Automatic human stress detection based on webcam photoplethysmographic signals — Source link ⧉

Choubeila Maaoui, Frédéric Bousefsaf, Alain Pruski

Institutions: University of Lorraine

Related papers:

- Remote plethysmographic imaging using ambient light

- Human movement intentions based on EEG using brain computer interfaces

- Application of non-linear SVM and LDA in sEMG gesture recognition

- Mental stress detection using physiological signals based on soft computing techniques

- Novel approach for stress recognition using EEG signal by SVM classifier

Automatic human stress detection based on webcam photoplethysmographic signals

Choubeila Maaoui, Frédéric Bousefsaf, Alain Pruski

Automatic human stress detection based on webcam photoplethysmographic signals

Choubeila Maaoui, Frédéric Bousefsaf, Alain Pruski

Automatic human stress detection based on webcam photoplethysmographic signals

Choubeila MAAOUI[1], Frederic BOUSEFSAF [1] and Alain PRUSKI[1]

[1]Laboratoire de Conception, Optimisation et Modélisation des Systèmes (LCOMS)
University of Lorraine, Metz, France.

E-mail: choubeila.maaoui@univ-lorraine.fr

Abstract   One of the goals of affective computing field is to provide to computers the ability to recognize automatically the affective state of the user in order to have more intuitive human-machine communication. This paper aims to detect automatically stress user when he is interacting with computer. The developed system is based on instantaneous pulse rate (PR) signal extracted from imaging photoplethysmography (PPG). Seven features from time and frequency domain are extracted from PR signal and processed by learning pattern recognition systems. Two methods based on SVM (Support Vector Machine) and LDA (Linear Discriminant Analysis) are used and compared to classify the user's emotional state. A computer application based on "Stroop color word Test" is developed to elicit emotional stress in the subject. The proposed method can achieve the overall average classification accuracy of 94.42% and 91.10% with SVM and LDA respectively. Current results indicate that our approach is effective for stress classification.

Keywords Stress detection, Pulse Rate, Non-contact sensor, Photoplethysmography, Support Vector Machines (SVM) and Linear Discriminant Analysis (LDA).

1.   Introduction

Stress is a physiological response to a situation perceived as threatening. It is also defined as the "Fight or Flight" response, originally described by the physiologist W. Cannon [1].

Currently stress is one of the common factors involved in several medical diseases. Chronic stress has become a serious problem affecting different life situations and carrying a wide range of health-related diseases, including cardiovascular disease, cerebrovascular disease, diabetes and immune deficiencies [2]. In health-care and well-being research the problem of stress management has been receiving an increasing attention [3, 4, 5]. Continuous monitoring of an individual's stress level is essential for understanding and managing personal stress. Several researchers are attempting to identify and measure stress through different modalities (facials expressions, voice, gestures and physiological markers). However, it has been proved that the physiological signal approach is more efficient than other methods [6]. This method is non-invasive and stress impact of the body can be easily analyzed using physiological signals. A number of physiological markers of stress have been identified, including electrodermal activity, heart rate, various indices of heart rate variability (HRV), blood pressure, temperature, muscle tension, and respiration [7]. Among the several physiological signals, the HRV signal is frequently used as an efficient measure to identify the stress.

In this paper, we propose a new operating technique of instantaneous PR signal for stress detection. This technique assesses the instantaneous PR remotely by observing photoplethysmography on human faces with a low-cost webcam which is considered as a non-contact and non-invasive sensor.

The remainder of this paper is organized as follows: first, we present the related works on stress detection based on HR signal. In section 3, we describe methodology to extract instantaneous PR signal using webcam. Next, we specify the features extraction method to clas-

sify our emotional categories. Then, in section 5, we expose our experimental results. Finally, conclusion and future works are presented in the last section.

2.   State of the art

Stress detection has been considered from different points of view and approaches. The work presented by Andren and Funk [8] provides a system that is able to compute the stress level of an individual by the manner and rhythm in which a person types characters on a keyboard or keypad. Furthermore, Dinges et al. [9], Metaxas et al. [10] provide a study of stress detection based on facial recognition. Also, stress detection of humans has been a well-researched topic in the area of speech signal processing [11] [12].   Gaggioli et al. [13] developed Decision Support System for automatic classification of stress levels using biosensor data, like electrocardiogram (ECG), breathing rate, electroencephalogram (EEG) and behavioral data.

