



HAL
open science

Parameter-free adaptive step-size multiobjective optimization applied to remote photoplethysmography

Richard Macwan, Yannick Benezeth, Keisuke Nakamura, Randy Gomez,
Yadong Wu, Alamin Mansouri

► **To cite this version:**

Richard Macwan, Yannick Benezeth, Keisuke Nakamura, Randy Gomez, Yadong Wu, et al.. Parameter-free adaptive step-size multiobjective optimization applied to remote photoplethysmography. IEEE International Conference on Biomedical and Health Informatics, 2018, Las Vegas, United States. hal-01678241

HAL Id: hal-01678241

<https://u-bourgogne.hal.science/hal-01678241>

Submitted on 9 Jan 2018

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.

Parameter-free adaptive step-size multiobjective optimization applied to remote photoplethysmography

Richard Macwan¹, Yannick Benezeth¹, Keisuke Nakamura², Randy Gomez², Yadong Wu³, Alamin Mansouri¹

¹ Le2i FRE2005, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté, Dijon, France

² Honda Research Institute, Japan, 8-1 Honcho Wako-shi Saitamaken, Japan

³ Southwest University of Science and Technology, Mianyang, Sichuan, China

Abstract—In this work, we propose to reformulate the objective function of Independent Component Analysis (ICA) to make it a better posed problem in the context of Remote photoplethysmography (rPPG). In recent previous works, linear combination coefficients of RGB channels are estimated maximizing the non-Gaussianity of ICA output components. However, in the context of rPPG a priori knowledge of the pulse signal can be incorporated into the component extraction algorithm. To this end, the contrast function of regular ICA is extended with a measure of periodicity formulated using autocorrelation. This novel semi-blind source extraction method for measuring rPPG has the interesting property of being free from manual parameter adjustment. The tedious selection of the step-size parameter in the gradient-ascent algorithm has been advantageously replaced by an adaptive step size. Our method has been validated against our large in-house video database *UBFC-RPPG*.

I. INTRODUCTION

Photoplethysmography (PPG) is a gold standard for monitoring the proper functioning of the cardiovascular system by measuring changes in blood volumes in tissues. Remote photoplethysmography (rPPG) has recently been developed to measure, without any contact, physiological signs only based on the ambient light and a video camera. This has the advantage of reducing user constraints while simultaneously eliminating the need for expensive and specialist hardware. Most rPPG methods share a common pipeline-based framework where regions of interest are first detected and tracked over frames, RGB channels are then combined to estimate the pulse signal, which is then filtered and analyzed to extract physiological parameters such as heart rate or respiration rate. Two recent review papers have been published presenting an overview of the wide range of rPPG methods [17] and [13].

The pulsatile information is mixed in the light reflected by the tissue with other signals such as incoming light changes or shadow casting variations. This mixed signal is then captured by the camera. This suggests that Blind Source Separation (BSS) techniques, such as Independent Component Analysis (ICA) can separate the different sources and isolate the pulse signal [16]. If the time varying color traces for n channels captured by the camera are represented as $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, which is an instantaneous linear mixture of the original m independent signals denoted as $\mathbf{c} = (c_1, c_2, \dots, c_m)^T$, then the process of mixing can be formulated as $\mathbf{x} = \mathbf{A}\mathbf{c}$, where the matrix $\mathbf{A}_{n \times m}$ represents

the linear memoryless mixing of the channels. The goal of ICA, then, is to estimate the unmixing matrix $\mathbf{W}_{m \times n}$ to recover all the independent components from the observed signal with no knowledge of \mathbf{A} and \mathbf{c} . The recovered signal $\mathbf{s} = (s_1, s_2, \dots, s_m)^T$ is given by $\mathbf{s} = \mathbf{W}\mathbf{x}$ [11].

