ROBUST MOTION ARTIFACT REDUCTION OF PHOTOPLETHYSMOGRAPHIC SIGNALS WITH TRAJECTORY SPACE CIRCULAR MOTION MODEL

Tzu-Ming Harry Hsu[†], Wei-Yu Chen[†], Kuan-Lin Chen[†], Mong-Chi Ko[†], You-Cheng Liu^{*}, An-Yeu (Andy) Wu[†]

†Dept. of Electrical Engineering *Graduate Institute of Electronics Engineering National Taiwan University, Taipei 106, Taiwan

ABSTRACT

Heart rate (HR) retrieval using photoplethysmographic (PPG) signals is recently a trend due to the portability and relatively low price of wearable biosensors. However, motion artifacts (MA) can severely corrupt the PPG signals, especially during intensive physical activities, and result in errors in HR estimation. To address the problem of MA's interference in PPG signals, we propose a framework to not only eliminate the periodical MA, but also remove irregular MA, by the cascaded process of decomposing, MA recognition, projecting with our proposed Trajectory Space Circular Motion (TSCM) model, and Bayesian filtering. Evaluation of our framework is done by Bland-Altman agreement and correlation analysis against the reference HR readings from simultaneously recorded electrocardiogram (ECG). The outcome of our framework is quantitatively verified, and is close to the ground truth and reliable for applications in practical MA removal for PPG signals obtained with wearable biosensors. The mean absolute error estimated from the dataset of 12 subjects during intensive activities is found to be 2.21 BPM while the Pearson correlation coefficient between the estimated HR and ground-truth HR is r = 0.995.

Index Terms— Pulse Oximeter, Photoplethysmography (PPG), Wearable Biosensor, Motion Artifact (MA), Multivariate Singular Spectrum Analysis (MSSA), Kalman Filter, Particle Filter

1. INTRODUCTION

PPG is a non-invasive optical measurement technique that can be used to detect blood volume changes in the microvascular bed of tissue. The most recognized waveform feature on the PPG signal is the peripheral pulse, which is synchronized to each heartbeat. Nowadays, we can conveniently monitor HR in modern wearable devices with PPG sensors. However, the PPG signal is very susceptible to the MA during physical exercise.

To estimate HR in beats per minute (BPM) precisely, the MA reduction technique plays an important role in corrupted PPG reconstruction. The method of measuring HR from PPG

signals is also an important issue because it may affect the accuracy severely.

The existing methods [1, 2, 3] have limited application

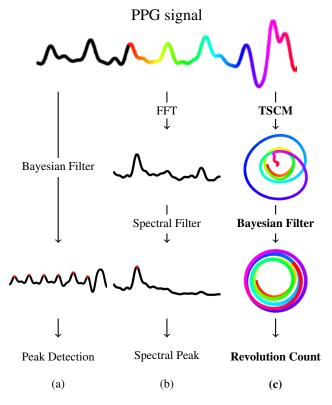


Fig. 1: A comparison between TSCM and classical HR retrieval methods (best viewed in color). (a) is a direct time domain method which implements Bayesian filter with time domain state space models. (b) is a frequency domain method, which is not feasible when MA is not spectral-separable, and besides the filter requires parameter tuning and the parameters can vary in a wide range when the types of activity changes. (c) is our proposed TSCM model which projects the PPG signal into the *trajectory space* where the dynamics can be easily depicted as a circular motion around the origin. (Corresponding points are labeled in the same color, and for clear visualization, some parts of the trajectory are not shown.)

in restoring corrupted PPG signals for the reason that the frequency band of accelerometer (ACC) signals may overlap with PPG signals'. The performance of recovered PPG signal is limited when MA's spectrum is close to that of HR. To solve this problem, we developed a signal decomposition technique called multivariate singular spectrum analysis (MSSA), which is used to decompose multi-channel signals. After MSSA, we can remove the MA component in PPG signals as much as possible even they overlap in frequency domain.

On the other hand, our method of measuring HR from PPG signals is different from the classical methods since the disadvantage of peak detection in time or frequency domain is too rough to estimate HR. For example, if the spectral peak corresponding to MA is larger than the one corresponding to HR, we will get wrong HR value when detecting peak in frequency domain. So we develop an algorithm called *Trajectory Space Circular Motion* model, which is a stable framework, mapping a sequence of PPG signal to the trajectory space just like the circular motion. After a revolution, we consider it as one heartbeat. Figure 1 is a diagram that illustrates the comparison between TSCM and classical HR retrieval methods. In addition to estimating HR precisely, it can also revise minor error due to irregular MA noise.

