Remote assessment of the Heart Rate Variability to detect mental stress

Frédéric Bousefsaf, Choubeila Maaoui, Alain Pruski

Laboratoire de Conception, Optimisation et Modélisation des Systèmes (LCOMS)

7 rue Marconi, 57070 Metz, France

Abstract—In the present paper, we introduce a new framework for detecting workload changes using video frames obtained from a low-cost webcam. The measurements are performed on human faces and the proposed algorithms were developed to be motion-tolerant. An interactive Stroop color word test is employed to induce stress on a set of twelve participants. We record the skin conductance and compare these responses to the stress curve assessed by a webcam-derived heart rate variability analysis. The results offer further support for the applicability of stress detection by remote and low-cost means, providing an alternative to conventional contact techniques.

Keywords—heart rate variability; non-contact; emotions; mental stress detection; Virtual Reality; webcam

I. INTRODUCTION

Recognizing an emotion by its physiological signature is a field of research that presents a particular interest on the last ten years. Understanding emotions can be useful, particularly during virtual reality (VR) therapies, where emotions are feedbacks that regulate the virtual environment level and intensity [1]. These physiology-driven virtual scenes are particularly employed to treat patients with anxiety disorders and phobia, like soldiers and war veterans [2]. Physiological parameters like Heart Rate (HR) and Heart Rate Variability (HRV) are reliable inputs to emotion recognition [3]. However, contact sensors can be limited in some scopes of application where a specialist must install and monitor them. When dealing with motor rehabilitation or serious games, contact sensors can disturb the interaction and may be intrusive to the privacy of the patient.

Non-contact measurements of physiological parameters can be achieved using thermal infrared imaging, a technology employed by Pavlidis et al. to collect physiological data on human faces [4]. Similarly, Doppler radars are non-contact sensors that were used to detect heartbeats [5] and respiration signals [6]. More recently, digital cameras and webcams were employed on the face to detect the blood volume pulse [7] and compute HR and breathing rate. The principle, based on PhotoPlethysmoGraphy (PPG) consists in observing light variations on the skin to recover the cardiovascular pulse wave. Sun et al. [8] have compared performances between a low-cost webcam and a high-sensitivity camera to assess HR and pulse rate variability. They conclude that the functional characteristics of a 30 fps webcam are comparable to those of a 200 fps camera when interpolating signals to improve the time domain resolution [8]. The HRV is a parameter used in affective computing and psychophysiology to give an index of the Autonomic Nervous System (ANS) activity in order to detect workload changes in real time [9]. Its spectral analysis can provide the sympathovagal balance, a ratio that reflects reciprocal changes of sympathetic and vagal outflows [10]. The HRV tends to be rhythmic and ordered in positive emotional states and follows the respiration by a phenomenon called Respiratory Sinus Arrhythmia. In contrast, the HRV tends to be chaotic and disordered in states of anger, anxiety or sadness. These rhythmic variations provide a state known as cardiac coherence [11]. Assessment of physiological signals by non-contact technologies is particularly advantageous in applications that need to understand feelings and sentiments of a patient. Emotions are feedbacks that can be employed on virtual reality-based applications, particularly on VR exposure therapy [12]. The real-time monitoring of physiological signals helps the therapist that supervises the exposure session to customize and optimize virtual stimuli and situations. In these psychophysiological experiments, contact sensors may generate a bias by interfering with the user, resulting practically by an erroneous emotional state estimation [4].

Machine learning algorithms can be employed on parameters computed from physiological signals to automatically recognize emotional states [3], [13] and particularly stress [14]. Only the cardiac pulse wave can be used to detect mental stress and workload changes [9], [15], [16]. Multiple modalities [17] are employed to improve the recognition accuracy compared to single modality approaches.

In this paper, we propose a new method that recovers the instantaneous HR signal from video frames of human faces using a low-cost high definition webcam. The HRV is then employed to detect mental stress on a set of 12 participants that were performing an interactive version of the Stroop color test.

