Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate

Frederic Bousefsaf, Choubeila Maaoui, Alain Pruski

▶ **To cite this version:**

Frederic Bousefsaf, Choubeila Maaoui, Alain 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. 10.1016/j.bspc.2013.05.010 . hal-01323841

HAL Id: hal-01323841
https://hal.science/hal-01323841

Submitted on 20 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate

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.

* Corresponding author. Tel.: +33 387 547 652; fax.: +33 387 547 301
E-mail addresses: frederic.bousefsaf@univ-lorraine.fr (F. Bousefsaf), choubeila.maaoui@univ-lorraine.fr (C. Maaoui), alain.pruski@univ-lorraine.fr (A. Pruski).

Color figures are intended to be reproduced in color on the Web and in black-and-white in print.

Abstract
Photoplethysmographic signals obtained from a webcam are analyzed through a continuous wavelet transform to assess the instantaneous heart rate. The measurements are performed on human faces. Robust image and signal processing are introduced to overcome drawbacks induced by light and motion artifacts. In addition, the respiration signal is recovered using the heart rate series by respiratory sinus arrhythmia, the natural variation in heart rate driven by the respiration. The presented algorithms are implemented on a mid-range computer and the overall method works in real-time. The performance of the proposed heart and breathing rates assessment method was evaluated using approved contact probes on a set of 12 healthy subjects. Results show high degrees of correlation between physiological measurements even in the presence of motion. This paper provides a motion-tolerant method that remotely measures the instantaneous heart and breathing rates. These parameters are particularly used in telemedicine and affective computing, where the heart rate variability analysis can provide an index of the autonomic nervous system.

Keywords
Instantaneous Heart Rate
Continuous wavelet transform
Non-contact
Respiratory sinus arrhythmia
Photoplethysmography
Webcam

1. 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 in virtual therapies, where emotions are feedbacks that regulate the virtual environment level and intensity. Physiological parameters like Heart Rate (HR) and Heart Rate Variability (HRV) are reliable inputs to emotion recognition [1–3]. However, contact sensors can be limited in some scopes of application where a specialist must install and monitor them. When dealing with serious games, contact sensors can disturb the interaction and may be intrusive to the privacy.

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 like heart and respiratory rates, perspiration, supraorbital and periorbital blood flow [4–7]. Similarly, Doppler radars are non-contact sensors that were used to detect heartbeats [8,9] and respiration signals [10]. More recently, digital cameras and webcams were employed on the face to detect Blood Volume Pulse (BVP) [11–15] and compute HR and Breathing Rate (BR). The principle, based on PhotoPlethysmoGraphy (PPG) consists in observing light variations on the skin to recover the cardiovascular pulse wave. This optical technique is mainly implemented in contact pulse oximeters where infrared wavelengths are employed to detect the pulse wave. Considered in this case as noise, ambient light is now an illumination source used for PPG exploitation via high sensitivity cameras and webcams. The main drawback of this technique is that PPG signals are susceptible to motion-induced artifacts, particularly when dealing with webcams and ambient light. Independent component analysis, a blind source separation method, has been proposed by Poh et al. [13] to remove noise artifacts from face imaging PPG signal. Standards of measurement recommend the use of ECG sensors to measure HRV [16]. However, it has been shown that pulse rate variability derived from PPG signals can be a good surrogate of HRV at rest [17]. Sun et al. [14,15] 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 [15].

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 [19]. Its spectral analysis can provide the sympathovagal balance, a ratio that reflects reciprocal changes of sympathetic and vagal outflows [18]. The HRV tends to be rhythmic and ordered in positive emotional states and follows the respiration by a phenomenon called Respiratory Sinus Arrhythmia (RSA). 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 [20,21]. The use of non-contact means to detect physiological signals is particularly advantageous in affective computing, where the objective is to induce emotions like stress or fear for example. In these psychophysiological experiments, contact sensors may generate a bias by interfering with the user, resulting practically by an erroneous emotion classification [4]. Currently, published methods effectively recover HR, BR and HRV spectral components over a given time period. However, few attempts [15] have been made to measure the instantaneous HR (iHR) with a webcam, especially when considering head motion artifacts.

The immediate objective of this study was to provide a motion-tolerant method that reliably recovers the instantaneous pulse and breathing rates using a low-cost webcam. The presented methodology was developed to overcome signal variations generated by natural head movements. Firstly, we describe the approach, where robust image and signal processing are introduced to gather exclusively skin pixels that contain PPG information. A wavelet filtering algorithm was elaborated to recover both the instantaneous HR and BR. Secondly, we validate the accuracy of the proposed approach using approved contact probes. Webcam PPG signals were remotely recorded from 12 healthy volunteers during a set of two experiments, specifically at rest and during motion. Remote measurements of the instantaneous heart and breathing rates were respectively compared to those acquired from BVP and chest belt sensors.

2. Methods

2.1. Experimental procedure

Two experiments were conducted indoors to evaluate the iHR assessment method on 12 healthy volunteers (Table 1) of both gender and various ages. The skin type reported in Table 1 corresponds to a visual estimation of the participant skin color using the Fitzpatrick chromatic scale [22], defined between I for white skins and V for black skins. For practical purposes, categories I and II were regrouped in one set. All participants gave their informed consent before the beginning of a session. Each experiment in a session lasted 35 seconds where participants were sitting on a chair in approximately 1 meter from the webcam (see Fig. 1). Sunlight intensities were previously recorded in all sessions as the experiments were performed on different moments of the day. Source signals [see Fig. 2 (e)] were recorded with a C++ based software and analyzed offline with MATLAB (The MathWorks, Inc.). The purpose of the first experiment was to evaluate the proposed iHR assessment method in quiet, still and calm conditions. In the second experiment, participants had to follow a set of three randomly predefined head movements, particularly two horizontal and one vertical rotations. The purpose of this second experiment was to evaluate the method under motion conditions. Head rotations were recorded using the bounding box of the face, delivered by the face detection algorithm. The iHR and the respiration signals were respectively recorded along both experiments using contact sensors (see Fig. 1).

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 320x240 pixels in order to keep an acquisition frequency of 30 frames per second. The maximum webcam resolution is 1280x800 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. The light intensity was measured with a light meter (model TENMA 72-6693 by TENMA) before each experiment. A finger BVP and a chest belt respiration sensor (BVP-Flex/Pro and Resp-Flex/Pro by Thought Technologies Ltd.) were used to measure respectively the HR and the BR at 256 Hz.

