Experiences with adaptive statistical models for biosignals in daily life

Stijn de Waele, Gert-Jan de Vries, Mark Jäger
Philips Research
High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
stijn.de.waele@philips.com

Abstract

We discuss the merits of adaptive statistical models for biosignals in a daily life context. Processing of this type of signals poses a number of challenges. First, it is clear that an adaptive model is needed to tailor for the differences in physiology between individuals, as well as adapt to someone's current physiological state. Second, in a daily life setting we use unobtrusive measurement devices, which will lead to reduced signal quality compared to the laboratory setting. Third, low-power portable sensors allow for only limited data storage and data transmission. Two techniques to address these challenges are discussed in detail: the usage of the cumulative histogram and parametric models. We show applications to electroencephalogram (EEG), electrocardiogram (ECG) and skin conductance (SC) signals and we advise on how to obtain the most reliable results.

1. Introduction

In this paper, we consider biosignal processing algorithms for use in daily life situations. We intend to apply these algorithms in a range of consumer applications such as relaxation devices, circadian rhythm management and attention measurement. The usage of these algorithms in consumer applications implies that the algorithms must function in real-time and without (expert) user intervention.

The requirements on biosignal processing are different from those of medical biosignal processing. To take into account the variation in physiological responses across different activities (e.g. desk work, travel) as well as differences between individuals, an adaptive model is needed. Moreover, the signal quality will be reduced compared to a laboratory or hospital bed environment. This is due to artifacts caused by increased movement, as well as the requirement for unobtrusive measurements. Finally, small and low-power sensors are used that are limited in memory size and wireless transmission bandwidth.

In this paper we will discuss our experiences with biosignal processing algorithms for electroencephalography (EEG), skin conductance (SC) and electrocardiogram (ECG) measurements in the daily life context. We consider the processing chain from pre-processing (including artifact removal), feature extraction and, finally, the combination of these features into estimates of the user's mental state.

An element that we see recurring in robust algorithms is adaptive statistical models, of which we will discuss cumulative histograms and parametric models in more detail. The cumulative histogram is used for robust normalization of features and for outlier detection. Parametric models have the advantage that they are more accurate (in particular under reduced signal quality) and offer the possibility of data reduction than non-parametric models. We focus here on time series models for spectral estimation.

The hardware that we used to gather our measurement data are our Emotions Measurement Platform [1] and the Nexus-10 [2]. Our companion paper at this conference [1] describes the Emotions Measurement Platform and the results of the first user tests.

The outline of the paper is as follows. In section 2, we discuss definitions and basic results related to adaptive statistical models. Algorithms for Skin Conductance and ECG using these models are discussed in sections 3 and 4, respectively. In section 0, we discuss in more detail spectral peak detection from EEG. Finally, we draw conclusions in section 6.

2. Adaptive statistical models

Statistical models are a parametric representation of the distribution function of a stochastic variable. In this context, the stochastic variable is typically a signal as a function of time, also known as a time series. We consider two ways to describe this type of signal x as a function of time n: $x = x_1, \dots, x_n, \dots, x_N$ (. The number of available observations is denoted N. In section 2.1, we introduce the time series model as a means to describe autocorrelations, or correlations between the signal at time lag r, x_n and x_{n-r}; in section 2.3 we introduce the cumulative histogram, which describes the distribution function of x.

2.1. Time series models

Time series models provide a method to determine the correlation structure of the signal x. The correlation structure can be described by either the autocovariance function or the power spectrum. The autocovariance function R for a stationary stochastic signal x_n is defined as the correlation between x_n and x_{n-r} [3]:

978-1-4244-4799-2/09/$25.00 ©2009 IEEE

$$R(\mathrm{r}) = \mathrm{E}\{x_n x_{n-\mathrm{r}}\}. \quad (1)$$

A stationary stochastic signal is a signal for which the statistical properties are time-independent. Therefore, the autocorrelation function does not depend on the time n, as we consider x to be a stationary signal. The Fourier Transform $F\{\}$ of the autocorrelation function is the power spectrum h:

$$h(f) = F\{R\} \quad (2)$$

As the Fourier Transform is a one-to-one mapping, this means that the power spectrum and the autocorrelation function contain exactly the same information.

