

# Periodic Variance Maximization using Generalized Eigenvalue Decomposition applied to Remote Photoplethysmography estimation

Richard Macwan, Serge Bobbia, Yannick Benezeth, Julien Dubois, Alamin Mansouri  
LE2I EA7508, Arts et Métiers  
Univ. Bourgogne Franche-Comté

{richard.macwan, serge.bobbia, yannick.benezeth, julien.dubois, alamin.mansouri}@u-bourgogne.fr

## Abstract

*A generic periodic variance maximization algorithm to extract periodic or quasi-periodic signals of unknown periods embedded into multi-channel temporal signal recordings is described in this paper. The algorithm combines the notion of maximizing a periodicity metric combined with the global optimization scheme to estimate the source periodic signal of an unknown period. The periodicity maximization is performed using Generalized Eigenvalue Decomposition (GEVD) and the global optimization is performed using tabu search. A case study of remote photoplethysmography signal estimation has been utilized to assess the performance of the method using videos from public databases UBFC-RPPG [1] and MMSE-HR [31]. The results confirm the improved performance over existing state of the art methods and the feasibility of the use of the method in a live scenario owing to its small execution time.*

## 1. Introduction

In physiological signal measurement, quasi-periodicity is a common property shared by many types of signals such as electrocardiographic, electromyographic, electroencephalographic and photoplethysmography signals. Extraction of such underlying source signals from multichannel physiological signal mixtures has been generally performed using Independent Component Analysis (ICA) in the context of Blind Source Separation (BSS) in problems such as extraction of electrocardiogram (ECG) signals, separation of fetal ECG, heart rate estimation using remote photoplethysmography (rPPG) and speech analysis. The independent components mixed into the multichannel sensor data are typically separated by maximizing independence, using metrics such as non-gaussianity, kurtosis, or mutual information [8].

However, BSS methods fail to profit from this quasi-periodic information which is a very common property among many physiological signals. Furthermore, the periodicity constraint can also mitigate against small motion and illumination disturbances.

This periodicity criterion can be incorporated as a priori information to successfully extract the quasiperiodic physiological signal sources. For instance, in remote photoplethysmography measurement, this exploitation of quasiperiodicity can benefit in scenarios such as remote, and possibly long-term, monitoring of geriatric patients and infants, patients with severe physical trauma, computer users and drivers. This criterion is also exploited in extracting fetal ECG signals from maternal ECG signal recordings[23] and can be adapted to extraction of other afore-mentioned physiological signals. It is worth noting that the exploitation of periodicity is not enough in scenarios where the perturbing signals are themselves periodic, e.g. heart rate measurement in fitness scenarios. In such a case, additional constraints would be needed to extract the desired physiological signals.

In this paper, the iterative subspace decomposition procedure of [23] is enhanced to extract the underlying quasi-periodic signal of an unknown period embedded into the signal recordings. This new algorithm, aptly coined as Periodic Variance Maximization (PVM), is applied to remote photoplethysmography to extract the cardiac signal embedded in the RGB temporal traces. The PVM algorithm aims to find the unknown period of the desired signal, by combining two approaches. First, the iterative subspace decomposition procedure, that estimates a periodicity maximizing basis for a given frequency, and second, a global optimization algorithm of tabu search to find the frequency with the highest global periodicity over the search space. The proposed method can be used to extract any desired quasi-periodic signal of an unknown period from a mixture of signals, for any type of physiological signal mea-

surement scenarios where periodic motion is not involved.

We validate our methodology with the application to remote photoplethysmography (rPPG) and analyse its performance against two public databases, the UBFC-RPPG [1] database and the MMSE-HR [31] database. The rest of the paper is organized as follows. In section 2 we present an overview of periodicity maximization algorithms and state of the art rPPG methods. The algorithm is described in section 3 followed by performance analysis in section 4.

## 2. Previous Work

Sameni *et al.* have exploited the periodic nature of ECG signals to separate multichannel fetal and maternal ECG recordings using periodic component analysis. They achieve this separation by maximizing a measure of periodicity [22]. They further expanded on their work by proposing a generalized deflation framework to separate the target signal based on not only periodicity, but any other quantifiable properties, such as SNR, stationarity, and spectral contrast, from noisy multichannel recordings [23]. However, in their work, the period of the desired signal to be separated is fixed and known, calculated from the maternal ECG.

Tsalaile *et al.* provide an improvement to this method, by allowing the sequential blind source extraction of quasi-periodic signals having *time-varying* periods, by diagonalizing time-lagged autocorrelation matrices at time-varying lags [21]. Despite the ability of their method to handle periods varying over time, these periods are still known, and are a part of the a priori information. On the contrary, in this paper, we propose to extract the quasi-periodic signal of an *unknown* period from the signal mixtures.

In the field of speech analysis, Saul *et al.* have used an eigenvalue method to analyze and enhance weak periodic signals which is insensitive to phase thanks to the use of Hilbert transforms. They perform the periodicity measurements using efficient sinusoidal fits to extract the fundamental frequencies, albeit with extensive auditory preprocessing, habitual to the domain of speech analysis which are computationally intensive and may hurt its feasibility for real-time applications.