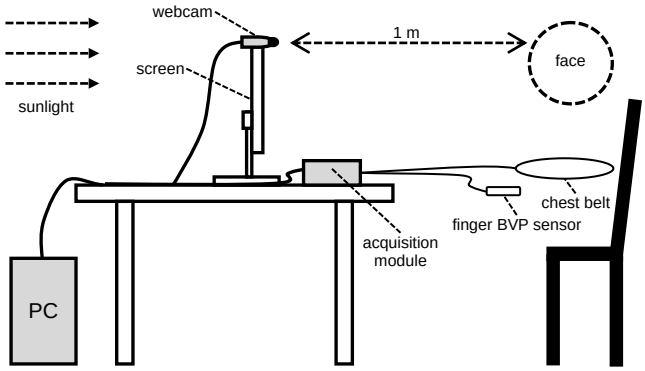

Fig. 1. Experimental setup. The subject is seated in front of a screen at a distance of approximately 1 m from the webcam.

2.3. Photoplethysmography pulse extraction

2.3.1. Overview

The overall system is composed with both image and signal processing. Changing the native resolution of the webcam by working with HD frames results in an increase of pixels that contain the signal. Given that HD acquisition requires more computational time, Pan, Tilt and Zoom parameters were employed to provide a partial frame of interest from the global HD frame. 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 in the input frame [Fig. 2 (a)]. The method presented in this study run on a mid-range computer in real time. Preprocessing operations are applied on the original frame to isolate skin pixels that contain true PPG information. Thus, to properly collect PPG pixels, a skin detection mask [Fig. 2 (b)] is applied on the input frame. A spatial averaging operation is computed on these segmented pixels while the others are discarded. A set of *t* frames gives a raw signal of *t* points [Fig. 2 (e)]. A continuous wavelet transform [Fig. 3 (b)] is performed on this signal to remove trends and noise. Then, the maxima are detected and the interbeat intervals computed, forming the iHR signal. The respiration is reconstructed (see Fig. 5) using another wavelet filtering operation from the interpolated iHR signal.

Subject n°	Gender	Age (years)	Skin Type[a]	HR at rest	Light [b] intensity (lux)
1	M	24	IV	73	1400 / 1400
2	M	25	IV	60	1500 / 1800
3	M	22	I – II	47	1150 / 1150
4	M	27	I – II	64	1250 / 1250
5	M	23	I – II	76	100 / 100
6	M	26	IV	62	300 / 300
7	F	22	III	80	1800 / 120
8	M	24	I – II	71	300 / 350
9	M	79	I – II	50	2000 / 2000
10	M	22	IV	76	300 / 400
11	M	25	I – II	78	1400 / 1500
12	F	22	III	80	250 / 200
Mean ± S.D.		28 ± 16		68 ± 11	979 ± 683 / 881 ± 705

[a] Using Fitzpatrick's chromatic scale [22].
[b] At rest / During motion

Table 1. Subjects and illumination characteristics

2.3.2. Image processing

The face is automatically detected using a cascade of boosted classifier on each frame. The method was originally proposed by Viola and Jones [23] and is employed using OpenCV 2.1 library. The algorithm returns the bounding box of the face and particularly the top-left point coordinates along with the width and height of the box. The PTZ parameters are computed using these bounding box coordinates. An iterative algorithm is applied where the zoom is continuously adjusted according to the surface of the bounding box [Fig. 2 (a)]. Pan and tilt parameters are regulated to track the face over time when the head of the user moves. A skin detection [Fig. 2 (b)] is performed on the face to detect and gather skin pixels that contain the PPG signal. The filter is established in the YC_bC_r color space [24] by setting a threshold on the 3 channels as follows:

$$\begin{cases} Y > 80 \\ 77 < C_b < 127 \\ 133 < C_r < 173 \end{cases} \quad (1)$$

Where Y is the luma, C_b the blue chroma and C_r the red chroma of the color space. It has already been used to detect skin color on pictures or videos [24].

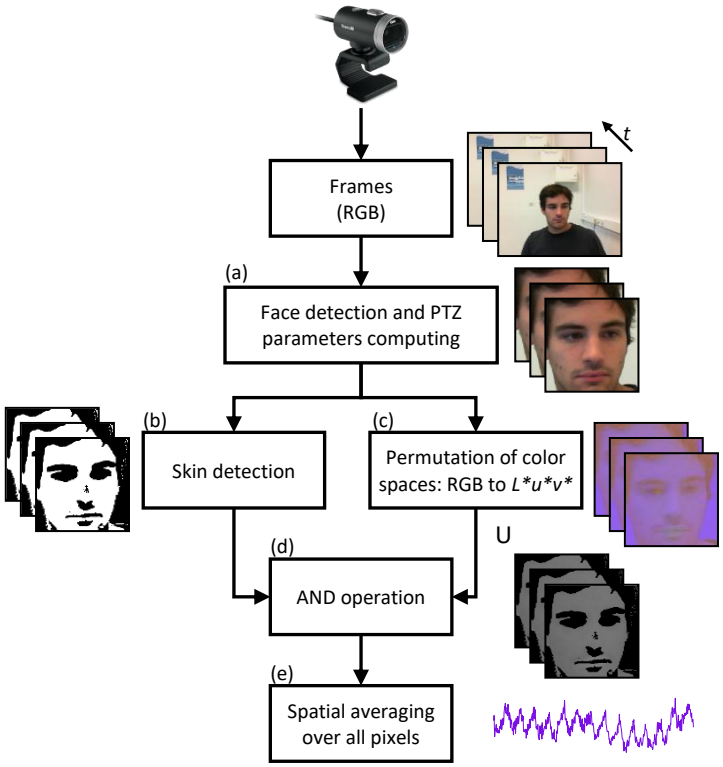

Fig. 2. Processing algorithm overview. (a) Pan, Tilt and Zoom parameters are computed to zoom and track the face. (b) Pixels that contain PPG information are isolated by a skin detection. (c) The RGB color space is converted to the CIE *L*u*v**. (d) The *u** frame is combined with the skin detection frame by a combinational AND operation. (e) A spatial averaging step is performed to transform a set of frames into a single raw signal [11,12].