II. METHODS

A. Extraction of the cardiac pulse from webcam frames

The overall system is composed with both image and signal processing (see Fig. 1). Working with HD frames leads to increase the number of skin pixels that contains the PPG signal. Given that HD acquisition requires more computational time, Pan, Tilt and Zoom parameters were employed in order to keep a frame rate of 30 fps. The face is automatically detected using a cascade of boosted classifier on each frame with OpenCV 2.1 library. Pan and tilt parameters are set to track the face over time while the zoom parameter is computed to adjust the size of the face on the input frame [Fig. 1 (a)]. Preprocessing operations are applied on the original frame to isolate skin pixels that contain the PPG signal. A skin detection mask [Fig. 1 (b)] is employed to properly collect these pixels. The CIE LUV color space was chosen in the present study [Fig. 1 (c)] to reduce fluctuations due to light variation or head
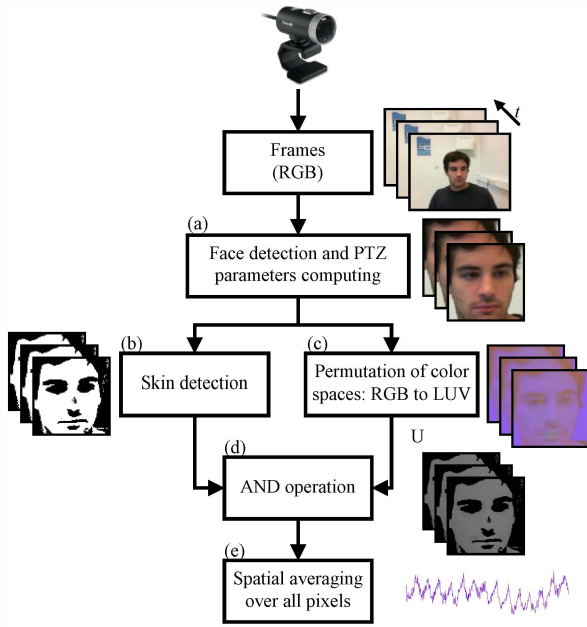
Fig. 1. Processing algorithm overview. (a) Pan, Tilt and Zoom parameters are computed to zoom and track the face. (b) Pixels that contain PPG information are isolated by a skin detection. (c) RGB color space is converted to CIE LUV. (d) The U frame is combined with the skin detection frame by a combinational AND operation. (e) A spatial averaging step is performed to transform a set of frames into a single raw signal.

movements. Oxy- and deoxyhemoglobin have a better absorption coefficient between 540 and 577 nm, corresponding to green/yellow colors wavelength [18]. The PPG information will be more influent when using the U component, which represents a red to green color indicator. A spatial averaging operation is computed using U pixel intensities that exist on the skin detection. A set of n frames gives a raw signal of n points [Fig. 1 (e)]. A Continuous Wavelet Transform filtering [Fig. 2 (b)] is employed to remove trends and high frequency noise of the raw signal [Fig. 2 (a)] in the 0.65–3 Hz frequency band. A custom algorithm was developed to detect peaks [Fig. 2 (c)] and compute the instantaneous HR trace.
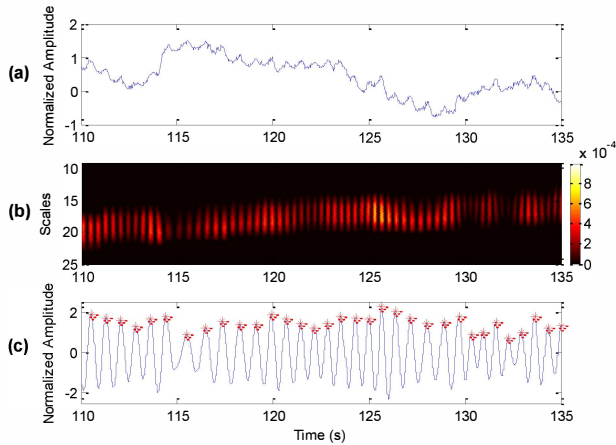

Fig. 2. Signal processing. The raw signal (a) is filtered using its wavelet representation (b). This particular plot is a scalogram, representing the percentage energy for each wavelet coefficient. The filtered signal (c) is detrended and denoised by computing the inverse wavelet transform.

B. An application to mental stress detection

1) Procedure

Twelve students (two females and ten males) from the laboratory participated in this study. All participants gave their informed consent before the beginning of a session. Each experiment lasted five to six minutes. The computer work task has already been applied in various studies and is based on an interactive version of the Stroop color word test [15]. Briefly, the participant has 3 seconds to click on the colored box that corresponds to the word printed in the center of the monitor [Fig. 4 (a)]. Some words are printed in a color not denoted by the name (incongruent, e.g. the word "green" printed in a blue ink) while the others are printed on the right color (congruent, e.g. the word "pink" written in pink).

