Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam

Ming-Zher Poh*, Daniel J. McDuff, and Rosalind W. Picard

***Abstract*—We present a simple, low-cost method for measuring multiple physiological parameters using a basic webcam. By applying independent component analysis on the color channels in video recordings, we extracted the blood volume pulse from the facial regions. Heart rate (HR), respiratory rate, and HR variability (HRV, an index for cardiac autonomic activity) were subsequently quantified and compared to corresponding measurements using Food and Drug Administration-approved sensors. High degrees of agreement were achieved between the measurements across all physiological parameters. This technology has significant potential for advancing personal health care and telemedicine.**

***Index Terms*—Autonomic nervous system, blood volume pulse (BVP), heart rate variability (HRV), independent component analysis (ICA), noncontact, photoplethysmography (PPG), remote sensing, respiration.**

I. Introduction

THE OPTION of monitoring a patient's physiological signals via a remote, noncontact means has promise for improving access to and enhancing the delivery of primary healthcare. Currently, proposed solutions for noncontact measurement of vital signs, such as heart rate (HR) and respiratory rate (RR), include laser Doppler [1], microwave Doppler radar [2], and thermal imaging [3], [4]. Noncontact assessment of HR variability (HRV), an index of cardiac autonomic activity [5], presents a greater challenge and few attempts have been made [6]–[8]. Despite these impressive advancements, a common drawback of the aforementioned proposals is that the systems are expensive and require specialist hardware.

Photoplethysmography (PPG) is a low-cost and noninvasive means of sensing the cardiovascular blood volume pulse (BVP) through variations in transmitted or reflected light [9]. Although PPG is typically implemented using dedicated light sources (e.g., red and/or infrared wavelengths), Verkruysse *et al.* showed that pulse measurements from the human face are attainable with normal ambient light as the illumination source [10]. However, the study lacked rigorous physiological and mathematical models amenable to computation; it relied instead on manual segmentation and heuristic interpretation of raw images with minimal validation of performance characteristics.

Recently, we developed a robust method for automated computation of HR from digital color video recordings of the human face [11]. In this letter, we extend this methodology to quantify multiple physiological parameters. Specifically, we extract the BVP for computation of HR, RR, as well as HRV. To the best of our knowledge, this is the first demonstration of a simple, low-cost method for noncontact HRV measurements.

II. Theory

Independent component analysis (ICA) is a relatively new technique for uncovering independent signals from a set of observations that are composed of linear mixtures of the underlying sources [12]. The underlying source signal of interest in this study is the BVP that propagates throughout the body. During the cardiac cycle, volumetric changes in the facial blood vessels modify the path length of the incident ambient light such that the subsequent changes in amount of reflected light indicate the timing of cardiovascular events. By recording a video of the facial region with a webcam, the red, green, and blue (RGB) color sensors pick up a mixture of the reflected plethysmographic signal along with other sources of fluctuations in light due to artifacts. Given that hemoglobin absorptivity differs across the visible and near-infrared spectral range [13], each color sensor records a mixture of the original source signals with slightly different weights. These observed signals from the RGB color sensors are denoted by $y_1(t)$, $y_2(t)$, and $y_3(t)$, respectively, which are the amplitudes of the recorded signals at time point t. We assume three underlying source signals, represented by $x_1(t)$, $x_2(t)$, and $x_3(t)$. The ICA model assumes that the observed signals are linear mixtures of the sources, i.e., ok but why 3 source signals?

$$\mathbf{y}(t) = \mathbf{A}\mathbf{x}(t) \tag{1}$$

where the column vectors $\mathbf{y}(t) = [y_1(t), y_2(t), y_3(t)]^T$, $\mathbf{x}(t) = [x_1(t), x_2(t), x_3(t)]^T$, and the square 3×3 matrix $\mathbf{A}$ contains the mixture coefficients a_{ij}. The aim of ICA is to find a demixing matrix $\mathbf{W}$ that is an approximation of the inverse of the original mixing matrix $\mathbf{A}$ whose output $W \approx A^{-1}$

$$\hat{\mathbf{x}}(t) = \mathbf{W}\mathbf{y}(t) \tag{2}$$

is an estimate of the vector $\mathbf{x}(t)$ containing the underlying source signals. To uncover the independent sources, $\mathbf{W}$ must maximize

Manuscript received July 15, 2010; revised September 15, 2010; accepted September 28, 2010. Date of publication October 14, 2010; date of current version December 17, 2010. This work was supported by the Nancy Lurie Marks Family Foundation and by a Center for Integration of Medicine and Innovative Technology (CIMIT) award. *Asterisk indicates corresponding author.*

*M.-Z. Poh is with the Harvard-MIT Division of Health Sciences and Technology (HST), Cambridge, MA 02139 USA, and also with the Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: zher@mit.edu).