A permutation between the native RGB and the CIE *L*u*v** color spaces is performed concurrently to the skin detection. Oxy- and deoxyhemoglobin have a better absorption coefficient between 540 and 577 nm, corresponding to green/yellow colors wavelength [26]. Thus, PPG fluctuations are more influent in the *u** channel, which represents a red to green color indicator. The CIE *L*u*v** is a perceptual color space proposed by the International Commission on Illumination—the CIE abbreviation comes from its French denomination: Commission Internationale de l'Eclairage. The color space is device-independent, perceptually uniform and is indirectly computed from the CIE XYZ space, a derived version of the RGB color space [25], as in Eq. (4). In Eq. (2), *L** represents the lightness while *u** and *v** are the chromaticity components. The color space is non-linear but reversible. The distance between two points indicates how different the colors are in luminance, chroma and hue. The *L**, *u** and *v** coordinates are calculated from the CIE XYZ space by the following formulas:

$$L^* = \begin{cases} 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 & \text{for } \frac{Y}{Y_n} > 0.008856 \\ 903.292\left(\frac{Y}{Y_n}\right) & \text{for } \frac{Y}{Y_n} \leq 0.008856 \end{cases}$$

$$u^* = 13(L^*)(u' - u_n) \qquad (2)$$

$$v^* = 13(L^*)(v' - v_n)$$

Where the subscript *n* in Eq. (2) denotes the D65 [25] white point coordinates: Y_n = 1.0, u_n = 0.197939 and v_n = 0.468311. *u'* and *v'* are computed with the X, Y and Z coordinates:

$$\mathrm{u}' = \frac{4\mathrm{X}}{\mathrm{X} + 15\mathrm{Y} + 3\mathrm{Z}}$$

$$\mathrm{v}' = \frac{9\mathrm{Y}}{\mathrm{X} + 15\mathrm{Y} + 3\mathrm{Z}} \qquad (3)$$

The XYZ space is derived from the RGB color space by the following formula:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} . \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (4)$$

A spatial averaging operation [Fig. 2 (d) and (e)] is then computed using *u** pixel intensities that exist on the skin detection, forming the raw signal *x(t)*. The reversed version of this signal is used for further analysis.

2.3.3. Signal processing

2.3.3.1. Continuous Wavelet Transform (CWT) to filter the pulse wave

The CWT constructs a time-frequency representation of a signal and has been used to denoise or refine peaks [27] and significant points [28] in PPG signals. The CWT uses, like the Fourier Transform, inner products to measure the similarity between an analyzing function and a signal. Unlike the Fourier Transform and the short time Fourier Transform, the CWT can detect rapid changes in frequencies in time due to its variable window width. These advantages have resulted in CWT being increasingly used for biological signals analysis [7,29,30]. The non-stationary PPG signal is convolved with a child wavelet $\psi_{\tau,s}$, representing a scaled and shifted version of a mother wavelet ψ:

$$CWT_x^{\psi}(\tau, s) = \int_{-\infty}^{\infty} x(t)\psi_{\tau,s}(t)\,dt$$
$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{|s|}}\psi\left(\frac{t-\tau}{s}\right) \quad (5)$$

Where $\psi_{\tau,s}$ is the child wavelet, scaled by *s* and dilated by τ. ψ is the referent mother wavelet. Analyzed frequencies are linked to the scale parameter. Increasing *s* and dilating the wavelet size results to cover a larger signal in the time domain, leading to analyze lower frequencies. A large set of standard mother wavelets is available in the literature, and the choice depends on the application requirements and on the input signal properties. The Morlet wavelet has already been used to analyze PPG signals [30] and was employed in this study. The original signal can be reconstructed from the wavelet transform via the following inverse formula:

$$x(t) = \frac{1}{C_\psi}\int_0^{\infty}\int_{-\infty}^{\infty}\frac{1}{s^2}CWT_x^{\psi}(\tau, s)\frac{1}{\sqrt{|s|}}\psi\left(\frac{t-\tau}{s}\right)d\tau\,ds$$
$$C_\psi = \int_0^{\infty}\frac{|\hat{\psi}(\zeta)|^2}{|\zeta|}d\zeta < \infty \quad (6)$$

C_ψ is the admissibility condition and $\hat{\psi}$ the Fourier transform of ψ.

The DC component is removed [see Fig. 3 (a)] to reveal detailed information [7] on lower scales prior to performing the CWT:

$$x_{AC}(t) = x(t) - \delta \quad (7)$$

Where δ is the mean of the raw PPG signal *x(t)*. The CWT is then computed within an operational frequency band, set to [0.65, 3] Hz corresponding to 40–180 bpm [Fig. 3 (b)].

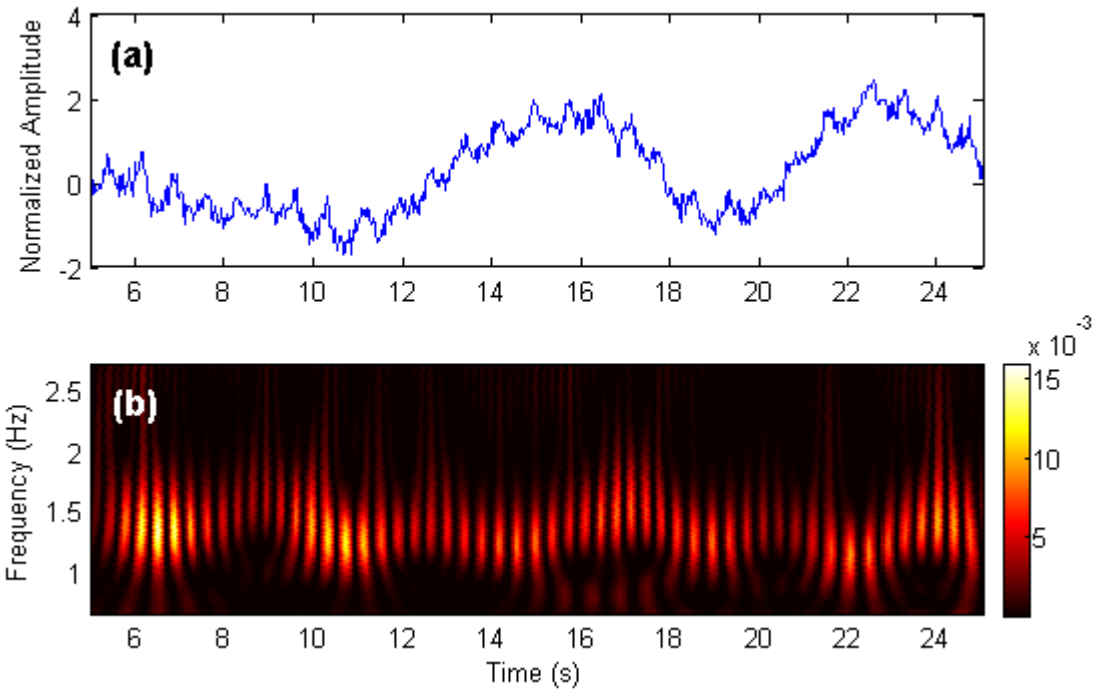

Fig. 3. A fragment of the raw signal (a) of the subject 12 (female, age = 22, skin type III) and its CWT computed between 0.65 and 3 Hz. This spectrogram (b) is a particular plot that represents the percentage energy for each coefficient, where pulse wave oscillations are distinctly identifiable around 1.5 Hz.

