



Remote detection of mental workload changes using cardiac parameters assessed with a low-cost webcam



Frédéric Bousefsaf*, Choubeila Maaoui, Alain Pruski

Laboratoire de Conception, Optimisation et Modélisation des Systèmes (LCOMS), Université de Lorraine, Bâtiment ISEA (Institut Supérieur d'Electronique et d'Automatique), 7 rue Marconi, 57070 METZ Technopôle, France

ARTICLE INFO

Article history:

Received 15 April 2014

Accepted 20 July 2014

Keywords:

Pulse rate variability
Photoplethysmography
Mental workload changes
Webcam

ABSTRACT

We introduce a new framework for detecting mental workload changes using video frames obtained from a low-cost webcam. Image processing in addition to a continuous wavelet transform filtering method were developed and applied to remove major artifacts and trends on raw webcam photoplethysmographic signals. The measurements are performed on human faces. To induce stress, we have employed a computerized and interactive Stroop color word test on a set composed by twelve participants. The electrodermal activity of the participants was recorded and compared to the mental workload curve assessed by merging two parameters derived from the pulse rate variability and photoplethysmographic amplitude fluctuations, which reflect peripheral vasoconstriction changes. The results exhibit strong correlation between the two measurement techniques. This study offers further support for the applicability of mental workload detection by remote and low-cost means, providing an alternative to conventional contact techniques.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Stress has repeatedly been associated with an increased risk for cardiovascular disease by primarily impacting blood pressure [1]. Depression, for example, corresponds to a risk factor for coronary heart disease [2]. Stress also impairs working memory and general cognitive function [3,4]. The association between affective states and computers has been popularized by Picard [4] who herein created the affective computing scientific domain. In these kinds of human–machine interactions, the computer is able to quantify affective states, stress and emotions [5] by using behavioral information and physiological parameters of the subject. Herein, stress detection and particularly mental workload changes are used to regulate the user-interface or the virtual environment to facilitate interactions [6].

Quantifying stress by its physiological signature is a field of research that presents a particular and increasing interest, where physiological parameters like Heart Rate (HR) and Heart Rate Variability (HRV) are reliable inputs to quantify different forms of stress [7–10]. However, contact sensors can be limited in some scopes of application where a specialist must install and monitor them [11]. In psychophysiological experiments, contact sensors

may generate a bias by interfering with the user, resulting practically by an erroneous estimation [12].

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 [7]. Its spectral analysis can provide the sympathovagal balance, a ratio that reflects reciprocal changes of sympathetic and vagal outflows [13]. The HRV tends to be rhythmic and ordered in relaxed and calm 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 when enduring stress. These rhythmic variations provide a state known as cardiac coherence [14], where the HRV regularity can be quantified using entropy-based algorithms [15]. Assessment of physiological signals by remote technologies is particularly advantageous in applications that need to understand feelings and sentiments of a 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 [12]. Similarly, Doppler radars are non-contact sensors that were used to detect heartbeats [16] and respiration signals [17]. More recently, digital cameras and webcams were employed on the face to detect the blood volume pulse [18–21] and compute heart rate and breathing rate. The principle, based on PhotoPlethysmography (PPG) consists in observing light variations on the skin to recover the cardiovascular pulse wave. The main drawback of this

* Corresponding author. Tel.: +33 387 547 652; fax: +33 387 547 301.
E-mail address: frederic.bousefsaf@univ-lorraine.fr (F. Bousefsaf).



Fig. 1. Protocol of the experiment, composed by three stress sessions and three relax sessions.

technique is that PPG signals are susceptible to motion-induced artifacts [22], particularly when dealing with webcams and ambient light. Standards of measurement recommend the use of ECG sensors to measure HRV [23]. However, it has been shown that Pulse Rate Variability (PRV) derived from PPG signals can be a good surrogate of HRV computed using ECG [24–26]. Sun et al. have compared performances between a low-cost webcam and a high-sensitivity camera to assess HR and PRV. 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 [19]. Here, a low-cost webcam can be a good surrogate to conventional contact sensors when assessing the cardiovascular pulse wave. This particular signal can be used to evaluate the ANS by observing changes in the period of the peaks and by observing fluctuations in the amplitudes or in the baseline of the signal [27–29].

We have recently developed a robust method to extract the PRV signal using the u^* channel of the CIE $L^*u^*v^*$ color space combined to a skin detection, an essential step that improves signal to noise ratio [20]. Then, we have employed this method to quantify mental workload changes using PRV-derived parameters [21]. In this paper, we extend this methodology by proposing a new filtering technique that was developed to remotely and robustly recover the instantaneous HR signal concurrently to photoplethysmographic amplitudes fluctuations from video frames acquired by a low-cost webcam. Orchestrated by the ANS, a peripheral vasoconstriction appears under stressful situations and leads PPG amplitudes to decrease [27]. We have employed these parameters to form a curve that represents mental workload changes for each of the 12 participants that were performing a computerized and interactive version of the Stroop color test.

The main contributions of this article are: (1) to provide a filtering technique based on the continuous wavelet transform of the raw PPG signal to automatically track HR variations on time using an adaptive window, and (2) to estimate mental workload changes of a participant by computing a set of basic parameters extracted from the instantaneous heart rate and the PPG amplitude fluctuations.

Firstly, we describe the approach, where a continuous wavelet transform filtering method was developed to precisely recover cardiac parameters of all participants. Secondly, we specify the protocol and the modalities used to induce stress during the experiments. Then, we detail how we have computed the parameters extracted from the HR series to form the mental workload curves.

2. Methods

2.1. Experimental procedure

Twelve students (two females and ten males, 22–27 years) 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 [8]. Briefly, the participant has 3 s to click on the colored box that corresponds to the word printed in the center of the monitor (Fig. 2). 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 in the right color (congruent, e.g., the word “pink” written in pink).

The participants performed three sessions (see Fig. 1) of the color word test, i.e. a one minute training session (TS) 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.

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 and stimulated [30]. They gave two sets of five responses: one set for the two stress sessions and one set for the three relaxation sessions. 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 activity was concurrently recorded using a skin conductance sensor.

2.2. Materials

A low-cost HD webcam (Lifecam Cinema by Microsoft) was used in these experiments. The resolution of the device is reduced to 320×240 pixels in order to keep an acquisition frequency of 30 fps. The maximum webcam resolution is 1280×800 pixels. The three RGB channels are encoded with 8 bits per pixel. It is important to note that auto white balance is disabled in these experiments. White balance locally regulates colors and generates non-desired artifacts in webcam PPG signals. A finger skin conductance sensor (SC-Flex/Pro by Thought Technologies Ltd.) was used to measure the electrodermal activity at a sampling frequency of 256 Hz. PPG signals [see Fig. 3(e)] were recorded with a C++ based software and analyzed offline with MATLAB (The MathWorks, Inc.).

2.3. Preprocessing operations

The overall system is composed with both image and signal processing [20]. Briefly, the raw PPG signal $x(t)$ is obtained using a spatial averaging operation [Fig. 3(e)] on the merged frame [Fig. 3(d)] computed using the skin detection mask [Fig. 3(b)] and the u^* component frame of the CIE $L^*u^*v^*$ color space [Fig. 3(c)]. It has been shown that using such a component improves the robustness of the system in presence of noise induced by motion or light artifacts [20]. The skin detection mask was developed to collect only PPG pixels that contain the pulse wave signal. The filter is established in the YCbCr color space by setting an empiric threshold on the 3 channels [20]. A set of t frames gives a raw signal of t points [Fig. 3(e)].