The participants performed three sessions (see Fig. 3) of the color word test, i.e. a one minute introductory session (IS) to familiarize the user with the virtual interface and two stress sessions (SS). Each session are separated by a one minute relaxation session (RS). In the first SS, the participant has one minute to click on 35 correct boxes. A wrong click decrements the value by one and a loudly error sound is played. A horizontal progress bar is added under the central word, giving the remaining time of the session. Additionally, a vertical progress bar is added to the right of the word, indicating the remaining time to click. The second SS last one minute and is identical to the first SS, except that the positions of all color boxes are randomized on each click. This time, the user must click on 40 correct boxes. A stressful music is played during both SS and an alarm siren is launched the 10 last seconds.
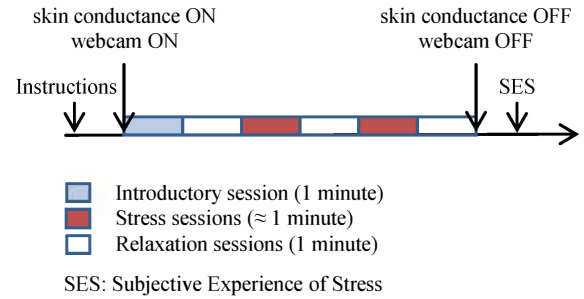

Fig. 3. Protocol of the experiment.

At the end of the session, the participants were asked to report their subjective experiences of stress via a 5-point Likert scale (1 = not at all, 5 = extremely). The following parameters were used: stressed, tensed, exhausted, concentrated, stimulated. They gave two sets of five responses: one set for the two stress sessions [Fig. 4 (a)] and one set for the three relaxation sessions [Fig. 4 (b)]. This rating technique is used to control the correlation between physiological responses and perceived exertion. Finally, a last question was asked to appreciate the effects of the randomized process on participants between the first and the second SS. The electrodermal response (EDR) was concurrently recorded using a skin conductance sensor (SC-Flex/Pro by Thought Technologies Ltd.).

Fig. 4. Screenshots of the interactive application: the Stroop color word test (a) and the first relaxation video (b) that starts right after the introductory session.

2) HRV analysis to detect stress

The processing operations detailed in section II.*A*. were employed to assess the instantaneous HR trace [Fig. 5 (a)]. This irregularly time-sampled signal was interpolated using a 15 Hz cubic spline function. Stress periods generally increase the mean HR. To uncover this trend, a 400-point (27 seconds) moving average was computed on the interpolated HRV signal [Fig. 5 (b)]. The HRV tends to be rhythmic and ordered in RS and chaotic and disordered in SS [11]. Thus, to detect these fluctuations, an empirical third-order derivative was computed on the HRV [Fig. 5 (c)]. Frequencies variations are additionally detected using the HRV high frequency ratio [Fig. 5 (d)]. A custom algorithm, wavelet-based, was developed to compute this sympathovagal index.

These three parameters are summed up and the result is smoothed using a 400-point moving average, giving a single representation of the stress state via a curve (see Fig. 6). Minima and maxima points are detected, corresponding to relax and stress extremities respectively.

III. RESULTS

Stress signals were compared to EDRs to assess the agreement with the camera measurements. A typical example is presented in Fig. 6 where the stress curve generated using the HRV signal is in close agreement with the EDR of the participant. Boxplots are employed to represent the differences on mean values for each session (see Fig. 7).
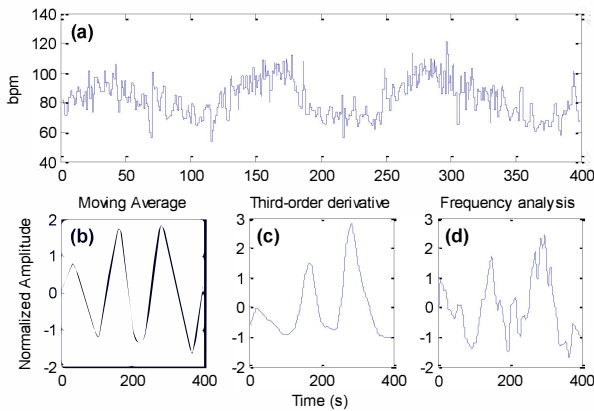

Fig. 5. Results of HRV (a) analysis. A moving average operation (b) is computed on the interpolated signal to detect trend variations. The third-order derivate of the signal (c) is employed to assess rhythmic fluctuations. The wavelet-based frequency analysis (d) gives an index of the ANS.
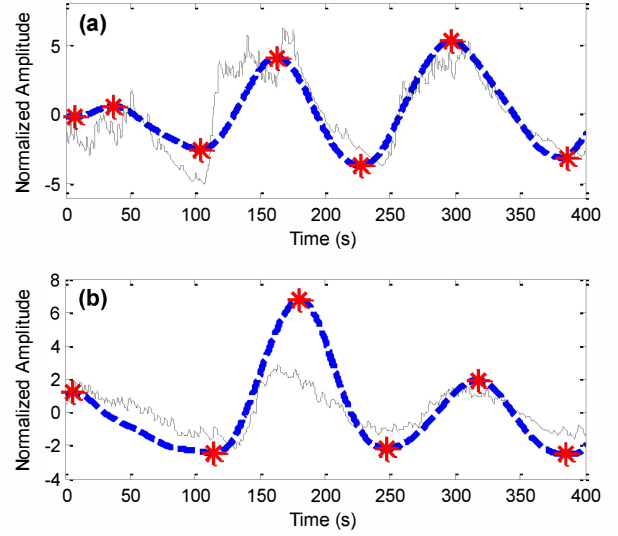