The wavelet transform coefficients are filtered using the wavelet energy curve [see Fig. 4 (a)], its maxima in the frequency axis corresponding to the averaged heart rate. The pulse wave presents stronger amplitudes than those generated by noise and trends. A weighted product is applied between the energy curve and the CWT coefficients [see Fig. 4 (b)]. The inverse transform is computed as in Eq. (6) to reconstruct a denoised and detrended version of the raw signal [see Fig. 4 (c)]. The cascade of these two operations—weighting the CWT and reconstructing the signal by inverse CWT—is employed to filter the PPG signal in the operational band [0.65, 3] Hz. The filtered signal is interpolated with a 256 Hz cubic spline function. A custom algorithm was developed to detect peaks and compute the interbeat intervals (IBIs). The iHR series derived from the webcam PPG recordings are computed using the IBI time series [Fig. 5 (a)].

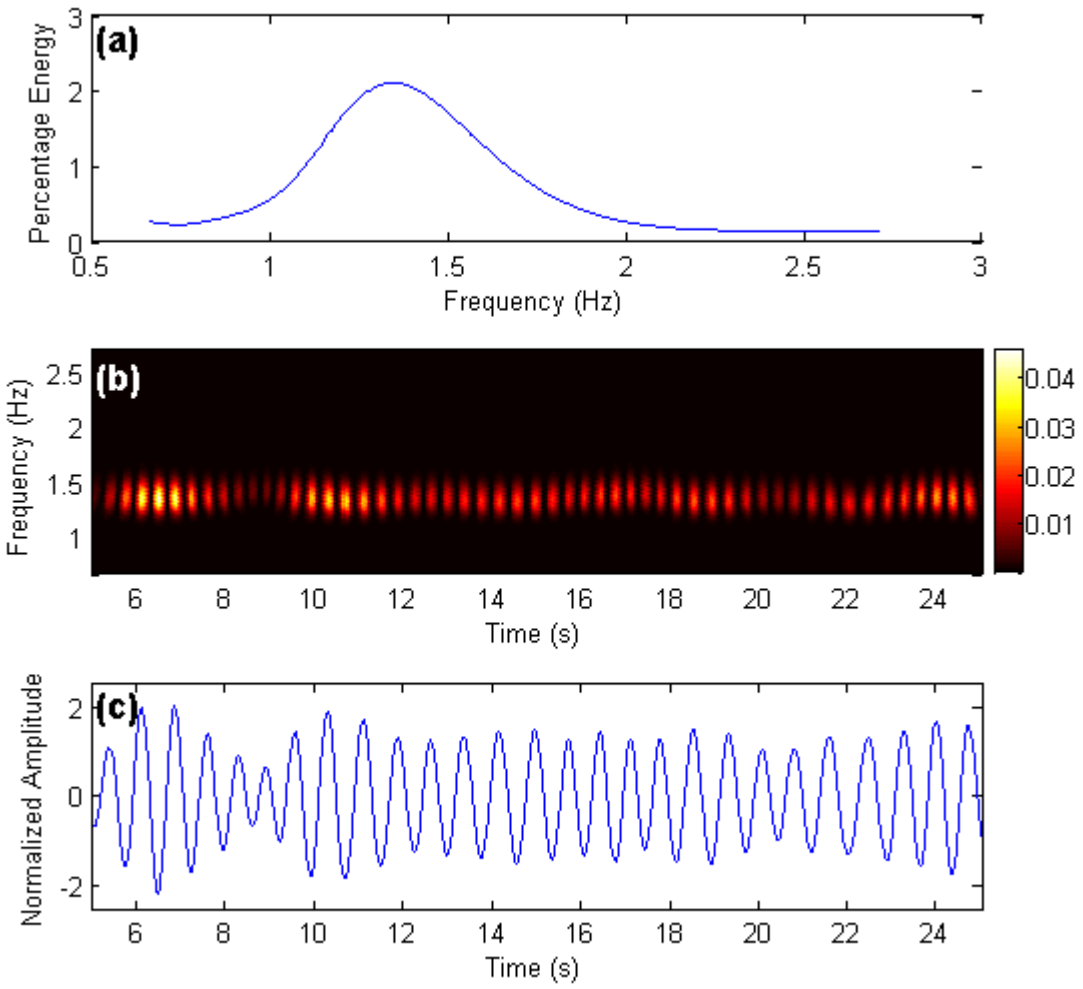

Fig. 4. The energy curve (a) is used to filter the spectrogram (b). The reconstruction gives a detrended and denoised version (c) of the raw signal.

2.3.3.2. Respiration signal assessment

A 30 Hz linear interpolation is performed on the irregularly time-sampled iHR series and a new CWT [Fig. 5 (b)] is performed between 0.15 and 0.4 Hz. This frequency band is widely used by HRV analysis communities and gives an image of the RSA [18]. The respiration is in close range with the HRV by RSA on rested persons and is equally recovered [Fig. 5 (c)] using the processing algorithms presented in section 2.3.3.1, in the operational band [0.15, 0.4] Hz. The instantaneous BR is computed by detecting respiratory cycles and measuring intercycles intervals on the respiration signal.