In this paper, we propose a method to extract periodic or quasi-periodic signals of unknown period embedded into multi-channel temporal signal recordings. The proposed technique is applied to remote photoplethysmography in order to extract the cardiac signal embedded in the RGB temporal traces. The quasi-periodic cardiac signal is mixed in the light reflected by the tissue with other signals such as changes in incident light or motion induced shadow casting varia-

tions. This mixed signal is then captured by the camera. The proposed PVM algorithm aims to find the quasi-periodic cardiac signal from the mixture. Moreover, contrary to the known problem of fetal ECG extraction [23], the period is unknown and consequently it is an interesting problem for us.

Recently, research on remote photoplethysmography measurements has been on the rise where different classes of methods have been proposed to extract the quasi-periodic cardiac signal embedded in RGB temporal traces built from sequential video frames [25]. RPPG measurement algorithms can be classed under various categories. First, blind or semi-blind source separation/extraction methods based on Independent Component Analysis (ICA) [19], [20] and [17], [18], [30] and constrained ICA [13], [15], [14]. Second, methods exploiting unique characteristics of skin and its chrominance properties [3], [4], [29]. And last but not the least, smart ROI based methods where the spatial characteristics of images are utilized to construct the rPPG signal [1], [10], [5].

To the best of our knowledge, the combination of principal component analysis with periodicity maximization has not been combined for extracting a quasi-periodic signal of unknown period. The Periodicity Variance Maximization algorithm is presented in the next section.

## 3. Method

The proposed method aims to extract the quasi-periodic signal embedded in the recorded signals which is marked by high periodicity, or in other words *high periodic variance*, corresponding to its fundamental period  $\tau^*$ . For a centered temporal signal,  $\mathbf{y} = \mathbf{y}(t)$ , we define high periodic variance at period  $\tau$  as the property that corresponds to a high variance  $\sum \mathbf{y}\mathbf{y}^T$ , as well as a high *lagged* variance  $\sum \mathbf{y}\mathbf{y}_\tau$ , where  $\mathbf{y}_\tau = \mathbf{y}(t + \tau)$ . Typically, a periodic or a quasi-periodic signal exhibits high periodic variance at  $\tau$  corresponding to its fundamental frequency  $f = 1/\tau$ . Typical biomedical signals such as electrocardiography, electromyography, neural and photoplethysmography signals are quintessentially quasi-periodic and at times periodic.

The method takes as input temporal traces  $\mathbf{x} \in \mathbb{R}^M$  of length  $N$  and  $M$  channels, centered by subtracting the channel-wise temporal means, to extract the most periodic signal of an unknown period, using  $\mathbf{y} = \mathbf{w}^T \mathbf{x}$ . Here,  $\mathbf{w} \in \mathbb{R}^M$  is the optimum weighting vector which gives the desired signal  $\mathbf{y}$  with the highest periodic variance, obviously corresponding to the fundamental period of the desired signal. To this end, we use Generalized EigenValue Decomposition (GEVD) on the pair of the lagged covariance and covariance matrices of  $\mathbf{x}$

defined as

$$P_{\mathbf{x}} = \mathbf{x}\mathbf{x}_{\tau}^T, \quad C_{\mathbf{x}} = \mathbf{x}\mathbf{x}^T \quad (1)$$

where  $\mathbf{x}_{\tau}$  are the temporal traces centered and lagged by  $\tau$  seconds. Estimating the unknown fundamental period is not an uncommon problem in a broad range of applications. An optimization scheme can be used to estimate the unknown fundamental period  $\tau^*$  of the desired signal by maximizing a periodicity metric similar to the one defined in [23] given by

$$\mathcal{P}(\tau, \mathbf{w}) \doteq \frac{E_t\{\mathbf{y}\mathbf{y}_{\tau}\}}{E_t\{\mathbf{y}^2\}} = \frac{\mathbf{w}^T P_{\mathbf{x}} \mathbf{w}}{\mathbf{w}^T C_{\mathbf{x}} \mathbf{w}} \quad (2)$$

Intuitively, the periodicity metric  $\mathcal{P}$  represents the extent of periodic information in the signals since it is the ratio of the lagged covariance matrix to the covariance matrix in a different basis. If the constituent signals were completely periodic,  $P_{\mathbf{x}}$  and  $C_{\mathbf{x}}$  would be equivalent giving  $\mathcal{P} = 1$ . Also, to ensure that the generalized eigenvalues are real,  $P_{\mathbf{x}}$  needs to be symmetrized using  $P_{\mathbf{x}} = (P_{\mathbf{x}} + P_{\mathbf{x}}^T)/2$ . This symmetrized  $P_{\mathbf{x}}$  matrix combines the two way variances between the channels, for instance, variances between the channels  $(R, G)$  and  $(G, R)$ . This symmetrization represents the overall lagged covariance among the channels to be maximized, to make it as similar as possible to  $C_{\mathbf{x}}$ .

$$W^T P_{\mathbf{x}} W = D, \quad W^T C_{\mathbf{x}} W = I \quad (3)$$

where the diagonal matrix  $D$  contains the generalized eigenvalues corresponding to the eigenvectors  $W = [\mathbf{w}_1, \dots, \mathbf{w}_N]$ , the eigenvalues being real and sorted in the ascending order along the diagonal.