Fig. 2. Screenshots of the interactive application: during the Stroop color word test (left picture) and the first relaxation video (right picture) that starts right after the training session.

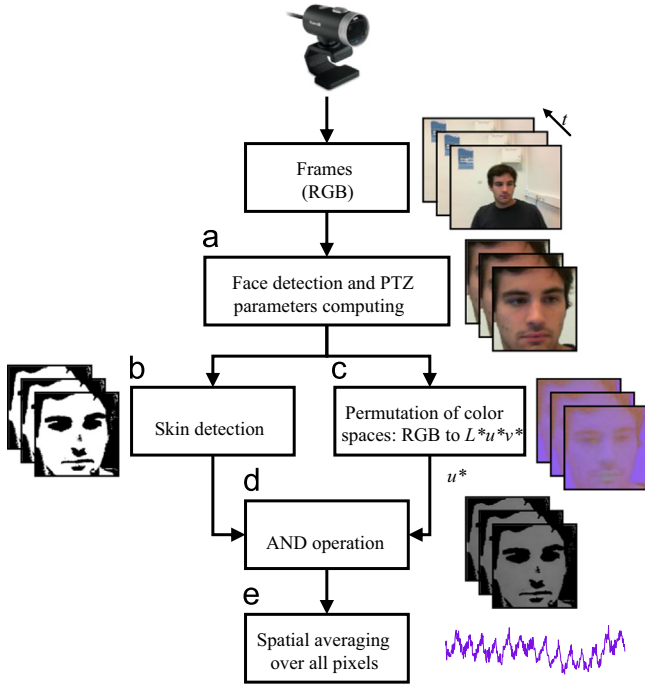


Fig. 3. Processing algorithm overview [20]. (a) Pan, Tilt and zoom parameters are computed to zoom and track the face on the input frame. (b) Pixels that contain PPG information are isolated by a skin detection. (c) The RGB color space is converted to the CIE $L^*u^*v^*$ color space. (d) The u^* frame is merged with the skin detection mask by a combinational AND operation. (e) A spatial averaging step is performed to transform a set of frames into a single raw signal. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.4. Continuous wavelet transform filtering to assess the cardiac pulse wave

The DC component is primarily removed to reveal detailed information [31] on lower scales prior to performing the Continuous Wavelet Transform (CWT):

$$x_{AC}(t) = x(t) - \mu \quad (1)$$

where μ is the mean of the raw PPG signal $x(t)$ [Fig. 3(e)]. The webcam acquisition generates irregularly time-sampled signals. To avoid these irregularities, the normalized signal $x_{AC}(t)$ is resampled using a 30 Hz cubic spline function. A CWT filter was developed to remove trends and high frequency noise of the raw signal in the [0.6–4 Hz] frequency band using the Morlet wavelet. This wavelet has already been used as a mother wavelet to analyze biomedical signals and particularly blood flow signals [32] and was employed in this study. A typical example is presented in the scalogram from

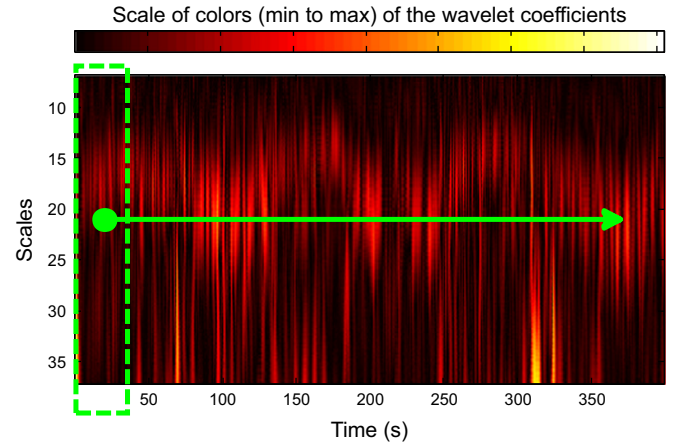


Fig. 4. The continuous wavelet transform of the PPG signal is computed between 0.6 and 4 Hz, corresponding respectively to scales 37 and 7 in the present scalogram. This particular plot represents the absolute value of each wavelet coefficient. The pulse wave oscillations are distinctly identifiable around the scale 20, corresponding to 1.22 Hz in the frequency domain. A 30 s moving window, represented in dashed green lines, runs through the signal by a constant 3 seconds step. This figure is an extract from the signals of subject #12. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 4, where high scales correlate better with low frequencies and low scales with high frequencies of the signal by respectively stretching or compressing the wavelet. A mask is then computed from the wavelet coefficients to attenuate effects generated by high frequency noise or low frequency trends. An iterative algorithm is adopted and uses the wavelet coefficients included in a 30 s moving window (see Fig. 4) that runs through the entire CWT with a constant 3 s step. For each of these regions of interest, the global energy is computed using the following formula:

$$E = \sum |CWT(x_{AC})^T| \quad (2)$$

where $CWT(x_{AC})$ is the continuous wavelet transform of the resampled PPG signal in the region of interest. The summation is computed for each scale on the CWT and gives a scalar by scale, forming the global energy plot (see Fig. 5). We select the point that presents the maximum amplitude in the energy axis, and use the corresponding scale for further processing.

Based on the reference point location, we construct a trapezoidal weighting window (Fig. 6) for each scale using the following formulas:

$$\begin{aligned} &1 && \text{if } RP_L < sc < RP_H \\ &\frac{RP_H - sc}{RP_H - RP_L} + 1 && \text{if } RP_H < sc < RP_{HH} \\ &\frac{sc - RP_L}{RP_L - RP_{LL}} + 1 && \text{if } RP_L < sc < RP_{LL} \end{aligned} \quad (3)$$

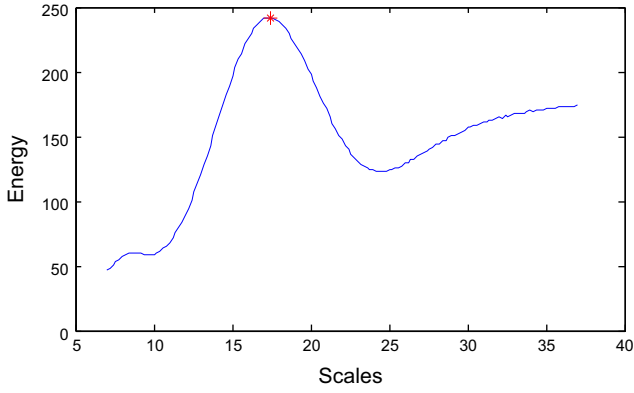


Fig. 5. A typical example representing the global energy plot of a CWT region of interest. The reference point, displayed with a red star in the figure, corresponds to the scale that presents the maximum energy. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