Fig. 6. Results of the stress detection for the participant #1 (a) and #8 (b). The dashed-line plot corresponds to the webcam-derived stress signal and the solid-line plot to the EDR trace.

Significant differences were observed between the RS and the SS on the subjective questionnaire (Table I) on different stress factors, i.e. stressed, tensed, concentrated and stimulated. In contrast, the exhausted factor presents no significant variation.
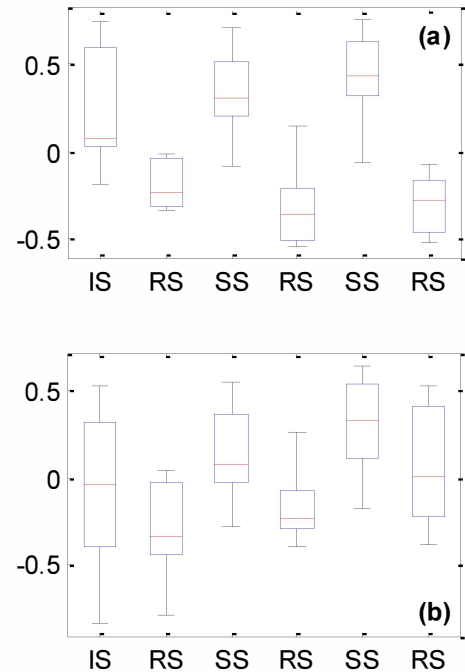

Fig. 7. Boxplots representing mean values for each sessions. Both webcam measurements (a) and EDRs (b) indicate an increase in mean values during the two stress sessions, compared to the three relaxation sessions. This plot represents data for all the twelve participants.

350

TABLE I.   RESULTS OF THE SUBJECTIVE EXPERIENCE OF STRESS.

Variable	RS	SS
Stressed	1.92 (0.99)	3.75 (0.75)
Tensed	1.58 (0.90)	3.58 (0.90)
Exhausted	1.25 (0.62)	1.83 (0.83)
Concentrated	2.75 (1.42)	4.58 (0.67)
Stimulated	2.92 (1.38)	3.92 (0.79)

These values represent the mean (SD) for each factor on all participants, scaled from 1 (not at all) to 5 (extremely).

IV.   DISCUSSION

Remote measurement of the HR and HRV is a powerful tool for monitoring and assess the mental state of a person. We choose to use an affordable technology to measure these physiological parameters. The results presented in this study demonstrate the feasibility of using the HRV response derived from a low-cost webcam to assess mental workload changes. The processing algorithms presented in this study are robust to light deficiency and motion-tolerant (see Fig. 8 for a typical example). The boxplot of means printed in Fig. 7 gives an estimation of the stress curves. These curves tend to decrease, as for EDRs, during the relaxation sessions. In contrast, the curves tend to increase during stress sessions. The relatively high disparity on the webcam data [Fig.7 (a)] indicates that participants were more stressed during the IS and SS than during the RS, in accordance with their subjective experience of stress.
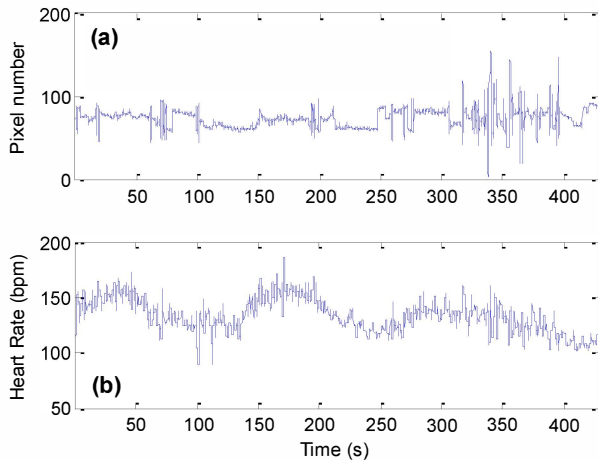

Fig. 8. Resistance to motion: Horizontal head movements are recorded (a) during the experiment. The instantaneous HR (b) is robust to artifacts.

