

RESEARCH ARTICLE

Non-contact heart rate estimation based on singular spectrum component reconstruction using low-rank matrix and autocorrelation

Weibo Wang^{1☯*}, Zongkai Wei^{1☯}, Jin Yuan^{1‡}, Yu Fang^{1‡}, Yongkang Zheng^{2‡}

1 Electrical Engineering and Electronic Information, Xihua University, Chengdu, China, **2** State Grid Sichuan Electric Power Research Institute, Chengdu, China

☯ These authors contributed equally to this work.

‡ JY, YF and YZ also contributed equally to this work.

* wangweibo@mail.xhu.edu.cn



OPEN ACCESS

Citation: Wang W, Wei Z, Yuan J, Fang Y, Zheng Y (2022) Non-contact heart rate estimation based on singular spectrum component reconstruction using low-rank matrix and autocorrelation. PLoS ONE 17(12): e0275544. <https://doi.org/10.1371/journal.pone.0275544>

Editor: Felix Albu, Valahia University of Targoviste: Universitatea Valahia din Targoviste, ROMANIA

Received: May 6, 2022

Accepted: September 19, 2022

Published: December 30, 2022

Copyright: © 2022 Wang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The UBFC-rPPG dataset is a third-party source that are focused specifically on rPPG analysis. Researchers can apply for permission to access the data by contacting the owner. His email is: yannick.benezeth@u-bourgogne.fr. The authors confirm they had no special access privileges to the data.

Funding: This work is supported by the National Natural Science Foundation of China (No. 61901393) and the National Natural Science Foundation of China (No. 61571371). There was no

Abstract

The remote photoplethysmography (rPPG) based on cameras, a technology for extracting pulse wave from videos, has been proved to be an effective heart rate (HR) monitoring method and has great potential in many fields; such as health monitoring. However, the change of facial color intensity caused by cardiovascular activities is weak. Environmental illumination changes and subjects' facial movements will produce irregular noise in rPPG signals, resulting in distortion of heart rate pulse signals and affecting the accuracy of heart rate measurement. Given the irregular noises such as motion artifacts and illumination changes in rPPG signals, this paper proposed a new method named LA-SSA. It combines low-rank sparse matrix decomposition and autocorrelation function with singular spectrum analysis (SSA). The low-rank sparse matrix decomposition is employed to globally optimize the components of the rPPG signal obtained by SSA, and some irregular noise is removed. Then, the autocorrelation function is used to optimize the global optimization results locally. The periodic components related to the heartbeat signal are selected, and the denoised rPPG signal is obtained by weighted reconstruction with a singular value ratio. The experiment using UBFC-RPPG and PURE database is performed to assess the performance of the method proposed in this paper. The average absolute error was 1.37 bpm, the 95% confidence interval was -7.56 bpm to 6.45 bpm, and the Pearson correlation coefficient was 98%, superior to most existing video-based heart rate extraction methods. Experimental results show that the proposed method can estimate HR effectively.

1 Introduction

Heart rate (HR) is an important indicator to measure human physiological activities, monitoring human health and emotional state. It has been widely used in cardiovascular disease diagnosis, health assessment, and emotional detection [1–5]. The traditional contact HR measurement methods include electrocardiogram (ECG) and photoplethysmography (PPG).

additional external funding received for this study. The funders had roles in study design and analysis.

Competing interests: The authors have declared that no competing interests exist.

Although the measurement accuracy of the two methods is high, the detection type that relies on specific sensors to contact the subjects' skin is not suitable for patients with skin damage and newborns. In contrast, the non-contact heart rate measurement using microwave Doppler or computer vision technology has attracted more and more research attention due to its non-contact advantages. LU G et al. [6] measured HR using Doppler radar. Pavlidis et al. [7] successfully extracted the HR of the subjects by analyzing the facial thermal infrared images. However, these devices are expensive and require complex hardware support, making it challenging to promote practical applications.

Based on the principle of PPG, remote photoplethysmography (rPPG) can collect the facial skin color changes of subjects by the camera in a non-contact way to extract the PPG signal and detect the HR and its changes related to cardiac activity [8]. Verkruyse et al. first verified the possibility of using facial video to measure HR and found that the color intensity of the facial area collected by ordinary cameras had periodic changes correlated with blood volume pulse (BVP) [9]. Compared with non-contact detection methods such as microwave Doppler and thermal infrared, the cameras reduce the detection cost and operation complexity, which has obvious advantages in unconstrained scenes such as neonatal monitoring [10], fatigue driving judgment [11], and pressure monitoring [12]. However, because the color change of the human face caused by cardiovascular activity is subtle, it will be affected by noise such as illumination change and motion artifact, which makes the measurement accuracy disturbed. Many research methods on rPPG denoising have been proposed in recent years to solve this problem. Blind source separation (BSS) technology is usually used to remove the noise contained in rPPG. Poh et al. decomposed the original RGB three-channel signals into three independent source signals through independent component analysis (ICA) and extracted the HR from the second component [13]. Lewandowska et al. proposed a method based on principal component analysis (PCA) for HR measurement and evaluated the measurement effect under different regions of interest (ROI), different combinations of color channels, and different illumination conditions. The results showed that the accuracy of this method was affected by the environmental illumination variables [14]. Different from the assumption of linear mixing for BSS, based on the two-color reflection model, Haan et al. eliminated the interference caused by the mirror reflection component through the difference between the two orthogonal chromaticity signals. In addition, they also proposed a rPPG method based on the normalized BVP vector signal in the normalized RGB color space to improve the motion robustness [15, 16]. Li et al. proposed an anti-interference method of normalized least mean square (NLMS) adaptive filter to correct illumination change [17]. Chen et al. reduced the influence of environmental light changes by decomposing green channels with the ensemble empirical mode decomposition (EEMD) of the Hilbert-Huang transform [18]. Wang et al. used image redundancy to offset the influence of the facial movement [19]. Kumar et al. solved the problem of low signal-to-noise ratio (SNR) caused by deep skin color and weak illumination conditions by using a weighted average to combine the skin color change signals of different tracking regions of the face [20]. Based on the multi-task convolutional neural network, Yue et al. proposed a framework combining empirical mode decomposition and permutation entropy to reduce the impact of face jitter and shooting environment [21]. Niu et al. designed a transfer learning strategy to estimate HR from the spatiotemporal representation of HR information [22]. In addition, the author further improves the HR estimation method of face videos based on channel and spatio-temporal attention mechanism [23]. Song et al. designed a new pulse wave generation framework based on generative adversarial network to improve waveform quality, thereby improving the accuracy of heart rate detection [24].

This paper proposes a non-contact HR measurement method (LA-SSA) using low-rank sparse matrix decomposition and autocorrelation function for SSA decomposition component

selection and reconstruction. The structure of the article is as follows. Section 2 introduces the basic framework of HR detection and elaborates on the basic principle of the LA-SSA method. Section 3 introduces the database information, evaluation index, and the experimental results on the UBFC-RPPG and PURE database. Concluding remarks are given in Section 4.

2 Methods

Based on SSA decomposition components, the proposed method introduces low-rank sparse matrix decomposition and autocorrelation function in component reconstruction, which reduces the interference of irregular noise contained in the signal and retains the periodic components of HR correlation. It realizes the accurate extraction of pulse signals containing HR information from facial videos. The framework of this study is shown in Fig 1. Firstly, face tracking and skin detection are performed for each frame of the face video, and the detected facial skin is regarded as the ROI. Then ROI is divided into the three RGB channels. It is known that the HR-related information contained in the G channel is more prosperous than that contained in the other two channels. When the video sample is seriously disturbed by noise, the introduction of the other two channels may bring more noise than the HR information. Therefore, only the G channel is selected in this paper to extract the HR-related information. Spatially averaged over all pixels in the ROI to reduce the noise, and the average pixel value of each frame in the G channel is used to form the original facial signal. The original signal is smoothed and preprocessed by detrending, normalization, and five-point moving average filtering. In preprocessing, EEMD is introduced for rough denoising of the signal. Based on the results of rough denoising, the LA-SSA algorithm is used for further denoising of the signal to extract the HR information.

