HAL
open science

Heart rate estimation using remote photoplethysmography with multi-objective optimization

Richard Macwan, Yannick Benezeth, Alamin Mansouri

To cite this version:

Richard Macwan, Yannick Benezeth, Alamin Mansouri. Heart rate estimation using remote photoplethysmography with multi-objective optimization. Biomedical Signal Processing and Control, 2019, 49, pp.24 - 33. 10.1016/j.bspc.2018.10.012 . hal-01939085

HAL Id: hal-01939085

https://u-bourgogne.hal.science/hal-01939085v1

Submitted on 30 Nov 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.

Heart rate estimation using Remote Photoplethysmography with Multi-objective Optimization

Richard Macwan, Yannick Benezeth, Alamin Mansouri
Le2i UMR6306, CNRS, Arts et Métiers
Univ. Bourgogne Franche-Comté, Dijon, France
richard.macwan@u-bourgogne.fr, yannick.benezeth@u-bourgogne.fr,
alamin.mansouri@u-bourgogne.fr

1. Introduction

Photoelectric plethysmography or photoplethysmography (PPG) was first introduced in 1937 by Hertzman where variations in the light absorption of human skin were measured by a photoelectric cell [11] placed under a finger illuminated by a light source above it. Since then PPG has been used widely because of its ease of usage, low cost and non-invasiveness. This non-invasiveness has, however, been superseded by that of remote photoplethysmography, henceforth referred to as rPPG, which aims at measuring the same parameters, but *sans contact*.

Verkrussysse *et al.* [28] demonstrated the extraction of remote PPG signals using videos from a simple consumer level camera and that the strongest photoplethysmographic signal was manifested in the G channel of the RGB temporal traces. RGB temporal traces are generated by frame wise quantification, e.g. spatial averaging of skin pixels from the face, and concatenating them. Current research focuses on extracting robust rPPG signals from simple web cameras for which Blind Source Separation (BSS) using Independent Component Analysis (ICA) has been used in many different works [4, 20, 22, 23].

ICA is a statistical technique for decomposing a multivariate signal into constituent signals assuming that the input signals are independent [14]. The problem of rPPG measurements is posed as a signal separation problem where the rhythmic cardiac pulse, appearing as variations in skin color, is assumed to be linearly mixed into the temporal traces of color data from cameras. If the time varying color traces for n channels are represented as $\mathbf{x} = (x_1, x_2, ..., x_n)^T$, which is an instantaneous linear mixture of the original m independent signals denoted as $\mathbf{c} = (c_1, c_2, ..., c_m)^T$, then the process of mixing can be formulated as $\mathbf{x} = \mathbf{Ac}$, 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 demixing 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 $\mathbf{s} = (s_1, s_2, ..., s_m)^T$ is given by $\mathbf{s} = \mathbf{Wx}$ [18].

Owing to this linear formulation, ICA suffers from two unavoidable ambiguities [2, 7]. First, the order of the independent components is indeterminable. A different permutation of the columns of $\mathbf{W}$ will give the same independent components. Second, the exact amplitude and sign of the independent components is also indeterminable. In spite of these limitations, ICA is being frequently used for rPPG measurements.

In this work, we propose to reformulate the objective function of ICA to make it a better posed problem by making two 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*. This requirement is not uncommon and is manifested in various applications. For instance, the On-Off simulation scheme of fMRI experiments [18].

Second, we know that the blood volume pulse embedded in the RGB temporal traces is by definition 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. To compensate for the negative values, the mean of squared autocorrelation is used as the periodicity measure. The rPPG pulse extraction is accomplished by using a multi-objective optimization approach to maximize both mean squared autocorrelation and negentropy [13], a measure of non-gaussianity fit for remote photoplethysmography. Our new method, Multi-objective optimization using Autocorrelation and ICA (MAICA), was validated with our inhouse database UBFC-RPPG, comprising of two datasets of 9 and 46 videos respectively, which is made publicly available along with the ground truth. To the best of our knowledge, this is the first dataset specifically geared towards rPPG analysis. Apart from the inhouse database, the algorithm was also validated against the MMSE-HR database. To gain an intuition on the advantage of incorporating periodicity information into the ICA source separation algorithm, a quick analysis of one of the difficult videos, where traditional ICA is not entirely successful, from our UBFC-RPPG database is presented here. The analysis was carried over sequential temporal windows of 30s and the corresponding weighting matrices that extract the rPPG signal of ICA in figure 1a are compared with those obtained by MAICA in figure 1c. It can be seen from figure 1d that the weighting matrices that simultaneously maximizes negentropy and periodicity indeed result in a heart rate closer to the ground truth heart rate. Moreover, the SNR obtained for the measured signal was improved from -10.89 for ICA to -3.39 for MAICA (also seen by the overall increase in SNR in table 1), proving the advantage of incorporating the periodicity information in the algorithm.

An overview of the related work is presented in section 2 followed by the formulation of the periodicity measure and multi-objective optimization in section 3 and the experiments and results are presented in section 4.

2. Previous Work

As mentioned earlier, rPPG signals have been successfully extracted using simple web cameras using the G channel which contains the relatively stronger plethysmographic signal [28]. One of the first works that used ICA for rPPG measurements comprised of using RGB temporal traces from a simple web camera to extract the cardiac pulse, albeit with limited success under the presence of movement artifacts [22]. Usage of more color channels using a five band camera (RGBCO) with ICA was also investigated by investigated proposing that the Cyan, Green, Orange (CGO) channels were better suited for rPPG measurements [20]. ICA's known caveat of the indeterminacy of the order of the estimated components calls for a heuristic to choose the correct component. Where Poh *et al.* [22] simply selected the second obtained component after applying ICA on RGB temporal traces, McDuff *et al.* [20] chose the signal with the peak of greatest power of the normalized Fast Fourier Transform (FFT) spectrum between 40 and 180 bpm.

In a somewhat related work, Zhang *et al.* [31] used ICA to simultaneously measure heart rate and eye blink frequency, and consequently to separate the eye blink and cardiac signals. In another work, Lewandowska *et al.* [16] used principal component analysis (PCA) and proper channel selection to extract the rPPG pulse signal. Bousefsaf *et al.* take a different approach for motion tolerant rPPG

