where α is an adaptive factor and computed as (Sakhnov et al., 2009):

$$\alpha = \frac{RMS_{max} - RMS_{min}}{RMS_{max}} \tag{27}$$

RMS_{min} and RMS_{max} are initialized by the first second of the EMG signal, i.e., the first 10 windows. The RMS_{min}, RMS_{max} and corresponding T are then updated for every incoming window before thresholding. A control factor, called inactive counter (*IC*), is defined as the number of consecutive windows with *RMS* value which is smaller than T. If the *RMS* value of the incoming window i, i.e. $RMS(i)$ is larger than T, it is classified as active; If $RMS(i)$ is smaller than T, and $IC > 3$, the window i is classified as inactive, otherwise it's still labeled as active.

4 RESULTS AND DISCUSSION

4.1 Signal Denoising

In order to test the validity of the proposed denoising approach, three ECG (101, 103, 118) and two EMG (*brux*1, *brux*2) noiseless signals are chosen arbitrarily from the MIT-BIH arrhythmia database (Moody and Mark, 2001) and the CAP sleep database (Terzano et al., 2001) respectively through PhysioNet (Goldberger et al., 2000). Then simulations for several different cases are carried out. A white Gaussian noise $n(t)$ is added to the noiseless signal $f(t)$, obtaining the synthetic noisy signal $x^*(t)$ (such as 101^*, 103^*, 118^*, $brux1^*$, $brux2^*$). The quantitative evaluation is assessed by the signal-to-error ratio (*SER*) (Blanco-Velasco et al., 2008)

$$SER = \frac{\sum_{t=0}^{L-1} f^2(t)}{\sum_{t=0}^{L-1} [f(t) - x'(t)]^2} \tag{28}$$

where $x'(t)$ is the denoised version of $x^*(t)$. The signal-to-noise ratio (*SNR*) is given by:

$$SNR = \frac{\sum_{t=0}^{L-1} f^2(t)}{\sum_{t=0}^{L-1} n^2(t)} \tag{29}$$

Fig.2 and Fig.3 illustrate the waveforms of the original (clean) signal, noisy signal, and denoised signal using the proposed EMD-based approach, the traditional EMD-based approach and the wavelet-based approach, for the ECG and the EMG respectively. Here, the traditional EMD-based approach (Jing-tian et al., 2007) and the wavelet-based approach (Tikkanen, 1999) are implemented for comparison. With respect to the latter, the 4-level discrete wavelet transform with the Daubechies (db4) and Symlets (sym2) basis are used for the ECG and the EMG signals respectively. The soft-thresholding proposed by Donoho (Donoho, 1995) is then applied to the discrete wavelet transform coefficients with fixed thresholds for each level. As can be seen from Fig.2 and Fig.3, the proposed EMD-based approach yields good results in terms of visual quality and outperforms the traditional EMD-based approach and the wavelet-based approach for both the ECG and the EMG cases.

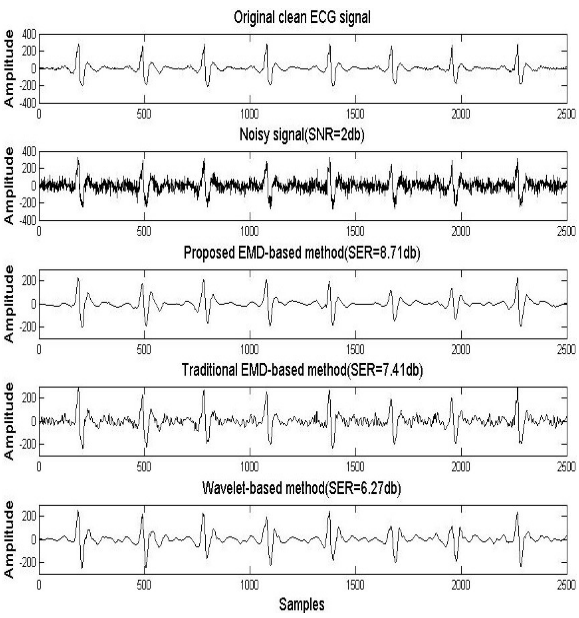

Figure 2: Example of ECG denoising for white Gaussian noise

Table 1: Signal denoising experiments carried out over several records.

