



HAL
open science

Probabilistic Signal Quality Metric for Reduced Complexity Unsupervised Remote Photoplethysmography

Yannick Benezeth, S Bobbia, K Nakamura, R Gomez, J. Dubois

► **To cite this version:**

Yannick Benezeth, S Bobbia, K Nakamura, R Gomez, J. Dubois. Probabilistic Signal Quality Metric for Reduced Complexity Unsupervised Remote Photoplethysmography. IEEE International Symposium on Medical Information and Communication Technology (ISMICT), May 2019, Oslo, Norway. 10.1109/ISMICT.2019.8744004 . hal-02143077

HAL Id: hal-02143077

<https://u-bourgogne.hal.science/hal-02143077>

Submitted on 29 May 2019

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.

Probabilistic Signal Quality Metric for Reduced Complexity Unsupervised Remote Photoplethysmography

Y. Benezeth¹, S. Bobbia¹, K. Nakamura², R. Gomez², J. Dubois¹

¹ImViA EA 7535, Univ. Bourgogne Franche-Comté, Dijon, France

²Honda Research Institute Japan Co., Ltd., 8-1 Honcho, Wako-shi, Saitama, Japan

Abstract—Remote photoplethysmography (rPPG) is a recent technique for estimating heart rate by analyzing the pulsatility of skin hue using regular cameras. To determine the quality of the measurement, many existing methods are based on the signal-to-noise ratio (SNR) calculated in the frequency domain. However, the Fast Fourier Transform (FFT) operation is performed with a minimal complexity of $\mathcal{O}(n \log n)$. Therefore, the use of this quality metric in an unsupervised rPPG framework in which this metric is estimated a large number of times will tend to greatly increase the complexity of the solution. In this paper, we propose a new probabilistic formulation of a cardiac signal quality index, with lower complexity, based on the Bayesian information criterion (BIC) that encapsulates the characteristic shape of the rPPG signal. The results of this study, obtained on a public database, have demonstrated that the proposed probabilistic metric outperforms the regular SNR metric with a lower computation complexity.

Index Terms—Biomedical monitoring, heart rate measurements, signal quality metric.

I. INTRODUCTION

Since 2010 and pioneering work of Poh *et al.* [5], remote photoplethysmography has become a very active field of research. In recent years, several methods have been proposed to estimate the physiological signal with fewer and fewer environmental constraints. Thus, an expensive configuration is no longer necessary [14], since a consumer grade webcam can be used to make the measurements. The basic principle of remote photoplethysmography stems from reflective photoplethysmography where the light reaching a camera is modulated by the blood pulsations of skin. The rhythmic beating of the heart results in the pulsating blood volume which in turn results in minute changes in the color of the skin which can be quantified using different signal processing techniques to generate a cardiac signal.

The problem of rPPG measurement has been attacked from different directions. In the initial research related to rPPG measurement, the blood volume pulse (BVP) signal was usually extracted from the RGB temporal traces using blind source separation (BSS) such as ICA [5] or PCA [6] or more recently PVM [17]. Another class of methods focuses on exploiting the physical characteristics of skin tissue and its distinctive interaction with light. Methods such as CHROM [2], POS [8] and PBV [7] fall into this category and are based on a detailed

formulation of models based on physiological properties of the skin. The principal advantages of these methods lies in their computational simplicity, owing to their analytic formulation.

As discussed in different review papers (*e.g.* [18], [19]), many rPPG methods implement a general pipeline-based framework: regions of interest (ROI) are detected and tracked over frames, RGB channels are then combined to estimate the pulse signal, which is filtered and analyzed to extract physiological parameters such as heart rate or respiration rate. This pipeline-based framework emphasizes the importance of the common first step of ROI segmentation. Usually, ROI segmentation is based on the result of classical face detection [5] and tracking algorithms and possibly refined with skin pixel classification [9], [20]. As an alternative to this pipeline approach, data-driven methods exploit the pulse-signal as a feature to segment the ROI in an unsupervised manner using voxels [9] or temporal superpixels [1], [3] video segmentation. They are called *unsupervised* rPPG methods to emphasize the difference with methods that require a trained classifier to determine the ROI.