However, in this paper, the proposed method focuses on stress detection based only on pulse rate signal. Considering this, we present only works related to this topic. PR signal has been widely studied and analyzed for stress evaluation. Jovanov et al. [14] presented a stress monitoring system based on a distributed wireless architecture implemented on intelligent sensors. HR was recorded along different locations in individual body by means of sensors located beneath clothes. Kharthikeyan et al. [15] developed a method for human stress detection through short-term electrocardiogram and heart rate variability signals. Giakoumis et al [16] have used skin conductance and electrocardiogram signals for stress detection. They proposed novel subject-dependent bio-signal features on suppressing between-subject variability. Sloten et al. [17] concluded that the heart rate is significantly higher with a mental task than in the control. Melillo et al. [18] analyzed HRV signal extracted from ECG and extracted nonlinear measures which vary in subject under stress due to university examination.

Heart rate variability contains abundant information on the status of the person and can be derived from ECG or PPG. However, the stress evaluation using the ECG seems to be unsuitable as a long-time monitoring method since several electrodes have to be directly attached to a human body to acquire the ECG data. These techniques can cause infections, as in the case of very premature infants with fragile skin. For these particular cases, non-contact and non-invasive techniques are principally advantageous.

3.   Methodology

The autonomic nervous system (ANS) of human beings comprises three separate sub-systems: the sympathetic, parasympathetic and enteric divisions. The parasympathetic subsystem predominates under resting conditions and causes HR and blood pressure to drop. The sympathetic subsystem, which is active during periods of exertion, stress, or emergency, causes increases in HR, cardiac output and blood flow in the muscles.
In the following section, we describe our developed approach to extract the cardiac pulse wave based on photoplethysmography and using a simple webcam. Further details can be found in Bousefsaf [19].

3.1   Extraction of cardiac pulse from webcam frames

The first step of our approach is based on automatic face detection using a cascade of boosted classifier on each frame with OpenCV 2.1 library [20]. The Pan, Tilt and Zoom parameters are computed and regulated in order to track the face over time [Fig. 1 (a)]. For extraction of PPG information, skin segmentation is applied on the original frame to isolate skin pixels that contain the PPG signal [Fig. 1 (b)]. The CIE L*u*v* color space was chosen in the present study [Fig. 1 (c)] to reduce fluctuations due to light variation or head movements.

Only the u* component, which represents the important PPG fluctuations, is used [Fig. 1 (d)] to compute a spatial averaging operation. A set of n frames gives a raw signal of n points [Fig. 1 (f)].

In order to remove trends and high frequency noise of the raw signal [Fig. 2 (a)], we have developed a Continuous Wavelet Transform filtering [Fig. 2 (b)] in the 0.65–3 Hz frequency band. The instantaneous PR trace is computed from peaks which are detected using a custom algorithm [Fig. 2 (c)]. The details of the filtering method have already been published in [21]. Briefly, a mask is applied around the maximum wavelet energy to remove all

non-relevant coefficients that are present in the cardiac frequency band but which not belong to the cardiovascular pulse wave.
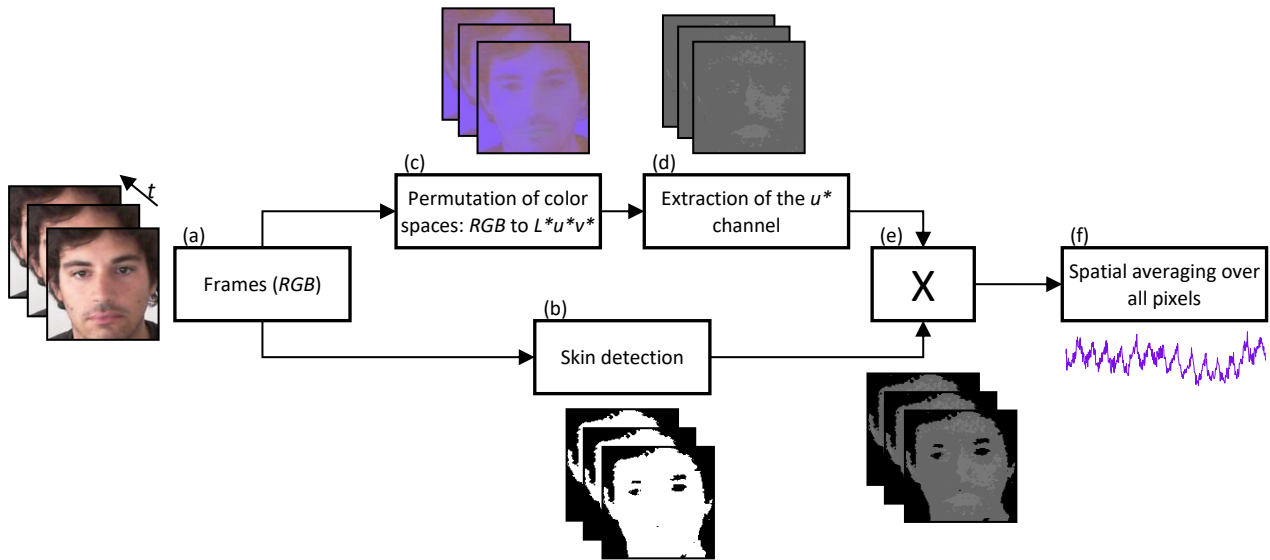