D. J. McDuff and R. W. Picard are with the Media Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: picard@media.mit.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2010.2086456

the non-Gaussianity of each source. In practice, iterative methods are used to maximize or minimize a given cost function that measures non-Gaussianity.

III. Methods

A. Experimental Procedure

This study was approved by the Institutional Review Board, Massachusetts Institute of Technology. Our sample featured 12 participants of both genders (four females), different ages (18–31 years) and skin color. All the participants provided their informed consent. The experiments were conducted indoors and with a varying amount of ambient sunlight entering through windows as the only source of illumination. Participants were seated at a table in front of a laptop at a distance of approximately 0.5 m from the built-in webcam (iSight camera). During the experiment, participants were asked to keep still, breathe spontaneously, and face the webcam while their video was recorded for one minute. All videos were recorded in color (24-bit RGB with three channels $\times$ 8 bits/channel) at 15 frames per second (fps) with pixel resolution of 640 $\times$ 480 and saved in AVI format on the laptop. We also recorded their BVP and spontaneous breathing using an FDA-approved finger BVP sensor and chest belt respiration sensor (Flexcomp Infiniti by Thought Technologies Ltd.), respectively at a sampling rate of 256 Hz.

B. Recovery of BVP from Webcam Recordings

All the video and physiological recordings were analyzed offline using custom software written in MATLAB. Fig. 1 provides an overview of the stages involved in our approach to recover the BVP from the webcam videos. We utilized the Open Computer Vision library [14] to automatically identify the coordinates of the face location in the first frame of the video recording using a boosted cascade classifier [15]. The algorithm returned the x- and y-coordinates along with the height and width that define a box around the face. We selected the center 60% width and full height of the box as the region of interest (ROI) for our subsequent calculations.

The ROI was then separated into the three RGB channels [see Fig. 1(b)] and spatially averaged over all pixels in the ROI to yield a red, blue, and green measurement point for each frame and form the raw signals [see Fig. 1(c)] $y_1(t)$, $y_2(t)$, and $y_3(t)$, respectively. Each trace was 1 min long. The raw traces were detrended using a procedure based on a smoothness priors approach [16] with the smoothing parameter $\lambda = 10$ (cutoff frequency of 0.89 Hz) and normalized as follows:

$$y_i'(t) = \frac{y_i(t) - \mu_i}{\sigma_i} \tag{3}$$

for each $i = 1, 2, 3$ where μ_i and σ_i are the mean and standard deviation of $y_i(t)$, respectively. The normalized raw traces are then decomposed into three independent source signals using ICA [see Fig. 1(d)] based on the joint approximate diagonalization of eigenmatrices (JADE) algorithm [17]. ICA is able to perform motion-artifact removal by separating the fluctuations caused predominantly by the BVP from the observed raw signals [11]. However, the order in which ICA returns the independent components is random. Thus, the component whose power spectrum contained the highest peak was then selected for further analysis.