A time series model is an accurate model for the correlation structure. A useful model structure is that of the Autoregressive-Moving Average (ARMA) model [4]:

$$x_n + a_1 x_{n-1} + \cdots + a_p x_{n-p} = \varepsilon_n + b_1 \varepsilon_{n-1} + \cdots + b_n \varepsilon_{n-q,} \quad (3)$$

where ε is a white noise signal with variance σ_ε^2 and p and q are the autoregressive order and the moving average order, respectively.

The autocorrelation function and power spectrum can be derived from the parameters $a_1, a_2, \ldots, a_p$ and $b_1, b, \ldots, b_q$ [4]. The ARMA model can be used to approximate any true power spectrum, given that sufficient parameters are used. This is very similar to the Taylor approximation, where a polynomial is used to approximate a large class of functions. This result still holds if we restrict the model to the pure Autoregressive (AR) ($q=0$) or pure moving average (MA) model.

To obtain an accurate model for a large variety of data, we need a procedure for type selection (AR, MA or ARMA), order selection (the number of parameters $p+q$) and robust parameters estimation. With recent advances in time series analysis, all of these procedures are available [5]. With this result, time series analysis has become an attractive alternative for the Fast Fourier Transform (FFT) or (windowed) periodogram, which is a non-parametric approach to spectral estimation.

From statistical theory, we derive the requirement for accurate parameter estimators that it is asymptotically equivalent to the maximum likelihood estimator [3]. The estimators used for time series analysis all meet this requirement, which explains their increased accuracy. No such methodology is available for non-parametric approaches such as the FFT or wavelets.

Most biosignals are not stationary in the strict sense; however, we can consider them to be *piecewise* stationary. This means that a time series models is determined for a short epoch and then updated regularly. For example, for the heart rate signal, we can determine a model for the last 2 minutes, and update this model every 10 seconds. The changing characteristics are visualized by means of a time-frequency plot of the power spectrum calculated from the model parameters.

The increased accuracy of the time series model can be exploited in 2 ways. Either we obtain more accurate results using the same amount of data or we obtain results with the same accuracy using less data.

A further advantage of time series models is that it results in a significant data reduction: a large data set is represented by a limited number of model parameters. This can be used to reduce the required data storage and wireless transmission bandwidth.

2.2. Spectral peak detection

Using the time series models introduced above, we can accurately derive the location of spectral peaks. We use this for spectral peak detection from heart rate data (see section 4) and for spectral peak detection from EEG (see section 0).

For the application of peak detection the AR model alone is typically already quite accurate, so we do not consider MA or ARMA models. This reduces the numerical complexity of the algorithm.

The power spectrum for the AR model can be expressed directly in terms of the AR parameters. However, here we use the expression in terms of the poles λ_k :

$$h(f) = \frac{1}{\sigma_\varepsilon^2} \frac{1}{\prod_{k=1}^{p} |(1 - \lambda_k e^{-2\pi i f})|^2} \quad (4)$$

The poles can be calculated directly from the parameters a_k [4]. Real-valued poles correspond to a peak located at frequency f=0. For complex-valued poles, the phase of the pole corresponds to the peak location.

AR modeling is done using the Matlab toolbox ARMASA [6]. In this toolbox, the Burg algorithm is used for parameter estimation and CIC as order selection criterion [5].

2.3. Cumulative histogram

The normalized cumulative histogram $H(k)$ is defined as the number of observations with a value of at most k:

$$H_x(k) = \frac{1}{N} |\{x_n : x_n \leq k\}|, \quad (5)$$

where $|.|$ is the number of elements in a set. The value of x where $H(x) = P$ is referred to as the P quantile (or: the $100P$ percentile). The cumulative histogram is an estimate for the cumulative distribution function of x. Combined with the estimated time series model we obtain a quite complete statistical description of the process at hand.