Fig.1. Processing algorithm overview. (a) Pan, Tilt and zoom parameters are computed to zoom and track the face on the input frame. (b) Pixels that contain PPG information are isolated by a skin detection. (c) The RGB color space is converted to the CIE L*u*v* color space. (d) The u* frame is extracted and (e) merged with the skin detection mask by a combinational AND operation. (f) A spatial averaging step is performed to transform a set of frames into a single raw signal.

3.2 *Stress Elicitation*

One of the most challenging points in this research is to obtain accurate physiological signal related to mental stress of a human subject.

The database was collected based experiments from 12 students from the laboratory, with ages from 22 to 27 years. All participants gave their informed consent before the beginning of a session. The protocol used to induce mental stress is the same than the one presented in [21]. We have used a succession of 1 min. relaxation videos and stress sessions based on an interactive and computerized version of the Stroop color word test.

The subject has 3 seconds to click on the correct colored box corresponding to the word printed in the screen center. Some words are printed in a color not denoted by the name (such as the word "pink" displayed in pink) while others are printed on the right color (like the word "green" displayed in blue). Several stressors, like alarm siren, error sound and progressive bar indicating the remaining time, are used in this experience as.
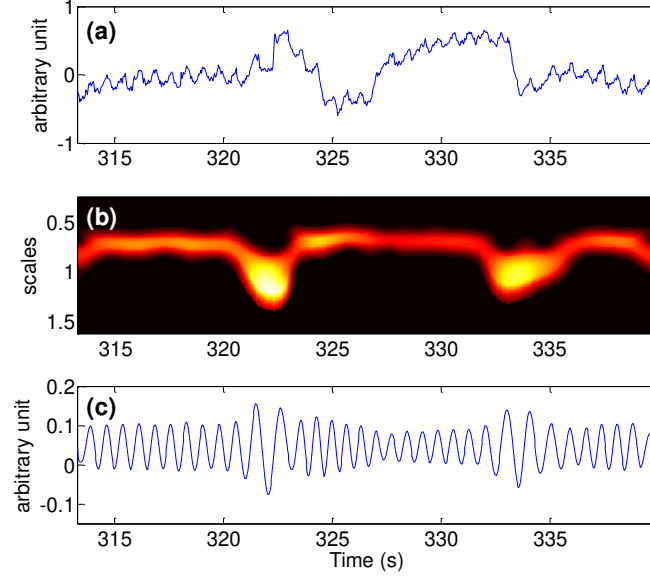
Fig.2. Signal processing. The raw signal (a) is filtered using its wavelet representation (b). This particular plot is a scalogram, representing the percentage energy for each wavelet coefficient. The filtered signal (c) is detrended and denoised by computing the inverse wavelet transform.

4. Data analysis

4.1 *Features selection*

One key research problem in the Affective Computing area is the mapping between affective states and physiological states. One promising solution to this challenge consists in computing multiple features of the physiological signal to indicate the correlation between a data and the internal stress state. In our work a total of 7 features are extracted from the signal.