Table 1 compares the major difference among other works and ours. Most works mainly focus on Regular [1] or Irregular [2, 3] MA removal only, while ours handle both and calculate BPM on the brand-new domain directly.

Table 1: A comparison among existing methods and ours.

Method	Regular MA removal	Irregular MA removal	HR retrieval method
[1]	Adaptive noise cancellation	N/A	Time domain (peak detection)
[2]	N/A	Wavelet transform	Frequency domain (spectral peak)
Ours	MSSA	TSCM	Trajectory space (revolution count)

2. PROPOSED FRAMEWORK

To address the problem of blending MA into PPG signals, we have devised a framework that takes the ACC readings into account. The database investigated [4] includes a two-channel PPG signal and a three-channel ACC signal, and however our framework can easily generalize to multi-channel PPG signal. This work is based on the signal segment of a time window of t_w seconds retrieved from a continuously measured signal set. There is as well a t_i seconds time increment between consecutive windows, creating some overlap between them and helping stabilize the HR readings. In our experiments, we will set $t_w = 8$ and $t_i = 2$.

As illustrated in Fig. 2, our framework consists of several parts, namely preprocessing, signal decomposition, TSCM

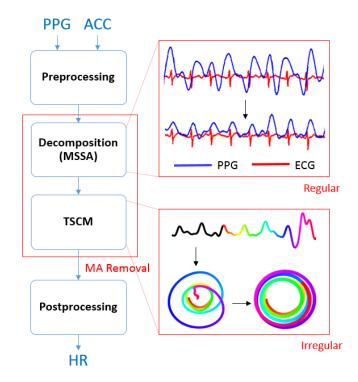


Fig. 2: A flow chart of our framework with highlights in the MSSA and TSCM block.

model, and post-filtering. In the signal decomposition section, we will extract some main components from the joint signal of PPG and ACC, and judge whether they correspond to the frequency of MA or HR. In the TSCM model, we try to project the signal to a trajectory space, in which we can obtain stable features and reconstruct the failure introduced by the remaining irregular MA. Finally, the outcome from the last steps will be regarded as observations of the true HR during the post-filter phase, and a filter will be then applied to the observations to obtain the true HR readings.

2.1. Notation

Hereby we consider the input signals, which are of sampling frequency f_s , are properly segmented into length t_i . The raw data will contain:

1. A k_P -channel PPG signal ensemble

$$\mathbf{S}_P = [\mathbf{s}_{P1}, \mathbf{s}_{P2}, \dots, \mathbf{s}_{Pk_P}]. \tag{1}$$

2. A k_A -channel ACC signal ensemble ($k_A = 3$)

$$\mathbf{S}_A = [\mathbf{s}_{A1}, \mathbf{s}_{A2}, \dots, \mathbf{s}_{Ak_A}]. \tag{2}$$

Some other auxiliary signals will as well be useful:

1. The ACC magnitude s_{A0}

$$s_{A0}^{2}(t) = \sum_{m=1}^{k_{A}} s_{Am}^{2}(t), \quad 1 \le t \le L, \quad (3)$$

where $L = f_s t_i$ is the length of every channel in the ACC signal.

2. The joint signal ensemble of PPG and ACC $S_J \equiv [S_P, S_A]$.

2.2. Preprocessing

Preceding our main algorithm, we must assure the signal is well-preconditioned and free from noise, trend, inter-channel delay, and scaling difference. Hence we proceed the following preprocessing procedures.

2.2.1. De-noising and De-trending

A Savitzky-Golay FIR smoothing filter [5] of narrower window (approximately 0.25 second) is used to eliminate the noise while a window of a longer duration (approximately 3 seconds) is applied to eliminate the trend.

2.2.2. Alignment

Since the sensors may be installed on different sites on the subject, we hereby consider a time difference between difference signal channels. The time lag can be 12 sample points per centimeter on a sampling rate of 125 Hz, and these asynchronously recorded signals can result in difficulties when attempting to decompose the signal set.