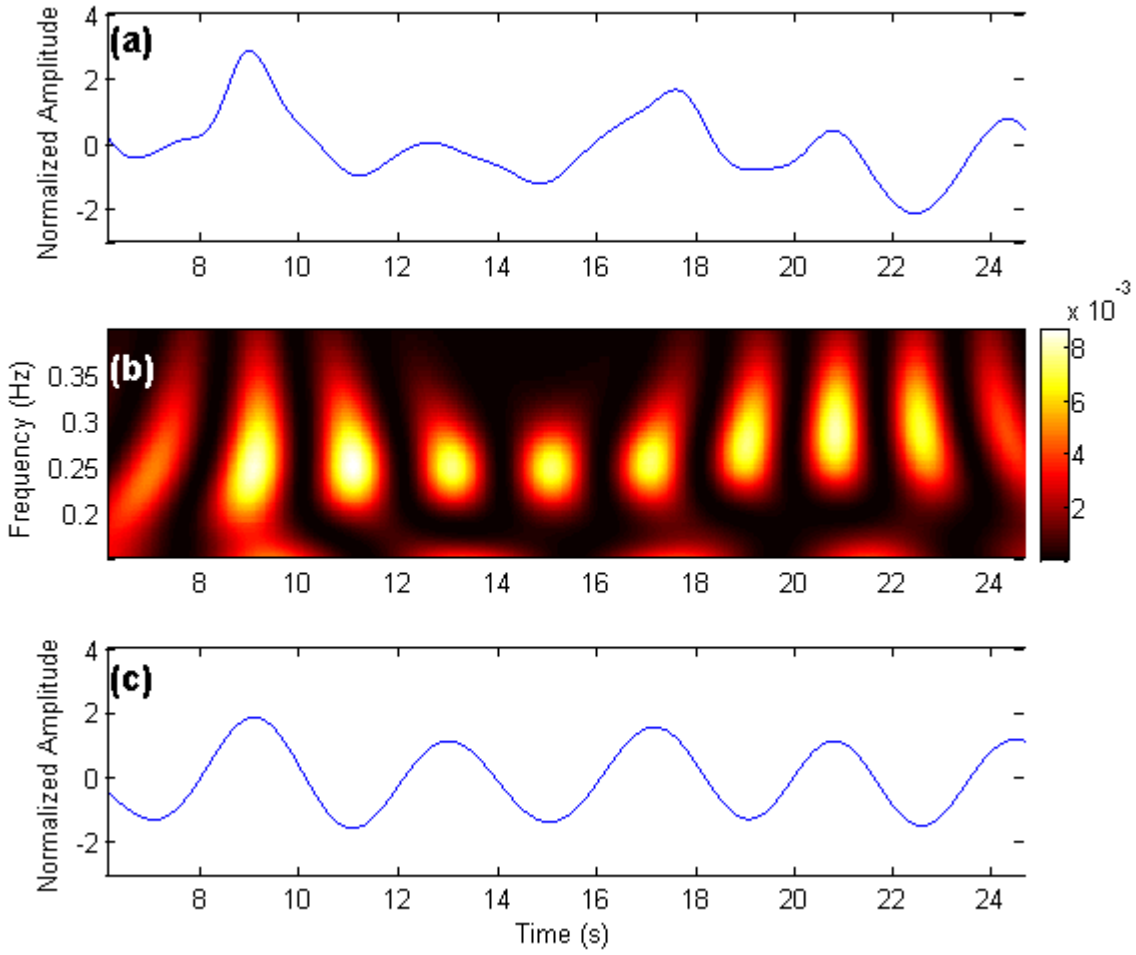

Fig. 5. The iHR series (a) is resampled and a new CWT is computed between 0.15 and 0.4 Hz (b). The weighted reconstruction gives the respiration signal (c).

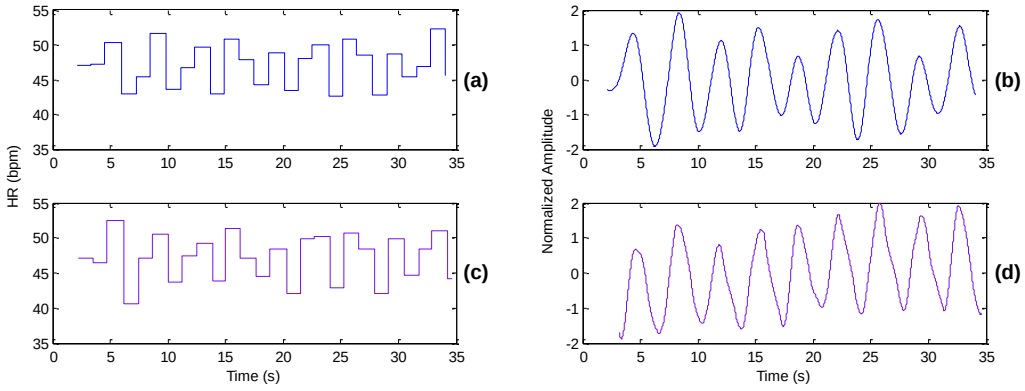

Fig. 6. A typical example representing the iHR series computed with the webcam PPG signal (a) and the corresponding respiration signal (b) extracted from the iHR time trace. The iHR series recorded with the finger BVP sensor (c) and the respiration signal measured by the chest belt sensor (d) are in close agreement with readings from the webcam.

3. Results

Image and signal processing detailed in the previous section were employed to compute the iHR series and respiration traces of all subjects. These two physiological parameters were simultaneously recorded by contact sensors. A typical example from subject 3 (male, age = 22 years and skin type = I–II) is presented in Fig. 6 where respective signals are directly comparable. Pearson's correlation coefficients along with a statistical analysis and Bland-Altman plots were used to quantify the level of agreement between physiological measurements by the remote and contact techniques. Results of the analysis are summarized in Table 2, where performances between motion and non-motion tests are revealed. HR_{mean} and BR_{mean} are obtained by averaging all values on the instantaneous HR and BR traces respectively. Similarly, HR_{SD} and BR_{SD} are the sample Standard Deviation (SD) of the instantaneous HR and BR series respectively.

At rest	contact sensors Mean ± SD	webcam Mean ± SD	Correlation
HR_{mean}	68.28 ± 11.38	68.26 ± 11.35	1.00
HR_{SD}	4.55 ± 1.96	4.38 ± 1.90	0.98
BR_{mean}	16.40 ± 4.14	16.21 ± 3.84	0.99
BR_{SD}	1.76 ± 0.98	2.34 ± 1.23	0.81
During motion			
HR_{mean}	66.53 ± 11.63	66.54 ± 11.62	1.00
HR_{SD}	4.31 ± 1.44	4.6 ± 1.44	0.94
BR_{mean}	16.71 ± 2.64	16.71 ± 2.79	0.98
BR_{SD}	2.21 ± 0.94	2.64 ± 0.69	0.46

HR_{mean} is the average over all heart rate values on the instantaneous HR signal and HR_{SD} his standard deviation, both expressed in bpm. Similarly, BR_{mean} and BR_{SD} are respectively the average and the SD of the instantaneous BR signal. Both are expressed in breaths/min

Table 2. Results of the HR and BR statistical analysis for experiments at rest and during motion.