The cumulative histogram can be used for signal scaling: By calculating the cumulative histogram of a signal value, we obtain a value between 0 and 1.

Quantiles are generally preferred over linear averages, because they are less sensitive to a few large outliers.

We use the cumulative histogram in a number of related applications:

- Peak detection from ECG data
- Outlier removal: Signal values x_i can be considered as outliers if $H_x(x_i)$ is smaller than $f_{low}H_{low}$, where H_{low} is a histogram value ($0<H_{low}<1$) and f_{low} is a multiplication factor smaller or equal to 1. Similarly, we can detect outliers based on an upper bound $f_{high}H_{high}$, with $f_{high}\geq 1$.
- Event detection, e.g. highly emotional arousing events based on skin conductance.
- Signal normalization: The reduction to the fixed domain [0..1] allows for normalization. This is useful for cross person comparison (as applied in, e.g., [15]), but as well for sensor fusion, when combining signals from different modalities, see below.

In real time applications we apply a sliding window version of the cumulative histogram. It relates the last measured value x_n with its predecessors within time interval τ (in number of samples):

$$H_x^{\tau}(x_N) = \frac{1}{\tau}|\{x_n : x_n \leq x_N \ \wedge N-\tau \leq n < N\}|. \quad (6)$$

Typically, the cumulative histogram has to be initialized when starting a measurement for a certain user. The time needed for initialization depends on the time interval τ taken into account. For peak detection and outlier removal, the duration of the initialization is quite short (around 1 minute). For emotional event detection, the initialization time depends on the type of events we want to detect. The window size of the windowed cumulative histogram should be at least the typical time interval between the events that we want to detect. Personalization is ensured by determining a new histogram for each user.

The cumulative histogram is a non-parametric description of the data. This has the advantage that there is no dependency on the quality of the used model. However, a parametric model has the potential to be more accurate, as discussed in the previous section. This accuracy can also translate into shorter initialization times.

3. Skin conductance analysis

In the human skin there are two types of sweat glands: the eccrine and apocrine glands [9][10][11]. The eccrine glands cover the complete body but are most densely populated on the palms of the hands and soles of the feet. These glands contain a sweat duct that is filled to a variable level with sweat, thereby influencing the electrical conductivity of the skin, as measured by skin conductance sensors.

Eccrine glands' primary function is thermoregulation, but they are also influenced by emotional stimuli as they are linked to the sympathetic part of the autonomous nervous system (ANS). Whereas on other parts of the body than the hands and feet, the density of eccrine glands is lower, eccrine glands all over the body are believed to reflect emotions [10].

3.1. Measurement location

Because of the higher density of eccrine glands on the palmar and sole areas, these areas are usually taken as measurement locations. In daily life, however, measurement at the feet poses a big challenge because the large sweat production by the enclosure in shoes potentially fills the sweat ducts completely, leaving no room for measuring emotional deviations. On the other hand, measurement at the palms, either on the fingers or thenar region, is considered very obtrusive in daily live, as we already touched upon.

We found the wrist to be a suitable location for daily life measurement. People are used to wearing watches at these body locations, therefore a measuring device at the wrist is not considered obtrusive. The signal retrieved from the wrist closely resembles the signal from the palmar area, for most people we studied, although the signal is smaller and shows smaller responses as well (see Figure 1). The sensors to pick up the signal from the wrist therefore need to be more sensitive.

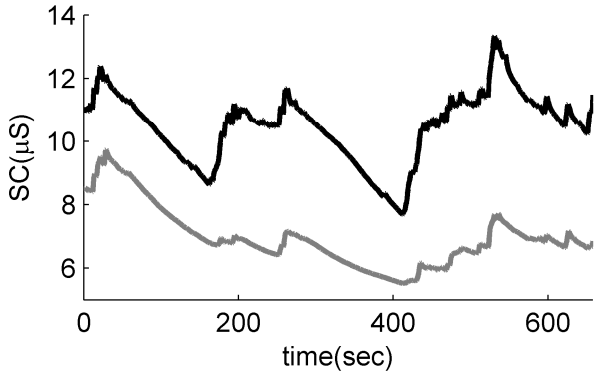

Figure 1. Example trace of skin conductance simultaneously measured at the fingers (black) and wrist (grey).