	SNR = 2 db			SNR = 6 db			SNR = 10 db		
Record	EMD	tEMD	WT	EMD	tEMD	WT	EMD	tEMD	WT
101^*	7.25	5.48	5.94	16.01	10.84	10.42	21.73	15.59	13.34
103^*	6.80	4.68	5.83	15.92	10.86	11.30	19.80	18.05	15.40
118^*	8.71	7.41	6.27	13.6	11.32	12.25	23.05	16.30	14.09
$brux1^*$	8.25	2.41	3.26	18.58	2.65	4.08	28.38	3.19	4.25
$brux2^*$	6.05	2.05	1.53	11.00	2.95	1.72	15.75	3.83	1.80

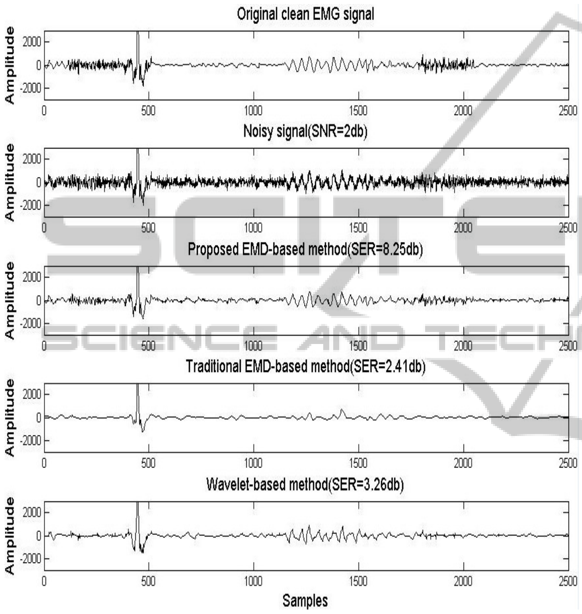

Figure 3: Example of EMG (the submentalis muscle) denoising for white Gaussian noise

Moreover, we study the behavior of the proposed method quantitatively, using multiple realizations of white Gaussian noise at different SNRs (2 db, 6 db and 10 db). The results are presented in Table.1 in terms of SER for the corresponding methods (EMD: proposed EMD-based, tEMD: traditional EMD-based and WT: wavelet-based). Again, as can be seen from Table.1, the proposed EMD-based method shows better ability to deal with the white Gaussian noise than the other two methods.

4.2 Real-time QRS Detection

We use both the synthetic and real ECG data to test the validity of proposed real-time QRS detection approach. The real ECG data were acquired for the analysis of emotional facial expressivity of Parkinson patients (Verbraeck, 2012). In the experiments for data acquisition, six emotions (happiness, sadness, anger, disgust, surprise, and fear) and a neutral state were induced via watching movies clips in a group of eight healthy subjects and seven persons suffering from Parkinson's disease. The EMG from two facial muscles (the levator labii superioris, and the orbicularis oculi), the ECG, as well as the face of each subject were recorded (Verbraeck, 2012). In total, the data used for validation in this study include 3 synthetic signals (101^*, 103^*, 118^*), 4 real signals from subjects with normal heartbeats ($N1, N2, N3, N4$) and 2 real signals from subjects with arrhythmia (A_1, A_2). All the data are noisy signals.

Table 2: The results of the QRS detection.

Record	No. of Beats	FP	FN	$Err(\%)$
101^*	1865	1	13	0.75
103^*	2084	1	15	0.77
118^*	2278	6	14	0.88
N_1	4431	5	39	0.99
N_2	4781	3	20	0.48
N_3	5288	2	9	0.21
N_4	5169	6	21	0.52
A_1	3924	12	93	2.68
A_2	4220	10	79	2.11

In Table.2, false positive (FP), indicates the method detects a beat when no beat is present; while FN, false negative, indicates that the method failed to detect a beat. The Err is the total detection failure rate, defined as:

$$Err = \frac{\sum FP + \sum FN}{\sum S} \times 100\% \tag{30}$$

where S is the total number of beats. Consequently, based on the obtained results, the proposed QRS detection approach is effective on noisy ECG signals.

