On the Diffusion Process for Heart Rate Estimation from Face Videos Under Realistic Conditions

Conference Paper · August 2017

DOI: 10.1007/978-3-319-66709-6_29

3 authors, including:

Jarek Krajewski
Rhenish University of Cologne
95 PUBLICATIONS **1,222** CITATIONS

SEE PROFILE

On the Diffusion Process for Heart Rate Estimation from Face Videos Under Realistic Conditions

Christian S. Pilz[1(✉)], Jarek Krajewski[2], and Vladimir Blazek[3]

[1] CanControls GmbH, Aachen, Germany
pilz@cancontrols.com
[2] Institute of Safety Technology, University of Wuppertal, Wuppertal, Germany
[3] Philips Chair for Medical Information Technology,
Helmholtz-Institute for Biomedical Engineering, RWTH Aachen University,
Aachen, Germany

Abstract. This work addresses the problem of estimating heart rate from face videos under real conditions using a model based on the recursive inference problem that leverages the local invariance of the heart rate. The proposed solution is based on the canonical state space representation of an Itō process and a Wiener velocity model. Empirical results yield to excellent real-time and estimation performance of heart rate in presence of disturbing factors, like rigid head motion, talking and facial expressions under natural illumination conditions making the process of heart rate estimation from face videos applicable in a much broader sense. To facilitate comparisons and to support research we made the code and data for reproducing the results public available.

1 Introduction

In general, the role of physiological states has a large impact on human state computing in computer vision, since it tells something about the affective nature of the human interacting with a machine or being monitored by optical sensors solely. During the last years, the task of measuring skin blood perfusion and heart rate measurements from facial images became part of top computer vision conferences [23,25,34,41]. Interestingly, all these contributions focus on how to cope with motion like head pose variations and facial expressions since any kind of motion on a specific skin region of interest will destroy the underlying blood perfusion signal in a way that no reliable information can be extracted anymore. Apart from being able to estimate vitality parameters like heart rate and respiration, monitoring functional survey of wounds as well as quantification of allergic skin reaction [4] are further topics of discovered employment scenarios of skin blood perfusion analysis. Recently, prediction of emotional states [28,36] and stress [6] became an interesting new achievement in this area, pushing the focus of this technology further towards human-machine interaction. The technical term of skin blood perfusion analysis by image sensors is known as Photoplethysmography Imaging, short PPGI. Although the PPGI measurement is accurate,

V. Roth and T. Vetter (Eds.): GCPR 2017, LNCS 10496, pp. 361–373, 2017.
DOI: 10.1007/978-3-319-66709-6_29

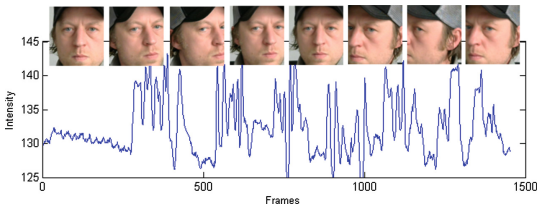

Fig. 1. The average green channel on a face region during rigid head motions. In the first 250 frames the user is in a resting state and the fine pulsation of blood volume is visible. After 300 frames the user started to move his head and the pulse signal gets lost. (Color figure online)

影响心率估计准确率的三个因素

it is bound to specific environmental conditions in order to operate accurately yet. The most significant factors influencing the operational performance are the sensors spectral response, user movements like rigid head motions and facial expressions as well as fast varying illumination conditions. Figure 1 illustrates the disturbing influence of head motions on the raw PPGI blood volume pulse signal. During head motion the pixel intensity distribution varies significantly, much more than the expectation which can be attributed to blood pulsation. This problem can be solved by a feature invariant under local intensity variations larger than blood pulsation and a process model that leverages the local invariance of the heart rate. The main contributions of this work are

- a novel formulation for the problem of estimating heart rate using low-cost camera sensor technology, and
- a spatio-independent feature representation invariant under large varying pixel intensity distribution induced by facial motions.

Initially, from the historical genuine up to the development of the current state of the art in computer vision, the methodology of PPGI will be reviewed. Followed by theoretical aspects, the model will be described in detail. Based upon conducted empirical field data collection the results will be presented and finally discussed.