3.2. Event detection

As discussed above, skin conductance is believed to reflect emotions. Emotions are generally considered as an affective state that has a short time span and a clear cause, and are therefore event based. Changes in SC can either be analyzed on the tonic level (global changes) or phasic level (local changes; individual skin conductance responses), see [1] for a more detailed description.

In our daily life research, we are usually not interested in detecting changes in skin conductance due to each and every event. We want to detect those moments in time that show extraordinary large reactions, in the context of daily life emotions that occur throughout a single, or a small number of, day(s).

In order to do so, we apply the cumulative histogram, in a windowed fashion, to the skin conductance signal. In addition we define an adaptive threshold in terms of a high quantile (e.g., 0.95) of the cumulative histogram

and trigger in case of sufficient exceedence. Setting the window size to the order of hours results in triggers at those moments throughout the day that the skin conductance is relatively high. These usually coincide with emotional events. The triggers can, for example, be used in research applications to popup a questionnaire, in order to probe emotion in daily life, or in user applications to create awareness of one's emotional states, or change the behavior or appearance of a user interface.

Current research comprises the execution of a user test to validate our method of detecting high arousing moments during daily life, using the method described above. First results indicate a good match between the trigger moments and the moments at which participants feel highly aroused. A more elaborate data analysis, however, still has to be carried out.

In a similar fashion as the above described event detection in skin conductance signals, events can be detected in fused sensor signals. In order to compare signals originating from different modalities it is advisable to convert these signals to a single domain, thereby enabling direct comparison. As described above, the cumulative histogram provides such a single domain. The signals to be fused (say x and y) can be merged in terms of their cumulative histograms, e.g., by taking a weighted average:

$$H_z = aH_x + bH_y, \quad (7)$$

where a and b are weighting coefficients.

In the context of emotions, sensor fusion can be applied to combine multiple physiological features that are known to reflect a single affective dimension. As an example, it is known that both skin conductance level and heart rate reflect arousal [8].

3.3. Artifact removal

Physical activity, as it occurs in daily life situations, can negatively affect the contact between the skin and electrodes of the measuring device. Loss of contact results in a very large resistance, therefore an infinitesimal (skin) conductance. By applying a lower bound on the SC one can exclude the signal in case the contact between skin and electrodes is lost. This lower bound should be chosen just above the bottom of the measuring range in order to reduce the range as less as possible, however be robust against the background noise that can be observed in case of loose contact.

Another artifact, we sometimes experience in the measured SC signal are glitches. These glitches manifest as single extreme values surrounded by the 'normal' SC signal. We filter out these outliers by using the window histogram H_x^τ with multiplication factor f_{high}, as described in section 2.3.

4. Heart rate variability analysis

Heart rate analysis consists of following steps:

- R-peak detection
- IBI outlier removal
- Heart rate variability analysis

For each of these steps we have made a substantiated choice of existing algorithms that are appropriate for the daily life context. For R-peak detection, we use the matched filter [7] or template matching. For noisy ECG data, this was found to outperform the frequently-used Pan-Tompkins approach [12], which tends to amplify noise due to the fact that high-order derivatives of the signal are taken.

Conversely, the matched filter suppresses the noise, as is illustrated in Figure 2. The peak signal to noise ratio (PSNR) is calculated as the match level at the detected peak location (here set to t=0) and the match level at a short time interval w just prior to the peak.

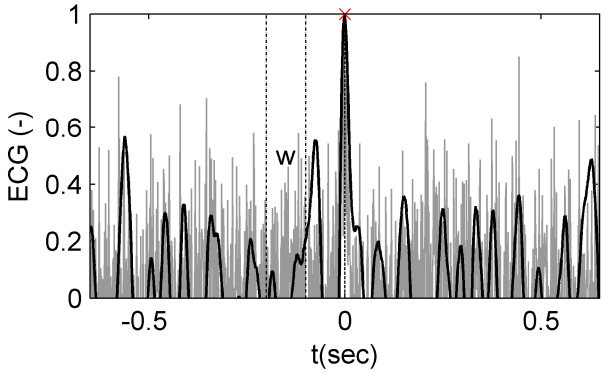

Figure 2 Peak match level (black) compared to the original ECG signal (grey) for noisy ECG data. In this example, the peak signal to noise ratio (PSNR) for the peak match level is 19.9, a factor of 2.8 higher than for the original signal.