Incorporating a priori information to guide the optimization process is an interesting approach in signal separation. Lu and Rajapakse [11] have used an existing reference signal to guide the separation process by using the method of Lagrange multipliers where the distance between the reference signal and the estimated signal is taken as the constraint. Based on this work, Tsouri *et al.* [19] proposed a constrained ICA based approach, using a rectangular pulse as a reference signal for rPPG measurements. However, the design of an accurate reference signal is prone to multiple complications such as the determination of the frequency of the reference signal. Macwan *et al.* [12] have proposed a more flexible formulation of constrained ICA where a priori knowledge about the periodicity of the blood flow signal is exploited to guide the component extraction process. However, it remains very difficult to select the threshold values for the constraints or the parameters of the optimization algorithm.

In this work, we propose to reformulate the objective function of ICA to make it a better posed problem by making three augmentations. First, we require only one component, i.e., the rPPG pulse from the mixture of the temporal traces. As a result, the problem of component *separation* can be modified into that of component *extraction*. Second, we know that the blood volume pulse embedded in the RGB temporal traces is by nature periodic (or at least pseudo-periodic). Consequently, we use the periodicity of the rPPG signal as an a priori information to help extract the most periodic component. To this end, we use autocorrelation as the measure of periodicity for guiding the ICA separation algorithm. The rPPG pulse extraction is accomplished by using a multi-objective optimization approach to maximize both mean squared autocorrelation and negentropy [8], a measure of non-gaussianity fit for remote photoplethysmography. Third, an adaptive step-size is used in the gradient-ascent optimization algorithm so that the method is free from manual parameter tuning.

The algorithm has been validated with the second dataset of our in-house database *UBFC-RPPG* [2], comprising of 46 videos, which is made publicly available.

II. PROPOSED METHOD

We exploit the inherent periodicity property of biomedical signals guiding the component extraction process to choose the component with the highest periodicity. To this end, we use autocorrelation along with negentropy formulated as a multi-objective optimization problem. The use of autocorrelation as a periodicity measure and formulation of the multi-objective optimization problem is presented next.

A. Autocorrelation as a periodicity measure

Autocorrelation is the correlation of a signal with itself at different lag times. For a time series signal $\mathbf{y} = [y_1, y_2, \dots, y_N]$ of N elements, its discrete autocorrelation r_k at lags $k \in [-(N-1), \dots, N-1]$ is given by

$$r_k = \sum_{j=0}^{N-1} y_j \odot y_j^k \quad (1)$$

where y_j^k is the j^{th} element of the signal \mathbf{y} lagged (or led if $k < 0$) by k units and padded with zeros to the left (or right if $k < 0$) and \odot is the element-wise multiplication operator. A periodic signal typically has a higher correlation with itself compared to a non-periodic one which can be quantified by the mean of the squared autocorrelation of the signal and consequently can be used as a measure of the periodicity of a signal.

To aid the use of autocorrelation as a periodicity measure and simplify its computation, two modifications need to be made. First, since the autocorrelation is symmetric, we only compute the correlation for lags $k \in [0, \dots, N-1]$. Second, since the correlation at lag 0 is always high, we set the autocorrelation to 0 at lag $k = 0$. Thus, the autocorrelation is given by $\mathbf{r} = [r_0, r_1, \dots, r_{N-1}]$ comprising of N values given by Eq. (1) and $r_0 = 0$.

B. Multi-objective optimization

In this section we describe the formulation of the rPPG extraction module using Multi-objective optimization with Autocorrelation as a periodicity measure and ICA now on referred to as MAICA.

Our two objective functions correspond to negentropy and autocorrelation respectively of the output $\mathbf{y} = \mathbf{w}^T \mathbf{x}$ where the ideal orthogonal row vector $\mathbf{w} \in \mathbb{R}^3$, obtained after optimization, extracts the desired component from the RGB temporal traces \mathbf{x} :

$$\underset{\mathbf{w}}{\text{maximize}} \quad J(\mathbf{y}), R(\mathbf{y}) \quad (2)$$

where $R(\mathbf{y}) = R(\mathbf{w}^T \mathbf{x})$, which is eventually a function of \mathbf{w} , is the mean squared autocorrelation given by