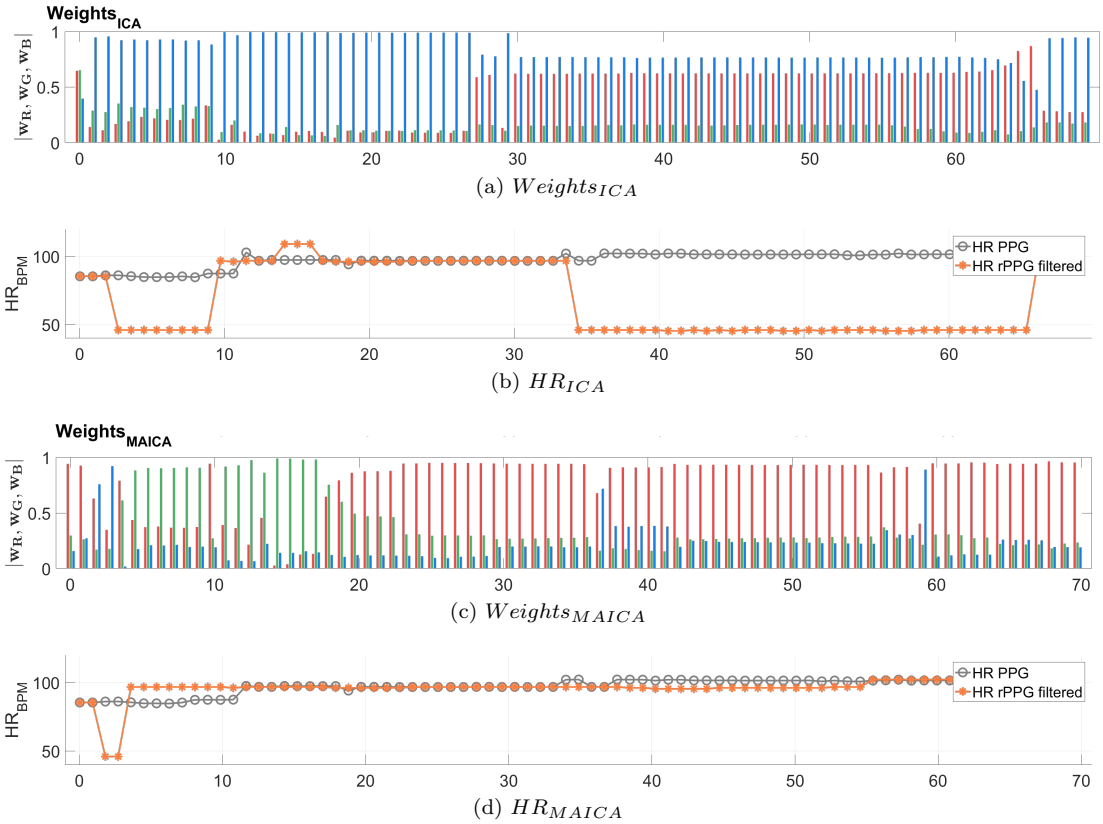


Figure 1: Window-wise weight analysis showing ICA weights and HR in (a) and (b) vs MAICA weights and HR in (c) and (d). For the exact same data, different values of $\mathbf{w} = |w_R, w_G, w_B|$ corresponding to the RGB channels can extract an accurate rPPG signal. Absolute values of the weights are shown for concise display.

signal estimation, using continuous wavelet transform to obtain instantaneous heart rate and heart rate variability. In another work, Gunther *et al.* [10] quasi-periodic process model to extract the heart rate signal with the help of a Toeplitz-structured matrix formulation and a markov process to model the slowly fluctuating reflected light, and finally extract the heart rate signal using the CVX [8], [9] optimization toolbox for Matlab.

De Haan *et al.* [4] introduced chrominance-based methods where two orthogonal chrominance signals were built from the RGB traces in addition to using skin-tone standardization to compensate for illumination variation of different skin colors. They further improved upon the chrominance based methods to show that the different absorption spectra of arterial blood happen along a specific vector in a normalized RGB space, termed as the Blood Volume Pulse vector [5]. Recently, they introduced a mathematical model that incorporates the relevant optical and physiological properties of skin reflection using which they proposed a new algorithm based on the Plane-Orthogonal-to-Skin (POS) which is a plane orthogonal to the skin-tone in the temporally normalized RGB space, suitable for rPPG pulse extraction [30].

Another class of methods focuses on smart ROI selection paradigms. Bobbia *et al.* used temporal superpixels to extract candidate pulse signals which were then merged into a rPPG signal. Li

et al. [17] used face tracking and Normalized Least Mean Square adaptive filtering methods to compensate against motion artifacts. In a related work, Wang *et al.* [29] attempted to extract the rPPG signal by constructing pixel based rPPG sensors to estimate a robust pulse signal using motion compensated pixel-to-pixel pulse extraction based on optical flow vectors.

Many new rPPG measurement algorithms have been introduced recently. An overview of a wide range of rPPG methods has been provided by Sun and Thakur [24] demonstrating the research on rPPG and showing its ubiquity and widespread acceptance. In a similar work, McDuff *et al.* provide a review on state of the art PPG imaging considering measurements other than pulse rate under realistic conditions such as presence of motion artifacts [21]. Kranjec *et al.* also provide a comparative study on heart rate (HR) and heart rate variability measurement between remote PPG using RGB cameras and other methods such as capacitively coupled ECG and HR from speech [15], highlighting their advantages, disadvantages and feasibility.

Incorporating a priori information to guide the optimization process is an interesting approach in signal separation. Lu and Rajapakse [18] 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 to be minimized. Based on this work, Tsouri *et al.* proposed a constrained ICA based approach, using a rectangular pulse as a reference signal for rPPG measurements, [26] which to the best of our knowledge is the closest contribution to ours.

Using a sufficiently accurate reference signal calls for one of two possible requirements. One possibility is to compare the extracted rPPG signal to reference signals of different frequencies (as done in [26]) which is computationally taxing and 30 times slower than traditional ICA. The other possibility is to continously update the frequency of the reference signal during the optimization process which highly reduces the probability of convergence. In rPPG measurements, a PPG signal is such an example the synthesis of which depends critically on the required frequency, even more so than on the actual shape of the signal.

3. Proposed Method

Quasi-periodicity is a ubiquitous property exhibited by various physiological signals such as electrocardiographic, electromyographic, electroencephalographic and photoplethysmography signals. However, this property has been unfairly ignored by most signal extraction problems, especially in remote photoplethysmography measurement scenarios. This inherent property of biomedical signals can be exploited by using a periodicity metric that can guide the component extraction process to choose the most periodic component. Of course, this periodicity metric needs to be combined with the ICA objective function in the case of traditional rPPG measurement, thereby converting the BSS ICA problem to *semi*-Blind Source Extraction (BSE) problem, since we are not entirely blind anymore with regards to the type of signal that needs to be extracted.

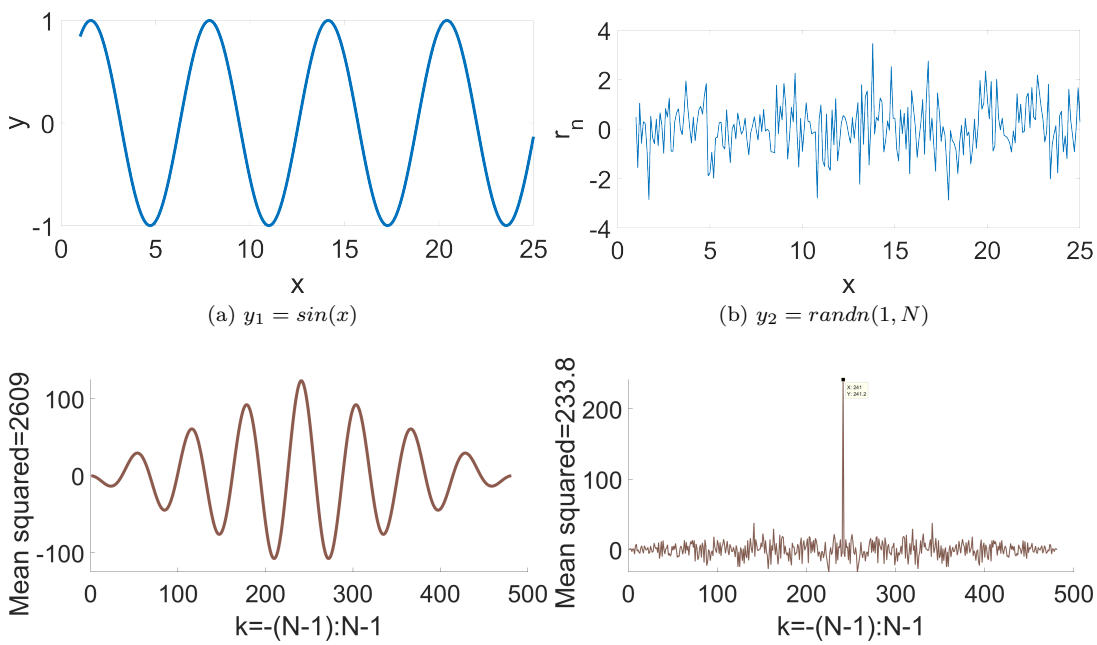
One possibility is to use autocorrelation as a constraint to nudge the algorithm towards selecting components having periodicity higher than a given threshold. However, formulation of such a constraint is critically dependent on the threshold which is quite complex to select, especially in stochastic scenarios such as that of rPPG measurements. A better solution would be to use autocorrelation as one of the objective function thereby maximizing 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.

