33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Towards Mental Stress Detection Using Wearable Physiological Sensors

Jacqueline Wijsman, Bernard Grundlehner, Hao Liu, Hermie Hermens, and Julien Penders

Abstract— **Early mental stress detection can prevent many stress related health problems. This study aimed at using a wearable sensor system to measure physiological signals and detect mental stress. Three different stress conditions were presented to a healthy subject group. During the procedure, ECG, respiration, skin conductance, and EMG of the trapezius muscles were recorded. In total, 19 physiological features were calculated from these signals. After normalization of the feature values and analysis of correlations among these features, a subset of 9 features was selected for further analysis. Principal component analysis reduced these 9 features to 7 principal components (PCs). Using these PCs and different classifiers, a consistent classification accuracy between stress and non stress conditions of almost 80% was found. This suggests that a promising feature subset was found for future development of a personalized stress monitor.**

I. INTRODUCTION

The second most frequently occurring type of work-related health problems in the European population is 'stress, depression or anxiety' [1]. Of the sickness absence for one month or more, 25% was caused by stress, depression or anxiety. These figures indicate that stress is a major financial and social problem in European society.

Chronic mental stress can cause health problems which include for example hypertension [2], cardiovascular diseases [3], increased likelihood of infections [2] and depression [4]. If mental stress could be detected in an early stage, stress related health problems could be prevented.

Stress is known to activate the sympathetic nervous system (SNS) [4]. Much research has already been done on the detection of stress from physiological parameters that are influenced by the SNS. Examples are muscle activity, heart rate, heart rate variability, skin conductance and pupil diameter [5]–[8]. Other studies have shown that a combination of these physiological parameters facilitates differentiation between stressful situations and situations without stress [8]–[14]. These studies all measure signals in different situations (Stroop test, mental arithmetic, movie and sound fragments, car driving) than we did. We tried to mimic daily work stress with our newly developed test protocol that consisted of problem solving puzzle tasks and a memory task done in team effort.

Manuscript received June 20, 2011

J. Wijsman is with Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Drienerlolaan 5, 7522 NB Enschede, The Netherlands and with Holst Centre/imec, High Tech Campus 31, 5656 AE Eindhoven, The Netherlands; Email: Jacqueline.Wijsman@imec-nl.nl, Telephone: +31 404020400, Fax: +31 404020699

B Grundlehner, H. Liu, and J. Penders are with Holst Centre/imec; Email: {Bernard.Grundlehner, Hao.Liu, Julien.Penders}@imec-nl.nl

H. Hermens is with Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente; Email: H.J.Hermens@utwente.nl

A system that can measure stress levels based on physiological signals will create interesting applications in real life situations. Ultimate goal is to design an easily wearable wireless system that can measure real time stress levels. A possible application of such a system could be stress prevention at work.

The goal of this study is to find the physiological signals and features that show the most distinct reaction to mental stress in the conditions included in our protocol. Once these features are identified, it should be possible to construct a reliable stress measure out of these features.

A short description of the test protocol, signal acquisition system, data processing, methods for feature extraction and statistical analysis are presented in section II. Results of the analysis of the physiological signals during the stress tests, in comparison to relaxing periods, are presented in section III. The discussion can be found in section IV. Finally, conclusions are drawn in section V.

II. METHODS

A. Experimental protocol

The experimental protocol is only discussed briefly here. For a detailed description of the protocol, and a validation based on questionnaire answers, the reader is referred to [15].

A total number of 30 healthy subjects were recruited to participate in a protocol specially designed for this study. The ages of the subjects were in the range 19-53 (mean = 33.1; SD = 7.87); 25 subjects were male and 5 female. The test was performed on a PC in a quiet room.

The subjects answered some general questions first, and filled in the perceived stress scale (PSS) questionnaire. Then they performed a reference contraction, followed by exposure to three different stress conditions: a calculation task (the Norinder test, 2:30 min), a logical puzzle task (3:00 min) and a memory task (approximately 5:00 min). All three tasks were done under time pressure and with distracting news fragments that were heard through headphones. Social pressure was induced in the memory task by telling the subjects that their performance would be included in a group result and published to colleagues afterwards. In between the stress conditions, 2:00 min resting periods were scheduled to make sure that one condition would not influence the next. Furthermore, a questionnaire had to be completed before and after each of the conditions.