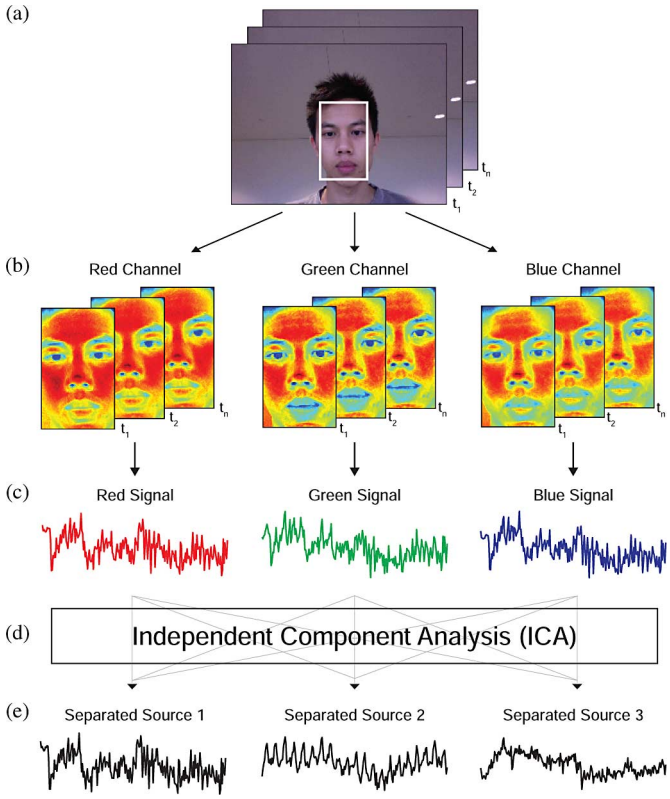

Fig. 1. Recovery of the BVP waveform. (a) Face within the first video frame is automatically detected to locate the ROI. (b) ROI is decomposed into red, green, and blue channels for each frame and spatially averaged to form (c) the raw signals. After the raw signals are detrended and normalized, ICA is applied to separate three independent sources. In this example, the BVP is visible in the second source signal.

C. Quantification of Physiological Parameters

The separated source signal was smoothed using a five-point moving average filter and bandpass filtered (128-point Hamming window, 0.7–4 Hz). To refine the BVP peak fiducial point, the signal was interpolated with a cubic spline function at a sampling frequency of 256 Hz. We developed a custom algorithm to detect the BVP peaks in the interpolated signal and applied it to obtain the interbeat intervals (IBIs). To avoid inclusion of artifacts, such as ectopic beats or motion, the IBIs were filtered using the noncausal of variable threshold (NC-VT) algorithm [18] with a tolerance of 30%. HR was calculated from the mean of the IBI time series as $60/\overline{\text{IBI}}$.

Analysis of HRV was performed by power spectral density (PSD) estimation using the Lomb periodogram. The low-frequency (LF) and high frequency (HF) powers were measured

as the area under the PSD curve corresponding to 0.04–0.15 and 0.15–0.4 Hz, respectively, and quantified in normalized units (n.u.) to minimize the effect on the values of the changes in total power. The LF component is modulated by baroreflex activity and includes both sympathetic and parasympathetic influences [19]. The HF component reflects parasympathetic influence on the heart through efferent vagal activity and is connected to respiratory sinus arrhythmia (RSA), a cardiorespiratory phenomenon characterized by IBI fluctuations that are in phase with inhalation and exhalation. We also calculated the LF/HF ratio, considered to mirror sympatho/vagal balance or to reflect sympathetic modulations.

Since the HF component is connected with breathing, the RR can be estimated from the HRV power spectrum. When the frequency of respiration changes, the center frequency of the HF peak shifts in accordance with RR [20]. Thus, we calculated RR from the center frequency of the HF peak f_{HFpeak} in the HRV PSD derived from the webcam recordings as $60/f_{\mathrm{HFpeak}}$. The respiratory rate measured using the chest belt sensor was determined by the frequency corresponding to the dominant peak f_{resppeak} in the PSD of the recorded respiratory waveform using $60/f_{\mathrm{resppeak}}$.

IV. Results

Using the techniques detailed in Section III, we extracted the BVP waveforms from the webcam recordings via ICA. A typical example of the recovered BVP recordings is shown in Fig. 2(a) along with the BVP recorded with the Flexcomp sensor. It is evident that the two signals are in close agreement and their respective IBI signals are comparable [see Fig. 2(b)]. Since the IBI series is irregularly time-sampled, we utilized the Lomb periodogram to obtain the PSD to avoid resampling and inferring probable replacement values for excluded samples. The resulting spectra are presented in Fig 2(c). Both spectra are comparable and exhibit a dominant HF component. A second example of HRV assessment is shown in Fig. 2(d)–(f). Once again, the BVP and IBI signals are similar and the HRV power spectra both exhibit a dominant LF component.