For these so-called unsupervised techniques, the quality of the different ROIs is estimated in order to identify skin areas [9] or to favor areas where information is predominant [3], [12]. The signal-to-noise ratio (SNR), defined as the ratio of the power of the main pulsatile component and the power of background noise, is generally used for this purpose. However, because of the video decomposition in voxels or temporal superpixels, a large number of FFT which has a complexity of $\mathcal{O}(n \log n)$, has to be calculated. Moreover, with this frequency domain index, the shape of the temporal signal is not considered. However, in the case of rPPG signals, the temporal form of the signal is characteristic and is an indication of signal quality.

From these observations, we propose a probabilistic framework to model the rPPG signals and derive a signal quality metric in the temporal domain. Interestingly, the model encapsulates the difference in duration between the growth and decay phases of the different stages of the cardiac cycle and has a complexity of $\mathcal{O}(n)$.

The rest of the paper is organized as follows. The unsupervised rPPG framework is described in section II. Then the probabilistic rPPG signal quality metric is described in section III. Finally, experiments and results are presented in section IV while conclusion is presented in section V.

Supported by the Conseil Régional de Bourgogne Franche Comté through the plan d'actions régional pour l'innovation (PARI) and the European Union through the PO FEDER-FSE Bourgogne 2014/2020 programs.

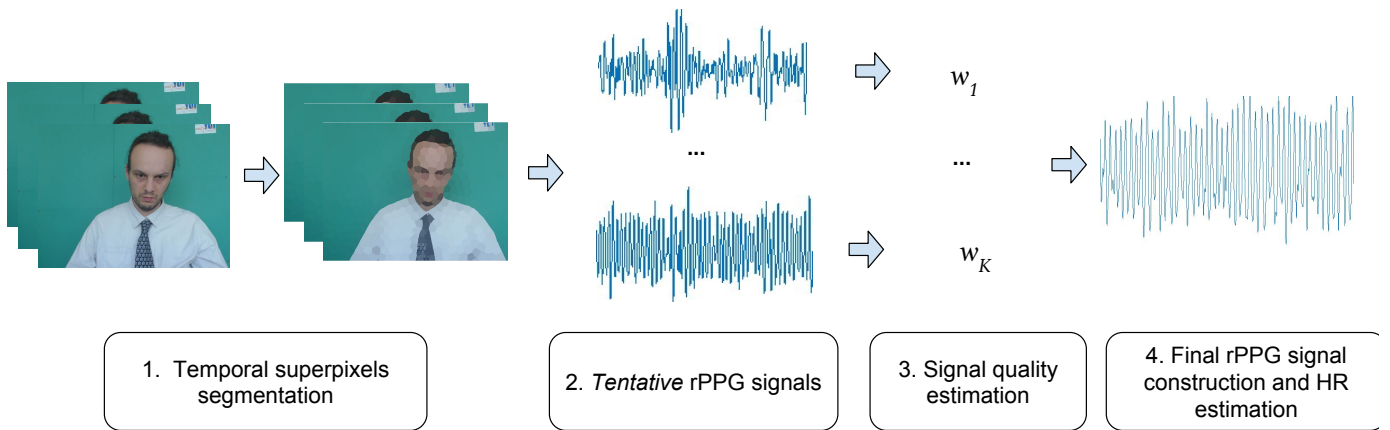


Fig. 1. Flowchart of the temporal superpixel-based heart rate estimation framework: (1) Input video stream is decomposed into temporal superpixels. (2) Tentative rPPG signal is extracted from each temporal superpixel. (3) A signal quality index is estimated for each ROI. (4) A weighted average of all the tentative rPPG signals is finally computed.