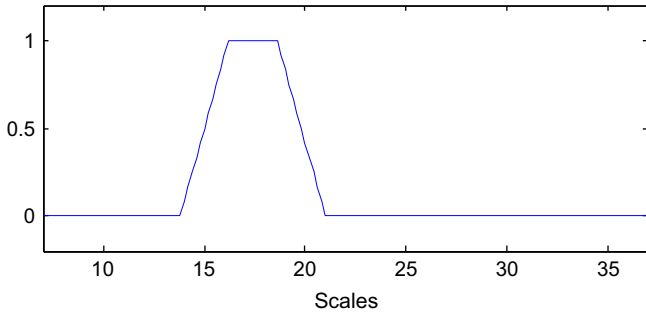


Fig. 6. The weighting window of the CWT region of interest, centered on the reference point.

sc corresponds to the scales. If we take into account the example presented before, sc scans all scales from 7 to 37 (corresponding to 0.6–4 Hz, see Fig. 4). RP_L , RP_{LL} , RP_H and RP_{HH} are a set of limits that are empirically determined using the reference point (RP in Eq. 4) location. Herein, RP_L and RP_H are a set of two limits where the weighting window (Fig. 6) is set to 1. Then, a linear cut is employed between RP_H and RP_{HH} and between RP_L and RP_{LL} to avoid abrupt breaks at the border of the window, in order to form a trapezoidal window instead of a rectangular window.

$$RP_L = RP - [4 \times (sc_{max} - sc_{min}) / 100]$$

$$RP_{LL} = RP - [12 \times (sc_{max} - sc_{min}) / 100]$$

$$RP_H = RP + [4 \times (sc_{max} - sc_{min}) / 100]$$

$$RP_{HH} = RP + [12 \times (sc_{max} - sc_{min}) / 100] \quad (4)$$

where sc_{min} is the lowest scale and sc_{max} the highest scale, respectively 7 and 37 in this particular case. RP corresponds to the reference point (see Fig. 5).

A weighting window is computed for each region of interest (Fig. 4) and is combined, coefficient by coefficient, to the CWT of the following region of interest by an element-wise multiplication.

At this stage, a new weighting window is constructed from the weighted CWT [Fig. 7(b)]. This technique was developed to track frequency fluctuations of the cardiac pulse wave signal in the scale domain, by automatically adapting the weighting window location over time. When all the weighting windows are computed, a weighting matrix is generated where the windows have successively been inserted one beside the other [Fig. 8(a)].

The weighting matrix presented in [Fig. 8(a)] exhibits the trends followed by the cardiac pulse rate of the participant. Nevertheless and to assess precisely the pulse rate variability, we need to perform a beat to beat analysis and assess the instantaneous pulse rate. Thus, this weighting matrix was employed only for filtering purpose on the entire signal. The result is presented in [Fig. 8(b)], where the CWT is weighted using the weighted matrix from [Fig. 8(a)].

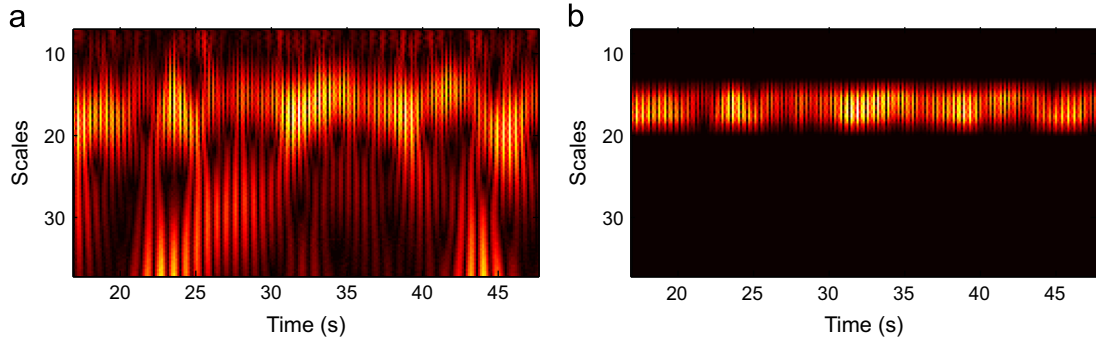


Fig. 7. (a) The wavelet coefficients of a region of interest. (b) The representation is weighted using the weighting window computed with the previous region of interest.

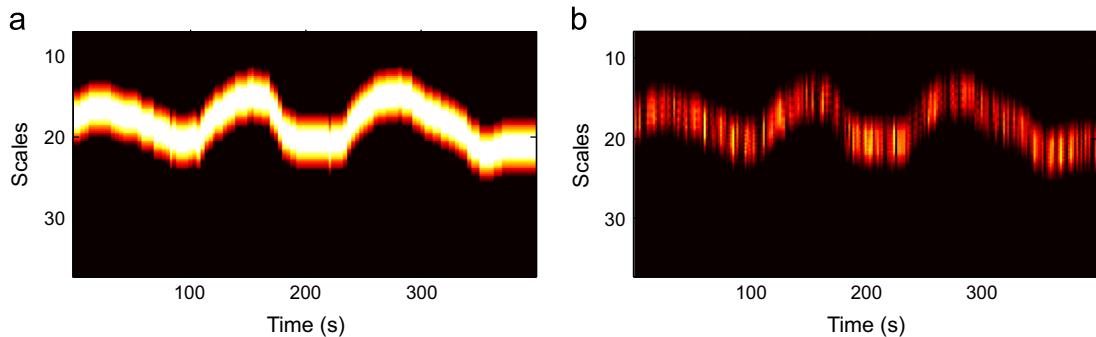


Fig. 8. All the weighting windows have been placed inside a weighting matrix (a). The pulse wave frequency changes are distinctly identifiable in this representation. The weighting matrix is applied on the entire CWT of the PPG signal (b). The inverse transform of the resulting coefficients is performed to remove trends and high frequency noise of the raw PPG signal.

The inverse transform of the weighted coefficients [Fig. 8(b)] gives a refined representation of the signal, without trends and high frequency artifacts [Fig. 9(b)]. To precisely assess the peaks, this signal is interpolated with a cubic spline function at a sampling frequency of 256 Hz. An existing peaks detection algorithm (*findpeaks* function in MATLAB) was employed to find local maxima, assess the InterBeat Intervals (IBI) and compute the instantaneous pulse rate (see Fig. 10). The minimum peak separation threshold was defined and fixed empirically. The peaks amplitudes were recorded and used to form mental workload curves (see Section 2.5).

Residual ectopic beats are detected and removed using a threshold condition defined empirically. Briefly, if a difference between two consecutive IBIs higher than 40 bpm occurs, the ectopic beat is removed and replaced by the average of the preceding and following beats (see the typical example presented in Fig. 10).