2 Related Work

The term Photoplethysmography, short PPG, dates back to the late first half of the 20th century, when Molitor and Kniazak [31] recorded peripheral circulatory changes in animals. One year later, Hertzman [16] introduced the term Photoelectric Plethysmograph as "the amplitude of volume pulse as a measure of the blood supply of the skin". Hertzman's instrumentation comprised mainly of a tungsten arc lamp and a photomultiplier tube. An advancement to the classical PPG is the camera based PPGI method introduced by the pioneering work of Blazek [5]. The basic principle behind the measurement of blood volume changes in the skin by means of PPG is the fact that hemoglobin in the blood absorbs

specific frequency bands of light many times more strongly than the remaining skin tissues. Since the first published visualisation of pulsatile skin perfusion patterns in the time and frequency domain by Blazek [5], classical signal processing methods are applied commonly to extract reasonable information out of the perfusion signals [17,35,42]. Since it is realized that motion of the skin area of interest [17] and later micro motion of the head due to cardiac activity [3,29] inherently induces artifacts into the extracted signal, especially when lighting is neither uniform nor orthogonal, canceling motion artifacts during signal processing became an important aspect for reliable skin blood perfusion measurements [30]. From the basic early idea of compensating the motion of the skin area of interest by optical flow methods directly in the image plane [17], Poh *et al.* [35] regarded the problem solution for facial videos as a blind source separation task using Independent Component Analysis (ICA) over the different color channels, whereby Lewandowska *et al.* [24] compared ICA against Principal Component Analysis additionally. However, in case the underlying signal basis is majoritarian Gaussian, ICA will not be able to determine a proper de-mixing matrix [8]. This happens exactly when the skin blood perfusion signals contain harmonics beside their fundamental frequency. De Haan and Jeanne [14] and De Haan and Van Leest [15] proposed to map the PPGI-signals by linear combination of RGB data to a direction that is orthogonal to motion induced artifacts. A recent alternative, which does not require skin-tone or pulse-related priors in contrast to the channel mapping algorithms, determines the spatial subspace of skin-pixels and measure its temporal rotation for signal extraction [44]. Tulyakov *et al.* [41] proposed matrix completion to jointly estimate reliable regions and heart rate estimates whereby Li *et al.* [25] applied an adaptive least square approach to extract robust pulse frequencies. Both reported performance gains similar to De Haan and Jeanne [14]. Only Wang *et al.* [44] reports significantly more accurate results especially under head motion and talking.

3 Methodology

信号由心跳频率、光照、头部运动和面部运动组成。心跳可以看成是随机振荡器的准周期形式，其他三种可以用维纳过程来表示

The underlying system of measuring heart rate from face regions using conventional camera technology is based upon a diffusion process. The entire process itself consists of independent single processes; the heart frequency, the illumination and the users head movement and facial motion. The periodic event of heart frequency appears in form of a stochastic resonator. The illumination as well as the head movement and facial motion is represented as a Wiener process [45], whereby a violation of the smoothness criterion yields to a generalized Poisson process [9]. The general solution of the corresponding stochastic differential equations is given by Itō's lemma [18]. If the resonator's fundamental frequency is known, the solution yields to a general time-discrete linear dynamic system [21]. However, in case the resonator's fundamental frequency is unknown, the problem is given as latent state of the frequency. This results in a Markov process, whereby the latent states are time-discrete linear dynamic systems. The closed form solution to this problem is described by Bloom and Bar-Shalom [7]. The

advantage of this kind of formulation is that the case of non-uniform sampling as well as missing observations is naturally included in the model [20]. The basic idea of methodology is inspired by the work of Särrkä [37] and Särrkä *et al.* [38].