HR series were strongly correlated across statistical parameters, exhibiting $r = 1.00$ for means and $r > 0.9$ for standard deviations, independently of motion ($p < 0.001$ in all cases). Moreover, a beat to beat Bland-Altman analysis was employed to compare the two measurement techniques. The differences between estimates from contact and remote measurements were plotted against the averages of both systems for respectively HR and BR at rest [Fig. 7 (a) and (c)] and during motion [Fig. 7 (b) and (d)]. Means are represented by dash-dot lines and 95% limits of agreement (± 1.96 SD) by dashed lines on plots from the figure 7. Specifically, the mean biases were −0.03 bpm with 95% limits of agreement −4.22 to 4.16 bpm for the HR and −0.27 breaths/min with 95% limits of agreement −5.36 to 4.82 breaths/min for the BR at rest. During motion, the biases were 0.02 bpm and −0.01 breaths/min for respectively HR and BR. The corresponding 95% confidence intervals were −4.96 to 4.99 bpm and −4.6 to 4.58 breaths/min. In addition the beat to beat Root Mean Square Error (RMSE) and Pearson's correlation coefficients were computed between iHR traces acquired from webcam [Fig. 6 (a)] and reference [Fig. 6 (c)] sensors. Results are summarized in Table 3, exhibiting strong correlation between measurement methods, where $r > 0.85$ ($p < 0.001$) independently of motion. The RMSE between both HR readings were 1.97 and 2.33 bpm for experiments at rest and during motion respectively.

	At rest Mean ± SD	During motion Mean ± SD
RMSE	1.97 ± 0.62	2.33 ± 0.73
r	0.889 ± 0.046	0.853 ± 0.056

The RMSE is expressed in bpm. *r* represents the Pearson's correlation coefficient. Both are computed beat to beat between the HR traces of the webcam and the reference sensor.

Table 3. Beat to beat RMSE and correlation results of the instantaneous HR series

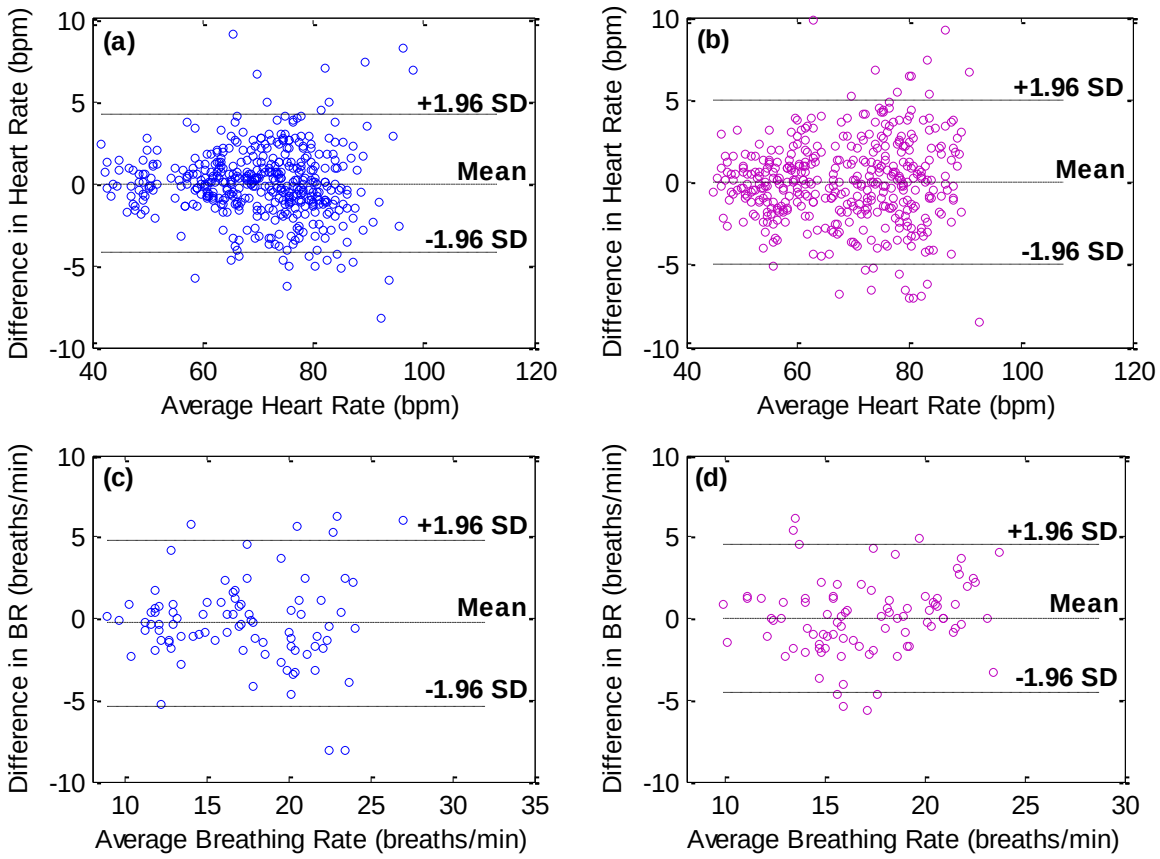

Fig. 7. Beat to beat Bland-Altman plots showing the differences in HR at rest (a) and during motion (b) and the differences in BR at rest (c) and during motion (d) plotted against respective averages of both measurement techniques. 95% limits of agreement are represented by dashed lines.

4. Discussion

The HRV is a physiological measurement used in several domains, especially in affective computing and personal health care. The present study demonstrates that the instantaneous HR can be assessed robustly using a low-cost HD webcam on human faces, even in presence of motion. Statistical and beat to beat analysis reveals that webcam-derived HR and BR are in close agreement with reference sensors. The respiration signal is recovered using the method presented in section 2.3.3.2. The breathing rate, computed by averaging all value in BR readings, is strongly correlated across measurement means. Nevertheless, these results demonstrate that BR measurements are in

closer agreements with the chest belt sensor during no motion tests, where the RSA phenomenon is generally more pronounced, facilitating the respiration assessment.

The skin detection (see Fig. 8 for a typical example) is particularly effective during head rotations, where undesired pixels (hair, background) are going to add noise in raw signals. The eyes are only considered during their closed position, when skin pixels are visible. During the motion experiments, subjects were asked to turn their heads following a set of predefined rotation: approximately 30° horizontally and 20° vertically. These values were defined to avoid the OpenCV face tracker to lose the face. In addition, the skin filter is built to discard pixel values that are too dark or bright, containing probably noisy information. The skin detection is sensitive to illumination parameters and an adaptive model [31] should be developed in the future to increase the robustness of the system. The *u** component of the CIE *L*u*v** color space was used in the presented study to reduce inherent effects due to light variation or head movements. A typical example is presented in Fig. 9 where strong fluctuations due to head movements appear at 4, 12 and 17 seconds. These fluctuations impact the G [Fig. 9 (b)] channel when using the native RGB space but are non-existent on the *u** component [Fig. 9 (c)] of the CIE *L*u*v** color space.