HR signal analysis method can be categorized into time domain and spectral domain analysis. In our case, time domain features include mean value of HR signal (*meanHR*), standard deviation of the HR (*StdHR*), first derivative of HR (*DerHR*) and Root Mean Square of the Successive Differences (*RMSSD*) given by the following equations:

$$meanHR = \bar{I} = \frac{1}{N-1}\sum_{n=2}^{N} I(n) \tag{1}$$

$$StdHR = \sqrt{\frac{1}{N-1}\sum_{n=2}^{N}\left[I(n) - \bar{I}\right]^2} \tag{2}$$

$$DerHR = \frac{1}{N-1}\sum_{n=1}^{N-1}[I(n+1) - I(n)] \tag{3}$$

$$RMSSD = \sqrt{\frac{1}{N-2}\sum_{n=3}^{N}[I(n) - I(n-1)]^2} \tag{4}$$

Moreover, in the spectral domain we have used low frequency band (LF) [0.04Hz, 0.15Hz], high frequency band (HF) [0.15Hz, 0.35Hz] and the ratio LF/HF which is interpreted as a measure of the relative sympathetic to parasympathetic activity of the autonomic nervous system. The LF power is generated mainly from sympathetic activity and the HF power, in contrast, is derived from vagal or parasympathetic activity. Thus, LF/HF ration represents a good indicator of sympathetic-vagal balance. It is shown that mental stress increases LF activity and decreases HF activity, as a result the LF/HF ratio increases during stressful situa-

tions [22]. All the heart rate series are irregularly time-sampled in nature. A 30 Hz linear interpolation was employed to resample the traces before computing the fast Fourier transform to extract the LF and HF frequency bands.

All used parameters are computed using 30 seconds moving windows that are spaced with a fixed 2 seconds step.

4.2   Stress classification

After all features were extracted, they were provided as input to learning systems which were trained to differentiate the stress state from the relax state. Two popular learning algorithms were employed: Support Vector Machine (SVM) and Linear Discriminant analysis (LDA).

4.2.1   *Support Vector Machine*

Support Vector Machine is a classifier that performs classification by constructing a high-dimensional hyper-plane [23]. The constructed high-dimensional hyper-plane is optimized to separate the testing data into two classes. SVM also allows different types of kernel functions to transform testing data points into a higher dimensional space and make the transformed data easier to be classified. Since SVM has recently become a popular machine learning technique for classification [24], we are interested in investigating its performance with our testing dataset.

4.2.2   *Linear Discriminant Analysis*

LDA aims to find linear combinations of the input features that can provide an adequate separation between two classes. LDA (sometimes called Fisher Discriminant Analysis) projects all the data points into new space, normally of lower dimension, which maximizes the between-class separability while minimizing their within-class variability. Further details about LDA can be found in Krzanowski [25].

5.   Experimental results

Signals from 12 experimental subjects were collected and 2376 features vectors were extracted, since each participant generated data under three non-stress segments and three stress segments [21]. Seven attributes (*meanHR*, *StdHR*, *DerHR*, *RMSSD*, *HF*, *LF* and *LF/HF*) were determined for each data entry.

The prediction performance was evaluated: 1500 remaining samples were used to train the classifier and 876 samples were used as the test samples. The goal was to develop and train a system that predicts the participant's emotional state. SVM and LDA were trained to build independent models, which could be used to identify an unknown emotional state from a set of 7 features. Performance of such systems is commonly summarized in a confusion matrix, which contains information about the actual and the estimated classifications generated by each one of the systems.

5.1 *SVM results*

SVM implementation was done using libSVM [26]. Different kernel functions can be chosen during the SVMs classification, corresponding to the different transformed feature spaces. They play an essential role in the SVMs classification since they determine the feature spaces in which data examples are classified and can directly affect the SVMs classification results and performances. The used SVM parameters are:

- Kernel type: A radial basis function (RBF) kernel and linear kernel;
- Cross-validation: In order to compute the accuracy of each SVM, we perform 10-fold cross-validation.

Table 1 gives the percentage of correctly classified examples for each subject using SVM method for each kernel (RBF and Linear kernels). We can note that the best recognition rate is obtained using RBF kernel for the most and all subjects. The RBF kernel per-

forms better than linear kernel.

Table 2 presents the recognition rate accuracy based on SVM method with RBF kernel and using time features (*meanHR*, *StdHR*, *DerHR*, *RMSSD*), frequency features (*LF*, *HF*, *LF/HF*) and all features separately. We have obtained 87.67%, 82.21% and 94.40% of recognition rate respectively. The best recognition rate is achieved using all information (time and frequency information).

Table 3 and Table 4 give the confusion matrices for the set of all subjects, using RBF and Linear kernel respectively. This matrix shows the effectiveness of classification method with SVM. We can observe that the best recognition rate is always obtained for the corresponding and correct emotional state for all kernels. However, the best rate is obtained using RBF kernel. Thus, we choose to use RBF kernel.