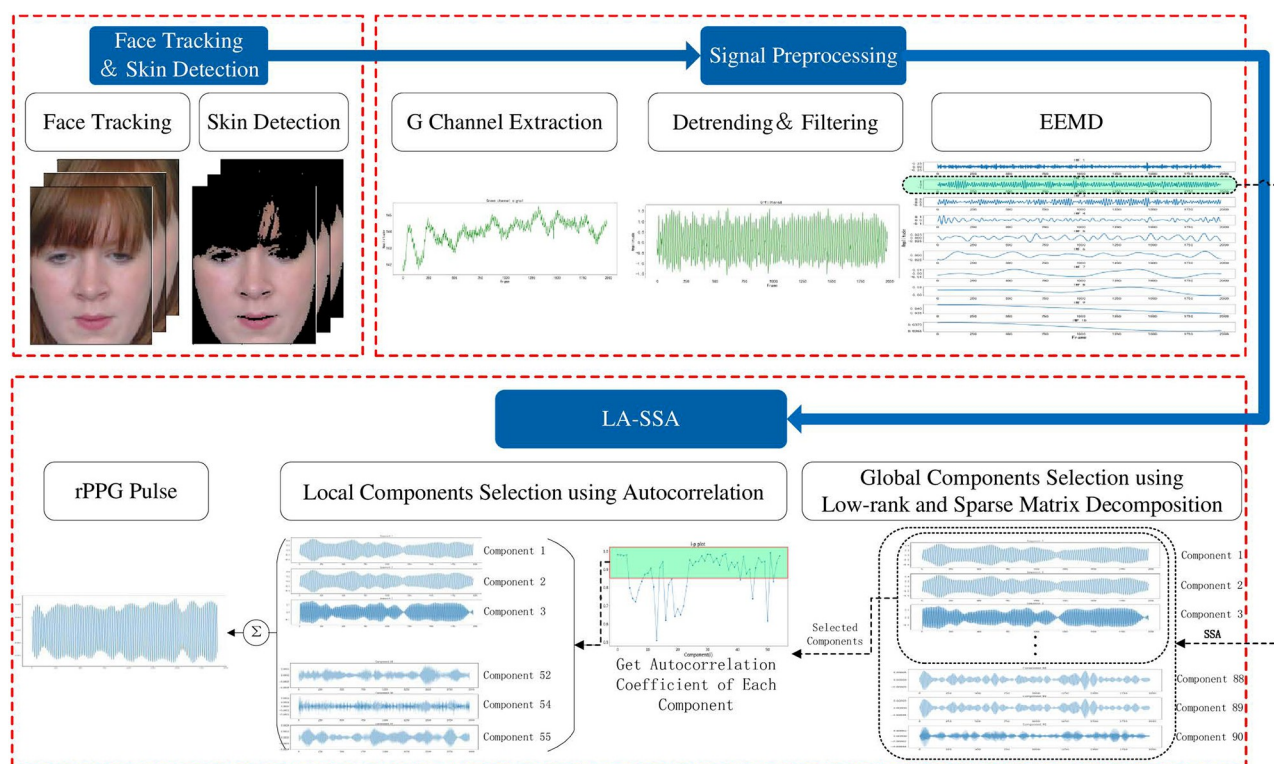


Fig 1. Face video HR estimation framework.

<https://doi.org/10.1371/journal.pone.0275544.g001>

2.1 Face tracking

Since head movement and background noise greatly influence pulse waves, accurate face detection and tracking are the keys to collecting high-quality original pulses. In the literature, the method of Viola-Jones combined with Kanade-Lucas-Tomasi is mainly used to realize the rapid detection of the human face. However, this method has good stability only when the measured object keeps the head still, which is greatly affected by the head movement and prone to false and missed detection. In order to ensure stable face detection in a more realistic environment, this paper uses the multi-task convolutional neural network (MTCNN) [25] model for face detection. The model is based on the idea of candidate box and classification. Through the cascade of P-Net (Proposal Network), R-Net (Refine Network), and O-Net (Output Network), it can realize fast and efficient face and feature point detection.

In face detection, the MTCNN face detection model is used to track the face of each frame in the video. If the face region is not detected in the current frame, the face region of adjacent frames is used. However, the results obtained by face detection still contain some non-skin areas, including hair, eyebrows, nostrils, and a small number of background areas, which cannot provide any helpful information related to HR. On the contrary, the eye-blinking and slight movement of lips will introduce artifacts into pulse signals. Therefore, the skin detection algorithm based on RGB-H is applied to exclude non-skin areas as much as possible after obtaining the face area. A frame of the face and skin detection results are shown in Fig 2.

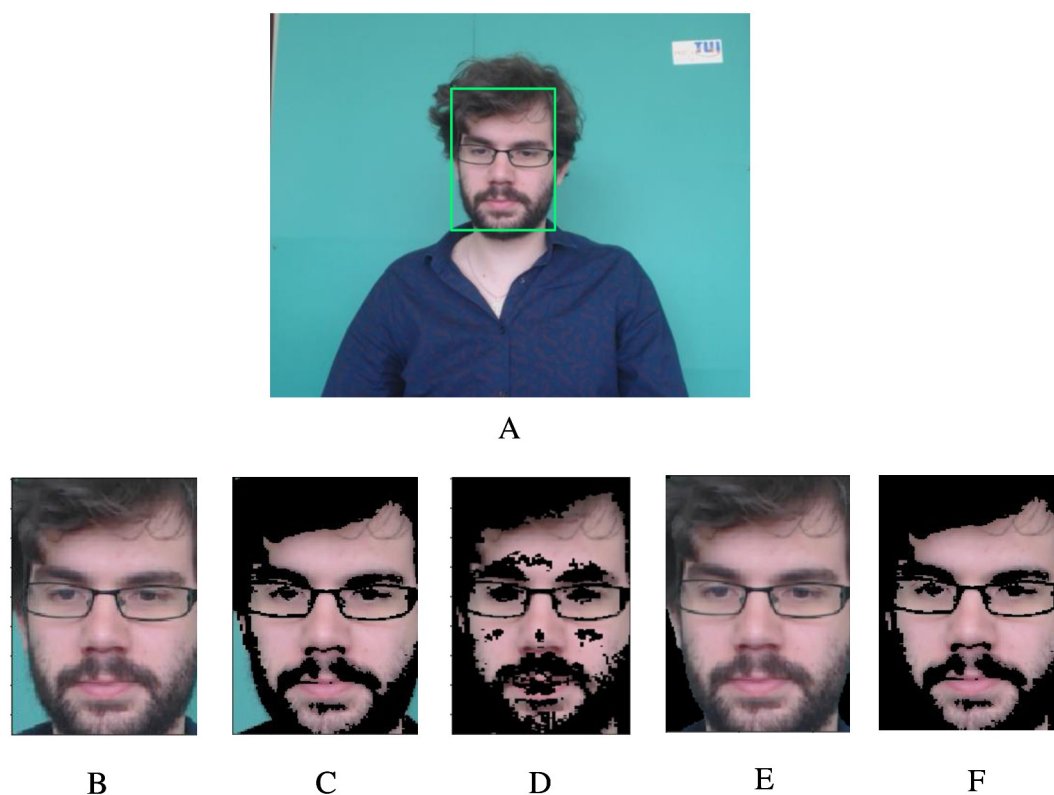


Fig 2. Face and skin detection results. A: The result of face detection. B: The cropped face. C~F: The skin detection results of RGB, YCrCb, YCrCb+OTSU, RGB-H.

<https://doi.org/10.1371/journal.pone.0275544.g002>

2.2 Skin detection based on RGB-H threshold

After the face is detected, ROI is identified in the face to extract information related to cardiac activity. The selection of ROI directly affects the accuracy and reliability of the HR extraction algorithm. Usually, the whole face, the rectangular area of the face cut with a particular proportion, and the forehead, nose, or cheek area can be selected as ROI. Rapczynski et al. proved that the whole facial skin and forehead could obtain more accurate HR information [26]. More skin pixels can improve SNR to obtain a clearer rPPG signal and better HR estimation. Since the forehead may be covered by hair, and the deflection of the head will affect the extraction of ROI, this paper selects the entire facial skin area as ROI to extract HR-related signals. The current skin detection algorithms mainly include threshold-based, model-based, region-based [27], etc. Considering that the skin detection method based on threshold can effectively and accurately detect the skin with less computation and shorter time. By comparing the detection effect of four skin detection algorithms based on the RGB threshold, YCbCr threshold, YCbCr space Cr component + OTSU segmentation, and RGB-H threshold, this paper selects the RGB-H threshold segmentation algorithm to detect the skin of each frame. The whole face skin area detected is used as ROI to realize HR detection.

$$\begin{cases} R > 95 \wedge G > 40 \wedge B > 20 \wedge \\ \max(R, G, B) - \min(R, G, B) > 15 \wedge \\ R > G \wedge R > B \wedge \\ 0 \leq H \leq 60 \end{cases} \quad (1)$$

The H and RGB values of each pixel are compared to the thresholds to judge whether the pixel is a skin pixel. The final skin detection result is obtained by removing non-skin pixels outside the selection criteria.

2.3 Preprocessing

In order to minimize the noise component in the signal before LA-SSA decomposition and reconstruction, based on the characteristics of LA-SSA single-channel input, after comparing the HR estimation accuracy of LA-SSA using different inputs such as G channel, signal obtained by CHROM method, ICA filtering results, and EEMD decomposition screening signal, EEMD is introduced into the preprocessing to realize the rough denoising of the signal in this paper. Proposed by Wu et al. [28], the EEMD algorithm is an adaptive time-frequency analysis method. It is based on EMD algorithm and can decompose nonlinear and non-stationary signals into a finite number of intrinsic mode functions (IMFs) according to the time scale characteristics of the signal itself, without setting any basis functions in advance. Meanwhile, by adding white noise to the original signal, this method maps different time scale components to the reference time scale related to white noise, and the white noise is removed by multiple mean, which effectively solves the problem of modal aliasing in EMD. This method has more evident advantages in tracking physiological signals such as pulse period than the existing stationary methods. At present, many studies have used the EEMD algorithm to solve the problem of noise removal in rPPG signals [18, 29, 30].