2.5. Mental workload changes detection using the instantaneous heart rate trace

Two parameters are derived from the previously quantified instantaneous HR trace. In order to suppress time-sampled irregularities, the signal is interpolated with a 15 Hz cubic spline function. The HR tends to increase gradually during the stress sessions, just like in the typical example presented in Figs. 8 and 10. To recover these trends, a 20 s two-sided moving average (5) is computed on the interpolated HR signal (Fig. 11).

$$MA(n) = \frac{1}{N} \sum_{k=0}^{N-1} y_{n-k} \quad (5)$$

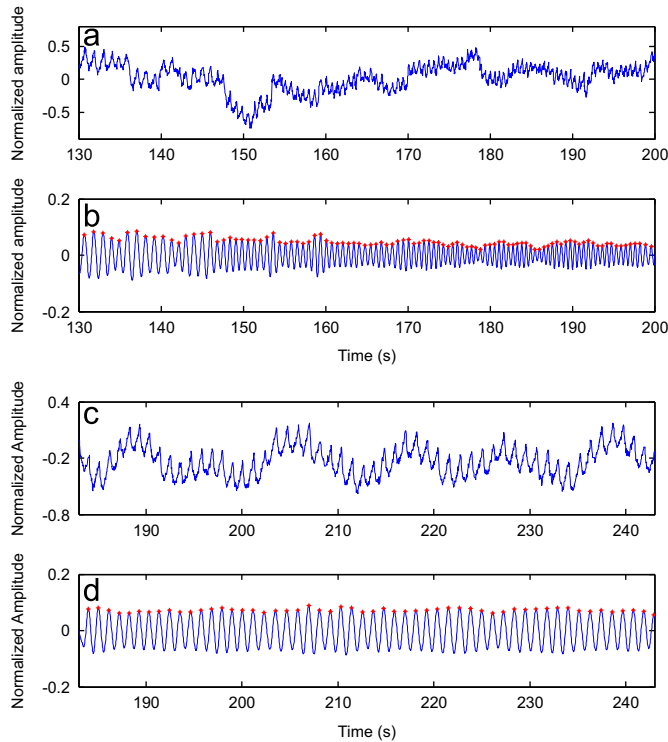


Fig. 9. Two typical examples representing the PPG signal before (a) and (c) and after the filtering process (b) and (d). Herein, red stars correspond to the peaks detected by the algorithm. In the first case, we can visually observe that the HR of the participant increases while the intensities of the PPG signal decrease. In this recording from participant #11, the second stress session was precisely launched at $t=135$ s. In contrast, the HR of the subject presented in the second case is slower and more stable and no significant fluctuations in intensities are notable here. This particular recording was extracted from participant #4 data, where the relax session was launched around $t=180$ s. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

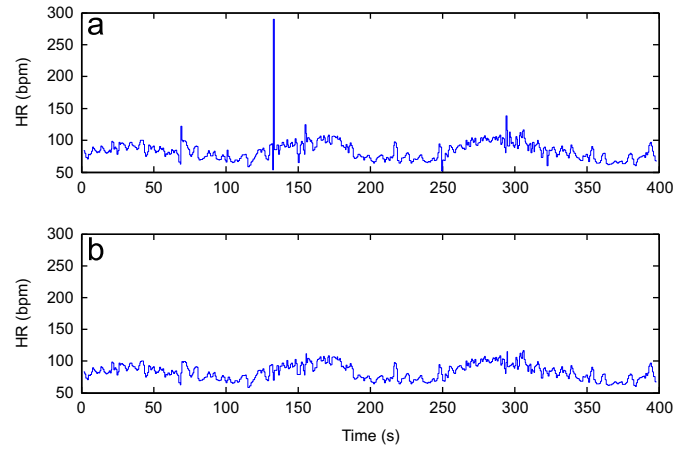


Fig. 10. The instantaneous pulse rate trace is formed after detecting local maxima, followed by the computation of the IBIs. Similarly to Fig. 8, the trend is distinctly identifiable in this figure. Residual ectopic beats (a) are removed using a set of thresholds (b). These signals are from participant #12.

With MA the two-sided moving average of the instantaneous HR signal, y corresponds to the interpolated instantaneous HR signal (see Fig. 10). The moving average is computed on the entire signal, represented by the n index.

Finally and to compute the mental workload curve (see Fig. 12), the two parameters were normalized (6) and combined (7). This sum produces a new curve, smoothed using a 20 s two-sided moving average filter. Herein, we propose a simple yet efficient approach which is based on observations: all these parameters vary concurrently and the combination is only employed to magnify their simultaneous increases and decreases.

$$\delta = \frac{\delta - \mu}{\sigma} \quad (6)$$

$$y_{workload} = MA + ampl \quad (7)$$

With $y_{workload}$ the resulting curve that represents mental workload changes (Fig. 12), $ampl$ corresponds to the PPG amplitude fluctuation parameter (see Fig. 11). μ and σ correspond respectively to the mean and standard deviation of the parameter to be normalized, represented in Eq. 6 by the δ variable.

3. Results

Workload signals were compared to electrodermal responses to assess the agreement with the camera measurements. A typical example is presented in Fig. 12 where the workload curve generated using the PRV signal is in close agreement with the raw electrodermal trace of the participant. The skin conductance level, the tonic component of the electrodermal activity, was computed using a 20 s two-sided moving average filter.

To be compared, both webcam and skin conductance level signals were normalized between -1 and 1 (Fig. 12). Boxplots are employed to represent the differences on mean and derivative of both webcam and electrodermal measurements for each session (see Figs. 13 and 14).

A statistical analysis was used to quantify the level of agreement between physiological measurements by the webcam and the contact skin conductance sensor. Results of the analysis are presented in Table 1 and plotted on Figs. 13 and 14, where performances between the two measurements techniques are revealed. Results of the subjective experiences questionnaires are reported in Table 2. Significant differences were observed on the

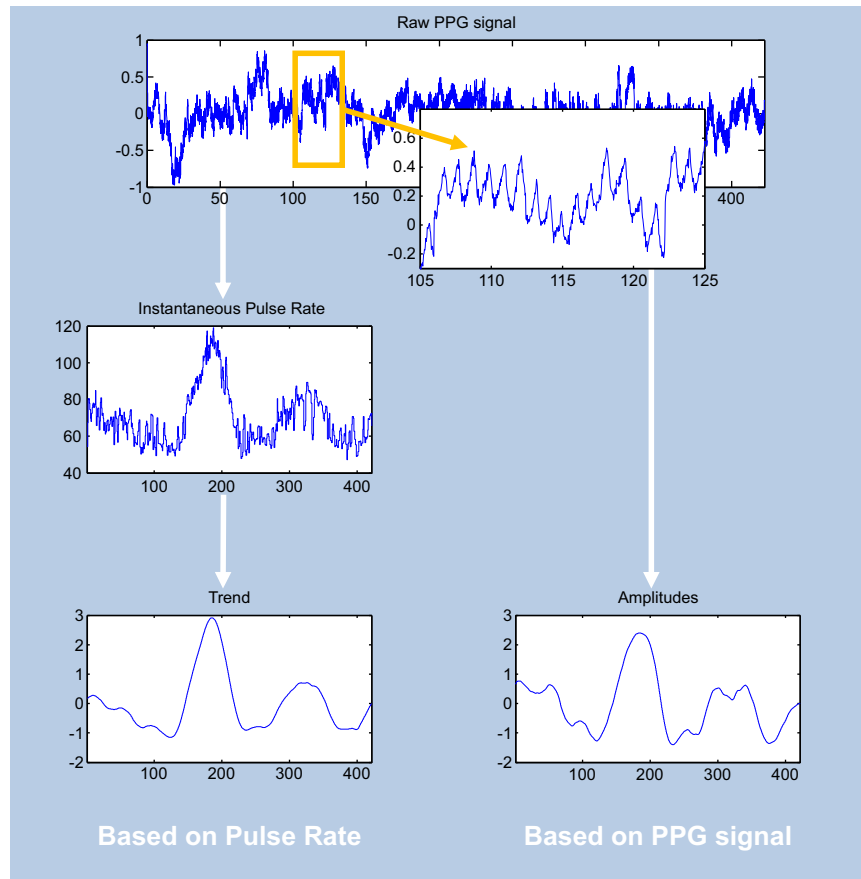


Fig. 11. Two parameters are computed from the raw PPG signal to quantify mental workload changes: the trend, by a 20 s two-sided moving average on the instantaneous pulse rate signal and the PPG amplitude fluctuation, which reflects peripheral vasoconstriction changes.