II. UNSUPERVISED RPPG FRAMEWORK

The overview of the proposed method is shown in Fig. 1. The algorithm can be decomposed into four main steps. First, the input video frames are decomposed into several temporal superpixels using the IBIS method [1]. The segmentation step is performed by implicitly identifying the superpixel boundaries. Hence, only a fraction of the image is used to perform the segmentation which reduces greatly the computational burden of the process. The segmentation quality remains comparable to state of the art methods while computational time is significantly reduced. Second, a *tentative* rPPG signal is estimated from each temporal superpixel. The term *tentative* means that the estimate can also be done on background superpixels. It is important to note here that this approach can be used with any existing rPPG method. For the sake of simplicity, the CHROM method [2] has been selected for this work. Indeed, this method consists of a simple linear combination of the RGB channels and is therefore very fast to calculate. In our case, this computational simplicity is a very interesting property because this operation is performed for all temporal superpixels.

Finally, a signal quality measure, *i.e.* a pulsatility measure, is estimated for each *tentative* pulse signal to determine contributive signals and a weighted average of all the signals is computed where the weights are given by the pulsatility measure. Since the most pulsatile signals provide high quality information, *i.e.* higher weight, this method implicitly selects regions of interest that represents living skin tissue and favors areas where the pulse trace is more predominant. This method is employed using a sliding window and consecutive resultant signals are then overlapped.

III. HIDDEN MARKOV MODELS FOR RPPG SIGNAL QUALITY MEASUREMENT

A. rPPG signal modeling

In order to measure the quality of rPPG signals and the weights in the fusion process, it is possible to estimate signal-to-noise ratio (SNR) defined as the ratio of the power of the main pulsatile component and the power of background noise, computed in dB due to the wide dynamic range of the signals. This approach has been used successfully in previous works [3], [12]. However, the estimation of the rPPG signal quality of all the *tentative* signals with a frequency-domain measurement, imposes the computation of a large number of FFT which has a complexity of $\mathcal{O}(n \log n)$. Moreover, with the SNR, the shape of the temporal signal is not considered. However, in the case of rPPG signals, the temporal form of the signal is characteristic and is an indication of the quality of the information. Fig. 2 presents a cardiac pulse signal that highlights the systole and diastole phenomena as well as the dicrotic notch. However, it is often not possible to observe the dicrotic notch in many remote PPG scenarios. In this case, the measured cardiac signal has two distinctive phases: rapid growth corresponding to systole followed by a slow decrease corresponding to the diastole. This difference in duration between the growth and decay phases is not captured by signal quality metrics in the frequency domain (such as SNR).

In this paper, we use a probabilistic framework to model the rPPG signal and derive a signal quality metric in the temporal domain. This approach has been used successfully in BayesHeart [4] where hidden Markov models are used to model temporal patterns of the different stages of the cardiac cycle. The signal is modeled by an alternation of two states

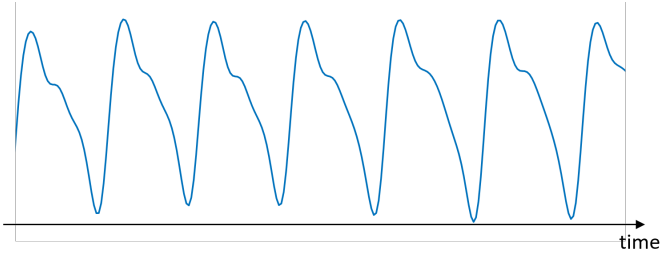


Fig. 2. Characteristic waveform of PPG with systolic and diastolic phases and dicrotic notches.

(growth or decay) and four possible discrete observations (see Fig. 3). O_1 represents increasing observations, O_2 represents local maximum observations, O_3 represents decreasing observations and O_4 represents minimum observations. A two-states cyclic hidden Markov model with a left-to-right state transition structure is used for the model. The hidden Markov model can be characterized by three parameters: $\lambda = \{A, B, \pi\}$. A represents the state transition probability distribution, B the observation symbol probability distribution and π the initial state distribution. The estimation of the transition and emission parameters is performed with the Baum-Welch method on a training dataset and the initial distribution π is determined from the average cycle time in the training data set.