Firstly, ROI is separated into the three RGB channels. The pixels in the ROI are spatially averaged and the process is repeated for each video frame to form the original facial signal which contains the rPPG signal. The prior smoothing method [31] is used to detrend the original signal to eliminate the low-frequency trend in the signal. The detrended signal is normalized and then smoothed by five-point moving average filtering. The standard deviation of

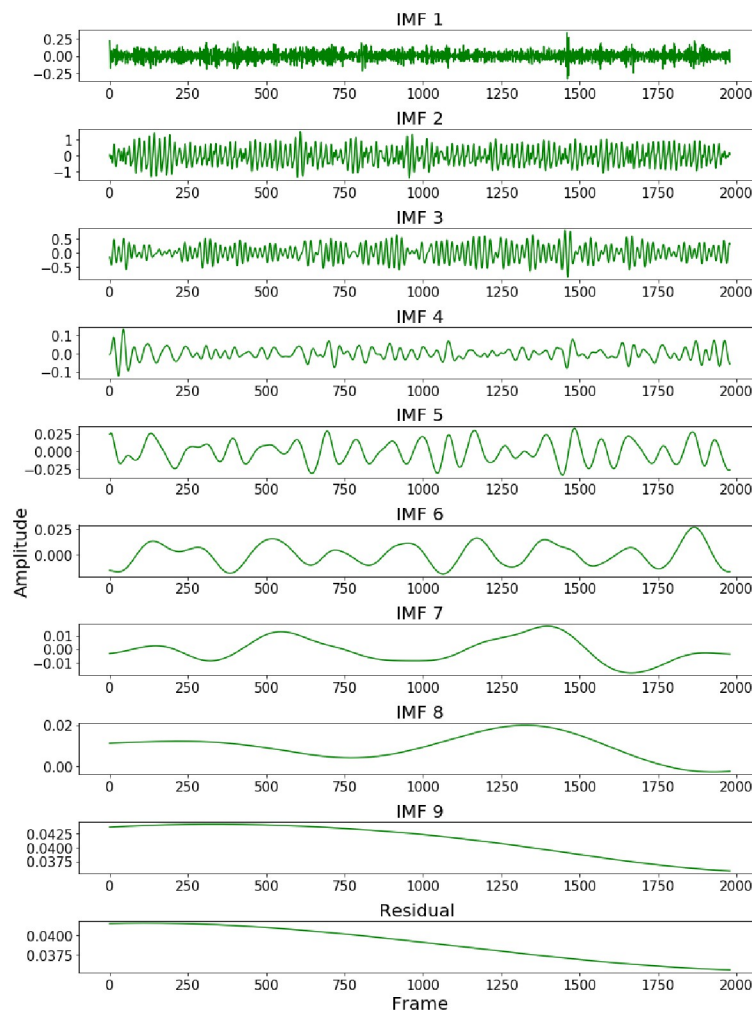


Fig 3. The IMFs and the residual of the filtered signal.

<https://doi.org/10.1371/journal.pone.0275544.g003>

additional noise is set to 0.05, and the average number of ensembles is set to 100. The filtered signal is decomposed by EEMD, and nine IMFs and one residual are obtained, as shown in Fig 3. Since the IMF related to the HR signal cannot be identified in the time domain, the FFT of nine IMFs is carried out. The component that the spectrum contains the highest peak in the heart rate range (0.7 Hz—3 Hz) is selected as the preprocessing output result to realize the rough denoising of the signal. It can be seen from Fig 4, that the target signal is IMF2. Fig 5 shows the extracted G-channel signal and the preprocessed signal obtained after EEMD decomposition and screening. It can be seen that the preprocessed signal is relatively smooth as a whole, and the low-frequency trend is removed, achieving the effect of coarse signal denoising.

2.4 Proposed method: LA-SSA

SSA is a non-parametric decomposition technique, and its calculation is different from the wavelet series expansion or AR algorithm. It neither needs to use the basic function nor assume a specific model. It decomposes the time series into multiple variable components

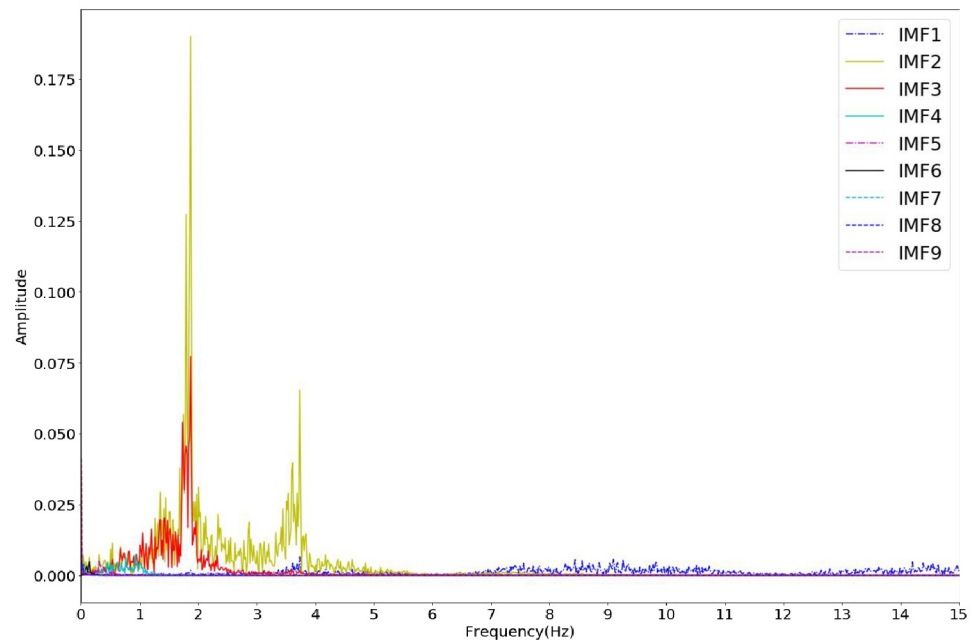


Fig 4. Frequency domain analysis of the decomposed IMF components.

<https://doi.org/10.1371/journal.pone.0275544.g004>

based on the singular value decomposition (SVD) of the time series trajectory matrix, without any prior knowledge of the time series, and achieves the effect of adaptive noise reduction of the signal by grouping, screening and reconstructing the target components [32]. In this section, based on the denoising results of EEMD, the method named LA-SSA is proposed to further denoise the rPPG signal. In the LA-SSA method, the preprocessed signal is decomposed by the SSA method to obtain several related components. Then the signal reconstruction is carried out with less noise by selecting the appropriate components using low-rank sparse matrix decomposition and the autocorrelation function, thereby improving the quality of the rPPG signal. The main steps of the proposed LA-SSA method are as follows:

1. Trajectory matrix construction

Assuming that the length of time series $x(n)$ is N , the length of the sampling window is set as L ($1 < L < N$), and the number of windows is $K = N - L + 1$. The trajectory matrix \mathbf{X} can be expressed as:

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{bmatrix} \quad (2)$$

Usually, L is a quarter of the data length [33]. When the data is relatively long and periodic, the best choice for the value of L is the quarter of the longest period in the data. In the rPPG signal, the longest fluctuation cycle caused by cardiovascular-related activities is about 12 s [34]. Therefore, the value of L is set as the number of data points for 3s in this paper.

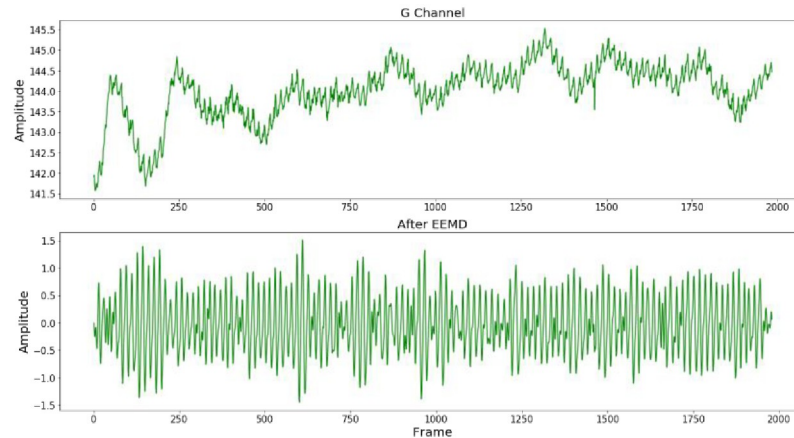


Fig 5. The waveforms of Green channel and after pretreatment.

<https://doi.org/10.1371/journal.pone.0275544.g005>

2. Singular value decomposition

\mathbf{X} is decomposed by SVD and expressed as the sum of d component matrices. Then \mathbf{X} can be represented as:

$$\mathbf{X} = \sum_{i=1}^d \mathbf{X}_i = \sum_{i=1}^d \sqrt{\lambda_i} u_i v_i^T (i = 1, 2, \dots, d) \quad (3)$$

Where d is the number of non-zero singular values and $d \leq \min(L, K)$, \mathbf{X}_i is the i -component matrix of \mathbf{X} , $\sqrt{\lambda_i}$, u_i and v_i are the i th singular value of \mathbf{X} and the corresponding left and right singular vectors, respectively. $\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_d}$ is the singular spectrum of trajectory matrix \mathbf{X} . It is known that the larger the singular value is, the more significant the contribution of the corresponding component matrix to the trajectory matrix is. So operations like eliminating noise by reconfiguring a new signal removing partial component matrix \mathbf{X}_i can be performed.

3. Components filtration

(a). Global optimization based on low-rank sparse matrix decomposition

In the singular spectrum, it is believed that the component signals obtained by component matrixes corresponding to the first s larger singular values contain the primary information of the signal, while the components corresponding to the smaller singular values mainly reflect the noise interference and other components. Fig 6. shows the original signal before decomposition, and the reconstructed signal components corresponding to the first 5 singular values. And the ratio curve of singular values is shown in Fig 7. Obviously, the contribution to reconstructing the original signal increases as the proportion of singular values increases. Therefore, the rPPG waveform can be reconstructed by screening the first s components to eliminate noise components unrelated to HR information. It is essential to determine a proper value of s , if s is too large, part of the noise component will be mixed into the reconstructed rPPG waveform, reducing noise reduction performance, and if s is too small, some helpful information related to HR will be eliminated. In this paper, the low-rank sparse matrix decomposition is employed to get the best approximate low-rank matrix \mathbf{A} of \mathbf{X} , and the rank of \mathbf{A} is the value of s [35]. Then the global optimization of component matrixes can be realized by selecting