To align the different channels of signals, we consider the unbiased cross-correlation function between two time series \mathbf{s}_0 and \mathbf{s} , which are both of length L, and \mathbf{s}_0 is the alignment reference:

$$(\mathbf{s}_{0} \star \mathbf{s})(t) \equiv \mathbf{E} \left[s_{0}(\tau) s(\tau + t) \right]$$

$$= \frac{1}{L - |t|} \sum_{\tau} s_{0}(\tau) s(\tau + t).$$
(4)

The correlation $(\mathbf{s}_0 \star \mathbf{s})$ (t) will have a maximum value at $t = t_d$, where t_d is close to zero and is the point in time that these two signals are best aligned, or,

$$t_d = \arg\max_{t} \left(\mathbf{s}_0 \star \mathbf{s} \right) (t). \tag{5}$$

We can then shift the elements in s into an aligned signal s^* as:

$$s^{\star}(t) = s\left(t + t_d\right). \tag{6}$$

In our implementation, we take the acceleration magnitude s_{A0} as reference and align all channels of signals in S_P .

2.3. Decomposition

The goal in this subsection is to remove regular MA. As illustrated in the upper-right corner of figure 2, even the original PPG signal is affected by MA, our methods can still efficiently find the intrinsic PPG consistent to ECG signal, that is, the ground truth.

2.3.1. MSSA

To decompose multiple coupled signals, we proposed an approach the based on the famous MSSA [6] which is an non-parametric spectral estimation of multiple time series in a vector space of a predetermined dimension $M \ (= L/2 \ \text{usually})$. The technique is used to separate latent components from a collection of l signal ensembles, each with k_1, k_2, \ldots, k_l channels respectively, or

$$\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_l]$$

$$= [\mathbf{s}_{11}, \mathbf{s}_{12}, \dots, \mathbf{s}_{21}, \dots, \mathbf{s}_{lk_l}]$$
(7)

Here we introduce an operator \mathcal{S} mapping from a signal s to a M-delayed random process embedding for later inductions, as:

$$(S_M \{\mathbf{s}\})(\tau) \equiv [s(\tau), s(\tau+1), \dots, s(\tau+M-1)]^T$$
(8

Then, from the embedding of each signal in S, we construct a lag-covariance matrix C as:

$$\mathbf{C} \equiv \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1l} \\ \mathbf{C}_{21} & \mathbf{C}_{22} & \cdots & \mathbf{C}_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{l1} & \mathbf{C}_{l2} & \cdots & \mathbf{C}_{ll} \end{bmatrix}, \tag{9}$$

where C_{mn} is the lag-covariance sub-matrix regarding the signal ensemble S_m and S_n . The entries of such matrices are:

$$\left(\mathbf{C}_{mn}\right)_{ab} = \mathbf{E}\left[\left(\mathcal{S}_{M}\left\{\mathbf{s}_{ma}\right\}\right)\left(\mathcal{S}_{M}\left\{\mathbf{s}_{nb}\right\}\right)^{T}\right], \qquad (10)$$

where $\mathbf{E}\left[\cdot\right]$ denotes expectation value. Then, we decompose the lag-covariance matrix C as

$$\mathbf{C} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T,\tag{11}$$

where $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_{kM})$ contains the eigenvalues sorted in descending order, \mathbf{U} is a set of orthonormal basis spanning \mathbb{R}^{kM} , and $k = \sum_{m=1}^l k_m$.

2.3.2. Component Feature Extraction

We have now obtained some orthogonal components of the signal sorted in descending importance, and the next step is to judge which components correspond to MA, while others correspond to HR. According to the decomposition in the last part, we assume the MA, which spans the MA-space $\mathcal{V}_{\rm MA}$, will be orthogonal to the intrinsic PPG signal, which spans the HR-space $\mathcal{V}_{\rm HR}$. Under this assumption, it is relatively easy to remove MA by projecting our input signal ensemble onto $\mathcal{V}_{\rm HR}$. However, to classify the eigenvectors obtained in the last section, we need some quantitative information about each eigenvector. For the sake of clarity, we drop the subscript

g describing the number of the eigenvector here. We proposed the *eigenvalue ratios* as follows.