For a set of temporal signals  $\mathbf{x} \in \mathbb{R}^M$ , these generalized eigenvectors have three important properties. First, they correspond to a change of basis onto which the projection  $W^T \mathbf{x}$  of the original signals  $\mathbf{x}$  are entirely uncorrelated. This is evident from the fact that the covariance matrix in this new space given by  $(W^T \mathbf{x})(\mathbf{x}^T W)$  from equation 3 is diagonal, ensuring maximum variance of the data. Second, they also diagonalize the lagged covariance matrix  $P_{\mathbf{x}}$ , making the projections of  $\mathbf{x}$  and  $\mathbf{x}_{\tau}$  on this new basis entirely uncorrelated as well, which is possible only when the lagged signal  $\mathbf{x}_{\tau}$  is extremely similar to the original signal  $\mathbf{x}$  at that given lag. In fact, the magnitude of the generalized eigenvalues actually reflects the amount of similarity between the original and the lagged signals. And finally, the first eigenvector  $W_1$  corresponding to the largest generalized eigenvalue that maximizes the ratio defined in equation 2, also referred to as the Rayleigh Quotient [24]. Consequently, the projection of the original signals  $\mathbf{x}$  to the new basis represented by  $W_1$  captures the maximum periodic content. Owing to these properties of GEVD, it can be used in an

iterative algorithm to extract the components exhibiting high periodicity and simultaneously containing the maximum information.

### 3.1. Iterative Periodic Variance Maximization

We can now proceed with the formulation of an optimization scheme that iteratively maximizes the periodicity of the projected signals over the range of frequencies corresponding to the specific application. For instance, in the context of rPPG measurements, this range corresponds to the human heart rate. Evidently, the optimizer needs to estimate the optimum value of the pair  $(\tau, \mathbf{w})$  which maximizes the periodicity metric  $\mathcal{P}$  equation 2.

The most periodic signal can then be obtained by estimating the weighting matrix  $\mathbf{w}^*$  that maximizes  $\mathcal{P}$  over the frequency search space  $\mathcal{F} \in [f_{min}, f_{max}]$ . Of course, owing to the time domain formulation of our problem, the optimization needs to be performed over the period  $\tau$ , with a step of  $1/F_s$  seconds, over the temporal search space  $\tau \in [\tau_{min}, \tau_{max}]$  corresponding to the frequency range  $[f_{max}, f_{min}]$ . Here  $F_s$  is the sampling rate of the recorded signals. An implicit advantage to this formulation in the time domain is the restricted search space corresponding to discrete integer time lags of  $[\tau_{min}F_s, \tau_{max}F_s]$  which facilitates the use of this method in a live scenario where the sampling rate is not exceedingly high.

For instance, in the application of rPPG measurement, the frame rate is typically between 20 to 30 frames/second for conventional cameras. Coupled with the limited range of human heart rates  $\mathcal{F} \in [40, 200] \text{ bpm}$ , or  $[.3, 1.5]$  second, this makes the live implementation of this method undoubtedly feasible. Indeed, the search space depends on  $F_s$ , facilitating finer searches at higher frame rates. This two-variable optimization of the pair  $(\tau, \mathbf{w})$  can be simplified into two following steps.

#### 3.1.1 Estimate the basis that maximizes periodicity

The first core step comprises of estimating the weighting matrix  $\mathbf{w}^*$ , which is nothing but a change of basis, that maximizes the periodicity metric in equation 2 for a given time period  $\tau$ . This sub-method was adapted from [23] and is listed in algorithm 1. It starts with a GEVD step that orders the components by their periodicity, based on the magnitude of the generalized eigenvalues. The decomposed signals are then projected onto the new subspace represented by  $W$ . Next, denoising is performed on the first  $R$  signals by using a wavelet denoiser, using the parameters similar to [22].

---

**Algorithm 1** Optimum Basis Estimation

---

```
1: procedure OBE( $\mathbf{x}, \tau$ )  $\triangleright$  Estimate the best  $\mathbf{w}$  for  
    $\tau$   
2:   for  $k = 1 : K$  do  
3:     Calculate  $P_{\mathbf{x}}$  and  $C_{\mathbf{x}}$   $\triangleright$  Using equation 1  
4:      $W \leftarrow \text{GEVD}(P_{\mathbf{x}}, C_{\mathbf{x}})$   
5:      $\mathbf{s} \leftarrow W^T \mathbf{x}$   $\triangleright$  Project onto new basis  
6:      $[\mathbf{s}_1 \dots \mathbf{s}_R] \leftarrow G([\mathbf{s}_1 \dots \mathbf{s}_R])$   $\triangleright$  Smooth R  
       channels  
7:      $\mathbf{x} \leftarrow W^{-1} \mathbf{s}$   $\triangleright$  Back project to original space  
8:   end for  
9:    $W = [\mathbf{w}_1 \dots \mathbf{w}_M]$   $\triangleright$  M channels  
10:   $\mathbf{y} = \mathbf{w}_1^T \mathbf{x}$   $\triangleright$  Most periodic component  
11:   $\mathbf{w}^* = \mathbf{w}_1$   
12:  return  $\mathbf{w}^*$   
13: end procedure
```

---

And lastly, the mixture of the projected and denoised signals are back-projected onto the original space. The above steps are repeated for  $K$  iterations. The output of this core step is the periodicity maximizing basis  $\mathbf{w}_1$  which is the generalized eigenvector corresponding to the highest generalized eigenvalue of the pair  $(P_{\mathbf{x}}, C_{\mathbf{x}})$ . Furthermore, for  $K > 1$ , the temporal signals are filtered to enhance periodic components. The impact of different values of  $K$  and  $R$  on the eventual estimation of the most periodic signal is presented in section 4 with the application of rPPG measurement.