$$R(\mathbf{y}) = E\{\mathbf{r}^2\} \quad (3)$$

with $\mathbf{r} = [r_1, r_2, \dots, r_N]$ being the autocorrelation and r_k is given by Eq. (1). $J(\mathbf{y})$ is the generic contrast function for ICA defined by [9] as $H(\mathbf{y}_{gauss}) - H(\mathbf{y})$. $H(\cdot)$ is the differential entropy and \mathbf{y}_{gauss} is a random variable with a

variance equal to that of the output signal \mathbf{y} . In practice, an approximation of negentropy is used for ease of computation and flexibility given by

$$J(\mathbf{y}) \approx \rho [E\{G(\mathbf{y})\} - E\{G(v)\}]^2 \quad (4)$$

where ρ is a positive constant, v is a Gaussian variable having zero mean and unit variance. $G(\cdot)$ can be any non-quadratic function as suggested by [8].

A simple way to incorporate the a priori information in the optimization problem is to scalarize the multi-objective optimization by forming a single-objective optimization such that the solutions to the scalarized problem are the set of feasible solutions commonly known as Pareto optimal solutions [5]. The linearly scalarized version of the multi-objective contrast function then becomes

$$\underset{\mathbf{w}}{\text{maximize}} \quad \bar{J}(\mathbf{y}) + \bar{R}(\mathbf{y}) \quad (5)$$

where $\bar{J}(\mathbf{y}) = s_1(J(\mathbf{y}) - J_{min})$ and $\bar{R}(\mathbf{y}) = s_2(R(\mathbf{y}) - R_{min})$ are the normalized versions of the respective objective functions in order to compensate for the disparities in scale and $s_1 = \frac{1}{J_{max} - J_{min}}$ and $s_2 = \frac{1}{R_{max} - R_{min}}$. Ideally, the boundary values of the objective functions correspond to their global maximum and minimum values. However, to emulate a live scenario as much as possible, all the processing are performed over a temporal window of 30 seconds. Consequently, J_{min} , R_{min} , J_{max} and R_{max} were calculated by using a sinusoidal signal emulating an ideal blood volume pulse, $y_s = \sin(t)$ where t corresponds to the time coordinates of the current temporal window. This was done by taking the maximum values of the objective functions over the frequency range of human heart rates, $\mathcal{F} \in [0.7, 3]$ Hz. The minimum values were calculated in the same manner for a uniform random signal y_r for the temporal window t .

$$J_{max} = \max_{\mathcal{F}} J(y_s), R_{max} = \max_{\mathcal{F}} R(y_s) \quad (6)$$

$$J_{min} = \min_{\mathcal{F}} J(y_r), R_{min} = \min_{\mathcal{F}} R(y_r) \quad (7)$$

Finally, the entire problem is solved using a first-order iterative optimization algorithm procedure with the following objective function:

$$\mathcal{L} = \bar{J}(\mathbf{y}) + \bar{R}(\mathbf{y}) \quad (8)$$

To find the maximum of \mathcal{L} in Eq. (8), gradient-ascent iteratively adapt \mathbf{w} with

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mu_1 \mathbf{J}'_{\mathbf{w}_k}(\mathbf{y}) + \mu_2 \mathbf{R}'_{\mathbf{w}_k}(\mathbf{y}) \quad (9)$$

where k is the iteration index, μ_1 and μ_2 are the step-size for negentropy and autocorrelation part of the objective function, $\mathbf{J}'_{\mathbf{w}_k}(\mathbf{y})$ and $\mathbf{R}'_{\mathbf{w}_k}(\mathbf{y})$ are the first derivatives of $\bar{J}(\mathbf{y})$ and $\bar{R}(\mathbf{y})$ at step k w.r.t \mathbf{w} .

Following the above equations, the optimization procedure converges to the optimum point \mathbf{w}^* representing the final weighting matrix which is then used to obtain the final rPPG signal.