3.1 The Diffusion Process

In mathematics and physics any kind of process producing the same output from a given starting condition or initial state is called a deterministic system. This means in the development of future states of the system there's no randomness involved. Like physical laws they are described by differential equations. Unfortunately, in nature when studying biological systems this is often a rarely case. Here, this can be expressed as a sequence of random variables. The random variables which correspond to various times may be completely different. The only assumption is that these different quantities take values in the same space. In probability theory this phenomena is known as stochastic process. Recalling the ordinary differential equation [1] of the form

$$\frac{d\boldsymbol{x}(t)}{dt} = F\boldsymbol{x}(t) \tag{1}$$

Adding a noise term as function yields to a stochastic differential equation

$$\frac{d\boldsymbol{x}(t)}{dt} = F\boldsymbol{x}(t) + L\boldsymbol{w}(t) \tag{2}$$

where the Itô process [18] defines the solution to the equation.

The discrete-time approximation of a linear stochastic differential equation yields to [32,37]

$$\begin{aligned} \boldsymbol{x}(t_i + 1) &= e^{\Delta t F}\boldsymbol{x}(t_i) \\ &\quad + \int_{t_i}^{t_{i+1}} e^{(t-s)F} L\boldsymbol{w}(s)\mathrm{d}s, \end{aligned} \tag{3}$$

with $\Delta t = t_{i+1} - t_i$, the Wiener process $w(t)$ with spectral density W and the covariance of the stochastic integral

$$Q_i = \int_0^{\Delta t} A_i L W L^T A_i^T \mathrm{d}\tau, \tag{4}$$

with $A_i = e^{\Delta t F}$, which results to the discrete-time model [10,13,19,21,47]

$$\boldsymbol{x}(t_{i+1}) = A_i\boldsymbol{x}(t_i) + \boldsymbol{q}(t_i) \tag{5}$$

$$\boldsymbol{y}(t_i) = H_i\boldsymbol{x}(t_i) + \boldsymbol{e}(t_i) \tag{6}$$

with process noise

$$\boldsymbol{q}(t_i) = N(0, Q_i) \tag{7}$$

and measurement noise

$$\boldsymbol{e}(t_i) = N(0, \Sigma_i). \tag{8}$$

A_i and H_i corresponds to the state transition and the measurement model respectively. Although, the dynamic of the model is continuous, the measurements are at discrete time steps of a linear Gaussian process.

3.2 The Feature Space

Under natural conditions, illumination is not static as well as not uniform distributed over facial skin regions. The intensity distribution of skin pixels varies significantly when an user starts to move the head as illustrated in Fig. 1. Assuming that the image source S produces random variables X with an associated probability density function $p(x)$, a suitable feature representation for the skin blood perfusion, invariant with respect to the location and large magnitude differences of pixel intensities can be expressed as the average quantization error d of an uniform quantizer and an associated set of decision intervals $\{I_k\}_{k=1}^{M}$, such that $I_k = \{b_{k-1}, b_k)$ for $k = 1, 2, ..., M$, where b_{k-1} and b_k represent the upper and lower limits of the pixels values at each region. Each interval I_k is represented by a quantized value y_k, the mean value of the each region, which implements the mapping $x \in I_k \Rightarrow y = y_k$ [12,33]

$$D = \sum_{k=1}^{M} \int_{b_{k-1}}^{b_k} (x - y_k) p(x) \mathrm{d}x. \tag{9}$$

The quantizers amount of bits has to be chosen smaller than the bits used for the sensors capacitor discharge voltage being read as pixel intensity. Since the contribution of the perfusion signals magnitude inside each code word interval is assumed to be not equal, each average interval distortion is normalized by standard deviation

$$F = \sum_{k=1}^{M} \frac{\int_{b_{k-1}}^{b_k} (x - y_k) p(x) \mathrm{d}x}{\sqrt{\int_{d_k} (d_k - \mu_{d_k})^2 p(d_k) \mathrm{d}d_k}} \tag{10}$$

with

$$u_{d_k} = \int_{d_k} d_k p(d_k) \mathrm{d}d_k. \tag{11}$$

A strategy of normalizing the local intensities is given by a suitable surrounding potential. This can be either realized as the ratio of the green channel and the total amount of RGB information on a given pixel, $G/(R + G + B)$ or for any monochrome intensity by its neighboring values, for example using the Moore neighborhood. As general pre-processing step skin segmentation can be performed by thresholding the chroma components Cr and Cb in the YCrCB color space.