Firstly we partition the eigenvector **u** as:

$$\mathbf{u}^{T} = \left[\mathbf{u}_{1}^{T}, \mathbf{u}_{2}^{T}, \dots, \mathbf{u}_{l}^{T} \right]. \tag{12}$$

Then we can define some quantities using this eigenvector and the original lag-covariance matrix C to provide a measure of how much a signal ensemble have contributed on the appearance of the component \mathbf{u}_{a} .

$$\begin{cases} \bar{\lambda}_0 \equiv \lambda_0^{-1} |\lambda| \\ \bar{\lambda}_m \equiv \lambda_0^{-1} |\mathbf{u}_m^T \mathbf{C}_{mm} \mathbf{u}_m|, \end{cases}$$
(13)

where

$$\lambda = \mathbf{u}^T \mathbf{C} \mathbf{u}$$

$$= \sum_{m} \sum_{n} \mathbf{u}_m^T \mathbf{C}_{mn} \mathbf{u}_n$$
(14)

is the eigenvalue associated with this eigenvector, and

$$\lambda_0 \equiv \sum_{m} \sum_{n} \left| \mathbf{u}_m^T \mathbf{C}_{mn} \mathbf{u}_n \right| \tag{15}$$

is proposed as the absolute eigenvalue which is better to measure the relevance between this eigenvector and the original signal since two signals can be negatively correlated when they are out of phase. In such scenario, negative values will result in the cross terms in the definition of λ , cancelling the positive self-correlation values given by diagonal terms, and as a result yielding a much smaller eigenvalue λ . Defining such a λ_0 can sum up the positively correlating and the negatively correlating parts thus present a more accurate measure. $\bar{\lambda}_m$ is thus a number called the eigenvalue ratio which is smaller than unity and provides a measure on how much the signal ensemble \mathbf{S}_m has contributed to the appearance of this component \mathbf{u} .

One can collect the ratios and the frequency of a signal component

$$\mathbf{\Phi}\left(\mathbf{u}_{g}\right) \equiv \left[\bar{\lambda}_{g0}, \bar{\lambda}_{g1}, \dots, \bar{\lambda}_{gl}, f_{g}\right] \tag{16}$$

as a feature of such an eigenvector.

2.3.3. De-motion Scheme

In general, on applying the feature extraction to the joint signal S, $\bar{\lambda}_{g1}$ will describe the contribution of S_P and $\bar{\lambda}_{g2}$ will describe that of S_A . Thus, we can identify which components are in \mathcal{V}_{HR} while others are in \mathcal{V}_{MA} and should be removed.

2.3.4. Signal Recovery

Now we can reconstruct the signal from remaining components by first project the signals onto the eigenvectors, giving principal components:

$$\mathbf{A}_{qm} = \mathbf{u}_{qm}^{T} \left(\mathcal{S}_{M} \left\{ \mathbf{s}_{m} \right\} \right) \tag{17}$$

Using these principal components, we can create the reconstructed components of the signal that belong to certain eigenvectors.

$$\tilde{s}_{m}(t) = \sum_{g} I_{g}^{\star} \mathbf{E} \left[u_{gm} \left(t - \tau + 1 \right) A_{gm} \left(\tau \right) \right]$$
 (18)

All channels of the PPG signal S_P are treated with the recovery process.

2.4. Proposed TSCM Model

2.4.1. Projection Method

The core concept of this framework is the *Trajectory Space Circular Motion* (TSCM) model, which discards the traditional peak detection techniques that rely on a small number of points. The classical methods are too sensitive to outliers and noises and can be catastrophic when MA is introduced.

At the same time, a method named Local Circular Motion (LCM) [3] model suggests we plot the signal s(t) versus a delayed signal s(t+d). However, it is hard to choose the delayed parameter d in the LCM model, and hence we further extend the definition into our TSCM model as a stable framework, which, at the same time, can determine the delay.

We proposed the trajectory space projection that projects a signal s as $s \to \psi$ into the *trajectory space* and ψ , the trajectory space embedding, is defined as follows:

$$\psi \equiv \mathbf{E}\left[\left(\mathcal{S}_{1}\left\{\phi\right\}\right)\left(\mathcal{S}_{L}\left\{\mathbf{s}\right\}\right)\right],\tag{19}$$

where S is previously defined as the mapping to a delayed random process embedding, and ϕ is a *trajectory* defined as:

$$\phi(t) = e^{\mathbf{j}\omega t}, \quad 1 \le t \le N, \tag{20}$$

where N is a predetermined trajectory length. It is called a trajectory since if we plot the real part of ϕ versus its imaginary part, the resulting graph will be a circular motion around the origin, resembling the trajectory of an object around a centripetal force source. In practice, we will project the signal s to a set of trajectory in trajectory space with multiple angular velocities ω . In order to refine these trajectories, we adopt Kalman filter. Kalman filter has been proposed to treat corrupted PPG signal in previous works [7], and however the application of Kalman filter was limited to time domain, where the dynamics of PPG signal is not clear. On the other hand, as we have proposed the TSCM model, it is ready for Kalman filters to be applied on since its dynamics can be described as an arc of a circle [3].