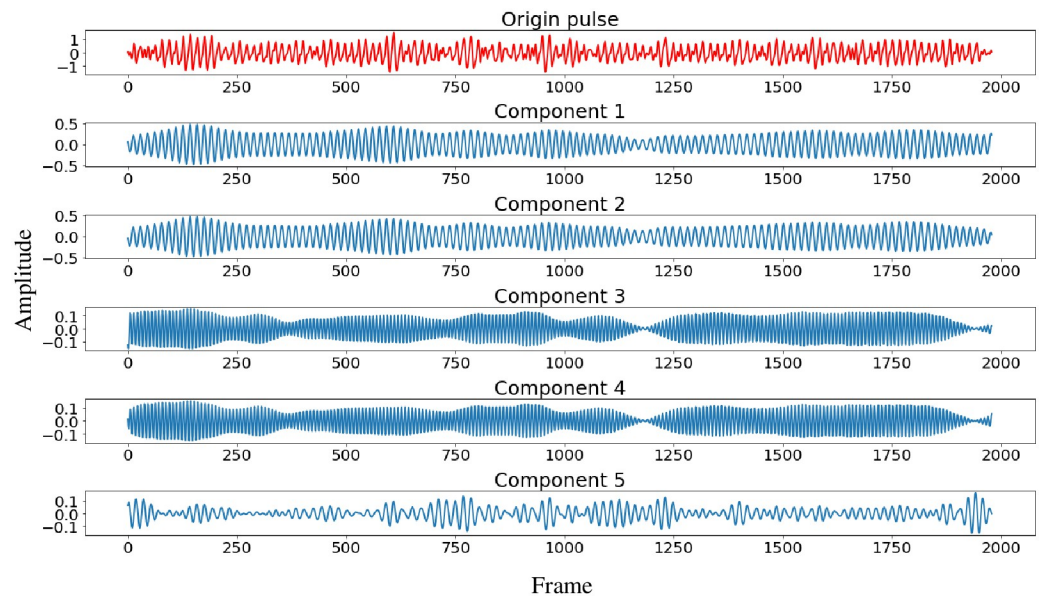


Fig 6. Original signal and its first five components.

<https://doi.org/10.1371/journal.pone.0275544.g006>

component matrixes corresponding to the first s singular values. Suppose \mathbf{X} is affected by random (sparse) noise. In that case, its low-rankness will be destroyed and become full-rank, which makes \mathbf{X} contains much redundant information besides HR information, thus affecting the accuracy of heart rate detection. then through the low-rank and sparse matrix decomposition, the denoise problem can be described as:

$$\begin{cases} \min_{\mathbf{A}, \mathbf{E}} \|\mathbf{A}\|_* + \eta \|\mathbf{E}\|_0, \eta = 1/\sqrt{\max(m, n)} \\ \text{s.t. } \mathbf{X} = \mathbf{A} + \mathbf{E} \end{cases} \quad (4)$$

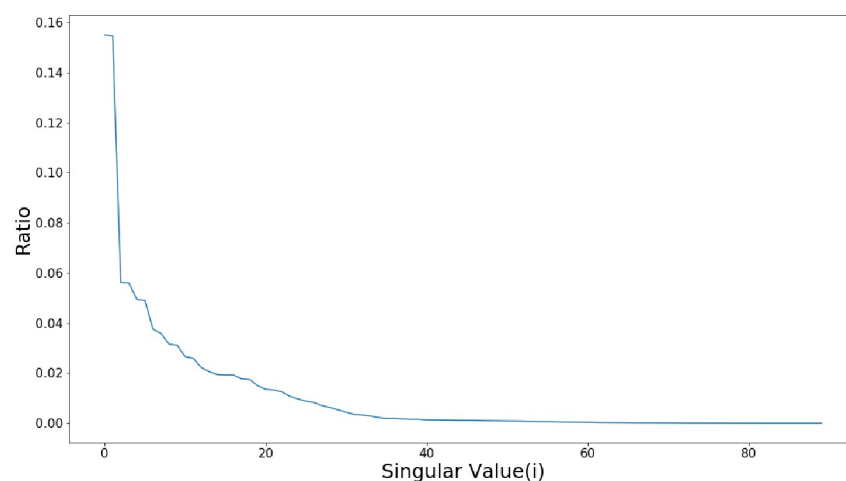


Fig 7. Proportion curves of singular values.

<https://doi.org/10.1371/journal.pone.0275544.g007>

Table 1. EALM algorithm.

| | |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Input: | X. |
| Initialization: | $\mathbf{A}_0 = \mathbf{E}_0 = 0$, \mathbf{Y}_0 shown in Eq (6), $\mu_0 > 0$, $\rho > 1$. |
| Process: | Repeat until convergence: Repeat until convergence: Given \mathbf{E} , optimize \mathbf{A} : $\mathbf{A}_{k+1} = D_{1/\mu}(\mathbf{X} - \mathbf{E}_k + \mu_k^{-1}\mathbf{Y}_k)$. Given \mathbf{A} , optimize \mathbf{E} : $\mathbf{E}_{k+1} = T_{\lambda/\mu}(\mathbf{X} - \mathbf{A}_{k+1} + \mu_k^{-1}\mathbf{Y}_k)$. Update $\mathbf{Y}_{k+1} = \mathbf{Y}_k + \mu_k(\mathbf{X} - \mathbf{A}_{k+1} - \mathbf{E}_{k+1})$. Update $\mu_{k+1} = \rho\mu_k$. Update $k = k + 1$. |
| Output: | $\mathbf{A}_k, \mathbf{E}_k$. |

<https://doi.org/10.1371/journal.pone.0275544.t001>

Where \mathbf{A} and \mathbf{E} represent low-rank matrix and sparse noise matrix, respectively, $\|\cdot\|_*$ and $\|\cdot\|_0$ denote kernel norm and zero norm of the matrix entries, respectively, m and n represent the number of rows and columns of matrix \mathbf{X} , respectively. The exact augmented Lagrange multiplier algorithm (EALM) solves this optimization problem (Table 1). The Lagrange function is defined as:

$$L(\mathbf{A}, \mathbf{E}, \mathbf{Y}, \mu) = \|\mathbf{A}\|_* + \eta\|\mathbf{E}\|_0 + \langle \mathbf{Y}, \mathbf{X} - \mathbf{A} - \mathbf{E} \rangle + \frac{\mu}{2} \|\mathbf{X} - \mathbf{A} - \mathbf{E}\|_F^2 \quad (5)$$

Where μ is the penalty coefficient, $\langle \mathbf{Y}, \mathbf{X} - \mathbf{A} - \mathbf{E} \rangle = \text{Tr}(\mathbf{Y}^T(\mathbf{X} - \mathbf{A} - \mathbf{E}))$, and the initial value of Lagrange multiplier \mathbf{Y} can be expressed as:

$$\mathbf{Y}_0 = \text{sgn}(\mathbf{X}) / \max(\|\text{sgn}(\mathbf{X})\|_2, \eta^{-1}\|\text{sgn}(\mathbf{X})\|_\infty) \quad (6)$$

Where $\|\cdot\|_\infty$ and $\|\cdot\|_2$ represent infinite norm and 2-norm of the matrix entries, respectively. Then the optimal approximate low-rank matrix of \mathbf{X} can be iteratively solved by the following EALM method:

Where $T_\tau(x) = \text{sgn}(x) \max(|x| - \tau, 0)$, $D_\tau(\mathbf{M}) = \mathbf{U} T_\tau(\Sigma) \mathbf{V}^T$. The value of coefficient ρ determines the convergence rate, which is usually between 1.1 and 2. It has been proven in [36] that the Lagrange multiplier \mathbf{Y} is sufficient to guarantee the linear convergence of the EALM algorithm when $\mathbf{X} - \mathbf{A} - \mathbf{E}$ is continuously differentiable. The rank of \mathbf{A}_k is the value of s , then the global optimization of component matrixes by preserving corresponding first s component matrixes of \mathbf{X} can be performed.

For the reserved first s component matrixes, each matrix \mathbf{X}_i is reduced to the corresponding time series component with length N by diagonal averaging (the first 5 corresponding time series are as shown in Fig 6). Given that the dimension of the component matrix \mathbf{X}_i is $L \times K$, further define $L^* = \min(L, K)$, $K^* = \max(L, K)$, and transform the matrix \mathbf{X}_i into a sequence $[z_1, z_2, \dots, z_N]$ of length N by the following diagonal average formula:

$$z_k = \begin{cases} \frac{1}{k} \sum_{m=1}^k x_{m,k-m+1} & 1 \leq k < L^* \\ \frac{1}{L^*} \sum_{m=1}^{L^*} x_{m,k-m+1} & L^* \leq k \leq K^* \\ \frac{1}{N-k+1} \sum_{m=k-K^*+1}^{N-K^*+1} x_{m,k-m+1} & K^* < k \leq N \end{cases} \quad (7)$$

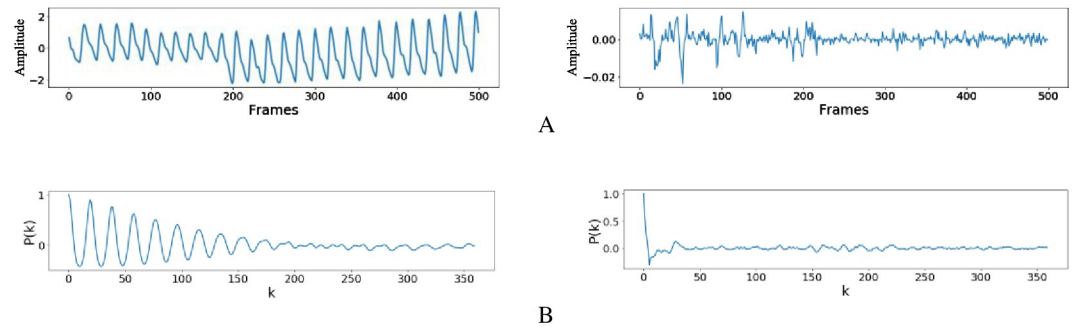


Fig 8. The autocorrelation of HR-related signal and noise signal. A: The HR-related signal(left), the noise signal(right). B: The autocorrelation of HR-related signal(left), The autocorrelation of noise signal(right).