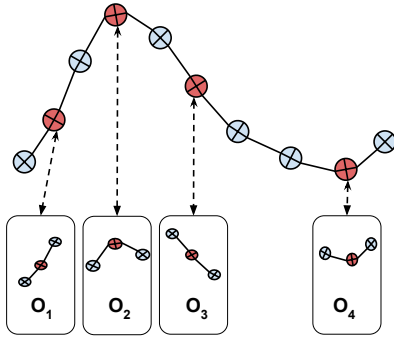


Fig. 3. Possible discrete observations.

B. Cardiac signal quality measure

The quality metric of the rPPG signal is then determined using the Bayesian Information Criterion \mathcal{B} [22] which provides the likelihood of the observations O given the model parameters λ :

$$\mathcal{B}_\lambda = -2 \cdot \ln(P(O|\lambda)) + k \cdot (\ln(n) - \ln(2\pi)). \quad (1)$$

where n is the number of points in the signal. Fig. 4 shows the model likelihood, as defined in (1) and a regular SNR metric estimated on all *tentative* rPPG signals. Thus, it is possible to observe that in both cases the face is correctly selected and that the areas that contribute the most (cheekbones and forehead) have a better score.

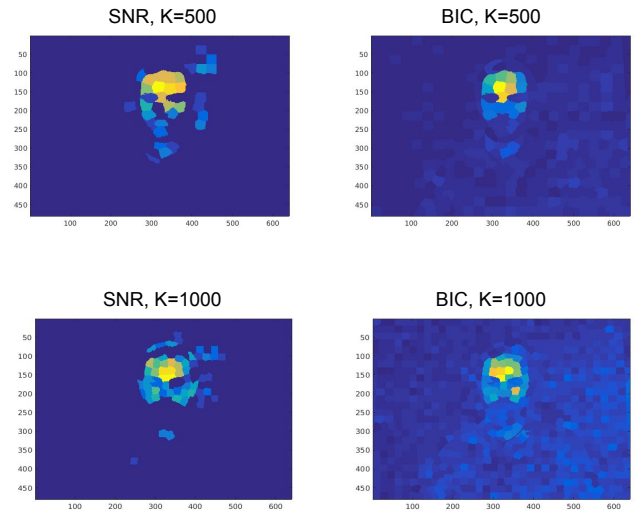


Fig. 4. Comparison of SNR (left) and BIC (right) as a cardiac signal quality measure. Blue means low pulsatility measures and yellow/orange means high. K is the number of temporal superpixels.

Finally, in order to determine the quality of the *tentative* pulse signals, we use two two-state models. The first model is trained on a set of rPPG signals and is denoted λ_s . The second model is an equiprobable model and is denoted λ_n which models a random signal. The quality index v is eventually defined by:

$$v = 1 - \frac{\mathcal{B}_{\lambda_s}}{\mathcal{B}_{\lambda_n}}. \quad (2)$$

We have empirically observed that it is often better to keep only the $x\%$ best contributions. We therefore define \tilde{v} which represents the set of v such that:

$$\tilde{v} = \{v, v > \max_i(v_i) \cdot \frac{x}{100}\}, \quad (3)$$

with i the i^{th} temporal superpixel and $\max_i(v_i)$ the most important contribution over all temporal superpixels. Weights are then obtained normalizing \tilde{v}_i with:

$$w_i = \frac{\tilde{v}_i}{\sum_{i \in K} \tilde{v}_i}, \quad (4)$$

with K the number of temporal superpixels. Finally, the fusion is carried out using the hidden states $H_i(t)$ determined by the Viterbi method to obtain, for each temporal superpixel, the most probable sequence of states that led to the provided observations. This method acts as a filtering and helps to mitigate disturbances. The final pulse signal $S(t)$ is constructed with the weighted average of the state sequences of the different temporal superpixels weighted by the weights w_i :