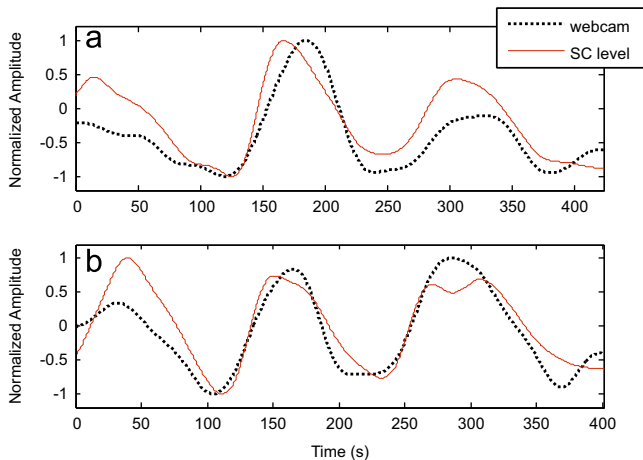


Fig. 12. Results of the mental workload detection for the participant #11 (a) and #12 (b). Black plots correspond to the webcam-derived workload signal and red plots to the skin conductance level, derived from the raw electrodermal activity signal. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

questionnaires between the relaxation and stress sessions following the stressed, tensed, concentrated and stimulated factors. In contrast, the exhausted factor presents no significant variation. Also, results of the questionnaires have been reported in Table 1. These described values are means of the most four relevant factors. Fig. 13 represents the mean value of respectively the webcam [Fig. 13(a)] and the skin conductance [Fig. 13(b)]

measurements during all sessions. A significant increase in the averages between stress and relaxation sessions is observable, independently of the measurement technique. Similarly, the evolution of the mental workload signal, i.e. the derivative of the responses, is quantified using the webcam [Fig. 14(a)] and skin conductance sensor [Fig. 14(b)]. The boxplots of derivatives present significant differences between the three RS and the two SS while the training session is located between them.

The global measurements of all relaxation and stress sessions are plotted in Fig. 15, for both contact and remote sensors and for the questionnaires. Specifically, the webcam-derived median was -0.3 normalized units (n.u.) for all RS and 0.33 n.u. for the two SS [Fig. 15(a)]. Medians of the derivative measurements present a significant increase, -0.011 to 0.015 n.u. for relaxation and stress sessions respectively [Fig. 15(c)]. Similar tendencies were extracted from skin conductance level signals, where medians vary between -0.27 and 0.32 for respectively RS and SS [Fig. 15(b)]. The derivatives measurements evolve from -0.009 to 0.016 between relax and stress sessions [Fig. 15(d)]. Also, the questionnaires were averaged to only obtain a single value per participant and per session. Medians vary between -0.5 and 0.5 for RS and SS respectively [Fig. 15(e)].

Pearson's correlation coefficients were used to quantify the level of agreement between physiological measurements by the remote and contact techniques (see Table 1). The time series, i.e. the skin conductance level and the mental workload signal derived from the webcam (Fig. 12) are correlated, except for participants #2 and #7. In addition, correlation coefficients for mean and derivative measurements were computed (always between webcam and skin conductance recordings) and tend to follow those computed between time series.

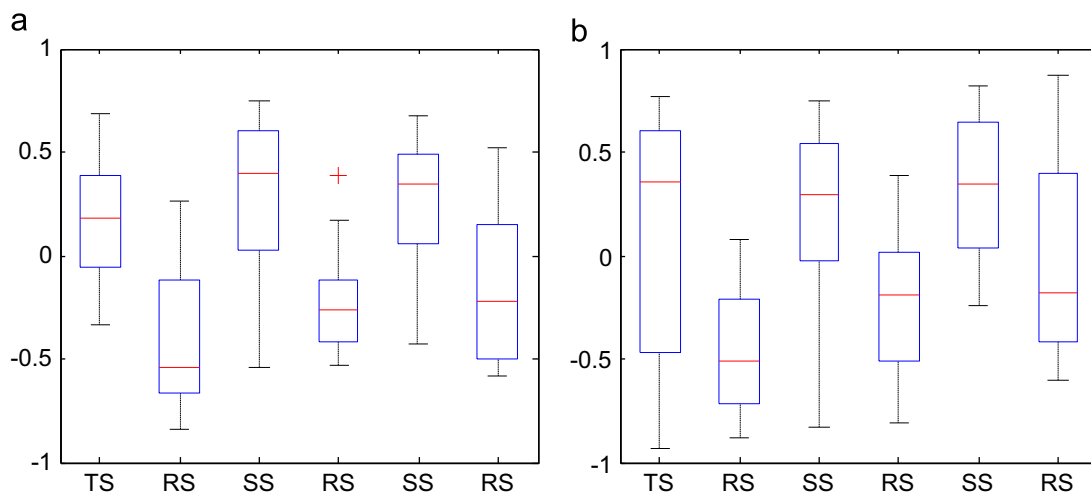


Fig. 13. Boxplots representing mean values for each sessions. Both webcam (a) and electrodermal (b) measurements indicate an increase on mean values during the two stress sessions, compared to the three relaxation sessions. This plot represents data for all the twelve participants.

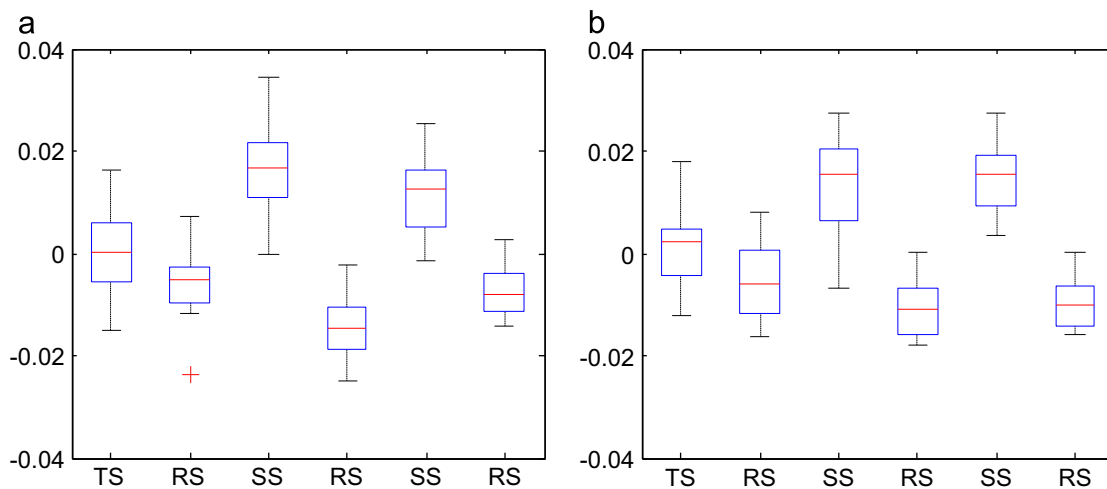


Fig. 14. Boxplots representing the webcam (a) and electrodermal (b) signal evolution. Both traces are computed using derivative operations.