4.3 EMG Segmentation

The proposed EMG segmentation approach is applied to both the synthetic data ($brux1^*$, $brux2^*$) and the real EMG signals acquired in (Verbraeck, 2012). Fig.4 shows the example of segmentation for an EMG signal from the levator labii superioris using the proposed approach. As can be seen, the validity of the

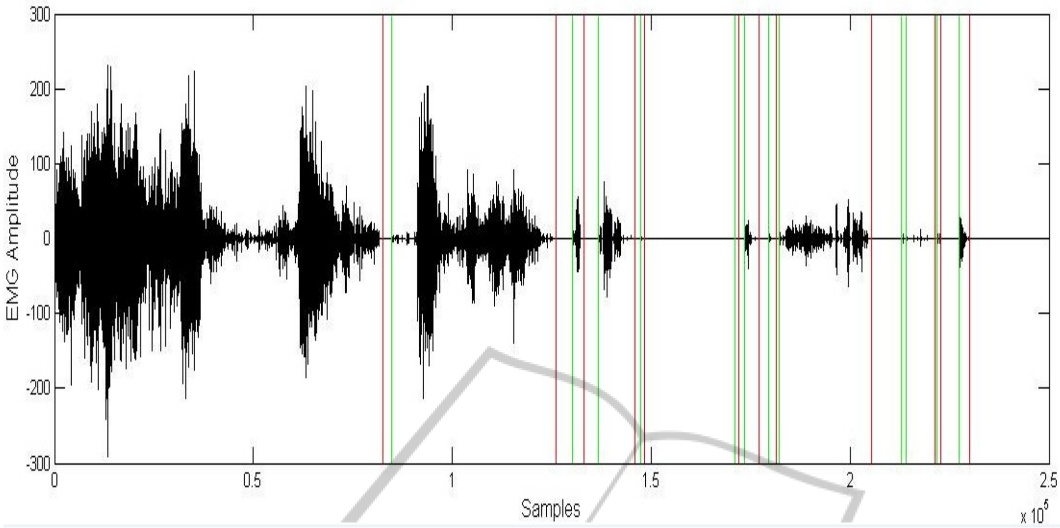

Figure 4: Example of EMG (the levator labii superioris muscle) segmentation on real data.

proposed approach has been confirmed: the EMG signals are segmented into several active segments between the green and the following red lines. Even short and weak active segments have been detected.

5 CONCLUSIONS

This paper introduces novel approaches of physiological signal processing for emotional feature extraction. Both the ECG and the EMG signals are addressed. Firstly an EMD-based denoising approach is proposed where the noise levels of all IMFs are effectively estimated and processed to successfully achieve the denoising. The validity of the proposed method is confirmed through several experiments. Results indicate that it is an effective tool for physiological signals denoising, especially for the case of white Gaussian noise. In addition, a real-time QRS detection approach is proposed which can be applied directly on the noisy ECG signals. Finally, an adaptive thresholding approach is employed for EMG segmentation. Both approaches for QRS detection and EMG segmentation are validated using synthetic and real physiological data resulting in good performances.

ACKNOWLEDGEMENTS

The research work reported in this paper is supported by a CSC-VUB scholarship grant, and the VUB-interdisiplinary research project ”Objectifying assessment of human emotional processing and expression for clinical and psychotherapy applications (EMO-App)”.

REFERENCES

Agrafioti, F., Hatzinakos, D., and Anderson, A. K. (2012). Ecg pattern analysis for emotion detection. *IEEE Transactions on Affective Computing*, 3:102–115.

Andrade, A. O., Nasuto, S., Kyberd, P., Sweeney-Reed, C. M., and Kanijn, F. R. V. (2006). Emg signal filtering based on empirical mode decomposition. *Biomedical Signal Processing and Control*, 1:44–55.

Aoi, M., Kamijo, M., and Yoshida, H. (2011). Relationship between Facial Expression and Facial Electromyogram (f-EMG) Analysis in the Expression of Drowsiness. In *International Conference on Biometrics and Kansei Engineering*, pages 65–70.

Behbahani, S. and Dabanloo, N. (2011). Detection of qrs complexes in the ecg signal using multiresolution wavelet and thresholding method. In *Computing in Cardiology, 2011*, pages 805–808.

Blanco-Velasco, M., Weng, B., and Barner, K. E. (2008). Ecg signal denoising and baseline wander correction based on the empirical mode decomposition. *Computers in Biology and Medicine*, 38:1–13.