5.2 *LDA results*

We have tested Linear Discriminant Analysis for stress detection in order to reduce the dimension of the features. Table 5 shows the percentage of correctly classified examples. We have obtained 91.10% of rate recognition. We can observe that this rate is lower than that obtained using SVM based RBF kernel

Table 6 gives the confusion matrix for the set of all subjects, using LDA method. The best recognition rate is always obtained for the corresponding and correct emotional state. Also, we note that the obtained accuracy rates in this table are lower than those obtained using SVM method with RBF kernel (Table 3). SVM methods with RBF kernel performs better than LDA method.

Figures 3, 4 and 5 present the results of the features signals separation. These features were projected down to a two dimensional space (Fisher features). Fisher transformation is often used to get a good representation of multidimensional class data in a two dimensional space. As expected, we can observe significant variations in the positions of data points for each state (stress, relax). The data are separated into well-defined clusters.

Obviously, merging the features of all subjects does not refine the information related to target state (Figure 6). There is confusion between relax and stress data, which explains the obtained LDA rate as 91.10% for all subjects case.

Table 1. Stress detection using SVM classification (RBF and linear kernel) for 12 subjects.

	SVM_RBF (%)	SVM_Linear (%)
Subject1	97.47	100.00
Subject2	92.93	80.81
Subject3	96.46	88.38
Subject4	87.88	93.94
Subject5	95.96	88.38
Subject6	96.46	97.98
Subject7	83.33	87.37
Subject8	92.43	85.35
Subject9	98.99	97.47
Subject10	96.97	87.88
Subject11	87.88	73.82
Subject12	99.50	100.00
All subjects	94.40	91.32

Table 2. Rate accuracy of different combination of parameters using SVM with RBF kernel for all subjects

parameters	Rate accuracy (%)
Time domain parameters	87.67
Frequency domain parameters	82.21
All parameters	**94.40**

Table 3. Confusion matrix for stress detection using RBF Kernel

	stress	relax
stress	**93.35%**	6.65%
relax	4.64%	**95.36%**

Table 4. Confusion matrix for stress detection using Linear Kernel

	stress	relax
stress	**88.97%**	11.03%
relax	8.84%	**91.16%**

Table 5. Stress detection using SVM and LDA methods

	LDA	SVM-RBF
All subjects	91.10%	94.40%

Table 6. Confusion matrix for stress detection using LDA

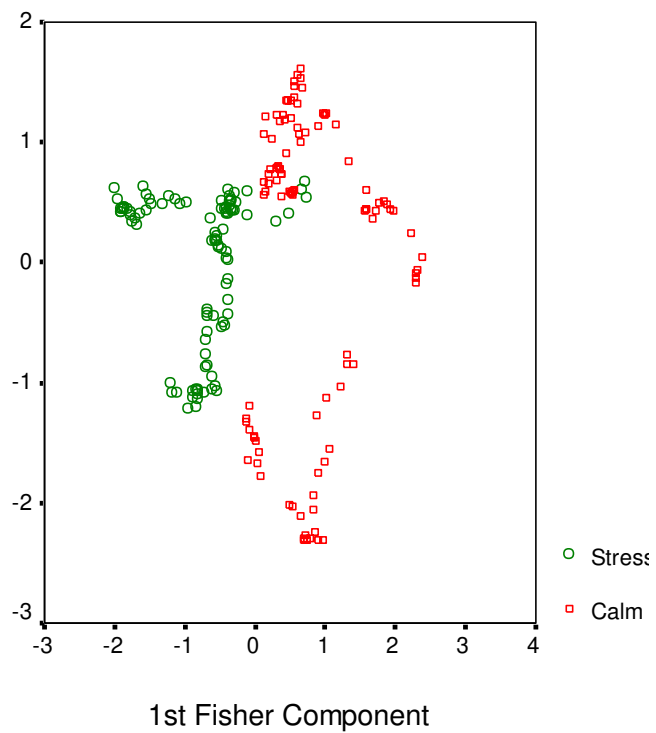|	stress	relax
stress	**86.6%**	13.4%
relax	6.4%	**93.5%**