Finally, we can represent the output of this step concisely as a function  $\mathcal{W}(\tau)$  and consequently rewrite the periodicity metric equation 2 solely as a function of  $\tau$

$$\mathcal{P}(\tau) = \frac{\mathcal{W}(\tau)^T P_{\mathbf{x}} \mathcal{W}(\tau)}{\mathcal{W}(\tau)^T C_{\mathbf{x}} \mathcal{W}(\tau)} \quad (4)$$

This objective function can now be maximized by using an appropriate optimization scheme the details of which are presented in the next subsection.

### 3.1.2 Optimize the periodicity metric $\mathcal{P}(\tau)$

To select an appropriate optimization scheme, the objective function  $\mathcal{P}(\tau)$  in equation 4 warrants some examination. First, it is evident that for temporal signals of a specific sampling rate, it is a continuous function over the corresponding temporal search space. The complexity, however, lies in the calculation of the derivative  $\mathcal{W}'(\tau)$  owing to the presence of the GEVD step. In the general sense, differentiation of an eigendecomposition represents the change in the eigenvalues with respect to change in the original data. Although, such a differentiation is mathematically possible, it is

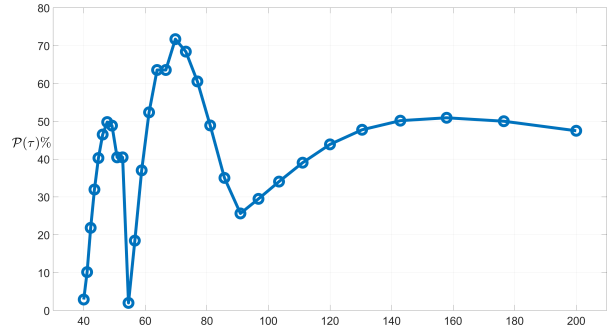


Figure 1: Typical distribution of  $\mathcal{P}(\tau)$

admittedly non-trivial [16] [27] and even more complicated with respect to the period  $\tau$ , which points towards the use of a derivative free optimization approach.

Another feature of interest was the shape of the objective function. To assess this, application to the problem of rPPG signal estimation was chosen. Using videos from UBFC-RPPG and MMSE-HR databases the objective function values of  $\mathcal{P}(\tau)$  for several videos from UBFC-RPPG and MMSE-HR databases were examined. Figure 1 shows the variation of  $\mathcal{P}$  with respect to  $\tau$ , calculated over a window of 15 seconds for a typical video, taken from the MMSE-HR database. It can be observed that the variation of  $\mathcal{P}(\tau)$  is prone to multiple local maximas. This is expected behavior owing to the product  $\mathbf{x}\mathbf{x}_\tau^T$  in the calculation of  $\mathcal{P}(\tau)$ . Specifically, a signal with a real period of  $\tau^*$  seconds will exhibit a relatively higher value for  $\mathcal{P}(\tau)|_{\tau=\tau^*}$ . Accompanied by this global maximum are local maximas exhibited at  $\tau = n\tau^*, n \in \mathbb{Z}$ , albeit lower than  $\mathcal{P}_{\tau=\tau^*}$ , and with magnitudes inversely proportional to  $n$ . This appearance of multiple local maximas, with decreasing magnitudes, is similar to the plot of the autocorrelation function. Additionally, the advantage of a limited search space, suggests the use of a global optimization scheme for best results. Algorithms that have the ability to escape local optimas serve as an appropriate solution. Consequently, the tabu search algorithm [6] was used to perform the global optimization of  $\mathcal{P}(\tau)$ .

### 3.2. Periodic Variance Maximization applied to RPPG signal estimation

We present here the applicability of the PVM algorithm implemented in the context of rPPG signal estimation using sequential video data from conventional cameras. The workflow of the procedure as depicted in Fig. 2 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 on the detected face was performed for tracking to crop the face based on facial landmarks. Skin detection as formulated by Conaire *et al.* [2] 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 [26] to remove low frequency trends in the signal. Centering was then performed so that the obtained signal  $\mathbf{y}$  in  $\mathbf{y} = \mathbf{W}\mathbf{x}$  is zero-mean. After the rPPG signal was obtained using PVM, 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 a 30 second moving window using a step size of 0.5 second for the UBFC-RPPG datasets. This 30 seconds window size was chosen as a trade-off between speed and availability of enough data for analysis. However, for the MMSE-HR database a 15 second window had to be chosen because of the relatively small length of several constituent videos. The window-wise heart rate estimations were then smoothed using a Kalman filter. This 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.

## 4. Results and Discussion

The PVM algorithm was validated using the UBFC-RPPG [1] database containing two datasets, labeled SIMPLE and REALISTIC, comprised of 9 (about 21k frames) and 46 (about 94k frames) videos respectively, and the MMSE-HR [31] database comprising of 97 (about 105k frames) videos. The UBFC-RPPG database was acquired using a Logitech C920 web camera placed at a distance of about 1m from the subject. The videos were recorded with a frame 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 comprising of the PPG waveform as well as the PPG heart rates. The experimental setup with sample images from both the databases is depicted in figure 3. An overview of the framework along with a symbolic graph of the core PVM algorithm is also presented in figure 2.