The described feature is a spatio-independent representation under the assumption of a phase synchronized perfusion phenomena, although in detail this is not the case [40].

3.3 The Periodicity

Any band-limited zero-mean periodic signal with stationary frequency f can be approximated to an arbitrary precision with a truncated Fourier series

$$c(t) = \sum_{n=1}^{N} a_n cos(2\pi n f t) + b_n sin(2\pi n f t) \tag{12}$$

with a_n and b_n the Fourier coefficients. The fundamental frequency f and the harmonic components with frequencies as multiples of their base, nf [22,46].

Physiological signals like the perfusion phenomena are not stationary and they are of quasi-periodic nature. The major source of non-stationary in the perfusion signal is caused by the varying fundamental frequency resulting in a function of time $f(t)$. Another source of aperiodicity of the signal are small changes in amplitudes and phases in the harmonics. This can be modeled by an additional white noise component $e_n(t)$ with spectral density q_n to each harmonic component

$$\begin{aligned} c(t) &= \sum_{n=1}^{N} a_n cos(2\pi n f(t) t) \\ &\quad + b_n sin(2\pi n f(t) t) + e_n(t). \end{aligned} \tag{13}$$

However, this representation is sensitive to changes in frequency. When t is large, any change in the frequency causes a large change in signal $c(t)$. Discontinuities in the frequency will also cause the signal $c(t)$ to be discontinuous.

Recalling the classical mechanics of circular motion, the system of a single harmonic oscillator yields to a 2nd order differential equations [11]

$$\frac{d^2 c_n(t)}{dt^2} = -(2\pi n f)^2 c_n(t) \tag{14}$$

with the solution

$$c_n(t) = a_n cos(2\pi n f t) + b_n sin(2\pi n f t) \tag{15}$$

where the constants a_n and b_n are set by the initial conditions of the differential equation. Accounting for non-stationary frequency as a function of time and changes in amplitude and phase leads to the differential equation

$$\frac{d^2 c_n(t)}{dt^2} = -(2\pi n f(t))^2 c_n(t) + e_n(t) \tag{16}$$

for each harmonic component. The major advantage of such a stochastic representation of a resonator is, even when the frequency is discontinuous the signal remains continuous. Figure 2 shows a single stochastic oscillator with time-varying frequency and amplitude. The stochastic state space for the resonator signal yields to

$$\frac{dx(t)}{dt} = F_0(f(t))x(t) + Le(t), \tag{17}$$

$$c(t) = Hx(t). \tag{18}$$

Since the frequency is unknown, the state space depends on an additional latent variable

$$\frac{dx(t)}{dt} = F_0(\theta)x(t) + Le(t), \tag{19}$$

$$c(t) = H(\theta)x(t). \tag{20}$$

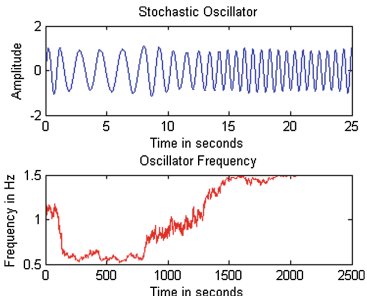

Fig. 2. A simulated trajectory of a stochastic oscillator with frequency trace in a range typical for a human in resting state.

such that

$$\theta \in \Omega = \{\theta^1, \ldots, \theta^S\} \tag{21}$$

forming a Markov chain with transition matrix Π with transition probabilities

$$P(\theta_t^i | \theta_{t-1}^j) = \Pi_{ij}. \tag{22}$$

The solution is given by computing the Gaussian mixture approximation to the joint posterior distribution of the latent variables and states [7].