Tapping errors were recorded and must be analyzed in future works to observe if errors have an impact on the participants' mental workload. The number of recognized emotions must be increased by integrating other modalities. Thus, analysis of video records will be the subject of future works to propose a multimodal emotion recognition framework. The capability to understand human emotion is promising, particularly in adaptive VR, where emotions are feedbacks that regulate the VR environment level and intensity.

REFERENCES

[1] D. Kukolja, S. Popovic, B. Dropuljic, et al., "Real-time emotional state estimator for adaptive virtual reality stimulation," in Proc. 5th Int. Conf. Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience, San Diego, CA, USA, Jul. 2009, pp. 175–184.

[2] A.A. Rizzo, J. Difede, B. O. Rothbaum, et al., "VR PTSD exposure therapy results with active duty OIF/OEF combatants," Medicine meets Virtual Reality, vol. 17, pp. 277–282, 2009.

[3] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, "Physiological signals based human emotion Recognition: a review," in IEEE 7th Int. Colloq. Signal Process. Appl., Arau, Malaysia, 2011, pp. 410–415.

[4] I. Pavlidis, J. Dowdall, N. Sun, et al., "Interacting with human physiology," Computer Vision and Image Understanding, vol. 108, pp. 150–170, Oct. 2007.

[5] V. Vasu, C. Heneghan, T. Arumugam, and S. Sezer, "Signal processing methods for non-contact cardiac detection using Doppler radar," in IEEE Workshop on Signal Process. Syst., San Francisco, CA, Oct. 2010, pp. 368–373.

[6] C. Gu, R. Li, C. Li, and S. B. Jiang, "Doppler radar respiration measurement for gated lung cancer radiotherapy," in IEEE Topical Conf. Biomed. Wireless Technol. Netw. Sens. Syst., Phoenix, AZ, Jan. 2011, pp. 91–94.

[7] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," IEEE Trans. Biomed. Eng., vol. 58, pp. 7–11, Jan. 2011.

[8] Y. Sun, S. Hu, V. Azorin-Peris, et al., "Noncontact imaging photoplethysmography to effectively access pulse rate variability," J. Biomed. Opt., vol. 18, 061205, Jun. 2013.

[9] A. Hoover, A. Singh, S. Fishel-Brown, and E. Muth, "Real-time detection of workload changes using heart rate variability," Biomedical Signal Processing and Control, vol. 7, pp.333–341, 2012.

[10] B. M. Appelhans and L. J. Luecken, "Heart rate variability as an index of regulated emotional responding," Review of General Psychology, vol. 10, pp. 229–240, 2006.

[11] W.A. Tiller, R. McCraty and M. Atkinson, "Cardiac coherence: a new, noninvasive measure of autonomic nervous system order," Altern. Ther. Health Med., vol. 2, pp. 52–65, 1996.

[12] K. Cosic, S. Popovic, D. Kukolja, et al., "Physiology-driven adaptive virtual reality stimulation for prevention and treatment of stress related disorders," Cyberpsychol. Behav. Soc. Netw., vol. 13, pp. 73–78, 2010.

[13] R. B. Knapp, J. Kim, and E. André, "Physiological signals and their use in augmenting emotion recognition for Human-Machine Interaction," in Emotion-Oriented Systems, Springer-Verlag, 2011, pp. 133–159.

[14] J. Zhai and A. Barreto, "Stress recognition using non-invasive technology," in Proc. 19th Int. Florida AI Res. Soc. Conf., Melbourne Beach, Florida, USA, May 2006, pp. 395–401.

[15] A. H. Garde, B. Laursen, A. H. Jorgensen, and B. R. Jensen, "Effects of mental and physical demands on heart rate variability during computer work," Eur. J. Appl. Physiol., vol. 87, pp. 456–461, 2002.

[16] K. Hercegfi, "Heart rate variability monitoring during Human-Computer Interaction", Acta Polytechnica Hungarica, vol. 8, pp. 205–224, 2011.

[17] J. Gonzalez-Sanchez, M. E. Chavez-Echeagaray, R. Atkinson, W. Burleson, "ABE: an agent-based software architecture for a multimodal emotion recognition framework," in Proc. 9th Working IEEE/IFIP Conf. Softw. Archit., Boulder, Colorado, USA, June 2011.

[18] B. L. Horecker, "The absorption spectra of hemoglobin and its derivatives in visible and near infrared regions," J. Biol. Chem., vol. 148, pp. 173–183, 1943.