Fig.3. Data projection on to the two first Fisher features for subject1.
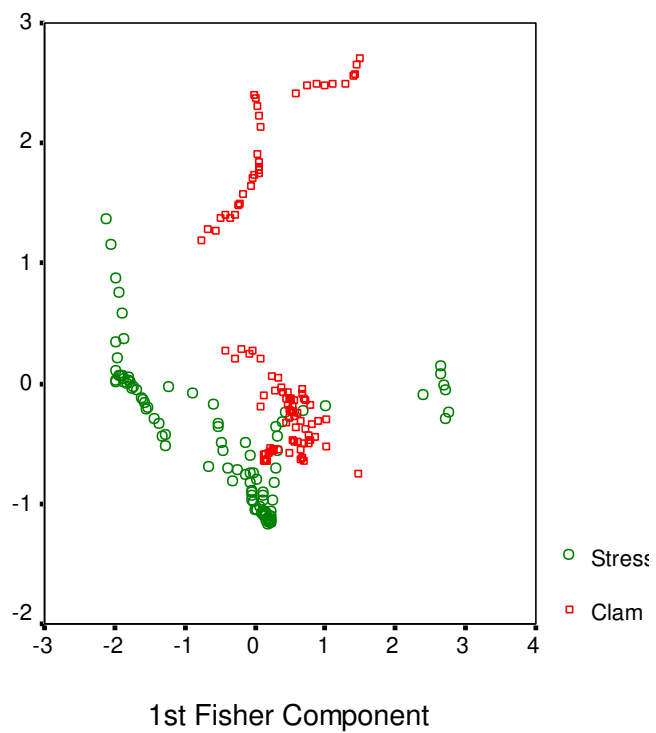

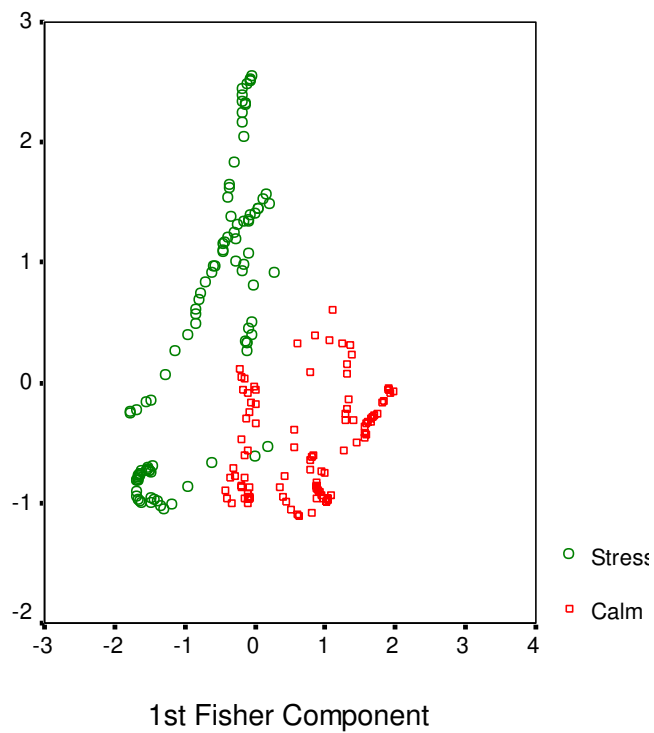Fig.4. Data projection on to the two first Fisher features for subject2.

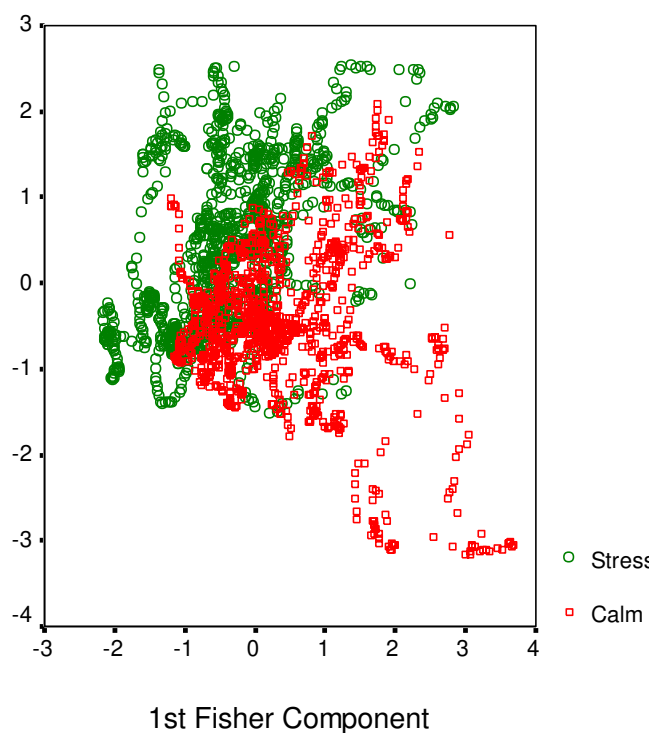Fig.5. Data projection on to the two first Fisher features for subject3.