3.4 The Drift

Despite the underlying pure perfusion signal, remote sensing signals of human skin encapsulates further components which can be described as time varying bias and jumps. The bias function $x(t)$ is assumed to be smooth and slow varying and can be assigned to general changes of natural illumination conditions as well as slow head movements and facial motion. Since there's no further distinct knowledge about the drift process it is modeled as Wiener process

$$\frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2} = w(t), \tag{23}$$

where $w(t)$ corresponds to a white noise process with spectral density q_w and has continuous sample path [45]. However, sudden strong changes in illumination, fast head movements and facial motion causing discontinuities and therefore violating the smoothness criterion of the drift. This yields to a general Markov process that modulates the intensity function of an inhomogeneous Poisson counting process

$$X(t) = X_0 + \sum_{i:s_i \leq t} \xi_i \tag{24}$$

where $X(t)$ is a real-valued stochastic process and (s_i, ξ_i) are the events of a two dimensional Poisson process on $[0, T] \times \mathbb{R}$. The intensity of this process is given by $\lambda(s, y) = \lambda h(y)$, where $\lambda > 0$ describing how frequently the jumps of X occur

and $h_\theta(y)$ is the jump density describing magnitudes of jumps ξ_i of process X. Here $\theta \in \Theta$ is an unknown parameter defining the distribution of jumps. The parameters θ and λ have a prior density $\pi(\theta, \lambda)$. As well the initial condition X_0 is assumed to have a prior density π_0. λ will have a time varying intensity $\lambda = \lambda(t)$. The Bayesian approach leads conveniently to the combined parameter and state estimation via the posterior expectation given the observation Y

$$\left(\widehat{X(t), \theta, \lambda}(t)\right) := \mathbb{E}_{\pi,\pi_\vartheta}[X(t), \theta, \lambda(t)|Y_1, ...Y_k] \tag{25}$$

for $\tau k \leq t < \tau(k+1)$ [27]. Figure 3 shows a simulated trajectory of a Wiener process and its realization modulated by a Poisson process.

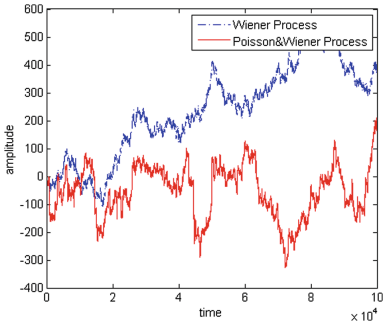

Fig. 3. A simulated trajectory of a Wiener process and its realization modulated by a Poisson process with $\lambda = 1.95$ and $\theta = 35$.

4 Experiments

In order to evaluate the proposed model, empirical data is collected under natural environmental conditions with a typical 24 bit low-cost webcam, a Logitech HD C270, as well as reference ground truth measurements using a common finger pulseoximeter, a CMS50E PPG device. 25 users were asked to perform video recordings in two sessions resulting in a total amount of 50 videos. The first session is selected to be even-tempered without any kind of larger head or body movements and facial expressions. During the second session, participants were free to move their head naturally while remaining seated. Typical movements included tilting the head sideways, nodding the head, looking up/down and leaning forward/backward. Some participants also made facial expressions, or started to talk. The recording illumination environment was chosen as daylight office scenario without any additional lighting. The duration of each session is approximately one minute. The frame rate was fixed to 15 fps in average and the corresponding time stamps for each frame were captured too. The finger pulseoximeter data for each session and participant was stored for later comparison. Bland-Altman [2] and correlation plots were used for combined graphical and statistical interpretation of the two measurement techniques. For every

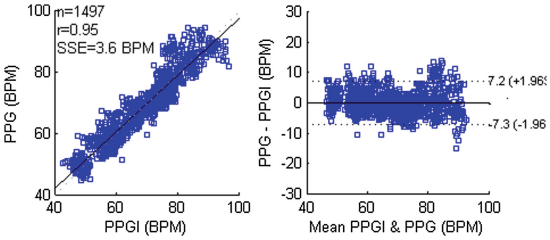

Fig. 4. Correlation and Bland-Altman [2] plots of PPGI diffusion process estimated heart rate against CMS50E PPG finger-pulseoximeter reference of 25 users in resting state.