B. Physiological recordings

One lead electrocardiography (ECG) and respiration were measured using a wireless chest belt. ECG was measured with commercial gel electrodes; respiration was measured

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

TABLE I
OVERVIEW OF FEATURES EXTRACTED FROM THE MEASURED SIGNALS

Number	Feature	Abbreviation	Meaning
1	Heart rate	HR	Mean heart rate
2	Standard deviation interbeat intervals	SDNN	Mean standard deviation of the interbeat intervals
3	Low frequency heart rate variability	LF	Heart rate variability in the 0.04-0.15 Hz band
4	High frequency heart rate variability	HF	Heart rate variability in the 0.15-0.4 Hz band
5	LF/HF ratio heart rate variability	LFHF	Ratio of the low and high frequency of heart rate variability
6	Skin conductance level	SCL	Mean level of skin conductance
7	Skin conductance response rate	SCRR	Mean number of skin conductance responses per second
8	Skin conductance second difference power	SCdiff2	Signal power in the second difference of the skin conductance signal
9	Ohmic perturbation duration skin conductance	OPD	Relative time of responsiveness of the skin conductance signal
10	Respiration frequency	RespFreq	Mean respiration frequency
11	RMS of the respiration signal	RMSResp	Root mean square value of the respiration signal for estimating tidal volume changes
12	RMS of the EMG signal	RMSEMG	Normalized root mean square value as percentage of the EMG reference contraction
13	Static load	Static	10th percentile of rank ordered EMG RMS values
14	Median load	Median	50th percentile of rank ordered EMG RMS values
15	Peak load	Peak	90th percentile of rank ordered EMG RMS values
16	Gaps/min	Gaprate	Average number of EMG gaps per minute
17	Relative time with gaps	Gaptime	Percentage of time in which EMG gaps occurred
18	Mean EMG frequency	MNF	Mean frequency of the magnitude of the EMG frequency spectrum
19	Median EMG frequency	MDF	Frequency at which the surface on the left side equals that of the right side of the magnitude of the EMG frequency spectrum

with a piezoelectric film sensor from SleepSense. A wireless hand sensor was used for measuring skin conductance (SC) by applying a constant voltage of 0.5V DC across the palm of the hand and measuring the change in current. The wireless sensor nodes used for this study are based on the body area network platform developed within imec. Details can be found in [16]. Electromyography (EMG) signals were measured bipolarly from the upper trapezius muscles of both shoulders with commercial gel electrodes. Details on the procedure of EMG recording can be found in [15]. ECG and respiration were recorded at a sampling frequency of 250 Hz. The SC was recorded at 100 Hz. The EMG was recorded at 1000 Hz.

C. Feature calculation

An overview of the features that were calculated from the measured signals can be found in Table I. LF, HF, and LFHF were calculated after applying a Hanning window on the interpolated heart rate signal. For details on the calculation of SCdiff2 and OPD, see [13]. RespFreq was determined as the main frequency component of the power spectral density of the respiration signal. Details on the calculation of the EMG related features can be found in [15].

The features were calculated with a sliding window of 120 seconds that moved over the signals. The length of the window was equal to the length of the shortest condition: the two-minute rest condition. For most of the features, the window moved in steps of 1 second. However, for HR and SDNN the window moved from one beat to the next beat, instead of in steps of 1 second. For every position of the window, all feature values were calculated for that particular time window. The obtained feature values were normalized by calculating the z-score as in (1), with x being the original feature value, z the normalized value, and μ and σ the mean and standard deviation of the feature values, respectively.

$$z = \frac{(x - \mu)}{\sigma} \qquad (1)$$

The 20 second reference contraction was excluded from calculation of μ and σ, because it is a physical exercise and probably influenced the physiological signals in a non psychological way.