3.1. Autocorrelation as a periodicity measure

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

$$r_k = \sum_{j=0}^{N-1} y_j \odot \overset{k}{y}_j \tag{1}$$

where $\overset{k}{y}_j$ is the jth element of the signal $\mathbf{y}$ lagged (or led if $k < 0$) by k units and padded with zeroes 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. Figure 2 depicts the high correlation of a periodic sinusoid compared to that of a uniform random signal with the mean of the squared autocorrelation being much higher than that of the random signal.



(a) $y_1 = sin(x)$

(b) $y_2 = randn(1, N)$

(c) Autocorrelation of y_1

(d) Autocorrelation of y_2

Figure 2: Autocorrelation of a sinusoid vs a random 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 $k \in [0, \cdots, N-1]$. 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 $\mathbf{r} = [r_1, r_2, \cdots, r_{N-1}]$ comprising of N values given by equation 1 and $r_0 = 0$. Keeping in mind that r_k is a scalar,

equation 1 can be rewritten in matrix notation as

$$r_k = \mathbf{y}[\overset{k}{\mathbf{y}}]^T = \overset{k}{\mathbf{y}}\mathbf{y}^T \tag{2}$$

where $\overset{k}{\mathbf{y}}$ is again the signal $\mathbf{y}$ lagged by k units and $k \in [1, \cdots, N-1]$. Furthermore, to simplify the derivation of the autocorrelation, $\overset{k}{\mathbf{y}}$ can be rewritten as yT_k where T_k is a toeplitz-like matrix that incorporates the lagging at lag k and padding with zeroes of the signal and is given by

$$T_k = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} 0_{N-k,k} & I_{N-k} \\ 0_{k,k} & 0_{k,N-k} \end{bmatrix} \tag{3}$$

T_k is an $N \times N$ matrix composed of the first $N-k$ rows made up of $(N-k) \times k$ zeroes and an identity matrix of size $N-k$, the rest of the elements being zero. Thus, r_k becomes

$$r_k = \mathbf{y}T_k\mathbf{y}^T \tag{4}$$

making its differential with respect to $\mathbf{y}$ easier to calculate.

The optimization is implemented using a Newton-method like approach while constraining the output to have unit variance with the help of the Lagrange multipliers method. This calls for the calculation of the first and second derivatives of the autocorrelation with respect to the weighting matrix $\mathbf{w}$ which, since they are not trivial, are detailed in the appendix.

3.2. Multi-objective Optimization using Autocorrelation and ICA

Multi-objective optimization problems are generally formed by relaxing the constraints in a constrained optimization problem and interpreting them as additional objective functions. In this section we describe the formulation of the rPPG extraction module using Multi-objective optimization with Autocorrelation as a periodicity measure and ICA (MAICA).

Objective functions. 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 $\mathbf{x}$.

$$\begin{aligned} \text{Maximize} \quad & J(\mathbf{y}),\ R(\mathbf{y}) \\ \text{Subject to} \quad & h(\mathbf{w}) = 0 \end{aligned} \tag{5}$$

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

$$R(\mathbf{y}) = E\{\mathbf{r}^2\} \tag{6}$$

with $\mathbf{r} = [r_1, r_2, \cdots, r_N]$ being the autocorrelation and r_k is given by equation 2. $J(\mathbf{y})$ is the generic contrast function for ICA defined by [14] as $H(\mathbf{y}_{gauss}) - H(\mathbf{y})$. $H(.)$ 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

$$J(\mathbf{y}) \approx \rho[E\{G(\mathbf{y})\} - E\{G(v)\}]^2 \tag{7}$$

where ρ is a positive constant, v is a Gaussian variable having zero mean and unit variance. G(.) can be any non-quadratic function as suggested by [13]. A good general purpose function is given by

$$G(y) = \frac{\log \cos(ay)}{a} \tag{8}$$

with $1 \leq a \leq 2$. Finally, as suggested by [18] the constraint, $h(\mathbf{w}) = E\{\mathbf{y}^2\} - 1 = 0$ was introduced to make sure that objective functions $J(\mathbf{y})$ and $R(\mathbf{y})$ and the weighting vector $\mathbf{w}$ are bounded.

Linear Scalarization. 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 [6]. The linearly scalarized version of the multi-objective contrast function then becomes

$$\begin{aligned} \text{Maximize} \quad & \overline{J}(\mathbf{y}) + \overline{R}(\mathbf{y}) \\ \text{Subject to} \quad & h(\mathbf{w}) = 0 \end{aligned} \tag{9}$$

where

$$\overline{J}(\mathbf{y}) = \frac{J(\mathbf{y}) - J_{min}}{J_{max} - J_{min}} = s_1(J(\mathbf{y}) - J_{min}) \tag{10}$$

and

$$\overline{R}(\mathbf{y}) = \frac{R(\mathbf{y}) - R_{min}}{R_{max} - R_{min}} = s_2(R(\mathbf{y}) - R_{min}) \tag{11}$$

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 was performed over a temporal window of 30 seconds. Using the actual boundary values in this case would not conform to the emulation.

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, $\mathcal{F} \in [.7, 3]$ Hz. The minimum values were calculated in the same manner for a uniform random signal y_r for the temporal window t.

$$J_{max} = \max_{\mathcal{F}} J(y_s), \; R_{max} = \max_{\mathcal{F}} R(y_s) \tag{12}$$

$$J_{min} = \min_{\mathcal{F}} J(y_r), \; R_{min} = \min_{\mathcal{F}} R(y_r) \tag{13}$$

Lagrange multipliers and Multi-objective Optimization. Finally, the entire problem was expressed as a Lagrange multiplier [1] formulation where a Newton-method like approach was used for optimization. The augmented Lagrangian method was used because of its wider applicability and robustness against the equality constraints owing to the penalty term that punishes violations to those constraints, making it more stable than the classical method of Lagrange multipliers. The Augmented Lagrangian for equation 9 is given as

$$\mathcal{L} = [\overline{J}(\mathbf{y}) + \overline{R}(\mathbf{y})] - \lambda h(\mathbf{w}) + \frac{1}{2}\gamma \left\| h(\mathbf{w}) \right\|^2 \tag{14}$$

where λ is the lagrange multiplier for the constraint $h(\mathbf{w})$, $\overline{J(\mathbf{y})}$ and $\overline{R(\mathbf{y})}$ are given by equations 10 and 11 respectively. $\|.\|^2$ denotes the Euclidean norm and the term $\frac{1}{2}\gamma\|.\|^2$ is the penalty term that makes sure that the optimization problem is held at the condition of local convexity assumption: $\nabla^2_{ww}\mathcal{L} > 0$. To find the maximum of $\mathcal{L}$ in equation 14 a Newton-like learning method was used to iteratively adapt $\mathbf{w}$