In our Emotions Measurement Platform, peak detection is done on the sensor nodes [1]. In this way, the bandwidth needed for wireless data transmission is strongly reduced compared to transmission of the raw ECG signal. The quality of the detected peak, as expressed by the PSNR is also transmitted, to allow for signal quality evaluation at the receiver side.

The next step is IBI outlier removal. We use quantiles of the sliding window histogram H_x^τ as described in section 2.3, with H_{low}=10%, f_{low}=0.7, H_{high}=90% and f_{high}=1.4. This approach is effective in removing the influence in the IBI signal of missed beats and incorrectly detected peaks.

Finally, heart rate variability analysis is done by using the estimated time series model to estimate the power in the high frequency (HF) band as defined in the heart rate variability guidelines paper [13], ranging from 0.04 to 0.15 Hz. The power in this band is known to vary with parasympathetic nervous system (PNS) activity [14].

5. Spectral peak detection from EEG

The most dominant peak in the EEG spectra is the alpha-peak, located in the frequency band between 8 and 13 Hz [17]. The power in the alpha band, as well as the alpha peak location, carry important information for characterization of cognitive performance [18].

Spectral peak detection from the heart rate signal can be in much the same way as in EEG analysis. Here, we discuss peak detection from EEG in more detail, as special care has to be taken in the correct pre-processing of this data.

5.1. EEG preprocessing

We consider the problem of estimating peak position in the alpha frequency band from EEG data. The original sampling frequency of the signal is f_s = 200 Hz.

Possible EEG pre-processing steps include low-pass filtering and downsampling. Low-pass filtering is done for the purpose of removing high frequency content. Furthermore, downsampling can be applied in case we are not interested in high-frequency content. Typically, we are not interested in frequencies above 50 Hz, so, using Shannon's sampling theorem, we can perform downsampling to 100 Hz (a factor of 2 in the current example).

Downsampling has the well-known benefit of reducing the computational load as well as data bandwidth. We will now show that is also leads to improved accuracy of the estimated peak location, by comparing the accuracy of the estimated peak location when using filtering only (**method 1**) versus filtering + downsampling (**method 2**).

First, we will discuss the effect of the 2 pre-processing techniques on spectral estimation using theoretical results from time series analysis. Numerical results are given in the next subsection.

Both method 1 and 2 contain the first step of low-pass filtering. The impact of this filter is to reduce the power in the higher frequencies. In absolute terms, practically no power is left, since the power has been reduced by 5 to 20 dB. However, Maximum Likelihood parameters estimation does not evaluate the spectrum in absolute terms, but rather tries to obtain the best fit in logarithmic terms. So, a large part of parameters in the AR model will be spent to correctly describe the high-frequency area. This will reduce the model quality in the frequency range of interest. The extreme low power area introduced artificially by filtering is difficult to model with a limited number of parameters.

In method 2, downsampling is part of the preprocessing. This means that the high-frequency components have been removed from the signal. All modeling effort is now focused on the frequency range of interest. Therefore, it is expected that we can obtain a more accurate result, while using less model parameters to achieve it.

In summary, method 2 is expected to yield more accurate power spectra. At the same time, a considerable reduction in computation effort is achieved due to the reduced number of samples.

5.2. Simulations and experimental data

To test the two methods for spectral estimation as introduced in the previous section, a simulation study was done. A simulation experiment is the most appropriate tool here because in simulations the true spectrum is known. This allows us to calculate the quality of the estimated spectra. The generating process is the sum of 2 AR(2) processes, with the main peak frequency at 10 Hz and a smaller, secondary peak at 20 Hz. This simulation example is selected to have a close agreement with the experimental data that are discussed below. The number of observations N generated equals 2000, or 10 seconds at the sampling rate of 200 Hz. The number of simulation runs is 1000.