$$S(t) = \sum_{i \in K} w_i H_i(t). \quad (5)$$

IV. EXPERIMENTS AND RESULTS

The in-house UBFC-RPPG database [12] is used to evaluate the proposed cardiac signal quality metric in the unsupervised rPPG framework. The *REALISTIC* subset of this database, used in this work, comprises of 46 videos where the subjects were required to play a time sensitive mathematical game in order to vary the heart rate and simultaneously emulate the scenario of the typical activity of using a computer. All the videos were taken under ambient light with limited illumination variations. The experimental setup with sample images is depicted in Fig. 5. The UBFC-RPPG database is made publicly available along with the ground truth data from the pulse oximeter for rPPG measurement analysis¹.

The video frames were obtained with a custom C++ application using a Logitech C920 web camera placed at about 1m from the subject with a resolution of 640x480 in 8-bit uncompressed RGB format at approximately 30 frames per second. A CMS50E transmissive pulse oximeter was used to obtain the ground truth PPG data.

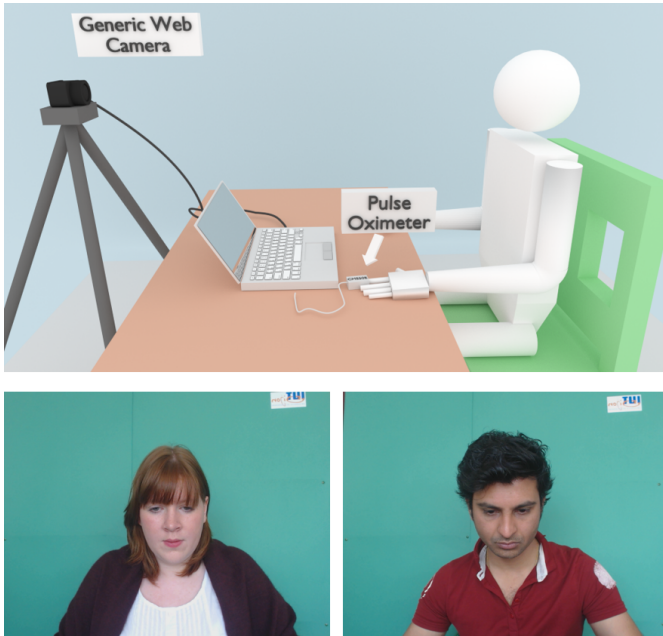


Fig. 5. Experimental Setup (top) and sample images from the UBFC-RPPG database (bottom).

The estimation of the model parameters is performed with the Baum-Welch method [21] on the *SIMPLE* subset of the UBFC-RPPG dataset. This subset is composed of 7 simple videos in which the volunteers were asked to sit still. The transition matrix A and the emission matrix B are initialized with:

$$\begin{aligned} A &= \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}, \\ B &= \begin{pmatrix} 0.6 & 0.1 & 0.3 & 0.1 \\ 0.3 & 0.1 & 0.6 & 0.1 \end{pmatrix}. \end{aligned} \quad (6)$$

¹<https://sites.google.com/view/ybenezeth/ubfcrppg>

Thus, the probability of transition from one state to the other is set to 20% while the probability of observation of the symbols are function of the shape of the signal. Since the most represented symbols with a rPPG signal are growth and decay, the emission probabilities are set to 60% for the observation of a growth in state 1 (*i.e.* growth) and 30% for a continuous decay in the state 1. These probabilities are reversed in state 2 (*i.e.* decay). The probabilities of observing a concave or convex curve variation are necessarily lower and are thus set to 10%.