The feature values of the different conditions were then determined from the normalized values. For rest conditions, the values were taken that were calculated when the 120 second window was exactly over the rest condition. For the stress values, the values were used that were calculated when the 120 second window was over the last two minutes of the stress condition.

D. Analysis methods

From the 19 features that were extracted from the signals, it was expected that some would show high correlations, since some are based on the same physiological processes. Therefore, a selection was made based on these correlations

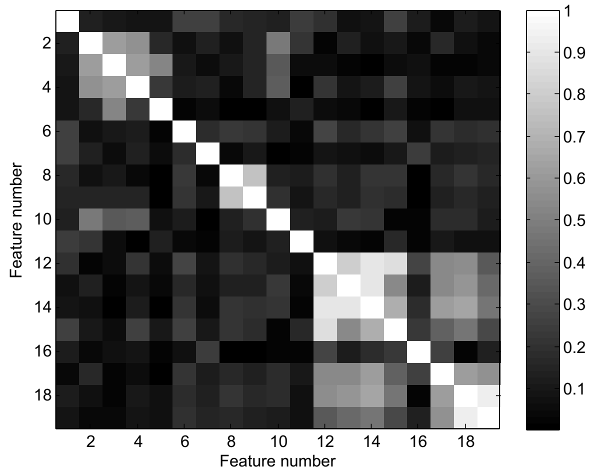

Fig. 1. Absolute correlation values among the features. The feature numbers correspond to the numbers in Table I.

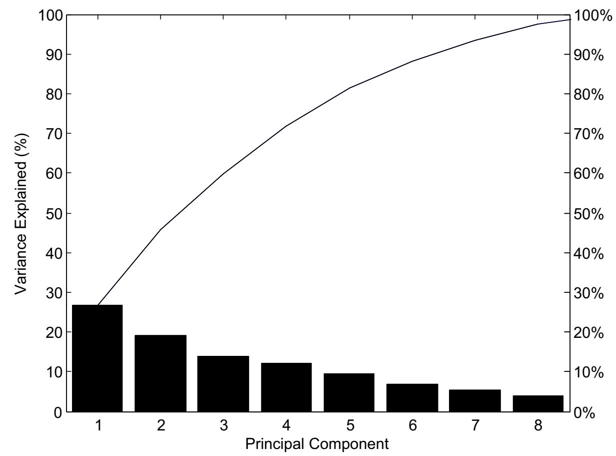

Fig. 2. Percentages of variance explained by the different PCs

and prior knowledge to reduce the number of features and construct a non redundant feature set for further analysis.

Next, principal component analysis (PCA) was applied to the feature subset to remove the last correlations between features. This is important for the classification to be done later. In general, for classification it is important to reduce the number of variables of the problem as much as possible to get a density of data points as high as possible in the multi-dimensional feature space.

The number of principal components (PCs) for the next step of analysis was chosen such that at least 90% of the variance was explained. These PCs were used to classify the three stress conditions and the rest conditions following each of the stress conditions.

A 5-fold cross-validation was performed five times to evaluate the classification performance. Four different classifiers were used to investigate the differences in performance: a Linear Bayes Normal Classifier, a Quadratic Bayes Normal Classifier, a K-Nearest Neighbor Classifier, and a Fisher's Least Square Linear Classifier.

III. RESULTS

Some subjects were excluded from analysis because of poor signal quality (1 poor respiration signal, 5 poor SC signals, 5 poor EMG signals, 9 subjects excluded in total), incomplete data due to failing sensor nodes (2 subjects), or distractions due to other people being present in the room during the experiment (1 subject). Excluding these subjects left a database with 18 complete recordings of good quality.

The absolute correlation values among the normalized 19 features are shown in Fig. 1. For some features, high correlations were found; for example, SDNN, LF, and HF correlate strongly as do RMSEMG, Static, Median, and Peak.