We were able to determine RR from the HRV power spectrum by locating the center frequency of the HF peak. Fig. 3(a) presents an IBI time series and its corresponding PSD [see Fig. 3(b)]. The center frequency of the HF peak was 0.23 Hz (14 breaths/min) and corresponds to the fundamental breathing rate computed from the PSD [see Fig. 3(d)] of the measured respiratory signal using a chest belt sensor [see Fig. 3(c)].

The level of agreement between the physiological measurements by our proposed method and reference sensors was accessed using Pearson's correlation coefficients ($n = 12$). Correlation scatter plots for each measured parameter are shown in Fig. 4. The webcam-derived physiological measurements were strongly correlated across all parameters with $r = 1.0$ for HR, $r = 0.92$ for HF and LF, $r = 0.88$ for LF/HF, and $r = 0.94$ for RR ($p < 0.001$ for all). The root-mean-squared error of the HR, HF, LF, LF/HF, and RR was 1.24 beats/min, 12.3 and 12.3 n.u., 1.1, and 1.28, respectively.

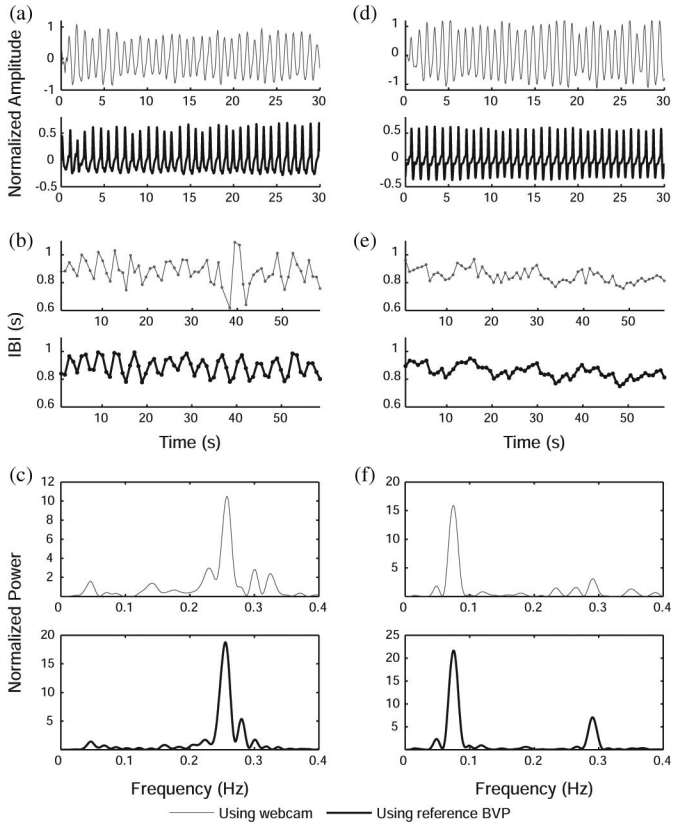

Fig. 2. HRV assessment using a webcam (thin gray lines) in comparison with a finger BVP sensor (thick black lines). (a) BVP waveform (selected source signal was smoothed using a five-point moving average filter and bandpass filtered, 0.7–4 Hz). (b) IBIs formed by extracting the peaks from the BVP waveforms. (c) Normalized Lomb periodogram of the detrended IBIs exhibiting a dominant HF component. (d)–(f) Examples of a recording exhibiting a dominant LF component.