Table 1

Results of the statistical analysis for both stress and relaxation sessions.

#	Global measurements during RS					Global measurements during SS					Correlation between time series	Correlation between mean measurements	Correlation between derivative measurements
	Webcam		Skin conductance		SES ^c	Webcam		Skin conductance		SES			
	μ^a	δ^b	μ	δ		μ	δ	μ	δ				
1	-0.40	-0.015	-0.37	-0.012	-1	0.19	0.020	0.65	0.016	0.38	0.8	0.81	0.93
2	0.05	-0.004	-0.19	-0.004	-0.5	0.20	0.012	0.30	0	0.5	< 0.1	< 0.1	< 0.1
3	-0.27	-0.006	-0.48	-0.011	0	0.38	0.006	-0.22	0.01	0.5	0.36	0.31	0.41
4	-0.39	-0.011	0.03	-0.013	-0.5	-0.30	0.019	0.23	0.023	0.5	0.62	0.5	0.95
5	0.16	-0.01	-0.07	-0.003	0.75	0.58	0.017	0.40	0.012	1	0.17	< 0.1	0.65
6	-0.07	-0.011	-0.21	-0.012	-0.88	0.40	0.017	0.33	0.007	0.25	0.52	0.37	0.90
7	-0.20	0	0.00	-0.002	-0.63	0.64	-0.001	0.44	0.015	0.75	< 0.1	< 0.1	0.28
8	-0.40	-0.01	-0.43	-0.006	-0.63	-0.09	0.012	-0.26	0.017	0.5	0.55	0.63	0.54
9	-0.30	-0.013	-0.33	-0.006	0.25	0.28	0.013	-0.15	0.004	0.25	0.64	0.67	0.41
10	-0.30	-0.023	0.07	-0.006	-0.5	0.49	0.008	0.64	0.019	0.5	0.73	0.73	0.26
11	-0.58	-0.011	-0.49	-0.012	-0.25	-0.08	0.024	0.32	0.022	0.25	0.86	0.88	0.95
12	-0.49	-0.011	-0.33	-0.015	-0.38	0.51	0.019	0.42	0.019	0.38	0.88	0.94	0.88

^a μ corresponds to the average of mean values during the two stress sessions or the three relaxation sessions.

^b δ corresponds to the average of derivative values during the two stress sessions or the three relaxation sessions.

^c SES: Subjective Experiences of Stress, corresponding to the average of the stressed, tensed, concentrated and stimulated factors.

Particular discrimination plots (Fig. 16) were employed to demonstrate that mean and derivative parameters can be used to linearly discriminate the two classes. The derivative, plotted on the horizontal axis, correspond to a better discriminant parameter than the average, plotted on the vertical axis.

Table 2
Results of the Subjective Experiences of Stress.

Variable	RS	SS
Stressed	−0.54 (0.5)	0.38 (0.38)
Tensed	−0.71 (0.45)	0.29 (0.45)
Exhausted	−0.88 (0.31)	−0.58 (0.42)
Concentrated	−0.12 (0.71)	0.79 (0.33)
Stimulated	−0.04 (0.69)	0.46 (0.4)

These values represent the mean (SD) for each factor on all participants, scaled from −1 (not at all) to 1 (extremely). **RS** represents the three relaxation sessions and **SS** the two stress sessions.

4. Discussion

Remote measurement of the HR and PRV is a powerful tool for monitoring and assess the mental state of a person [12]. We choose to use an affordable technology to measure these physiological parameters. The results presented in this study demonstrate the feasibility of using the cardiac response derived from a low-cost webcam to assess mental workload changes. The processing method presented in this study is motion-tolerant and robust to light deficiency [20]. The instantaneous pulse rate can be properly assessed even in presence of strong motion artifacts (see Fig. 17). In this typical example, the participant was moving his head at $t \approx 315$ s [Fig. 17(a)]. The processing algorithms detailed in Section 2 were efficient to compensate motion fluctuations by generating a correct weighting matrix. This way, both the PPG amplitude fluctuation and the pulse rate variability [Fig. 17(b)] could be accurately assessed and the mental workload properly determined.

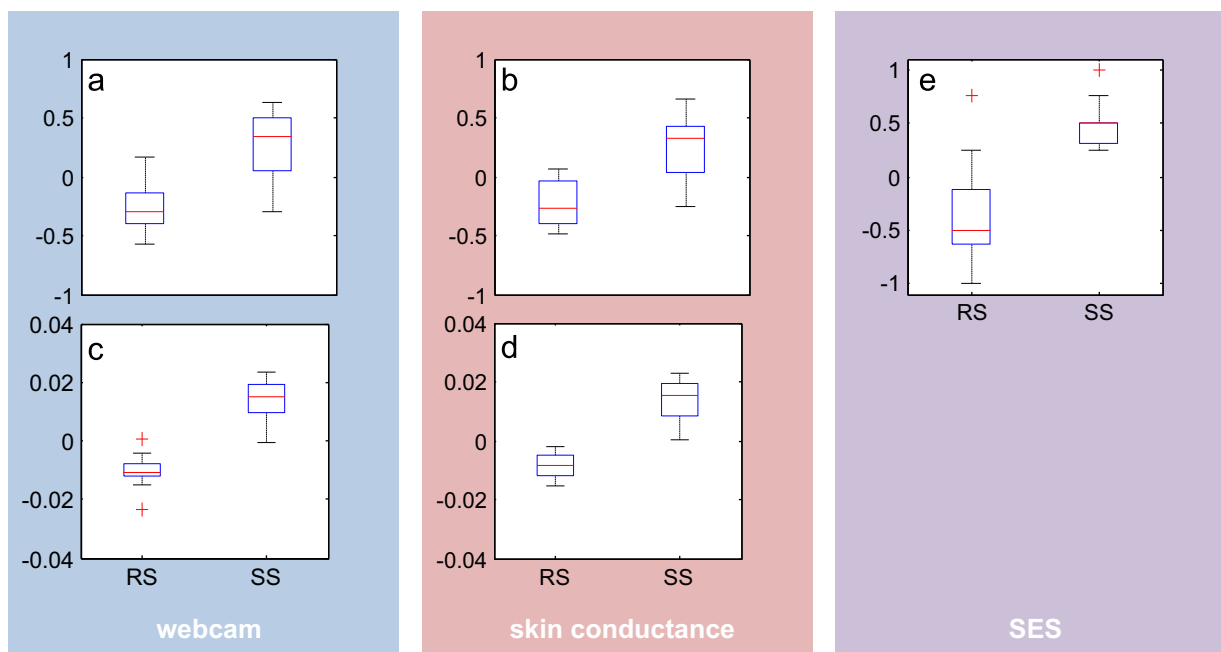


Fig. 15. Boxplots representing global average measurements of means in figures (a) and (b) and derivatives in figures (c) and (d) for the three relax sessions and the two stress sessions. The mean values of the stressed, tensed, concentrated and stimulated factors of the questionnaires are presented in (e).