C. Adaptive step-size

Most BSS algorithms use a fixed step-size parameter. However, use of a fixed step-size creates several problems. To maximize the objective function rapidly, the step-size should be set to a large value when the objective function is small. To get a precise weighting matrix \mathbf{w} , on the contrary, the step-size should be set to a small value. In (9), the step-size selection is even more critical because it defines the relative importance between negentropy and autocorrelation in the update procedure. Several adaptive step-size have been proposed in the literature based, for instance, on an exact line search of μ [22] or solving the optimum step-size that minimizes the cost function with an iterative process based on a gradient method [6]. In this work, we use the adaptive step-size proposed by Nakajima *et al.* [14] where the step-size is set anti-proportional to the first gradient of the cost function. The adaptive step-size method are formulated as follows

$$\mu'_1 = \frac{J(\mathbf{y})}{2 \left\| J'_{\mathbf{w}_k(\mathbf{y})} \right\|^2} \quad \text{and} \quad \mu'_2 = \frac{R(\mathbf{y})}{2 \left\| R'_{\mathbf{w}_k(\mathbf{y})} \right\|^2} \quad (10)$$

where $\|\cdot\|^2$ means the Frobenius norm. Even if the complete procedure is almost entirely free from any manual parameter adjustment, we experimentally observe that a threshold on the value of μ_1 can be useful because the derivative of the negentropy is sometimes quite unstable.

D. System Framework

The workflow of the procedure as depicted in Fig. 1 is presented here. Temporal RGB traces, $\mathbf{x} = [x_1, x_2, x_3]^T$ where each x_m , $m \in [1, 2, 3]$, corresponds to a temporal trace of size N of each channel were generated by frame-wise spatial averaging of the skin pixels. To obtain these skin pixel averages, face detection and tracking was first performed using the Viola-Jones and the Kanade-Lucas-Tomasi implementations provided by the computer vision toolbox of MATLAB. Then, corner detection in the detected face was performed for tracking to crop the face based on facial landmarks. Skin detection as formulated by Conaire *et al.* [3] was then performed to select the candidate pixels which were then spatially averaged to obtain a triplet of RGB values per frame and concatenated to obtain the RGB temporal traces.

These temporal RGB traces were then detrended [18] to remove low frequency trends in the signal. Next, after normalization, two additional preprocessing steps which are generally recommended for ICA to simplify calculations were performed. First, *centering* was performed so that the obtained signal \mathbf{y} in $\mathbf{y} = \mathbf{W}\mathbf{x}$ is zero-mean. Next, *whitening* was performed to ensure that the components were uncorrelated and their variances equal to unity. The traces were then passed to the rPPG extraction module where the MAICA algorithm was used to extract the rPPG signal.

After the rPPG signal was obtained, the per window heart rate was calculated from the highest peak of the FFT filtered within the acceptable range of heart rate $\mathcal{F} \in [0.7, 3]$ Hz over

	MAE	SNR	r
MAICA	3.18	0.06	0.90
MAICAfix	3.41	0.06	0.89
ICA	6.02	-1.11	0.79
PCA	9.65	-3.45	0.67
Green	7.73	-2.78	0.68
CHROM	3.81	-0.93	0.87
POS	4.73	-1.60	0.80
G-R	9.79	-3.10	0.65

TABLE I: Performance comparisons between the various methods using Mean Absolute Error (MAE), Signal-to-Noise Ratio (SNR) and Pearson's correlation coefficient (r)

a 30 second moving window using a step size of 0.5 second for our in-house datasets. All the processing was performed over a 30 seconds window, using the weighting matrix \mathbf{w}_k obtained at window k as an initial estimate for calculation of \mathbf{w}_{k+1} at the next window. This 30 seconds window size was chosen as a trade-off between speed and availability of enough data for convergence. The window-wise heart rate estimations were then smoothed using a Kalman filter. The Kalman filter helped to remove spurious outliers resulting from sudden variation in illumination and/or motion. We present the results of the experiments in the next section.