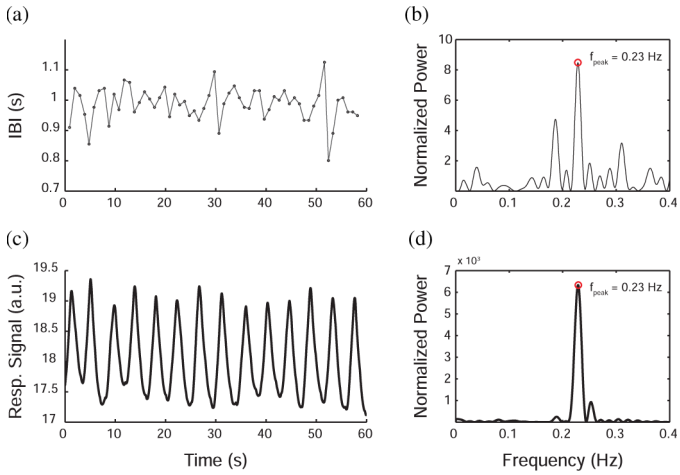

Fig. 3. Comparison of RR assessment between a webcam (thin gray lines) and a chest belt respiration sensor (thick black lines). (a) IBI series from webcam and its (b) normalized Lomb periodogram showing HF power (0.15–0.4 Hz) centered at 0.23 Hz. (c) Respiration waveform measured by the chest belt sensor and its (d) normalized Lomb periodogram showing the fundamental respiration frequency of 0.23 Hz.

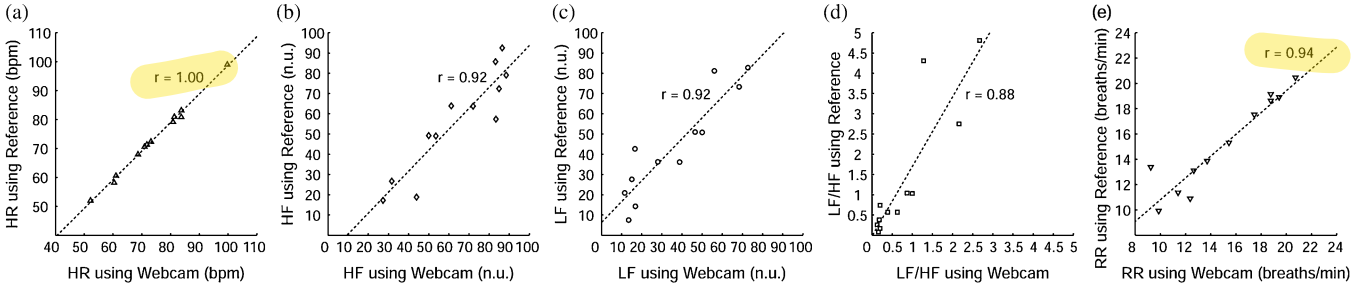

Fig. 4. Scatter plots comparing measurements of (a) HR, (b) HF, (c) LF, (d) LF/HF, and (e) RR between a webcam and reference sensors (finger BVP for HR and HRV measures, chest belt respiration sensor for RR). $p < 0.001$ for all correlations.

TABLE I
SUMMARY OF OVERALL RESULTS

Statistic	Heart Rate (bpm)	Respiratory Rate (breaths/min)	Heart Rate Variability		
			LF (n.u.)	HF (n.u.)	LF/HF
Mean error	0.95	0.12	7.53	7.53	0.57
SD of error	0.83	1.33	10.17	10.17	0.98
RMSE	1.24	1.28	12.3	12.3	1.1
Correlation coefficient	1.00	0.94	0.92	0.92	0.88

All analyses performed on one-minute recordings from 12 participants.

V. DISCUSSION

On the basis of the results from the present study (see Table I), we have demonstrated the feasibility of using a simple webcam to measure multiple physiological parameters. This includes vital signs, such as HR and RR, as well as correlates of cardiac autonomic function through HRV. Our data demonstrate that there is a strong correlation between these parameters derived from webcam recordings and standard reference sensors. Regarding the choice of measurement epoch, a recording of 1–2 min is needed to assess the spectral components of HRV [5] and an averaging period of 60 beats improves the confidence in the single timing measurement from the BVP waveform [9]. The face detection algorithm is subject to head rotation limits. About three axes of pitch, rotation, and yaw, the limits were $32.6° \pm 4.84°$, $33.4° \pm 2.34°$, and $18.6° \pm 3.75°$ from the frontal position.

The results must be considered in light of several limitations of the present study. First, the webcam video sampling rate fluctuated around 15 fps due to the use of a standard PC for image acquisition, causing misalignment of the BVP peaks compared to the reference signal. The performance could be boosted, if each video frame was time stamped and the signals were resampled. Second, the video sampling rate is much lower than the recommended rates ($\geq$250 Hz) for HRV analysis. By interpolating at 256 Hz to refine the peaks in the BVP and improve timing estimations, we achieved the high correlations in Table I. We acknowledge that PPG beat-to-beat variability can be affected by changes in the pulse transit time, which is related to arterial compliance and blood pressure, but it has been shown to be a good surrogate of HRV at rest [21]. Another limitation of this system is that only three source signals can be recovered. However, our results suggest that this is sufficient to obtain accurate measurements of the BVP. Our findings should motivate extensive validation and continued systematic exploration of these variables.