The error in the alpha peak location of 10 Hz was determined for the selected models. For method 1, the average error is 0.19 Hz; for method 2, the error is 0.10 Hz: a reduction in error by a factor of 2. A set of 5 typical simulation results is given in figure 3 for method 1 (top) and method 2 (bottom).

It was found that for method 1, the selected model order was considerably larger than for method 2, which results in a reduced accuracy of the spectrum in the frequency range of interest. This validates the expectation based on theoretical arguments as expressed in the previous section.

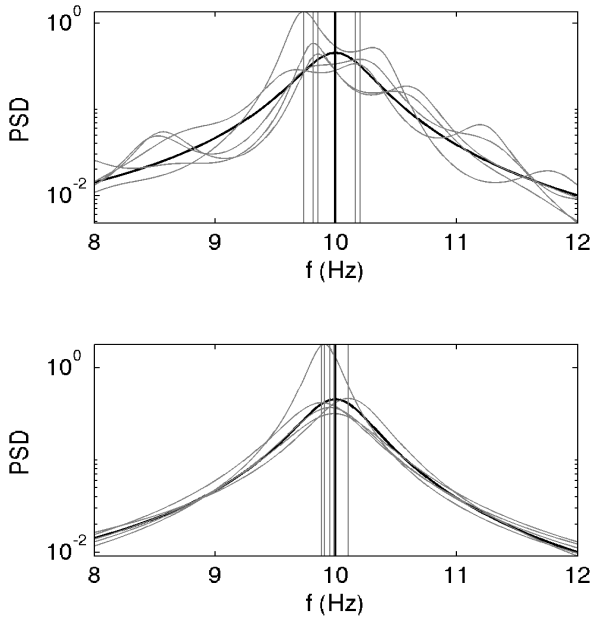

Figure 3 Detected peak location in the alpha frequency band using the selected autoregressive model for a set of 5 typical simulation runs for method 1 (no downsampling; top graph) and method 2 (with downsampling; bottom graph).

The two methods introduced in section 5.1 were also compared for real EEG measurements. Ongoing EEG was recorded for 10 seconds at scalp location C3-A1 (according to internationally accepted 10-20 system

[19]) with 'wet' electrodes and the NeXus-10 amplifier. From previous measurements it was known that besides the alpha peak at 10 Hz the subject also shows a secondary peak at 20 Hz during a resting eyes-closed condition.

The experimental data was processed as described in section 5.1. Figure 4 compares the spectra obtained with both methods. The benefits of downsampling are clearly visible. The upper spectrum (no downsampling) shows a significantly higher distortion compared to the lower spectrum (with downsampling). Without downsampling, it is difficult to recognize the secondary peak around 20 Hz in the upper spectrum of figure 3.

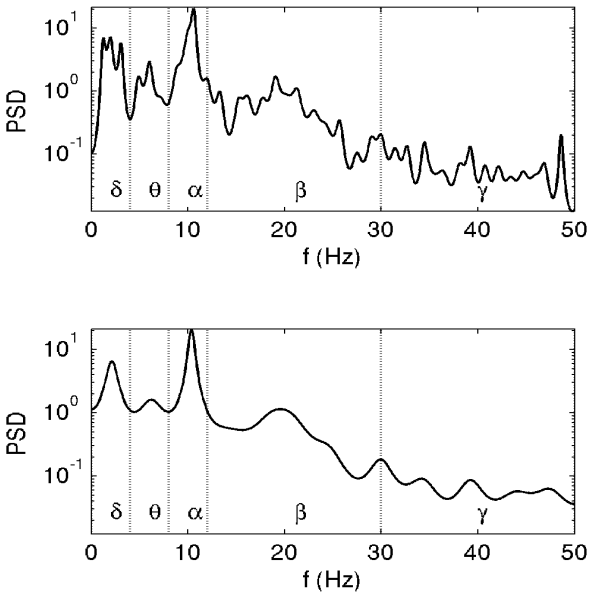

Figure 4 Estimated spectrum with the selected autoregressive model for experimental EEG data for method 1 (no downsampling; top graph) and method 2 (with downsampling; bottom graph). The two methods show the same qualitative behavior as in the simulation experiment.