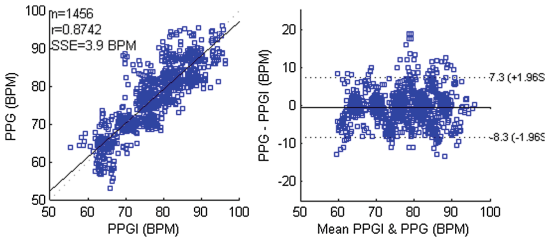

Fig. 5. Correlation and Bland-Altman [2] plots of PPGI diffusion process estimated heart rate against CMS50E PPG finger-pulseoximeter reference of 25 users performing head rotations.

video recording a standard Viola-Jones [43] face finder was used to determine the analysis region of interest. The extracted feature, 6 bit quantized green channel, was feed into the vector valued representation of the diffusion process on a frame by frame basis. On every estimated pulse trace a spectral peak is determined by the Lomb periodogram [26,39]. The frame duration was set to 10 s with an overlap of 90%. The correlation and Bland-Altman plots for the resting and head motion condition are reported in the following Figs. 4 and 5 respectively.

To obtain further insides about the potential strength of the diffusion process model, the approach is compared against the recently published Spatial Subspace Rotation (SSR) [44] and the baseline ICA approach [35][1]. Figure 6 compares an users estimated pulse signal under rigid head motions and the corresponding spectrogram for the three methods. The hear rate frequency for the ICA methods nearly gets lost completely. For the SSR method the frequency trace of heart rate is better visible but cannot compete against the diffusion process model where the heart rate is very clear over the entire sequence of head movements. Interestingly, the heart rate gets slightly increased during the head rotation phases. The detailed correlation coefficients and squared errors of prediction for all approaches are provided for the two data sessions in Table 1. The primary

[1] We also reimplemented other methods [14,25,41], since their code is not available. Unfortunately, we obtained worse results.

signal processing and frequency estimation is the same used during the diffusion process experiments. ICA performs worst and is not able to provide reliable heart rate information during head motion. Although SSR performs better it cannot compete against the robustness of the diffusion process.

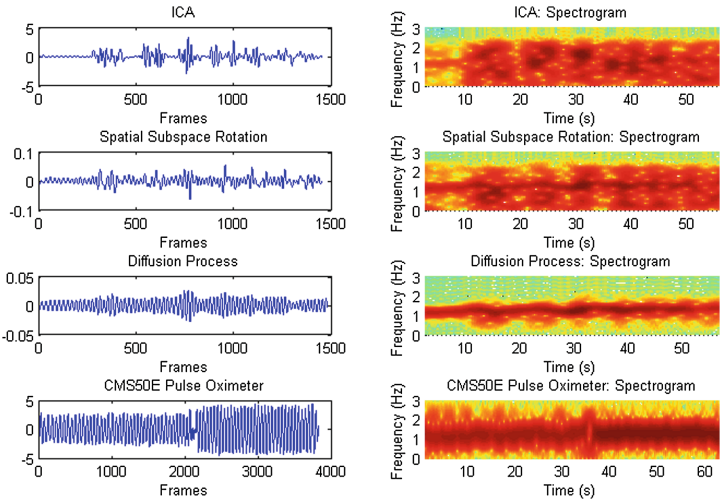

Fig. 6. Comparison of an users estimated blood volume pulse signal. Under rigid head motions and the corresponding spectrogram for the ICA [35], the SSR [44], the diffusion process model and the reference finger pulseoximeter. These estimates are based upon the video illustrated in Fig. 1.

Table 1. Pearson's correlation coefficient and squared errors of prediction of ICA [35], the SSR [44] and the diffusion process (DP).

Type	ICA	SSR	DP
Resting	0.61/7.8	0.78/4.8	0.95/3.6
Head rotation	0.21/14.6	0.47/7.6	0.87/3.9

5 Conclusions

In this work, we have derived a holistic signal interpretation of heart rate estimation from face videos under realistic conditions. The closed form solution of the corresponding stochastic differential equations yields to a diffusion process where the exact estimate of the source separated heart rate signal is obtained via the posterior distribution of the process. The recursive nature of the underlying Bayesian inference scales the processing to a very fast approach. Further, we considered a spatio independent intensity feature. We compared the model

against two approaches on face videos under resting as well as head and facial motion scenarios under natural illumination conditions. Measurements on a 25 user experiment showed clearly superior robustness of the diffusion process modelling, although the uncertainty of prediction still gets slightly increased during natural head motion. We conclude that an entirely invariant process model still depends on a more robust feature representation.