2.5. Post-filtering

The filtered trajectories can be recovered to HR readings by counting phase revolutions. The collection of trajectories

$$\bar{\boldsymbol{\Psi}} = \left\{\bar{\boldsymbol{\psi}}_n\right\}_{n=1}^{N_{\omega}} \tag{21}$$

will yield N_{ω} HR readings, namely

$$\bar{\mathbf{\Omega}} = \left[\bar{\omega}_1, \bar{\omega}_2, \dots, \bar{\omega}_{N_{\omega}}\right],\tag{22}$$

and we can view these number as measurements for true $f_{\rm HR}$. To decide which $\bar{\omega}$ fits the true $f_{\rm HR}$ best, we perform particle filter to trace HR with state and measurements defined as

$$\begin{cases} \mathbf{x} = [f_{\text{HR}}], \\ \mathbf{z} = \frac{1}{2\pi}\bar{\mathbf{\Omega}}. \end{cases}$$
 (23)

A final HR reading by this post-filtering mechanism will be the result of our architecture.

3. EXPERIMENTAL RESULTS

3.1. Data and Parameter Setting

The dataset [4] includes the recordings from 12 subjects with a two-channel PPG signal with a tri-axis accelerometer. For every subject, the PPG signal is recorded from wrist by two pulse oximeters with green LEDs (wavelength 609 nm), and the distance between them is 2 cm. The accelerometer is as well installed on the wrist, embedded in a wristband as the oximeters do. An ECG signal is simultaneously recorded and all the signals are sampled at 125 Hz.

Over the course of recording, the subjects walked or ran on the treadmill with the following speeds:

- 1. Walk (1 \sim 2 km/h), for 0.5 minutes
- 2. Run (6 \sim 8 km/h), for 1 minute
- 3. Run (12 \sim 15 km/h), for 1 minute
- 4. Run (6 \sim 8 km/h), for 1 minute
- 5. Run (12 \sim 15 km/h), for 1 minute
- 6. Walk (1 \sim 2 km/h), for 0.5 minutes

The subjects were asked to purposely use the hand with the wristband to pull clothes, wipe sweat on forehead, and push buttons on the treadmill, in addition to freely swing.

In MSSA, the embedding length M=500, which is the half of the sample numbers in a report window. In the particle filters, all σ have been set to $3~(\mathrm{BPM})$ while the covariance matrix is $\Sigma=5^2~(\mathrm{BPM}^2)$. The HR range of interest is set as $f_{\min}=30~(\mathrm{BPM})$ to $f_{\max}=200~(\mathrm{BPM})$, and N_{ω} is chosen in a dynamic manner while its maximum value is 20.

3.2. Performance Measure

The ground truth HR is obtained using the simultaneously recorded ECG, with the number of its cardiac cycles and durations, which is denoted as $f_{\rm tru}$. The performance of the result of our framework, denoted as $f_{\rm est}$, can be evaluated by two importance numbers, which are mean absolute error

$$e_{\rm abs} = \mathbf{E}\left[|f_{\rm tru} - f_{\rm est}|\right],\tag{24}$$

and the mean percentage error

$$e_{\rm per} = \mathbf{E} \left[\frac{|f_{\rm tru} - f_{\rm est}|}{f_{\rm tru}} \right],$$
 (25)

where the expectation is taken over all report windows. Other than these two error measures, we also include Bland-Altman plot [8] to assess the agreement between the ground truth and our estimate. Besides agreement, the Pearson product-moment correlation coefficient is used to represent the correlation

3.3. Results

3.3.1. Effectiveness Verification

Figure 3 is the Bland-Altman plot, the *Limit of Agreement* (LOA) is -7.59 (BPM) through 6.65 (BPM) while Fig. 4 shows the Pearson correlation coefficient to be r=0.995