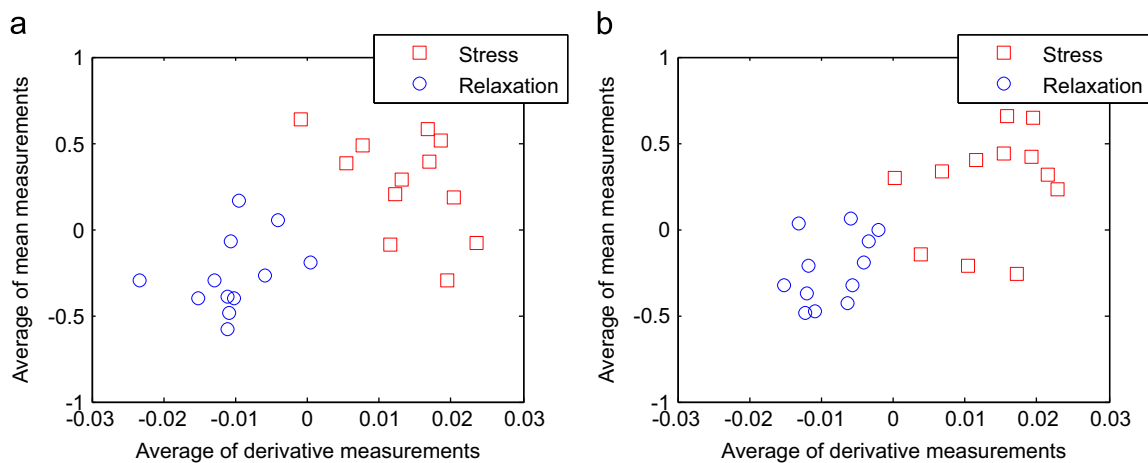


Fig. 16. Plots representing averages of derivative versus averages of mean for both webcam (a) and skin conductance recordings (b). Red squares represent values during the two stress sessions and blue circles during the three relax sessions. Each respective symbol represents a participant. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

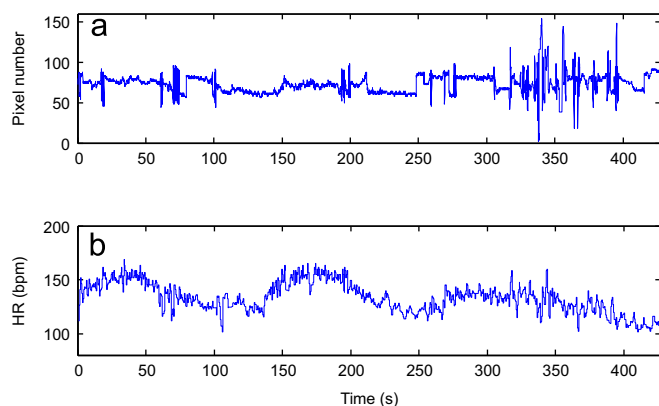


Fig. 17. Resistance to motion: horizontal head movements are recorded (a) during the experiment where strong fluctuations can be observed at $t \approx 315$ s. The instantaneous HR trace (b) is robust to these motion artifacts and is properly assessed. These signals are from subject #9.

A simple, yet efficient, mental workload detection is proposed in this study where we have employed 2 parameters derived from webcam PPG signals, i.e. the trend of the instantaneous HR, assessed by a moving average and the amplitude fluctuation of the PPG signal that reflects peripheral vasoconstriction changes, which is modulated by sympathetic nervous system activity [27]. There are certain limitations that should be pointed out when considering the combination approach. The sum we employed may not be the best method and it is not assured that parameters extracted from HR and PPG evolve in a linear way. A system identification should be conducted to address this issue in future work.

Generally, a significant increase on mean and derivative values is perceptible between relaxation and stress sessions, independently of the measurement techniques (see Fig. 15). The derivative and the mean can be used as parameter employed to directly discriminate calm and stress states without using particular machine learning algorithms (Fig. 16).

Results from Table 2 indicate that participants were effectively stimulated by the interactive stress test we propose in this study. Also, the significant difference between RS and SS on the stress factor indicates that our induction protocol, based on the Stroop test, was quite effective. The boxplot of means printed in Fig. 14 gives an estimation of the mental workload curves computed with the webcam measurements. As for the skin conductance level, these curves tend to decrease during relaxation sessions and, in contrast, tend to increase during stress sessions. Our results indicate that participants seem to be less stressed during the training session than during stress sessions. These variations are probably driven by the addition of stressors like sounds and timers (see the experimental protocol presented in Section 2.1). Nevertheless, the relatively high disparity on the data (Figs. 14 and 15) indicates that participants were more stressed during the TS and SS than during RS, in accordance with their subjective experiences of stress (see Table 2). Accordingly to the correlation coefficients presented in Table 1, subjects #2 and #7 presented no correlation between time series. This can be explained by the global trend of the signals (see Fig. 18 for the particular case of subject #7) which evolves in the inverse way, while local fluctuations are indeed correlated. For subject #7, the Pearson's correlation is equal to 0.7 if we compute the operation on detrended signals.

A limitation to note is that we perform these experiments on a set composed only by 12 participants. It is obvious that a larger sample should permit to ascertain the clinical validity of the method. Herein, the purpose of this study is to evaluate the feasibility of such a method, which is relatively innovative.

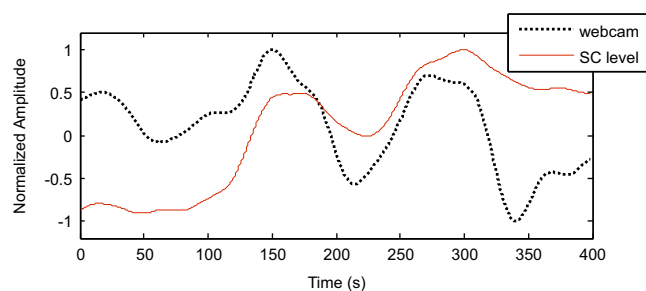


Fig. 18. Result of the mental workload detection for the participant #7. The black dotted line corresponds to the webcam-derived workload signal and the red one to the skin conductance level, derived from the raw electrodermal activity signal. The Pearson's correlation between these two series was < 0.01 . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Tapping errors were recorded during the stress tests and will be analyzed in future works in order to observe their impact on the participant 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 detect stress by non-contact means is promising, particularly in affective computing, where the stress level can be used as an input that regulate the environment parameters.