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \eta(\mathcal{L}''_{\mathbf{w}_k})^{-1}\mathcal{L}'_{\mathbf{w}_k} \tag{15}$$

where k is the iteration index, η is the positive learning rate to avoid uncertainty in convergence and $\mathcal{L}'_{\mathbf{w}_k}$ is the first derivative of $\mathcal{L}$ at step k w.r.t $\mathbf{w}$ given by

$$\mathcal{L}'_{\mathbf{w}_k} = s_1\bar{\rho}E\{\mathbf{x}G'_y(\mathbf{y})\} + s_2E\{R'_{\mathbf{w}}(\mathbf{w})\} - \lambda E\{\mathbf{xy}\} \tag{16}$$

where the sign of $E\{G(\mathbf{y})\} - E\{G(v)\}$ gives the value of $\bar{\rho} = \pm\rho$, $G'_y(\mathbf{y})$ and $R'_{\mathbf{w}}(\mathbf{w})$ are the first derivatives of $G(\mathbf{y})$ and $R(\mathbf{w})$ w.r.t $\mathbf{y}$ and $\mathbf{w}$ respectively. The Hessian $\mathcal{L}''_{\mathbf{w}_k}$ in equation 15, is calculated as

$$\mathcal{L}''_{\mathbf{w}_k} = s_1\bar{\rho}\mathbf{R_{xx}}E\{G''_y(\mathbf{y})\} + s_2E\{R''_{\mathbf{w}}(\mathbf{w})\} - \lambda\mathbf{R_{xx}} \tag{17}$$

the inversion of which is not problematic because $\mathbf{R_{xx}}$ being the covariance matrix of the whitened and centered signal $\mathbf{x}$ is an identity matrix. $G''_y(\mathbf{y})$ and $R''_{\mathbf{w}}(\mathbf{w})$ are second order derivatives and $\mathcal{L}''_{\mathbf{w}_k}$ is of size 3×3. The first and second derivatives of $R(\mathbf{w})$ are not trivial and are presented in the appendix. The optimum multiplier λ^* is obtained iteratively based on a gradient-ascent method [18]:

$$\lambda_{k+1} = \lambda_k + \gamma h(\mathbf{w}_k) \tag{18}$$

Following the above equations, the optimization procedure converges to the optimum point defined by the doublet $(\mathbf{w}^*, \lambda^*)$ representing the tuned parameter and final weighting matrix $\mathbf{w}^*$ which is then used to obtain the final rPPG signal.

3.3. System Framework

The workflow of the procedure as depicted in figure 3 is presented here. Temporal RGB traces, $\mathbf{x} = [x_1, x_2, x_3]^T$ where each x_m, $m \in [1...3]$, corresponds to a temporal trace of size N of each channel and was generated by spatial averaging of the pixels (face-cropped or skin-segmented). To this end, 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. 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 using a smoothness priors approach, with the regularization parameter λ set to 500, proposed by Karjalainen and Pasi [25] to remove low frequency trends in the signal. Next, after normalization, two additional preprocessing steps are generally recommended for ICA to simplify calculations. First, *centering* was performed so that the obtained signal $\mathbf{y}$ in $\mathbf{y} = \mathbf{Wx}$ 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 $\mathcal{F} \in [.7, 3]$ Hz over a 30 second

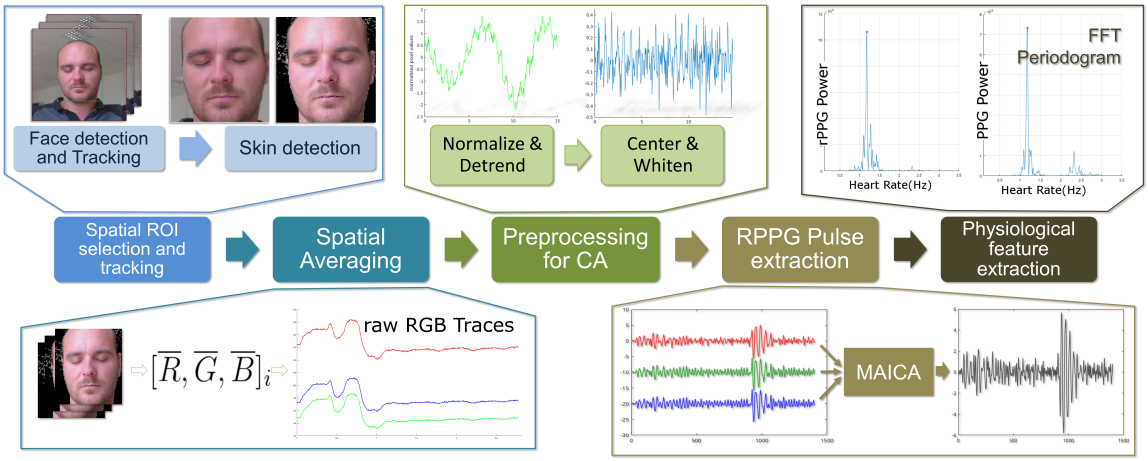


Figure 3: Flowchart of the proposed method

moving window using a step size of 0.5 second for our in-house datasets. Although, ICA is known to work much better with signals of longer duration, as mentioned earlier, all the processing was performed over a 30 second 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 second window size was chosen as a trade-off between speed and availability of enough data for convergence. On the other hand, a 15 second window was needed for the MMSE-HR dataset owing to the shorter length of many constituent videos. The window-wise heart rate estimations were then smoothed using a Kalman filter.

The Kalman filter helped to remove spurious outliers resulting from abrupt variations in illumination and/or motion. The constant velocity motion model was used in the HR measurement with the motion and measurement noise variances set to $\begin{bmatrix}1 & 1\end{bmatrix}$ to reflect a standard deviation of 1 BPM. These values were fixed for the analysis of all the videos in the three different datasets that we used. They were chosen to represent a margin of 1 BPM in the motion and measurement noise models and were verified by trial and error in their ability to discard spurious outliers. The next section presents the experiments and results.

4. Experiments and Results

Two video databases, namely UBFC-RPPG and MMSE-HR were used to test the MAICA method. The inhouse UBFC-RPPG database comprises of two datasets. The first, labeled as SIMPLE, comprises of 9 videos (about 21k frames), where the subjects were requested to relax and close their eyes. The second dataset comprises of 46 videos (about 94k frames), labeled as REALISTIC, 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. Both the datasets comprise of subjects with skin colors varying from dark to light tones. 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

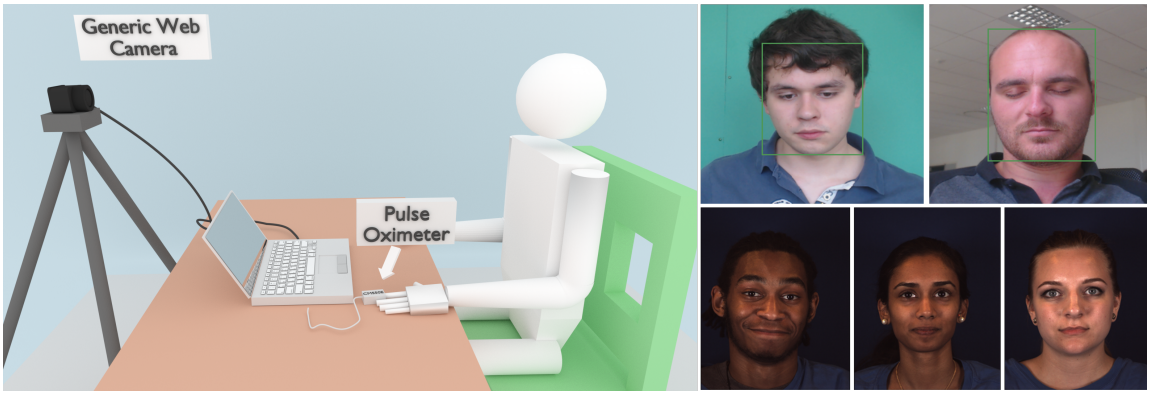


Figure 4: Experimental Setup (left) and sample images from the SIMPLE and REALISTIC datasets from the UBFC-RPPG database (top-right) and the MMSE-HR (bottom-right) database

oximeter for rPPG measurement analysis[1].

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. The experimental setup with sample images from UBFC-RPPG database are shown in figure 4 depicting the lighting conditions.