References

1. Arnold, I.: Ordinary Differential Equations. MIT Press, Cambridge (1973)
2. Bland, J., Altman, D.: Statistical methods for assessing agreement between two methods of clinical measurement. Lancet **327**(8476), 307–310 (1986)
3. Blanik, N., Blazek, C., Pereira, C., Blazek, V., Leonhardt, S.: Wearable photoplethysmographic sensors: past and present. In: Proceedings of the SPIE 9034, Medical Imaging: Image Processing (2014)
4. Blazek, C., Hülsbusch, M.: Assessment of allergic skin reactions and their hemodynamical quantification using photoplethysmography imaging. In: Proceedings of 11th International Symposium CNVD, Computer-Aided Noninvasive Vascular Diagnostics, vol. 3, 85–90 (2005)
5. Blazek, V.: Optoelektronische Erfassung und rechnerunterstützte Analyse der Mikrozirkulations-Rhythmik. Biomed. Techn. **30**(1), 121–122 (1985)
6. Blazek, V., Blanik, N., Blazek, C., Paul, M., Pereira, C., Koeny, M., Venema, B., Leonhardt, S.: Active and passive optical imaging modality for unobtrusive cardiorespiratory monitoring and facial expressions assessment. Assessment. Anesth Analg. **124**, 104–119 (2017)
7. Bloom, H., Bar-Shalom, Y.: The interacting multiple model algorithm for systems with markovian switching coefficients. IEEE Trans. Autom. Control **33**(8), 780–783 (1988)
8. Cardoso, J.: High-order contrasts for independent component analysis. Neural Comput. **11**(1), 157–192 (1999)
9. Cox, D.: Some statistical methods connected with series of events. J. Roy. Stat. Soc. **17**(2), 129–164 (1950)
10. Durbin, J., Koopman, S.: Time Series Analysis by State Space Methods. Oxford University Press, Oxford (2001)
11. Feynman, R., Leighton, R., Sands, M.: The Feynman Lectures on Physics, vol. 1. Addison-Wesley, Boston (1963). Chap. 21
12. Gray, R., Neuhoff, D.: Quantization. IEEE Trans. Inf. Theory **44**(6), 2325–2383 (1998)
13. Grewal, M., Andrews, A.: Kalman Filtering Theory and Practice Using Matlab. Wiley Interscience, Hoboken (2001)
14. de Haan, G., Jeanne, V.: Robust pulse-rate from chrominance-based rppg. IEEE Trans. Biomed. Eng. **60**(10), 2878–2886 (2014)
15. de Haan, G., van Leest, A.: Improved motion robustness of remote-ppg by using the blood volume pulse signature. Physiol. Meas. **3**(9), 1913–1926 (2014)
16. Hertzman, A.: Photoelectric plethysmography of the fingers and toes in man. Exp. Biol. Med. **37**(3), 529–534 (1937)
17. Hülsbusch, M.: A functional imaging technique for opto-electronic assessment of skin perfusion. Ph.D. thesis, RWTH Aachen University (2008)