The MMSE-HR database is inclined towards research on emotion elicitation and recognition. As a result, the videos in this database comprise of a large number of facial expressions and movements, which

Table 1: Performance metrics

	UBFC-RPPG						MMSE-HR	
	SIMPLE			REALISTIC				
	MAE	SNR	r	MAE	SNR	r	MAE	r
PVM	0.93	2.01	<b>0.99</b>	4.47	<b>-0.22</b>	0.82	<b>4.38</b>	0.82
ICA	0.67	2.70	0.98	6.02	-1.11	0.79	5.84	0.67
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	<b>3.04</b>	0.99	<b>3.70</b>	-0.32	0.87	5.59	<b>0.83</b>
POS	<b>0.67</b>	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

eventually aids in assessing our method thoroughly. The MMSE-HR database does not provide ground truth PPG waveforms, only the ground-truth heart rates obtained using the BIOPAC 150 data acquisition system at 1 kHz, calculated using contact based ECG electrodes.

Table 1 shows the performance comparisons between PVM and other state of the art methods, viz., ICA [8], PCA [11], Green [28], CHROM [3], Plane Orthogonal to Skin (POS) [29], and G-R [7]. The ICA implementation used for our analysis has been adapted from FastICA [9]. Furthermore, the analysis of all the methods was performed using exactly the same pre and post processing steps like normalization, filtering and smoothing. The exact metrics of ICA from related state of the art methods such as [19], [20] and [17] could not be used because they all use their own private databases which were inaccessible to the public. However, the core algorithm of ICA remains the same making the metrics in Table 1 applicable. Furthermore, comparison with smart ROI selection methods such as [1], [12], [10] and [30] was not deemed relevant in order to limit the comparison amongst source separation/extraction methods, and the fact that our method can easily be incorporated into any ROI selection framework.

The metrics used in our analysis are mean absolute error (MAE) in beats per minute (bpm), 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 MAE was calculated as the window-wise mean of  $|HR_{rPPG} - HR_{PPG}|$ , averaged per video. 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 worth mentioning that the MMSE-HR database does not provide the ground truth waveforms,