The MMSE-HR database was also used to test our method, the samples of which are equally shown in figure 4. The dataset comprises of 97 usable videos (about 105k frames) of varying length, at 25 frames per second of varied skin colors. Although, the main objective of the MMSE-HR database was for emotion elicitation, it does provide the video data and HR ground truth data for our validation. Also, MMSE-HR only provides the heart rate, and not the actual PPG waveform of the ground truth as compared to our UBFC-RPPG dataset. Thus, when using the UBFC-RPPG database, a more meaningful validation and comparison can be achieved since the same algorithm can be used to extract the heart rate from both the PPG and the rPPG waveforms.

Figure 5 shows a typical rPPG signal from a simple video and its heart rate estimation, HR_{rPPG} vs HR_{PPG}, the heart rate estimation of the ground truth signal using the Fast Fourier Transform (FFT). The high correlation between the rPPG and PPG signal is clearly visible in the figure, and so is the correlation between HR_{rPPG} and HR_{PPG}.

Figure 6 shows the correlation comparisons between ICA [22] and MAICA for the two databases where HR_{PPG} and HR_{rPPG} are plotted against each other. The metrics PRECIS 2.5 and PRECIS 5 show the percentage of windows where $\delta = |HR_{rPPG} - HR_{PPG}| < 2.5$ and 5 beats per minute (bpm) respectively. MAE corresponds to the average mean absolute error between HR_{rPPG} and HR_{PPG} in bpm calculated over all the windows for all videos. As the name suggests, the analysis of the SIMPLE dataset, which can be thought of emulating a patient at rest, was not challenging and both ICA and MAICA performed well. However, MAICA did result in fewer outliers than ICA.

The REALISTIC and the MMSE-HR datasets were understandably more challenging. However, MAICA demonstrated a better correlation between HR_{PPG} and HR_{rPPG}. It is also evident that

[1] http://ilt.u-bourgogne.fr/benezeth/projects/UBFCRPPG

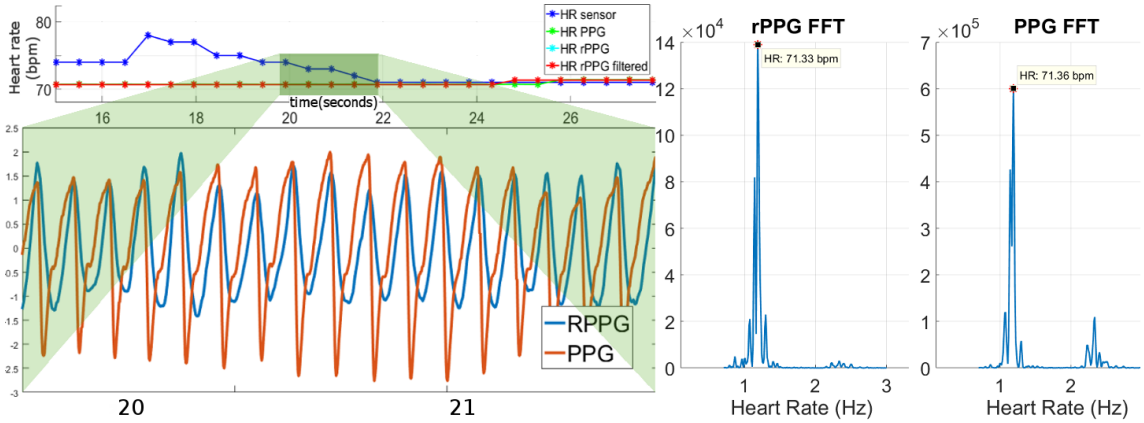
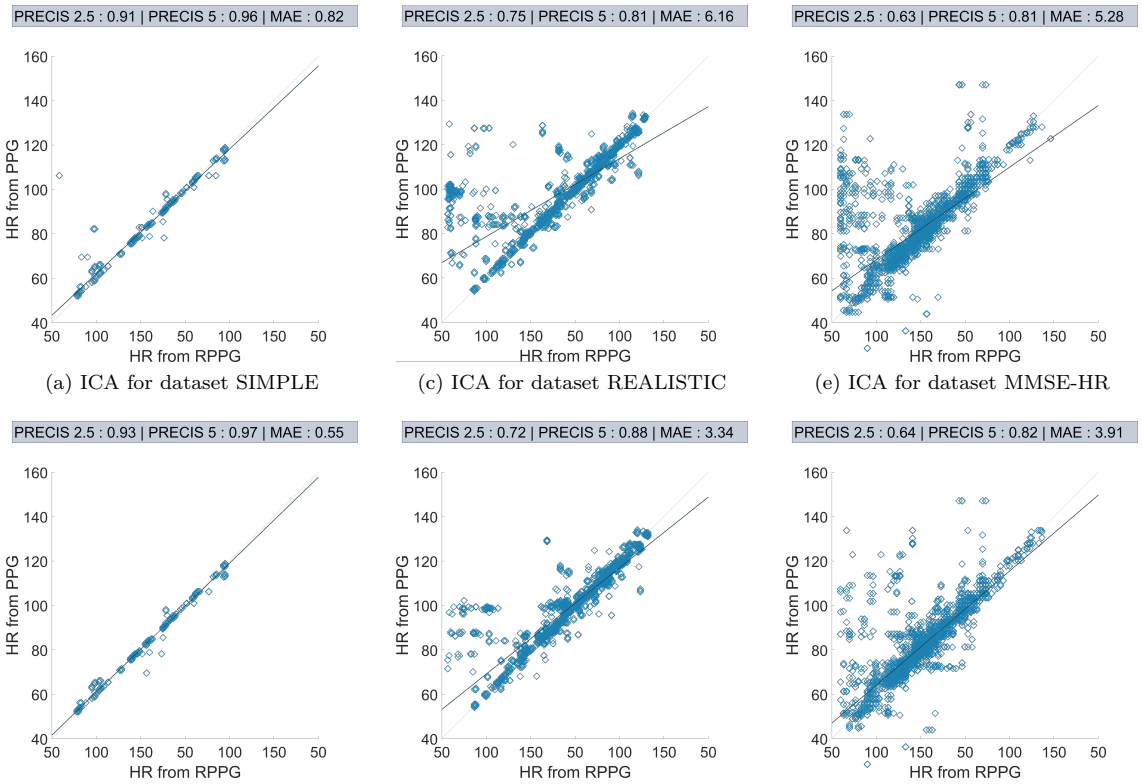


Figure 5: Extracted rPPG signal for video 10 of the SIMPLE dataset of the UBFC-RPPG database. HR_sensor is HR from the finger sensor. HR_rPPG_filtered is HR_rPPG after Kalman filtering



(a) ICA for dataset SIMPLE (c) ICA for dataset REALISTIC (e) ICA for dataset MMSE-HR

(b) MAICA for dataset SIMPLE (d) MAICA for dataset REALISTIC (f) MAICA for dataset MMSE-HR

Figure 6: Correlation Analysis for rPPG

	UBFC-RPPG						MMSE-HR	
	SIMPLE			REALISTIC				
	MAE	SNR	r	MAE	SNR	r	MAE	r
MAICA	**0.55**	2.98	**0.99**	**3.34**	**-0.26**	**0.89**	**3.91**	0.86
ICA	0.82	2.73	0.98	6.16	-1.23	0.76	5.28	0.70
PCA	2.04	-1.43	0.97	9.65	-3.45	0.67	9.15	0.49
Green	9.86	-1.61	0.29	7.73	-2.78	0.68	10.65	0.47
CHROM	0.72	**3.04**	0.99	3.81	-0.93	0.87	5.59	**0.83**
POS	0.67	2.57	0.99	4.73	-1.60	0.80	5.77	0.82
G-R	0.67	1.97	0.99	9.79	-3.10	0.65	8.56	0.58

Table 1: Performance comparisons between the various methods using Mean Absolute Error (MAE), Signal-to-Noise Ratio (SNR) and Pearson's correlation coefficient (r)

the lesser number of outliers were a direct cause of better convergence towards the correct rPPG signal. The MAICA method also provided stronger peaks in the FFT periodogram along with better HR estimates. It also provides a better performance than the Self-Adaptive-Matrix Completion method introduced in [27] whose MAE is 7.6 bpm as compared to our 5.2 bpm. However, this comparison is not very clear since we do not use the same ROI selection method as them. We also provide more rPPG data using our skin masks.