<https://doi.org/10.1371/journal.pone.0275544.g008>

(b). Local preferences based on autocorrelation functions

It is known that HR-related signals in rPPG signals are essentially periodic (or at least quasi-periodic). Therefore, based on global optimization results, the periodicity of HR-related signals is used as a priori information to select the most periodic components of the first s signals. Thus the autocorrelation coefficient is adopted as the screening criterion of the periodic metric. The autocorrelation coefficient $P_i(k)$ of the i th component is defined as follows:

$$P_i(k) = \frac{1}{\sigma_i^2} \sum_{t=0}^{n-k} \{z_i(t) - \mu_i\} \{z_i(t+k) - \mu_i\} \quad (i = 1, 2, \dots, s) \quad (8)$$

Where $z_i(t)$ is obtained from the corresponding component matrix \mathbf{X}_i by diagonal averaging, μ_i and σ_i^2 are the average and variance of $z_i(t)$, respectively, $z_i(t+k)$ represents the sequence obtained by shifting the elements in $z_i(t)$ backward to k positions. By definition, when $k = 0$, $P_i(k)$ takes the maximum of 1. Fig 8 depicts the autocorrelation within 12 s of the HR-related signal and the noise signal obtained by low-rank sparse decomposition. It can be seen that if the signal is periodic or quasi-periodic, some peaks will appear in the k th order autocorrelation. And the more periodic the signal, the greater the peak. Therefore, for the first s components screened by low-rank sparse matrix decomposition, the autocorrelation peak of the components related to HR is often more significant than that of the components related to intermittent noise.

Then the maximum autocorrelation coefficient ρ_i of each component can be defined as follows:

$$\rho_i = \max_{2 \leq p \leq J_i} P_i(k_p) \quad (i = 1, 2, \dots, s) \quad (9)$$

Where k_p is the displacement order of the p th peak in $P_i(k)$ and J_i is the number of all peaks. Each component can obtain the maximum autocorrelation coefficient corresponding to the current component through Eq (9) (as shown in Fig 9). Fig 10 shows the maximum autocorrelation coefficient of the first s components obtained by global optimization, and the shadow part represents the components which ρ_i greater than 0.85. Usually, component $\rho_i > 0.8$ is selected for reconstruction, which can reduce the HR estimation error of the reconstructed signal [37]. In this paper, the threshold is set to 0.85. The corresponding component with the ρ_i higher than the threshold is retained to eliminate the noise signal with relatively weak periodicity.

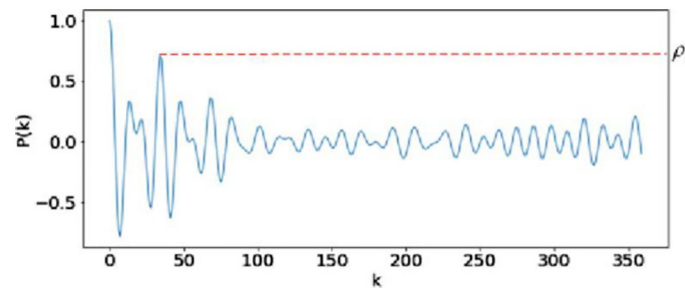


Fig 9. Autocorrelation coefficient $P_i(k)$ and maximum autocorrelation coefficient ρ_i of component i .

<https://doi.org/10.1371/journal.pone.0275544.g009>

4. Signal reconstruction

The components obtained by autocorrelation screening are weighted and superposed according to the proportion of singular values, and the reconstructed signal x_{rc} can be expressed as:

$$x_{rc} = \sum_{i \in S} w_i z_i \quad (10)$$

Where w_i is the proportion of singular value corresponding to z_i in the total singular value, the denoised rPPG signal can finally be obtained after the above steps. Fig 11 shows the actual performance of the proposed LA-SSA method. As shown in Fig 11, the waveform denoising effects are compared about the G-channel signal after five-point moving average filtering, EEMD denoising results, and LA-SSA restored waveform. The red box represents the interference of irregular noise caused by motion or illumination changes on the waveform, and the green box represents the denoising effect. It can be seen that the denoising results of the filtered G-channel signal or EEMD are still subject to the interference of noise to varying degrees. However, the noise caused by motion or illumination changes is

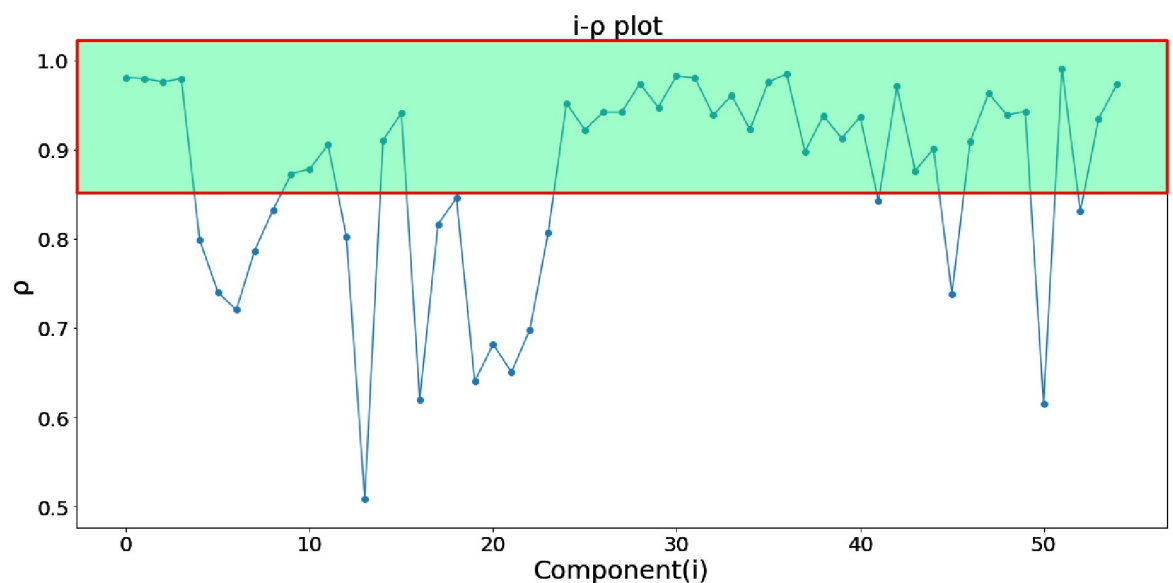


Fig 10. Screening of the first s maximum autocorrelation coefficients.

<https://doi.org/10.1371/journal.pone.0275544.g010>

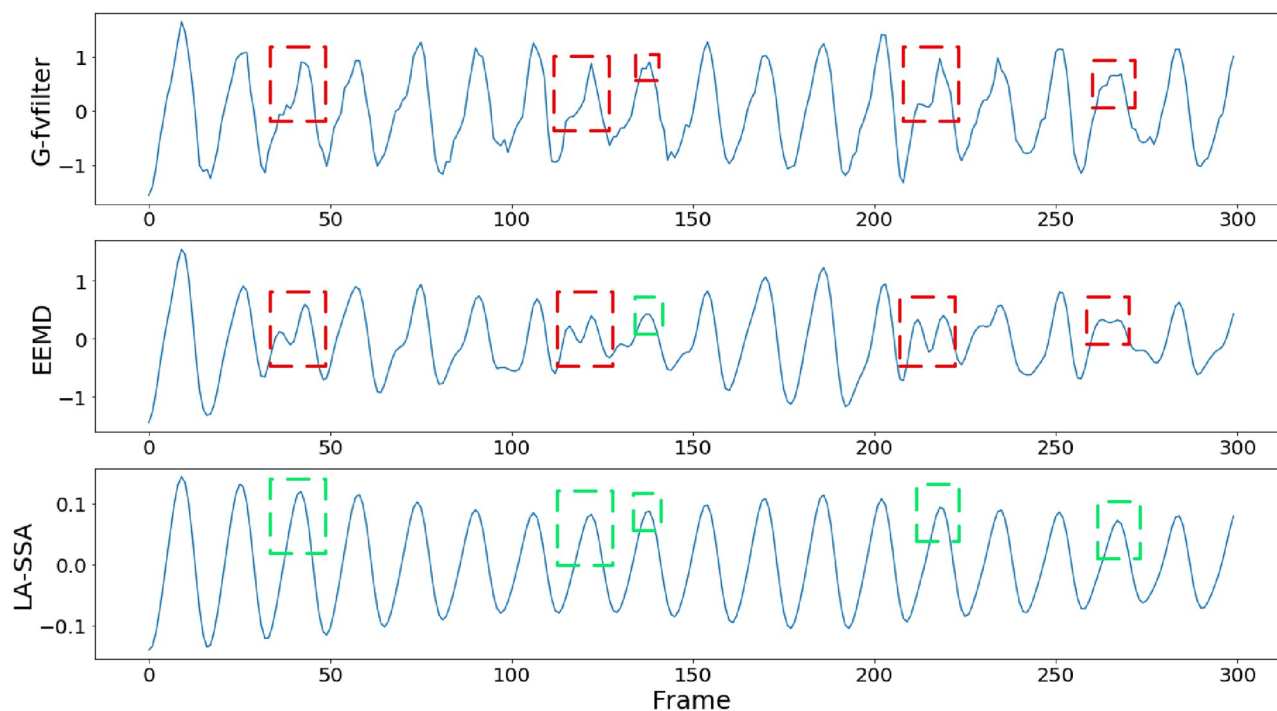


Fig 11. Waveform comparison of filtered G-channel, EEMD and LA-SSA.

<https://doi.org/10.1371/journal.pone.0275544.g011>

significantly reduced after using LA-SSA method. Fig 12 compares the rPPG obtained by LA-SSA recovery with the reference PPG waveform and gives the spectrum of the two signals. It can be seen that the high correlation between rPPG and PPG signals and the consistency of HR obtained by the two signals. The results show that LA-SSA effectively removed the irregular noise contained in rPPG and improved the accuracy of HR estimation.