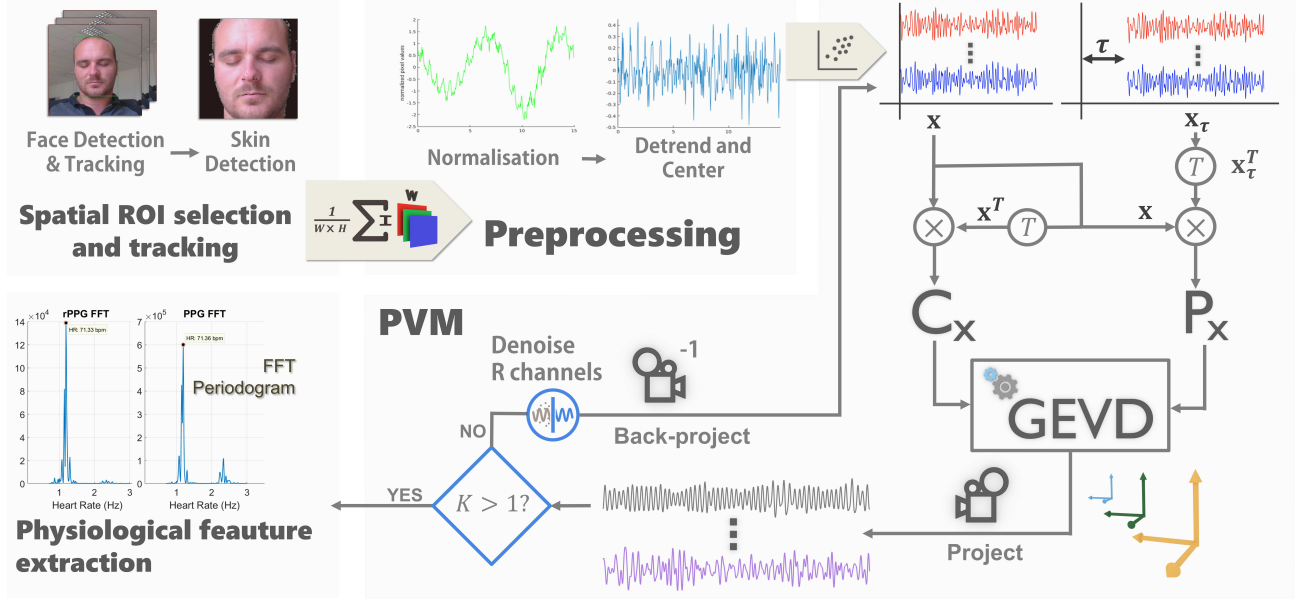


Figure 2: System framework and Periodic variance maximization overview

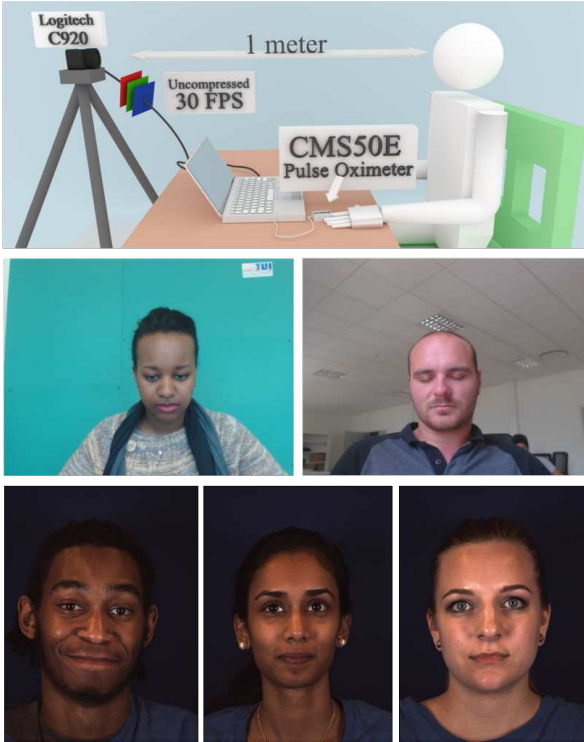


Figure 3: Experimental Setup (top) and sample images from the UBFC-RPPG REALISTIC and SIMPLE (middle) and MMSE-HR (bottom) databases

thereby obliging the use of the main pulsatile component of the RPPG instead of the PPG for the SNR

calculation. In this case, the SNR just represented the strength of the main pulsatile component, which, although being a useful metric by itself, is not entirely appropriate for the comparison with the SNR for other databases which were calculated in a different manner. As a result, the SNR values for the MMSE-HR database are not really relevant and are omitted. The MAE values, however, are relevant since it is calculated as the difference between rPPG and PPG heart rates, which are provided as ground truth for the MMSE-HR database.

Figure 4 presents the overall correlation analysis between the PVM method and the CHROM method which in our experience is one of the most consistent methods for rPPG estimation. The comparisons were performed using window-wise calculations between HRs from PPG against RPPG obtained from all the videos in each dataset for the skin-segmented pixel data for one particular run. It is worth mentioning that the MAE values in figure 4 and table 1 differ from those in figure 5, which are averaged over 20 executions, but are obviously in range. Moreover, differences in range of  $10^{-2}$  bpm are inconsequential.

With the UBFC-RPPG database, the analysis of the SIMPLE dataset, as the name suggests, was relatively simple and efficient, with an average MAE of less than 1 bpm. This can be attributed to the fact that the subjects were generally relaxing, mostly with their eyes closed, which resulted in minimal motion artifacts. On the other hand, the REALISTIC dataset was slightly more challenging since the subjects were actually work-