Fig.6. Data projection on to the two first Fisher features for all subjects.

6. Discussion

The HRV is a physiological measurement used in several domains, especially in affective compu-

ting. In the present study, we have demonstrated that the instantaneous pulse rate signal, extracted from human faces in frames that were recorded using a low-cost webcam, can be used efficiently to detect and classify the emotional state of the user.

In the previous work [21], we have shown a strong correlation between the webcam PPG signal and contact skin conductance level traces for stress evaluation. However, with the proposed approach, we confirm the feasibility and effectiveness of the system to detect and classify automatically the emotional states using PPG signals derived from webcam and processed by machine learning algorithms.

For classification, we have compared two methods: SVM and LDA. The obtained results show that SVM based on a RBF kernel performs better than LDA. The used data are not linearly separable.

SVM method classifies with less accuracy and sensitivity when time domain and frequency domain features are processed separately. However, when all features are combined, the classification rate can achieve score that exceeds an accuracy of 94 per cent. The combined features perform much better than separate features derived from time (87%) and frequency (82%) domains.

The system we proposed in this study was able to efficiently classify two different affective states, but can nevertheless not achieve a 100% of recognition rate. The proposed method depends on several parameters:

- The inductor: we have noted that the used inductor did not have the same effect on every subject. Other inductors can be tested in order to better induce stress or even emotions, such as realistic virtual environment [27].
- Selected features: this step plays an important role in pattern recognition systems. In order to translate the recorded physiological signals to stress or affective states, other features can be computed. "Features selection methods" can be employed to reduce dimensionality of data and choose only the features that give a decent recognition contribution.
- Classification method: alternative classification methods can be tested, like for example: neural networks, naïve Bayes or decision tree to only retain the method that gives the best recognition rate.

7. Conclusion

The results of the current study show that pulse rate analysis derived from webcam based photoplethysmography could be effective in automatic stress detection. The proposed approach, based on non-contact assessment of the emotional state of a computer user is performed with two independent classifiers; the support vector machine and linear discriminant analysis. SVM method achieved the highest level (94.40%) of success in identifying user stress states, using 7 features extracted from the PR signal into time and spectral domains. Therefore, this work confirms the potential of integrated digital signal processing and machine learning algorithms to differentiate key emotional states from PPG signal derived from webcam.

As future work, we intend to incorporate other means of emotional state recognition using non-contact sensor (webcam) such as pupil diameter and facial expression in order to improve our system.

References

1. W. Cannon, "Bodily Change in Pain, Hunger, Fear and Rage: An Account of Recent Research into the Function of Emotional Excitement". *Appleton, New York, USA*, 1915.

2. J. Cacioppo, L. Tassinary, G. Berntson, *Handbook of Psychophysiology*. Cambridge University Press, New York, 2000.

3. Y. Ayzenberg, J. H. Rivera, R. Picard, "Feel: frequent EDA and event logging – a mobile social interaction stress monitoring system". *In CHI'12 Extended Abstracts on Human Factors in Computing Systems,* New York, NY, USA, 2012, pp. 2357–2362.

4. J. Bakker, L. Holenderski, R. Kocielnik, M. Pechenizkiy, N. Sidorova, "Stess@work: From measuring stress to its understanding, prediction and handling with personalized coaching". *In Proceedings of ACM SIGHIT Int. Health Informatics Symposium*, *ACM Press*, 2012, pp. 673–678.

5. Y. Deng, D. F. Hsu, Z. Wu, C-H. Chu, "Feature selection and combination for stress identification using correlation and diversity". *In 12th Int. Symp. on Pervasive Systems, Algorithms and Networks (ISPAN)*, 2012, pp. 37–43.

6. V. Malhotra, M. K. Patil, *Mental Stress Assessment of ECG Signal using Statistical Analysis of Bio-Orthogonal Wavelet Coefficients*. International Journal of Science and Research, 2013, 2(12): 430-434.

7. T. Vrijkotte, L. J. van Doornen, E J. de Geus, "Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability". *Hypertension*, 2000, 35(4):880 –886.