3 Experiments

3.1 Database

UBFC-RPPG database [38]: This dataset contains 50 videos, which are composed of two parts. The first part (marked as SIMPLE) contains eight videos under ideal conditions, and participants are required to sit still with their eyes closed during recording. The second part (marked REALISTIC) contains 42 real-world videos. Subjects were asked to play a time-sensitive mathematical game to increase pulse frequency and maintain HR diversity while simulating typical human-computer interaction scenarios. The video was captured by the Roche C920 HD Pro camera, which was placed 1 meter away from the subjects. The videos are not compressed in 8-bit RGB format. The frame rate of 30 fps and the spatial resolution of 640×480 pixels. Each video takes about 2 minutes, and the Contec Medical CMS50E collects PPG pulse signals at a sampling rate of 60 Hz. In this paper, only the REALISTIC videos are used, each video with 30s window length and 1s step to extract HR.

PURE database [51]: This dataset consists of 10 persons (8 males and 2 females) performing different, controlled head motions in front of a camera. The head motions contain steady (S, The subject was sitting still and looks directly into the camera avoiding head motion), talking (T, the subjects were asked to talk while avoiding additional head motion), slow translation

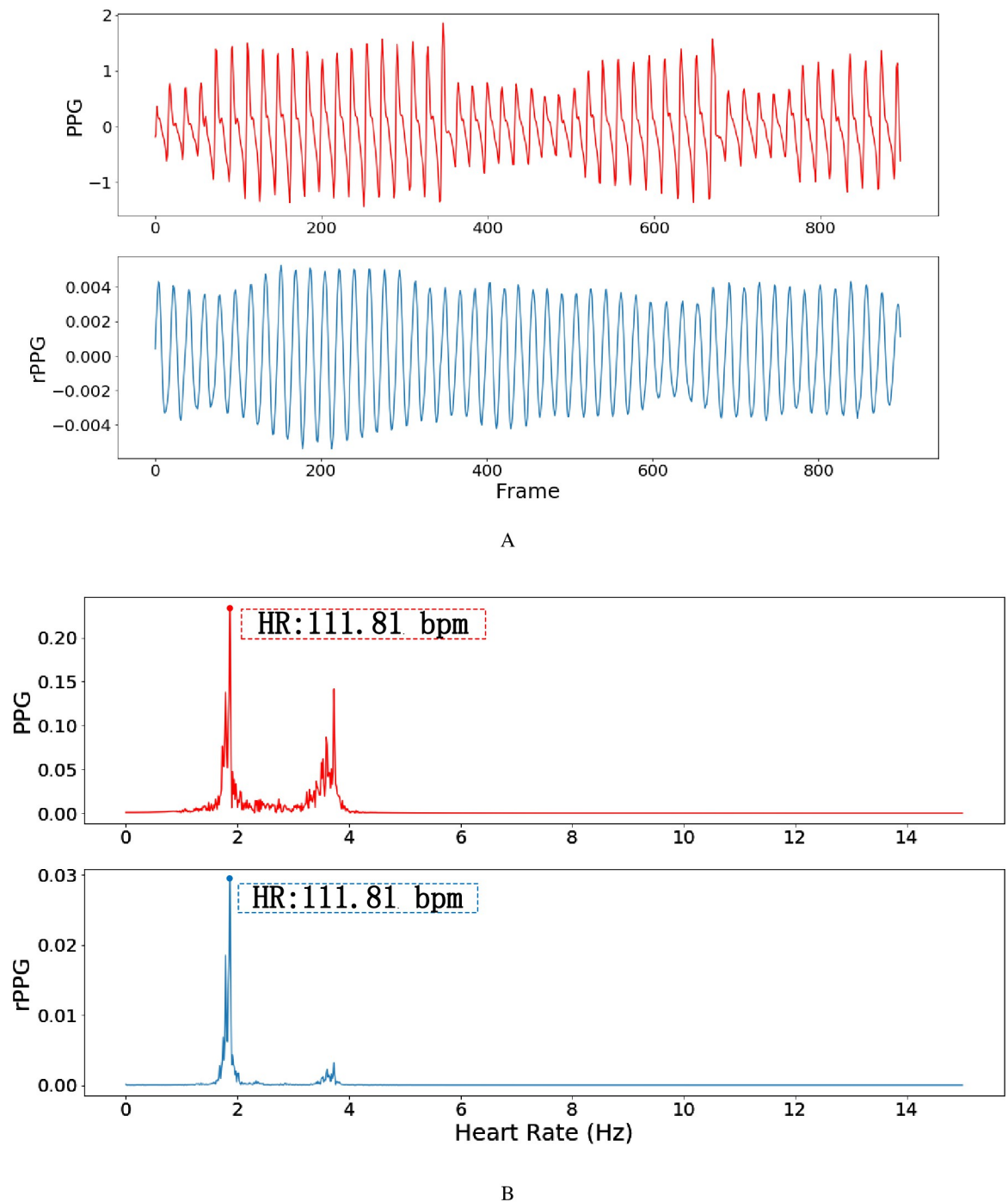


Fig 12. Comparison of LA-SSA result with reference PPG signal. A: Waveform comparison. B: Spectrum comparison.

<https://doi.org/10.1371/journal.pone.0275544.g012>

(ST, the images recorded by the camera were displayed on screen and shown to the subjects. A moving rectangle of the size of the face was added to the image, and the subjects were asked to keep their face inside), fast translation (FT, has the same setup as slow translation, except twice the speed of the moving target), small rotation (SR, different targets that were placed at 35 cm around the camera. The subjects were told to look at these targets in a predefined sequence.

They were asked to move not only their eyes but orient their head), and medium rotation (MR, has the same setup as for small rotation, but with targets placed 70 cm around the camera resulting in average head angle of 35°). The test subjects were placed in front of the camera with an average distance of 1.1 meters, resulting in a total of 60 video sequences. Each video takes 1 minute and is recorded by the ECO274 CVGE camera with a resolution of 640×480 pixels and a frame rate of 30 fps. The PPG pulse signal is recorded by the Contec CMS50E pulse oximeter with a sampling rate of 60 Hz.

The images of subjects appearing in the paper have been agreed by the subjects.

3.2 Experimental results

In this paper, the following indicators are used to evaluate the performance of the HR measurement method:

- **Mean Absolute Error(MAE)**: The mean value of absolute error between HR_{rPPG} estimated by recovered rPPG signal and HR_{PPG} estimated by reference PPG signal, reflecting the actual situation of the error between the measured heart rate and the proper heart rate.
- **Root mean square error(RMSE)**: The quadratic error between HR_{rPPG} and HR_{PPG} . It is susceptible to outliers, reflecting the stability of the algorithm. The lower the value is, the more stable the algorithm is.
- **Pearson correlation factor(r)**: The correlation between HR_{rPPG} and HR_{PPG} was used to evaluate the correlation between predicted heart rate and real heart rate.

In the LA-SSA method, the maximum autocorrelation coefficient ρ_i is introduced in SSA component selection to measure component periodicity. The weak periodic noise in the rPPG signal can be removed by setting the ρ_i threshold to select the periodic or quasi-periodic components related to the HR signal. In order to investigate the influence of the appropriate threshold on the accuracy of HR estimation under the condition of $\rho_i > 0.8$. The threshold was further set to 0.8, 0.85, and 0.90 for component screening and reconstruction, respectively. As shown in Table 2, it can be seen that when the threshold is set to 0.85, the reconstructed signal has the highest accuracy for HR estimation. Therefore, 0.85 is used as the standard for periodic components screening in this paper.

In order to improve the quality of the LA-SSA input signal, make the noise components in the signal are minimized before LA-SSA decomposition and reconstruction. In this paper, the performance of the recovered rPPG signal with different input signals to estimate HR is compared. The green channel [14], the chromaticity signal obtained through CHROM [15], the signal obtained through FastICA filtering [13], and the signal obtained through EEMD decomposition [18] and screening were used as the input of LA-SSA, respectively. The accuracy of the HR estimation is shown in Table 3. The results show that the signal processed by EEMD filtering shows higher accuracy for HR estimation after LA-SSA decomposition and reconstruction compared with the other three methods.

Table 2. Effect of different threshold on HR estimation accuracy.

| Threshold | MAE | RMSE | r |
|-------------|-------------|-------------|-------------|
| 0.80 | 1.63 | 3.95 | 0.98 |
| 0.85 | 1.37 | 3.61 | 0.98 |
| 0.90 | 1.89 | 4.02 | 0.97 |

<https://doi.org/10.1371/journal.pone.0275544.t002>

Table 3. Performance comparison of LA-SSA HR estimation using different input signals.

| Method | MAE | RMSE | r |
|--------------------|-------------|-------------|-------------|
| G+LA-SSA | 4.15 | 11.09 | 0.79 |
| CHROM+LA-SSA | 5.04 | 14.65 | 0.65 |
| FastICA+LA-SSA | 5.24 | 14.73 | 0.66 |
| EEMD+LA-SSA | 1.37 | 3.61 | 0.98 |

<https://doi.org/10.1371/journal.pone.0275544.t003>

Fig 13 shows the comparison of HR values obtained by the proposed method and the contact measurement method. The linear regression graph is shown in Fig 13A. It can be seen that the data points are concentrated near the linear regression line, and the slope of the linear regression line is close to 1, indicating that the estimated HR is highly correlated with the actual reference HR. Fig 13B is the Bland-Altman consistency analysis diagram. The blue center line represents the relative average error between the heart rate measurement and reference values. Two virtual red lines represent the confidence interval of 95% confidence [$\mu - 1.96\sigma$, $\mu + 1.96\sigma$], and only the points between the virtual lines are considered to be highly credible. The results show that most of the HR values obtained by the proposed method are in the confidence interval, indicating that the HR measurement values are highly consistent with the reference values.