REFERENCES

[1] S. Ulyanov and V. Tuchin, "Pulse-wave monitoring by means of focused laser beams scattered by skin surface and membranes," in *Proc. SPIE*, Los Angeles, CA, 1884, pp. 160–167.
[2] E. Greneker, "Radar sensing of heartbeat and respiration at a distance with applications of the technology," in *Proc. Conf. RADAR*, Edinburgh, U.K., 1997, pp. 150–154.
[3] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 8, pp. 1418–1426, Aug. 2007.
[4] J. Fei and I. Pavlidis, "Thermistor at a distance: Unobtrusive measurement of breathing," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 988–998, Apr. 2009.
[5] M. Malik, J. Bigger, A. Camm, R. Kleiger, A. Malliani, A. Moss, and P. Schwartz, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, pp. 354–381, 1996.
[6] S. Suzuki, T. Matsui, S. Gotoh, Y. Mori, B. Takase, and M. Ishihara, "Development of non-contact monitoring system of heart rate variability (hrv)-an approach of remote sensing for ubiquitous technology," in *Proc. Int. Conf. Ergonom. Health Aspects Work Comput.*, San Diego, CA, 2009, pp. 195–203.
[7] G. Lu, F. Yang, Y. Tian, X. Jing, and J. Wang, "Contact-free measurement of heart rate variability via a microwave sensor," *Sensors*, vol. 9, pp. 9572–9581, 2009.
[8] U. Morbiducci, L. Scalise, M. De Melis, and M. Grigioni, "Optical vibrocardiography: A novel tool for the optical monitoring of cardiac activity," *Ann. Biomed. Eng.*, vol. 35, pp. 45–58, Jan. 2007.
[9] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, pp. R1–R39, Mar. 2007.
[10] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Expr.*, vol. 16, pp. 21434–21445, Dec. 2008.
[11] M. Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Expr.*, vol. 18, pp. 10762–10774, May 2010.
[12] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, pp. 287–314, 1994.
[13] W. G. Zijlstra, A. Buursma, and W. P. Meeuwsen-van der Roest, "Absorption spectra of human fetal and adult oxyhemoglobin, de-oxyhemoglobin, carboxyhemoglobin, and methemoglobin," *Clin. Chem.*, vol. 37, pp. 1633–1638, Sep. 1991.
[14] A. Noulas and B. Kröse, "EM detection of common origin of multi-modal cues," in *Proc. ACM Conf. Multimodal Interfaces*, 2006, pp. 201–208.
[15] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Conf Comput. Vis. Pattern Recognit.*, 2001, pp. I-511–I-518.
[16] M. P. Tarvainen, P. O. Ranta-Aho, and P. A. Karjalainen, "An advanced detrending method with application to HRV analysis," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 2, pp. 172–175, Feb. 2002.
[17] J.-F. Cardoso, "High-order contrasts for independent component analysis," *Neural Comput.*, vol. 11, pp. 157–192, 1999.

[18] J. Vila, F. Palacios, J. Presedo, M. Fernandez-Delgado, P. Felix, and S. Barro, "Time-frequency analysis of heart-rate variability," *IEEE Eng. Med. Biol. Mag.*, vol. 16, no. 5, pp. 119–126, Sep./Oct. 1997.

[19] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A. C. Berger, and R. J. Cohen, "Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control," *Science*, vol. 213, pp. 220–222, Jul. 1981.

[20] T. Brown, L. Beightol, J. Koh, and D. Eckberg, "Important influence of respiration on human RR interval power spectra is largely ignored," *J. Appl. Physiol.*, vol. 75, pp. 2310–2317, 1993.

[21] E. Gil, M. Orini, R. Bailón, J. Vergara, L. Mainardi, and P. Laguna, "Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions," *Physiol. Meas.*, vol. 31, pp. 1271–1290, 2010.