Table 1 shows the accuracy comparisons between ICA and MAICA and other state of the art methods, viz., PCA [16], Green [28], CHROM [4], POS [30], and G-R [12]. The metrics used are apart from MAE are signal-to-noise ratio (SNR) and Pearson's correlation coefficient (r) between heart rate calculated using the rPPG signal, HR_{rPPG} and the heart rate calculated using the ground truth PPG waveform, HR_{PPG}. The windowed method is computationally more taxing, owing to the smaller window length, but is more realistic. 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. However, it is to be noted that the MMSE-HR database does not provide the ground truth waveforms, thereby obliging the use of the main pulsatile component of the RPPG instead of the PPG for the SNR calculation. As a result, the SNR values for the MMSE-HR database are not really relevant and are omitted in table 1. It is worth mentioning here that although the MAICA is not faster than the ICA since one the autocorrelation objective function has to be additionally calculated, its novelty lies in the combination of the periodicity information with the notion of independence used by the ICA algorithm. Of course, optimizations can be easily performed for calculating these objective functions which can improve the computation times drastically. This combination can be also formulated as a Constrained ICA problem as has been done in [19].

The analysis of the SIMPLE dataset, as the name suggests, was quite easy for almost all the methods. However, MAICA did remove the few outliers that are present in the case of ICA. This is reflected in the low MAE and high SNR values as shown in table 1. This can be attributed to the fact that the subjects were generally relaxing, mostly with their eyes closed, which resulted in

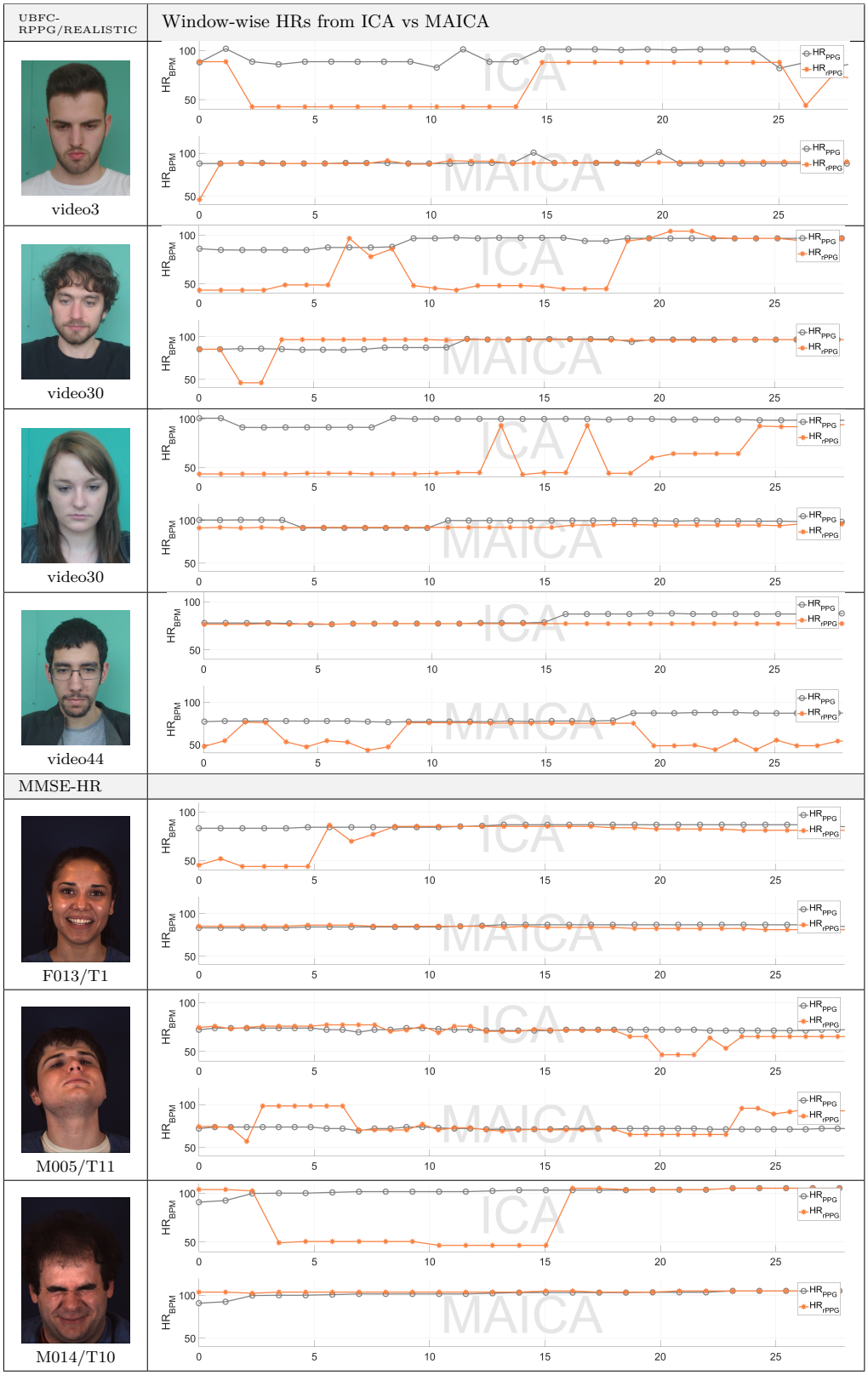


Figure 7: ICA vs MAICA for certain videos from the UBFC-RPPG\REALISTIC and MMSE-HR datasets. The FFT periodogram was used to perform the HR measurements over temporal windows of 30s and 15s respectively for the two datasets owing to the shorter lengths of videos in the latter dataset.

minimal motion artifacts. On the other hand, the REALISTIC dataset was more challenging since the subjects were actually working on the computer and were only requested to keep their hand still for the PPG sensor. Similarly, the MMSE-HR database was challenging owing to it being an emotion elicitation database. There were many instances where the subjects laughed out loud and exhibited considerable movements. This resulted in the usual problems arising from movement of the subjects. Consequently, the presence of outliers was more pronounced for both the REALISTIC dataset and the MMSE-HR database, which MAICA was able to reduce. This can be seen in table 1, where the pearson correlation coefficient was also closer to unity which is also evident from figure 6 where the fitting line was closer to the 45° line as compared to ICA. Finally, in figure 7 we present the window-wise heart rate comparisons between ICA and MAICA, notably for videos where ICA is not entirely successful. The combination of the autocorrelation function with the ICA negentropy function in the premise of multi-objective optimization enables the extraction of a more accurate rPPG signal closer to the ground truth. Videos from the UBFC-RPPG SIMPLE dataset are not presented in this figure since there was not much difference between the performance between ICA and MAICA owing to the simplicity of these videos. With regards to the MMSE-HR dataset, it actually proved to be useful in highlighting the versatility of our method in presence of motion artifacts arising from facial expressions, as is visible in the snapshots in the figure. It is to be noted that almost all the videos of the MMSE-HR dataset manifested changing facial expressions and moving faces.

5. Conclusions and future work

In this paper we presented a novel semi blind source extraction method, MAICA, for the application of rPPG measurements using multi-objective optimization 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 component. The periodogram of the extracted signals was also consistently closer to that of the PPG. Our method can also be combined with other methods like smart ROI selection [17] to further obtain better rPPG estimations.