Based on these correlations and prior knowledge, a subset of 9 features was chosen for further analysis. This subset included:

- **HR** Chosen because it did not correlate to any other feature and has shown to react to stress in other studies (for example [17]).
- **SDNN** Chosen from the correlating features SDNN, LF, HF, LFHF. SDNN gave promising results before [13].
- **SCL** Did not correlate with other features, so this feature was included on itself.
- **SCRR** Like SCL, SCRR did not correlate with other features, but can contribute to results.
- **SCdiff2** OPD and SCdiff2 correlated strongly. SCdiff2 was chosen arbitrarily.
- **RespFreq** Did not correlate with other features, but might react to stress.
- **Peak** Chosen from RMS, Static, Median, and Peak. Peak load gave the best result from this subset in an earlier analysis on EMG signals [15].
- **Gaptime** This feature also showed promising results in an earlier analysis [15].
- **MNF** From MNF and MDF, MNF gave a slightly better result in an earlier analysis [15].

Next, PCA was performed with these 9 features. The percentage of explained variance by the different PCs is shown in Fig. 2. The 1st to the 7th PC were selected for further analysis as together they explain 93.6% variance.

These seven PCs were input for classification between three rest and three stress conditions, treated as two classes. The guessing rate for this classification problem was thus 0.5. The mean error rate and its standard deviation give information about how well the classification succeeded and how constant the error rate is. The resulting error rates and standard deviations can be found in Table II.

IV. DISCUSSION

The goal of this study was to find the physiological signals and features that showed the most distinct reaction to mental stress in the conditions of our protocol. A selection of 9 features was made that gave a good representation of the physiological reaction to mental stress. These 9 features were reduced to 7 PCs that showed promising results in classifying mental stress from periods without stress.

TABLE II
ERROR RATES AND STANDARD DEVIATIONS OF CLASSIFICATION OF REST AND STRESS CONDITIONS WITH DIFFERENT CLASSIFIERS

Classifier	Error rate	Standard deviation
Linear Bayes Normal	0.2167	0.0250
Quadratic Bayes Normal	0.2222	0.0207
K-Nearest Neighbor	0.2370	0.0168
Fisher's Least Square	0.2074	0.0140

The results from different classifiers varied slightly, but they were all in the same order of magnitude. The error rates were just above 0.2 for all four classifiers. The similarity of results means that the dataset is consistent and it makes a strong case for the result being valid. The result implies that a good subset of features was chosen. Also, good classification results indicate a possibility for classifying individual cases, which is needed for a personalized stress detection system.

Compared to other papers on classification of mental stress situations, our results were somewhat worse. Classification accuracies of other studies range from 80% [12] to 97.4% [11]. However, other studies only used one type of stressor in their protocols and therefore they were able to tune their algorithms to detecting the single type of mental stress induced by that stressor. In our study, three different stressors were used that were all included in the analysis. Still, an accuracy of almost 80% was achieved in differentiating these different mental stress conditions from rest conditions.

Something that can be improved in our analysis method is the way the feature values were normalized. The feature values calculated over the entire length of the protocol were used. By using this method of normalization, not only the baseline value, but also the reactivity from rest to stress conditions was normalized. However, to apply this method, a long recording time is needed. It is preferable for future use to have only a short time of baseline measurement that is sufficient for normalizing all future recordings. Investigations are needed to find a short, but representative, protocol that can be used for calibration of the physiological features to be calculated in the remainder of the recording. Because the reactivity needs to be normalized as well, some type of standard stressor could be included in this calibration protocol. This must be a stressor that is easy to apply and also triggers reactivity in a majority of people.

Another limitation of our study is the controlled environment in which it was performed. In daily life conditions, physiological signals are influenced by other factors than stress, for example physical activity. For application of a stress detection system in daily life, one must be able to distinguish between changes in physiology caused by psychological factors and changes caused by other factors.

V. CONCLUSIONS

A subset of 9 physiological features was found that can be used for mental stress detection. The features were extracted from ECG, respiration, SC, and EMG signals. PCA indicated that the feature subset could be expressed as 7 PCs. These PCs were used for classification of cases into rest or stress conditions. A classification accuracy of almost 80% was found. This promising result indicates that this feature subset can be used for stress detection in the future. The high classification accuracy also indicates that the features are suitable for individual stress detection.