In the PURE database, there are six different setups corresponding to different levels of head movement noise. Table 4 shows the performance for different noise level of LA-SSA HR estimation. It can be seen that when the setup is steady, the result of HR estimation by the LA-SSA method is best, and the performance may be a little poor when the setup is talking, fast translation and small rotation. However, from the overall results, the proposed method performs well in varying degrees of motion noise in the video.

Finally, Tables 5 and 6 show the performance comparison of the proposed method with other methods on the UBFC-RPPG database and PURE database. In the UBFC-RPPG database, the MAE, RMSE, and Pearson correlation coefficients of the LA-SSA method were 1.37 bpm, 3.61 bpm, and 0.98, respectively, which were better than those of all other algorithms in

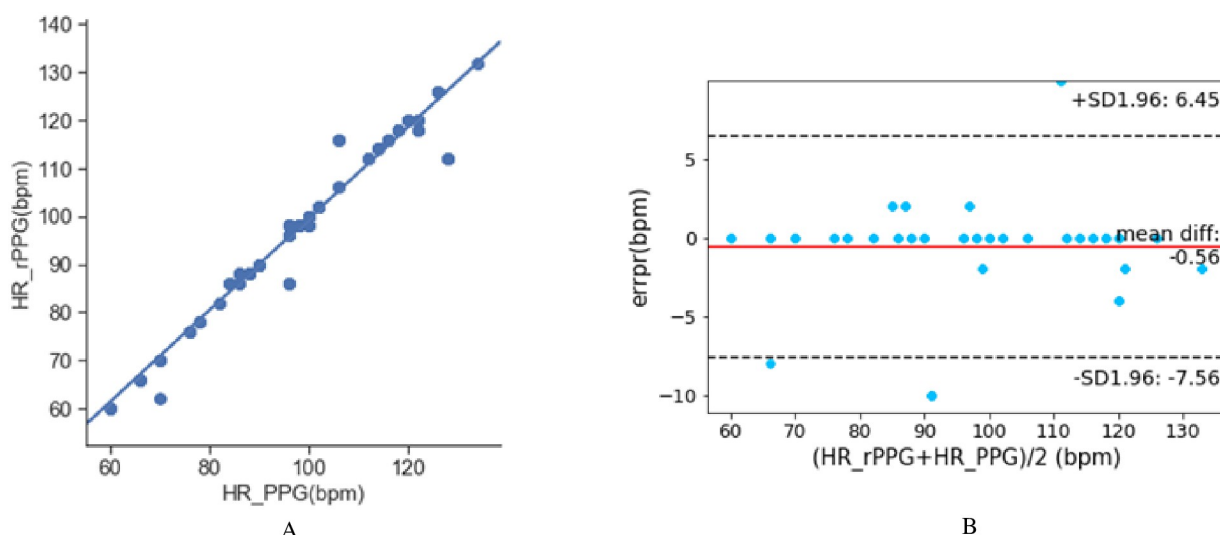


Fig 13. The comparison of HR values obtained by the proposed method and the contact measurement method. A: Linear regression diagram. B: Bland-Altman diagram.

<https://doi.org/10.1371/journal.pone.0275544.g013>

Table 4. Performance for different noise level of LA-SSA HR estimation.

| Setup | MAE | RMSE | r |
|------------------|-------------|-------------|-------------|
| Steady | 1.28 | 6.53 | 0.96 |
| Talking | 3.58 | 9.40 | 0.94 |
| Slow Translation | 2.45 | 9.20 | 0.91 |
| Fast Translation | 3.53 | 9.83 | 0.90 |
| Small Rotation | 3.72 | 10.65 | 0.93 |
| Medium Rotation | 2.63 | 9.30 | 0.93 |

<https://doi.org/10.1371/journal.pone.0275544.t004>

Table 5. Performance of different methods for HR estimation on database UBFC-RPPG.

| Method | MAE | RMSE | r |
|----------------------|-------------|-------------|-------------|
| Cpr+fine [40] | 2.10 | 3.43 | - |
| PulseGAN [24] | 2.09 | 4.42 | 0.97 |
| CK [42] | 2.29 | 3.80 | 0.98 |
| MAICA [41] | 3.34 | - | - |
| POS [42] | 2.43 | 6.60 | 0.93 |
| ICA [39] | 6.02 | - | 0.79 |
| GREEN [42] | 4.47 | 11.59 | 0.84 |
| CHROM [24] | 3.10 | 6.84 | 0.93 |
| cICA [44] | 3.14 | - | 0.91 |
| TSP+CHROM [38] | - | 2.38 | 0.96 |
| Bousefsafet al. [46] | 5.45 | 8.64 | - |
| G-R [43] | 9.79 | - | 0.65 |
| PVM [39] | 4.47 | - | - |
| BCG [45] | 27.92 | 37.96 | 0.25 |
| Tsou et al. [47] | 1.29 | 8.73 | - |
| Lee et al. [48] | 5.97 | 7.42 | 0.53 |
| PCA [14] | 9.65 | - | 0.67 |
| EEMD+LA-SSA | 1.37 | 3.61 | 0.98 |

<https://doi.org/10.1371/journal.pone.0275544.t005>

the table. In the PURE database, the MAE, RMSE, and Pearson correlation coefficients of the LA-SSA method were 2.87 bpm, 7.61 bpm, and 0.96, respectively, which were also better than those of most other algorithms in the table. The results show that the proposed method further removes the irregular noise in rPPG, which effectively improves the accuracy of HR estimation. Several deep learning methods were tested on the two databases. The length of video samples truncated by these deep learning methods is very different from that used in our process.

Table 6. Performance of different methods for HR estimation on database PURE.

| Method | MAE | RMSE | r |
|--------------------|-------------|-------------|-------------|
| ICA [49] | 24.1 | - | - |
| CHROM [24] | 3.82 | - | 0.97 |
| NMD-HR [49] | 8.68 | - | - |
| LICVPR [17] | 28.2 | - | -0.38 |
| Zhao et al. [50] | 3.09 | - | - |
| EEMD+LA-SSA | 2.87 | 7.61 | 0.96 |

<https://doi.org/10.1371/journal.pone.0275544.t006>

4 Conclusion

In this paper, a new method for SSA components selection and reconstruction is proposed. Based on the contribution of the SSA component to the original signal, the low-rank sparse matrix decomposition is used to select the appropriate reconstructed component in the global component signals. Then combined with the autocorrelation function, the weak periodic noise signal in the reconstructed component is eliminated, and the irregular noise such as face motion and light change in the rPPG signal is effectively removed. Tested on the UBFC-RPPG and PURE dataset, the experimental results verify the best performance of the method. In addition, the autocorrelation function in component filtering is based on the periodic strong and weak screening criteria for noise removal. The motion noise may be identified as the HR correlation component when the human body is in periodic motion, such as fitness. Then the algorithm's accuracy will be affected to a certain extent. In the subsequent study, we will focus on exploring the recognition and removal method of strong periodic noise. We will do some studies on the approaches based on deep learning.

Acknowledgments

The authors would like to thank all the experimental participants. The authors would like to thank Dr. Benezeth of University of Burgundy who provided the video data for conducting this study.

Author Contributions

Conceptualization: Zongkai Wei.

Data curation: Zongkai Wei, Jin Yuan, Yu Fang.

Formal analysis: Zongkai Wei.

Funding acquisition: Weibo Wang, Yongkang Zheng.

Investigation: Weibo Wang.

Methodology: Zongkai Wei.

Project administration: Weibo Wang.

Resources: Weibo Wang.

Software: Zongkai Wei, Jin Yuan, Yu Fang.

Supervision: Weibo Wang.

Validation: Zongkai Wei.

Visualization: Zongkai Wei.

Writing – original draft: Zongkai Wei.

Writing – review & editing: Zongkai Wei.

References

1. Liu S, Lan X, Yuen P C. Remote photoplethysmography correspondence feature for 3D mask face presentation attack detection[C]. Proceedings of the European Conference on Computer Vision (ECCV). 2018: 558-573.
2. Chen X, Cheng J, Song R, et al. Video-based heart rate measurement: Recent advances and future prospects[J]. IEEE Transactions on Instrumentation and Measurement, 2018, 68(10): 3600–3615. <https://doi.org/10.1109/TIM.2018.2879706>