Boudraa, A. O., Cexus, J. C., and Saidi, Z. (2005). Emd-based signal noise reduction. *Signal Processing*, 1:33–37.

Donoho, D. L. (1995). Denoising by soft-thresholding. *IEEE Transactions on Information Theory*, 41:613–627.

Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaption via wavelet shrinkage. *Biometrika*, 81:425–455.

Farfán, F., Politti, J., and Felice, C. (2010). Evaluation of emg processing techniques using information theory. *Biomedical engineering online*, 9(1):72.

Ghosh, P. K., Tsiartas, A., and Narayanan, S. (2011). Robust voice activity detection using long-term signal variability. *Audio, Speech, and Language Processing, IEEE Transactions on*.

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. (2000). Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220.

Guralnik, V. and Srivastava, J. (1999). Event detection from time series data. In *Knowledge Discovery and Data Mining*, pages 33–42.

Hamedi, M., Salleh, S.-H., and Swee, T. T. (2011). Surface electromyography-based facial expression recognition in bi-polar configuration. *Journal of Computer Science*, 7(9):1407–1415.

Huang, N. E., Shen, Z., and Long, S. R. (1999). A new view of nonlinear water waves: The hilbert spectrum 1. *Annual review of fluid mechanics*, 31(1):417–457.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, 454:903–995.

Jing-tian, T., Qing, Z., Yan, T., Bin, L., and Xiao-kai, Z. (2007). Hilbert-huang transform for ecg de-noising. In *Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on*, pages 664–667.

Karagiannis, A. and Constantinou, P. (2009). Noise components identification in biomedical signals based on empirical mode decomposition. In *IEEE EMBS International Conference on Information Technology Applications in Biomedicine*, pages 1–4.

Kim, J. (2007). Bimodal emotion recognition using speech and physiological changes. *Robust speech recognition and understanding*, pages 265–280.

Kohler, B.-U., Hennig, C., and Orglmeister, R. (2002). The principles of software qrs detection. *IEEE Engineering in Medicine and Biology Magazine*, 21:42–57.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological psychology*, 84(3):394–421.

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., and Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear-and sadness-inducing films. *Psychophysiology*, 44(5):787–806.

Moody, G. B. and Mark, R. G. (2001). The impact of the mit-bih arrhythmia database. *Engineering in Medicine and Biology Magazine, IEEE*, 20(3):45–50.

Özgünen, K. T., Çelik, U., and Kurdak, S. S. (2010). Determination of an optimal threshold value for muscle activity detection in emg analysis. *Journal of Sports Science and Medicine*, 9:620–628.

Pantelopoulos, A. and Bourbakis, N. (2008). A survey on wearable biosensor systems for health monitoring. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 4887–4890.

Picard, R. W. (2000). *Affective computing*. MIT press.

Sakhnov, K., Verteletskaya, E., and Simak, B. (2009). Approach for energy-based voice detector with adaptive scaling factor. *IAENG Internat. J. Comput. Sci*, 36(4).

Terzano, M. G., Parrino, L., Sherieri, A., Chervin, R., Chokroverty, S., Guilleminault, C., Hirshkowitz, M., Mahowald, M., Moldofsky, H., Rosa, A., et al. (2001). Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (cap) in human sleep. *Sleep medicine*, 2(6):537–553.

Tikkanen, P. E. (1999). Nonlinear wavelet and wavelet packet denoising of electrocardiogram signal. *Biological Cybernetics*, 80(4):259–267.

Üstündağ, M., Gökbulut, M., Şengür, A., and Ata, F. (2012). Denoising of weak ecg signals by using wavelet analysis and fuzzy thresholding. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 1(4):135–140.

Van Boxtel, A. (2010). Facial emg as a tool for inferring affective states. In *Proceedings of Measuring Behavior 2010*, pages 104–108.

Van Gerven, S. and Xie, F. (1997). A comparative study of speech detection methods. In *Eurospeech*, volume 97.

Verbraeck, F. (2012). Objectifying human facial expressions for clinical applications. Master's thesis, Vrije Universiteit Brussel, Belgium.

Wu, Z. and Huang, N. E. (2004). A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, 460:1597–1611.