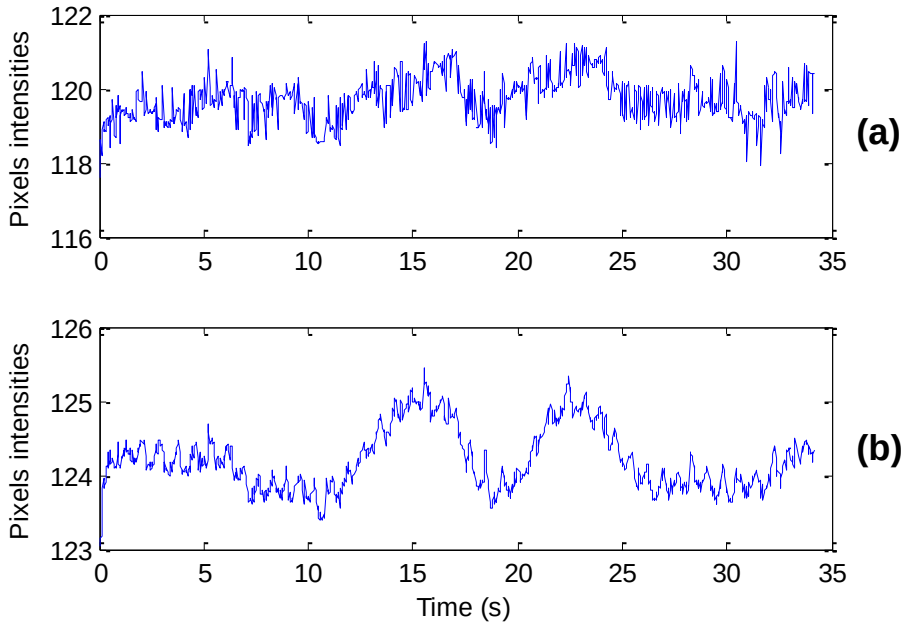

Fig. 8. The instantaneous HR cannot be properly assessed using the PPG signal (a) recorded without the skin filter. The cardiac pulse is visually identifiable when the skin detection is enabled (b). These signals are two green channel recordings from subject 12 (female, age = 22, skin type III).

Vision-based systems are particularly sensitive to the illumination source position. Distance, skin color and light conditions are parameters that can degrade raw signals. For example, if the source is sunlight coming through a window perpendicularly to the subject, all head rotations are going to generate strong fluctuations in the PPG signal by impacting sub-regions of pixel values due to shadow and edge variations. Illumination parameters, like propagation, type—natural and/or artificial—location and diffusion must be considered in these optophysiological systems.

The auto-exposure should be enabled to deliver clear and bright input images and improves the robustness of the skin detection filter presented in this paper. On the contrary, white-balance that locally regulates pixel colors must be disabled and manually controlled to avoid additional random noises in signals. The number of pixels that composes the face relies on the webcam resolution over a given distance. Therefore, the PTZ algorithm presented in this study was implemented to acquire HD frames while preserving an operational frame rate of 30 fps. All these preprocessing operations are used to refine input signals to properly and precisely assess the iHR.

The ANS continuously regulates the heart rate by affecting the sinoatrial node with both sympathetic and parasympathetic fibers. When breathing, the RSA will increase and then decrease the heart rate by modulating the parasympathetic nervous system [18]. Thus, the RSA gives a measure of the parasympathetic component and is less pronounced on seniors or people who suffer from cardiovascular diseases. It has been widely used to measure the cardiac vagal control in response to stress stimuli [32]. Our results indicate that the respiration signal assessment is more effective on rested and calm subjects than participants who are performing controlled motion tests. Activation of the sympathetic nervous system leads RSA to disappear [20]. Participants who performed the motion experiment were probably affected more by stress than during the static experiment.

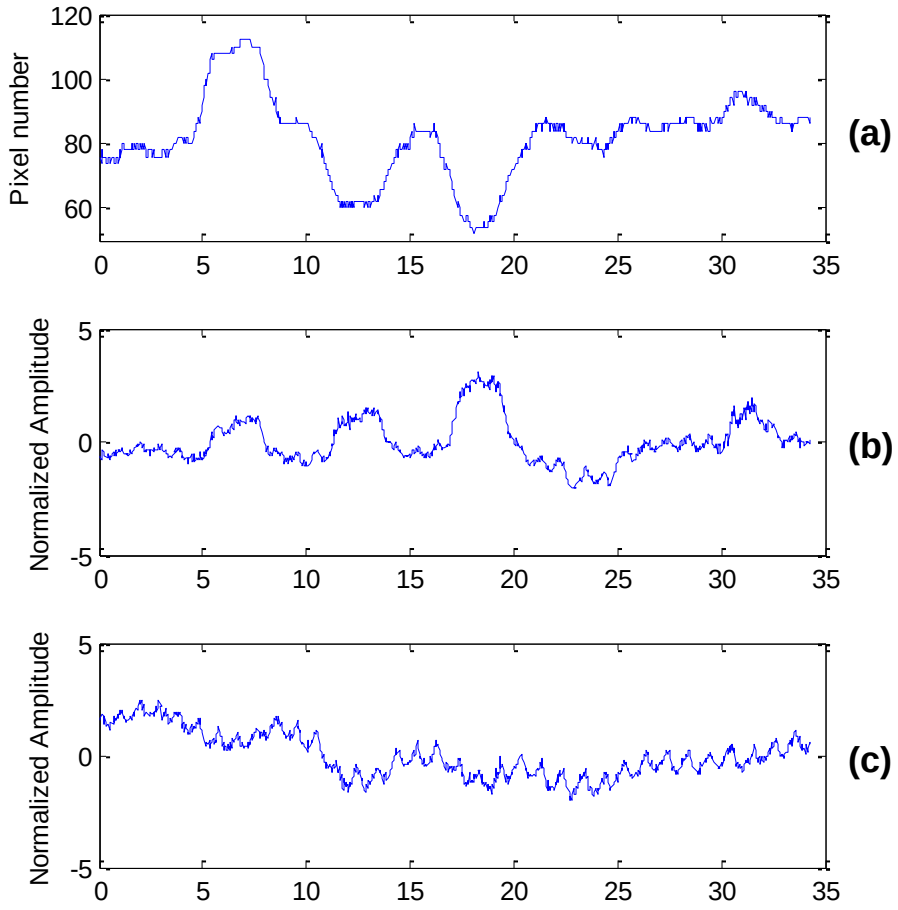

Fig. 9. The movements of the head are recorded using the x-coordinate of the face tracker (a) and generate fluctuations in the G component (b). These signals are from subject 6 (male, age = 26, skin

type IV).The noise due to movements is strongly attenuated using the *u** component (c) of the CIE *L*u*v** color space.

