Computers in Biology and Medicine 53 (2014) 154–163

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

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.

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).

http://dx.doi.org/10.1016/j.compbiomed.2014.07.014
0010-4825/

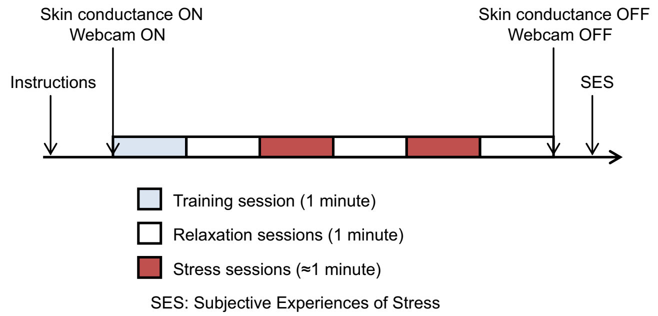

Fig. 1. Protocol of the experiment, composed by three stress sessions and three relax sessions.

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)].