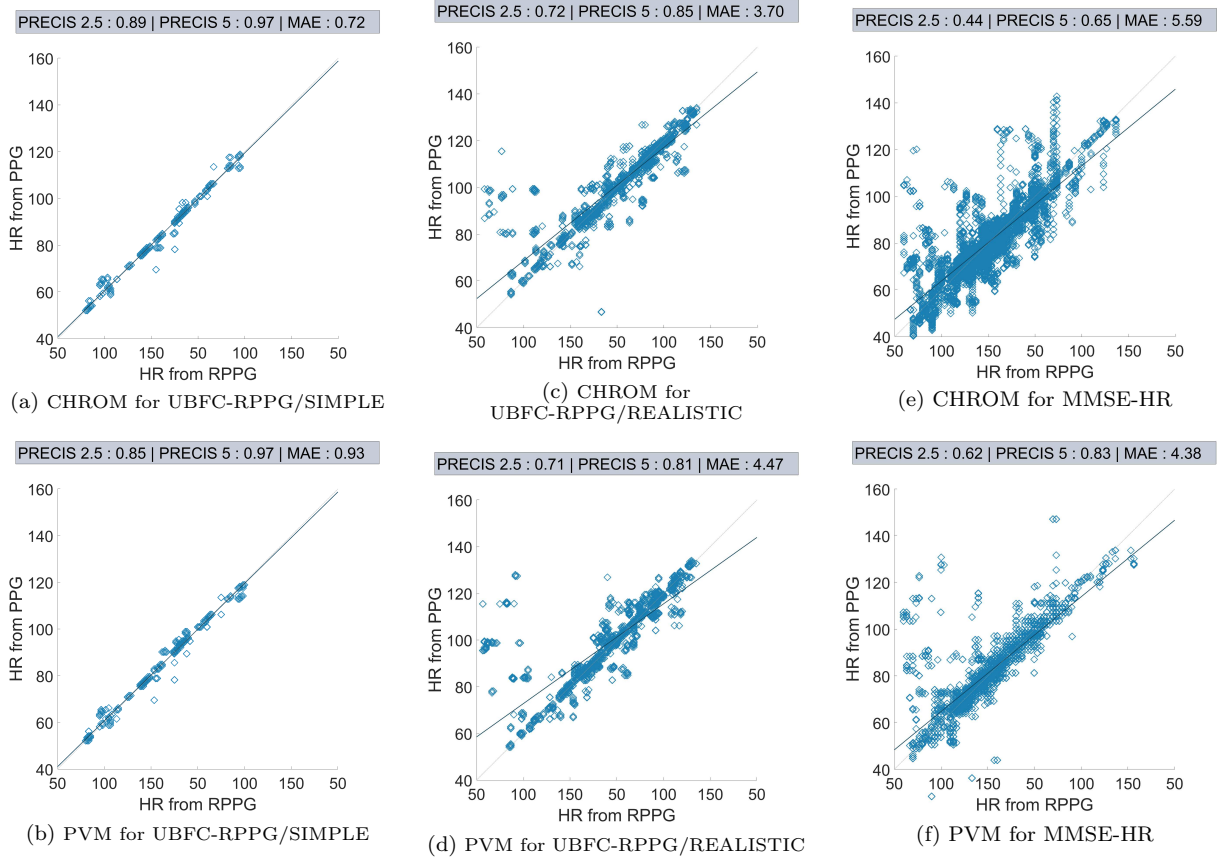


Figure 4: Correlation comparison plots for CHROM vs PVM. The metrics PRECIS 2.5 and PRECIS 5 represent the percentage of windows where  $\delta = |HR_{rPPG} - HR_{PPG}| < 2.5$  and 5 bpm respectively.

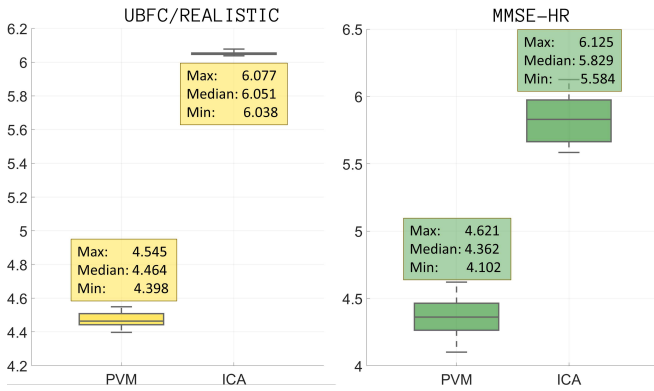


Figure 5: Box plot of MAE from the PVM and ICA algorithms over 20 observations for the two databases

ing on the computer and were only requested to keep their hand still for the PPG sensor. The performance of the PVM method is similar to that of the CHROM method, albeit with a slightly higher average MAE.

One possible hypothesis for this inconsistency in the MAEs can be as follows. Since the REALISTIC dataset was recorded under ambient light, the effect of the fluorescent light source in certain videos was more pronounced, most probably due to lesser ambient light. The 50 Hz flicker of fluorescent lights is a well known problem in video recordings where each video frame is exposed at different light pulses. The discrepancy between our 30 Hz frame rate and the 50 Hz flicker might result in a perturbed periodic signal which overshadows the cardiac signal in certain cases and is selected by PVM. The CHROM method is able to overcome this problem because of the projection of RGB signals onto a different subspace where only the reflections specific to the skin properties are enhanced. On the other hand, all the videos in the MMSE-HR database were recorded with a more sophisticated lighting setup, thus avoiding such an issue and resulting in a better performance for the PVM method. However, the estimation and elimination of this background flicker is in itself not a difficult issue and can be solved with an appro-

appropriate filtering scheme, which incidentally is one of the issues we aim to address in our future work.

The MMSE-HR database was challenging owing to it being an emotion elicitation database. There were many instances where the subjects laughed out loud, exhibited considerable movements, while regularly manifesting various facial expressions. This resulted in the usual problems arising from movement of the subjects and the face ROIs. However, the PVM method outperformed the CHROM method by a larger margin for this database. This highlights the relative robustness of the PVM method against motion variances as well. The better performance of the PVM method can also be attributed to the sophisticated lighting setup of the MMSE-HR database. This is also evident from the correlation plot having a narrower spread and the fitting line being slightly closer to the 45° line PVM as compared to CHROM.

**Effect of the parameters  $K$  and  $R$  on the overall result** The value of  $K$  depends on the number of dimensions of the desired subspace, which is 2 in our case. However, for our application, a value of  $K = 1$  gives optimum results, which makes this algorithm suitable for use in live scenarios. Similarly, the performance of the method with respect to  $R \in [1 : M]$  where  $M = 3$  is the number of channels, was also performed. It was equally observed the algorithm was able to obtain the most periodic component even by denoising just one channel thanks to the efficiency of the GEVD step to successfully sort components in order of decreasing periodicity.

Finally, to assess the consistency of the overall algorithm, and because of the use of the tabu search global optimization, it was worth assessing the performance of the method over multiple runs. Figure 5 shows the box plot comparing the MAEs with the PVM and ICA algorithms for the two challenging datasets: UBFC-RPPG/REALISTIC dataset with subjects working on a computer under ambient light, and the MMSE-HR dataset with subjects exhibiting facial expressions under indoor lighting. These tests on the datasets were performed 20 times. The PVM method has a consistent performance, giving MAEs in the range [4.39, 4.55] bpm and [4.10, 4.62] bpm for the UBFC-RPPG/REALISTIC and MMSE-HR datasets respectively. The consistency analysis for the other baseline methods was not required because only the PVM and ICA methods are optimization based. The box plots for the UBFC-RPPG/SIMPLE are not shown since they exhibit MAEs of less than 1 bpm for both PVM and ICA methods and their comparison was deemed inconsequential.