18. Itô, K.: On Stochastic Differential Equations, vol. 4. Memoris of The American Mathematical Society (1951)
19. Jazwinski, A.: Stochastic Processes and Filtering Theory. Academic Press, New York (1970)
20. Jones, R.H.: Fitting multivariate models to unequally spaced data. In: Parzen, E. (ed.) Time Series Analysis of Irregularly Observed Data. LNS, vol. 25, pp. 158–188. Springer, New York (1984). doi:10.1007/978-1-4684-9403-7_8
21. Kalman, R., Bucy, R.: New results in linear filtering and prediction theory. Trans. ASME-J. Basic Eng. **83**, 95–108 (1961)
22. Khintchine, A.: Korrelationstheorie der stationären stochastischen Prozesse. Springer-Mathematische Annalen **109**, 604–615 (1934)
23. Lam, A., Kuno, Y.: Robust heart rate measurement from video using select random patches. In: IEEE International Conference on Computer Vision, pp. 3640–3648 (2015)
24. Lewandowska, M., Ruminski, J., Kocejko, T., Nowak, J.: Measuring pulse rate with a webcam - a non-contact method for evaluating cardiac activity. In: Proceedings of the FedCSIS, Szczecin, Poland, pp. 405–410 (2011)
25. Li, X., Chen, J., Zhao, G., Pietikinen, M.: Remote heart rate measurement from face videos under realistic situations. In: IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH (2014)
26. Lomb, N.: Least-squares frequency analysis of unequally spaced data. Astrophys. Space Sci. **39**(2), 447–462 (1976)
27. Makhnin, O.: Filtering for some stochastic processes with discrete observations. Ph.D. thesis, Department of Statistics and Probability, Michigan State University (2002)
28. McDuff, D., Gontarek, S., Picard, R.: Remote measurement of cognitive stress via heart rate variability. In: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2957–2960 (2014)
29. Moço, A., Stuijk, S., de Haan, G.: Ballistocardiographic artifacts in PPG imaging. IEEE Trans. Biomed. Eng. **63**(9), 1804–1811 (2015)
30. Moço, A., Stuijk, S., de Haan, G.: Motion robust PPG-imaging through color channel mapping. Biomed. Opt. Express **7**, 1737–1754 (2016)
31. Molitor, H., Knaizuk, M.: A new bloodless method for continuous recording of peripheral change. J. Pharmacol. Exp. Theret. **27**, 5–16 (1936)
32. Øksendal, B.: Stochastic Differential Equations. Springer, Heidelberg (2003)
33. Oliver, B., Pierce, J., Shannon, C.: The philosophy of PCM. Proc. IRE **36**, 1324–1331 (1948)
34. Osman, A., Turcot, J., Kaliouby, R.E.: Supervised learning approach to remote heart rate estimation from facial videos. In: 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, pp. 1–6 (2015)
35. Poh, M., McDuff, J., Picard, R.: Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Opt. Express **18**(10), 10762–10774 (2010)
36. Ramirez, G., Fuentes, O., Crites, S., Jimenez, M., Ordonez, J.: Color analysis of facial skin: detection of emotional state. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 474–479 (2014)
37. Särkkä, S.: Recursive Bayesian inference on stochastic differential equations. Ph.D. thesis, Helsinki University of Technology (2006)
38. Särkkä, S., Solin, A., Nummenmaa, A., Vehtari, T., Vanni, F.L.: Dynamic retrospective filtering of physiological noise in BOLD fMRI. NeuroImage **60**(2), 1517–1527 (2012)

39. Scargle, J.: Studies in astronomical time series analysis. II - statistical aspects of spectral analysis of unevenly spaced data. Astrophys. J. **263**(1), 835–853 (1982)
40. Teplov, V.: Blood pulsation imaging. Ph.D. thesis, Department of Applied Physics, University of Eastern Finland (2014)
41. Tulyakov, S., Pineda, X.A., Ricci, E., Yin, L., Cohn, J., Sebe, N.: Self-adaptive matrix completion for heart rate estimation from face videos under realistic conditions. In: Computer Vision and Pattern Recognition (2016)
42. Verkruysse, W., Svaasand, L., Nelson, J.: Remote plethysmographic imaging using ambient light. Opt. Express **16**(26), 21434–21445 (2008)
43. Viola, P., Jones, M.: Robust real-time object detection. Int. J. Comput. Vis. **57**, 137–154 (2001)
44. Wang, W., Stuijk, S., de Haan, G.: A novel algorithm for remote photoplethysmography: spatial subspace rotation. IEEE Trans. Biomed. Eng. **63**(9), 1974–1984 (2015)
45. Wiener, N.: The average of an analytical functional and the brownian movement. Proc. Nat. Acad. Sci. USA **7**(1), 294–298 (1921)
46. Wiener, N.: Generalized harmonic analysis. Acta Mathematica **55**, 117–258 (1930)
47. Zakai, M.: On the optimal filtering of diffusion processes. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **11**(3), 230–243 (1969)