III. EXPERIMENTS AND RESULTS

The in-house UBFC-RPPG database [2] is used to test the MAICA method. This database comprises of 46 videos where the subjects were required to play a time sensitive mathematical game in order to vary the heart rate and also 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 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 a distance of about 1m from the subject with a resolution of 640x480 in 8-bit uncompressed RGB format at 30 frames per second. A CMS50E transmissive pulse oximeter was used to obtain the ground truth PPG data.

Table I shows the accuracy comparisons between the proposed method MAICA with adaptive step-size, with a fix step-size (MAICAfix), ICA[15] and other state of the art methods, viz., PCA [10], Green [20], CHROM [4], POS [21], and G-R [7]. The metrics used are Mean Absolute Error (MAE), signal-to-noise ratio (SNR) and Pearson's correlation coefficient (r) between heart rate calculated using the rPPG signal and the heart rate calculated using the ground truth PPG waveform [1]. The SNR (dB) was calculated as the ratio of the power of the main pulsatile component of the PPG to that of the background noise to accommodate the wide dynamic range of the signals.

Even if the SNR values are quite low, it is interesting to observe that the MAE values are acceptable for most state-

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

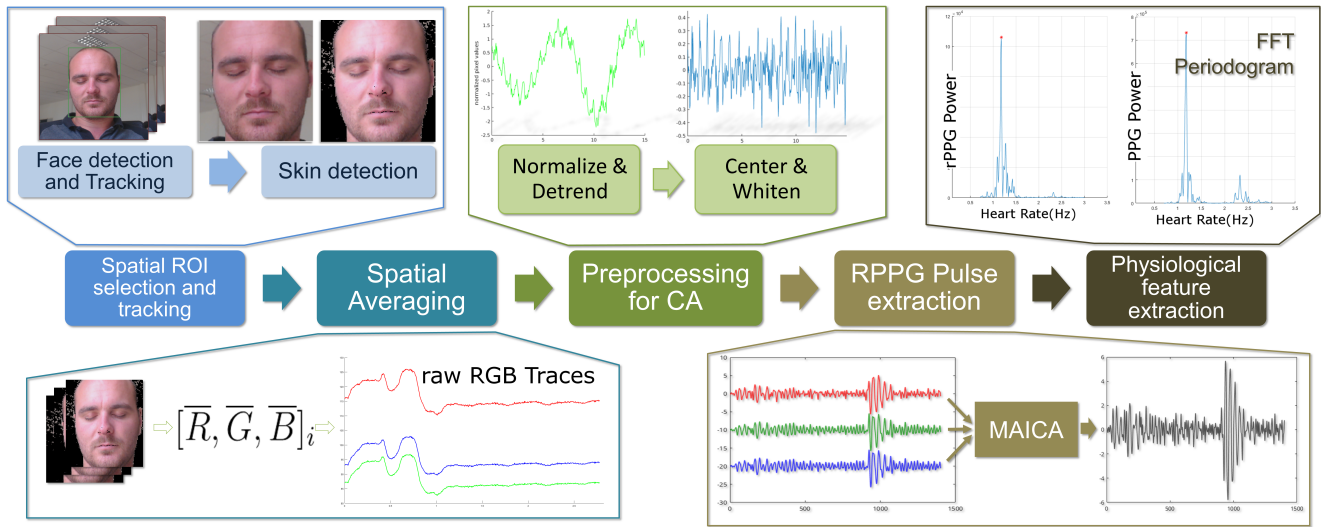


Fig. 1: Flowchart of the proposed method

of-the-art methods. The CHROM method [4] is undoubtedly the most reliable rPPG techniques from the literature because it systematically outperforms other methods. Then, it can be observed that the addition of the autocorrelation term greatly improves ICA. Finally, the results obtained with the adaptive step-size are only slightly better than the ones with a fixed step-size, but the proposed method has the undeniable advantage of not having to empirically fix the optimal value of the threshold.