## 5. Conclusions and Future Work

A Periodic Variance Maximization algorithm was presented which extracts the most periodic signal of an unknown period from a mixture of temporal recordings. The PVM algorithm is fairly generic and can be used in any problem domain where periodicity maximization is called for. It can also be combined with smart-ROI selection methods to exploit the spatial and/or facial features to enhance the resulting signal. Its vulnerability, however, lies with scenarios having periodic motion, e.g. fitness based rPPG signal estimation. Motion compensation schemes, which itself is another subject of research, can be beneficial in such cases. Future work comprises of addressing the issue of perturbation arising from background light flicker and extending the algorithm to the multi-linear case using tensorial analysis to estimate rPPG signal strength across the skin region.

## References

- [1] S. Bobbia, Y. Benezeth, and J. Dubois. Remote Photoplethysmography Based on Implicit Living Skin Tissue Segmentation. *PRL*, 2016. 1, 2, 5
- [2] C. Ó. Conaire, N. E. O'Connor, and A. F. Smeaton. Detector adaptation by maximising agreement between independent data sources. *CVPR*, 2007. 5
- [3] G. De Haan and V. Jeanne. Robust pulse rate from chrominance-based rPPG. *IEEE Transactions on Biomedical Engineering*, 60(10):2878–2886, 2013. 2, 5
- [4] 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. 2
- [5] L. Feng, L. M. Po, X. Xu, Y. Li, C. H. Cheung, K. W. Cheung, and F. Yuan. Dynamic ROI based on K-means for remote photoplethysmography. *ICASSP*, 2015-Augus:1310–1314, 2015. 2
- [6] F. Glover. Paths for Integer Programming. *Computers and Operations Research*, 13(5):533–549, 1986. 4
- [7] M. Hülbusch. Ein bildgestütztes, funktionelles Verfahren zur optoelektronischen Erfassung der Hautperfusion. *Anesthesia & Analgesia*, page 145, 2008. 5
- [8] A. Hyvärinen. Fast and robust fixed-point algorithm for independent component analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 10(3):626–634, 1999. 1, 5
- [9] A. Hyvärinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, 13(4-5):411–430, 2000. 5
- [10] M. Kumar, Ashok Veeraraghavan, and Ashutosh Sabharwal. DistancePPG: Robust non-contact vital signs monitoring using a camera. *Biomedical Optics Express*, 6(5):1565, 2015. 2, 5
- [11] M. Lewandowska, J. Ruminski, T. Kocejko, and J. Nowak. Measuring pulse rate with a webcam; A non-contact method for evaluating cardiac activity. *FedCIS*, pages 405–410, 2011. 5



- [12] X. Li, J. Chen, G. Zhao, and M. Pietikäinen. Remote heart rate measurement from face videos under realistic situations. *Proceedings of the IEEE Computer Society Conference on CVPR*, pages 4264–4271, 2014. 5
- [13] W. Lu and J. C. Rajapakse. ICA with Reference. *Neurocomputing*, 69(16-18):2244–2257, 2006. 2
- [14] R. Macwan, Y. Benezeth, and A. Mansouri. Remote photoplethysmography with constrained ICA using periodicity and chrominance constraints. *BioMedical Engineering OnLine*, 17(1):22, 2018. 2
- [15] R. Macwan, Y. Benezeth, A. Mansouri, K. Nakamura, and R. Gomez. Remote Photoplethysmography measurement using Constrained ICA. In *E-Health and Bioengineering Conference*, 2017. 2
- [16] J. Magnus. On Differentiating Eigenvalues and Eigenvectors. *Econometric Theory*, 1:179–191, 1985. 4
- [17] 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. 2, 5
- [18] D. J. McDuff, J. R. Estep, 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. 2
- [19] 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. 2, 5
- [20] 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. 2, 5
- [21] R. Sameni, , et al. Sequential blind source extraction for quasi-periodic signals with time-varying period. *IEEE Transactions on Biomedical Engineering*, 56(3):646–655, 2009. 2
- [22] R. Sameni, C. Jutten, and M. Shamsollahi. Multi-channel Electrocardiogram Decomposition Using Periodic Component Analysis. In *IEEE Transactions on Biomedical Engineering*, pages 1935–1940, 2008. 2, 3
- [23] R. Sameni, C. Jutten, and M. B. Shamsollahi. A deflation procedure for subspace decomposition. *IEEE Transactions on Signal Processing*, 58(4):2363–2374, 2010. 1, 2, 3
- [24] G. Strang. *Linear Algebra and its Applications*. New York, 3rd edition. 3
- [25] 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. 2
- [26] 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. 5
- [27] N. P. Van Der Aa, H. G. Ter Morsche, and R. R. M. Mattheij. Computation of eigenvalue and eigenvector derivatives for a general complex-valued eigensystem. *Electronic Journal of Linear Algebra*, 16:300–314, 2007. 4
- [28] W. Verkrusse, L. O. Svaasand, and J. S. Nelson. Remote plethysmographic imaging using ambient light. *Optics express*, 16(26):21434–21445, 2008. 5
- [29] W. Wang, A. Den Brinker, S. Stuijk, and G. De Haan. Algorithmic Principles of Remote-PPG. *IEEE Transactions on Biomedical Engineering*, PP(99):1–1, 2016. 2, 5
- [30] 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, 2015. 2, 5
- [31] 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. 1, 2, 5