5. Summary

Monitoring physiological signals via noncontact means presents a greater challenge in personal health care, telemedicine and affective computing. In this study, we propose to remotely assess mental workload changes using the amplitude fluctuation of the photoplethysmographic signal concurrently with the pulse rate variability. These particular signals provide an estimation of the autonomic nervous system state and are formed using video frames of human faces, recorded with a low-cost webcam. Robust image and signal processing are introduced to collect only pixels that contain photoplethysmographic information. We have used a continuous wavelet transform filter to denoise and detrend signals in order to detect peaks and compute interbeat intervals. In order to validate the proposed method we have recorded, concurrently to the webcam curves, electrodermal activity during an interactive game that was developed to successively stress and relax the subject. The results exhibit a strong correlation between the trends of the webcam and contact skin conductance level traces and offer further support for the applicability of mental workload detection by remote and low-cost means, providing an alternative to conventional contact techniques.

Conflicts of interest

None declared.

References

- [1] T.G.M. Vrijkotte, L.J.P. van Doornen, E.F.C. de Geus, Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability, *Hypertension* 35 (2000) 880–886.
- [2] S.J. Bunker, D.M. Colquhoun, M.D. Esler, et al., Stress and coronary heart disease: psychosocial risk factors, *Med. J. Aust.* 178 (2003) 272–276.
- [3] G.E. Prinsloo, H.G. Laurie Rauch, M.I. Lambert, et al., The effect of short duration heart rate variability (HRV) biofeedback on cognitive performance during laboratory induced cognitive stress, *Appl. Cognit. Psychol.* 25 (2010) 792–801.

- [4] J.A. Healey, R.W. Picard, Detecting stress during real-world driving tasks using physiological sensors, *IEEE Trans. Intell. Transp. Syst.* 6 (2005) 156–166.
- [5] J. Zhai, A. Barreto, Stress recognition using non-invasive technology, in: *Proceedings of the 19th International Florida Artificial Intelligence Research Society Conference*, Melbourne Beach, Florida, USA, 2006, pp. 395–401.
- [6] K. Hercegfi, Heart rate variability monitoring during human–computer interaction, *Acta Polytech. Hung.* 8 (2011) 205–224.
- [7] A. Hoover, A. Singh, S. Fishel-Brown, E. Muth, Real-time detection of workload changes using heart rate variability, *Biomed. Signal Process. Control* 7 (2012) 333–341.
- [8] A.H. Garde, B. Laursen, A.H. Jorgensen, B.R. Jensen, Effects of mental and physical demands on heart rate variability during computer work, *Eur. J. Appl. Physiol.* 87 (2002) 456–461.
- [9] 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, *Biomed. Eng. Online* 10 (2011) 96.
- [10] K. Yoshino, K. Matsuoka, Personal adaptive methods to assess mental tension during daily life using heart rate variability, *Methods Inf. Med.* 51 (2012) 39–44.
- [11] J. Kranjec, S. Begus, G. Gersak, J. Drnovsek, Non-contact heart rate and heart rate variability measurements: a review, *Biomed. Signal Process. Control* 13 (2014) 102–112.
- [12] I. Pavlidis, J. Dowdall, N. Sun, et al., Interacting with human physiology, *Comput. Vis. Image Underst.* 108 (2007) 150–170.
- [13] B.M. Appelhans, L.J. Luecken, Heart rate variability as an index of regulated emotional responding, *Rev. Gen. Psychol.* 10 (2006) 29–240.
- [14] W.A. Tiller, R. McCraty, M. Atkinson, Cardiac coherence: a new, noninvasive measure of autonomic nervous system order, *Altern. Ther. Health Med.* 2 (1996) 52–65.
- [15] C. Liu, K. Li, L. Zhao, et al., Analysis of heart rate variability using fuzzy measure entropy, *Comput. Biol. Med.* 43 (2013) 100–108.
- [16] V. Vasu, C. Heneghan, T. Arumugam, S. Sezer, Signal processing methods for non-contact cardiac detection using Doppler radar, in: *Proceedings of the IEEE Workshop on Signal Processing Systems*, San Francisco, CA, 2010, pp. 368–373.
- [17] C. Gu, R. Li, C. Li, S.B. Jiang, Doppler radar respiration measurement for gated lung cancer radiotherapy, in: *IEEE Topical Meeting on Biomedical Wireless Technologies, Networks and Sensing System*, Phoenix, AZ, 2011, pp. 91–94.
- [18] M.-Z. Poh, D.J. McDuff, R.W. Picard, Advancements in noncontact, multi-parameter physiological measurements using a webcam, *IEEE Trans. Biomed. Eng.* 58 (2011) 7–11.
- [19] Y. Sun, S. Hu, V. Azorin-Peris, et al., Noncontact imaging photoplethysmography to effectively access pulse rate variability, *J. Biomed. Opt.* 18 (2013) 061205.
- [20] F. Bousefsaf, C. Maaoui, A. Pruski, Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate, *Biomed. Signal Process. Control* 8 (2013) 568–574.
- [21] F. Bousefsaf, C. Maaoui, A. Pruski, Remote assessment of the heart rate variability to detect mental stress, in: *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, Venice, Italy, 2013, pp. 348–351.
- [22] H. Han, J. Kim, Artifacts in wearable photoplethysmographs during daily life motions and their reduction with least mean square based active noise cancellation method, *Comput. Biol. Med.* 42 (2012) 387–393.
- [23] Task Force of the European Society of Cardiology, North American Society of Pacing and Electrophysiology, Heart rate variability: standards of measurement, physiological interpretation, and clinical use, *Circulation* 93, 1996, pp. 1043–1065.
- [24] E. Gil, M. Orini, R. Bailon, et al., Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions, *Physiol. Meas.* 31 (2010) 1271–1290.
- [25] P. Shi, S. Hu, Y. Zhu, A preliminary attempt to understand compatibility of photoplethysmographic pulse rate variability with electrocardiographic heart rate variability, *J. Med. Biol. Eng.* 28 (2008) 173–180.
- [26] S. Lu, H. Zhao, K. Ju, et al., Can photoplethysmography variability serve as an alternative approach to obtain heart rate variability information? *J. Clin. Monit. Comput.* 22 (2008) 23–29.
- [27] E. Peper, R. Harvey, I.-M. Lin, et al., Is there more to blood volume pulse than heart rate variability, respiratory sinus arrhythmia, and cardiorespiratory synchrony, *Biofeedback* 35 (2007) 54–61.
- [28] M. Nitzany, A. Babchenko, B. Khanokh, D. Landau, The variability of the photoplethysmographic signal – a potential method for the evaluation of the autonomic nervous system, *Physiol. Meas.* 19 (1998) 93–102.
- [29] K.-S. Yoo, W.-H. Lee, Mental stress assessment based on pulse photoplethysmography, in: *Proceedings of the IEEE 15th International Conference Symposium on Consumer Electronics*, Singapore, 2011, pp. 323–326.
- [30] N. Hjortskov, D. Rissen, A.K. Blangsted, et al., The effect of mental stress on heart rate variability and blood pressure during computer work, *Eur. J. Appl. Physiol.* 92 (2004) 84–89.
- [31] D. Shastri, A. Merla, P. Tsiamyrtzis, I. Pavlidis, Imaging facial signs of neurophysiological responses, *IEEE Trans. Biomed. Eng.* 56 (2009) 477–484.
- [32] M. Wacker, H. Witte, Time-Frequency techniques in biomedical signal analysis: a tutorial review of similarities and differences, *Methods Inf. Med.* 52 (2013) 279–296.