We also present two datasets of 9 and 46 videos respectively recorded under ambient light. This database is aimed towards testing of rPPG algorithms and is made publicly available. The SIMPLE dataset is not very challenging and can be used as a starting point for basic rPPG algorithms. The REALISTIC dataset emulates an everyday working scenario and can be used as a benchmark for applications aimed at working in live environments.

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. We have not considered constraints related to the unique manner of light scattering on skin due to its physical properties. Such constraints might help to improve the method further. Finally, we average the pixel data (cropped or skin segmented) to obtain a single value and thus loose any spatial information. Higher order analysis which preserves the spatial relationships between pixel neighborhoods can be an important avenue to look into.

6. References

[1] D. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1982. 7

[2] P. Comon. Independent component analysis, A new concept? *Signal Processing*, 36(3):287–314, 1994. 1

[3] C. Ó. Conaire, N. E. O'Connor, and A. F. Smeaton. Detector adaptation by maximising agreement between independent data sources. *CVPR*, 2007. 8

[4] G. De Haan and V. Jeanne. Robust pulse rate from chrominance-based rPPG. *IEEE Transactions on Biomedical Engineering*, 60(10):2878–2886, 2013. 1, 3, 12

[5] G. de Haan and A. van Leest. Improved motion robustness of remote-PPG by using the blood volume pulse signature. *Physiological measurement*, 35(9):1913–1926, 2014. 3

[6] K. Deb. Multi-objective optimization using evolutionary algorithms: an introduction. *Multi-objective evolutionary optimisation for product design and manufacturing*, pages 1–24, 2011. 7

[7] D. Djuwari, D. Kant Kumar, and M. Palaniswami. Limitations of ICA for Artefact Removal. *IEEE Medicine and Biology Society*, 5:4685–4688, 2005. 1

[8] M. C. Grant and S. P. Boyd. Graph Implementations for Nonsmooth Convex Programs. In V. D. Blondel, S. P. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, pages 95–110. Springer London, London, 2008. 3

[9] M. C. Grant and B. Stephen. CVX: Matlab software for disciplined convex programming, version 2.1, 2014. 3

[10] J. Gunther, N. Ruben, and T. Moon. Model-based (passive) heart rate estimation using remote video recording of moving human subjects illuminated by ambient light. *International Conference on Image Processing (ICIP)*, pages 2870–2874, 2015. 3

[11] A. B. Hertzman. Photoelectric Plethysmography of the Fingers and Toes in Man. *Experimental Biology and Medicine*, 37(3):529–534, 1937. 1

[12] M. Hülsbusch. An image-based, functional method for the opto-electronic detection of skin perfusion. *Anesthesia & Analgesia*, 105(6S Suppl):S31, 2008. 12

[13] A. Hyvarinen. Fast and robust fixed-point algorithm for independent component analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 10(3):626–634, 1999. 2, 7

[14] A. Hyvärinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4-5):411–430, 2000. 1, 6

[15] J. Kranjec, S. Beguš, G. Geršak, and J. Drnovšek. Non-contact heart rate and heart rate variability measurements: A review. *Biomedical Signal Processing and Control*, 13(1):102–112, 2014. 4

[16] M. Lewandowska, J. Ruminski, T. Kocejko, and J. Nowak. Measuring pulse rate with a webcam; A non-contact method for evaluating cardiac activity. *2011 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pages 405–410, 2011. 2, 12

[17] X. Li, J. Chen, G. Zhao, and M. Pietik. Remote Heart Rate Measurement From Face Videos Under Realistic Situations. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 4264–4271, 2014. 4, 14

[18] W. Lu and J. C. Rajapakse. ICA with Reference. *Neurocomputing*, 69(16-18):2244–2257, 2006. 1, 2, 4, 7, 8

[19] R. Macwan, Y. Benezeth, and A. Mansouri. Remote photoplethysmography with constrained

ICA using periodicity and chrominance constraints. *BioMedical Engineering Online*, 17(1):22, 2018. 12

[20] D. McDuff, S. Gontarek, and R. Picard. Improvements in remote cardiopulmonary measurement using a five band digital camera. *IEEE Transactions on Biomedical Engineering*, 61(10):2593–2601, 2014. 1, 2

[21] D. J. Mcduff, J. R. Estepp, A. M. Piasecki, and E. B. Blackford. A Survey of Remote Optical Photoplethysmographic Imaging Methods. In *37th Annual International Conference of the IEEE EMBC*, pages 6398–6404, 2015. 4

[22] 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*, 18(10):10762–10774, 2010. 1, 2, 10

[23] M. Z. Poh, D. J. McDuff, and R. W. Picard. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Transactions on Biomedical Engineering*, 58(1):7–11, 2011. 1

[24] Y. Sun and N. Thakor. Photoplethysmography Revisited: From Contact to Noncontact, from Point to Imaging. *IEEE Transactions on Biomedical Engineering*, 63(3):463–477, 2016. 4

[25] M. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen. An advanced detrending method with application to HRV analysis. *IEEE Transactions on Biomedical Engineering*, 49(2):172—-175, 2002. 8

[26] 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*, 21(7):071105, jul 2016. 4

[27] S. Tulyakov, X. Alameda-pineda, E. Ricci, L. Yin, J. F. Cohn, and N. Sebe. Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2396–2404, 2016. 12

[28] W. Verkruysse, L. O. Svaasand, and J. S. Nelson. Remote plethysmographic imaging using ambient light. *Optics express*, 16(26):21434–21445, 2008. 1, 2, 12

[29] W. Wang, S. Stuijk, and G. De Haan. Exploiting spatial redundancy of image sensor for motion robust rPPG. *IEEE Transactions on Biomedical Engineering*, 62(2):415–425, jan 2015. 4

[30] W. Wang, S. Stuijk, and G. De Haan. A Novel Algorithm for Remote Photoplethysmography: Spatial Subspace Rotation. *IEEE Transactions on Biomedical Engineering*, 63(9):1974–1984, 2016. 3, 12

[31] C. Zhang, X. Wu, L. Zhang, X. He, and Z. Lv. Simultaneous detection of blink and heart rate using multi-channel ICA from smart phone videos. *Biomedical Signal Processing and Control*, 33:189–200, 2017. 2

[32] Z. Zhang, J. M. Girard, Y. Wu, X. Zhang, P. Liu, U. Ciftci, S. Canavan, M. Reale, A. Horowitz, H. Yang, J. F. Cohn, Q. Ji, and L. Yin. Multimodal Spontaneous Emotion Corpus for Human Behavior Analysis. *IEEE Conference on CVPR*, pages 3438–3446, 2016.