5. Conclusion

This study presents image and signal processing techniques to remotely assess the instantaneous HR. Using a skin detection filter, the proposed method selects only skin pixels that contain PPG information. The *u** component of the CIE *L*u*v** color space is used to increase robustness on motion and light variations. A typical example is presented in Fig 9. A wavelet-based filtering operation is then applied to detrend and denoise raw signals. It has been shown that trends in PPG signals must be removed in order to compute the power spectra density of the HRV [33]. The instantaneous HR trace is assessed after detecting peaks and computing IBIs. Results from both the static and motion experiments show that measurements acquired using the webcam are closely comparable to those from the reference sensors. The shape of the HR series has to be properly recovered to apply time domain algorithms that used the HRV outline in order to detect stress [20]. This study put emphasis on instantaneous HR assessment by a webcam in the time domain that is motion-tolerant. A spectral analysis of the HRV is conceivable to compute sympathovagal index, an estimate that gives the sympathetic versus parasympathetic balance [21].

References

[1] L. Li, J.-H. Chen, Emotion recognition using physiological signals, Lecture Notes in Computer Science 4282 (2006) 437–446.

[2] C.Maaoui, A. Pruski, Emotion recognition through physiological signals for human-machine communication, Cutting Edge Robotics 2010 (2010) Available: http://www.intechopen.com/books/cutting-edge-robotics-2010/emotion-recognition-through-physiological-signals-for-human-machine-communication

[3] S. Jerritta, M. Murugappan, R. Nagarajan, 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, C. Puri, J. Fei, M. Garbey, Interacting with human physiology, Computer Vision and Image Understanding 108 (2007) 150–170.

[5] J. Fei, I. Pavlidis, Thermistor at a distance: Unobtrusive measurement of breathing, IEEE Trans. Biomed. Eng.57 (2010) 988–998.

[6] I. Pavlidis, J. Levine, Thermal image analysis for polygraph testing, IEEE Eng. Med. Biol. Mag. 21 (2002) 56–64.

[7] D. Shastri, A. Merla, P.Tsiamyrtzis, I. Pavlidis, Imaging facial signs of neurophysiological responses, IEEE Trans. Biomed. Eng. 56 (2009) 477–484.

[8] G. Lu, F. Yang, X. Jing, J. Wang, Contact-free measurement of heartbeat signal via a Doppler radar using adaptive filtering, in: Int. Conf. Image and Signal Process., Xiamen, China, Apr. 2010, pp. 89–92.

[9] V. Vasu, C. Heneghan, T. Arumugam, 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.

[10] C. Gu, R. Li, C. Li, 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.

[11] C. Takano, Y. Ohta, Heart rate measurement based on a time-lapse image, Med. Eng. Phys. 29 (2007) 853–857.

[12] W. Verkruysse, L. O. Svaasand, J. S. Nelson, Remote plethysmographic imaging using ambient light, Opt. Expr. 16 (2008) 21434–21445.

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

[14] Y. Sun, C. Papin, V. Azorin-Peris, R. Kalawsky, S. Greenwald, S. Hu, Use of ambient light in remote photoplethysmographic systems: comparison between a high-performance camera and a low-cost webcam, J. Biomed. Opt. 17 (2012) 037005.

[15] Y. Sun, S. Hu, V. Azorin-Peris, R. Kalawsky, S. Greenwald, Noncontact imaging photoplethysmography to effectively access pulse rate variability, J. Biomed. Opt. 18 (2013) 061205.

[16] 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) 1043–65.

[17] 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.

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

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

[20] D. Servant, R. Logier, Y. Mouster, M. Goudemand, Heart rate variability. Applications in psychiatry (La variabilité de la fréquence cardiaque. Intérêts en psychiatrie), L'Encéphale 35 (2009) 423–428.

[21] 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.
[22] T. B. Fitzpatrick, Sun and skin (Soleil et peau), J. Med. Esthétique 2 (1975) 33–34.
[23] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Kauai, HI, Dec. 2001, pp. I-511–I-518.
[24] T. M. Mahmoud, A new fast skin color detection technique, World Academy Science Eng. Technol. 19 (2008) 498–502.
[25] C. Poynton, The CIE system of colorimetry, in: Digital Video and HDTV: Algorithms and Interfaces, Morgan Kaufmann Publishers Inc., San Francisco, 2003, pp. 211–231.
[26] B. L. Horecker, The absorption spectra of hemoglobin and its derivatives in visible and near infrared regions, J. Biol. Chem. 148 (1943) 173–183.
[27] S. Soni, Y. Namjoshi, Delineation of raw plethysmograph using wavelets for mobile based pulse oximeters, in: Proc. of 5th Innovative Conf. Embedded Syst. Mobile Commun. Comput.,Kuppam, India, Jul. 2010, pp. 74–84.
[28] T. Peterek, M. Prauzek, M. Penhaker, A new method for identification of the significant point in the plethysmografical record, in: Proc. 2nd Int. Conf. Signal Process. Syst., Dalian, China, Jul. 2010, pp. VI-362–VI-364.
[29] P. Leonard, T.F. Beattie, P.S. Addison, J.N. Watson, Standard pulse oximeters can be used to monitor respiratory rate, Emerg. Med. J. 20 (2003) 524-525.
[30] P. S. Addison, J. N. Watson, A novel time–frequency-based 3D Lissajous figure method and its application to the determination of oxygen saturation from the photoplethysmogram, Meas. Sci. Technol. 15 (2004) L15–L18.
[31] M. Soriano, B. Martinkauppi, S. Huovinen, M. Laaksonen, Adaptive skin color modeling using the skin locus for selecting training pixels, Pattern Recognition 36 (2003) 681–690.
[32] J. D. Lane, R. A. Adcock, R. E. Burnett, Respiratory sinus arrhythmia and cardiovascular responses to stress, Psychophysiology 29 (1992) 461–470.
[33] S. A. Akar, S. Kara, F. Latifoglu, V. Bilgiç, Spectral analysis of photoplethysmographic signals: The importance of preprocessing, Biomedical Signal Processing and Control 8 (2013) 16–22.