We give below the matrices obtained after the training:

$$\begin{aligned} A &= \begin{pmatrix} 0.8694 & 0.1306 \\ 0.1298 & 0.8702 \end{pmatrix}, \\ B &= \begin{pmatrix} 0.9192 & 0.0808 & \sim 0 & \sim 0 \\ \sim 0 & 0.0480 & 0.8176 & 0.1345 \end{pmatrix}. \end{aligned} \quad (7)$$

Table I shows the heart rate estimation accuracy comparison between the proposed signal quality metric based on the probabilistic framework and the regular SNR metric for the fusion process. The evaluation metrics used are Mean Absolute Error (*MAE*), Root Mean Square Error (*RMSE*), Pearson's correlation coefficient (r) and *Precision* between heart rate calculated using the rPPG signal and the heart rate calculated using the ground truth PPG waveform. The *Precision* metric represents the percentage of estimations where the absolute error is under a threshold (2.5 or 5 bpm).

TABLE I
RESULTS FOR HEART RATE ESTIMATION WITH THE PROPOSED BIC SIGNAL QUALITY METRIC COMPARED TO THE REGULAR SNR METRIC.

Metrics	SNR	BIC
Precision 5 bpm	0.940	0.975
Precision 2.5bpm	0.899	0.951
RMSE	4.00	2.41
MAE	2.05	1.21
r	0.83	0.82

In order to obtain these results, the CHROM method [2] was selected for its simplicity, we use a resolution of 150 temporal superpixels and $x = 5\%$. Values in bold indicate the best result for each metric. It can first be noted that the average error *MAE* and the squared error *RMSE* are both lower with BIC. That means that the number of errors is lower and that the wrong estimates have values which are closer from the ground truth with BIC. However, the results are very close, and the heart rate estimation is very good in both cases using this dataset.

Fig. 6 shows the robustness to motion of the two signal quality indices. The background is generally very efficiently eliminated by the SNR metric and the difference between the weights is larger. However, the movements and deformation of superpixels tend to bring out false signal information induced by the variations of superpixel boundaries. On the contrary, BIC quality metric can be more resilient to these disturbances.

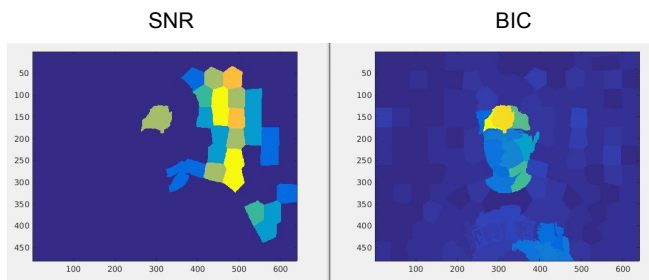


Fig. 6. Example of results obtained with SNR (left) and BIC (right).

V. CONCLUSION

In the present study, we have described and evaluated a new cardiac signal metric in an unsupervised rPPG framework in which the regions of interest segmentation is performed by implicitly selecting living skin tissue via their distinct pulsatility feature. Photoplethysmogram signals are estimated with the weighted fusion of several *tentative* rPPG signals computed on a set of temporal superpixels. The cardiac signal quality metric is based on a probabilistic framework where the characteristic shape of the temporal signal is considered in the hidden Markov model. It has been shown that the fusion of several rPPG signals estimated on several regions of interest generally improves the quality of the resulting cardiac signal (e.g. [12]). However, this approach can be very intensive in calculations because the operations performed on each area are repeated several times. Therefore, it is interesting to note that the complexity of the proposed metric is $\mathcal{O}(n)$ whereas the complexity of the FFT (necessary for the SNR) is $\mathcal{O}(n \log n)$.

Based on a publicly available dataset of 46 videos, the results of this study have demonstrated that the probabilistic metric outperforms the regular SNR metric with a lower computation complexity.

With regard to future developments, it would be interesting to quantify the saving of computing time more precisely and to study the possibility of using a four-states model by considering the dicrotic notch for scenarios where the signal is of good quality (for example in controlled environments with professional cameras). Next, we also plan to examine the use of this measure to analyze other biomedical signals.