REFERENCES

[1] European Commission, *Health and safety at work in Europe - A statistical portrait*. Luxembourg: Publications Office of the European Union, 2010.

[2] A. Steptoe, "Invited review: The links between stress and illness," *Journal of Psychosomatic Research*, vol. 35, no. 6, pp. 633–644, 1991.

[3] C. Schubert, M. Lambertz, R. Nelesen, W. Bardwell, J.-B. Choi, and J. Dimsdale, "Effects of stress on heart rate complexity–a comparison between short-term and chronic stress," *Biological Psychology*, vol. 80, no. 3, pp. 325–332, 2009.

[4] C. Tsigos and G. P. Chrousos, "Hypothalamic-pituitary-adrenal axis, neuroendocrine factors and stress," *Journal of Psychosomatic Research*, vol. 53, pp. 865–871, 2002.

[5] U. Lundberg, R. Kadefors, B. Melin, G. Palmerud, P. Hassmen, M. Engstrom, and I. E. Dohns, "Psychophysiological stress and EMG activity of the trapezius muscle," *International Journal of Behavioral Medicine*, vol. 1, pp. 354–370, 1994.

[6] N. Hjortskov, D. Rissén, A. Blangsted, N. Fallentin, U. Lundberg, and K. Søgaard, "The effect of mental stress on heart rate variability and blood pressure during computer work," *European Journal of Applied Physiology*, vol. 92, no. 1, pp. 84–89, June 2004.

[7] M. L. Peters, G. L. R. Godaert, R. E. Ballieux, M. van Vliet, J. J. Willemsen, F. C. G. J. Sweep, and C. J. Heijnen, "Cardiovascular and endocrine responses to experimental stress: Effects of mental effort and controllability," *Psychoneuroendocrinology*, vol. 23, no. 1, pp. 1–17, Jan. 1998.

[8] J. Zhai and A. Barreto, "Stress detection in computer users based on digital signal processing of noninvasive physiological variables." in *IEEE Engineering in Medicine and Biology Society*, vol. 1, 2006, pp. 1355–1358.

[9] J. Healey and R. Picard, "Smartcar: detecting driver stress," in *Proc. 15th Int Pattern Recognition Conf*, vol. 4, 2000, pp. 218–221.

[10] M.-H. Lee, G. Yang, H.-K. Lee, and S. Bang, "Development stress monitoring system based on personal digital assistant (pda)." *Conf Proc IEEE Eng Med Biol Soc*, vol. 4, pp. 2364–2367, 2004.

[11] J. Healey and R. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156–166, June 2005.

[12] J. Zhai, A. B. Barreto, C. Chin, and C. Li, "Realization of stress detection using psychophysiological signals for improvement of human-computer interactions," in *Proc. IEEE SoutheastCon*, 2005, pp. 415–420.

[13] B. Grundlehner, L. Brown, J. Penders, and B. Gyselinckx, "The design and analysis of a real-time, continuous arousal monitor," in *Proc. Sixth Int. Workshop Wearable and Implantable Body Sensor Networks BSN 2009*, 2009, pp. 156–161.

[14] C. Setz, B. Arnrich, J. Schumm, R. L. Marca, G. Trster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable EDA device." *IEEE Trans. Inform. Technol. Biomed.*, vol. 14, no. 2, pp. 410–417, Mar 2010.

[15] J. Wijsman, B. Grundlehner, J. Penders, and H. Hermens, "Trapezius muscle EMG as predictor of mental stress," in *Wireless Health 2010*, ser. WH '10. New York, NY, USA: ACM, 2010, pp. 155–163.

[16] L. Brown, B. Grundlehner, J. van de Molengraft, J. Penders, and B. Gyselinckx, "Body area network for monitoring autonomic nervous system responses," in *Pervasive Computing Technologies for Healthcare*, 2009.

[17] F. W. Kok, H. G. Westenberg, J. H. Thijssen, and J. M. van Ree, "Endocrine and cardiovascular responses to a series of graded physical and psychological stress stimuli in healthy volunteers." *European Neuropsychopharmacology*, vol. 5, no. 4, pp. 515–522, Dec 1995.