6. Conclusions

We have shown the application of parametric models as well as the cumulative histogram in the biosignal processing. Parametric time series models have the potential to provide highly accurate models when used correctly. Furthermore, they provide a compact description of the data, thus reducing the requirements on data storage and transmission.

For EEG, the application of downsampling as a preprocessing step both increases model accuracy and reduces the numerical load of signal processing algorithms.

We have shown the various applications of the cumulative histogram. Its primary strength is the detection of extreme signal values relative to a current state, either to detect outliers (unwanted) or to detect alterations of the current state (wanted extraordinary values). The application to skin conductance measurements has been discussed in more detail. By excluding signal components that are not related to emotional events, we improve the robustness to the disturbances introduced by unobtrusive sensors and the daily life setting.

References

[1] J. Westerink, M. Ouwerkerk, G.-J. de Vries, S. de Waele, J. van den Eerenbeemd, and M. van Boven, Emotion measurement platform for daily life situations, International Conference on Affective Computing & Intelligent Interaction (ACII), Amsterdam, The Netherlands, September 2009.

[2] Nexus-10 Amplifier, www.tmsi.nl.

[3] M. B. Priestley, Spectral Analysis and Time Series, Academic Press, London, 1981.

[4] P. Stoica and R. L. Moses, Introduction to Spectral Analysis, Prentice Hall, Upper Saddle River, 1997.

[5] P. M. T. Broersen, Automatic autocorrelation and spectral analysis, Springer-Verlag, London, 2006.

[6] P. M. T. Broersen, ARMASA Matlab toolbox, http://www.mathworks.com/matlabcentral/fileexchange/1330.

[7] H. Vincent Poor, An introduction to signal detection and estimation, 2nd ed, Springer, New York, 1994.

[8] L. R. Brody, and J. A. Hall, Chapter 22: Gender, Emotion, and Expression. From M. Lewis, and J. M. Haviland-Jones (eds.), Handbook of Emotions, 2nd edition. The Guilford Press, 2000.

[9] W. Boucsein, Electrodermal activity. New York, NY, USA, Plenum Press, 1992.

[10] M. E. Dawson, A. M. Schell, and D. L. Filion. Chapter 8: The electrodermal system. From Handbook of Psychophysiology, 2nd edition. Cambridge University Press, 2000.

[11] K. Sato, The physiology, pharmacology, and biochemistry of the eccrine sweat gland. Review of Physiology, Biochemistry and Pharmacology, 79:51-131, 1977.

[12] J. Pan, W. J. Tompkins, A Real-Time QRS Detection Algorithm, IEEE Trans. Biomedical Engineering, 32(3):230-236, 1985.

[13] Taskforce of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, Heart rate variability, Standards of measurement, physiological interpretation and clinical use, European Heart Journal 17:354-381, 1996.

[14] P. Grossman, and E. W. Taylor, Toward understanding respiratory sinus arrhythmia: relations to cardiac vagal tone, evolution and biobehavioral functions, Biol. Psychol, 74:263-85, 2007.

[15] A. G. Money and H. Agius, Analysing user physiological responses for affective video summarisation. Displays, 30(2):59-70, 2009.

[16] J. T. Cacioppo, L. G. Tassinary, G. Berntson (Eds), Handbook of Psychophysiology, Cambridge University Press, 2000.

[17] B. J. Fisch, Fisch & Spehlmann's EEG primer: Basic principles of digital and analog EEG, Elsevier, Amsterdam, 1999.

[18] W. Klimesch, EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis, Brain Research Reviews, 29:169-195, 1999.

[19] H. H. Jasper, The ten-twenty electrode system of the international federation, Electroencephalography and Clinical Neurophysiology, 10:371-3375, 1998.