IV. CONCLUSIONS AND FUTURE WORK

In this paper we presented a parameter-free adaptive step-size multiobjective optimization method applied to remote Photoplethysmography with mean squared autocorrelation and negentropy as the objective functions. The method provides better results than other state of the art methods while removing the extra step for choosing the best parameter values. Our method can also be combined with other methods like smart ROI selection [2] to further obtain better rPPG estimations. The assumption that the most periodic component is the cardiac pulse signal does not hold in scenarios with periodic motion e.g. in fitness. Our method can thus benefit from motion compensation and improve HR estimations in a more realistic scenario with motion disturbances.

REFERENCES

- [1] S. Bobbia, Y. Benezeth, and J. Dubois. Remote Photoplethysmography Based on Implicit Living Skin Tissue Segmentation. *IEEE Int. Conf. on Pattern Recognition (ICPR)*, 2016.
- [2] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, and J. Dubois. Un-supervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognition Letters*, 2017.
- [3] C. Ó. Conaire, N. E. O'Connor, and A. F. Smeaton. Detector adaptation by maximising agreement between independent data sources. *IEEE int. conf. on CVPR*, 2007.
- [4] G. De Haan and V. Jeanne. Robust pulse rate from chrominance-based rPPG. *IEEE Trans. on Biomedical Engineering*, 60(10):2878–2886, 2013.
- [5] K. Deb. Multi-objective optimization using evolutionary algorithms: an introduction. *Multi-objective evolutionary optimisation for product design and manufacturing*, pages 1–24, 2011.
- [6] S. Douglas and A. Cichocki. Adaptive step size techniques for decorrelation and blind source separation. In *IEEE int. conf. on Signals, Systems & Computers*, 1998.
- [7] M. Huelsbusch. *An image-based functional method for opto-electronic detection of skin-perfusion*. PhD thesis, RWTH Aachen dept. of EE, 2008.
- [8] A. Hyvärinen. Fast and robust fixed-point algorithm for independent component analysis. *IEEE Trans. on Neural Networks and Learning Systems*, 1999.
- [9] A. Hyvärinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, 2000.
- [10] M. Lewandowska, J. Ruminski, T. Kociejko, and J. Nowak. Measuring pulse rate with a webcam; A non-contact method for evaluating cardiac activity. *Federated Conf. on Computer Science and Information Systems*, 2011.
- [11] W. Lu and J. C. Rajapakse. Constrained independent component analysis. In *Advances in Neural Information Processing Systems*, 2000.
- [12] R. Macwan, Y. Benezeth, A. Mansouri, K. Nakamura, and R. Gomez. Remote photoplethysmography measurement using constrained ica. *IEEE int. conf. on E-Health and Bioengineering*, 2017.
- [13] D. J. McDuff, J. R. Estep, 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*.
- [14] H. Nakajima, K. Nakadai, Y. Hasegawa, and H. Tsujino. Blind source separation with parameter-free adaptive step-size method for robot audition. *IEEE trans. on audio, speech, and language processing*, 2010.
- [15] 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.
- [16] M. Z. Poh, D. J. McDuff, and R. W. Picard. Advancements in non-contact, multiparameter physiological measurements using a webcam. *IEEE Trans. on Biomedical Engineering*, 2011.
- [17] Y. Sun and N. Thakor. Photoplethysmography Revisited: From Contact to Noncontact, from Point to Imaging. *IEEE Trans. on Biomedical Engineering*, 2016.
- [18] M. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen. An advanced detrending method with application to HRV analysis. *IEEE Trans. on Biomedical Engineering*, 2002.
- [19] G. R. Tsouri, S. Kyal, S. Dianat, and L. K. Mestha. Constrained independent component analysis approach to nonobtrusive pulse rate measurements. *Journal of Biomedical Optics*, 2012.
- [20] W. Verkruysse, L. O. Svaasand, and J. S. Nelson. Remote plethysmographic imaging using ambient light. *Optics express*, 2008.
- [21] W. Wang, A. Den Brinker, S. Stuijk, and G. De Haan. Algorithmic Principles of Remote-PPG. *IEEE Trans. on Biomedical Engineering*, 2016.
- [22] V. Zarzoso and P. Comon. Optimal step-size constant modulus algorithm. *IEEE Trans. on communications*, 2008.