3. Tadic M, Cuspidi C, Grassi G. Heart rate as a predictor of cardiovascular risk[J]. *European journal of clinical investigation*, 2018, 48(3): e12892. <https://doi.org/10.1111/eci.12892> PMID: 29355923
4. Mather M, Thayer J F. How heart rate variability affects emotion regulation brain networks[J]. *Current opinion in behavioral sciences*, 2018, 19: 98–104. <https://doi.org/10.1016/j.cobeha.2017.12.017> PMID: 29333483
5. Angeli F, Reboldi G, Verdecchia P. Heart failure, pulse pressure and heart rate: Refining risk stratification[J]. *International Journal of Cardiology*, 2018, 271: 206–208. <https://doi.org/10.1016/j.ijcard.2018.07.072> PMID: 30029799
6. Guohua Lu, Fang Yang, Xijing Jing, et al. Contact-free measurement of heartbeat signal via a doppler radar using adaptive filtering. 2010 International Conference on Image Analysis and Signal Processing, 2010: 89–92.
7. Pavlidis I, Dowdall J, Sun N, et al. Interacting with human physiology[J]. *Computer Vision and Image Understanding*, 2007, 108(1–2): 150–170. <https://doi.org/10.1016/j.cviu.2006.11.018>
8. Wang W, Brinker B D, Stuijk S, et al. Algorithmic principles of remote-PPG[J]. *IEEE Trans Biomed Eng*, 2017, 64(99): 1479–1491. <https://doi.org/10.1109/TBME.2016.2609282> PMID: 28113245
9. Verkruysse W, Svaasand L O, Nelson J S. Remote plethysmographic imaging using ambient light[J]. *Optics Express*, 2008, 16(26): 21434–21445. <https://doi.org/10.1364/oe.16.021434> PMID: 19104573
10. Aarts L A M, Jeanne V, Cleary J P, et al. Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit—A pilot study[J]. *Early human development*, 2013, 89(12): 943–948. <https://doi.org/10.1016/j.earlhumdev.2013.09.016> PMID: 24135159
11. Huang P W, Wu B J, Wu B F. A heart rate monitoring framework for real-world drivers using remote photoplethysmography[J]. *IEEE journal of biomedical and health informatics*, 2020, 25(5): 1397–1408. <https://doi.org/10.1109/JBHI.2020.3026481>
12. Salahuddin L, Cho J, Jeong M G, et al. Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings[C]. 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2007: 4656–4659.
13. Ming-Zher, Poh, Daniel, et al. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation[J]. *Optics Express*, 2010, 18(10): 10762–10774. <https://doi.org/10.1364/OE.18.010762>
14. Lewandowska M, Nowak J. Measuring pulse rate with a webcam[J]. *Journal of Medical Imaging and Health Informatics*, 2012, 2(1): 87–92. <https://doi.org/10.1166/jmihi.2012.1064>
15. De H G, Jeanne V. Robust pulse rate from chrominance-based rPPG[J]. *IEEE Transactions on Biomedical Engineering*, 2013, 60(10): 2878–2886. <https://doi.org/10.1109/TBME.2013.2266196>
16. De Haan G, Van Leest A. Improved motion robustness of remote-PPG by using the blood volume pulse signature[J]. *Physiological measurement*, 2014, 35(9): 1913–1926. <https://doi.org/10.1088/0967-3334/35/9/1913> PMID: 25159049
17. Li X, Chen J, Zhao G, et al. Remote heart rate measurement from face videos under realistic situations [C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014: 4264–4271.
18. Chen D Y, Wang J J, Lin K Y, et al. Image sensor-based heart rate evaluation from face reflectance using Hilbert–Huang transform[J]. *IEEE Sensors Journal*, 2014, 15(1): 618–627. <https://doi.org/10.1109/JSEN.2014.2347397>
19. Wang W, Stuijk S, De Haan G. Exploiting spatial redundancy of image sensor for motion robust rPPG [J]. *IEEE transactions on Biomedical Engineering*, 2014, 62(2): 415–425. <https://doi.org/10.1109/TBME.2014.2356291> PMID: 25216474
20. Kumar M, Veeraraghavan A, Sabharwal A. DistancePPG: robust non-contact vital signs monitoring using a camera[J]. *Biomedical Optics Express*, 2015, 6(5): 1565–1588. <https://doi.org/10.1364/BOE.6.001565> PMID: 26137365
21. Yue H, Li X, Cai K, et al. Non-contact heart rate detection by combining empirical mode decomposition and permutation entropy under non-cooperative face shake[J]. *Neurocomputing*, 2020, 392: 142–152. <https://doi.org/10.1016/j.neucom.2018.09.100>
22. Niu X, Han H, Shan S, et al. Synrhythm: Learning a deep heart rate estimator from general to specific [C]. 2018 24th International Conference on Pattern Recognition (ICPR). IEEE, 2018: 3580–3585.
23. Niu X, Zhao X, Han H, et al. Robust remote heart rate estimation from face utilizing spatial-temporal attention[C]. 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019). IEEE, 2019: 1–8.
24. Song R, Chen H, Cheng J, et al. PulseGAN: learning to generate realistic pulse waveforms in remote photoplethysmography[J]. in *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(5): 1373–1384. <https://doi.org/10.1109/JBHI.2021.3051176> PMID: 33434140

25. Zhang K, Zhang Z, Li Z, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. *IEEE Signal Processing Letters*, 2016, 23(10): 1499–1503. <https://doi.org/10.1109/LSP.2016.2603342>
26. Rapczynski M, Werner P, Saxen F, et al. How the region of interest impacts contact free heart rate estimation algorithms[C]. 2018 25th IEEE International Conference on Image Processing (ICIP), 2018: 2027–2031.
27. Saxen F, Alhamadi A. Color-based skin segmentation: an evaluation of the state of the art[C]. 2014 IEEE International Conference on Image Processing (ICIP), 2014: 4467–4471.
28. Wu Z, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. *Advances in adaptive data analysis*, 2009, 1(01): 1–41. <https://doi.org/10.1142/S1793536909000047>
29. Zhang Y, Liu B, Zhang Z. Combining ensemble empirical mode decomposition with spectrum subtraction technique for heart rate monitoring using wrist-type photoplethysmography[J]. *Biomedical Signal Processing and Control*, 2015, 21: 119–125. <https://doi.org/10.1016/j.bspc.2015.05.006>
30. Zhang Y, Dong Z, Zhang K, et al. Illumination variation-resistant video-based heart rate monitoring using LAB color space[J]. *Optics and Lasers in Engineering*. 2021, 136: 106328. <https://doi.org/10.1016/j.optlaseng.2020.106328>
31. Tarvainen M P, Ranta-Aho P O, Karjalainen P A. An advanced detrending method with application to HRV analysis[J]. *IEEE Transactions on Bio-medical Engineering*, 2002, 49(2): 172–175. <https://doi.org/10.1109/10.979357> PMID: 12066885
32. Elsner J B. Analysis of Time Series Structure: SSA and Related Techniques[J]. *Journal of the American Statistical Association*, 2002, 97(460): 1207–1208. <https://doi.org/10.1198/jasa.2002.s239>
33. Elsner J B, Tsonis A A. Singular spectrum analysis: a new tool in time series analysis[M]. Springer Science & Business Media, 1996.
34. Davis M J, Hill M A, Kuo L. Local regulation of microvascular perfusion[M]. *Microcirculation*. Academic Press, 2008: 161–284.
35. Wang D, Yang X, Liu X, et al. Detail-preserving pulse wave extraction from facial videos using consumer-level camera[J]. *Biomedical Optics Express*, 2020, 11(4): 1876–1891. <https://doi.org/10.1364/BOE.380646> PMID: 32341854
36. Bertsekas D P. Constrained optimization and Lagrange multiplier methods[M]. Academic press, 1982.
37. Sugita N, Akay M, Akay Y M, et al. Noise reduction technique for single-color video plethysmography using singular spectrum analysis[J]. in *IEEE Journal of Biomedical and Health Informatics*, 2020, 24(6): 1788–1795. <https://doi.org/10.1109/JBHI.2019.2949883> PMID: 31714244
38. Bobbia S, Macwan R, Benezeth Y, et al. Unsupervised skin tissue segmentation for remote photoplethysmography[J]. *Pattern Recognition Letters*, 2019, 124: 82–90. <https://doi.org/10.1016/j.patrec.2017.10.017>
39. Bobbia S, Macwan R, Benezeth Y, et al. Periodic variance maximization using generalized eigenvalue decomposition applied to remote photoplethysmography estimation[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018: 1332–1340.
40. Artemyev M, Churikova M, Grinenko M, et al. Robust algorithm for remote photoplethysmography in realistic conditions[J]. *Digital Signal Processing*, 2020, 104: 1–9. <https://doi.org/10.1016/j.dsp.2020.102737>
41. Macwan R, Benezeth Y, Mansouri A. Heart rate estimation using remote photoplethysmography with multi-objective optimization[J]. *Biomedical Signal Processing and Control*, 2019, 49: 24–33. <https://doi.org/10.1016/j.bspc.2018.10.012>
42. Song R, Zhang S, Cheng J, et al. New insights on super-high resolution for video-based heart rate estimation with a semi-blind source separation method[J]. *Computers in Biology and Medicine*, 2019, 116: 103535. <https://doi.org/10.1016/j.compbiomed.2019.103535> PMID: 31760272
43. M. Huelsbusch. An image-based functional method for opto-electronic detection of skin-perfusion[D]. PhD thesis, RWTH Aachen dept. of EE, 2008.
44. Macwan R, Benezeth Y, Mansouri A. Remote photoplethysmography with constrained ICA using periodicity and chrominance constraints[J]. *Biomedical engineering online*, 2018, 17(1): 1–22. <https://doi.org/10.1186/s12938-018-0450-3> PMID: 29426326
45. Balakrishnan G, Durand F, Guttag J. Detecting pulse from head motions in video[C]. 2013 IEEE Conference on Computer Vision and Pattern Recognition, 2013: 3430–3437.
46. Bousefsaf F, Pruski A, Maaoui C. 3D convolutional neural networks for remote pulse rate measurement and mapping from facial video[J]. *Applied Sciences*, 2019, 9(20): 4364. <https://doi.org/10.3390/app9204364>

47. Tsou Y Y, Lee Y A, Hsu C T, et al. Siamese-rPPG network: Remote photoplethysmography signal estimation from face videos[C]. Proceedings of the 35th annual ACM symposium on applied computing. 2020: 2066-2073.
48. Lee E, Chen E, Lee C Y. Meta-rPPG: remote heart rate estimation using a transductive meta-learner [C]. in European Conference on Computer Vision. Springer, 2020: 392–409.
49. Demirezen H, Erdem C E. Remote photoplethysmography using nonlinear mode decomposition[C]. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018:1060-1064.
50. Zhao C, Mei P, Xu S, et al. Performance evaluation of visual object detection and tracking algorithms used in remote photoplethysmography[C]. 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2019: 1646-1655.
51. Stricker R, Müller S, Gross H M. Non-contact video-based pulse rate measurement on a mobile service robot[C]. The 23rd IEEE International Symposium on Robot and Human Interactive Communication. IEEE, 2014: 1056-1062.