8. J. Andren, P. Funk, "A case-based approach using behavioral biometrics to determine a user's stress level". *In Workshop proceedings of the 6th International Conference on Case Based Reasoning*, 2005, pp. 9–17.

9. D. F. Dinges, S. Venkataraman, E. L. McGlinchey, D. N. Metaxas, "Monitoring of facial stress during space flight: Optical computer recognition combining discriminative and generative methods". *Acta Astronautica*, 2007, 60 (4–7): 341–350.

10. D. Metaxas, S. Venkataraman, C. Vogler "Image Based Stress Recognition Using a Model-Based Dynamic Face Tracking System". *Computational Science-ICCS, 2004*, pp. 813-821.

11. H. J. M. Steeneken, J. H. L. Hansen, "Speech under stress conditions: Overview of the effect on speech production and on system performance". *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1999, pp. 2079-2082.

12. A. Mesleh, D. Skopin, S. Baglikov, A. Quteishat, "Heart Rate Extraction from Vowel Speech Signals". Journal of Computer Science and Technology, 2012, 27 (6), pp. 243-1251.

13. A. Gaggioli, P. Cipresso, S. Serino, G. Pioggia, G. Tartarisco, G. Baldus, D. Corda, M. Ferro, N. Carbonaro, A. Tognetti, D. De Rossi, D. Giakoumis, D. Tzovaras, A. Riera, G. Riva, "A Decision Support System for Real-Time Stress Detection During Virtual Reality Exposure". *Studies in Health Technology and Informatics, IOS Press*, 2014, 196: 114-120.

14. E. Jovanov, A. O'Donnell Lords, D. Raskovic, P. G. Cox, R. Adhami, F. Andrasik, "Stress monitoring using a distributed wireless intelligent sensor system". *IEEE Eng. Med. Biol. Mag.*, 2003, 22(3), pp.49-55.

15. P. Karthikeyan, M. Murugappan, S. Yaacob, "Detection of human stress using short-term ECG and HRV signals". *Journal of Mechanics in Medicine and Biology,* Vol. 13, No. 02, pp. 1-29, 2013.

16. D. Giakoumis, D. Tzovaras, G. Hassapis, "Subject-dependent biosignal features for increased accuracy in psychological stress detection". *Int. J. Hum.-Comput. Studies*, 2013,71(4): 425-439.

17. J. V. Sloten, P. Verdonck, M. Nyssen, J. Haueisen, "Influence of mental stress on heart rate and heart rate variability". *International Federation for Medical and Biological Engineering Proceedings*, 2008, pp. 1366-1369.

18. P. Melillo, M. Bracale, L. Pecchia, "Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination". *BioMedical Engineering OnLine*, 2011, 10(96).

19. F. Bousefsaf, C. Maaoui, A. Pruski, "Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate". *Biomedical Signal Processing and Control*, 2013, 8 (6), pp. 568–574.

20. G. Bradski, T. Darrell, I. Essa, J. Malik, P. Perona, S. Sclaroff, C. Tomasi, http://sourceforge.net/projects/opencvlibrary/, 2006.

21. F. Bousefsaf, C. Maaoui, A. Pruski. "Remote detection of mental workload changes using cardiac parameters assessed with a low-cost webcam". *Computers in Biology and Medicine*, 2014, 53, pp. 154–163.

22. E. Minakuchi, E. Ohnishi, J. Ohnishi, S. Sakamoto, M. Hori, M. Motomura, J. Hoshino, K. Murakami, T. Kawaguch, "Evaluation of mental stress by physiological indices derived from finger plethysmography". *Journal of Physiological Anthropology* 2013, 32:17.

23. C. J. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition". *Data Mining and Knowledge Discovery*, 1998, 2:121–167.

24. M. K. Das, S. Ari, "Electrocardiogram beat classification using S-Transform based feature set", *Journal of Mechanics in Medicine and Biology,* Vol. 14, No. 05, 2014.

25. W. J. Krzanowski, "Principles of multivariate analysis: a user's perspective". *Rev edition. Oxford Statistical Science Series,* 2000.

26. C. C. Chang, C. J. Lin, "*LIBSVM :* a library for support vector machines". *ACM Transactions on Intelligent Systems and Technology (TIST*), 2001, 2 (3), 27.

27. D. VÄSTFJÄLL, "The Subjective Sense of Presence, Emotion Recognition, and Experienced Emotions in Auditory Virtual Environments". Cyberpsychology & Behavio, Vol 6, N° 2, 2003.