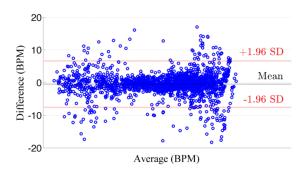


Fig. 3: The Bland-Altman plot of the estimated HR

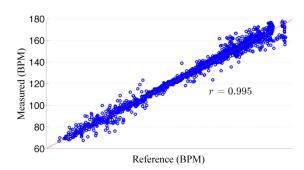


Fig. 4: The Pearson correlation of the estimated HR

Table 2: $e_{\rm abs}$ (BPM) of all subjects of different methods

Subject #	1	2	3	4	5	6	7	8	9	10	11	12	Avg.
TROIKA [4]	2.29	2.19	2.00	2.15	2.01	2.76	1.67	1.93	1.86	4.70	1.72	2.84	2.34
Ours	2.43	2.32	1.66	1.79	1.31	2.28	2.07	1.37	1.32	5.29	2.42	2.21	2.21

Table 3: $e_{\rm per}$ of all subjects of different methods

Subject #	1	2	3	4	5	6	7	8	9	10	11	12	Avg.
TROIKA [4]	1.90%	1.87%	1.66%	1.82%	1.49%	2.25%	1.26%	1.62%	1.59%	2.93%	1.15%	1.99%	1.79%
Ours	1.93%	2.14%	1.39%	1.52%	0.93%	1.78%	1.51%	1.14%	1.04%	3.26%	1.58%	1.53%	1.65%

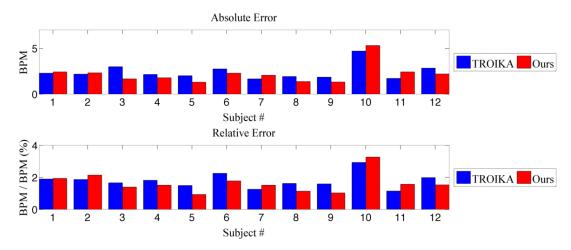


Fig. 5: The error of different methods on different subjects

3.3.2. Performance Comparison

Table 2 and 3 list out the mean absolute error $e_{\rm abs}$ using our proposed framework. Averaging over all subjects, we have an overall 2.21 (BPM) absolute error, and a 1.65% percentage error.

4. CONCLUSIONS

In this work, we have proposed a complete framework to eliminate both irregular and regular MA in a multi-channel PPG signal recorded when the subjects are under intensive activities. The framework involves an improved MSSA with *eigenvalue ratios* defined to eliminate regular MA, as well as the TSCM model which fixes irregular behavior of cardiac cycle by projecting the source signal onto a *trajectory space* where the dynamics are clear. Experiment results on all subjects have shown that our framework is competitive to the state-of-the-art method.

5. REFERENCES

 R. Yousefi, M. Nourani, and I. Panahi, "Adaptive cancellation of motion artifact in wearable biosensors," in 2012 Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (EMBC), Aug 2012, pp. 2004–2008.

- [2] M. Raghuram, K.V. Madhav, E.H. Krishna, N.R. Komalla, K. Sivani, and K.A. Reddy, "Dual-tree complex wavelet transform for motion artifact reduction of ppg signals," in 2012 IEEE International Symposium on Medical Measurements and Applications Proceedings (MeMeA), May 2012, pp. 1–4.
- [3] SM Hong, BH Jung, and D Ruan, "Real-time prediction of respiratory motion based on a local dynamic model in an augmented space," *Physics in medicine and biology*, vol. 56, no. 6, pp. 1775, 2011.
- [4] Z. Zhang, Z. Pi, and B. Liu, "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic (ppg) signals during intensive physical exercise," *IEEE Transactions on Biomedical Engineering*, vol. PP, no. 99, pp. 1–1, 2014.
- [5] Abraham. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures.," *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.
- [6] Robert Vautard and Michael Ghil, "Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series," *Physica D: Nonlinear Phenomena*, vol. 35, no. 3, pp. 395–424, 1989.
- [7] S Seyedtabaii and L Seyedtabaii, "Kalman filter based adaptive reduction of motion artifact from photoplethysmographic signal," in *Proceedings of World Academy of Science, Engineering and Technology*. Citeseer, 2008, vol. 27.
- [8] J Martin Bland and DouglasG Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *The lancet*, vol. 327, no. 8476, pp. 307–310, 1986.