Appendix A. Derivatives of $R(\mathbf{w})$

Here we present the first and second derivatives of $R(\mathbf{w})$ needed by the Lagrange multipliers method. We follow the convention that the derivative of a scalar w.r.t a column vector is a column vector of the same size as that of the vector. The first derivative of $R(\mathbf{w})$ in

$$R(\mathbf{y}) = E\{\mathbf{r}^2\} \tag{A.1}$$

can be obtained as follows considering squared autocorrelation as $\mathbf{r}^2 = [r_1^2\, r_2^2 \ldots r_N^2]$.

$$R'(\mathbf{w}) = -E\{\frac{\partial}{\partial \mathbf{w}}(\begin{bmatrix} r_1^2 & r_2^2 & \cdots & r_N^2 \end{bmatrix})\} \tag{A.2}$$

where the derivative of the squared autocorrelation $\mathbf{r}^2$ is then obtained using the chain rule of derivatives. Also, we have $\mathbf{y} = \mathbf{w}^T\mathbf{x}$ giving $\frac{\partial \mathbf{y}}{\partial \mathbf{w}} = \mathbf{x}$.

$$\begin{aligned}
\frac{\partial(\mathbf{r}^2)}{\partial \mathbf{w}} &= \mathbf{x}\frac{\partial(\mathbf{r}^2)}{\partial \mathbf{y}} && \text{(A.3)}\\
&= \mathbf{x}\begin{bmatrix} \frac{\partial(r_1^2)}{\partial y_1} & \frac{\partial(r_2^2)}{\partial y_1} & \cdots & \frac{\partial(r_N^2)}{\partial y_1} \\ \frac{\partial(r_1^2)}{\partial y_2} & \frac{\partial(r_2^2)}{\partial y_2} & \cdots & \frac{\partial(r_N^2)}{\partial y_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial(r_1^2)}{\partial y_N} & \frac{\partial(r_2^2)}{\partial y_N} & \cdots & \frac{\partial(r_N^2)}{\partial y_N} \end{bmatrix} && \text{(A.4)}\\
&= \mathbf{x}\begin{bmatrix} 2r_1\frac{\partial r_1}{\partial y_1} & \cdots & 2r_N\frac{\partial r_N}{\partial y_1} \\ \vdots & \ddots & \vdots \\ 2r_1\frac{\partial r_1}{\partial y_N} & \cdots & 2r_N\frac{\partial r_N}{\partial y_N} \end{bmatrix} && \text{(A.5)}\\
&= 2\mathbf{x}\begin{bmatrix} r_1\frac{\partial r_1}{\partial y_1} & \cdots & r_N\frac{\partial r_N}{\partial y_1} \\ \vdots & \ddots & \vdots \\ r_1\frac{\partial r_1}{\partial y_N} & \cdots & r_N\frac{\partial r_N}{\partial y_N} \end{bmatrix} && \text{(A.6)}
\end{aligned}$$

The size of $\frac{\partial(\mathbf{r}^2)}{\partial \mathbf{w}}$ is then $3 \times N$ from the product of $\mathbf{x}_{3\times N}$ with the jacobian of size $N \times N$. Consequently, its expectation ends up having a size of 3×1 since it is nothing but a temporal mean over N samples. The jacobian in equation A.6 can be concisely expressed as $\begin{bmatrix} r_1\frac{\partial r_1}{\partial \mathbf{y}} & r_2\frac{\partial r_2}{\partial \mathbf{y}} & \cdots & r_N\frac{\partial r_N}{\partial \mathbf{y}} \end{bmatrix}$ where each column is the product of the derivative $\frac{\partial r_1}{\partial \mathbf{y}}$ and the scalar r_k and is of size $N \times 1$. Deriving $r_k = \mathbf{y}T_k\mathbf{y}^T$ w.r.t $\mathbf{y}$ using the product rule of differentiation,

$$\begin{aligned}
\frac{\partial r_k}{\partial \mathbf{y}} &= \mathbf{y}\frac{\partial}{\partial \mathbf{y}}(T_k\mathbf{y}^T) + \mathbf{y}T_k\frac{\partial}{\partial \mathbf{y}}(\mathbf{y}^T)\\
&= \mathbf{y}\frac{\partial}{\partial \mathbf{y}}(\mathbf{y}T_k^T) + \mathbf{y}T_k\\
&= \mathbf{y}T_k^T + \mathbf{y}T_k = \mathbf{y}(T_k^T + T_k) && \text{(A.7)}
\end{aligned}$$

where $\frac{\partial}{\partial \mathbf{y}}(T_k\mathbf{y}^T) = T_k^T$ comes from the fact that the differential of $T_k\mathbf{y}^T$, a vector, will remain the same even when it is transposed and the derivative is computed element-wise. This result is owing to the fact that T_k is not symmetric. If it were symmetric, then the result would have been $2\mathbf{y}T_k$.

For conciseness, we will represent the sum $T_k + T_k^T$ as $\mathbf{T}_k$. Finally to be consistent with our convention, using the same argument of the differential being immutable under transpositions, the row vector $\frac{\partial r_k}{\partial \mathbf{y}}$ can be transposed into a column vector and the matrix $\frac{\partial \mathbf{r}}{\partial \mathbf{y}}$ can be built as

$$\frac{\partial \mathbf{r}}{\partial \mathbf{y}} = \begin{bmatrix} r_1\mathbf{T}_1\mathbf{y}^T & \cdots & r_N\mathbf{T}_N\mathbf{y}^T \end{bmatrix} \tag{A.8}$$

giving $R'(\mathbf{w})$ in equation A.2 as

$$R^{'}(\mathbf{w}) = -2\mathbf{x}E\left\{\left[r_1\mathbf{T}_1\mathbf{y}^T \quad \cdots \quad r_N\mathbf{T}_N\mathbf{y}^T\right]\right\}$$

which can be further simplified to

$$R'(\mathbf{w}) = -2\mathbf{x}\left[\ \mathbf{T}_1\mathbf{y}^T \quad \cdots \quad \mathbf{T}_N\mathbf{y}^T\ \right]\mathbf{r}^T/N \tag{A.9}$$

since the expectation is a temporal mean the element-wise multiplication with r_k can be replaced by multiplication with the vector $\mathbf{r}^T$ which also simplifies the computation.

Next, to simplify the calculation of the second derivative of $R(\mathbf{w})$, we perform column-wise matrix multiplication in equation A.9, omitting the scalar multiplication and division, to obtain

$$R'(\mathbf{w}) = -\mathbf{x}\left[\ \mathbf{T}_1 r_1\mathbf{y}^T + \quad \cdots \quad + \mathbf{T}_N r_N\mathbf{y}^T\ \right] \tag{A.10}$$

$$= -\mathbf{x}\sum_{k=1}^{N}\mathbf{T}_k r_k\mathbf{y}^T \tag{A.11}$$

And since differentiation and summation are interchangeable based on the sum rule, $R''(\mathbf{w})$ can be obtained by

$$R''(\mathbf{w}) = -\mathbf{x}\sum_{k=1}^{N}\frac{\partial(\mathbf{T}_k r_k\mathbf{y}^T)}{\partial\mathbf{w}} \tag{A.12}$$

$$= -\mathbf{x}\sum_{k=1}^{N}\frac{\partial(\mathbf{T}_k r_k\mathbf{y}^T)}{\partial\mathbf{y}}\frac{\partial\mathbf{y}}{\partial\mathbf{w}} \tag{A.13}$$

$$= -\mathbf{x}\left(\sum_{k=1}^{\mathbf{N}}\frac{\partial(\mathbf{T_k r_k y^T})}{\partial\mathbf{y}}\right)\mathbf{x}^T \tag{A.14}$$

The derivative of $\mathbf{T}_k r_k\mathbf{y}^T$ w.r.t $\mathbf{y}$ is then obtained by the product rule of differentiation.

$$\frac{\partial(\mathbf{T}_k r_k\mathbf{y}^T)}{\partial\mathbf{y}} = \mathbf{T}_k\frac{\partial r_k}{\partial\mathbf{y}}\mathbf{y}^T + \mathbf{T}_k r_k \tag{A.15}$$

$$= \mathbf{T}_k\left(\frac{\partial r_k}{\partial\mathbf{y}}\mathbf{y}^T + r_k\right) \tag{A.16}$$

which is of size $N \times N$. Consequently, the size of $R''(\mathbf{w})$ turns out to be 3×3 since the sum of $\frac{\partial(\mathbf{T}_k r_k\mathbf{y}^T)}{\partial\mathbf{y}}$ over N samples is also of size $N \times N$. These formulas of $R'(\mathbf{w})$ and $R''(\mathbf{w})$ are then used in the implementation of the multi-objective optimization algorithm.