REFERENCES

- [1] S. Bobbia, D. Luguern, Y. Benezeth, K. Nakamura, R. Gomez and J. Dubois, "Real-Time Temporal Superpixels for Unsupervised Remote Photoplethysmography", IEEE Conf. on Computer Vision and Pattern Recognition Workshops, 2018.
- [2] G. de Haan and V. Jeanne, "Robust pulse rate from chrominance-based rppg". IEEE Trans. on Biomedical Engineering, 2013.
- [3] S. Bobbia, Y. Benezeth, J. Dubois, "Remote Photoplethysmography Based on Implicit Living Skin Tissue Segmentation", IEEE International Conference on Pattern Recognition, 2016.
- [4] F. Xiangmin and J. Wang, "Bayesheart: A probabilistic approach for robust, low-latency heart rate monitoring on camera phones", ACM International Conference on Intelligent User Interfaces, 2015.
- [5] M.Z. Poh, D.J. McDuff and R.W. Picard, "Non-contact automated cardiac pulse measurements using video imaging and blind source separation", Optics Express, 2010.

- [6] M. Lewandowska et al., "Measuring pulse rate with a webcam - a noncontact method for evaluating cardiac activity", in Computer Science and Information Systems (FedCSIS), Federated Conference on, 2011.
- [7] G. de Haan and A. van Leest, Improved motion robustness of remote-ppg by using the blood volume pulse signature, Physiological Measurement, 2014.
- [8] W. Wang et al., Algorithmic principles of remote-ppg, IEEE Trans. Biomedical Engineering, 2017.
- [9] W. Wang et al., Unsupervised subject detection via remote ppg, IEEE Trans. on Biomedical Engineering, 2015.
- [10] G. Gibert et al., Face detection method based on photoplethysmography, in Advanced Video and Signal Based Surveillance, 2013.
- [11] F. Bousefsaf, C. Maaoui, A. Pruski, "Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate", Biomedical Signal Processing and Control, 2013.
- [12] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, J. Dubois, "Unsupervised skin tissue segmentation for remote photoplethysmography", Pattern Recognition Letters, Elsevier, 2017.
- [13] W. Wang, S. Stuijk and G. de Haan, "A Novel Algorithm for Remote Photoplethysmography: Spatial Subspace Rotation", IEEE Trans. on Biomedical Engineering, 2015.
- [14] W. Verkrusysse, L. O. Svaasand, J. S. Nelson, "Remote plethysmographic imaging using ambient light", Optics express, 2008.
- [15] B.S. Kim and S.K. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis", IEEE Trans. on Biomedical Engineering, 2006.
- [16] Y. Sun, S. Hu, V. Azorin-Peris, S. Greenwald, J. Chambers and Y. Zhu, "Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise", Journal of Biomedical Optics, 2011.
- [17] R. Macwan, S. Bobbia, Y. Benezeth, J. Dubois, A. Mansouri, "Periodic Variance Maximization using Generalized Eigenvalue Decomposition applied to Remote Photoplethysmography estimation", IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018.
- [18] Y. Sun and N. Thakor, "Photoplethysmography revisited: from contact to noncontact, from point to imaging". IEEE Trans. on Biomedical Engineering, 2016.
- [19] D.J. McDuff, J.R. Estepp, A.M. Piasecki and E.B. Blackford, "A Survey of Remote Optical Photoplethysmographic Imaging Methods", int. conf. of the IEEE Engineering in Medicine and Biology Society, 2015.
- [20] R. Macwan, Y. Benezeth, A. Mansouri, Heart rate estimation using remote photoplethysmography with multi-objective optimization. Biomedical Signal Processing and Control, Elsevier, 2019
- [21] L.E. Baum, "An Inequality and Associated Maximization Technique in Statistical Estimation for Probabilistic Functions of Markov Processes", in Proceedings of the Third Symposium on Inequalities, 1972.
- [22] G.E. Schwarz, "Estimating the dimension of a model", Annals of Statistics, 1978.